Translation of Book

SITUATIONAL CONTROL: Theory and Practice

by

D. A. Pospelov
"Nauka" Publishers, Moscow, 1986

dated

November 26, 1991

TABLE OF CONTENTS

If this were so, it would not be bad, but if not bad, then that's how it would be, but since it is not so, then it is not thus. Such is the logic of things!

Lewis Carroll

FOREWORD

When I planned to write this book and was considering ways of how best to do it, the recollection of Borobudur came to mind.

Borobudur is a famous Buddhist temple complex located in the center of Java in a broad valley surrounded by mountains, some of which are active volcanos. When a traveler approaches Borobudur, he sees before him a hugh edifice rising upwards by step-like terraces. Each such step is completely covered on the outside by stone relief, here and there in grottos and recesses sit Buddhas and bodhisattvas. There are many kilometers of this stone book in which are quaintly interwoven mythological and fantastic plots depicting a series of events from the life of Buddha and bodhisattvas. Raising upwards along the galleries surrounding Borobudur, the traveller gradually loses his sense of reality among the thousands of sculptures and stone reliefs. And there arrives the moment when he comes out on the next ledge; it is a stone square which opens onto an astonishing view of the valley and the mountains. A stone chaos of phantoms remains below. And here, above, there is only nature in its primordial beauty and grandeur. But this terrace still is not the last. Raising still higher and higher, the traveler reaches the highest point of borobudur, where a giant stupa rears up. It is a vital attribute of the Buddhist structure. And already, those lower tiers, in which the stone chronicles of the real and fantastical worlds are frozen, cannot be seen from there.

Moving in this way, it is difficult to perceive the general idea of Borobudur, its hidden meaning, incarnate by unknown builders in the VII or IX centuries in stone. However, the XX century provides a new possibility for penetrating into the secret of this ancient structure. From a helicopter it is possible to see it immediately in unison with all of its tiers, integrally. An then the distinct plan of the structure suddenly appears. This cumbersome pyramid in its thrust upward reflects three worlds: the world of sensual forms, the world of concepts, and the world of abstract idea. That which the traveller, moving upward through the galleries of the temple, had to deduce from the mottled mosaic of his impressions, is immediately revealed in the view from the air.

Describing a given domain of science, it is possible also to travel two paths. One can take the reader-traveller through long galleries, filled with isolated facts, integral parts of all together somewhat complete, joined by general ideas and principles. But it is possible to describe this domain "from the altitude of a bird's flight," immediately formulating these fundamental ideas and principles, throwing aside all details, even if each of them is remarkable and interesting. However, it is possible to try to comibne both paths: the analytical and the synthetic.

This scheme is realized in the present book. To what extent this is good is for the reader to judge, not me. In any case the book is made up as though there are four levels.

On the first level, it presents a conceptual look at those questions which comprise the essence of situational control. The reading of these paragraphs does not require more than a low level of knowledge. And the entire first level, it seems to me, gives the reader, who is interested in only the essence of situational control, a sufficiently complete picture of the ideas and possibilities of its practical application.

On the second level (referring to those paragraphs of the book marked with a little star) it sets forth the methods and procedures used in situational control. These methods can be of use not only by those who want to use the situational control method. They can be of use also in other cases. Therefore, the second level is intended for a sufficiently wide circle of specialist who are intensely interested in developing methods of solving problems bearing on logic-linguistic models. For an understanding of these paragraphs entering into this level, requires reading through and analyzing the sections forming the above level.

The third level (those paragraphs marked by two little stars) is intended for specialists who want to use the method of situational modeling for practical purposes of solving substantive problems. In them are concentrated methodical and methodological recommendations and advise for constructing suitable systems. For the perusal of paragraphs of this level, it is desirable to read through also those paragraphs which relate to the first two levels. But, generally speaking, in the majority of cases, it is possible to skip only that information which is in the first level.

The fourth level is supplementary. Therefore, it is separated in the text of the book by brevier [8 point type]. Some situations and methods described in the preceding levels are commented on in its sections and it provides a summary of the literature. In the text of the primary three levels, there are no references to the literature. They distract the reader and clutter up the text. Therefore, on the fourth level there is complete coverage of sources used in the writing of the book, also included are those works in which their ideas closely resembled those discussed in the book. The fourth level concludes with a historical synopsis of the development of situational modeling.

Furthermore, the book has two peculiarities. In it there are quite enough history, literature citations, and antidotes interspersed throughout the text, especially in the first level. I have attached important meaning to them, although reader-purists may consider that they are out of place in a scientific monograph. They are interjected in order for me to make the necessary associations for the reader. And, this is all that I can say here in their justification.

Secondly, there are the examples. In them it is not uncommon to find statements which are important for understanding one or the other aspect of a given point which they illustrate. Therefore, it is desirable that they not be avoided in reading.

Now, it is necessary to say a few words about why this book is being published in the series of books, "Problems of Artificial Intelligence," whose emblem adorns the cover of the book. The concepts of situational control, as the reader will learn about from the historical essay in the Appendix, originated from where, at that moment, the new science of artificial intelligence appeared on the horizon. However, the logic of the birth of new scientific directions is such, that they are never conceived in a void, they don't arise like the Phoenix, from nothing. The new sciences are born in the midst of "old" acknowledged sciences. Only then, when this maturation reaches a certain point, will the birth of a new science or a new scientific paradigm occur.

Artificial intelligence, this idol of modern day, certainly was not an exception. The fact is that it subsequently became one, having originated in the depths of cybernetics, linguistics, psychology, and programming. If very concisely (and therefore inaccurately from the point of view of a purist) one were to define what today is understood by the theory of artificial intelligence, it would be possible to say that it is a science about knowledge, about how it is derived, represented in an artificial system, processed within that system, and used for the solution of practical problems. In other words, systems studied within the scope of artificial intelligence and created in the course of this science are systems the work of which rests on knowledge, reflected by the semantics and pragmatics of that external world in which intellectual systems operate.

In completing the foreword, I would like to note, that the basic conclusions of this book were formulated and definitively shaped in the course of many years of contacts with all the active participants of what could be called the "situational movement" (the term of G.P. Shchedrovitskiy).

I deeply thank all of my comrades and colleagues in this "movement." Numerous conferences, symposia, schools, and seminars on situational control united us, permitted us to work out common terminology and formulate conceptual models, forming the basis of the method of situational control. It is doubtful whether it is now possible to accurately determine, "who first said A," and I would not want, that from the text of the book the reader would assume, that its author adhered to the well known rule: "The author is not the one who first said A, but the one who first said I."[*]

---

[*] This is a Russian expression that does not translate very well. A is the first letter of the Cyrillic alphabet and Я , or translated into English "I", is the last.

Chapter 1

PRINCIPLES OF SITUATIONAL CONTROL

> *Various nations usually have many words to describe that which interests them the most.*
>
> *Ernest Renan*

1.1 Traditional Scheme of Control

The traditional theory of automatic control of technical objects which grew from the earlier existing theory of automatic control, has to do with such objects for which the control procedure in its most general form is represented as is shown on Figure 1.1.

On the object being controlled are acting (n + r) inputs, and the value of the inputs x_i can be estimated at any moment of time, while for the inputs w_j such possibility is absent. Most often this occurs because of an absence of required sensors. But it may happen that the measuring of inputs w_j is physically impossible or costs too much. In the ideal case there are no w_j inputs. The object has m outputs. It is assumed that between the input vector <X, W> and the output vector Y there exists a certain connection Y=f(X,W). This connection can be of a most diverse character: it can be functional, statistical, in the form of a non-single-valued mapping, and so forth. Important is only that changes in X and W affect somehow the values of Y. These values of Y are not

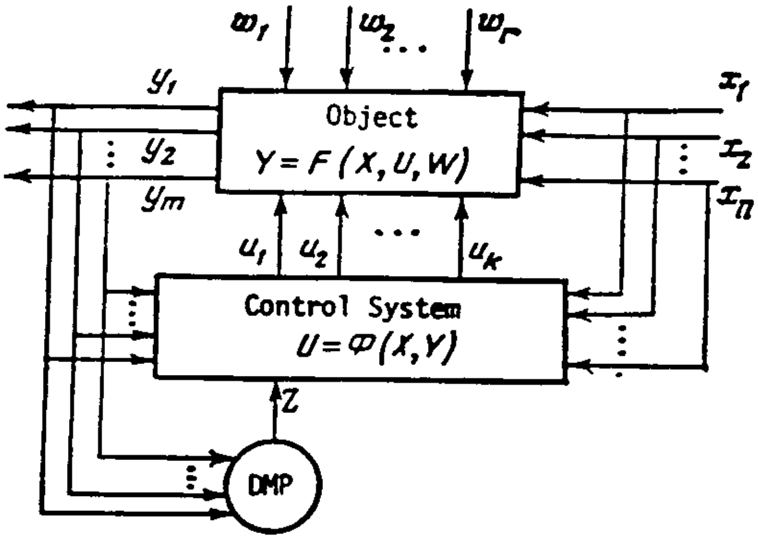


Figure 1.1

indifferent to a certain person, whom we shall henceforth conditionally call *decision-making person* (DMP), as it is customary in modern operations research theory. For the DMP, it is important to attain certain values of Y for any values of X and W. Vectors X and W cannot be affected by the DMP. For this reason it constructs the system of control so as to obtain from the object the values of Y that are of interest to it. This is achieved with the aid of special regulatory actions u_i supplied from the control system to the object. Just as vectors X and W, the values of the action vector U influence the values of vector Y. In other words, instead of the relation $Y=f(X,W)$ there occurs the relationship $Y=F(X,U,W)$. The problem of the DMP is to make the control system such, that with available information on the current values of X (while any information on the values of W is inaccessible to the control system, of course) it would then be possible to attain such values of Y, which would satisfy the DMP. This means that the control system implements a certain relationship of the form $U = \Phi(X,Y)$, with the aid of which it finds the required controlling actions on the object.

If the DMP is not satisfied with its brainchild, then it can enter corrections into its functioning. On Figure 1.1 this possibility is indicated in the form of supplying to the control system of a correction vector Z. If, however, the correction does not lead to the desired objective or is impossible, then the DMP can replace the control system.

Let us not go into any further consideration of the cited scheme. For us the following is important: in order that the DMP could construct a control system satisfactory to it, it must know two things: the set of preferred vectors Y and the way of stating the transformation f. Transformation f describes the functioning of the object of control, while the knowledge of the preferred values of Y tells about the understanding of the purpose of its functioning. And so, the DMP must know the *description of the object* (i.e., its structure and functioning), which it intends to control, and the *purposes of existence* of this object. Knowing this, it can formulate a criterion of control, in which are embodied the requirements toward the system of control. The transformation Φ on Figure 1.1 embodies in itself a criterion of control of the object. It can be different: maximization or minimization of some values of vector Y, their maintenance within certain given limits, prevention of certain combinations of these values or prohibition of the occurrence of certain time sequences of values of vector Y, and so forth. It is precisely the presence of the control criterion that enables the DMP to pose and to solve the traditional problem of controlling the object.

Thus originates the chain: description of the object of control -- description of the purposes of existence of the object of control -- formation of its control criterion -- design and creation of the system of control. It might be expected, that the theory of control would devote equal attention to all elements of the chain. But historically this did not occur. As a rule, specialists in the field of control have exerted efforts only toward finding a procedure of controlling an object, after the object itself and its control criterion had already been described in exact terms. Optimization of control was the central problem of the traditional theory of automatic control. And only at the last stage of its development was the attention of the specialists shifted to the problem of identification of the object of control and to the problem of revealing of the criteria of controlling it.

But here arose a number of difficulties related to the fact that systems had become the objects of control for which the familiar methods elaborated over decades turned out to be unapplicable. What was their nature?

1.2 Nontraditional Objects of Control

They are called variously: poorly defined or weakly structured, organizational or possessing *free will*. But regardless of appellation these new objects possess a number of properties unexpected for traditional control, which distinguish them from the familiar objects of the past. Let us enumerate their basic properties.

1. **Uniqueness.** Every object has such a structure and function, so that its control system must be built with consideration of all its qualities and it is not permitted to apply to it some standard model procedure of control. If for example, it was possible to design a system of controlling the motion for the automobile VAZ 2103, then it can be used for any specific automobile of such a model. If however, a control system has been designed for the ministry of public health of a certain Union Republic, then it may not be transferred without changes to a similar ministry of another republic. In transferring, it is necessary to consider all individual peculiarities of the

new object of control. This circumstance sharply raises the cost of the procedure of constructing the control system, because in fact it is necessary to create as many control systems, as there are objects that we wish to automate.

2. **Absence of a Formalizable Purpose of Existence.** For earlier objects, familiar to the theory of automatic control, it was always clear why the object was created whose control we are designing. As a rule, control people had to do with objects of an artificial origin, created by people for achieving goals understandable to them. A machine tool was to carry out a definite working of the billets, an airplane had to transport through the air passengers and freight. But not for all objects (even those created by man) is it possible just as precisely to formulate the purpose of their existence. Nowadays, when we want to control cities, branches of national economy, regions, ecosystems, we get into a most difficult situation when attempting to precisely formulate the purpose of existence of these objects. Even though created by people, they spring up not by somebody's plan, to solve some specific problem, but developed gradually owing to definite social-economic and historical reasons. Which, for example, are the purposes for the existence of supercities? Or, for what purpose did originate the ecosystem of the Azov-Black Sea basin? To such questions it is practically impossible to reply.

And this leads to extremely great complexities in the formation of a control criterion. For the control criterion in traditional control systems was most closelytied with the purpose of existence of the object. The control criterion of an aircraft was based on the attainment by it of its purpose of existence -- transport of people and loads by air, while the criterion of control of synthetic rubber production was based on considerations of raising the quality of the product. It is precisely because of this that in different automated control systems, created for objects of a new class, it is very often possible to observe the realization of different criteria of control.

3. **Absence of Optimality.** A result of what was discussed in the preceding two subheadings is the incompetence of formulation of the classical problem of optimization for objects of the new nature (type). Because of the absence of the purpose of existence (within the framework of the theory of control) it is not possible to construct an objective control criterion for

the considered objects. The criterion of control becomes subjective, totally dependent on the DMP.

This extremely important (for subsequent considerations) thought can be illustrated by the following example, which actually stands somewhat apart from the control problem proper. Let us measure, with the aid of people-experts, the dimension of some object. For example the length of the automobile VAZ 2105. Each specialist names a certain number, which, in his opinion, denotes the length of the automobile. The DMP (which also can estimate the length of the automobile) analyses the statements of the experts. If there is an expert who estimates the length of the automobile as 10 m or 50 m, then the DMP is justified to suspect his qualifications as an expert. Tossing aside such anomalous suggestions, the DMP can sum up the remaining numbers and find an arithmetic average from the obtained result. Such an averaging sort of objectivizes the result. If there had been many experts and they had a good eye-judgment, then the result of this expertise would be close to the true length of the automobile.

Let us point out two peculiarities of the procedure described by us above. Firstly, if the experts are already many, then the appearance of some new expert will not bring about any special changes in the result obtained by the DMP. In other words, such an expertise has the property of stability. Secondly, it is possible to check the quality of the expertise by taking some kind of measuring device the precision of which satisfies the DMP, and to make a measurement. Whatever closeness of the measurement data to the result obtained with the aid of experts will characterize the quality of the expertise and this allows to pose, for example, the problem of optimizing the formation of a collection of experts by their number or by some social or physiological characteristics.

Now let us consider an example of another expertise. A group of people became lost in a cave. After long wandering about they find themselves at a place, from which the path that led to it, branches out. Left and right are running off underground corridors. It is required to decide: which way to go? The leader of the group (the DMP) takes a poll. Proponents of moving along the left corridor state their considerations, their opponents state theirs. For the DMP the reasons cited by either do not sound too convincing, and so the DMP takes a vote. Let's assume that the majority of the participants of this expertise favored moving along the left corridor. The group then followed it. After several days, weakened by hunger and thirst, they get out to the surface. And since that time, the DMP is being tormented by the question: did it make the correct decision?

Clearly, there is no answer to this question. There would be one, if the group had also passed along the right corridor. Possibly it would have taken them at once to the surface, but it is extremely probable that they would have remained underground forever. To evaluate the correctness of the choice, its expedience, optimality in this case is possible only by having a plan of the cave, which means having *passed* through all of its corridors. In distinction from the situation with the measurement of the length of an automobile, there is no possibility to evaluate the quality of the accepted decision, if the alternative decisions have not been checked. Moreover, the expertise of the second type does not have the stability that characterized the previous example. If, after having made the decision or moving along the left corridor by an overwhelming majority of the experts, a man would have crawled out of the left corridor into the chamber where the group was standing, and this man, who also was looking for an exit, said that this path leads to a dead-end, then the whole result of the poll would collapse. And then the subsequent move would be continued along the right corridor.

A situation similar to the search of a path in a cave happens very often. In deciding on the selection of whatever characteristics of a future product, in making whatever decisions on the structure and methods of functioning of an automated control system, during any *willed* decision (albeit supported by a number of considerations of the DMP itself and of other experts) there always occurs a situation, described by us (above). From this it follows that in these cases it is impossible to speak about an optimality of the resulting decisions. It is not possible to demand, for instance, that the designers would create an optimal Automatic Control System for a branch of industry or a city. Such a formulation is incorrect and incompetent. The quality of a created system for controlling objects having this new nature can be evaluated only subjectively by the DMP itself or by a collection of them. Therefore it is more proper to talk about the expediency of the result of control and not about its optimality. It is only important that the DMP would not be afraid to make decisions in required cases.

In the one-act stageplay of Slavomir Mrozhek, which is called "Striptease", is described just such a situation, in which one of the DMPs is shirking to make a decision. Two persons, unknown to each other, are placed by some unfathomable means, into a room from which there is an exit through an open door. They need to make the decision: what to do? The following dialogue takes place:

> *First person:* What is freedom? It is the possibility of choosing. While I remain here, I know that I can exit through this door, until then I am free. But at the moment when I get up and exit from here I will thereby make a choice, i.e., limit the possibility of my behavior and loose my freedom. I shall become a slave of my departure.
>
> *Second person:* But, by staying in place and not going anywhere, you will also make a choice. You are choosing sitting and non-exit.
>
> *First person:* Not true. I sit, but I can exit. Having exited I exclude the possibility of sitting.

4. **Dynamicity.** In a certain sense the objects encountered presently by the theory of control are similar to living systems. With the passage of time their structure and functioning change. The objects so to speak evolve in time. At an enterprize new shops are built, new manufactures arise, technology changes. The ECM* network grows, individual networks begin to unit among each other, tending towards an allworld network of data processing. The structure of a branch of industry changes, the cities expand their boundaries, their building development changes, the flows of transport and people are displaced. And this dynamicity must be taken into account in the control systems of such objects. They willy-nilly must be adaptive, ready for change of their functioning.

5. **Incompleteness of Description.** As a rule, no collection of experts who know the object of control, are able to immediately provide the information which would be known beforehand to be sufficient to create a control system of the object. There exist several reasons why this is so. While describing an object of control of the old type, the controller always

---

* Electronic Computing Machines

knew about those assumptions which he adopted, while constructing the description. He could presume that the transfer function of the object is of such and such form, that the delay does not play a great role in its functioning, that the effect of parameters W_j is insignificant and may be ignored, and so forth. And if the control system created by him turned out to be not very good, then he knew, which assumptions he should waive. But in working with objects of the new nature (type) these assumptions cannot be formulated so clearly and easily. The controller in this case is almost entirely reliant on experts who know the object of control. And whatever the level of assumptions is, is actually suggested by them. But, not being specialists on control systems, the experts cannot evaluate that level of completeness of description, which is required by the specialist on control.

For, the description itself, the distinction in it of whatever aspects and peculiarities is tightly linked with the problem of control. And this cannot always be discerned by a person who is looking at the object of control with different eyes -- the eyes of a technologist. I will illustrate my idea by an example of a conversation of two controllers, two outstanding airmen, Antoin de Saint Exupery and his friend Guillaume. Guillaume explains to his colleague how to fly over Spain on the way from France to Africa:

> "That was a strange lesson in geography. Guillaume did not present me any information about Spain, he gifted to me its friendship. He did not speak of water bodies, or the numbers of population and head of cattle. He talked not about Guadise, but of the three orange trees, that grow on the edge of a field near Guadise. 'Beware, mark them on your chart. . .' And from that hour on these three trees occupied on my chart more space than the Sierra Nevada. He talked not about Lork, but about a small farm near Lork. Of the life of that farm. Of its owner. And the owner's wife. And that couple, lost in the expanses of the earth at over one thousand kilometers away from us, grew immensely in my eyes. Their house stood on a mountain slope, their windows shone from afar, like stars, -- just like lighthouse keepers these two were always ready to help people with their light. . . Thus, little by little, Spain on my map, under the lamp of Guillaume, because some fairytale country. I marked with crosses the landing strips and the dangerous traps. Marked off the farmer on the hill and the brook on the meadow. Carefully marked on the map the shepherdess with thirty lambs. . ." ("The People Planets").

This is a sample of how one professional describes an object to another professional. Here everything is important: what is being spoken of, at what level of detail, in what language. In the work of a controller with a complex object, the greatest complexities arise during contact with technologists who know the given object. Nonagreement of their opinions of it sometimes result in a complete non-understanding of each other, as result of which there occurs an incompleteness of description, which is needed for guidance by the designer of the control system.

Another reason of no small importance for incompleteness of object description is the ignorance of certain aspects of its functioning by the technologist himself. Certain situations, never encountered by him before, can obviously not be communicated to the control systems designer. Most often these are the various emergency situations. To illustrate possible consequences of breakdown let us recall the catastrophe with the power system of the USA that occurred several decades ago. The failure of the system is blamed on the control system. Automatic switching-off of lines and sources of power during overloads resulted in the loss of electric power over a considerable part of the country, and this led to enormous losses. As another example may serve the control system of a large ECM, its operational system. When its design specialists in the 1960s were creating multiprogram operational systems with developed systems of interruptions and priorities, they did not at all count on situations in which the control system would block itself because of the impossibility of coming out of interruptions. In that and in the other case the controllers were fooled by the incompleteness of the description of object functioning and the effect of signals u_t on this functioning.

The third reason for incompleteness of description can be the absence of a precise understanding of object functioning by the technologist himself. While issuing to the controller a large amount of information, he nevertheless does not tell him the most important one by which he himself makes a decision on the functioning of the object. He does this unconsciously, because the *most important information* may be accounted for by him only at the level of his own intuition.

I recall one case of which I was a witness. The World Health Organization (WHO) distributed in a number of countries the histories of the sickness of some sick persons from Scandinavian countries and asked leading psychiatrists of different countries to give a diagnosis of the sickness on the basis of the sickness histories. In the USSR these were reprinted and sent out to a considerable number of specialists. After a certain time they assembled all together in Moscow for a discussion of their diagnoses. A table of collected results was completed and hanged up for general viewing giving diagnoses for all the patients. This was an astonishing table! Almost for each case there was a scatter of diagnoses ranging from *practically healthy* to *the sickness of X is of the most aggravated form*. The authors of the diagnoses came out on the stage and substantiated their points of view. The most interesting thing was that, using the identical data from the history of the sickness, they arrived at almost opposite conclusions. For instance, from the fact that the patient refused to contact his relatives and his doctor, in one case followed the conclusion about a desire for isolation, non-contactibility, and in another, about a fully adequate form of behavior for the given situation. The arguments stopped in fact after the address of one very respected older lady specialist in the field of psychiatry. She said, literally, the following: *Colleagues, what are we arguing about? We all know that as soon as the patient comes to our office, we shall in the very first second determine -- is he sick or not. Except that we cannot tell, how this happens with us.* And everyone agreed with the speaker.

And finally, one other reason that leads to incompleteness of description of complex objects. This consists in the fact that many peculiarities of the functioning of the object, and sometimes also of its structure, cannot be described quantitatively. They admit only of a qualitative, word description. The transition from qualitative description to certain formal concepts must be carried out by the controller who is not always capable to solve such a complex problem.

6. **Presence of Free Will.** In many objects of control, people are elements of their structure. These are the so-called organizational systems. In distinction from all other elements that comprise objects, people function in it with consideration of their personal interests and goals. Their interests and goals can differ considerably from that which they are supposed to do for the point of view of the DMP. Their individual behavior is practically impossible to account for in creating systems of control, and

special methods to neutralize their influence on the functioning of an object of control are required. Otherwise situations may arise, similar to the one described in the following anecdote.

A certain X, a big boss, says to his subordinates: *I consider that post A should be occupied by a retired general of about 60 years of age.* One of the subordinates joyfully exclaims that he knows such a man. But X is obviously displeased. *You idiot, -- whispers into the employee's ear his neighbor, -- he simply wants to put his friend into this post.*

There exist some other peculiarities of the objects of the new type, into which the theory of control has begun to bump into since the end of the 50s years of our century. But even what has already been said is probably quite enough for the purpose of estimating the necessity in the new approach to an object of control in attempting to create a system governing it. But before we pass on to the description of the possible ways of solving this problem, we need to make an excursion into the psychology of thinking.

1.3 Psychological Premises

Psychologists have devoted much attention to the problem of studying the behavior of man during the solution of problems arising before him. Not few conceptual schemes have been created which explain the peculiarities of his behavior during the solution of problems different in their character. Let us consider three such schemes that received widest distribution and are of interest to us within the framework of our book.

The crudest system carries the appelation of stimulo-reactive theory. In its base lies the concept of the object of control as a "black box". The internal structure of the object is incomprehensible for the DMP. Observable are only the external input signals entering the object (its stimuli), and the output signals of the object (its reactions). The problem of the DMP consists in being able, while varying the values of the input signals and observing the resulting change of the output signals, to say something about f, i.e., the description of the functioning of the object.

In psychology such an approach to the behavior of a man carries the name *behaviorism*. It has justly been subjected to criticism, for it reduces the whole process of problem solving to a method of trials and errors with a gradual accumulation of probability information about the expedient behavior in a given situation. Nevertheless a number of control systems can be constructed on the basis of this very crude theory. We shall limit ourselves by just one example to illustrate this statement.

Example 1.1. Consider an object having one output channel and two input channels x_1 and x_2. To the inputs x_1 and x_2 may be supplied signals of two types 1 and 0, which corresponds to the stimulation or non-stimulation of the respective inputs. At the output may appear also two signals, 0 and 1, which corresponds to non-stimulation and stimulation of the output of the object. With a fixed excited input, the case when both inputs are excited is excluded; the value of the signal at the output is determined by the distribution (P_i; $1-P_i$), where $i = 1,2$. In other words, with the input x_i excited, a signal 0 will appear at the output of the object with a probability P_i, while a signal 1 will appear with a probability $1-P_i$. The task of the DMP consists in designing such a control system which would minimize the cases of non-excitation of the object's output.

If the values of P_i were known to the DMP, then the matter would be very simple. Let, for example, for x_i the distribution be (0.9; 0.1), and for x_2 of the form (0.01; 0.99). Then it is perfectly clear, that it is necessary to always excite input x_2 and only it. In so doing the cases of non-excitation of the object will be reduced to a minimum. But in the case of a "black box" the information about distributions (P_i; $1-P_i$) is not known a priori. One must try to construct such a control system which would for any initially unknown distribution achieve success by minimizing the number of non-excitations of the object. How this can be done, is shown on the scheme given in Figure 1.2. As the control system is used a "finite automaton of the probabilistic type". It works as follows. In all states of the left-hand group the automaton issues a signal x_1 to the object, and in all states of the right-hand group -- a signal x_2. A change of the states of the automaton takes place with the aid of analysis of the signal issuing from the object after supplying it one or another input signal. If this signal is equal to 0, there occurs a change of the current state of the automaton following the arrow shown on the figure by a dashed line or else the current state is retained as shown by the loop-shaped dashed curve. The choice of one or another occurs with equal probability. If the output signal from the object is 1, then there always occurs a change of the current state of the automaton in accordance with the solid arrow.

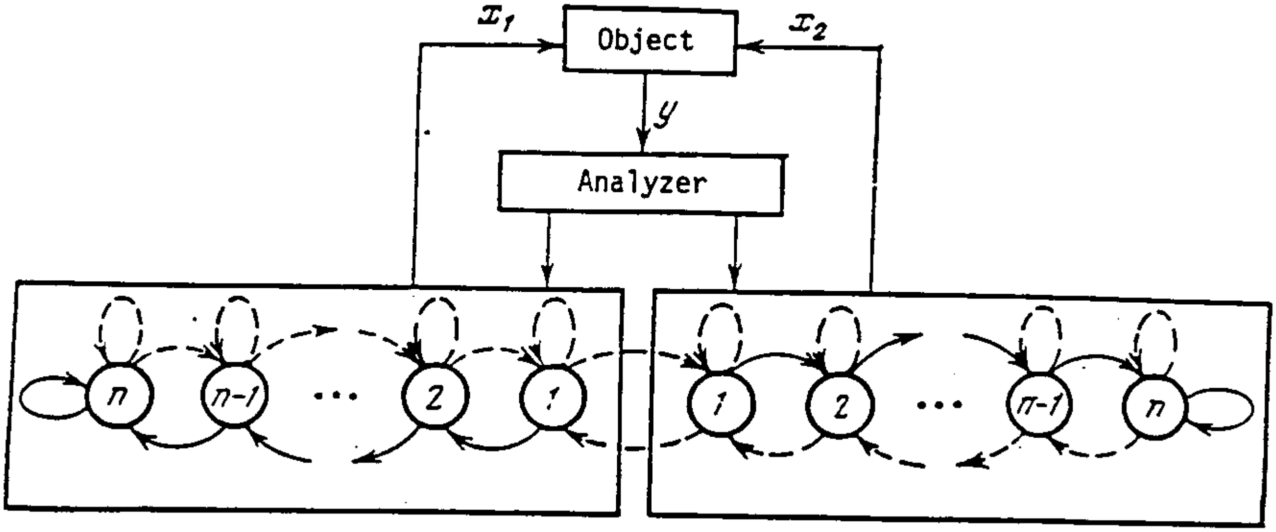


Figure 1.2

It is possible to prove strictly, that such an automaton, with increase in the depth of memory n, will, according to DMP's opinion, tend asymptotically toward the best possible functioning. But even without a strict proof this can be seen from the structure of the automaton. Let, for example, the object be described by distributions cited earlier: (0.9; 0.1) and (0.01; 0.99). And let our automaton at the initial moment be in the state 2 of the left-hand group. It issues a signal x_1 to the object. With a probability 0.9 in response a signal will be received that the object has not become excited. Then, with a probability 0.5 the automaton will remain in state 2 of the left-hand group, while with a probability 0.5 will go over into state 1 of the same group. In this manner the state 2 in the left-hand group will be preserved with a probability 0.45, while with the same probability the state 1 originates in the left-hand group, and only with a probability 0.1 the state 3 in the left-hand group. In the next cycle of operation the automaton again will issue signal x, and again will receive the signal about the nonexitation of the object with a probability 0.9. If it was in state 1 of the left-hand group, then this state will be retained with a probability 0.45, while the transition of the automaton to state 1 of the right-hand group will occur with a probability of 0.45. Reasoning for the other states of the left-hand group, in which the automaton may find itself, is similar. After two cycles of operation from the initial state (state 2 of the left-hand group) its next consecutive state will be, with a probability of 0.2025, the state 1 of the right-hand group, or, with a probability of 0.405. then state 1 of the left-hand group, or, with a probability of 0.2025, the state 2 of the left-hand group, or with a probability of 0.09, the state 3 of the left-hand group, or, with a probability of 0.01, the state 4 of the left-hand group. It is easy to calculate that with passage of time the probability of going deep into the left-hand group will continually decrease, while the probability of transferring to the right-hand group will increase.

But as soon as the automaton goes over to the right-hand group and starts to issue signal X_2, the situation will change sharply. Now, with a probability of 0.99 at each cycle of operation, it will move into the depth of states of this group and only with a probability of 0.005 will it retain its current state or change it by moving to the left. It is obvious that with passage of time, the automaton is without fail to get into the right group of states and, with large n, will practically never learn it. In so doing, it will find the correct methods of controlling the object solely on the basis of knowing the output signals, knowing nothing apriori about the internal structure of the object.

In spite of such an effective example, it should still be noted that the stimulus-reaction theory is too weak both for explaining complex forms of human behavior in solving problems as well as for the use of such an approach in control systems of complex objects. Even there, where initially it appeared that models of this type might lead to success, in the theory of pattern recognition (the most popular model of the stimulus-reaction type in this area is the perceptron), it turned out that absence of structurization of the problem (the object) invariably leads to insurmountable difficulties. Having for instance taught the perceptron to stably distinguish letter A from letter B and then showing it the combination AB we will puzzle it completely. For it cannot take apart the combination AB into the A and B with which it is familiar. This combination appears to it as an absolutely new object in no way connected with the earlier assimilated ones.

Another idea that has received considerable development in models of problem solving by man and which exerted substantial influence on the development of heuristic programming for ECM, bears the name "labyrinth theory". According to this theory there exists in front of the man who makes a decision a kind of labyrinth of possible paths. Using certain local criteria, he selects this on that continuation of the motion in a labyrinth of possibilities. The peculiarity of the labyrinth model consists in the fact that the man sees the labyrinth not as a whole, but only in a certain fixed neighborhood of the platform where he is located.

As an illustration of this may serve situations that take shape in various *positional games*. Let a certain chess-game be played out. Any position, that forms itself in the process of the play on the board, is a platform of a certain labyrinth, the corridors of which are all the possible moves, permitted by the game. And so, we find ourselves on a certain platform. The player who is about to make the next move, has the possibility of selecting any move (corridor of the labyrinth), allowed in the given position. If he could see the labyrinth in a birds eye view, then he could mark out a sequence of moves leading to a mate position or to a draw. But for man this is excluded. He can mentally extrapolate the development of the party only a few moves in advance, considering possible reply moves of the opponent. And this means that he can analyze only a certain part of the labyrinth, a certain neighborhood of the platform on which he presently is located. And he must make a decision on the basis of this local information. Therefore the rules by which the player is guiding himself in selecting his moves are inexact, heuristic. His selection will not certainly lead him to a position which brings victory in the game closer to him. He may permit an error and not consider something outside the analysable vicinity of the labyrinth.

An exact solution of the labyrinth model is achieved only then, when it becomes possible to analyze the entire labyrinth. An example of such a position is the game of crosses and zeros (tic-tac-toe) in a 3 by 3 field. It is not without reason that little kids are attracted by it. They have not yet realized that it is possible, apriori, to analyse all paths of the development of the game and to always win if the opponent permitted an error. With errorless playing the game of tic-tac-toe always ends in a stalemate.

Small finite labyrinths lead to a model of solution of the problem initiated by finite automata. Any platform of the labyrinth corresponds to some state of the automaton, and the corridors -- to transitions of the automatons from state to state under the influence of an input signal and with consideration of the state (the platform of the labyrinth where the automaton is at the moment). The input signal imitates those decisions which are made by the DMP in selecting a corridor of the labyrinth. In so doing the DMP can create a control system of the object also in the form of certain finite (deterministic on probabilistic) automaton. Such a control system is shown on Figure 1.3.

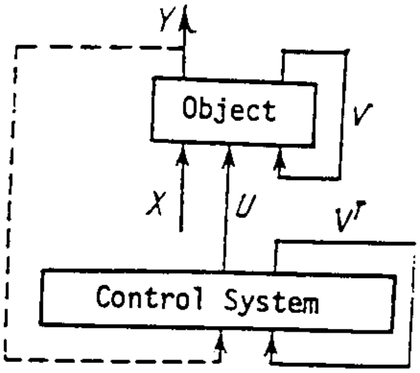


Figure 1.3

On it, V and V^1 reflect feedbacks characterizing the memory of the automata which imitate the object of control and the control system. If the DMP knows completely the structure of the automaton which imitates the object (i.e. known the automaton's graph of changes of state and of formation of output signals), then the feedback connection from the object to the control system is not needed. Therefore it is shown by a dashed line. The need in it occurs only then, when in the process of work it is necessary to carry out a *tuning* of the automaton which imitates the control systems, because of an incomplete apriori knowledge about the object of control. Compared with the system on Figure 1.1, there are no inputs characterized by vector W. This corresponds to the fact that in the labyrinth theory of behavior it is assumed that the effect of these ummeasurable parameters can be described by a certain probabilistic scheme of connection between vector (X, W) and vector Y.

Example 1.2. Let us illustrate the scheme shown in Figure 1.3 for a concrete example. For this we will select the simplest system of a conditioned reflex with two exciters: unconditional X_1 and conditional X_2. As is known, upon arrival at the input of the scheme of an unconditional exciter, a certain fixed reaction must be put out -- the signal y. Upon admitting to the input of the scheme of only a conditional exciter, the untrained scheme should not put out a signal y. Let, during a certain number of cycles (a certain time), the input of the scheme be given simultaneously two signals: x_1 and x_2. Then, after a certain time, a conditional reflex must occur. It consists in the fact that upon issuance to the input of the scheme of only the signal x_2 it will still put out a signal y, i.e., it reacts to a conditional exciter as the dog did in the experiments of I.P. Pavlov, which after a prolonged feeding with simultaneous bell ringing begins to secrete stomach juice in this presence of a bell. If now for a certain number of cycles only signal x_2 is fed into the input of the scheme, then the phenomenon of damping of the conditional reflex will occur. After a certain number of cycles the reflex must disappear and the moment will come when feeding x_2 to the input will not cause a signal y at the output.

On Figure 1.4 is shown a scheme of a probabalistic finite automaton which initiates the process described. The automaton consists of three blocks: a logic block LB, a magazine memory and a sensor, at the output of which occurs the signal ξ which takes on the values -1, 0, 1 with probabilities of 1/3. The automaton works as follows. At the initial moment the memory-magazine is empty. Let the depth of the magazine be n. If only signal x_1 is fed to the input of the automaton, then at the output of the LB a signal y appears and the contents of the magazine does not change. The sensor does not operate under the effect of x_1 alone. If the automaton

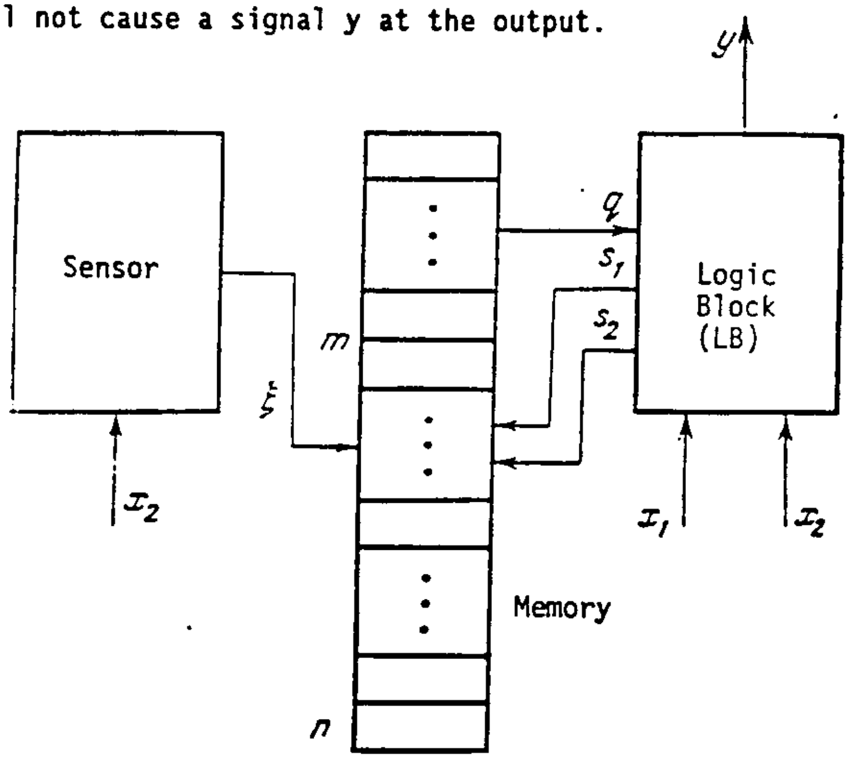


Figure 1.4

input gets two signals, x_1 and x_2, then the LB puts out a signal y. Moreover, the magazine receives the signal s_1, which causes the appearance of unity at the very lowest level of the magazine. On repeating the combination of signals x_1 and x_2, the signal s_1, moves this unity each time by one cell of the magazine upwards. During the combination of signals x_1 and x_2 the sensor of random signals is also triggered. For $\xi = 0$ it exerts no effect on the recording in the magazine. For $\xi = 1$ it puts out a signal that forces the unity in the magazine to move by another cell upwards. Finally, for $\xi = -1$ the signal from the sensor forces the unity in the magazine to move by one cell downward or to disappear from the magazine, if that unity was then at its lowest level. On acceptance at the input of the automaton of only the x_2 signal, the following occurs. If the unity in the magazine is higher than a certain level m, then the magazine produces a signal q, which forms a signal y at the output of the LB. If however, the level m in the magazine has not yet been reached, then signal q is not put out and there is no signal y at the output of the automaton. Moreover, in the presence of only x_2, the LB puts out into the magazine a signal s_2 that forces the unity in the magazine to go down by one cell. The sensor, when only x_2 is present, operates in the same way as in the presence of both x_1 and x_2.

Thus, the position of the unity in the magazine characterizes the degree of learning of the automaton with respect to the conditional reflex. The level m is the threshold of the reflex. By changing it, it is possible to imitate different predispostions of the system to the establishment of the reflex. The sensor brings in the needed probabilistic component. But since the mathematical expectation ξ is equal to zero, it does not bring any systematic distortion into the process.

As it follows from Figure 1.3, we assume that the imitating automaton and the object are described by similar (to the precision of probabalistic realization) processes of functioning. In other words, this scheme is sort of reversible. And the experimental animal can consider itself as the experimenter who establishes the conditioned reflex for the true experimenter (DMP)*. This situation is well reflected in the following anecdote.

---

* Decision-making person

In a cage there are two experimental monkeys. Into the room comes a scientific worker who experiments with the monkeys. At this moment one monkey says to the other: *Look what I have learned. I shall now pull this cord, the bell will ring and this funny man will give me a banana. True, this didn't happen all at once. It was necessary to work with him for some time.*

But in the majority of practically interesting cases the DMP does not attempt to have the control system imitate the object of control. In this there is no need. As a rule the control system can be made simpler than the object which it controls.

For example it is not difficult to show, that with complication of the object of imitation - with transition from the simplist system of conditional reflex to chains of such reflexes and networks of conditional reflexes, capable of modeling most complex behavioral actions, - the system of control which reproduces their peculiarities will remain the same finite automator that was used in example 1.2.

But what is to be done if the labyrinth turns out to be too big? Or completely infinite. For example, we are playing tic-tac-toe on an unlimited field with the condition that 1 crosses or zeros (1 is chosen by mutual agreement) in horizontal or vertical rows or on a diagonal result in winning. Such a labyrinth cannot be mapped onto a scheme of a finite automaton. In this case it is possible to utilize a hypothetical device easily imitated on an ECM* which bears the name "General Problem Solver" (GPS). Its functioning occurs in the following manner. Let there be a certain initial area of the labyrinth. This area (the current situation) can be described somehow. If once more the chess game is used as an example, then the description of the initial area may comprise, for example, of an enumeration of all the figures present on the board, with an indication of the fields on which they are located. A description of the final area will also be a description of some later position on the chess board. And the objective (goal) consists in searching such a sequence of moves which gradually transforms the initial description into the desired one (goal). For transformation of descriptions special operators are used from a certain fixed list stored in the GPS. The condition of their use is formed on the basis of a system of differences and

---

* Electronic Computing Machine

priorities. Each two non-coincident descriptions are in someway distinct from one another. This is fixed in the GPS in the form of a certain type of difference. For example, in a chess game one position can differ from another by the selection of the figures on the board or their position on the playing fields. The priority of differences defines the local rules of selection. They define which differences must be removed in the first order. Such priority of distinction allows the GPS, at each step if its operation, to select for application those operations (to select the corridors in the labyrinth) which removes the most priority requiring differences. With several operators being able to remove a given difference, the GPS may be also given a selection of priorities for the operators. Let us illustrate the operation of a GPS by an example which we shall need later on also.

Example 1.3. The 5 game is a truncated variant of the 15 game which many readers have enjoyed probably. The game takes place on a 2 x 3 field (Fig. 1.5), on which are placed five chips numbered from 1 to 5. One field remains free. It is required to move the chips from an initial disposition (for example, the one shown in Fig. 1.5) to a certain, initially prescribed position (for example one in which the chips are arranged in order, and the right lower cell is free). Let's now introduce the differences among the positions of this game and the operators for removing these differences. It may, for example be assumed that the positions have an elementary difference if in the corresponding cells of the field different chips are located, and to consider a difference of two positions the sum of elementary differences by which they are characterized. Then the position shown in Fig. 1.5 will differ from the goal position by 4 units, because the chips numbered 4 and 5 are already standing on the places which are specified for the goal position. But such a definition of differences is not very covenient for solving the problem posed. If, for example, in the position shown on Figure 1.5, the chips numbered 3 and 2 are interchanged, then the difference is retained, but there is no certainty that the initial position and the newly obtained one are "equally distant" (in the sense if number of transformations needed) from the goal position. It's more convenient to introduce a difference via the summary number of transpositions

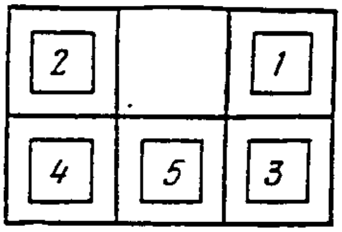


Figure 1.5

needed to transform the current situation into the desired one. (Transposition - that is the switching of two adjoining chips or the move of a chip into a vacant field). For example, in order that chip number 2 would be where required, one transposition is required - its move by one field to the right. For chip number 1, two transposition are required, and for chip number 3 - one. As already noted, chips 4 and 5 are already in place. Then the summary difference of the initial position from the goal one will be equal to 4 units. But if chips numbered 2 and 3 are switched, then the difference in transpositions will be a different one (in distinction from what was the case with a different definition of difference cited earlier). With such a change, the difference will equal 6 units. In other words, the newly obtained position will differ from the goal one by a greater distance than the initial one.

To remove the differences in the game above it is possible to use operators of four types: to move to the free field the lower, the upper, the left and the right chips. In a real position not all the operators can be used. With a free field at the center - three of the four, and with a free field in a corner - two of the four. The selection of one or another operators is determined by their priority. The priority, in turn, is defined by how much the given operator decreases the differences. For example, in the position shown in Fig. 1.5, there is no sense to move the chip numbered 5 (to use operator *move the lower chip to the free field* because this will result in increasing the difference. However, the use of operators: *move the left chip to the free field*, and *move the right chip to the free field* are equally rightfull here. The use of either of them will decrease the difference of the given position from the goal one by one unit.

The example 1.3 demonstrates one important characterstic of the GPS. The solution plans which GPS builds have the property of monotony. At each step of the solution an improvement takes place, a closer approach to the target. It is precisely this property that proved to be fatal for the GPS and for similar programs and systems. It is possible to cite the following analogy from mathematics. The gradient method allows one to find an extension of a function only under the condition that this function has one extremum. Then a monotonic motion toward the extremum point is possible. But in most cases a function has several extremums. Then the gradient method, which successively *improves* the value of a function, can lead us to any local extremum which will become a trap for us. Wherever we would move from the given point with a given step of the gradient, failure awaits us, for the values of the function in the entire accessible neighborhood are *worse* than the one at the point where we are.

For finding a global extremum a few methods were invented in the theory of optimization: changing the gradient step, motion *along a valley*, etc. In systems of the GPS type some techniques of improving the solution plans were also devised. For example, the transition from the original labyrinth to a more simple, more crude one. Planning the motion first along it and then making this plan more precise on the original labyrinth, so that with monotony of the crude plan the more precise plan may become non-monotonic. But unfortunately so far, it has not succeeded to find good procedures for constructing the crude plan. And as long as this has not been done, it is not possible to hope to build good models of decision making.

But the labyrinth theory has a far more important defect than the property of monotony which is inherent in practically all procedures of motion along a labyrinth. This defect consists in the apriori specification of the labyrinth or methods of its construction. But from where is this information derived. In the terms adopted in paragraph 1.1 this means that the DMP knows f and he needs only to build the procedure of control. But we have already stated that the creation of a description of the object of control is a problem not only more complex than the finding of the control procedure itself, but has not as yet any kind of standard methods of solution.

In solving various problems, we always encounter two cases in making decisions. Either we are facing a problem which, in principle, we are able to solve and all that is needed is to find a solution of the given specific problem, or else we encounter a completely unclear problem for which it is not even known from where to begin. Here a child is trying to construct from playing blocks that which is shown on a picture. If the box contains only six blocks on whose sides something is pictured, then his labyrinth is very large - it is possible to put together more than 48 thousand combinations. But the little child manages to cope with such a huge labyrinth. For he does not arrange the blocks chaotically, but relates then to the end platform of the labyrinth, by selecting for each block only one true side out of six. In this problem the labyrinth is evident. Once having learned to put together something, the child carries out with assurance a similar task for other blocks, even if there are many more than six. But now, there is another problem before him.

Let us refer to Figure 1.6 which shows the map of an island. On it are marked the towns and the railroads connecting them. Mister Brown lives in town A, located in the north of the island. During his vacation he posed before himself the following problem: to travel from town A to town B without making use of paved highways in his car, and passing only once through all the towns shown on the map.

Is this problem solvable? The labyrinth is before you. Try it before you read on.

Now it has become clear to you that a solution apparently does not exist, yet Mr. Brown solved the problem he posed to himself. How did he do it? Apparently he has heard something about the labyrinth and the model theories of thinking, but most likely he managed without them, not knowing that he *speaks in prose*. He knew that the problem is solvable. For this, it is sufficient to utilize the fact that between the towns B and Щ it is possible to use the sea route, which is not prohibited by the statement of the problem. Then the problem becomes solvable. If you yourself have guessed of such a solution, then you have experienced insight (enlightenment), and a feeling of happy astonishment over the discovered solution. In so doing you have

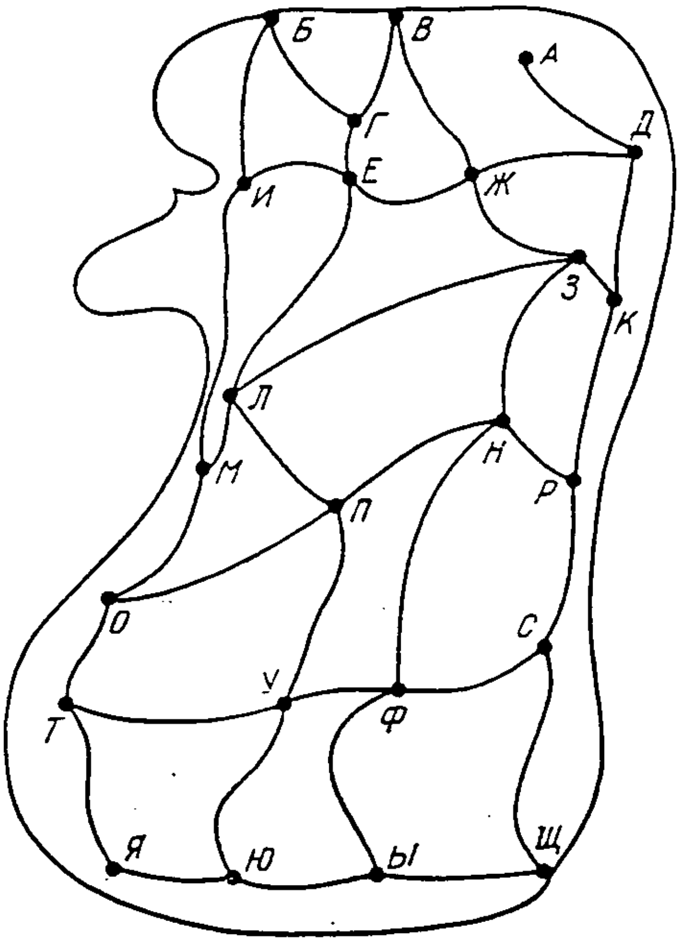


Figure 1.6

transferred from the imposed labyrinth of Figure 1.6 to a new one, built by yourself.

In the problem above, the prompt about the new labyrinth was contained in the map drawing. The towns B and Щ are located on the sea coast. But such a prompt may not be available. Such a situation exists in a popular psychological puzzle: *without breaking or bending the matches, put together, from six matches, four equal equilateral triangles.* In the puzzle all six matches are laid on the table in front of the experimenter, which specifies a labyrinth of all-possible positions of the matches is the plane of the table. The tested people are divided into two groups. Those who belong to the first group, having moved the matches around and being convinced that the problem is not solvable (in the given labyrinth!) refuse to solve it, thinking that they simply have been cheated. The tested people of the second group attempt to clarify the conditions under which the solution of the problem is possible. And they are rewarded for their efforts by constructing a new labyrinth of possibilities - a spatial one. In it they easily solve the posed experimental problem by constructing a regular triangular pyramid.

It is possible to assume that in a certain measure the labyrinth model describes the behavior of a man during problem solving, when he uses a standard pattern found earlier by himself or given to him by a teacher, or derived from books and study texts.

And so, in addition to searching all over the labyrinth for possibilities for the man during solution of non-trivial problems, most important is the procedure of the search of the labyrinth itself (or of its part), in which it is possible to find the needed solution. In the process of problem solving he must form this labyrinth and already after that seek a solution. It is precisely the finding of such a labyrinth that causes within a man a flow of positive emotions, the *joy of discovery*. The postulate formulated by us about the two-stage nature of solving any problem is the cornerstone of the third psychological theory of solving problems - the "model theory".

In the framework of this theory any solution of a certain problem consists of the sequence of several steps: description of the starting position of the problem, description of the goal position of the problem, the establishment of homomorphism between these descriptions or their reduction to a common language of description, establishment of a system of transformation of the descriptions, a search of succession of the transformations which lead from the original position to the final. Only the last step is reflected in the labyrinth theory of problem solving, while the first four steps in it are considered as being already realized.

Let us consider the essence of the first four steps. At the stage of describing the initial position it is possible to act in various ways. For an illustration of this let us again turn to chess. In some chess positions it is possible, as was explained earlier, to describe the position of every figure or the playing field. And this description will be complete and exhaustive. But the chess player will hardly use it for finding his move. It is for him too remote from the description of the goal position. Indeed, the final position in a chess game is not described in the terms of disposition of actual figures on the fields of the board. It is described in other terms, of the type "The king of the opponent is under threat, cannot leave the cell in which he is, in order to move over to a cell that is not under threat, and cannot abolish, using other figures, the threat of having the king taken". This means, that the description of the initial position which we adopted, and the description of the final position (the description of the fool's mate position is made in the same time as the description of mate) are not homomorphous to each other. They do not lie within the limits of the same labyrinth. And it is precisely this that constitutes the psychological difficulty of solving chess problems for inexperienced chess players. Experienced chess players, on the other hand, describe the initial position in the language of the same level as the language used for describing the concluding position. There, instead of individual figures and moves appear connections, forkings, doubled and passing pawns, forced moves and much other.

In chess there is not need to search for a system of transformation of descriptions. It is defined by the rules of the game. But in many cases it is necessary to do so. At competition of pastry bakeries, who must prepare a torte from a selection of components common for all, each participant of the competition formulates his own system of *transformation of positions*. And depending on how successful he does this, so successful will be the result of his work.

Ending this excursion into the psychology of problem solving by man, let us pause on one more question important to us. Up to now we have considered solutions of problems by a man in static conditions, outside of time and space. But for real objects of control such a limitation is hardly feasible, as a rule, decisions on control are adopted with allowance for the dynamics of functioning of the object of control. What can psychology tell us in this connection?

During observation of the behavior of humans occupied in controlling dynamic objects, psychologists have classified certain principles, by which humans guide themselves in making their decisions. And the main conclusion is, that in solving administration problems, in which the dynamics of the processes in the object is accounted for, the man builds their dynamic model. They sort of occur in his brain, retaining the relative time scale of the parts of the processes.

Depending on the *remoteness* of the controller from the process of control, the dynamic model can be different. Whereas the man on duty at a railroad marshalling yard, who controls the transportation of locomotives and wagons with the aid of railroad switches, can see the entire spatial picture of the transportation directly, the assistant station master perceives it with the aid of a mnemo scheme (mimic panel), while the train dispatcher - only by the graph of train traffic, drawn from discrete reports sent to him. In this manner the dynamic model of the process at the marshalling yard does not require of the attendant any special methods of representation except those related to direct observation over the real state of things. The assistant station master can observe and estimate directly only what is visible from the window of the room in which he is. The rest of the information is indirect, it is reflected on the mnemo scheme (mimic panel). For this reason the development of processes takes place partially in his head. He is forced to interpolate and extrapolate this development. To the greatest degree the interpolation and extrapolation of the processes takes place in the head of the train dispatcher or duty man for a division of the railroad. In his head the process of transposition of trains on the area which he controls is sort of reflected in the real time scale. With experienced duty persons and dispatchers this flow of processes at the object of control reflects itself sort of *by itself*, outside of active consciousness. Word communications arriving at the dispatcher or duty man are recorded by him to the level of representation, characteristic for an imitational model which operates in his head.

1.4 Of What Does Situational Control Consist?

And so, in working with objects which the theory of control has presently bumped into, the possibility of utilizing traditional methods and techniques for controlling them can apparently not be hoped for. As has already been made clear, the problem is in the description of the unique object of control, of the consideration in this description of not only its specific structure and functioning, but also of the behavior of the people, and of the possibility of evolution of the object in time. It is required to create such an approach which would give, in a single language, the possibility of describing both the object of control and its functioning, as well as the procedure of controlling it.

Let us first consider the formulation of the control problem, within the scheme of which are fitted in all the presently known problems of control of complex objects.

Definition 1.1. Let us define as the *current situation* at the object of control the totality of all knowledge on the structure of the object of control and of its functioning at the given moment of time.

Definition 1.2. Let us define as the *complete situation* the totality consisting of the current situation, the knowledge or the state of the system of control at the given moment and the knowledge or the technology of control.

In these two definitions, much requires further precision (elaboration). But at the conceptual level of presentation, it appears to be sufficient to have the same understanding of them as was given above. Let us denote the complete situation by S_i (where i is the distinctive number of the situation), and denote the current situation by Q_j (where j is the distinctive number of the situation). Let there be available, to the control system, n different methods of affecting the object of control (single-step solutions). Each such solution we shall denote as U_k (where k is the distinctive number of the effect). The elementary act of control can be represented in the following form

$$S_i \; ; \; Q_j \underset{U_k}{\Rightarrow} Q_1$$

The meaning of this relationship is as follows. If, at the object of control, a situation Q_j has arisen, and the state of the control system and the technological control scheme defined by S_i can utilize the effect U_k, then it is utilized, and the current situation Q_j is transformed into a new situation Q_1. Similar rules of transformation will be henceforth referred to as *Logico-Transformational Rules (LTR)* or *correlational rules*. The complete list of LTR's defines the possibilities of the control system affecting the processes occurring in the object.

Evidently, because of the finiteness of the number of different influences the entire set of possible complete situations will somehow fall into (n) classes, to each of which will correspond one of the possible influences (effects) on the object of control. In other words, there must exist such procedures, which would allow us to classify the complete situations in such a way that it would be possible to form from them as many classes as there are single-step solutions at the disposal of the control system. These procedures may be called *Procedures of Classification*. If, for some complete situations, it is impossible, because of a not too good knowledge of both the object as well as the effect of the influence on it, to point out a single one-step solution, then it is allowable for the present to include this situation into several classes.

But because of such an overlap of classes there will arise the problem of selection of one or another solution out of the number of possible ones for the given complete situation. Special *Procedures of Extrapolation* of the after effects of making one or another decision are needed to make such a selection. With their help it is possible, on the basis of the knowledge of the object of control and its functioning, to estimate ahead of time the results of applying the selected influence and to compare the obtained prognoses for all possible effects for the given situation.

If all the procedures above could be constructed, then the general scheme of the solution of the control problem would appear as is shown on Figure 1.7.

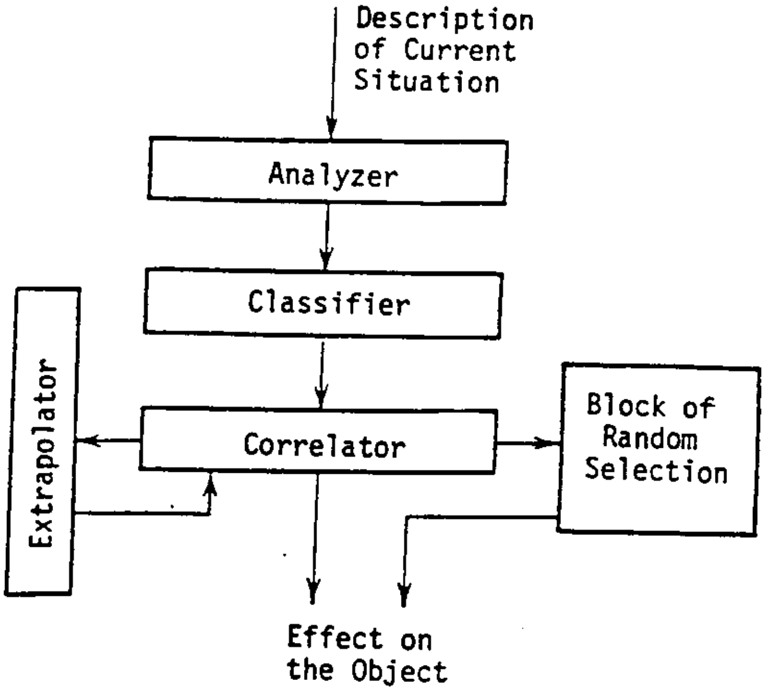


Figure 1.7

The description of a current situation which exists at the object of control, is supplied to the input of the "Analyzer". Its task consists in evaluating the messages and determining the need for the control system to interfere with the process occurring in the object of control. If the current situation does not require such interference, then the Analyzer does not transmit it for further processing. In the reverse case the description of the current situation enters into the "Classifier". Using the information stored in it, the Classifier relates the current situation to one or several classes, to which correspond single-step solutions. This information is transmitted to the "Correlator", in which are stored all LTR (Logico-transformational Rules). The Correlator determines the particular LTR which must be used. If this rule is the only one, it is issued for execution. If there are several rules however, then the selection of the best of them is carried out after the processing of preliminary solutions in the "Extrapolator", after which the Correlator gives out the solution of action upon the object. If neither the Correlator nor the Classifier are able to make a decision from the incoming description of the current situation, then the "Block of Random Selection" is activated and a selection is made of one of the influences that cause a not too great effect on the object, or else the system refuses to take any kind of action on the object. This says that the control system does not dispose of the necessary information about its behavior in the given situation.

In fact, because of the complexity of objects of control with which we are occupied, there is no hope that the initial knowledge about them and about methods of their control would be sufficiently complete. For this reason a control system of this type should in principle be an open system. It must have the possibility to correct its knowledge about the object and the method of controlling it. In the operation of such a control system there so to speak are two stages: the stage of learning and training and the stage of working. During the initial period, when the control system is just being created, a large amount of data are collected from technologists, who know well the object of control. With their help are formed the classes of situations and LTR. They determine the expediency of utilization of whatever influences on the object of control under one or another situation. Here the opinion of the experts may not coincide, which results in putting the identical situations under different classes with respect to control. With the aid of experts are also formed the procedures of extrapolation, methods of evaluating situations, resulting from the desired functioning of the object of control.

This means, that at the first stage the design-project of the system receives a considerable amount of subjective information about the object of control and the procedures controlling it. It is, so to speak, the collective experience of those people who, good or bad, were able to cope with their responsibilities while controlling the object. Here it is important, that in designing the system it is desirable to allow for the most insignificant, on first sight, information about the object. It frequently happens that precisely this peripheral knowledge plays a great role in the making of decisions on object control.

Following the stage of knowledge accumulation and formation of procedure in blocks, shown in Figure 1.7, the system can begin to operate. But is the process of its exploitation, especially in the early stages, it will either often make wrong decisions because of incompleteness of information and inexactness of procedures. In all these cases it is necessary to have a stage of make-up training of the control system.

Let us note our important characteristic of the systems of the class considered. The number of different complete situations is usually very great. Possibly they cannot even be enumerated ahead of time. Whereas the number of solutions is finite and not great. In other words, most frequently $|\{S_i\}|>>|\{U_k\}|$, where, as always, $|A|$ denotes the number of elements entering into set A. When steering a car, the driver perceives a huge number of complete situations, determined by external conditions on the drive path and by the condition of the automobile. Whereas the number of single-step decisions made by him is not large. The traffic regulator at a cross-roads reduces the entire set of situations in the vicinity of the cross-roads to two single-step decisions by switching the traffic light in a certain fixed direction. A secretary who is in the reception room of a big boss, instantaneously estimates the complete situation and makes one of three decisions, expressed usually in the standard phrases, *Please go in*, *Wait, he is busy now*, or *Unfortunately, he is not in and will not be in today.*

The greater is the number of possible situations and the less is the number of allowable single-step solutions, the more efficient does appear to be the operation of the control scheme, which we have shown in Figure 1.7. But even when $|\{S_i\}| \approx \{U_k\}|$, it is possible to utilize a similar scheme. Although its efficiency in that case will not be very high, still, possibly, this is the only way of controlling the object.

The control method, which we have described in general terms, is based on the introduction of the concept of situation, the classification of situations, and of there transformation. It is natural therefore to call it the method of situational control:, for, to counterbalance the epigraph [at the beginning] of the present chapter, in the Russian language the word, defining the essence of the described approach, is actually the only one. And only the great interest toward methods of control of this type has recently given birth to a series of names of methods, partially more wide, partially more error compared to that which is included into the concept of situational control (semiotic modelling, semiotic control, logico-linguistics methods of control, and so forth).

Let us now formulate a number of characteristics inherent in the methods of situational control.

1. Situational control demands great expenditures for the reaction of a preliminary base of data about the object of control, its functioning, and methods of controlling it. These expenditures are justified only when traditional ways of formalizing the description of the object of control and procedure of control are impossible to realize. In other words, if the object of control is such, that it is adequately described, for example, by a system of linear differential equations of the first degree with constant coefficient, then there is no need to utilize the method of situational control. Its use is justified only when the transitional formalization results in a problem of such dimensions, that its practical solution by known methods is impossible. An example being the case when the number of equations comprises several tens of thousands.

2. The description of situations which are formed at the object of control (current situations) should be carried out in and a language, in which would be reflected all basic parameters and relations needed for the classification of this description and the juxtaposition to it of a single-stage solution or control. Then it is necessary to correctly select the level of description which should be neither too detailed nor too crude. With a too detailed description a "noise effect" occurs, and particulars and facts that are not essential for control can considerably complicate the understanding of the essence of functioning of the object and can make the construction of the control system impossible. This case can be illustrated by the following example. Somebody, who has never seen a game of football describes his impressions of the game to another person who also does not have any conception of this game, in the following manner: *Twenty two people eleven of which are dressed in one uniform with numbers on the back, while eleven others in a different uniform, also with numbers on their back, move about on a field, which has the following characteristics (then follows an exact and detailed description of the field and its markings). Each of the players may move along the following trajectories, apparently related to his numbers (then follows a formal description of the dynamics of motion of the players on the field). Sometimes, simultaneously with the transposition of the players, a normal object is transposed (then follows a description of the ball), but with a certain probability there occurs a transfer of this object from one player to another of else its expulsion from the game field (then follows a*

description of a probabilistics process of ball transfer). In the latter case, special players return the ball to the field along certain trajectories (then follows a description of throwing the ball back in). And so forth. The question arises: could the listener discern, in this scientific exposition which describes sufficiently accurately the match, the essence of the football game? I think it would be very difficult. And what if the football match were to be described at the level of the language of statistical physics (this is also possible in principle), then what could be obtained from such a description?

3. The language used for description of situations must allow one to reflect not only the quantification of facts and relations that characterize the object of control, but also qualitative knowledge which cannot be formalized in the usual mathematical sense. For, the majority of information which the controller will be receiving from the technologist, will have approximately the following appearance: *In the majority of cases I think it is necessary to do X, if Y takes place. It appears to me that in situation X it would be nice to do Z, when X increases, then as a rule, Y decreases, but not too strongly,* and so forth. And it is necessary to learn to reflect these qualitative statements in the language of situational description. It should also be considered, that human statements about the object and methods of controlling it are incomplete. Special methods are needed to obtain from him all the information which is needed for making decisions on control (recall the case with the diagnostics of psychic illnesses, when in fact the histories of the illnesses did not contain any information). The well-known children's problem: *In a moving train consisting of 40 cars, three of which are passenger cars and the rest freight cars, there are three conductors and one guard; it is required to find how old the engineer is . . .* illustrates the above statement.

4. Classification of situations, their unification into classes, while using single-stage solutions, takes place on a subjective basis, since the initial information about the correspondence of one or another current situation to one or another solution is obtained from experts. The system kind of sums up the knowledge of individual experts and becomes a carrier of the collective knowledge of people. However, the procedures of classification must be constructed in such a manner that the classification itself would be

suited for those current situations, on which the system did not receive any information from experts. This causes the classification problem to become similar to the problem of forming concepts on the basis of teaching sequences. The system, having formed a certain concept, possesses already greater knowledge than that which was just put into it at the beginning by the experts, although this additional knowledge may turn out to be untrue, which may reveal itself in the process of its operation. Thus, *preconceived notions* may appear in the system, i.e., untrue concepts and unjustified generalizations. The system may behave itself like a woman, who says during a conversation: *I know you, all men lie.* But to the malicious question: *What is it with you, did you have to do with all the men*, she simple-heartedly replies: *Yes, I was acquainted with one like you....*

5. Analogical words can be said also about the formation of LTR[*]. They also are formed initially with the aid of information, communicated by experts. Precise definition of these rules, lead to the removal of conflicts in them, and the formation of new LTR's takes place in the process of operating the system. All this is also true for rules of extrapolation and for the appraisal of some or other current situations.

6. From the two last points above there follows a principal conclusion that systems of structured control cannot optimize the process of control itself. They are solely oriented toward such control, when the achieved results are not worse than the best results obtained by a human. But, as the practice of utilization of systems of this type has shown, most often the results given out by the system are better than those of humans. This occurs for a number of reasons. In particular, the system is not influenced by the emotions of a human controller, and in stress situations it makes decisions in the same manner as under ordinary conditions; the system forgets nothing and does not leave out solutions which for a human are priori *obviously understood* but did not enter into his head at the right time. Remember, how in the well-known story by A. P. Chekhov: *Would it be a goat's*

---

[*] Logico-transformational rules

leg to me?[*]. In this manner, the method of situational control may be considered to be heuristical.

7. Actually, for many real objects of control the single-step solutions do not define the strategies of control. In such objects it is necessary to form, as solutions, chains out of single-step solutions. For this, in the system of extrapolation, there must be foreseen special procedures of *glueing together* of single-step solutions. With their aid the more complext solutions or control are formed.

The indirected peculiarities are accounted for in the procedures and methods described in subsequent chapters. It should be noted that these methods are not the only ones possible. And any specialist is justified to conceive and suggest his approaches for solving the problems that are derived from the essence of the method of situational control.

It follows from the historical sketch of the development of situational control, given at the end of the book, it had originated much earlier than was formed the now actively developing scientific direction, which is called the theory of *artificial intelligence*. Situational control has much in common with the methods of artificial intelligence. In the course of presentation we shall point out these analogies, which will help the reader to utilize, from the arsenal of the theory of artificial intelligence, not only the terminology, but also those results that were obtained there, to enrich his own systems of control of complex objects.

1.5 Semiotic Models[**]

The goal of this section is to present the concept of a "semiotic model" which lies at the core of situational control and methods of semiotic control in general.

Let's begin with the definition of a formal model.

---

[*] This is some popular saying that is not explained in the dictionaries that are available.

[**] Asterisk refers to topics of the "second level" as explained by the author in the foreword to this book.

Definition 1.3. The term formal model refers to a four-term expression:

$$M = < T, P, A, \Pi >,$$

where T = set of basic elements, P = syntactic rules, A = system of axioms, Π = semantic rules.

Let's explain the essence of the elements that comprise a formal model. The set T consists of a finite or countable number of elements of any nature. It is sort of an alphabet of a formal system, a selection of elements, from which will be constructed all the other of its elements. No limitations are imposed on the elements of a formal system. It is important only that any two elements of T are in some way different from each other. In other words, there must exist such a construction (completed after a finite number of steps) procedure π, so that its application to the elements of set T gives an answer to the question: Are the two compared elements equal or not? The importance of this procedure may be illustrated on the example of set T, consisting of the letters of the Russian alphabet. The identification of the various renditions of the letter A, for example, is not such a simple matter. And sometimes, especially in rapid writing, it is rather difficult to identify, in a word, one or another letter.

Besides, there should exist a construction procedure π_2, which gives the answer to the question: Does the given element belong to T or does it not. For the same illustrative example this corresponds to an answer to the question of belonging of a certain grapheme to the list of letters of the Russian alphabet.

The syntactic rules P are used to construct, from the base elements, such combinations of them, which within the framework of the given formal system are considered *syntactically correct combinations*. There are no special limitations on the set of syntactic rules. The only requirement is the availability of a constructive procedure π_3, which would give a single-valued answer to the problem: Is the given combination of the base elements syntactically correct or not?

The system of axioms (A) is formed by any set of syntactically correct combinations. No special limitations are imposed on this set. In particular, it may coincide with the entire set of syntactically correct combinations.

Finally, the semantic rules (Π) otherwise known as rules of extraction expand the set of axioms, if it is possible, adding to them new, syntactically correct combinations. The set which results after application of semantic rules, carries the application of a set of "semantically correct combinations".

Figure 1.8, shows the result of formal system generation, which acts as an autonomous generator. The procedures of generation of a set of syntacticaly correct combinations, based on an arbitrary application of rules P, and the procedures of generation of a set of semantically correct combinations, based on the application of rules Π, operate independently.

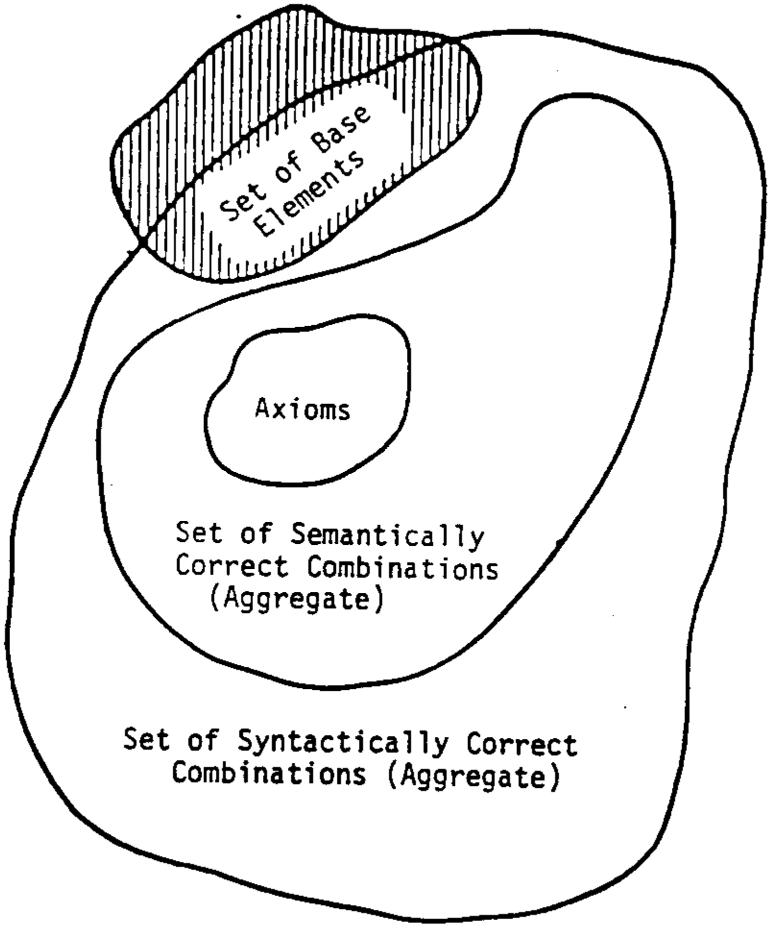


Figure 1.8

Definition 1.4. A formal model for which there exists constructive procedures π_1, π_2, and π_3, is called a "constructive formal model".

Definition 1.5. A constructive formal model, for which there exists a constructive procedure π_4, which gives a single-valued answer to the question - does the given syntactically correct combination belong to the set of semantically correct combinations is called a "solvable formal model".

Let us cite some illustrative examples.

Example 1.4. Lets turn back to the box of children's playing blocks, containing only six blocks. Each block is a base element. It is easy to convince yourself, that there exist π_1, and π_2, that allow to distinguish one block from another (by the non-coinciding picture on at least one edge of two compared cubes) and to distinguish the block from other selections (i.e., by non-coincidence of end points of the lines or boundaries between the colors or all six faces of a given block with all the other blocks of the selection, if there is not a simpler type characteristic of the other dimension of the block). The system of syntactic rules is such that it considers syntactically correct any combination of blocks, in which all six blocks are laid out in the form of a rectangular 2 x 3 or 3 x 2. It is obvious that there exists a construction preceedor π_3. The system of axioms considers with such a combination of blocks, which corresponds to one of the pictures included in the box (this picture is sort of fixed as the initial position). Rule π gives the possibility to obtain, from the initial picture, new pictures (I think, that the readers in their childhood constructed these rules themselves). And, finally, it is clear, that there exists a constructive procedure π_4, for the determination of all semantically correct combinations which is contained in the selection of six pictures, included in the selection. Therefore, we are having to do with a solvable formal model.

Example 1.5. Let's consider classical "propositional calculus". As elements T we shall use lower case Latin letter without indices and with lower indices, and also the symbols ⌉, &, V, and →. As P we shall take the following selection of rules; 1. Each base element is a syntactically correct combination. 2. If α is a syntactically correct combination (formula), then ⌉α also is a formula. 3. If α and β are formulas, then (α V β), (α & β), ($\alpha \rightarrow \beta$) are also formulas. Here α and β are metasymbols, not entering into T. The procedure π_3 obviously exists. Let us exercise, for example, the combination of base elements of the type (⌉ a V b).

This of course, is a formula, for a and b are formulas according to the first of the rules from P, ⌉a is a formula in accord with the second rule from P, and (⌉ a V b) is a formula by the third rule from P. In contrast to the preceeding combination, the combination (⌉ V a) is not a formula, because in P there are no rules, whose consective application would lead to its construction. An explicit expression of procedure π_3 is not given in order to give the reader the pleasure of writing it himself.

The system of axioms has the following appearance:

(1) $(\alpha \rightarrow (\beta - \alpha))$;
(2) $((\alpha \rightarrow (\beta \rightarrow \gamma)) \rightarrow ((\alpha \rightarrow \beta) \rightarrow (\alpha \rightarrow \gamma))$;
(3) $((\alpha \mathbin{\&} \beta) \rightarrow \alpha)$; (4) $((\alpha \mathbin{\&} \beta) \rightarrow \beta)$;
(5) $((\alpha \rightarrow \beta) \rightarrow ((\alpha \rightarrow \gamma) \rightarrow (\alpha \rightarrow (\beta \mathbin{\&} \gamma))))$;
(6) $(\alpha \rightarrow (\alpha \vee \beta))$; (7) $(\beta \rightarrow (\alpha \vee \beta))$;
(8) $((\alpha \rightarrow \gamma) \rightarrow ((\beta \rightarrow \gamma) \rightarrow ((\alpha \vee \beta) \rightarrow \gamma)))$;
(9) $((\alpha \rightarrow \beta) \rightarrow (\neg\beta \rightarrow \neg\alpha))$; (10) $(\alpha - \neg\neg\alpha)$;
(11) $(\neg\neg\, \alpha \rightarrow \alpha)$.

Evidently there are many axioms, and for a reader not accustomed to symbolic logic it is difficult to encompass them by a single glance. But that is not even needed. It is necessary to look at them as a set of randomly selected elements from correct combinations.

Finally, rule $\sqcap$ in the classicial propositional calculus in that form of it, in which it is considered in the present example, have the form: 1. If α is a semantically correct combination (a deducible formula), then in substituting in α any base element (everywhere it enters into α) by any formula, there again results a deducible formula; 2. If $(\alpha \rightarrow \beta)$ and α are deducible formulas, then β also is a deducible formula.

The question on the existence of a constructive procedure π_4, which would allow for any formula of propositional calculus to determine whether it is deducible or not, remains open at the present. But in reasonings which will follow after Definition 1.6, we shall give a positive answer to this question. And then it will become clear to the reader why in the given example precisely such an axiom was chosen.

The example cited, should apparently be sufficient. Let us note only that the calculation of predicators and the formal grammar, widely utilized in programming languages, are constructive formal systems. But in the general case they are not solvable formal systems.

Let us now consider a model of the type

$$L = \langle Z, D, H, V \rangle.$$

Here Z is a certain set, which we shall call a "set of interpreted values". Rules D, refined to as "rules of mapping (transformation)", establish the possibility of the mapping $T \rightleftarrows Z$, which is a multiple sign in both directions. These rules in a certain concrete realization give the mapping $T \rightarrow Z$, of single sign in one direction, i.e., they ascribe to each element of T a certain interpreting value. One or another realization of the mapping is given by H - the "rules of mapping", or else is given from without. Finally, the rules of interpretation V allow to ascribe to any syntactically correct combination of base elements a certain interpretive value, if the interpretive values of all the base elements entering into this combination are given.

Definition 1.6. Formal model M, for which is specified model L, is called an "interpreted formal model".

Example 1.6. For a statement calculation considered in Example 1.5, it is possible to take as a model of interpretation a model in which Z = {Truth, Lie}, the rules of mapping specify an arbitrary mapping of any elements from T, except ⌉, &, V, and → into any element of Z, the rule H are absent, and the rules V are defined by Table 1.1.

TABLE 1.1

a	b	⌉a	a&b	aVb	a→b
Lie	Lie	Truth	Lie	Lie	Truth
Lie	Truth	Truth	Lie	Truth	Truth
Truth	Lie	Lie	Lie	Truth	Lie
Truth	Truth	Lie	Truth	Truth	Truth

These rules allow to ascribe a certain interpretor value for any syntactically correct formula of propositional calculus. It is not difficult to verify that to all axioms enumerated in Example 1.5 these rules ascribe, for any interpretation of the base elements entering into them, the value

"Truth". The semantic rules, given in the same example, cannot change the value of the formulas to which they are applied. In this manner, all formulas which are deduced from the axioms also have always the Interpretor value "Truth". It is possible to prove, that the axioms and the rules of deduction, cited by us, posses the property of completeness in the sense that any formula of propositional calculus which has the value "Truth", is deducible for this set of axioms with the available semantic rules. And this means that the propositional calculus, considered by us, is a solvable formal system. Procedure π_4 consists of checking the fact that, for any interpretation of the base elements of the formula under check, it assumes the value "Truth". If this is not so, then the formula under check does not belong to the set of semantically correct formulas.

In formal systems, T, P, A, and $\sqcap$ remain unchanged. This means that if formal systems were used for creating situational models of control, then the language of description of situations (it is determined by T and P), the initial information about the object of control and laws of control (they are defined by specifying a set of axioms) and LTR* (they coincide with the semantic rule) would remain unchanged and immovable (unshakable , firm, stable). But this contradicts everything we have said about objects of control and the very system of situational control, its principle openness for teaching and additional teaching. A system of control, which is based on a formal model, must apriori have all the information, which remains unaltered for it during the entire period of control. All statements deduced in it at any moment of time, remain deduced forever.

It is clear, that for objects of control, different from the children's blocks in Example 1.4, such a situation does not take place. In a functioning control system, the languages of situation description may be corrected and the knowledge about the object and methods of its control may change. This means, that all elements entering into the definition of M may change in the process of its functioning.

---

* Logico-Tranformational Rules

Let us consider in this connection a model of the form

$$C = < M, \chi_T, \chi_P, \chi_A, \chi_\Pi >$$

Here χ_T, χ_P, χ_A, and χ_Π are respectively, the rules of variation of T, P, A, and Π.

Definition 1.7. Model C is called a *semiotic model*.

Such a name of C is related to the fact that, in distinction from formal models, in which the elements that form set T possess a rigid syntexis specified by procedure π_1, a rigid semantics defined by the interpretation procedure, and a rigid pregmatics (that's how the DMP treats this element from the standpoint of control process), in model C all these properties of elements t_i become accessible for change. But it is precisely this property that is possessed by *signs* - the elements of *sign*, or *semiotic systems*, studied in semiotics. Such systems are closely related to all human activity. It is namely the changeability and conditionality of signs that make this activity effective. Let us explain the above with a simple example. The signal of a bell may mean, that someone must lift the telephone receiver. But the pragmatics of the bell for me may be different depending on whether I am expecting to talk to someone or know that the receiver should be picked up in the next room. Instead of a bell, I may be called to the phone by knocking on the wall. The syntaxis of the sign has changed, but the semantics and pragmatics have remained. And exactly the same kind of bell may signify the end of work, which for an unchanged syntaxis gives a different semantics. And in countries, where a special fee must be paid for each connection of the line, people make an agreement that, if the phone rings four times, then the meeting will take place, but if five - it is delayed. Retaining the syntaxis of telephone calls, they change their semantics and pragmatics.

Man is surrounded by sign systems, his activity is permeated by them. He constantly creates such systems, making agreements with other partners on the syntaxis, semantix, and pragmatics of the signs. The rules of change, introduced by us in model C, reflect, in some measure, a similar process. The rules χ_A and χ_Π change the semantics and pragmatics of combinations. It could also be possible to change the model of interpretation. Then we would obtain a model similar to the sign systems of a man. But we shall not do it. For control systems it is quite significant to change the elements in formal models.

An alternate name for sign systems and models is *semiotic models*. We will hold ourselves to this name subsequently also.

Let us see what signify the rules of change in model C in the terminology of a control system. Let's begin with LTR*. Model M is determined and invariant. The rules χ_n gives the possibility to make LTR variable, for example, adaptive. Examples of such LTR may be LTR with a probabilistic returning of the type

$$S_i \; ; \; Q_j \underset{U_k}{\Rightarrow} (Q_{\ell 1} \, , \, q_1 \; ; \; Q_{\ell 2} \, , \, q_2 \; ; \; \ldots . \; ; \; Q_{\ell_r} \, , \, q_r).$$

In such an LTR the current situation Q_j, with a complete situation S_i, is transformed not into a single fixed current situation Q_ℓ, as earlier, but into 2 different current situations. The selection of one or another transformation is carried out by a probabilistic mechanism with accounting for the fact that q_i ($i = 1, 2, \ldots . r$) are probabilistics if realization of the corresponding transformation ($\sum_{i=1}^{r} q_i = 1$). Depending on success or non-success of utilization of such an LTR, there occurs a danger in the distribution. If for example, in the complete situation S_i there was made the transformation $Q_j \Rightarrow Q_{\ell i}$ and it turned out to be successful (for example, from the point of view of the DMP), then q_ℓ can increase by a certain amount under the condition of retention of normalization, i.e., under the condition of corresponding decrease of the other q_j. This method of changing LTR allows to tune semantic rules to that object for which the control system was designed. Instead of recalculation of the probabilistic scheme it is possible to change LTR by other methods too, and this will be addressed in the proper part of the book.

In each current moment if certain statements about the object, its state, and also the state of the control system are true (they) reflect the

---

* Logico-Transformational Rules

state of affairs. This knowledge can be expressed in the form of a selection of axioms. It can be assumed that everything, which comprises the description of the complete situation S_i at the moment of making decisions, is a selection of such axioms. They serve as a source of deduction of that decision U_k, which should be used in the given case. But, in another moment of time this knowledge will be different already. And the assertion that machine tool No. 1245 is free, which could pose as an axion in the preceeding cycle of control, could be false at the given cycle of control. The system of axioms continually "breathes" changes. And these changes can be explicitly shown, for example in the right parts of LTR, by writing them in the form

$$S_i \; ; \; Q_j \underset{U_k}{\Rightarrow} Q_\ell \; ; \; I_i$$

Here I_i are those changes which must be inserted into the description of the complete situation S_i after Q_j has become Q_ℓ. Of course, other methods of describing of changes of axioms defined by rules χ_A are also possible.

Finally, changes in syntactic rules may witness that the language of description of situations on the object of control, the complete knowledge about it and the control procedures turned out to be too poor. Into it must be introduced new methods of forming functions, of relations of type communication. Functions of such a change are carried out by rules χ_p. Rules χ_T add, to the list of initial concepts, relations or functions, new elements or delete from it those that turned out to be unneeded for purposes of control.

A Semiotic model can be represented in the form of a network, shown in Figure 1.9. Each vertix of the net represents a certain formal system, and the connections between the vertices define transition from one formal system to another under the effect of change χ^i. These changes may coincide with χ_T, χ_P, χ_A, or χ_Π or may be some combination of them. During one cycle of operation of a semiotic model, depending on the contained I_i, the model will either remain in the same state (in the framework of the same formal system) as before, or will transfer into a new state.

1.6 Data and Knowledges[*]

The concept of a semiotic model introduced by us and its interpretation shown in Figure 1.9 allows us to consider the important problem of artificial intelligence - the interaction of "data" and "knowledge" and a logical contradiction of knowledge[*]. This question is important for structional control also. But before we go directly to its discussion, let us note that until recent time, in the theory of programming of problems for systems of artificial intelligence, a lively discussion went on over two possible paths for constructing similar systems: the "procedural" and the "declarative" paths. In the procedural approach, in the center of attraction of the developer, was the procedure, the program, while the information for it played a secondary role. The program worked with the data, called them up as they were needed, assimilating them in the memory. The data remained somewhere on the periphery. It was considered that the basic work of the system lay in the area of creating the procedures.

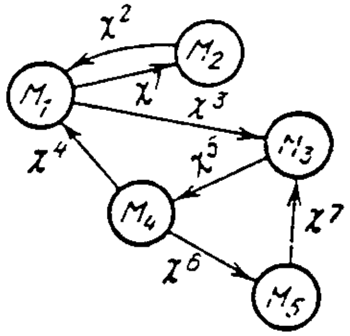


Figure 1.9

This situation began to change when the data became more complex. There appeared structurized data - lists, documents, semantic networks, frames. On some of them the discussion is still to come. They will play an important role in the pages of our book. Then came special means for storing the data - information banks and bases, to which began to be added the term "intellectual". This meant, that together with storing of data in bases and banks there occurs an elementary processing, carried out by special auxiliary programs. They carry out the search of data, their recording into a designated place and a series of other operations. The form of data recording into a storage place may not coincide with the form used by the program that solves the main problem, for these programs are many, while the storage place is only one. For this reason, service programs took upon themselves the

---

[*] The author actually uses the plural term "knowledges", which is okay, in Russian, but not in American.

function of translation of the query and of the output reply from one form of representation into another. Information complexified its form of representation, it became structurized, some portions of information began to be connected to others by some relations. Procedures related to its processing began to be more complex, become self-contained.

Many transformations, which at one time were performed by programs solving the main problem (e.g., search by pattern), would be performed now outside the main program. In many cases, the usual programs for processing of information by its direct conversion were replaced by special procedures for processing of information. Thus emerged a declarative approach, in which the operation with data (we are beginning to call them knowledges) became first and foremost and the procedures were driven into the background.

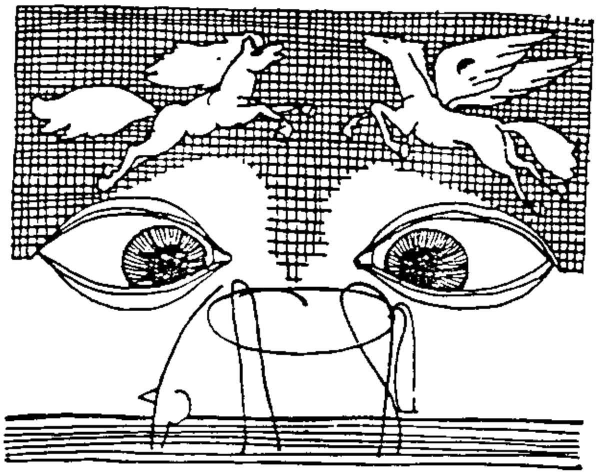


Figure 1.10

What are the main differences between data and knowledges? Traditional programming offers no answer to this question. However, we can find such an answer within the boundaries of the theory of semiotic models. Let us examine Figure 1.10. We can see a stylized outline of a horse grazing on a meadow, which image evokes in the mind of the looking individual a certain generalized picture of a "horse grazing on a meadow". This picture fits into a certain system of data on such a situation. The individual may, for instance, visualize that the motions made by the horse lead to its movement along the meadow, warding off annoying flies and swallowing grass fodder. The individual may determine that it is a horse of a certain breed (if, of course, he knows what it means), that it is a work horse, and, consequently, that the horse belongs to somebody. By means of certain associations, the individual can imagine a stable (not a concrete stable where the horse is kept, but a stable in general), a man who takes care of the horse, etc. The important point is, that the entire informational complex formed by the individual looking at the grazing horse seems to be "logically non-contradictory", because he can visualize a concrete situation (corresponding to the current situation in situational control) within the frame of which all the data can be included. And if we ask the individual "Is the horse flying?", his answer will be negative. Of course, such horses do not fly. However, the individual can easily visualize a flying horse. For instance, the Pegasus, if he ever heard anything about this horse. In the world of Greek mythology, where Pegasus exists, a horse can fly. Furthermore, if in the world of mythology the question is posed "By what means if Pegasus flying?", the answer will be "By means of wings", because Pegasus is a winged horse. But in the realm of Russian fairly tales, the horse does not even have wings. "Konek-Gorbunok" can fly without wings, transporting the hero of the tale to wherever needed.

If, however, knowledges about the horse are placed within the frame of a concrete situation, a logically contradictory system is created. In this system, the horse flies and does not fly, it flies only if it has wings and it can fly also without wings. Absurdity emerges. However, similar absurdity arises regularly in the path of man and he does not suffer at all. He functions calmly with such logically-contradictory images as, for instance, "the third half of the game started", "fulfillment of all wishes which have no boundaries", "a river is flowing and it is not flowing, it is made of moonlight silver", etc. Moreover, such absurd statements amuse the human mind, they are pleasant to man. This happens because all such information is not stored in the brain within the framework of one system. If that was the case, he would be deprived of the possibility to perform logical operations with this information, and consequently he would not be able to use the stored information for his pragmatic goals. Thus, the logically-contradictory information is organized in his brain in such a way as to avoid possible logical contradictions.

One of the possibilities of such an organization of information is shown in Figure 1.9. Each vertex of the network stores logically non-contradictory information which can fit within the frame of one situation, and this information can be reflected in a certain formal system. As far as this system is concerned one can make logically-valid conclusions which can be used in the presence of a similar situation. However, other situations are also possible where their own axioms and semantic rules, their own interpretations exist. In these situations (possible worlds), one can arrive at conclusions which would be meaningless in another situation.

Let us mention another important aspect permitting us to clarify the difference between data and knowledges. Any concept used by man appears to have two sides - an "extensional" and an "intensional". In science, a controversy as to their precise interpretation continues to this date. Therefore, our further discussions should not be regarded as a method for solving these arguments. Concepts of the "extensional" and the "intensional" will be considered pragmatic as accepted in the theory of representation of knowledges. Let us assume that the extensional of a certain concept consists of certain facts corresponding to this concept. If, for instance, we examine the concept of a "worker" for a concrete laboratory at a certain scientific institution, a list of all employees of the laboratory can be its extensional. The extensional of a microcalculator repair establishment in the city of Minsk can be a list of all such establishments in the given city at a given time. The extensional of a concept can be finite as it was the case in the above given examples, or it can also be infinite. In the latter case, it cannot be defined by a simple list. Here, we have to proceed in a different manner. For instance, we can define a certain characteristic rule which must be satisfied to determine the affiliation of the extensional to a certain concept. For example, when defining the extensional of the concept "even number", we can't simply enumerate all even numbers, but we can define the concept by means of the following characteristic rule: when a certain number is divided by 2 and the remainder is zero, the number belongs to the extensional of the "even number" concept. Such a rule can be used (and generally is used) for defining the "even number" concept. This definition will be the intensional of the given concept.

An obvious relationship exists between the declarative and the procedural representations on one hand and the extensional and the intensional on the other hand. The extensional is a set of concrete data defined in a declarative form. The intensional, as a rule, defines a certain procedure which permits the determination of the belonging of one or another concrete fact to a certain concept. The intensional selects knowledges and separates them from data, which are always defined extensionally.

Now, we can describe the difference between data and knowledges. Data are a complex of information which fit within the boundaries of a certain formal system if we take into account all possible interpretations of this system. In situational control, the multitude of data is the multitude of current situations, which can form the object of control (assumed to be constant) and information on the current state of the object and its control system. Knowledges are the information which is stored in all possible worlds together with conditions for transition from one world into another. In other words, knowledges are not only the multitudes of all current situation in objects of a given type, but also methods of transitions from one description of the object to another, methods for changing components of a formal system and, consequently, knowledges are all the factors which form the whole of a certain semiotic model.

1.7 A Plan for Further Presentation

Before we turn to the presentation of basic problems connected with the creation of situational control systems, let us outline our plan of presentation.

The subsequent chapters will deal first of all with the description of the situational control language which will be used for the description of situations. This language will permit the coordination with the language used for description of knowledges about the object of control and its control as well as the language used for description of the target of control.

Subsequently, we will discuss classification procedures which are very important in situational control. We will describe methods to supplement the descriptions of situations by means of knowledges about the object of control and its control methods, which are stored in the memory of the system. In relation to these aspects, we will discuss special pseudophysical logics used for supplementation of description and procedures of inductive conclusion, which will permit us to gain new knowledge during control.

Finally, we will discuss various approaches to the acceptance of solutions for control, which are traditional for situational control. We will describe not only traditional deductive systems for drawing conclusions but also nontraditional systems which are supported by the specific properties of semiotic models.

Chapter 2.

LANGUAGES FOR DESCRIBING THE OBJECT AND THE SITUATION

2.1 Introductory Remarks

In traditional control methods, the object of control is replaced, as a rule, by a certain *syntactical formal model*. Let us clarify what enters into this expression. Let us take a certain differential equation of the type

$$\ddot{x} + 2\dot{x} - 3e^{x} = t^{2}$$

Let us ask ourselves, what object is described by this equation? Unfortunately, it is impossible to find an answer if we don't know the object for the description of which the equation was formulated. The transition from a meaningful (indicative?) control problem to its mathematical analog excludes semantics of the initial problem, and removes the concrete nature of the problem.

When such a transition is undertaken, the mathematician begins to work with the problem. He is not concerned with the object of control. The mathematical reality concealed in the syntactical model is quite sufficient for him. He can investigate the finest nuances of behavior of integral curves for this equation, solve the problem of stability of its solutions, develop methods for the solution of similar types of equations. This is his kingdom where he can reign without the necessity of seeing and knowing the object of control. An analogous picture forms in the transition from the object of control to any syntactical (which is formal in its nature) mathematical model. Here lies the strength of mathematics, a broad applicability of its proposed methods. If the mathematician was able to find and investigate solutions of the equation describing diaphragm vibrations in the telephone receiver, and if the vibrations of aircraft wings, which finally result in the destruction (flutter), could be described by the same type of equations, then the solutions obtained for the vibrating diaphragms could be transferred in their entirety to the flutter phenomenon. Or, when multitudes of sets of various problems can be reduced to linear integral programming, the mathematician will give us a method to find the optimal solution to all of these problems regardless of semantics of the problem to be solved. However, as was discussed in Chapter 1, new objects do not give us such a possibility. The standardization of methods for solutions of problems arising here is practically impossible. We all remember the voices ringing decades ago among specialists about the notorious automatic control systems, how they were conceived by their first followers, a call for standardization of modules of these automatic control systems. If such units could be built the problem of design of automatic control systems, one of the most urgent problems of the past two decades, would be considerably simplified, and we could hope to formulate the theory of automatic control systems. Experience has proven that such hopes were elusive. Generally speaking, with the exception of the standard wages-calculating module and certain other units (personnel system, warehouse stock system, etc), nobody succeeded in designing standard modules. The reason for this, as mentioned earlier, is the fact that each object of automatic control systems is unique, and the development of its control system must be geared toward this object. Apparently, the only thing that can be standardized is the method for the design of the same-type ACS.

The uniqueness of the object of control (only such objects will be discussed in this book) demands special *semantic* and *pragmatic formal* models for description of its structure, functioning and characteristic features of its control.

Special language resources are needed for the description of such models. The language resources must be able to reflect semantics and pragmatics of the described phenomenon. Unfortunately, classical mathematics will not permit us to do that. Thus, its resources must be expanded. The easiest way to accomplish this lies in the field of mathematical logic, which runs close to the method used for the description of phenomena and drawing of conclusions from these descriptions by man. On the other hand, technologists who somehow managed to perform their task, can supply us with a lot of information on their work with automation. However, the submitted information will always be given in a natural language, which is far removed from the precise language of mathematics.

What happens when a student approaches a complex object the control of which has not been automated? What will an experienced control technologist do in this case? He can't give the student the precise control algorithm because he does not know it. However, he knows something at a certain intuitive level. He begins to explain the principles of control of an object. He instructs the student to follow the path taken by the technologist for realization of the control process. After a certain time, a miracle occurs. A common place miracle which has been happening millions of times in the history of mankind. The student has formulated in his brain the model of the object and the model of its control, and he himself becomes a control technologist. The explanations presented by the teacher in an ordinary natural language, and the observation of teacher's activity were sufficient for the student. Now, we are able to suggest the following natural-science hypothesis.

Hypothesis 2.1. The entire information on the control object and its control methods can be expressed by means of an ordinary natural language.

This hypothesis can't be strictly substantiated. It is possible that the hypothesis is incorrect for certain objects. Let us go back to Chapter 1 containing the case of diagnosis of psychological ailments. It is doubtful that the instantaneous "sensing" of the ailment by experienced physicians who practically unerringly arrive at the correct diagnosis, can be written down as text in the Russian language. Then, the problem of diagnosis would not exist. Nevertheless, such a hypothesis can be accepted for many objects. Consequently, there exists an interest in the natural language (and its representational means) which is characteristic of the situational control method.

However, for modern computers, which form the basis of the most flexible control system, a natural language in its entire volume is inaccessible. Therefore, the natural-language texts, which describe the object of control and the experience with its control, must be reduced to a level at which the obtained descriptions can be "loaded" into a semiotic system. This brings us to a certain statement which takes the form of a hypothesis.

Hypothesis 2.2. Any natural-language text dealing with matters stated in hypothesis 2.1 can be converted into the formal language of a semiotic model.

Apparently, as it was the case with hypothesis 2.1, hypothesis 2.2 is not always correct. However, the practical experience of the author and that of all developers of situational control systems confirms its correctness.

Thus, we arrive at the following statement of the basic problem connected with the description of a situation as well as the structure and functioning of the object of control: to find a method for natural-language description of necessary information which would enable us to realize basic procedures of situational control; classification, correlating the description of situations with single-step solutions with help of the DMP), and extrapolation. This is the problem to be discussed in the present chapter.

2.2 Roles and Relationships

Man must know how to model and describe the world around him in order to comprehend the world and accomplish his activities in the world. However, we know very little about the process by means of which man accomplishes that. Many details of this process remain unknown to us. Nevertheless, in the last decades, psychology, psycholinguistics and the theory of knowledge accumulated numerous facts which can form the basis for the formulation of the following hypothesis.

Hypothesis 2.3. When attempting to represent the external world and describe it, man distinguishes a finite set of *relationships*. These relationships connect the individual elements of a model of the world. The elements themselves, connected by relationships, appear in them as definite roles, the semantics of which are defined in the form of a relationship.

The hypothesis requires certain explanations. Apparently, there is no doubt that man (possibly all animals at a sufficiently high level of development) perceives his surrounding in a structured form. As proof we have the human language specifically intended for the description of the world in which man lives. Man distinguishes in the world individual concrete facts, objects, phenomena, processes, events, and relationships among them. The language preserves this structure. Facts, objects, phenomena, processes and events acquire their own distinctive names in the language, and the relationships are expressed either by means of special names of by purely linguistic syntactical means.

Semantics of relationships are inseparably linked with the roles played by the elements connected by a relationship. For instance, the relationship *superior -- subordinate* connects two persons (elements of a model) with their own definite roles; one role is called the *superior*, the other - the *subordinate*. In the sentence "Ivanov is the superior of Petrov", Ivanov and Petrov have such roles. Through comprehension of roles a person recognizes a situation of reality and can organize his goal-directed activities within it. To solve the latter problem, man utilizes roles of the type *means -- result*, *cause -- effect* and numerous other roles.

In accordance with hypothesis 2.3, there is a finite number of such roles, consequently there is a finite number of relationships distinguishable by man in the real world, because each such relationship represents a pair of interrelated roles. It is doubtful that the validity of such a statement could be proven. However, observations of language texts and methods of design of special languages for description of reality, which will be introduced in the coming paragraphs, seem to confirm this hypothesis. In conclusion of this paragraph, let us present the following list of roles (obviously incomplete): *an integer, a sign, a sign of a sign* (and recursively any number of times), *a superior, a number, a set, order, a comparison, a higher concept, a measure, an attributed sign* (e.g., a person among the people I know), *action, a sign of action, an accomplice, means, an object, an object of information* (e.g., "Ivan received information about his affairs"), *a source, a receiver, location, time, cause, circumstance, goal, result, a substitute* ("He drank the tea in my place").

As we shall see later, all the roles can be reflected in the languages proposed for description of situations.

2.3 The Natural Language And The Theory Of Control

Representatives of different sciences perceive the natural language in different ways. For the linguist, the language in itself is an object of study. For the psychologist, it is an instrument of thought and communication among people; for the philosopher, it is a model by means of which cognitive processes can be organized. The specialist in the theory of control also has his own concept of the language. For him the language presents above all the means to describe the objects to be controlled, the situations within which the objects must be controlled, and the procedures of control. The earlier formulated hypotheses on the use of a natural language in situational control initiate the search for the mechanisms by means of which a natural language can express concepts, relationships and roles that are important to control. The first step on the path to discovery of these mechanisms is to isolate in the vocabulary of a language the groups which carry a definite functional load in the description of objects, situations and procedures of control.

1. Concepts. Here we find a large group of lexical units of the language which is used for designation of uniform groups of facts, events, phenomena and other elements in the real world. The following words can serve as examples of concepts: a table, a crane, the end of a shift, an earthquake, a prize, etc. The important matter here is the fact that the word table does not denote a certain concrete object of the real world, but a table in general, any object which can be called a table. Similarly, the end of a shift is not a concrete definition, but an event described by a given combination of words that can take place anywhere at any time.

This means that concepts have a certain structure of their own, a certain set of necessary characteristics by means of which certain concepts can be distinguished from others, and the concrete elements of reality are designated by these words and word combinations. Indeed, this is true. The next chapter of this book will dwell on this question. For the time being, we will appeal to our intuition which will hint at what can be applied to the group of concepts.

In the future, concepts will be designated by Roman letters with or without an index. We will exclude only those letters which will be introduced for designation of other functional groups in the vocabulary of a natural language. Let us also have an understanding about the following designation. When a concept is used in relation to a certain concrete element of reality, small letters will be used, and when a concept refers to a class of uniform elements, capital letters will be used. For instance, in the sentence "A table can be of wood, iron or stone" the concept table will be represented by any capital Roman letter; but in the sentence "I like this table", the concept table will be designated by a small letter. Thus, the size of the letter will indicate whether we are dealing with a concrete concept or a class concept.

2. **Names.** Practically all words or word combinations can assume the role of names. Names serve as concrete definitions of any type of element of reality which enters into a class concept. Let us give you the following examples of names: Peter Petrovich Petrov, train number 33, carriage number 7, the number of a coat hanger at the theatre on Bronnaya, etc. Numbers will not always isolate a unique element from a class. There are hundreds if not thousands of Petrovs with the patronymic Peter Petrovich; there are tens of trains numbered 33 and carriages numbered 7, hundreds of numbers in the coat room of the Theatre on Bronnaya. Therefore, as it was the case with concepts, let us introduce two designations for names. Small i(s) with various indices will designate names of concrete elements, capital I(s) with various indices will designate names of subclasses of the class concept all elements of which assume this name. Let us analyze the phrase "My uncle Peter Petrovich Petrov". This phrase uses the class concept "uncle", which will be designated by any capital Roman letter. "My uncle" isolates from this class a certain subclass of persons who call me their nephew. This seems to be a name which should be designated by I_1. Finally, Peter Petrovich Petrov of this subclass stands for a concrete person and, according to our agreement, his name should be designated by a small indexed latter, for instance i_1. Sometimes, it is difficult to determine whether the name refers to a concrete element or whether it refers to a certain subclass of elements. Let us hope that this difficulty will vanish in the description of narrow problematic fields connected with the control of a concrete object.

3. **Relationships.** Relationships establish connections between concepts and names as well as other functional language groups which will be discussed now. In contrast to other functional groups, relationships can be given not only by means of the vocabulary of a language but also by means of grammatical connections expressed by means of the language. Due to the importance of relationships in the design of a situational control language and other relation-type[1] languages, let us describe in greater detail the groups of relationships of a language.

---

[1] From the English word *relation*. These languages emphasize that relationships play a major role.

3.1 Classification Relationships. These relationships serve as means to classify elements of the real world, to form classes of elements, to establish relations between classes as well as relations between classes and individual elements. The most important among these relationships is *to have a name*, which in the future will be designated by the letter ρ. This relationship enables us to isolate from a class concept the various subclasses of different strength down to the selection of a concrete element from a class of similar elements. Its effect had been demonstrated in the previous section dealing with names. This is the group of relationships that contains also the relationships *class -- subclass* and *element -- class*. In essence, these two relationships play an analogous role as the relationship *to have a name*. In a language, however, they can be extablished even without the use of any words and word combinations referring to names. For example, "Certain members of crew Nr 1", Some of the workers of an assembly shop", "One of the cars of the taxi-fleet column". Similar in spirit to these relationships are also relationships of this type: *family --species, subordinate concept -- higher concept*. As examples of their use we can name "Both the stationary and the traveling cranes belong to hoisting cranes", "A hammer is an example of a tool". The same group of relationships contains also the relationship *part -- whole* the essence of which becomes clear in the following examples: "These are the windows of my house", "The lathe has a stand", "The carburetor of a car engine constitutes one of the major parts of the system". In the future, all these relationships (as well as other types of relationships) will be indicated by small letters r with various indices. At the end of this section, we will present a list of certain relationships with such indices as will be used within the framework of this book.

3.2 Indicative Relationships. These relationships assign certain qualitative properties to concepts. All of them can be presented in the form of a successive composition of two relationships: *to show signs of or to have the value of a sign*. As examples we can take "The lock-through line is long" ("The lock-through line has length. This length is extensive"), "Gantry crane", "Weaving stand". These relationships describe a complex of attributes inherent to the concepts, they represent their characteristics. Some of them provide the concepts with a definite feature (e.g. to have length d for the concept of line, but for others this statement is not true, because they characterize only a given concrete element in a given location at a given time (e.g. "This house is white"; certainly, at another time, the house might have been repainted).

Indicative relationships are closely connected with concepts. The concepts can be regarded as a certain collection of definite attributive indicative relationships. A concept would be the sum of its defining indicators. The concept concentrates all common things that unite individual, concrete elements into a class. This communality seems to fluctuate within certain bounds of changing values of attributive concepts.

The indicative relationships encompass also various relationships of metaphoric meaning (e.g. "The plan is heavy").

3.3 Quantitative Relationships. They express quantitative characteristics of concepts. Quantitative relationships can be reduced to the composition of two relationships, *to have measure and to have the value of measure*. As examples of phrases where such relationships are realized we can name: "The dispatcher has two free buses at his disposal", "The distance between points A and B is ten kilometers". A special role is played by quantitative relationships where in place of a concrete value of measure a certain qualitative value is indicated: "There is plenty of time to carry out this work", "Machine tool Nr 7 is often at a standstill". This case will be singled-out and discussed in our analysis of the language after the analysis of relationships.

3.4 Comparative Relationships. These relationships compare two characteristics of a certain concept or a group of concepts by means of an indicative or quantitative relation. As examples we present: "This decision is more effective than the decision made by dispatcher Petrov", "The tanker is longer than the moorage", "Worker Ivanov works twice as fast as Smirnov".

3.5 Affiliation Relationships. These relationships are similar to classification relationships; the only difference is the fact that they connect two elements of the external world with only a situational relation. As examples we present: "Ivanov is a colleague of Petrov, they share the same profession", "Operations X and Y are now within the view of the dispatcher", "Sidorov is the brother of Kuptsov".

3.6 Time Relationships. They encompass the following types of relationships: *to be at the same time, to precede in time, to coincide at the start, to cross in time*, and numerous others. Included are also relationships ascribing to the elements a certain time of existence or a date: "The plan will be fulfilled this year", "Assembly of the workpiece ended at 8 A.M. on November 30, 1983".

3.7 Space Relationships. These relationships either establish the place of presence of a certain element of the real world or the relationship of elements among themselves in a certain space. As examples we present such relations as *to be in, to be between, to be to the left, to touch*, etc. The following phrases contain spatial relationships: "A wall divides the two workshops", "The vessel is to the right of moorings", "The workpiece is on the machine tool".

3.8 Causal Relationships. These relationships are used to reflect cause -- effect relations as well as relations which reflect the goal, motivation or preferences in the making acceptance of decisions and actions. As examples of phrases where such relations are revealed we offer: "Moving the vessels away from moorage enabled us to unload the tanker Baku", "Line Nr 6 was disconnected by the dispatcher to eliminate accidents", "When pressure rises above the norm, the temperature of the alloy abruptly increases".

3.9 Instrumental Relationships. These relationships reflect the pragmatic aspect of activity. The most important of them are -- *to serve as* or *to be the means for*. Here we include also relationships of the following type: *to be the instrument of, to be auxiliary means of, to contribute*, etc. Such types of relationships are present, for instance, in the following phrases: "The foreman felled the pine by means of an ax", "Mooring berth Nr 8 is used for unloading of dry cargo", "This process can be realized only in vacuum".

3.10 Informational Relationships. This group of relationships describes various aspects of sending and receiving of information, instructions, inquiries etc. As examples we present the following phrases: "Ivanov gave Sidorov the order to start well-drilling", "The Ministry did not receive in time the report on activities of Enterprise A".

3.11 Order Relationships. They describe the relative order of elements of the real world. They are expressed by words such as -- *the following, the nearest, the next in turn,* etc.

There are other groups of relationships (for instance, substitution relationships: "Ivanov's crew performed the unloading in place of Petrov's crew), but the main groups of relationships that are important to problems of control have been enumerated by us. Table 2.1 contains basic relationships. (In the future, the ordinal number of a relationship in the table will coincide with the index of the corresponding r_i).

Table 2.1

Nr	Name of Relationship	Nr	Name of Relationship
1	to have	32	to be to the right
2	to be an element of a class	33	to be in front
3	to be a subclass of a class	34	to be behind
4	family -- species	35	opposite
5	minor -- major (in hierarchy)	36	crossing in space
6	part -- whole	37	abutting
7	to have a sign	38	to be inside (in space)
8	sign -- value	39	to be upon
9	sign -- measure	40	to be above
10	measure -- value	41	to be below
11	equal to	42	to be between
12	in comparison with	43	to be in
13	more than	44	belong
14	more than or equal to	45	serve as
15	less than	46	to be means for
16	less than or equal to	47	to be an instrument of
17	not comparable	48	contributing to
18	comparability -- measure	49	cause -- effect
19	correlate	50	to be the goal
20	correlation -- value	51	representing the cause of
21	at the same time	52	to be the sender
22	to be earlier	53	to be the receiver
23	to be later	54	to be the source of information
24	to begin at the same time	55	to be next in turn
25	to end at the same time	56	sequence
26	adjoining in time from the left	57	to be closest to
27	crossing in time	58	to be a substitute

28	coinciding in time	59	to be in the state of
29	to be within (in time)	60	to have a name (designated by ρ)
30	coinciding in space	61	to have valuation (designated by ϵ)
31	to be to the left		

The list is not a complete enumeration of all static relationships of the Russian language (only static, because actions have not been discussed by us). Furthermore, it is easy to notice that not all the relationships in the table are interchangeable. This aspect will be discussed a little later. The meaning of the list in Table 2.1 can be called illustrative. We will use the relationships included in the table in examples to be given in this and the following chapters.

4. Actions. All of the dynamics of the external world are expressed by the language which words and word combinations which describe actions. There are various types of actions, and it is important to us to isolate the basic types. Letter d with various indices will be used to designate actions.

4.1 Imperatives. The actions included here are various orders and commands for the execution of certain actions: *switch on, fulfill a plan, lower the pressure,* etc. As a rule, these imperatives give the decisions that are issued by the control system to the controlled object.

4.2 Processes. This group contain lexical units which correspond to names of technological processes that can take place in the controlled object. Often, they will be called *process-actions*. As examples we can name: *handling cargo, regulating a valve, painting,* etc.

4.3 Conditions. This group contains words and word combinations which establish definite starter of the controlled object and the control system. As examples we can mention: *the malfunctioning automobile, the empty mooring berth,* etc. Conditions will often be called *condition-actions.*

The reader ought not be confused by our broad interpretation of the term *actions.* Such terminology is quite applicable to processes and states when texts contain such expressions as "Loading of the truck is taking place" or " The mooring berth is becoming free". Unfortunately, we don't have a complete and convenient classification of actions for the solution of applied problems. Numerous attempts at its creation had been made in this direction.

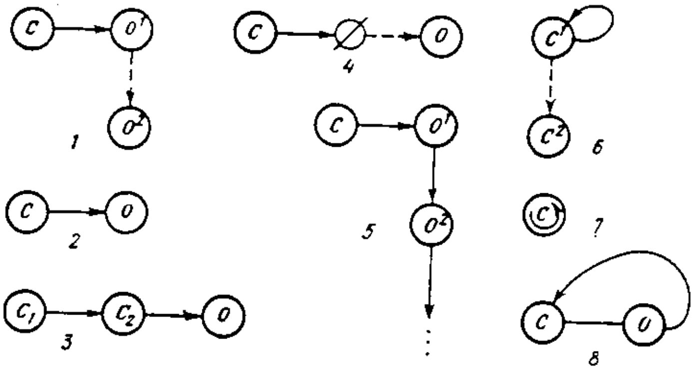


Figure 2.1

Figure 2.1 contains one of such classifications, Here, "C" denotes the subject performing the action and "O" denotes the object at which this action is directed. Superscripts correspond to states of the subject and the object, subscripts correspond to different subjects. Solid-line arrows symbolize active forces, broken-line arrows represent passive transitions. Let us cite some examples for various classes of actions shown in this figure. *1 - loading a vessel, 2 - observing altimeter reading, 3 - sending a signal to the regulator to change pressure, 4 - building a house* (as always, phi denotes vacuum), *5 - actuating a set of programs that controls a washing machine, 6 -*

changing observation conditions, 7 - checking the proper functioning of one's circuits, 8 - sending malfunction signals.

Just as in the case of static relationships, we can construct a table of dynamic relationships. In addition to the relationships defined by classes of actions depicted in Figure 2.1, this table will also contain relationships connected with the shifting of objects in time and space. Table 2.2 contains examples of dynamic relationships (numbering of these relationships consecutively follows the numbering in Table 2.1).

Table 2.2

Nr	Name of Relationship	Nr	Name of Relationship
62	moving towards	70	to be free
63	to decend	71	having as an object of action
64	to approach	72	bringing into a state
65	to age	73	moving from behind a certain object
66	to load	74	action -- location
67	to switch on	75	action -- time
68	hammering in	76	taking action
69	to be in good working order		

5. **Quantifiers.** A special group of words and word combinations is formed by the so-called *quantifiers.* Two of them have a very distinct meaning. One of them, which is usually called the *commonality quantizer* ("kvantor obshchnosti"), is designated by the inverted letter A and produces such statements as "All workers received their pay", "Every car in the taxi fleet has been checked by GAI". Its meaning denotes that certain information encompasses all the elements (without exception) of a defined class. The other quantifier, which is usually called the *existence quatizer* ("kvantor sushchestvovaniya"), is designated by the mirror image of letter E and produces such statements as "There are Zil 130 among the trucks waiting to be loaded", "This part is stored in at least one of the warehouses". In essence,

the existence quantizer denotes that certain information is true for at least one elements of the class under consideration.

The other quantifiers, which are widely used in natural languages as well as by technologists in description of processes at the object of control and control methods, do not posses such unambiguous semantics. Their examples follow: *often, many, soon, in a large time interval, only, even,* etc. We will devote several paragraphs to these inexact quantifiers in this chapter and in later chapters. At present, let us only demonstrate the complexities encountered here. In the phrase "Even Ivanov was able to do this work", the quantifier *even* seems to project all workers within the scale of expertise required for this particular work, and Ivanov stands on the lowest rung. However, in the phrase "Not even Petrov was able to do this work", the same quantifier concentrates not on the lowest but on the highest rung of the scale. In addition to the ability of this quantifier to isolate the lower and upper boundary values of the scale, it can also project completely different shades of meaning (for example, "He was not able even to get up"). The same difficulties are encountered in identification of all shades of meaning imparted by the quantifier *only* (just) to the natural-language text. Let us present a few examples: "The crane only (just) began to operate", "Only three students work on the construction job", "The work will be fulfilled only tomorrow, not today" "The work will be fulfilled (done) only tommorow and not the day after tomorrow". The meaning of such quantifiers as *many, often, close,* etc. is simpler and, a little later, we will describe the procedures for interpretation of similar quantifiers which are often encountered in the practice of control. Quantifiers will be designated by $\succ\!\!|_2$, preserving the tradition of inverted letters which originated in logic.

6. Modifiers. This group of words and word combinations practically belongs to the group of relationships which determine the criteria of elements. There are, however, several reasons why it is more convenient for us to analyze them separately. They include such lexical units as: *rapidly, cautiously, sea* (as in sea port), *freight* (as in freight traffic), etc. Letter m_i will be used to denote modifiers.

7. **Modalities.** They contain lexical units of the following type: *it is necessary, it is desirable, it is absolutely impossible, it is obligatory,* etc. These units will be designated by f_i.

8. **Valuations.** The following units belong here: *it is nice, it is expedient, it is harmful, it is attainable,* etc. Symbols v_i will be used for valuations.

The above-discussed eight functional groups for natural languages are not the only possible groups. One could separate the units in another way. Our division, however, allows us to design a language convenient for description of situations at the object of control and in the control system, as well as to perform the necessary procedures of classification, correlation and extrapolation.

2.4 Situational Control Language (SCL)

Here, we will present fundamentals of the situational control language. The nature of the description will be semiformal in order to avoid multiple-bracket formulas which can be very tiresome to the reader. In paragraph 2.5 we will describe a strict approach to the formulation of the lexicon of a similar-type language on the example of the SCL directed toward a definite problematic area; in paragraph 2.6 we will present a formal determination of basic syntactical structures of the SCL. After an analysis of languages similar in their ideology to the SCL and a comparison of their possibilities, we shall present a more precise definition of the semiotic model introduced in chapter 1, and we will indicate the place of SCL in this model.

A *simple nuclear structure* in the form of a triad (xyz) forms the basic unit of the SCL. The center is occupied by a certain relationship or action. The outer positions are accupied by concepts. If the center contain the relationship *to have a name*, a name will be in the right position; if the center contain the relationship *to have valuation*, a valuation will be in the right position. Finally, if the center contains the relationship *measure--value*, the concept *measure* will be in the left position and a numerical value will be to the right.

Example 2.1. Let us demonstrate how phrases describing a certain situation can be represented by simple nuclear structures of the SCL. Let us begin with the phrase "Part Nr 1244 is being machined on a lathe". This phrase would seem to contain two phrases in SCL: "The part has the name Nr 1244" and "The part is being machined on a lathe". To indicate the simultaneous existence of the two facts we will use the conjunction sign & which will always be omitted if there is no special necessity to use it. Then, if we introduce the designators: a - the part, b - the lathe, d - is being machined (here and now), i - Nr 1244, the statement expressed in the SCL will take the form $(a\rho i)(adb)$. Or, in another equivalent form, which does not require the repetition of certain symbols: $((a\rho i)(db)$. Another description: "100 VAZ 2103 cars have been produced". If q - car, i_1 - VAZ 2103, r_{59} - *to be in the condition*, d_1 - *to be finished*, u - *measure*, r_{18} - *measure--value*, r_1 - *to have*, v - a piece, then this description will take the form $(((q\rho i_1)r_1((u\rho v)r_{18}100))r_{59}d_1)$.

Here, the simple extreme-left nuclear structure says that the description deals with cars VAZ 2103, the structure $(u\rho v)$ means that the measure has the name *a piece*, and the abutting $r_{18}100$ means 100 pieces. The content in the large left bracket refers to the fact that we are dealing with 100 pieces of the VAZ 2103 car. Finally, the relationship is in a *condition*, and the condition *have been produced* enables us to obtain a representation of the information which was in the initial description.

The above-given example shows that, in the SCL of major significance is the operation of substitution of the left and the right terms of a simple nuclear structure of the whole simple nuclear structure. This process can be repeated indifinitely. In particular, in the last SCL description, such a substitution was done once to the left of r_1 and twice to the right of r_1.

A triad of the type $\succ\!\!\mid((\alpha_1,x)(\alpha_2 z)(\alpha_3 y)$ is called a *complex nuclear structure*. Certain quantifiers or modifiers can stand in place of α_1; in addition to them, modal operators can stand in place of α_2; only quantifiers can stand in place of $\succ\!\!\mid$. Several modifiers can stand in place of each α_i.

Example 2.2. Let us analyze the following description: "Often, when the pressure is above the norm, the temperature is also above the norm. Let $\succ\!\!\mid_1$ be the quantifier *often*, h - pressure, t - temperature, m - the modifier *above the norm*. Then, the following SCL notation will correspond to the above description: $\succ\!|_1$ $((mh)r_{49}(mt))$. Relationship r_{49} from Table 2.1 is the relationship *cause - effect. Relationship r_{48} - contributing to* - could be used in place of r_{49}, because a strict causal relationship between rise of pressure and rise of temperature has not been extablished in the initial description. To affirm such a strict relationship the description would have to read "When pressure rises above the norm the temperature always rises above the norm".

Another description: "The truck rapidly moved toward warehouse Nr 2". If we use a for the truck, c - for the warehouse, r_{62} - for the dynamic relationship *to move* k, m - for the modifier *rapidly*, the given description can be presented as $(a(mr_{62})(c\rho\text{Nr } 2))$. Let us note, that in this example, in place of the symbolic name of the warehouse, the direct meaning of the name was used in the notation, as it was done in the case with 100 pieces of VAZ 2103.

The given examples enable us to make descriptions in SCL of rather arbitrary descriptions of situations. And phrases of the description are translated into individual "phrases" of the SCL.

Example 2.3. Let us analyze the following description of a situation at an object. "Simultaneously with the loading of the Moscow--Vladivostok rolling stock (train) it is necessary to shunt train Nr 126 to the sixth track. Summon porter crew Nr 1 to platform Nr 6. A mail cart must be summoned to the same platform to receive mail from the mail car. Preferably from Ivanov's crew". The complex description of a situation can be presented in the following form, where Arabic numerals, which are not absolutely necessary for the use of SCL in a computer, will be used for a clearer notation. They separate some "clauses" of the SCL from others.

(1) $((s\rho i_1)r_{59}e)((n\rho i_2)(f_1d_1)(j\rho i_2))$,

(2) $((b\rho i_3)d_1(g\rho i_4))$,

(3) $((kd_1(g\rho i_4))r_{53}l)(lr_{38}(mv))(k(f_2r_2)q)$.

The following designators are used here: s - rolling stock (train), i_1 - "Moscow - Vladivostok", e - loading, n - train, i_2 - Nr 126, d_1 - to shunt, f_1 - it is necessary, j - track, b - porter crew, i_3 - Nr 1, i_4 - Nr 6, g - platform, k - mail cart, l - mail (noun), v - car, m - mail (adj. as in mail car), f_2 - preferably, q - Ivanov's crew. The relationships used in the SCL notation were taken from Table 2.1.

Thus, we can see that rather complex descriptions of situations forming at the object of control can be described in the situational control language. If the idea of this language is clear to the reader he can proceed to the description of a situation in the SCL in connection with the structure and functioning of the object of control. This is discussed in paragraph 2.7. Paragraphs 2.5 and 2.6 do not refer to the first level of presentation.

2.5 Development of the SCL Vocabulary

When working with a certain concrete problem area we encounter the first basic problem, namely, the development of the SCL vocabulary. The dictionaries of basic concepts, relationships, actions, valuations, quantifiers, modifiers, names and modal operators must be comprehensive and reflect all aspects of the object of control and its control methods, necessary for creation of a full-value situational control system. First of all this refers to the development of the following three dictionaries: dictionary of concepts, dictionary of relationships and dictionary of actions (operations). They represent the basic dictionaries for all situational control systems, because with the exception of the dictionary of names, the remaining dictionaries are universal. They are identical for all situational control systems and have a universal nature. The dictionary of names is open, and for every system one must simply indicate the principles of name formation in the given system. A similar observation refers also to the dictionary of values of measure or measures which are used in a given problem area.

Dictionaries of concepts and actions are totally determined by semantics of the problem area. The dictionary of relationships, however, displays a broader nature. One could conclude from hypothesis 2.3 that basically this is a finite dictionary. Certain experiments conducted with several European languages (Italian, English, Russian) seem to confirm this hypothesis. The experiments were conducted in the following manner. A rather extensive text from a universal field (as a rule, fiction) was taken. In a successive analysis of phrase after phrase, relationships encountered in these phrases were established. If a newly-established relationship was already included in a previously considered list, this relationship was omitted; if the relationship was new, an attempt was made at first to express it through a combination of previously discovered relationships. For instance, let us consider the following text. "When the combat began, its ordeals began, fears for one's life (emerged). In the midst of fighting, however, these fears disappeared somewhere; they seem to have vanished forever". The relationship of the two events "combat" and "fear" can be expressed by the relationship r_{24} *to begin at the same time*; one can, however, cross-over to events -- "beginning of combat and beginning of fears", "end of combat and end of fears". Then, in place of the relationship r_{24}, one can use the relationship r_{21} - *at the same time* for the first two new events and the relationship r_{22} - *to be earlier* for the two other new events. Using the conjuction of two simple nuclear structures, we receive a notation equivalent to that which was used in relationship r_{24}. This fact would serve as a signal to omit the relationship *to begin at the same time* from the basic list of relationships. If, however, the newly-discovered relationship can not be reduced to a combination of relationships previously-included in the list, the new relationship is entered on the list as a basic relationship. These experiments, involving a sufficiently large volumes of texts, produced an almost identical pattern for different languages. The list became saturated at the point of approximately 175-185 relationships and could not be enlarged further, which fact confirms hypothesis 2.3 referring to the description of the real world. Of course, the inner world of man can be described by quite a number of other relationships not included in the list that was compiled during the above-mentioned experiments.

The finite nature of the list of relationships is interesting only from the conceptual point of view. In practice, there is no particular need to abbreviate this list. The abbreviation results in an unnecessary complication of notations which are not very clear even in their original form. Therefore, in practice one tries to compile a sufficently comprehensive list of relationships which are significant to a given object of control and its control methods.

Now, let us present a classification of relationships which differs somewhat from the classification presented in the previous paragraph. This classification is supported by Figure 2.2. Here, A, B and C represent sets of *elements* inherent in the object of control or elements isolated during its control, (sets) of *properties* of the given elements, and (sets) of *operations* which can be realized in the object of control or in its control system. Designators R_{ij} correspond to classes of relationships which can be established between vertices of the diagram given in the figure. The meaning of these classes is self-evident. For instance, relationships encompassed by class R_{11} are relationships of the type *element - element*, and relationships forming class R_{31} are of the type *operation - element*. A similar interpretation can be applied to the remaining classes of relationships.

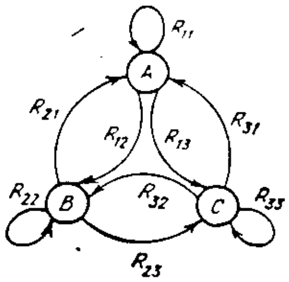


Figure 2.2

Now, if we use X to designate any of the vertices A, B or C, we obtain for X the diagram presented in Figure 2.3. Here, D denotes a set of measures with various names, and $\mathfrak{Z}$ - a set of values. The broken arrow line means that relationship r_8 can be realized only when X and B coincide and qualitative values of the criterion appear in place of values.

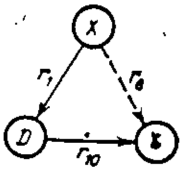


Figure 2.3

In paragraph 2.3 we isolated 11 classes of relationships; the diagram in Figure 2.2 establishes 9 of these types. In Table 2.3 we find connections between these divisions. Crosses were placed in the cells where intersections of divisions are not empty ?.

Here, there are no criterial relationships. They have been absorbed by properties. The property *color* is used in place of the relationship *to have a color*. Missing are also quantitative relationships which are defined by the diagram in Figure 2.3 and not by the diagram in Figure 2.2.

We will present examples of relationships which can characterize the aggregates marked by crosses in Table 2.3.

Table 2.3

Classes	Types								
	R_{11}	R_{12}	R_{13}	R_{21}	R_{22}	R_{23}	R_{31}	R_{32}	R_{33}
1. Classification Relationships	+				+				+
2. Comparison Relationships					+				+
3. Affiliation Relationships	+	+		+			+		+
4. Time Relationships									+
5. Space Relationships	+								
6. Causal Relationships	+			+			+	+	+
7. Instrumental Relationships			+			+			
8. Information Relationships	+			+			+		
9. Order Relationships	+				+				

For type R_{11}: classification relationship - "a gantry crane belongs to loading-unloading mechanisms"; affiliation relationship - "the boom is a part of the gantry crane"; space relationship - "the car is to the right of the bridge ramp"; causal relationships - "overheating causes a fire"; information relationship - "the director informed the dispatcher"; order relationships - "the following part is machined on the lathe". For type R_{12}: affiliation relationship - "the fire truck is a red object". For type R_{13}: causal relationship- "the barrier forces us to go around it", instrumental relationship - "nails are hammered-in by means of a hammer". For type R_{21}: affiliation relationship - "the fire truck is red", information relationship - "red is a danger signal". For type R_{22}: classification relationship - "? blue is a part of the painted ?", comparison relationship - "VAZ 2106 is more powerful than VAZ 2103", order relationship - "orange is next to red in the spectrum". For type R_{23}: instrumental relationship - "looseness of soil requires preliminary strengthening of the track". For type R_{31}: affiliation relationship - "addition operations are included in the command system of the computer", causal relationship - "clearing appeared because of tree cutting", information relationship - "onset of emergency signaling is proof of disturbances in operation of an object". For type R_{32}: causal relationship - "after polishing the surface became smooth". For type R_{33}: classification relationship - "oil change is included in the sequence of car maintenance operations", comparison relationship - "milling operation is more difficult than spot welding operation", affiliation relationships - "fastening of a part in the gripping device of the lathe is a step of the machining operation, time relationship - "the reactor tripped before pumps were engaged", causal relationship - "cable laying forced us to discontinue assembly operations, order relationship - "assembly comes next in the sequence of operations".

Example 2.4. As an example, let us present a dictionary of basic lexical units, which has been used in the development of a situational control system to help the section foreman of an instrument making plant to make decisions on work assignments in the section. The dictionary contains 17 basic relationships: *to be a name, to be a piece, to be a class, to be earlier, to be equal, to be larger, to be subordinate, to provide with, to prepare, subject--object, to move in space to a location...*(for the object of action), *to move in space to a location...*(for the subject of action), *to move in time on..., to be in the vicinity, to be simultaneous, to be the location, to be the time*. In this list, certain relationships are essentially actions, however, in the description of current situations, they seem to represent frozen actions.

Furthermore, the dictionary of basic vocabulary contained 158 basic concepts. Let us list several of them: *stock number, quality index, repairs, time fund, free-of-work period,* piece--operation, plan, period (date), *duration, defect, repair* (defect correction?), *shift, twenty-four hours, Saturday shift, set* (assembly), *production, shaft, price* (rate), *labor-consuming nature, mechanic, milling machine operator, lathe operator, warehouse man, truck, electric power truck, conveyor line, body* (framework), *bushing, key, broaching, die, punch, winding, transformer, etc.*

Finally, the dictionary contained 7 operation-decisions that the control system was able to make: *start production; remove from production; obtain fittings, materials, fixtures; distribute fittings, materials, fixtures; repair equipment; transfer the worker...to piece-operation...; announce a Saturday shift.*

Example 2.4 shows again that the dictionary of basic relationships depends much less on the specific features of the problem than the dictionaries of basic concepts and operation-decisions.

Let us note also that, in transition from descriptions using the natural language to descriptions in SCL, it is not always possible to draw a fine boundary between concepts and properties as well as between properties and operations. The boundary between relationships and operations as well as relationships and properties is also informal. This enables the designer of a system, who takes into consideration one or another pragmatic aspect, to assign one or another element to the necessary type of vocabulary.

2.6 A Formal SCL Syntax Model

SCL syntax can be established by a formal method as it is done when establishing the set of basic elements and syntactic rules in formal systems or formulating the syntax of programming languages for computers.

Definition 2.1. Set $X = \{x_i\}$, where x - any small leter of the Roman alphabet with the exception of d, f, i, m, n, o, p, v with i passing through natural series, is the *set of basic concepts of SCL.*

Definition 2.2. Set $I = \{i_j\}$, where j passes through natural series, is called the *set of names of concepts of SCL.*

Definition 2.3. Set $D = \{d_i\}$, where i passes through natural series, is called the *set of actions (operations) of SCL.*

Definition 2.4. Set $V = \{v_i\}$, where i passes through natural series, is called the *set of properties of concepts (objects) and actions of SCL.*

Definition 2.5. Set $M = \{m_i\}$, where i passes through natural series, is called the *set of modifiers of SCL.*

Definition 2.6. Set $\succ\!\mid \; = \{\succ\!\mid_i\}$, where i passes through natural series, is called the *set of quantifiers of SCI.*

Definition 2.7. Set $O = \{o_i\}$, where i passes through natural series, is called the *set of valuations of SCL.* Specifically, this set contains all real numbers.

Definition 2.8. Set $F = \{f_i\}$, where i passes through natural series, is called the *set of modalities of SCL.*

Definition 2.9. Set $R = \{r_i\} \cup \{\rho, \in\}$, where i passes through the natural series, is called the *set of relationships of SCL.* Relationships ρ and ϵ syntactically differ from r_i.

Definition 2.10. Set $X' = \{m_i x_j\}$, where $m_i \in M$ and $x_j \in X$, is called the *set of modified concepts of SCL.*

Defnition 2.11. Set $X'' = \{\succ\!\!|_i x_j\}$, where $\succ\!\!|_i \in \succ\!\!|$ and $x_j \in X$, is called the *set of quantified concepts of SCL.*

Definition 2.12. Set $D' = \{m_i d_j\}$, where $m_i \in M$ and $d_j \in D$, is called the *set of modified actions of SCL.*

Definition 2.13. Set $D'' = \{\succ\!\!|_i d_j\}$, where $\succ\!\!|_i \in \succ\!\!|$ and $d_j \in D$, is called the *set of quantified actions of SCL.*

Definition 2.14. Set $D''' = \{f_i d_j\}$, where $f_i \in F$ and $d_j \in D$, is called the *set of "modalized" actions of SCL.*

Definition 2.15. Set $V' = \{m_i v_j\}$, where $m_i \in M$ and $v_i \in V$, is called the *set of modified properties of SCL.*

Definition 2.16. Set $V'' = \{\succ\!\!|_i v_j\}$, where $\succ\!\!|_i \in \succ\!\!|$ and $v_j \in V$, is called the *set of quantified properties of SCL.*

These definitions fully formulate the set of basic elements of a formal system corresponding to the syntactical part of SCL. Let us introduce designators $\tilde{X}=X \cup X' \cup X''$, $\tilde{D}=D \cup D' \cup D'' \cup D'''$, $\tilde{V}=V \cup V' \cup V''$. To define the SCL syntax let us introduce also three operations: *concatenation* (assigning symbols to each other) which will not be notated in any special way, *conjuction and negation* with their customary designations. When introducing syntactically correct combinations to SCL we will use round brackets. Elements $\bar{\bar{X}}$, $\bar{\bar{D}}$ and $\bar{\bar{V}}$ will be designated by corresponding small letters with a tilde.

Definition 2.17. The set of simple nuclear structures of SCL consists of triads of this type: $(\bar{\bar{x}}_i r_j \bar{\bar{x}}_k)$, $(\bar{\bar{x}}_i \rho i_k)$, $(\bar{\bar{x}}_i \epsilon o_k)$, $(\bar{\bar{x}}_i r_j \bar{\bar{d}}_k)$, $(\bar{\bar{x}}_i r_j v_k)$, $(\bar{\bar{d}}_i r_j \bar{\bar{d}}_k)$, $(\bar{\bar{d}}_i \rho i_k)$, $(\bar{\bar{d}}_i \epsilon o_k)$, $(\bar{\bar{d}}_i r_j \bar{\bar{x}}_k)$, $(\bar{\bar{d}}_i \bar{\bar{r}}_j v_k)$, $(v_i r_j \bar{\bar{v}}_k)$, $(\bar{\bar{v}}_i r_j \bar{\bar{x}}_k)$, $(\bar{\bar{v}}_i r_j \bar{\bar{d}}_k)$. These structures are considered to be syntactically correct.

The definition is shown in Figure 2.4. It is easy to note the connections between this diagram and diagrams shown in Figure 2.2 and 2.3 as well as Table 2.3. Let us emphasize that this paragraph deals only with the fact that the 13 triads introduced by definition 2.17 exhaust all possible simple syntactical structures of SCL.

Definition 2.18. The set of complex nuclear structures of SCL consists of the following types of structures: $(\alpha_i(m_j r_j)\alpha_k)$, $(\alpha_i(\succ\!\!|_j r_j)\alpha_k)$, $(\succ\!\!|_l(\alpha_i r_j \alpha_k))$, $(f_l(\alpha_i r_j \alpha_k))$, $(\alpha_i \bar{r}_j \alpha_k)$, where α_i, α_k - any of the symbols permissible in accordance with definition 2.17 at a given place.

The latter structure contains $\bar{r}_j$ - a negative sign of the relationship r_j, which is interpreted as the presence of the relationship "not r_j".

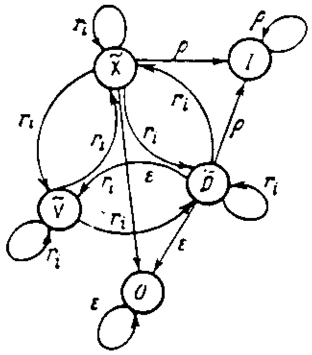


Figure 2.4

Definition 2.19. *Syntactically correct for SCL* are all nuclear structures corresponding to definitions 2.17 and 2.18 as well as structures obtained in the following way. If (β_1) and (β_2) are two syntactically correct structures, then $((\beta_1)r_j(\beta_2))$, and $((\beta_1)\rho i_k)$, $(\beta_1)\epsilon o_k)$, $((\beta_1)\&(\beta_2))$ are also syntactically correct structures. No other syntactically correct structures exist.

This definition fully establishes the SCL syntax. Let us note, that the conjunction can be omitted if it is replaced by the relationship *at the same time*.

2.7 Discrete Situational Networks (DSN)

The situational control language, which became familiar to us in previous paragraphs, gives us the means for description of situations forming at the object of control and enables us, as we shall see later, to make decisions with regard to control. In many problems a certain description of the structure of the object of control is also necessary. However, it is desirable that the language, that would permit us to reflect the necessary information on the structure and functioning of the object of control, is in agreement with the SCL that reflects situations forming at this structure.

Such languages will be discussed in detail in the chapter dealing with the problem of extrapolation of the development of situations and evaluation of the effect of various solutions for control. Here, we will describe a simple version of such a language based on the model of the object of control, which is called the *discrete situational network* (DSN).

Let us analyze three sets: the set of sources I, the set of outputs C and the set of resolvents P. All three sets will be considered finite for each concrete DSN. Any vertex of the DSN can be considered to be one of the elements of the above mentioned three sets. Let us introduce also a set with elements called *objects* and the set X with elements called *characteristics*. Two vertices of the DSN will be connected if the objects can cross over from one vertex to the other, bypassing the remaining vertices. The connection can be oriented or nonoriented.

Now, we will describe the basic elements of the DSN.

Definition 2.20. The functioning of the source consists of the generation (during each functioning cycle of the DSN) of one or several objects with given characteristics according to the law of their generation or the law of nongeneration of objects during certain functioning cycles.

Hence, it follows specifically that the DSN functions in discrete cycles. The law of generation of objects in sources of the DSN can be either definite (for example, according to a fixed schedule) or probabilistic with a known distribution.

Definition 2.21. The functioning of outputs consists of the removal from the DSN of objects that entered the output.

Resolvents of the DSN are divided into two types - passive and active resolvents.

Definition 2.22. The passive resolvent of DSN represents a delay of a definite number of cycles; the number is determined by the type of the object and by this resolvent.

Thus, the passive resolvent can imitate in the DSN the waiting-in-line or the loss of time spent for travel between vertices of the DSN if the path between them passes through a passive resolvent.

Definition 2.23. Objects that have entered the active resolvent either change their characteristics or an alternative on subsequent motion in the DSN is found for them, or both things are taking place.

It follows from the given definitions that sources should not have incoming connections, and outputs should not have outgoing connections. Passive resolvents can have an arbitrary number of incoming connections and only one outgoing connection. Active resolvents will have at least one incoming and one outgoing connection, but their number can also be greater.

In many cases, it will be useful to introduce also *positions*. Their role consists of establishing the position of objects in the network when the objects are outside of resolvents, sources or outputs. As a rule, the positions will be numbered. Figure 2.5 depicts a DSN. Sources I are presented as squares, passive resolvents P as circles, active resolvents AP as double circles and outputs as triangler. Positions are denoted by small black dots.

A graphic representation of the DSN is not absolutely necessary, but it is visually helpful. A DSN can be defined purely functionally by determining the laws of generation of various types of objects in sources, laws of transposition of objects in the DSN, laws of their handling in passsive and active resolvents. These subjects will not be discussed now; they will be discussed later in a chapter dealing with extrapolation of the development of situations. This paragraph has another goal. We will demonstrate by means of examples how the SCL can be used in the description of static situations forming at a given moment of time in the DSN.

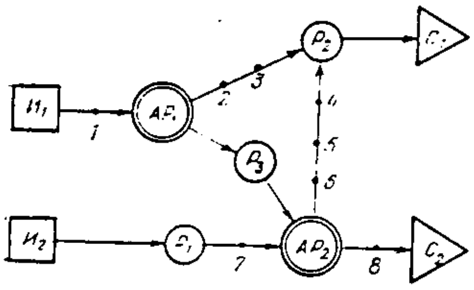


Figure 2.5

Before crossing over to examples, let us introduce the concept of an object in the DSN:

$$O_k = <t_1, t_2, \ldots, t_m, w_1, w_2, \ldots, w_n>.$$

Here, the first m characteristics of the object display a time nature (e.g., total time elapsed from the moment of emergence of the object in the DSN, total time lost for handling in active resolvents, time spent waiting for service in each of the active resolvents, etc), the remaining characteristics can be arbitrary. For transport-type objects of control, they can have the following meaning: type of the object, speed at a given moment, priority of transfer, priority of handling, overall dimensions and weight, destination of transfer. These characteristics can be different for other systems (e.g., when using the DSN to describe a document-flow system in organizational systems or in conveyer production). [i.e. conveyerized production line]

Example 2.5. Let us assume that traffic lights in the area of a city shown in Figure 2.6 must be controlled. The area includes two intersections with controlled traffic; the intersections are connected by streets with one-way traffic. These intersections are shown on the diagram as circles -- trafic-light signs. To describe the situations which can develop in such an object of control, let us shift from the contrete object of control to the DSN.

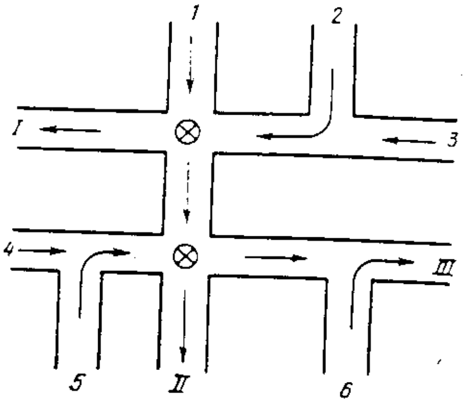


Figure 2.6

The resultant DSN is shown in Figure 2.7. The indices of sources in the diagram coincide with numbers given in the preceding diagram. Indices of outputs correspond to Roman numerals in the same diagram. The upper intersection corresponds to AP_1, the lower to AP_2. Positions have been chosen arbitrarily. They correspond to certain cross sections of main streets where cars are taken into account. Instead of a continuous flow of cars, which takes place in reality, the objects corresponding to cars in the DSN, move discretely across positions and main vertices of the network. On segments without positions, the movement occurs in one work cycle of the DSN. For purposes of control, source H_6 can be disregarded since it does not effect the switching of traffic lights in the controlled intersections. Positions 1, 2, 3, 6, 7 can be regarded as STOP lines before a corresponding intersection.

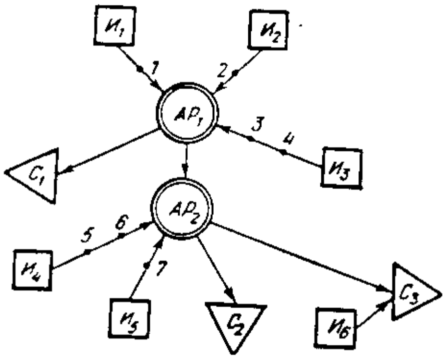


Figure 2.7

Let "a" be the concept of the car. The triad $(a\rho i_k)$ will be interpreted by us as a car with a state license number i_k which actually is its *name*. If b denotes the position, $(b\rho i_m)$ will denote the DSN position with the number i_m. Then, the notation $((a\rho i_k)r_{39}(b\rho i_m))$ corresponds to the fact that the car with the number i_k is in the position numbered i_m. For reasons of simplicity and clarity of the description, no cars will be placed in vertices. In other words, cars are found only in position of the DSN. Thus, in description of static situations (instantaneous photos of the position) we don't have to discuss sources or outputs; we can confine ourselves to positions and active resolvents. Active resolvents will be marked h and given the names 1 and 2. Active resolvents can be in two states. In the first state, the green traffic light is on in the horizontal direction of streets (Figure 2.6). In the second condition - a red light. These conditions will be designated as g_1 and g_2. Then, the notation $((h\rho 1)r_{59}g_2)$ means that, in the horizontal direction of the street passing through the intersection, the first traffic light is in such a condition as to display a red light.

Let us introduce two characteristics of cars: v - speed, p - type. Values of speed will be taken from the set {0,5,10,...140}, values of the type - from the set {q_1 - special, q_2 - ordinary}. Then, the notation $((a\rho i_1)r_{39}(b\rho 1))((a\rho i_1)r_7(v\rho 0))((a\rho i_1)r_7(p\rho q_2))((h\rho 1)r_{59}g_2)$ corresponds to a situation when an ordinary car with the number i_1 stays in position 1 before the first traffic light, because there is a red light in its direction.

Such a description can be given for each car in the DSN at a given moment. The total of such descriptions represents a complete discription of a static situation which exists at a given moment at the object of control.

Example 2.6. Let us discuss an enterprise which has a planning department, accounting, and a personnel department. A certain document flow takes place among these subdivisions. There are several types of standard documents and nonstandard documents. The characteristics of documents can appear as names of attributes, names of ?fillers? for these attributes, etc. Each document can be considered an object in the DSN with three sources, three outputs and three active resolvents. In other words, each subdivision of the enterprise seems to be divided into three parts: source, output and an active resolvent (we would not want them to be passive resolvents, since it would mean that, in certain instances, the documents are not being processed but are "piling-up"). We can introduce positions that correspond to stacks of documents awaiting their turn on the desk of one or another employee. If necessary, we can develop the structure of each subdivision, and make each document-handling employee an active resolvent. What is important here is the fact that the new DSN will not essentially differ from the DSN of the system of intersections, and a certain object, a document in our case, will not differ from a car. Only the meaning of its characteristics will be different, and different laws will be used to describe the functioning of the new DSN.

The given examples depict one important idea. If objects of control, which are different in their nature, can be represented in the form of a certain DSN, the description of current situations at these objects can be uniformly developed by means of universal SCL relationships of the type *to be in the position with the number..., to have a name..., to have priority...* etc. This fact lets us hope that it is possible to create a universal SCL (universal in its composition) for description of current situations in the DSN. Such hopes are not groundless. We will return to this problem in the chapter on extrapolation of situations.

2.8 Other Languages for Description of Situations

The situational control language is not the only language that can be used for description of current and complex situations. In this paragraph we will discuss three different languages that are suitable for this purpose even if they display certin shortcomings when used in situational control.

One of the languages, *the RX-code*, served as prototype in the development of the SCL. Only concepts and relationships are used in the RX-code language. Each concept, with the exception of basic concepts, is defined by the following type of expression:

$$X_n = r_1 x_1 r_2 x_2 \ldots r_{n-1} Y_{n-1}$$

The network given in Figure 2.8 is a graphic illustration of this expression. It seems that the concept x_n is defined through concepts x_1, $x_2 \ldots$, x_{n-1} by means of relationships r_1, $r_2, \ldots$, r_{n-1}.

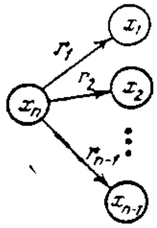


Figure 2.8

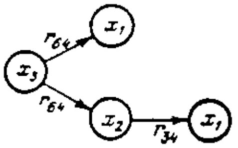


Figure 2.10

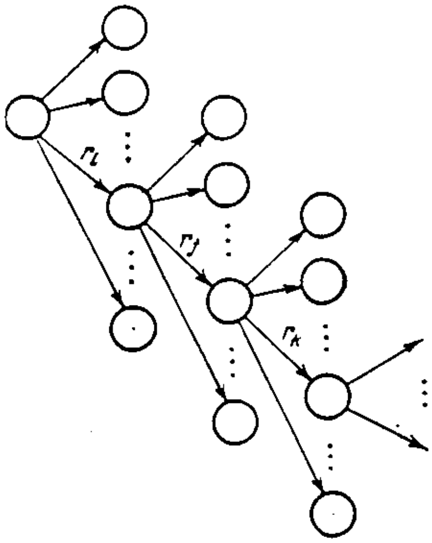


Figure 2.9

Example 2.7. Let us analyze the RX-code of the type $x_4 = r_{71}x_1r_{45}x_2r_2x_3$. If x_1-- power, x_2-- amplification, x_3-- device, and the relationships are taken from Tables 2.1 and 2.2, this code can be treated as the definition of the concept *power amplifier*.

When using RX-codes, it is possible to replace the x_i with RX-codes different from x_n, which will define these x_i. The process oof such a substitution can be repeated as many times as necessary. A fragment of the resulting tree-like structure is shown in Figure 2.9.

Contrary to notations in the situational control language, the RX-codes will not allow us to clearly indicate connections between x_i which are used to define the x_n. If x_n is the code of a certain situation, the components x_i entering into its description, do not interact in any way with each other. If their interaction must be reflected, the components must be duplicated in the description.

Example 2.8. Let x_3 be the code of a situation which is defined by the following description: "An ambulance approaches the intersection, a truck moves behind it". Let x_1 be an ambulance and x_2 be a truck. The description of the situation x_3 by means of the RX-code will be of this type

$x_3 = r_{64}x_1(r_{64}x_2r_{34}x_1)$.

The network given in Figure 2.10 graphically corresponds to this notation. We will not dwell on problems that call for the repetition of notations in RX-codes, because it is not our intention to endorse the use of this language in situational control, even through the language has had its advocates. Let us note only one feature of the RX-code language. It can be used for defining concepts which form classes of objects that are not interconnected. In other words, RX-codes can define classes by means of a simple enumeration of their elements. For example, if x_1 is a table, x_2-- a bed, x_3-- a secretary[1] etc., then

$X_n = r_4x_1r_4x_2r_4x_3...$

denotes the concept of *furniture*. Such a concept can't be introduced into the SCL unless we apply the operation of disjunction between simple nuclear structures of this language.

The second language for description of situations, which is still popular among situational control specialists, is the language of the *calculus of predicates of the first order*. This language can express everything that can be expressed by the SCL if we disregard modifiers and all quantifiers with the exception of those that are accepted in the calculus of predicates of quantizers $\forall$ and $\exists$. There are advantages in the calculus of predicates; they are related to the presence of the developed procedures of deduction conclusion. We will discuss this later. Assuming that the majority of readers is familiar with the classic calculus of predicates we won't waste tthe time to describe it and limit ourselves to only one example.

Example 2.9. Let us consider the description of the following situation. "Between the robot and the warehouse there is a pit to the left of which stands an excavator". Let us introduce the two-position predicate $P_1(x,y)$ with the value x positioned to the left of y. Let us introduce also the three-position predicate $P_2(x,y,z)$ with the valuey positioned between x

---

[1] i.e. piece of furniture called a secretary

and z. If x*--excavator, y*--the pit and z*--warehouse, then the conjunction $P_1(x^*,y^*)$ & $P_2(x^*,y^*,z^*)$ describes the initial verbally-expressed situation.

Lesser known among the developers of situational control systems is the third language for description of situations - *the universal semantic code* (USC). A triad of the type (SAO) represents the simple nuclear structure in the USC. In this triad, the S corresponds to the subject performing action A, and O is the object toward whiiich the given action is directed. For example, the statement "A crane is unloading a ship" corresponds to such a triad. Included in the USC is a closed system of operations that enable us to form more complex series from simple nuclear structures.

Example 2.10. Let us call S_1-- aircraft, S_2-- refueling truck, A_1-- to perform landing, A_2-- to move toward, O_2-- parking area Nr12. Then, the notation $(S_1A_1O_1)A_3(S_2A_2O_2)$ in USC will be interpreted as "The aircraft is landing and, at the same time, the refueling truck is moving toward the parking area Nr12". The line above O_1 denotes nonoccupancy of the position since the action *to perform landing* refers to the object S_1. The nonoccupancy of action A_3 in USC is interpreted as the relationship of simultaneity. If A_3 is interpreted as *cause - effect,* the notation $(S_1A_1O_1)A_3(S_2A_2O_2)$ will correspond to the test "The refueling truck is moving toward the parking area Nr 12 because the aircraft is landing".

The following principles were realized in the development of the USC.

1. A full explication of the meaning. In other words, any notation in USC has only one meaning and this meaning is at a certain metalevel. For instance, the first of the series given in Example 2.10 corresponds to the metameaning "Subject X performs action directed toward itself and, at the same time, subject Y performs action directed toward object Z". The concrete defining of the metameaning takes place with concrete interpretation of all elements included in the series. The requirement of a full explication of the meaning results in the necessary introduction into the USC of various means to give the notation a more precise meaning. For instance, the natural-language phrase "machine tools operate" can be interpreted in different ways; it can have different meanings. It can be regarded as a general statement referring to machine tools in general or as a statement referring to certain specific mechine tools, or as a statement referring to a certain moment in time, as well as a statement which is always valid. Special markers are introduced in

the USC for such expressions; they help to ascribe a more precise meaning to a given expression. Thus, the following four USC strings emerge: $\{S\}_{\text{ᗡ}}AO$ -- "These machine Tools are operating here and now (operate simultaneously)", $\{S\}_{\text{ꓘ}}AO$ -- "Certain machine tools are operating", $\{\{S^1\{\}_{\text{ꓘ}}AO$ -- "All the machine tools are operating". The interpretation of such strings in terms of machine tools denotes that S^1 in the strings is interpreted as *machine tool* and *A as operate.* Braces in the USC denotes plurality. Double braces correspond to the community quantizer of the calculus of predicates, and the line refers to the existence of such a calculus. Signs ꓘ and ᗡ are characters specific to the USC and refer to quantum and deictic statements (*always and everywhere* and *here and now*). In addition to these means, the USC uses symbols of general logic negation, modal operators and numerous special symbols. It is important to note, that the latest USC versions can have a specific meaning for each and every string permissible by the USC syntax if this is done at a certain metalevel.

2. The USC realizes the principle of independence from the subject area. With various interpretations of the symbols entering into the string, one can obtain descriptions which are suitable for any imaginable subject area. This situation is of fundamental nature. In the situational control language, the principle of complete explication of the meaning has not been fully realized. It seems that the relationships that form the basic set "are lumped together". They are concrete up to different levels of commonality which fact calls for the introduction of modifiers to relationships. In languages such as the calculus of predicates, the principle of explication of the meaning has not been realized at all. Notations in these languages make it impossible to compare their meaning at any level of commonality.

3. Semantics which is realized in the USC at the metalevel, i.e. without interpretation of symbols in the strings. This property of USC also differentiates this language from the language of the calculus of predicates where the meaning of predicate formulas depends on one or another agreement. In this respect, the situational control language with a given list of basic relationships and their interpretations approaches the possibilities of the USC.

4. As mentioned earlier, the calculus of predicates and other purely logical calculi have one advantage, namely, we can often formulate procedures of conclusions of a goal-oriented nature, which enable us eithe to obtain the necessary transformation with a finite number of steps or prove that the conclusion of interest to us can not be formulated. Unfortunately, the SCL is also inferior in this respect when compared to the logical calculus language. The USC contains a set of rules for transformation of series from one into the other. As of today, however, there is no proof of either the completeness of the set of the efficiency of the transformation procedure, which is defined as the ability to stop transformations after a finite number of steps if the goal is unattainable.

The above indicated features of USC make the language quite attractive for its utilization in situational control systems. We must mention, however, that the structure of USC strings (unlike that of SCL) is such that the interpretation of their meaning by man is quite difficult and, as of today, we have no effective algorithms to transform natural-language texts into USC notations.

2.9 Languages for Representation of Knowledges**

When working with computers, the described situational control language as well as the other languages are transformed into special languages for *representation of knowledges*, which are convenient for organization of computer procedures. In the *knowledge representing languages (KRL)* used at the present time, the basic unit is the *slot*. The slot has the following form {<name of the slot>;<f_1><v_1>; ...,<f_n><v_n>;<q_1>; ...;<q_m>}. Here f_i-- names of attributes that are characteristic of the given slot, v_i-- values of their attributes or sets of these values, and q_i-- various references to other slots.

Example 2.11. {<AUTOMOBILE>; <NUMBER SIGN><MMG 06-40, MMM 65-12>; <CONDITION><IN GOOD WORKING ORDER, IN MAINTENANCE>; <TYPE><GAZ-24, GAZ-24>; <COLUMN>,<GARAGE>, <DRIVER>}.

This slot has the name AUTOMOBILE. It contains information on two automobiles with the given numbers; one of them is being repaired, the other is in good working order. Both automobiles are of the same type. Certain other information pertaining to these objects can be found in slots with names COLUMN, GARAGE, and DRIVER. For example, the slot DRIVER may have the following form: {<DRIVER>; <FIO><Ivanov Ivan Petrovich, Sidorov Nikanor Stepanovich>;<NUMBER SIGN><MMG 06-40, MMM 65-12>}. When shifting from the slot AUTOMOBILE to the slot DRIVER, we can determine that, at the present time, Nikanor Stepanovich Sidorov is not on the course but he might be in the garage or on vacation or... We can't tell for sure what is happening to Nikanor Stepanovich at the given moment, we know only one thing. He is the driver of the motor vehicle MMM 65-12 which is being repaired at the present time.

A larger structural unit - the *frame* - is used to properly arrange the information stored in individual slots. As a rule, the frame structure has the following form: {<name of the frame><name of the slot><value of the slot><name of the slot><value of the slot>... }.

Let us note, that each frame represents some kind of a finished structure which, after the slots have been filled-out in one way or another by values, is transformed into the description of a concrete fact, event, phenomenon or process. Therefore, we can speak of *prototype frames and concrete frames*. Prototype frames store knowledges on the subject area, concrete frames complement these units of knowledge with real data. Data and knowledges in frame structures begin their co-existence which could not have been realized by previous means.

Let us note also that the names of slots are often called *role markings* or simply *roles*. This emphasizes the special feature of these names which can easily be demonstrated, for instance, in the structure of the WORK ASSIGNMENT frame: {<WORK ASSIGNMENT><WHO><value of the slot><WHERE><value of the slot><WHEN><value of the slot><FOR WHAT LENGTH OF TIME><value of the slot><WITH WHOM><value of the slot>}.

Furthermore, when shifting to a concrete frame, certain slots must be filled-out and others can be empty. For the purpose of our example, the role WITH WHOM can represent such a non-obligatory role.

The basic operation in frame languages is the *search by pattern*. A pattern represents a frame in which not all structural units are filled-out, but only these units which will be used to search for the necessary frames among the frames stored in the computer memory. For instance, a pattern can have the name of a frame as well as the name of a certain slot in the frame with indication of the value of the slot. Such a pattern will check the presence in the computer memory of the frame with the given name and the given value of the slot indicated in the pattern. Or, a pattern may contain the name of a certain slot and its value. Then, the search-by-pattern procedure will ensure the selection from the computer memory of all the frames that contain a slot with the same name and the same slot value as indicated in the pattern. Finally, a certain logical function can be assigned from the frame name, certain slot names and slot values, for instance: {<WHO><Ivanov V.V.><WHERE><Minsk><Kiev><WHEN><1.07.83><10.07.83>}. Such a pattern will select from the memory all the frames containing the slot <WHO> with the value <Ivanov V.V.>, the slot <WHERE> with the value <Minsk> or <Kiev>, as well as the slot <WHEN> with the value <1.07.83> or <10.07.83>. In this case, the name of the frame selected by the search-by-pattern procedure will be arbitrary.

In concrete realizations of frame languages for the search-by-pattern procedure, concrete restrtictions are imposed, nevertheless, a wide latitude is available to the user in manipulation of knowledges and data.

Other procedures, which are characteristic of frame languages, are procedures for entering data into slots as well as procedures for introduction of new prototype frames (i.e., new knowledges) and introduction of new connections among them.

The use of frame languages for control problems evoked the desire to introduce typology of basic prototype frames. One of the versions calls for the introduction of four types of frames: *the technology frame, the conflict frame, the production frame, and the indicator frame*. The first type of frame is used in the description of knowledges and data connected with the development of processes in the object of control and the technological models used in the control of a process under normal conditions. The second type of frame serves for classification of conflicting situations, methods of their detection, exposure of their causes and description of methods for elimination of conflict situations. The third type of frame describes the cause-effect relationships which form the basis of the process used in obtaining a certain end product or a result (situation), which is the direct goal of control. Finally, the fourth type of frame defines the structure of the indicators which are used to evaluate the process taking place in the object of control and the final results achieved by the control system.

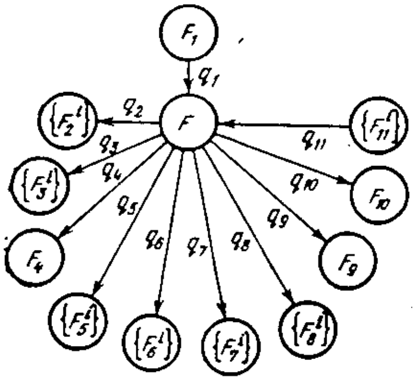


Figure 2.11

Example 2.12. Figure 2.11 shows the structure of a conflict frame where the following frame designators and references between frames are used: F - prototype frame, F_1 - concrete frame, which stores concrete data on the emerged conflict. These data are used in frame F for excitation of any type of connections. Set $\{F_2^i\}$ - all frames which must be analyzed when the conflict described in F_1 emerges. Set $\{F_3^i\}$ contains frames acting as generic frames for frame F. They contain information which defines the class of conflict situations that encompasses also the situation forming the value of slots in F_1. Frame F_4 - the technology frame to which the type of the examined conflict situation belongs. Sets $\{F_5^i\}$, $\{F_6^i\}$, $\{F_7^i\}$ and $\{F_8^i\}$ - sets of frames containing procedures for the analysis of conflicts which arise, correspondingly, in disturbed operation of structural subdivisions in the object of control, in deviation of indicators for semifinished and finished products, in deviation of indicators for material and power flows circulating in the object of control, and in disturbances of performance of the attending personnel. In essence, these groups of frames form the basic information on methods for elimination of the conflict. Frame F_9 contains valuations which enable us to accept only these causes of the conflict situation defined by F_1 which are described by the four groups of frames. In the same frame F_9, a valuation of the degree of conflict of the arising situation can be formulated. Frame F_{10} defines dynamics of the development of the conflict situation, stores the measures previously used to reduce the degree of the conflict of a situation. Finally, the set of frames $\{F_{11}^i\}$ corresponds to the set of frames storing information of interest to the analysis of causes of emergence of the conflict situation the type of which is described in frame F, and its concrete representation is the situation described in frame F_1. The movement along connections q_1 takes place by means of the procedure stored in the prototype frame F.

Such multiple-type systems of frames are similar in their nature to what is called by many authors the *semantic network*. This concept in contemporary literature has no clear-cut definition. Therefore, we will limit ourselves to a form-and-content discussion of semantic networks, mainly touching upon aspects that are of value for their utilization in situational control systems. Let us begin with two possible types of semantic networks: *extensional and intensional* semantic networks (ESN and ISN). Here, we will remind the readers who do no skip over certain parts of the text, that in paragraph 1.6 we briefly discussed the problem of extensional and intensional knowledges. In frame representations, the concrete frame and the prototype frame with its connections correspond to them to a certain degree.

The ESN reflects knowledge of a certain concrete fact or a certain concrete situation fixed in the control system. The ESN consists of vertices and their interconnections. The vertices correspond to certain objects which are present in a given fact or a given situation and are distinguished by means of individual names from a class of same-type objects. Connections correspond to the concrete relationships between the objects in a described situation or fact. Thus, the ESN fixes instantaneous prints of reality.

On the surface, the ISN does not differ from the ESN. Only here, names of classes (not individual names) correspond to vertices, and connections reflect the relationships which are always characteristic of these objects. Thus, certain rules, which are characteristic of that or another problem area, are fixed in the ISN.

Example 2.13. Let us examine an example of the ESN. Let us deal with a group of facts related to a set consisting of three numbers: 0, 1, 2. These facts refer to two relationships determined in the set of the given numbers. The first is the relationship of a strict order, the second relationship is defined by the addition operations. Let Φ_i denote names of these facts; then, the entire group of facts can be defined, for instance, in Tables 2.4 and 2.5.

The ESN, which reflects the facts stated in these tables, is depicted in Figure 2.12. Here, the network vertices marked as squares correspond to elements of the basic set 0, 1, 2; vertices marked as circles are facts reflected in Tables 2.4 and 2.5. Around them we find names of arguments of corresponding relationships. Names of relationships are written in diamondshaped vertices of the network. For reason of simplicity, connections from facts to names of relationships have no special markings in our ESN.

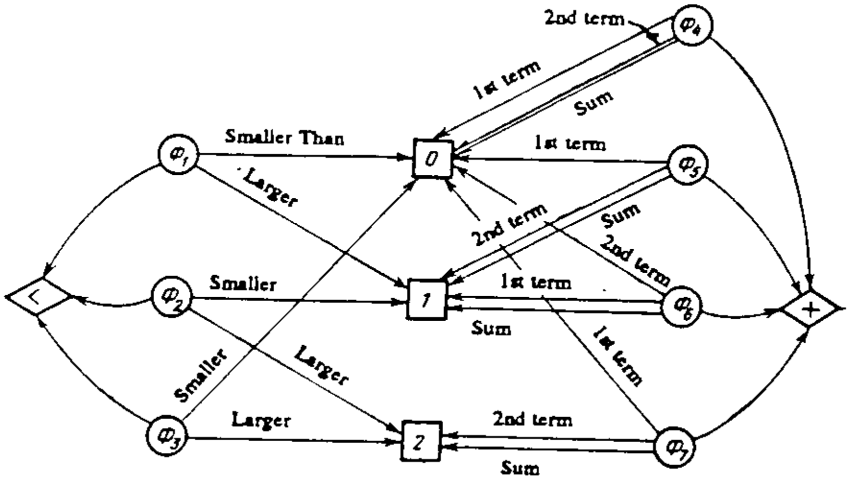


Figure 2.12

TABLE 2.4

Names of Facts	Name of Relationship Smaller (<)	
	Argument of Relationship	
	Smaller	Larger
Φ_1	0	1
Φ_2	1	2
Φ_3	0	2

TABLE 2.5

Names of Facts	Name of Addition Relationship (+)		
	Argument of Relationship		
	1st Term	2nd Term	SUM
Φ_4	0	0	0
Φ_5	0	1	1
Φ_6	1	0	1
Φ_7	0	2	2

Let us note, that in our example, the knowledge of the addition relationship is incomplete. We do not know the sum of 2 + 0 or the sum of 1 + 1. Further, we do not know that the addition relationship in our basic set can not be determined between elements 2 positioned in both terms and between elements 2 and 1 (by two methods) arranged in the first two positions characterizing the addition relationship.

Example 2.14. As an example of ISN, let us analyze the network depicted in Figure 2.13. It defines the concept detached service WORK ASSIGNMENT corresponding to the role frame which has been discussed in the beginning of the paragraph. The names of roles in the figure are marked on the curved lines, vertices of the network contain names of the classes of objects the elements of which can fill-out the vertices in extensionals.

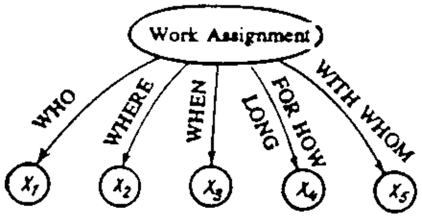


Figure 2.13

Thus, in knowledge representing languages one can isolate the extensional and the intensional components. The various languages are distinguished by the different means used for representation of intensional knowledges.

In conclusion of this paragraph, we will present a brief description of one of the knowledge representing languages that is called FRL (*frame representation language*). For the time being, the FRL is the only language which can be realized on Soviet computers. There are two versions for realization of FRL; one of them uses LISP ES as the base language, the other uses a version of the REFAL ES language. There are several situational control systems that use the FRL for representation of knowledges and their manipulation.

In the FRL each frame represents a structure with a maximal degree of nsertability amounting to five. This structure has the following form:

(name of the frame
    (name of the first slot
        (name of the first cell
            (first datum
                (first comment:

```
                     message)
             (second comment:
                     message)
            ...................
            i-th comment:
                     message))
         (the second datum
            ...................))
      j-th datum
            ...................)))
   (name of the second cell
   ........................)))
   (name of the k-th cell
   ........................))))
(name of the second slot
.............................))))
(name of the l-th slot
.............................)))))
```

In such a structure certain levels might be absent. For instance, comments and messages on certain data might be absent; only one slot might be present in a frame, and the slot might have only one cell; a cell might contain only one datum. The *indicator path* is used in the FRL to indicate the position of a certain portion of information in the frame. As an example we can name the indicator path searching for the message stored in the second comment to the third datum of the first cell of the seventh slot in the frame named I. The path will have the following form (I; s7; f1; d3; c2). Upon its analysis, it is easy to find the necessary "address" in the structure of the I frame. In Soviet realizations of the FRL the indicator path requires a clear indication of all names of the structural parts of the frame. This was done in order to use the standard search-by-pattern operation instead of registers in address search.

To operate with frames, the FRL contains four basic operations, FASSERT, FPUT, FGET and FREMOVE. The FASSERT operator enables us to introduce into the knowledge base the frame which represents the argument of this operator. For instance, let us take the notation

```
FASSERT (WORK ASSIGNMENT detached service
         (WHO
            (Ṡ VALUE
            (Ivanov Ivan Ivanovich)
            (Petrov Petr Petrovich)))
         (WHERE
            (Ṡ VALUE
            (Leningrad)
            (Baku)
            (Ashkhabad)))
         (WHEN
            (Ṡ VALUE
            (2.06.84
               (COMMENT
               (Order of i.o. director
               Sidorov Yevgeniy Stepanovich))))))
```

This frame has three slots with the names WHO, WHERE, WHEN. Each slot contains one cell with the standard name Ṡ VALUE. This name indicates that ordinary data are stored in the given cell. A comment has been added to the datum stored in the Ṡ VALUE cell of the WHEN slot; the comment says that the individual starting his detached service work assignment on June 2, 1984, is doing this in compliance with the order given by i.o. director Sidorov Yevgeniy Stepanovich. The FASSERT Operator is used to enter this frame into the knowledge base of the system permitting a retrieval by the name detached service WORK ASSIGNMENT.

The FPUT operator adds new data to a certain spot of the frame. The argument of this operator is the above introduced indicator path containing at the end a explicit notation of the part to be inserted. For instance, FPUT (WORK ASSIGNMENT detached service; WHO; Ṡ VALUE; Sidorov Yevgeniy Stepanovich; COMMENT; order given by director Petrov Iliy Alekseyevich) results in a change of content of the first slot of the WORK ASSIGNMENT frame. Now the slot will have the following form:

```
(WHO
   (Ṡ VALUE
   (Ivanov Ivan Ivanovich)
   (Petrov Petr Petrovich)
   (Sidorov Yevgeniy Stepanovich)
   (COMMENT
      (by order of director Petrov Iliy Alekseyevich))))
```

If we have to extract certain information from a frame, we use the FGET operator. The same indicator path is used as its argument. For instance, the operator FGET (WORK ASSIGNMENT; WHERE; Ṡ VALUE) will be used to extract from the WORK ASSIGNMENT frame the information stored in a cell of the WHERE slot. The information will be given in the form of a list ((Leningrad), (Baku), (Ashkhabad)).

Finally, the operator FREMOVE with the same indicator path serving as its argument enables us to remove unnecessary information from the frame. If, for instance, we use the operator FREMOVE (WORK ASSIGNMENT; WHERE; Ṡ VALUE Baku) in our frame, the datum Baku disappears from the corresponding cell in the slot WHERE, and the cell retains only Leningrad and Ashkhabad in its content. When the operator FREMOVE (WORK ASSIGNMENT is used, the frame with this name is removed from the knowledge base. (Certain FRL versions use the special operator FERASE to remove a frame).

Since true addressing in the computer memory is not directly connected to the addresing that forms indicator paths, the real location of frames and their substructures in the memory is not known to the user. The correlation between real addresses and indicator paths in entry and search of frames and their substructures is achieved by the control system of the knowledge base. This system has a standard "scavenger" that is put into operation either by means of special signals in the program or automatically after FREMOVE-type operators have completed their action.

In addition to Ṡ VALUE cells, the FRL has five types of cells discernable by the system as special cells. The cells of the first special type are called Ṡ DEFAULT. The content of these cells is used "in omission". Such use of information occurs during search of a slot for information indicated, for instance, in the indicator path of the FGET operator, when it is discovered that the information is missing. Then, the output will contain the information of the Ṡ DEFAULT cell instead. If the first slot of our WORK ASSIGNMENT frame had the form

```
(WHO
   (Ṡ VALUE
   (Ivanov Ivan Ivanovich)
   (Petrov Petr Petrovich))
   (Ṡ DEFAULT
      (For the present, no work assignment directives exist))),
```

then in the presence of the operator FGET (WORK ASSIGNMENT; Ṡ VALUE; Sidorov Yevgeniy Stepanovich), the indicator path would not find the necessary datum. When this fact is signalled, the system would automatically cross over to the content of the Ṡ DEFAULT cell of this slot and deliver the information that, at the present time there are no work assignment directives for Sidorov Yevgeniy Stepanovich.

Two types of special cells, ṠIF ADDED and ṠIF REMOVED, are used correspondingly for operators FPUT and FREMOVE. Data in these cells consist of names of certain procedures used for their retrieval and execution. The procedures are written in the same programming language that forms the basis of the FRL. As mentioned earlier, Soviet systems use LISP and REFAL as their base languages. When the FPUT operator is adding something to the Ṡ VALUE cell of a given slot, the procedures specified in the ṠIF ADDED cell are automatically put into operation. In the FRL, they are called *associated* procedures. Let us assume that we are required to store data on the number of people designated for a work assignment. The first slot of the WORK ASSIGNMENT frame may contain the following structure:

```
(WHO
   (Ṡ VALUE
   (Ivanov Ivan Ivanovich)
   (Petrov Petr Petrovich))
   (ṠIF ADDED
   (AMOUNT)))
```

Here, AMOUNT is the name of the procedure written, for instance, in LISP. This procedure operates in the following way: the number of new data annexed to the Ṡ VALUE cell of the WHO slot is added to data stored in the given location of the memory.

The ṠIF REMOVED is used in the same way as the ṠIF ADDED cell, but it is used in removal of data from the Ṡ VALUE cell by means of the FREMOVE operator. When in the above cited example certain individuals are deleted from the WHO list, and the structure of this word contains the ṠIF REMOVED cell storing the DECREASE datum, a procedure is actuated that will decrease the number of individuals designated for work assignment by the number indicated in the indicator path of the FREMOVE operator.

The ṠIF NEEDED cell may also contain names of procedures. These are also called associated procedures since they are not performed in the FRL but in the basic programming language. Procedures stored in the ṠIF NEEDED cell are performed when the indicator path of the FGET operator contains a reference to this cell and the name of the corresponding procedure. If the name of the procedure is not indicated all procedures stored in the ṠIF NEEDED cell are performed sequentially.

The associated procedures enable us to utilize all the computational-type possibilities that are available to the basic level programming system.

The last special cell is the cell called Ṡ REQUIRE. The cell stores names of certain predicates defined on the values of the given slot. The predicates acquire a true or a false value depending on the real cells and the data which are stored in the slot at the given time. This makes it possible, for instance, to receive answers to questions such as "Is it true that Moscow is not included among the cities listed in the WHERE slot?", "Is it true that no work assignments were given in May?", "Is it true that Ivanov Ivan Ivanovich and Petrov Petr Petrovich are the only individuals who left on work assignments in June of 1984?". The predicates, just as the associated procedures for checking the true value, retrieve appropriate procedures written in the basic programming language. As an answer to the action of the FGET operator the indicator path of which contains the Ṡ REQUIRE cell after the name of a certain slot, the system gives a list of true values of all predicates for the slot data that act as arguments in these predicates.

Until now we have examined isolated frames and operations with them. However, the FRL language provides the possibility to connect the frames into a common network. First of all, the frames can be interconnected by means of the family--kind relationship. In the FRL, the connections which define the transition of a kind to a family are called AKO-connections (abbreviation of a kind of, i.e., *one of*). Reverse connections are called INSTANCE-connections. The are used to memorize the upward path along AKO-connections in order to return to the initial frame. AKO-connections allow us to reduce the volume of information stored in the knowledge base. Information that is common to several frames is collected in a special frame where it is stored as a single sample. To enable the lower-level frames to retrieve this information, a special slot called AKO is provided in them. The data stored in the Ṡ VALUE cell of this slot consist of names of frames to which the given frame is connected by means of AKO-connections.

Assume for instance, that in addition to the WORK ASSIGNMENT frame the knowledge base stores also the RECOMBINATION frame. According to existing standards, if an individual departs for a work assignment, his wages must be taken into account by recombination. This means that the WHO slot will be common in WORK ASSIGNMENT and RECOMBINATION frames. To avoid duplication of information, one can set-up a special frame and call it the WORK-ASSIGNED INDIVIDUAL. The new frame will have the following form:

```
(WORK-ASSIGNED INDIVIDUAL
   (WHO
      (Ṡ VALUE
      (Ivanov Ivan Ivanovich)
      (Petrov Petr Petrovich)))
      (INSTANCE
      (Ṡ VALUE
      (WORK ASSIGNMENT)
      (RECOMBINATION)))
```

The WHO slots will be missing from WORK ASSIGNMENT and RECOMBINATION frames. Instead, they will have a slot called AKO where the Ṡ VALUE cell will contain the datum WORK-ASSIGNED INDIVIDUAL. In the presence of such cells in WORK ASSIGNMENT and RECOMBINATION frames, the operator FGET (WORK ASSIGNMENT;

WHO; Ṡ VALUE), for instance, will be implemented in the following way. The system will analyze the indicator path and find the frame called WORK ASSIGNMENT, but it will not be able to find the WHO slot. This will represent the signal to automatically turn over to the AKO slot of this frame. After the data stored in its Ṡ VALUE cell have been read, the system begins to look for the frame WORK-ASSIGNED INDIVIDUAL in which it searches for the WHO slot. After finding this slot, the system provides the information the requirements for which have been formulated in the indicator path. After the WORK-ASSIGNED INDIVIDUAL frame has been found the FGET operator records into its INSTANCE slot the name of the frame from which the transition by means of the AKO-connection originated. That is why we have highlighted the data in the INSTANCE slot of the WORK-ASSIGNED INDIVIDUAL by wide-spread type. Actually, no data are stored there until a certain operator crosses over to a given frame by means of the AKO-connection. Certain systems will already contain such data, and the operator arriving at the frame by the AKO-connection will only activate a certain datum by means of a special mark.

The FGET operator, as well as other FRL operators, will refuse to perform a given task after all the search possibilities imparted by AKO-connections and contents of Ṡ DEFAULT cells have been exhausted.

As of today, the existing versions of the FRL have poorly developed procedures for introduction into the frames of relations different from relations of the tree-type structure (*family--kind, whole--part, class--element*, etc). Lately, however, some successes has been noticed in this direction, which fact makes the FRL even more promising for utilization in situational-control and other systems that have been traditionally placed in the artificial intelligence sphere.

Although special types of frames, which are very convenient in practical systems, are not found in the FRL, another knowledge representing language, the KRL, contains seven different types of frames: *a basic frame, an abstract frame, a specialization frame, an individual frame, a declarative frame, a relationship frame and a conditions frame*. This language when compared to FRL contains also a larger amount of cells of different values. However, even this language has poorly developed means (at least the versions operating in computers) for an efficient organization and manipulation with network frame structures.

2.10 Transition from Texts in the Natural Language to an Internal Representation**

There is an obvious discontinuity between the description of situations in a natural language and internal representations of information about situations in the control system. Therefore, one of the difficult problems to be solved in the development of a situational control system is the transformation of verbal descriptions into internal representations. This problem is closely related to the problem of development of dialogue systems based on a natural language. However, for its solution it is not enough to utilize available linguistic processors safeguarding the dialog. It is further necessary to obtain specific information connected with the functioning of the Analyzer depicted in Figure 1.7. Let us remind the reader that this problem includes the classification of input information in accordance with the problems to be solved by the control system. There are three types of such problems: replenishment of the system with new information on the object of control or control methods, formulation of an answer to a certain inquiry based on the information already stored in the system, and search for the solution in a situation the description of which has already been entered into the system. The first two problems are auxiliary problems, the third is a basic problem. However, their division into three clases should be accomplished in transformation of the input text into internal representations. This is the stage during which the system must "comprehend" what is required of it. Therefore, the Analyzer can be regarded as a component of the linguistic processor.

Figure 2.14 depicts the traditional structure of a linguistic processor, where the analysis of the entering text is accomplished sentence by sentence as they enter into the system. At first, the dictionary supplies for all words of the sentence the necessary morphological and syntactical characteristics, followed by the gender, the singular or plural forms, the case, inflection, etc. Then, a syntactical analysis takes place; its goal is to reveal the syntactical structure of the sentence. Its structure has the form of a tree reminding us of the trees obtained in school during parsing of sentences. At this stage, the predicates and the name group of nouns as well as all other parts of the sentence that depend on them are found. At the semantic-analysis stage, the semantic structure of a sentence is developed, and the sentence is interpreted in terms used by the knowledge representing language. At this stage, the linguistic processor interacts not only with its dictionaries but also with the data and knowledge bases which are depicted in our figure as the Problem-Area Model Unit. After this interpretation it is the turn of the Analyzer which corresponds to the Pragmatic-Analysis Unit in Figure 2.14.

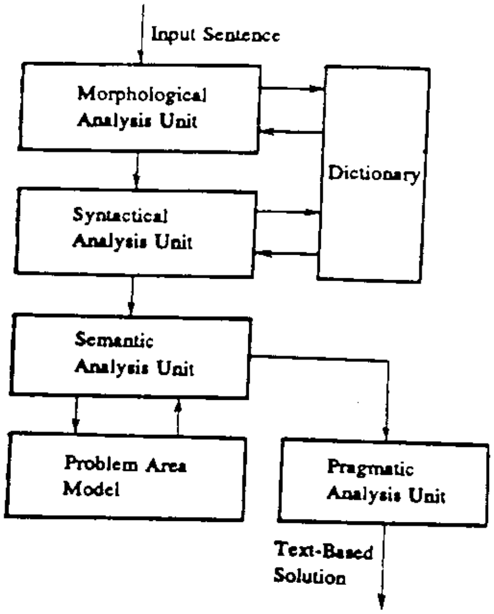


Figure 2.14

In real linguistic processors and dialogue systems the structure shown in Figure 2.14 can have different forms. The successive characteristics of individual analysis stages may become successive-parallel when certain stages begin before the preceeding stages have been completed. Some stages (e.g., the morphological analysis stage) may be absent.

In development of systems oriented toward narrow problem areas, we can completely abandon the structure shown in Figure 2.14. Furthermore, the existing languistic processors have the defficiency of handling only individual sentences without catching the meaning of the entire text; this fact severely limits the possibilities of such systems. This must be taken into account especially in the development of the type of control systems discussed by us. As a rule, the description of a situation has the character of a text consisting of a series of sentences. Here, we must recall the statement made by the noted Soviet psychologist A. N. Sokolyanskiy, who developed a method to teach a language to the blind and deaf-mute people. He wrote: "The text is the main element, never the individual word or the individual sentence. The name is not consolidated with the object but with the situation where the object is used"[1]. We are not going to develop here his ideas since this would consitute a deviation from the main theme of this paragraph. Comments to this paragraph contain bibliographic references where such problems are discussed in detail. We, however, will proceed to the description of methods for transition from a text into representations which are characteristic of the situational control language.

In §2.3 we had cited tables of static and dynamic relations, which can be used in describing situations arising at the object of control. We spoke of the fact that the number of base relationships of and type is finite and includes in itself no more than 200 relations. Some of them are shown in Tables 2.1 and 2.2. In §2.4, we had cited Table 2.3, which shows that in the set of relations it is possible to introduce a certain hierarchy (in this case a two-level one).

Now let us select a certain problem area and set up a dictionary of basic concepts, which will be utilized to describe situation characteristic for this area. For transition from verbal descriptions to notations in LSC it is necessary to learn how to find a correspondence between certain parts of the text and symbols used in the LSC. In §2.3 we have defined eight classes - concepts, names, relations, actions, quantifiers, modifiers, modalities, and estimates (valuations). In addition, for relations and actions we gave a more detailed classification inside these classes. With a small problem area (in the sense of the utilized dictionary for describing situations and decisions made on control), the required correspondence can be specified with the aid of special tables.

Example 2.15. Let us assume that Table 2.6 is stored in the memory of the system.

---

[1] Sokolyanskiy, A. N., Teaching a Language to the Blind and Deaf-Mute. Theses of the conference on machine traslations. M. 1958.

Table 2.6

LEXIC UNIT	FRAGMENT IN LSC	LEXIC UNIT	FRAGMENT IN LSC
TEMPERATURE	g	Because of	r_{73}
PRESSURE	b	Pump	a
IN THE COLLECTOR	$r_{38}c$	Pump	a
HAS RISEN	d_3	Necessary	f_1
HAS FALLEN	d_2	Turn on	$r_{72}d_2$
NORMAL	d_4	No. 5	ρi_1
STOP	d_1	No. 8	ρi_2
STOPS	d_1	No. 9	ρi_3
BECAUSE OF	r_{49}	Autocar	h

Let the following text have been given to the system drive: "Pressure in manifold No. 5 has dropped because of stoppage of pump No. 8. Necessary to turn on pump No. 8 on pump No. 9". Analyzing this text sequentially, it is possible to construct an internal concept that corresponds to it. First, instead of the word pressure is substituted b - its description in the internal language. Next are found fragments of representation for the lexemes "in the manifold" and No. 5". They are $r_{38}c$ and ρ_{i1}. For assembling, from the found fragments, of a description of type $(br_{38}(c\rho_{i1}))$, it is necessary to have certain additional information. It is stored in the system in the form of special *production rules*. A production rule (production) is an expression of the type $\alpha \rightarrow \beta$, which is read as follows: "If α, then β". Rules of this type will be the object of our discussion in the next chapter. The left part of this production sets a certain condition, the fulfillment of which assures the applicability of the given production. On the right side of the production is described a certain transformation which can be executed. An example of and a production is the rule "If there are available adjacently standing fragments x and ρy, where x belongs to a set of concepts and y to a set of names, then those fragments are unified into a construction $(x\rho y)$". Thus is formed from the fragments a syntactically correct notation in LSC.

Further analysis of a text gives new fragments, which combine with each other by means of production rules. For instance, if during an analysis the term *because of* is encountered, it is a priori not clear, which fragment of the ones in the table corresponds to the case at hand. It is clear that in the case of "because of pump stoppage" is proper to substitute the relation r_{49} for "because of". But it is just as necessary in the phrase "the auto car drove out from behind the pump" to replace *from behind* by the selection r_{73}[1]. The choice of one or another alternation is even here made by means of a special production rule. As a result of such procedure, there is formed, from the initial text, an internal representation in LSC, that has the following form:

$$\{((a\rho i_3)r_{59}d_1)r_{49}[(br_{38}(c\rho i_2))r_{59}d]f_1\{[c\rho i_2)r_{72}d_3]V[(c\rho i_3)r_{72}d_2]\}\}.$$

In place of the sign of disjunction, it is possible to introduce the corresponding relation. Note, that in the first proposition that part of it is formed first, which in the internal representation stands to the right of the relation z_{49}. It corresponds to a consequence. Whereas the part describing the cause, is formed later. Their correct position with respect to z_{49} is also attained due to special production rules.

The approach to the transformation of the text from the natural language into an internal representation, which was demonstrated in Example 2.15, differs in principle from the traditional approach of elucidation of the meaning of the text used in linguistics. In systems of control, not the meaning of the text itself is important, but what the system must do when it receives the text. In other words, the text must cause a definite reaction of the system. It is this approach that allows us, in many practical systems of

---

[1] In Russian, the expressions "because of" and "from behind of" are identical. In English this difficulty would not occur. (Translator)

natural language intercourse, to utilize not the structure shown in Figure 2.14, but a direct transition from the text to the internal representation. In the novel by Lawrence Stern, "Life and Opinions of Tristram Shandy," this situation is illustrated by the following words: "My mother...could use a difficult word for 20 years in a row, and also reply to it, if it was a verb, in all types and conjuctions, without troubling herself with questions about its meaning". In the comments to the present chapter, the reader will find referrals to literature in which are described systems based on direct transition from texts to internal representation. At the present time the practice of construction of translation blocks for transition from texts in a natural language to internal representation in systems of situational control is such, that narrowly specialized translations are required, that are totally oriented toward those problem areas for which they are intended.

§2.11. Formalization of Qualitative Characteristics*

Most difficulties are formalized during transition from a natural language to LSC and, the quantifiers, modifiers, evaluations, and modalities, which man uses purely quantitivly.

Examples of these types of descriptions inherent to man, may be stated as, "the losses from a bad regulation of temperature are quite considerable", "It is possible to switch on an extra engine but this is dangerous". What is hidden behind the phrases cited? Does it mean that the turbine gets out of operation *frequently*? Once a day, a month, a year? What actually are the losses that are *quite considerable*? What percentage do they comprise - 20 or 80? And what regulation of temperature is called *bad*? And finally, should or should not the extra engine be switched on, for this *can* be done, but is *dangerous*?

To these questions there is no straight answer. Mom's language does not operate with the precise meanings of the words high-lighted by italics in our text, and this is understandable, since these words by themselves, generally speaking, do not signify anything. Their meaning is completely defined by a certain concrete reality for the description of which they have been brought in. Hence the statements, "The product requires very many fastener elements" and "In the manufacture of the product very many enterprizes participate" are talking about different *very many's*. In the first instance are understood, apparently, thousands on tens of thousands of fastener elements, while in the second case the number of enterprizes is probably counted in the tens.

Mom, for understanding the meaning of such words and word combinations very often draws on the concept of the *norm*, that characterizes a given phenomenon, fact, or process. The very concept of a norm is also fluid, situative, and depends on concrete conditions in which the text or the speech is perceived. If in a conversation between two Moscow natives the phrase, "I live far from work" is associated with a time expenditure for transportation on the order of an hour, in the case of people living in the Academic Village near Novosibirsk, the same phrase indicates far less waste of time for travel to work. And this is related to the different norm of expenditures of Moscovites and inhabitants of the Academic Village.

Assume that a certain variable x takes on values from segment [a,b] (Fig. 2.15). And let the observations on the values of x allow us to construct the histogram, shown on the same Figure. Then the norm will be characterized by those values of x which have occurred *most often*. Man distinguishes a certain interval of values (x*,x**), inside of which he is ready to call all values *normal*. If values of x lie to the left of the norm interval, then man is inclined to consider them *smaller than norm*, while if the values happen to lie to the right of x**, then man interprets them as *surpassing the norm*. Thereby he operates with a certain scale in which the norm concept occupies a central position.

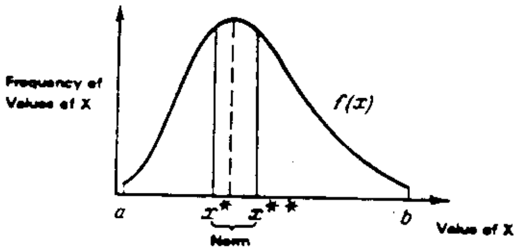


Figure 2.15

As an example of such a scale for estimating distances, consider the following sequence of words: *adjacent; very near; near; not far, not near* (the norm); *far; very far*. As another example, words may be used as a scale for estimating the frequency of occurrence of events or phenomena: *never; almost never; very seldom; seldom; not often, not seldom* (the norm); *often; very often; almost always; always*. The same type of scales can be formed for antonymic pairs of words: *big - small, much - little, strong - weak,* and so forth.

It's possible to carry out the following experiment, which has actually been performed with subjects of different age, social status, and nationality. The subject was issued a stack of cards on each of which was written one of the words or word combinations from the scale. Also, the subject was given a segment of the scale, at the left and right ends of which were situated the antonym words that closed the scale (for example, *never* and *always*, for the frequency scale), while in the middle of the scale was placed the word or word combination corresponding to the norm for the given scale. On the scale are marked divisions the number of which is greater than the number of the words forming the series-scale. The task of the subjects tested consists in indicating the intervals, which, in their opinion, correspond to the words of the scale. On Figure 2.16 is shown a result of such a test performed with a single subject on the construction of a *never - always* scale, when the scale included the words: *never; very seldom; seldom; not often, not seldom; very often; always*. With the concepts *never* and *always* the subjects, as a rule, relate only the terminal points of the segment. The other words are interpreted by them by certain segments, similar to that shown in our Figure.

Figure 2.16

If now a sufficient statistic is collected from a large group of subjects, it is possible to construct some generalized scale by the group of test subjects. Here a certain difficulty arises, since the limits of the intervals for words in the case of different subjects many turn out to be different, which forces us to transfer to a certain histogram, an example of which is shown on Figure 2.17. Five curves shown on it (for the words *never* and *always* there are practically no deviations) are called *functions of affiliation* if the words *very seldom; seldom; not often, not seldom; often;* and *very often.*

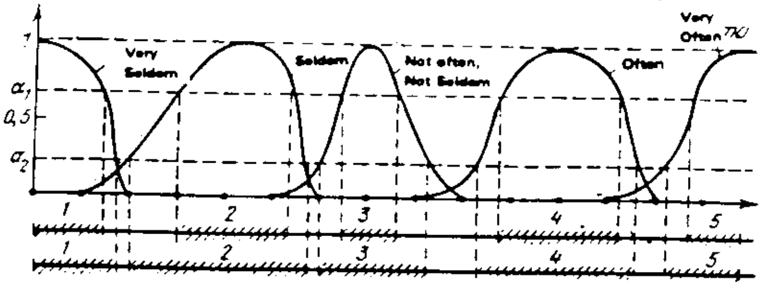


Figure 2.17

The concept of the affiliation function, which we originated in connection with the processing of experimental material, can be specified completely strictly.

A function of affiliation for a certain set W is the name given to a position $\mu(w)$, defined on all elements of set $W^1 \supseteq W$ and taking on values from segment [0,1].

The meaning of this definition can be interpreted in the following way. If, for a certain $w \in W^1$ $\mu(w)=0$, then it means that w does not belong to W; if for a certain $w \in W^1$ $\mu(w)=1$, then w does belong to W. If however for a certain w $0<\mu(w)<1$, then this means that w belongs to W with a certain possibility estimated by the quantity $\mu(w)$.

Example 2.16. On Figure 2.18, is shown a series of pictures (cartoons).

Figure 2.18

About the left one of these we have no doubts - it is a person. About the right one there is also no doubt. That, of course, is a cat. Let us form, from the set of images shown in Figure 2.18 a set W, which is interpreted as *images of cats*. Then, for the right image, the affiliation function for this set $\mu(w)=1$, and for the left image - $\mu(w)=0$. The other images, apparently, cannot be related with such deep conviction to a given set or be excluded from it. The value $\mu(w)$ obviously increases during an analysis of the images from left to right. But the problem arises, what numerical value of $\mu(w)$ should be assigned to each image. An answer to this question can be given only by human expertise. Having questioned a sufficient number of humans - experts, each of which gives for each image his numerical estimates, it is possible to average it over all the experts and to obtain for each image a value of the affiliation function for the set of cat images. It is of interest to note that the characteristics of the expertise depends much, as was said earlier, on the norm, which the experts adhere to. If in the given social group whiskers are not customary, then possibly all estimates of the expertise will be displaced toward an increase. Still more visually clear this can be demonstrated on an example of forming values of an affiliation function for such a set, for example, as old people depending on the actual age of the people. In a social group of long-livers the age of 50 would hardly receive a large value of the function of affiliation with this set. But in a social group where the life expectancy or the average is not great, the age of 50 may receive a very high degree of affiliation to the old people set.

Let's return to Figures 2.16 and 2.17. Could they be tied to each other? Let's introduce a *level of cutoff* α, when α assumes a definite value in the interval from 0 to 1. On Figure 2.17 are shown two levels of cutoff with $\alpha_1>0.5$ and $\alpha_2<0.5$. Those levels allow us to identify two systems of cross-hatched segments for α_1 and α_2 on the lower axes. Points inside of them correspond to values of frequency which belong, with a possibility exceeding the given level, to sets characterized by our five words. The "Empty" intervals of values contain such values of frequency about which it is not possible to say anything with the given level of possibility. The smaller α, the greater, naturally, is the area occupied by the cross-hatched segments. At a certain value α^* they fill the entire axis, and with further increase of α they begin to superimpose one upon another. This will mean that with a given level of possibility a certain relation [ratio?] of frequency may belong both to a set defined by some word on the scale, as well as to an adjacent set.

From this follows the method of selection of α^* which assures the relegation of all values on the scale to some single set.

Let's consider now a most important problem - the problem of a *universal scale*. As long as values of the affiliation function are going to be wholly determined by the semantics of a concrete problem area, it is difficult to make any kind of general conclusion or scales or to construct procedures of operating with them.

Let's transform the histogram shown on Figure 2.15 in the following manner (Figure 2.19). Let us map the section of the curve which lies to the right of the extremum point as a mirror image upwards, and then carry out an affine transformation of the resulting curve in such a way that the point of the former maximum, would have the coordinates (0.5;0.5), and the entire curve would be inscribed into a unit square so that b=1, a=0, f(b)=1, and f(a)=0. Now mark off on the ordinate axis segments corresponding to the scale of frequency obtained experimentally with a certain value of cutoff α^*. The conversion to the universal scale occurs as follows. Let a certain phenomenon or event have a frequency of encounter that lies inside segment $[\xi_1, \xi_2]$. Let us project this segment using the histogram f,(x) or the ordinate axis. Corresponding to it will be a segment evaluated by the characteristic *seldom*.

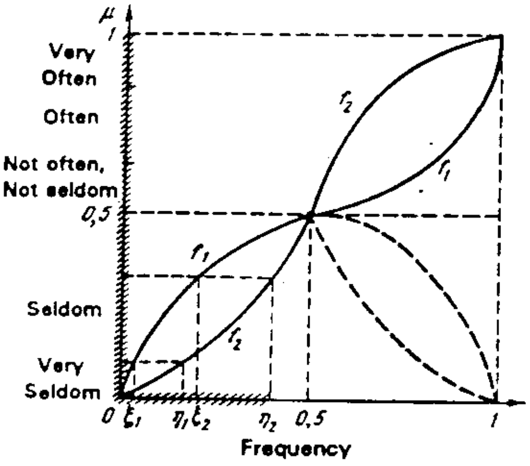


Figure 2.19

This means that the occurrence of an event defined by the histogram $f_1(x)$ with such a frequency is evaluated as a rare event.

If we have another event or phenemenon with its own histogram $f_2(x)$, then the true frequencies of its occurrence, evaluated as seldom, will be different. On Figure 2.19 the segment $[\eta_1, \eta_2]$ which does not coincide with segment $[\xi_1, \xi_2]$ corresponds to them.

Similar to the case of an inexact scale for frequencies it is possible to construct universal scales for other inexact scales. For example for the space scale along the abscissa axis, first are laid off distances measured in some natural units of distance measurement (for instance in meters), while along the ordinate axis are arranged segments that correspond to the words *very far, far* and so forth. Next occurs the transfer of the entire graph into a unit square with the condition that the maximum point of the histogram falls into the point (0.5;0.5). After that the transition from the inexact spatial scale to the universal scale occurs just as it was done for the frequency scale.

In general, if there is some scale that reflects a certain order in oral evaluations of some phenomenon or fact, and there exists a method of their quantitative evaluation, then, with availability of a histogram it is always possible to switch over to a universal scale not dependent on the concrete semantics of the given phenomenon or fact.

Let us now describe a different approach to the process of introducing a universal scale.

If the object of control is complex, and, in describing current situations on it and also for describing the knowledge or its functioning and its control, it is necessary to use qualitative descriptions, then it is clear that the model formed in the control system is only an approximate description of the object and of the process of controlling it. But not every approximate model is suited for it, but only one, for which the *postulate* of *consistency holds*.

The meaning of our statement can be illustrated with the aid of Figure 2.20. On it, W is the set of states w (taken in its widest sense) of the object of control. It can be treated as the set of all possible exact

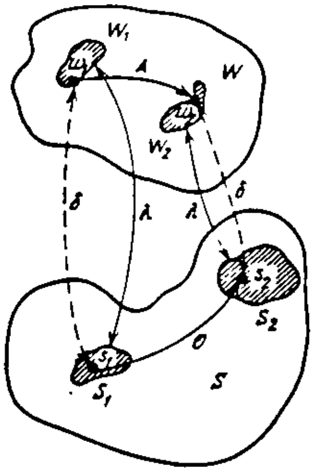


Figure 2.20

descriptions of its states. The set S is the set of descriptions of complete situations which are utilized by the control system for making decisions. Individual complete situations S_i do not exactly correspond to the true descriptions of that which characterizes the current state of the object with consideration of its specific w_i. Hence the mapping α is inexact. Instead of an exact single-valued on both sides mapping $\delta:w_i \to s_i$ we have towards both sides a multivalent mapping $\lambda:W_i \Rightarrow S_i$, where W_i and S_i are subsets of W and S, respectively. Here the mapping $\delta:w_i \to s_i$ is sort of "included" into mapping $\lambda:W_i \Rightarrow S_i$, i.e. it is one of the mappings of this multivalent mapping λ. Let O be some operator for transforming one complete description of one situation into another, while A--is a certain process of transition from a state w_1 of the object of control into state w_2. And let O imitate the transition A in the control system. Then the *model* of *description will be consistent*, if $\delta:w_1 \to s_1$ is included in $\lambda:W_1 \Rightarrow S_1$, $\delta:w_2 \to s_2$ is included in $\lambda:W_2 \Rightarrow S_2$ and the subsets W_1 and W_2 are such, that with the given criterion of control of the object (or with considerations are its expediency used in its control) all states, that fall under W_1 and W_2, require equal (with a certain acceptable precision) solutions on control. In other words, from λ is required only an "approximate truthfulness" of transmission of what is happening on the object of control, but the degree of approximation should not exceed a certain limit outside of which the decisions on control, adopted in the model, will not correspond to the required ones. This mapping should be a mirror with minor distorations, and not that "wild mirror" about which V. Nabokov had written in his time: "In one word, we had such a wild mirror and a whole collection of all kinds of 'netki', i.e. absolutely ridiculous objects: all such shapeless, multicolored, lumpy things, like some sort of minerals, but this mirror, which completely distorted ordinary objects now gets some real nourishment, i.e., if you placed such an incomprehensible and monstrous object so that it became reflected in the incomprehensible and monstrous mirror, the result was remarkable; no by no gave yes, everything became restored, eveything was fine..."

The mapping λ can be specified in the form of some matrix, in which states w_i; correspond to lines and complete situations s_i to columns. At the intersections of the lines and the columns it is possible to specify a number $0 \le m_{ij} \le 1$, which will characterize the possibility of replacing w_i by s_j or the reverse substitution. Then the lines of the matrix can be considered as the functions of affiliation of w_i or the set of complete situations. And the columns of the matrix -- as the function of affiliation of s_i or the set of time states of the object of control.

Example 2.17. Consider the Table-Matrix 2.7.

Table 2.7

Distance Meters	Very Near	Near	Not far, Not near	Far	Very Far
0,5	0,9	0,1			
2	0,05	0,65	0,2		
20		0,15	0,6	0,05	
100		0,05	0,4	0,8	
3000			0,05	0,7	0,2

This table reflects the opinion of a certain expert and was obtained in the process of a psychological experiment. In it the test subjects were able to write in their evaluations from segment [0,1] with a spacing of 0.05 in the cells of the matrix. The matrix assigns five functions of affiliation of the values of an exact metric scale of distances to word estimates of distances and five functions of affiliation of word estimates on an inexact scale of distances to exact metric values. On Figure 2.21 is shown, for illustration, the affiliation function for *not far, not near*. Its graph consists of five points defined by the third column of the matrix. The piecewise-linear approximation was done for better visualization.

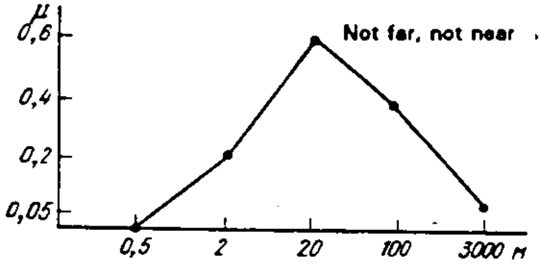


Figure 2.21

The mapping of λ, given by the matrix, may be called the *inexact relation (ratio?) of modeling*. It shows how one exact or inexact scale is modeled by some other scale. And, in particular, how does the transition occur from word evaluations of some qualitative characteristics to their numerical evaluations, which is the goal of the present paragraph.

A *diagram of modeling* can be associated with an inexact modeling relation.

$$\begin{tikzcd} A \arrow[r, Rightarrow, "T"'] \arrow[d, squiggly, "\lambda"'] & A' \\ B \arrow[r, Rightarrow, "Q"'] & B' \arrow[u, squiggly, "\delta"'] \end{tikzcd}$$

Here A and A' are states of some medium, and B and B' are associated states of the model that represents this medium. Transformation operators Q in the model must correspond in a coordinated manner to transformation operators T in the medium. For successful modeling it is necessary that the diagram of modeling be commutative.

Example 2.18. In the preceding example the test subjects constructed the matrix as if outside consideration of semantics. It is of interest to reveal its effect on the continuation of an inexact relation of modeling. In one of the experiments carried out, the test subjects had to write in evaluations into matrices similar to the ones used in the previous example, if they found themselves in the following three situations: going around a pit with slippery edges, going around an operating sprinkling automation vehicle, and crossing a road in front of rapid moving transport vehicles. The result of this experiment turned out to be practically the same matrix, which is recorded in example 2.17, but the evaluations during transition from one situation to another were as if shifted over to the right by one category during the transition from the more dangerous (third) situation, to which corresponded the matrix from Example 2.17, to the less dangerous second and first situations of the experiment. For instance, in the situation with the sprinkling automotive vehicle the distance of 0.5m with affiliation function of 0.85 received the evaluation *near*, while in the situation of walking around the pit with slippery edges with about the same degree of affiliation this distance was evaluated as "*not far, not near:*.

Such a linear shift, observed in the example above, allows us to put forth a very important hypothesis about how a change in the semantics of situations, in which an estimate of distances is made, leads to a homomorphic mapping (representation) of the affiliation function. In other words, if the scales, on which are specified our functions of affiliation or the inexact intervals, are transformed linearly with the aid of function $y=ax+b$, then the functions of affiliation do not change, but only the semantics of the situation to which they are related. The stated hypothesis must find corroboration in experiments with people for different, by interpretation, modeling relations (so far it has been verified only for relation of distance). If experiments will substantiate it, then with work with qualitative descriptions will be considerably simplified.

Let's enumerate at the end of the paragraph those problems in the solution of which the inexact relations of modeling may play an important role.

1. Calculation of affiliation functions of incompletely specified inexact linquistic scales by computing of incomplete or complete transitive closure of the modeling relation given by a matrix of the type considered in Example 2.17.
2. Carrying out a classification of inexact concepts for decreasing the dimensionability of the model space.
3. Possibility of approximation of new inexact concepts, given in the form of functions of distribution or as inexact intervals.
4. Construction of an hierarchic system of modeling relations which corresponds to the construction of an hierarchical system of scales for defining inexact concepts with different degree of generalization.
5. Construction of universal scales that take into account the concrete semantics of some or other situations.

CHAPTER 3

SUPPLEMENTATION OF SITUATION DESCRIPTIONS

Before there was anything,
there was nothing. I. Shtok

§3.1. Formulation of the Problem

When man perceives a communication in the form of a text in a natural language, there occurs a phenomenon so familiar to us, that we simply do not notice it. It consists in the fact that upon perceiving a communication we perceive a much greater volume of information than is contained directly in the communication. The enrichment occurs for three reasons. First, the communicator himself is received by us within the framework of a certain distinction which is what enriches it. For instance, if you are told, "turn on that switch there", you see what switch the speaker is pointing to, and carry out the instruction correctly. If this communication were not accompanied by the gesture, then the communication would not have been understood. In the question: "How, though, did this happen? Look!" the text remains totally incomprehensible outside the frame work of that real situation in which it was spoken. And so, to the information which was contained in the text, it is as a rule, necessary to still add information on the situation within the frame work of which the given text originated.

Second, the author can utilize the text for attaining some goals of his which are not explicitly expressed in the text. The well-known definition of the concept "a bore" as a person who when asked "How are you?" begins to tell in detail about his life during the past years, is based precisely on that. The object of the asker was not at all to obtain such information. He used the text of his question merely as a special marker-signal to express his consideration toward the person of the listener. To the same type of additional information is related the intention of the speaker, the mimic of his face, and much other.

Third, to the information, which is directly expressed in the text of the communication, there is immediately added information related to it and stored in the memory of the listener. Assume you heard the communication

"Before we landed, the airplane made five circles over the airfield. Then it descended sharply and landed. Only on land did I begin to feel how my feet were shaking". In this communication, between the facts expressed by the first two sentences, which follow each other directly in time, and the last sentence, there is a logical discontinuity. To the listener it does not have a great importance. Every man, familiar with flying in an aircraft, understands that after an aircraft had landed, there occurred another series of facts mandatory for the given situation: the aircraft taxied to the parking ramp, turned off its engines, the door was opened, the gangway was brought or the stair was lowered, passengers, and among them the speaker, descended to earth. And only after all of that took place did the "shaking of the feet" occur. And all this additional information helps the listener to correctly understand that which was communicated to him directly in the text.

If the conversation is about a control system, then the additional information generated during intercourse of people for the first two reasons, is inaccessible to it, if the system is not equipped, as a robot is, with special receptor channels: sight, hearing, etc. Therefore, if for the control system such information is indispensable, then it must be communicated to it together with that main communication that is being transmitted. As concerns enrichment of information by date stored in the system memory, this function can be imposed on the control system itself. Such a task we will call *supplementation of description.*

Within the framework of our book these will be considered only the problem of supplementing descriptions by means of special semistic systems. In other words, the basic problem which we shall discuss, is the construction of a special system that insures a supplementation of the description fed in at the input. In the most general form the system needed by us can be specified in the form of a system of *productions* of the type

$$\gamma;\ \alpha \Rightarrow \beta;\ \delta.$$

where γ denotes a certain consideration, the fulfillment of which allows us to utilize the production. Through α, is designated a fragment of that structure (description), which is subjected to transformation. The transformation itself consists in the exchange of fragment α for a new fragment β. Within the framework of solution of the problem of supplementation of descriptions, fragment β must be in a certain sense richer than fragment α. Finally, δ is a certain modifier of the condition γ. Upon application of the production this modifier changes the condition of applicability of the given rule or leaves it unchanged.

Various systems of supplementation of descriptions differ from each other by the manner in which the productions are organized and the appearance of the strategy of their application to the initial description and the intermediate descriptions obtained in the process of supplementation. The system of production may, in particular, form a certain logical system. Such a system must reflect the regularities inherent in the given problem area and the methods of constructing solutions on the basis of a description of situations in it. Here one may separate the productions into three types: *deductive*, *inductive*, and *traductive*. In productions of the first type the fact β is local, resulting from the carrying out of condition γ and simultaneous presence of fact α in the description being transformed. For inductive productions fact β is a more general fact than fact α fulfilling the condition γ. Finally, during a traductive production the facts β and α have an equal degree of generality.

Example 3.1. Let us consider productions in which the modifier δ is absent.

1. The parts storage bin of a machine tool is empty. If the bin is empty, the machine tool is idle.
2. The machine tool is idle. If the machine tool is idle, the bin is empty.
3. The bin is empty. If the machine tool is idle, it is necessary to take measures for filling the storage bin.

These three productions illustrate, respectively, the productions of the deduction, inductive, and traductive types.

Systems of productions possess a series of properties making them a most convenient means of describing a system of supplementation of descriptions and of its programmatic realization. Firstly, a system of productions possesses the quality of autonomy. Any one production can be taken out of it or added to it, while all the other productions remain unchanged. This property makes the system of productions flexible and easily adaptable to any changes in the problem area. Let us note that the adaptability of productions can also occur due to changes in γ and δ, which allows us, while keeping the production in the system, to vary its effect. Every production is a certain finished fragment of data on the problem areas, and in this sense it does not depend on other productions. But as a necessity, it is possible to set up connections between individual productions by a system of mutual references. Of course then, the autonomy which we are describing, will disappear.

Secondly, a system of productions which presents the procedures of transformation of descriptions is an asynchronous form. Productions furnish, in a natural way, the possibility of parallel processing. Each of them can be carried out independently of the others, which in time, may lead to a non-singluarness of the result of the transformation. But this is either not important (different variants of description supplementation are possible), or there are available, special procedures that permit only such parallelism, as does not lead to undesirable results.

Thirdly, systems of productions are easily mapped into systems of operators of programming languages. And some of the programming languages (for instance, the native language REFAL or such widespread language as PROLOG) use productions as their main operators.

Finally, models in the form of systems of productions encompass a wide class of different generating models, into which enter such familiar models, as formal grammars, propositional calculus and predicate calculus, network models and many others. In all presently known models of situational control, it was precisely the productional systems that were utilized for solving the problems of supplementation.

§ 3.2. Scenarios

Most often for supplementation, use is made of network models which have lately received the appellation of *scenarios*. A scenario is represented by a certain network having apexes which correspond to facts, and paths which describe relations of a special type. These relations possess the following property: if between apex x and y there exists a variety of paths π_1, π_2, ..., π_n, and present, are both facts a and b corresponding to apexes x and y, then there takes place, at least a population of facts corresponding to the apexes as one of the paths connecting x and y.

Cause-effect, *part-subpart*, *goal-subgoal* are examples of relations having the indicated properties.

Example 3.2. In Figure 3.1 a scenario is shown in which the role of the relation is played by *cause-effect*. The apexes in this scenario correspond to the following facts: 1- the machine tool is idle, 2- there is no worker at the work-place, 3- the machine tool is out of order, 4- there are no billets, 5- it is lunch time, 6- the worker has walked away from the machine tool during work-time, 7- an external supplier did not supply the factory with billets, 8- a machine tool which supplies the billets for the given machine tool, is out of order.

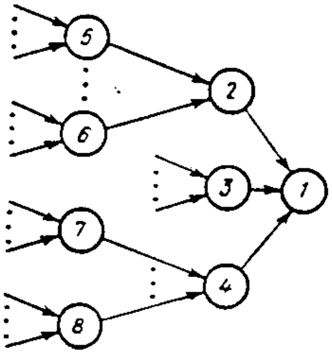


Figure 3.1

If the memory of the system contains such a scenario and the input of the system has received the text "lunch break", then the system can, with the aid of the scenario, supplement this description by adding to it the facts: "At the work-place there is no worker" and "The machine-tool is idle". We note that our scenario talks about a machine tool in general, i.e., is applicable to any machine tool.

Example 3.3. In Figure 3.2, another example of a scenario is shown in which the following facts correspond to the apexes: 1- the aircraft is approaching the airfield, 2- the aircraft asks permission to land, 3- the aircraft enters the waiting circle region, 4- the aircraft executes the flight in the waiting circle, 5- the aircraft receives permission to land, 6- the aircraft lands, 7- the aircraft taxies to the parking area, 8- a passenger

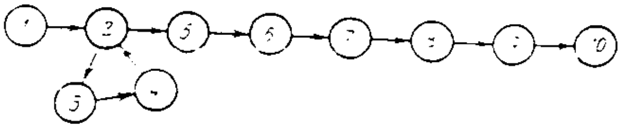


Figure 3.2

ramp is supplied to the aircraft, 9- passengers descend to earth, 10- passengers are at the baggage pick-up area.

In this scenario, as compared with the one we considered in Example 3.2, the areas characterize not the cause-effect relation, but an operational relation. The order of realization of states, characterized by the apexes, is the order of time. Strict orderliness is violated only by the alternative transition after apex 2. Depending on the real situation, these may either occur only once and then to the state characterized by apex 5, or else again will occur, one or more times, the states characterized by the sequences of apexes 2-3-4-2-3-4-2-...-2, after which will occur the transition to apex 5.

If the input system is the fact - the aircraft has landed, the passengers are waiting for the baggage, then the scenario will help to supplement this description by means of these facts which are present in it and which must definitely have occurred prior to the occurrence of facts corresponding to apexes 6 and 10. It is only about the facts corresponding to apexes 3 and 4, that the system cannot tell anything definite.

Example 3.4. Let us consider, finally, one more scenario, the structure of which is shown in Figure 3.3. The arrows on this scenario, map the relation *whole-part*, and the apexes are interpreted in the following manner: 1) electric vacuum cleaner, 2) upper part, 3) seal ring with filter, 4) lower part, 5) electric motor with fan assembly, 6) switch, 7) diffuser, 8) cord.

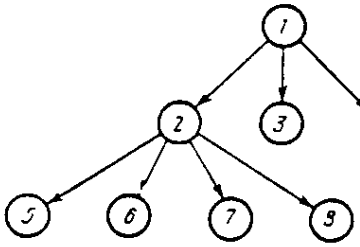


Figure 3.3

The scenario with the relation *whole-part* is used in supplementing descriptions differently, than the earlier considered scenarios. Let, for example, the input to the system be the description: "A vacuum cleaner was operating in the room". By means of the scenario we have, this description can be supplemented by an entire group of facts of the type "The switch of the vacuum cleaner was located in the room" (the same room is understood as was mentioned in the inputted description), "The cord of the vacuum cleaner existed at moment t of time" (the same time moment is understood that corresponded to the existence of the fact reflected in the initial description). In other words, in a number of cases it is permissible to carry over the statements about existing in space and time of a certain object, onto its component parts. Of course this is not always achievable. If, for instance, the statement "The train passes through a tunnel" is carried over to some car of the train, the transposition may turn out to be false. The given car at that moment of time may not have entered the tunnel yet.

Let now the system input be this description "The flow of air from the diffuser is blowing away the papers which are on the table". Using our scenario we would supplement the description, for example by this fact: "The vacuum cleaner is in the same room as the table with the papers at time moment t" (understood is the moment of time, to which corresponds the fact indicated in the initial description). As in the preceding case, the mechanical transportation of spatial and temporal statements from a part of the object to the entire object is of course not always possible.

Supplementation by a scenario based on the relation *whole-part* may refer not only to position in space and time. If, for instance, the input description has the form "The electric motor of the vacuum cleaner costs more than one hundred rubles", then a supplementing fact may be introduced into the description the fact: "The vacuum cleaner cost more than one hundred rubles". This new fact may be formulated on the basis of the scenario shown in Figure 3.3 and the supplementation rule on the summation of prices (values) during transition from a part to the whole (more precisely, not summation of value, but retention of the relation *to cost more*).

The additional rules we referred to in the preceding example, are rules of the production type. Their applicability in one or another case is related to the fulfillment of the condition of applicability of the production. Entering into the condition of applicability of production is also that necessary connection between object which is specified in the scenario. Thus, we arrived at the conclusion that there are possible two paths of supplementing descriptions: via the formal properties of relations used in the scenarios and by means of the semantics of the latter. Thus, in the first two examples, we had considered the relations "cause - effect" and "order relation". These relations are transitive. And this means, that without any regard of the semantics of facts a.) the presence in the initial description of fragments (a_1ra_2) and (a_2ra_3) allows us to supplement the description by the fragment (a_1ra_3). For example, from the statement that "A cylinder is a part of an engine" and that "An engine is a part of an automobile", it follows, that "A cylinder is a part of an automobile". In a similar way occurs the supplementation of a description if it employs a symmetry relation. If, for instance, the systems entry receives the description "The drilling machine is switched on simultaneously with the drilling machine". In other words, in the presence of fragment (a_1ra_2) with the symmetry relation r it is possible to automatically supplement a description by means of the fragment (a_2ra_1).

But supplements of this type, using only algebraic properties of relations entering into them, are not all that interesting. More important are rules that consider the semantics of the relations themselves and the semantics of the situation in which the supplement is being formed.

Example 3.5. Take from Table 2.2 the relations r_{71} - *to have as an object of action* and r_{74} - *action - place*. Let a be an aircraft, i_1 - board 1244, d - flight, b - city, i_2 - Odessa. The relation r_{40} taken by us from Table 2.1, means *to be a bore*. Let system input be the following information: "Aircraft board 1244 is flying over Odessa". The notation of this description has the form

$$((api_1)r_{71}d)((api_1)r_{40}(bpi_2))$$

Between the large parenthesis is omitted the relation ***simultaneously***. The semantics of r_{71} and r_{40} are such that the initial description can be supplemented by a fact of the following form: ($dr_{40}[bpi_2]$), which corresponds to a statement that the flight is taking place over Odessa.

The cited examples show that the scenarios by themselves do not yet solve the problem of supplementatation of descriptions. Indeed, to solve the problem of supplementation it is necessary to create productional systems, in which rules take into account the summations of the relations which enter into the description being supplemented. In the following paragraph (beginning with §3.4) we shall show how such systems are constructed for individual groups of relations. In the next paragraph, however, we will discuss the relation between scenarios and concepts with which we occupied ourselves in Chapter 2.

§ 3.3 Scenarios and Frames **

The purpose of this paragraph is to show that scenarios are easily mapped into those frame structures, which we have described in § 2.9. For better visualization we shall here utilize a graphic representation of a frame in the form shown in Figure 3.4. Each rectangle corresponds to a frame with a definite name. A frame consists of slots the names of which are shown in the smaller size rectangles. Slots can be terminal or nonterminal. In the latter case their names serve as names of the frame comprising the contents of the given slot. Terminal slots define a list of certain values corresponding to them. On Figure 3.5 is shown an example of such a structure. The names of the first slot of the frame KOMANDROVKA is the name of the frame KTO (who).

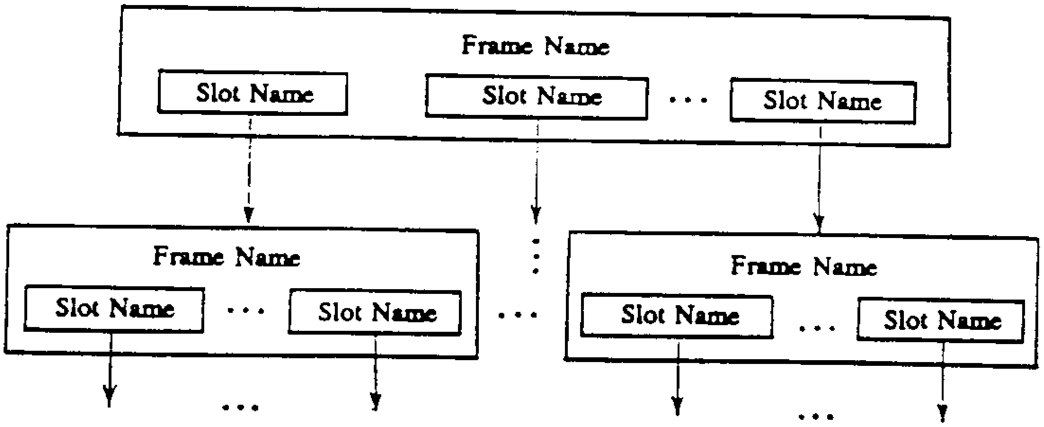


Figure 3.4

The slots of this frame are terminal. To each of them corresponds a list C(s) of elements forming the given slot. Slots with names WHERE TO and WHEN of the frame KOMANDIROVKA also are terminal. Such a structure may be realized in traditional languages of representation of knowledge (for example, with the

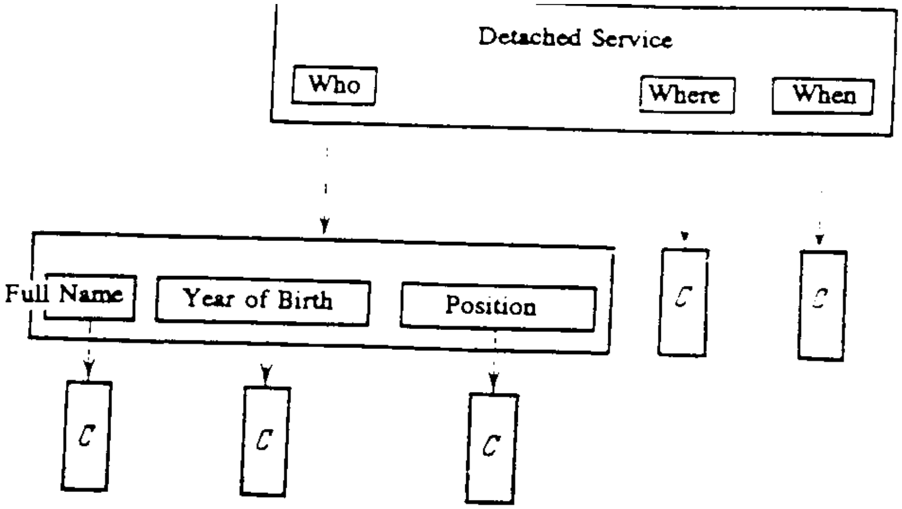


Figure 3.5

use of FRL, the basic characteristics of which were described by us in §2.9). However, for the sake of simplicity of the considered representation, it might be attempted to create a simpler and more effective language, full of defects inherent in languages of the FRL or KRL type.

In Example 3.2 - 3.4 we have looked at scenarios, which based themselves, respectively on cause - effect relations, operational relations and relations of the type part - whole. We will show, that scenarios of this type are easily mapped into structures of the forms considered by us.

Example 3.6 Let us return to the scenario described in example 3.2. Let us look at the network of frames depicted on Figure 3.6 (Terminal slots are line-shaded). It is easy to see that this network is equivalent to the scenario shown in Figure 3.1.

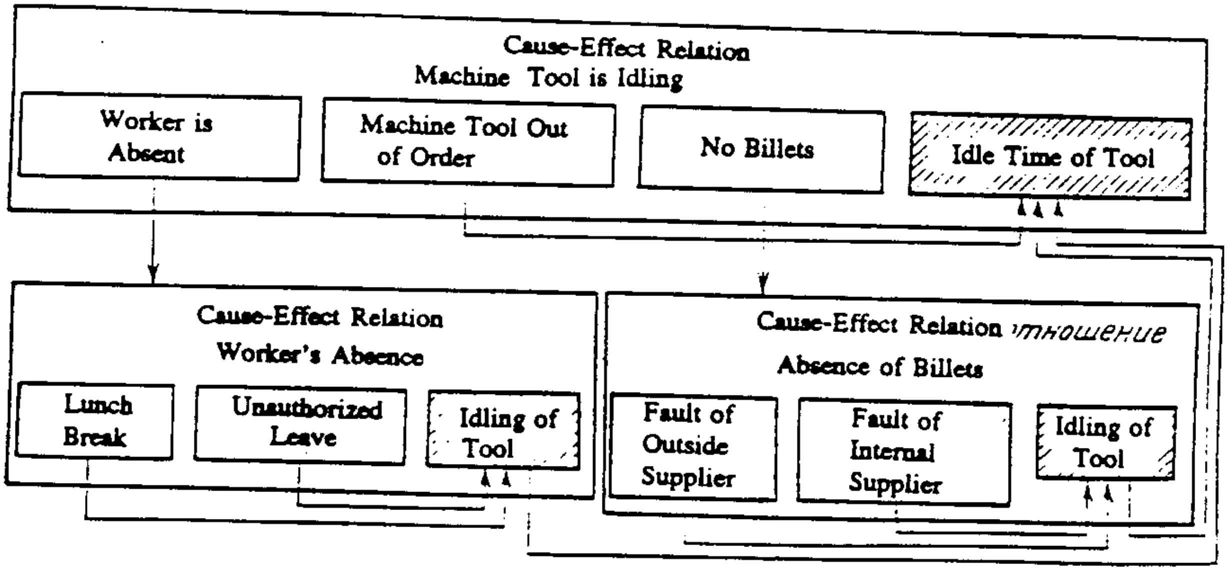


Figure 3.6

In Figure 3.7 a frame of the cause - effect relation is depicted in the general form. On it, in addition to slots CAUSE and EFFECT, are also the slots MECHANISM, ACTING FACTORS, and REGULARITY. Frames corresponding to the slot MECHANISM, describe cause - effect sequences (on Figure 3.7 these slots are marked by letters M with various indices), the realization of which leads to some or other effects, shown in the initial frame labeled X. Slots CAUSE (on the figure they are marked with letters П with various indices) show the direct cause causing one or several effects, described in the initial frame.

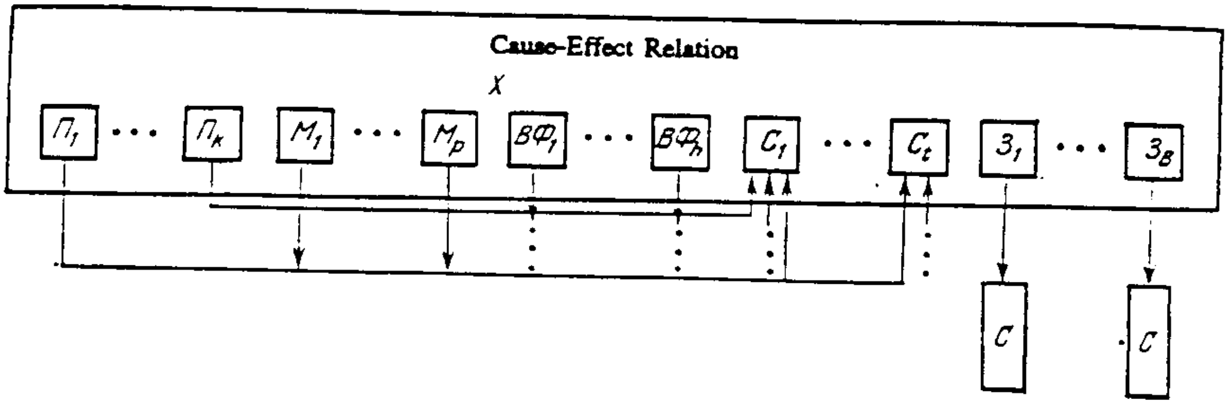


Figure 3.7

The effects (letters C(s) with various indices) - are the terminal slots. To them correspond certain final texts. The slots ACTING FACTORS (letters $B\phi$ with various indices) refer us to the frames in which are described factors serving as the cause of occurrence of the phenomenon described in the frame labelled X. Finally, slots corresponding to the REGULARITIES (letters Z with various indices), are always terminal. Their content reflects the character of the regularity existing between the causes and effects, described in the frame marked X.

Example 3.7. In Figure 3.8 is shown a mapping into a system of frames of the cause - effect scenario, to explain the fact of a sudden loss of blood. The meaning of the slot EFFECT in the initial frame SUDDEN BLOODLOSS is the name of this frame.

Let us now consider the representation of scenarios that are based on operational relations, in the form of a network of frames.

Example 3.8. In Example 3.3 we looked at the scenario shown in Figure 3.2 and related to the operation of landing of an aircraft at an

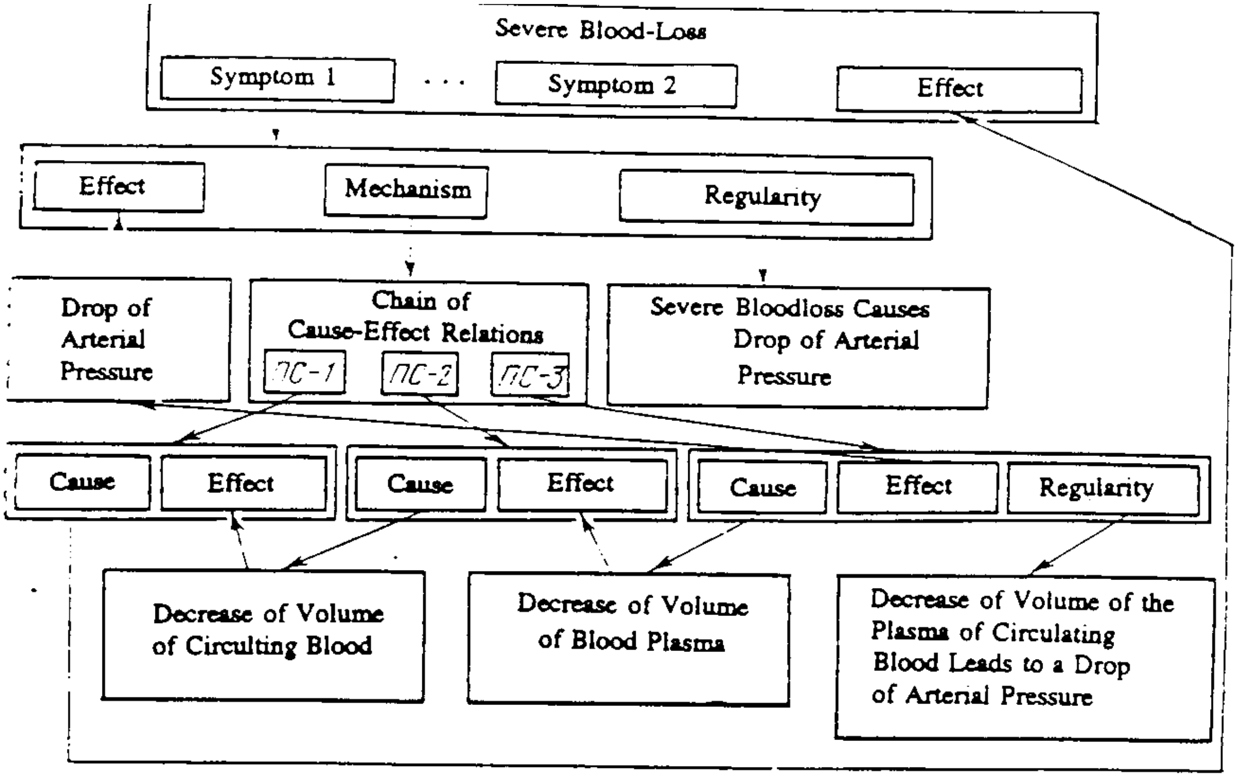


Figure 3.8

airport. On Figure 3.9 is shown a network of frames that corresponds, as is easy to see, to the scenario on Figure 3.2. Movement along the network of frames occurs by virtue of the slots CONDITION. They allow to carry out some or other operations, related to a given frame. The results of these operations affect the significance of the CONDITION slots in other frames, that are related to the given frames operationally. In the lower part of the drawing is shown the resulting frame, the realization of which corresponds to a complete fulfillment of the scenario.

We shall not proliferate the number of examples. We will only explain on the last one of these, the supplementation of the description of a situation. Let, as before, the system receive the text: "The aircraft has completed landing, the passengers are waiting for luggage". These two facts within the fact, are localized in the frames LANDING and TRANSITION TO THE LUGGAGE DEPARTMENT. This permits us, by using the network of frames, shown in Figure 3.9, and moving from the available results to the conditions of the subsequent frames, to find out that there existed conditions for operations connected with the debarkation of the passengers, and also for the regulation

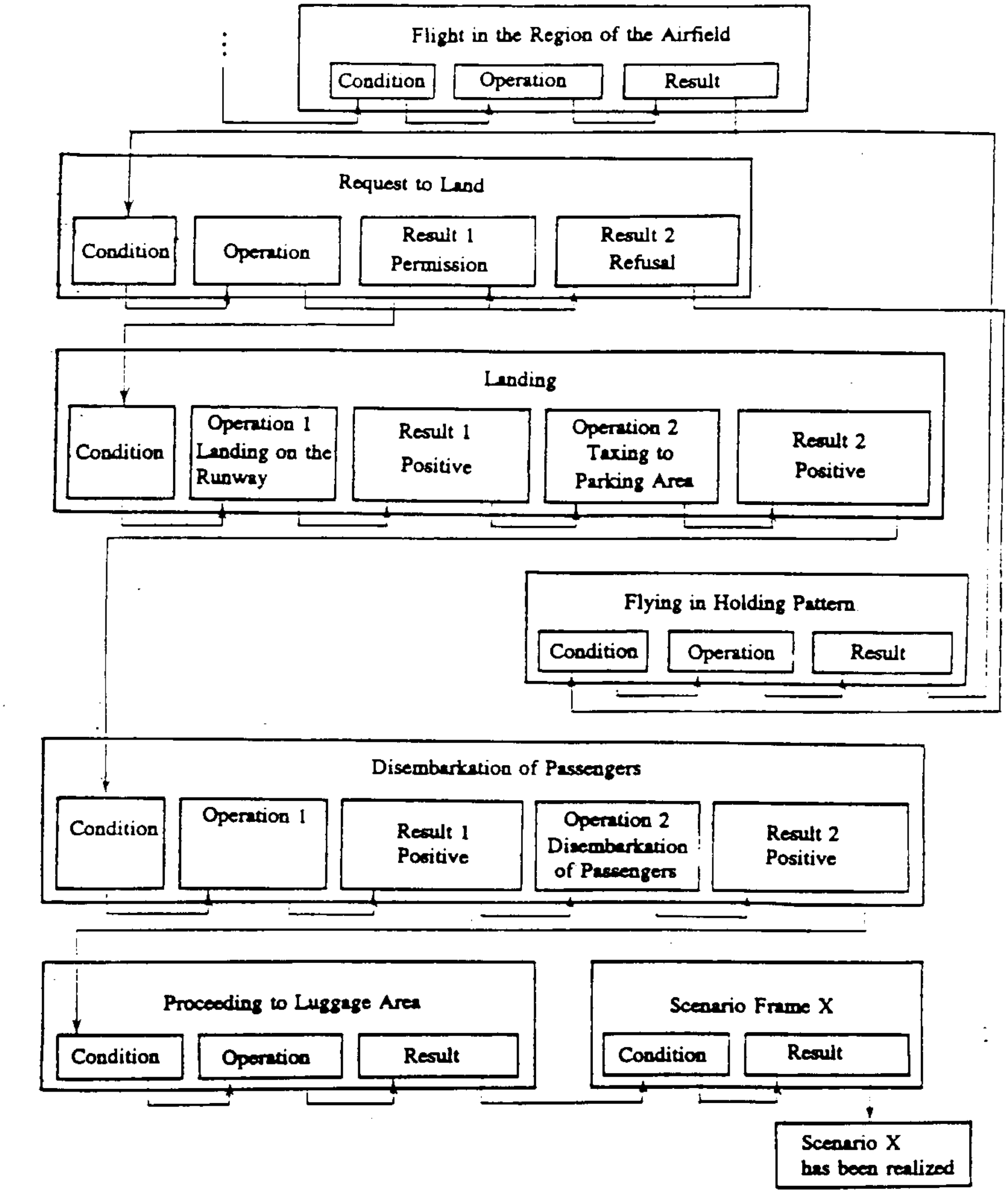

Flight in the Region of the Airfield
Condition
Operation
Result
Request to Land
Condition
Operation
Result 1 Permission
Result 2 Refusal
Landing
Condition
Operation 1 Landing on the Runway
Result 1 Positive
Operation 2 Taxing to Parking Area
Result 2 Positive
Flying in Holding Pattern
Condition
Operation
Result
Disembarkation of Passengers
Condition
Operation 1
Result 1 Positive
Operation 2 Disembarkation of Passengers
Result 2 Positive
Proceeding to Luggage Area
Condition
Operation
Result
Scenario Frame X
Condition
Result
Scenario X has been realized

Figure 3.9

of the entire scenario. If we now move along the network of frames from the end to the beginning, from conditions to results, which defined the given conditions, then we shall obtain the entire combination of facts we need.

And so, networks of frames are not only convenient for representation of knowledge, as discussed in detail in § 2.9, but also render an important service in solving the problem of supplementation of descriptions.

§ 3.4. Pseudophysical Logics*

We have already mentioned, that it would be desirable, for certain classes of relation, to construct deductive productional systems, which would allow us to carry out the required supplementation of descriptions without storing in the memory of the system the totality of a large number of scenarios. One of the classes of such deductive systems are the *pseudophysical logics*.

The application, "pseudophysical logics", comes from the fact that in their rules of derivation one used the properties of perception by man of the surrounding world, which possesses a number of subjective characteristics. For that reason they describe not the objective physical world, but its subjective perception by man.

In distinction from formal systems about which we spoke in § 1.5, the pseudophysical logics have a number of important characteristics. Let us note the principal ones of these.

1. Pseudophysical logics are logics of relationships. It is the latter that play the role of variables. Hence, pseudophysical logics are classified depending on the type of relationships utilized. The logic of time, studies the mutual relationship of temporal relationship, the logic of space - the spatial ones, the logic of actions - the relationships of the type *subject-action* or *action-space*, the causal logic - the mutual relations of the type *cause-effect*, the frequency logic - the relations of the type *repeatability-frequency*, and so forth. The objects, however, that are connected by relationships, appear in these logics only as the invariant part of the descriptions. Let us illustrate this by the following example. Let there be, at our disposal, two facts $(a^2_{22}b)$ and $(c^2_{21}b)$. From these, with any

interpretation of objects a, b, and c, it is possible to deduce the fact $(a^2_{22}c)$.

2. Pseudophysical logics and logics on *scales*. Scales may be of two types: *metric* and *topological*. Metric scales, in turn, are divided into *absolute* and *relative* ones. Topological scales are specified among facts projected onto them, relations of a nonstrict order or washed out relationships discussed in §2.11. Supposing, for example, we are discussing the spectral disposition of three objects a, b, and c along a straight line. If, at our disposal we have an absolute metric scale, then there is specified on it a certain scale and a fixed point of origin. Assume, for simplicity, that the measure of the scale is selected so that the objects we are discussing can be thought of as points located on the division of the scale. A situation of this type is shown in Figure 3.10a. With an absolute metric scale, all spatial relationships on the straight line are easily calculated

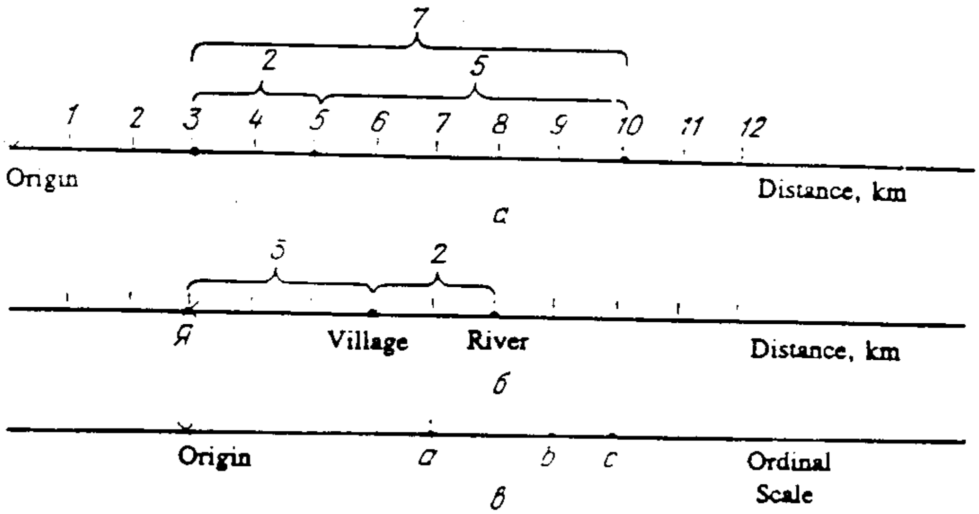


Figure 3.10

and pose no problems. One absolute scale differs from another only by the position of the origin in that the scale, which results in simple relationships for passing from one scale to another. On Figure 3.10b is shown a relative scale, and the point of origin is fixed only in the descriptions themselves, for example: "Two kilometers past the village is a river, but I still have three kilometers to get to the village". A relative metric scale can be converted easily to an absolute scale. Formulas for such conversions are trivial.

More interesting are the topological scales. They are closely connected to vague verbal estimates actively used by man for describing situations (including technological ones, for example). These verbal estimates determine only a certain order of distribution of facts on the scales. In Figure 3.10b an example is shown of mapping onto a topological temporal scale of three facts: (a) soon will come the moment when the steel will be ready, (b) after casting, the brigade will end its shift, (c) a short time after finishing work the specialist will leave the premises of the company. Topological scales are connected with the problem of formulation of qualitative descriptions and the perceptions of such descriptions. About that we have spoken in §2.11. The difference in scales also defines the difference in logics, which can be metric and topological.

3. Not only the facts are arranged on scales, but the deduction itself (the construction of productional rules) must take account of the ordering inherent in the discussions in the framework of pseudophysical logics. To illustrate this statement let us consider the following scheme of reasonings. Let a rotation α of a technology result in a decrease β of the quality of production. Let a technology violation γ be more serious than α (it is located farther from the origin on the scale of seriousness of technology violation than α is). Then, if α corresponds to β, the γ must correspond to such a determination of production, which on the scale of estimation of quality determination must be located farther (in any case not closer), than β. Arguments of this type are most important in checking rules of deviation for pseudophysical logics.

4. Pseudophysical logics contain, as their axioms, certain statements following from the perception of the world by man. These statements must be substantiated by results of corresponding psychological experiments. Axioms connect among each other relations of different nature, which allows man to carry out the substitution of some relations by others. One example is the substitution of spatial by temporal relationships: "Is the store far? - Fifteen to twenty minutes walk". Here the substitution was made

possible thanks to the axiomatic correspondence that relates the path, the time, and the average velocity of walking.

5. A system of pseudophysical logics is characterized by relations existing between individual logics. Some of the types of these relations are illustrated above. The axiom s = vt, where s - path, t - time, v - velocity, may be considered as relations of the temporal, spatial pseudophysical logics and the pseudophysical logic of action. An example of another kind of relation is that between the causal (cause-effect) logic and the logic of actions.

In constructing pseudophysical logics there should be considered three type of problems for the solution of which they are intended:

a) supplementation of descriptions of situations entering the systems memory with that knowledge on the object of control that is directly stored in the system, the prehistory of control and the rules of control of the object;

b) verification of reliability of inputted description of the situation, spotting contradictions in that description and of its compatibility with the information already stored in the system;

c) participation in forming decisions on control and verification of possibilities of realization of the selected controlling actions.

However, here we will discuss the pseudophysical logics only for problems of the first two types. In subsequent chapters, devoted to problems of generalization and classification of situations and to elaboration of decisions on control, we shall again return to pseudophysical logics.

Usually, one or another pseudophysical logic encompasses the circle of concepts of a man on some or other phenomena of the ambient world. Thus are generated the temporal logics that reflects regularities inherent in man in perceiving time and discussing it, the spatial logic, the causal (cause-effect) logic, the logic of actions, logic of goals, logic of estimates, and so forth. For problems of control of complex objects these six logics are of most interest. These logics are not independent. On Figure 3.11 is shown the mutual relationship among them. Unfortunately, to date in situational control only separate fragments of pseudophysical logics had been utilized, which is explained by the lack of their development. In recent years great success has

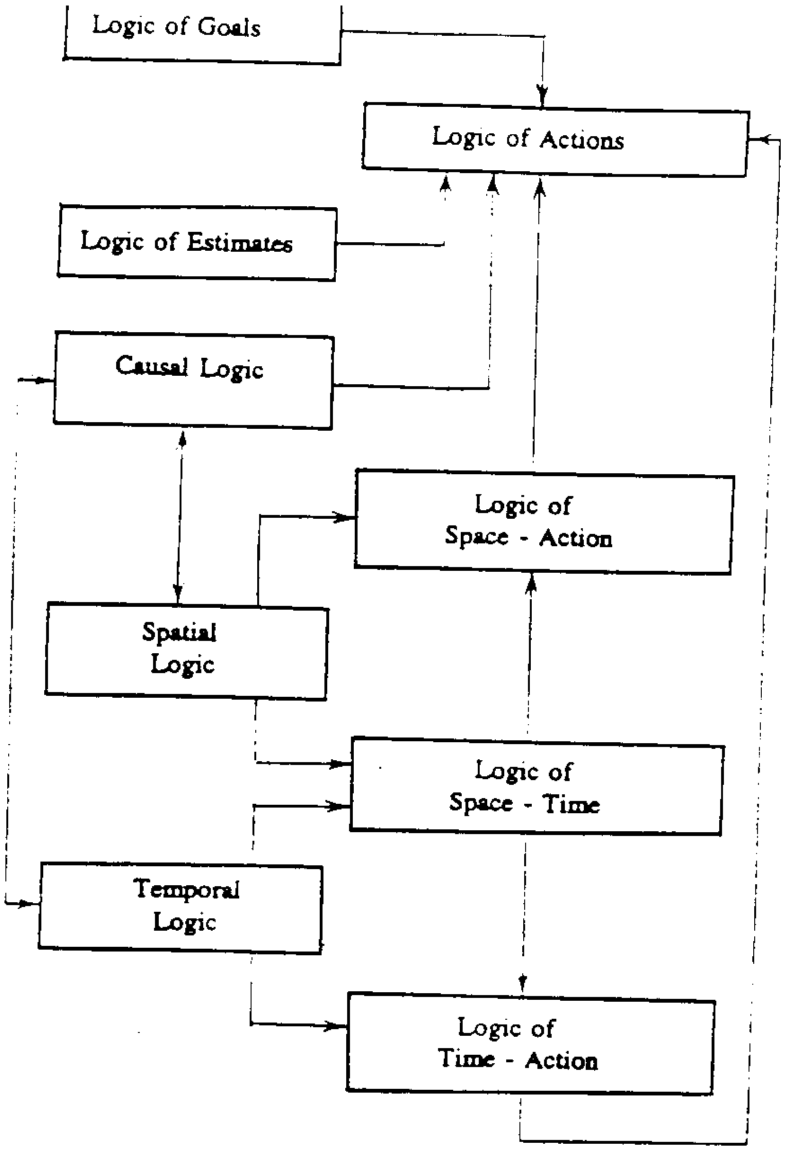


Figure 3.11

been achieved in this area. At the present time it can be assumed that the creation of temporal logic and of static spatial logic has been completed, and that fragments have been created for the logic time-action and for the logic of goals. But for causal logic, the logic space-action, and as consequences, for the complete logic of actions, the situation still remains not too good. As for the logic of estimates, it is presently undergoing, in its development, a stage of philosophical comprehension of those concepts which must be utilized in it in constructing this logic, which is aimed at solving problems of control. More detailed information on the paths of development of pseudophysical logics can be found in studies discussed under the comments to this chapter. In the following three paragraphs, as examples of constructing such logics and their utilization for purposes of supplementing descriptions, we will only touch upon the temporal, the spatial and the causal logics.

Every pseudophysical logic may be considered as a system having a structure shown in Figure 3.12. It shows that from the initial description, a certain structure of facts or events is separated, which is characteristic for the given pseudophysical logic (temporal structure, spatial structures, etc.). In it are identified certain units of the type of phenomena, events, processes, and facts, and are defined relationships among them from the considered group of relations. This part, essentially, does not pertain to pseudophysical logic proper, and its functioning is based on procedures described in §2.10. In the model of concepts, are reflected those basic regularities of perception that are characteristic for the control system (or of the man, if an imitation of his perception is taking place). For concrete pseudophysical logics this model becomes the temporal model, the model of

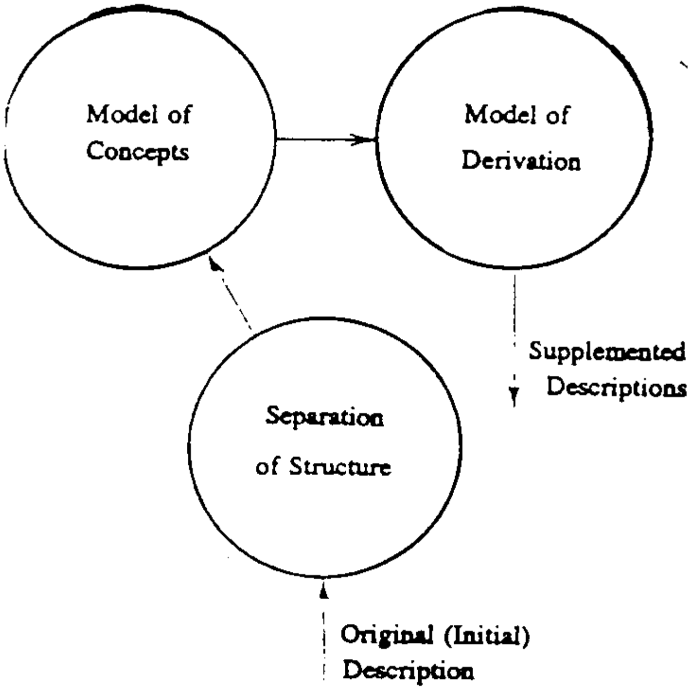


Figure 3.12

space, and so forth. Finally, the model of deduction contains rules, with the aid of which takes place the supplementation of situation description. The functioning of the model of representation and the model of deduction, and also the procedure of their construction will become clear from the material of the next three paragraphs.

§3.5 Temporal Logic**

We shall begin by considering the simplest pseudophysical logic - the temporal one. As it follows from Figure 3.12, we need to construct a model of representation of time and a model of deduction of new facts about time, which could follow from already available facts. Let me start with the first component.

We shall first formulate some general properties of time, which are convenient to adapt in constructing temporal logic aimed at its utilization in control systems, and here we will not consider the possible philosophical objections related to time as a philosophical category. In the comments to this chapter this is addressed more fully.

Let us consider a set T, whose elements we will call *moments of time*. Let us specify on T a complete strict order. The relation of a strict order on T specifies an ordinary scale for the moments of time t_i. Let us consider another set E, the elements of which e_j we will call *events*. Consider the relation $(e_j\ \tau\ ti)$, the meaning of which reduces to ascribing to an event e_j of a certain moment of time t_i. If for e_j there exists only one moment of time t_i, in which it occurs, then the event e_j in the given scale is call a *point* event. Besides point events, there may also exist *interval* events and *chain* events: the interval events have the property, that with the aid of the relation of projection onto the scale, they are juxtaposed to moments of time t_ℓ, $t_{\ell+1}$, ..., $t_{\ell+k}$, that form a continuous sequence* on that scale (i.e. there takes place: $t_\ell < t_{\ell+1} < \ldots < t_{\ell+k}$). For chain events the mapping, done with the aid of τ, is arbitrary. It is clear, that chain events breakup into a combination of interval ones and point ones, and that the interval events can be substituted by a combination of point events.

Let us now introduce the special operation of *discretization* D_n. If this operation is applied to set T with its specified relation of strict order, it is possible to import exactly n new elements t_i' and

---

*This means that between $t_{\ell+s}$ and $t_{\ell+s+1}$ for any s from 0 to 1+k+1 it is not possible to insert any element from T without violating the order.

$t_i < t_i' < t_2' < \ldots < t_n' < t_{i+1}$ between any pair of neighboring elements from T, such that $t_i < t_{i+1}$.

Let us introduce the operation of *rediscretization* R_n. It is constructed as follows. If R_n is applied to set T with its specified relation of strict order, then for any combination of n+2 elements of T that from a continuous sequence $t_i < t_{i+1} < \ldots < t_{i+n+1}$, there are excluded n central elements which by this fact are also excluded from set T. Here R_n is utilized sequentially, starting with the left-most element of T that forms the scale. After throwing out the first n elements, a new sequence of (n+2) elements is considered, of which the left element is the right element saved from the previous step of the previous sequence. We will explain the meaning of the introduced operations of discretization and rediscretization in the next example.

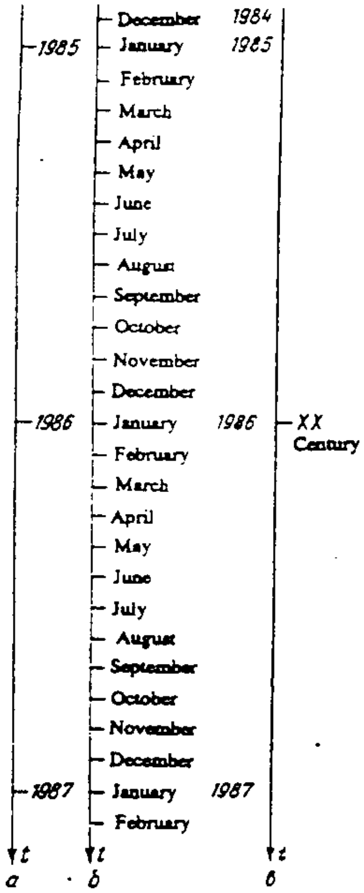


Figure 3.13

Example 3.9. On Figure 3.13a is shown as initial ordering scale. On it, the time moments are the years. On Figure 3.13b is shown the result of applying to this scale of operation, D_{11}. The new scale is arranged so, that between two old elements of T there have appeared eleven new elements for each pair. In distinguishing the elements, new names have been introduced. On Figure 3.13b is shown the result of applying the operation, R_{99}, to the original scale. With it, each continuous sequence of 101 elements has been transformed into a sequence of two elements to which new names have been given (XX century, XXI century).

Thus, by means of operations D_n and R_n there is generated an infinite set of scales for set T. Instead of the operations of discretization and rediscretization introduced by us, it is possible, in forming new scales, to utilize any monotonic function, because all such functions have the property of maintaining a strict order presented or the elements of set T. All scales of this type we shall henceforth call *absolute*.

If for some scale with number i (the numbering of scales may be given in an arbitrary manner, for example by means of a sequence of natural numbers) there was made a mapping of t_i elements of set E onto that scale, then in passing to a new scale t_i will change. If ej was a point event, then in operations R_n the point events are retained and are mapped into that moment of time which is retained in the new scale as the left moment in the treated sequence. If ej was an interval event, then after conversion it either shrinks, contracting toward the left end of the old interval, or (in the limit) becomes a point event. A chain event under operations of rediscretization becomes either a chain event with smaller intervals, or becomes an interval event, or else contracts into a point event.

Example 3.10. On Figure 3.14a is shown a point event e_j, corresponding to the statement "Konstantine Tsiolkovskiy was born on October 17, 1857". On Figure 3.14b is shown how this event got transformed into a mapping or a scale in which a transition from days to months was made. Now the point event e_j' corresponds to the statement: "Konstantine Tsiolkovskiy was born in September 1857". The interval event e_2, shown on Figure 3.14a, corresponds in the given case to the statement: "All of September 1857 a heavy rain poured without cease". On Figure 3.14d this same event is depicted as a point event (e_2').

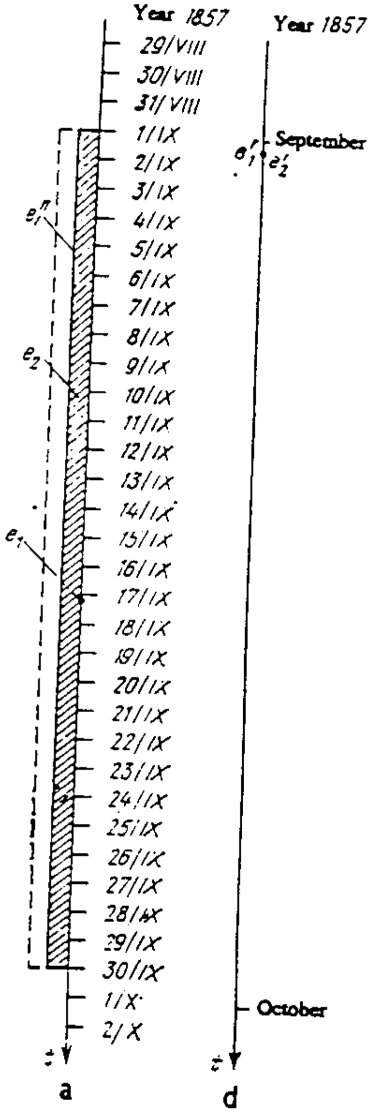


Figure 3.14

If an operation of discretization is performed on scales, then the mappings of events on scales may undergo the following changes: point events may turn into interval ones, or chain ones, interval events may be stretched out or become chain events, while chain events may acquire a longer length of their individual portions or increase the number of separate interval portions. Here are possible various combinations of the shown types of changes.

Example 3.11. If Figure 3.14d is used and a transition is made to the type of scale shown in Figure 3.14a, then as a result for e_1', we will obtain the interval event e_1'' shown by dots on Figure 3.14a. Thus, a comparison of the mapping of this event with the one that was on Figure 3.14a, shows, that a loss of information may occur, and that the operations D_n and R_n cannot be considered with respect to the mapped events as being mutually reversible.

The name which we ascribe to each moment of time t_i on one or another scale, reflects our conception about the order of sequence of these time moments. The starting point of initial count is usually selected as some event, for which there exists a mutual agreement. In human society there have been many diversified initial points of counting (origins). In pre-Peter the Great times, the year count was done from the beginning of the world, since Peter I we count years from the birth of Christ, inhabitants of Moslem countries conduct the count from the entrance of the Prophet into Mecca, and so forth. It is important that this counting origin for moments of time is fixed once and for all, and has an absolute meaning. This is why these types of scales are called *absolute scales.*

In addition to absolute scales, relative scales are also possible. On them, the origin is taken as that moment of time when a certain count takes place, specified in a statement, or when the statement itself is actualized. In the latter case we usually speak about a "point of speaking", which is then adopted as the origin of counting for moments of time. The following two examples illustrate the statements, in which the respective events are mapped onto relative scales: "Two years before the start of the Great Fatherland War, Germany seized Czechoslovakia" and "In five minutes dinner will be ready". In the first of these, the origin of count is taken as the starting moment of the Great Fatherland War, and in the second - that moment when the statement itself is actualized (spoken).

Absolute scales and relative scales are *metric* (as units of measurement of the distances between t_i and t_j, it is possible to take the number of time moments increased by unity, which forms a continuous sequence from t_i to t_j) and may have the same (by name) unit of measurement. However it is not always possible to establish a direct connection between them. Thus, for the first of the two statements above and knowing the fact that the Great Fatherland War began in 1941, it is possible to relate the relative scale with the absolute one, in which years are used as names of time moments. If the origin of the above relative scale is made to coincide with the moment of time named 1941 on the absolute scale, then the event "Germany seized Czechoslovakia" will receive an exact location named 1939 on the absolute scale. But for the second statement such a mapping is possible only in the case, if at the moment of its pronouncement one takes a look at a clock which indicates exact time and a calendar which marks the current date, month, and year. Without such information a mapping onto the absolute scale will turn out to be impossible. Thus, in any instance a transition from relative scales to absolute ones requires the involvement of special additional knowledge which must be stored in the memory of the system or must be observed by it.

We also note, that toward relative scales it is possible to apply the same operations D_n and R_n, which we have considered for absolute scales. The transformation of mappings of events on relative scales a result of applying D_n and R_n takes place according to the same rules as the earlier described transformations for absolute scales.

Let us note a circumstance of importance to us, namely, that for metric scales any interval event can be completely defined by two point events corresponding to time moments when the interval event begins and when it ends. These point events are called ***markers*** of the beginning and the end and are designated as μ_H^j and μ_K^j, where j is the index of the interval event ej. Chain events may be substituted by a sequence of interval events with different indexes, and each such interval event may be specified by a pair of its markers.

Besides metric scales there also exist, as was stated in §3.4, ***topological*** scales. In such scales moments of time cannot be juxtaposed with any names and it is not possible to introduce distances between these moments. Thus, on a topological scale it is possible to reflect only an order of certain moments of time. In mapping events onto topological scales, only their mutual disposition in time is fixed. A necessity in topological scales may arise for different reasons. Indistinct mapping of events on the scale or indistinct temporal relations established between events may occur. The former case is illustrated by the statement "At the beginning of May a fresh wind began to blow". Corresponding to the second case is, for instance, the statement "Shortly before the shot the neighbor came out of the criminal's room". The difference between these two types of indistinctness is the fact that during transition from the absolute scale, whose units are the days of the month, to the other, also absolute scale, whose units are the months themselves, it is possible to "hide" the indistinctness of the first statement. For statements of the second type this is principally impossible.

Between the topological scales introduced by us and the indistinct modeling relation discussed in §2.11 there exists a direct connection. The matrix defining this relation, constitutes essentially a method of transition from distinct scales (absolute and relative) to indistinct scales, which enter into the composition of topological scales. Hence, frequently, topological scales are called *fuzzy* or *indistinct*. We however, in our book shall stick with the introduced term, because it is not always that a topological scale may be interpreted as a selection of functions of application, characterizing the values of linguistic variables, as it was in §2.11. We also note, that the matrix, specifying an indistinct modeling relation, allows us to transform topological scales into certain metric ones. With the aid of a transformation of an indistinct modeling relation, which retains order, it is possible to introduce classes of topological scales (in analogy with scales, resulting for metric scales with the aid of operations D_n and R_n or by monotonic transformations).

To complete the model of representations in temporal logic we must also fix that selection of temporal relations, which will be used to described the temporal structure of situations.

Let me begin with point events. Our base relation $(e_j \tau_m t_i)$, where m - number of scale, we have introduced already. Let us also introduce the form relations: r_{21} - *simultaneous*, r_{22} - *to be earlier*, r_{23} - *to be later* and r_{26} -

to be immediately adjacent on the left[*]. It is clear that with the aid of the introduced relations it is possible to describe any temporal structure, into which,enter only point events.

Example 3.12. Consider the following text: "At 9 o'clock Paul, as usual mounts a horse and, accompanied by the heir, departs to review the troops. 10 o'clock. The usual plazaparade, the changing of the guard in the presence of the emperor... From 11 o'clock on their majesties were pleased to promenade mounted around the city: his imperial majesty with count Kutaysov, and her imperial majesty with the lady-in-waiting Protasova, 2nd". The city, as usual, becomes vacated until one P.M., everyone is afraid to meet Paul... After dinner "At 4 P.M. her imperial majesty with lady-in-waiting Protasova 2nd was pleased to have an excursion in a coach to the Novodevichiy Smol'nyy Monastery and returned at 6 o'clock. Later their imperial majesties were pleased to spend the time that evening with their imperial highnesses and important personages in the living room...".[**]

On Figure 3.15 is shown the mapping of the events mentioned in our text, to an absolute scale. That these events occurred on the 11th of March of 1801, is indicated in the source, from which was taken the text for the example. As we see, on the axes of years, months and days all events listed in the text are mapped all together into a single event on each of them. On the hours axis the events diverge. These events are either point events (e.g. Paul mounts a horse and departs to review the troops), or represent some set of events on the hours axis forming a continuous sequence (e.g. the city becomes empty). Some of the events are not tied in the text to exact hours (e.g. the return from the promenade). Finally, the event tied to the evening spending of time of the imperial couple with their children is not marked on the hours scale, since the text did not indicate either the beginning nor the end of that event. It is only clear that it began after six p.m. Such events are related to topological rather than absolute scales.

---

[*] Relation numbers r_{21}, r_{22}, r_{23}, and r_{26} coincide with the numbers assigned to them in Table 2.1.

[**] Eydel'man, H.Ya. The Border of Centuries, Moscow, Mysl' Publishers, 1982, pp. 268, 274, 275.

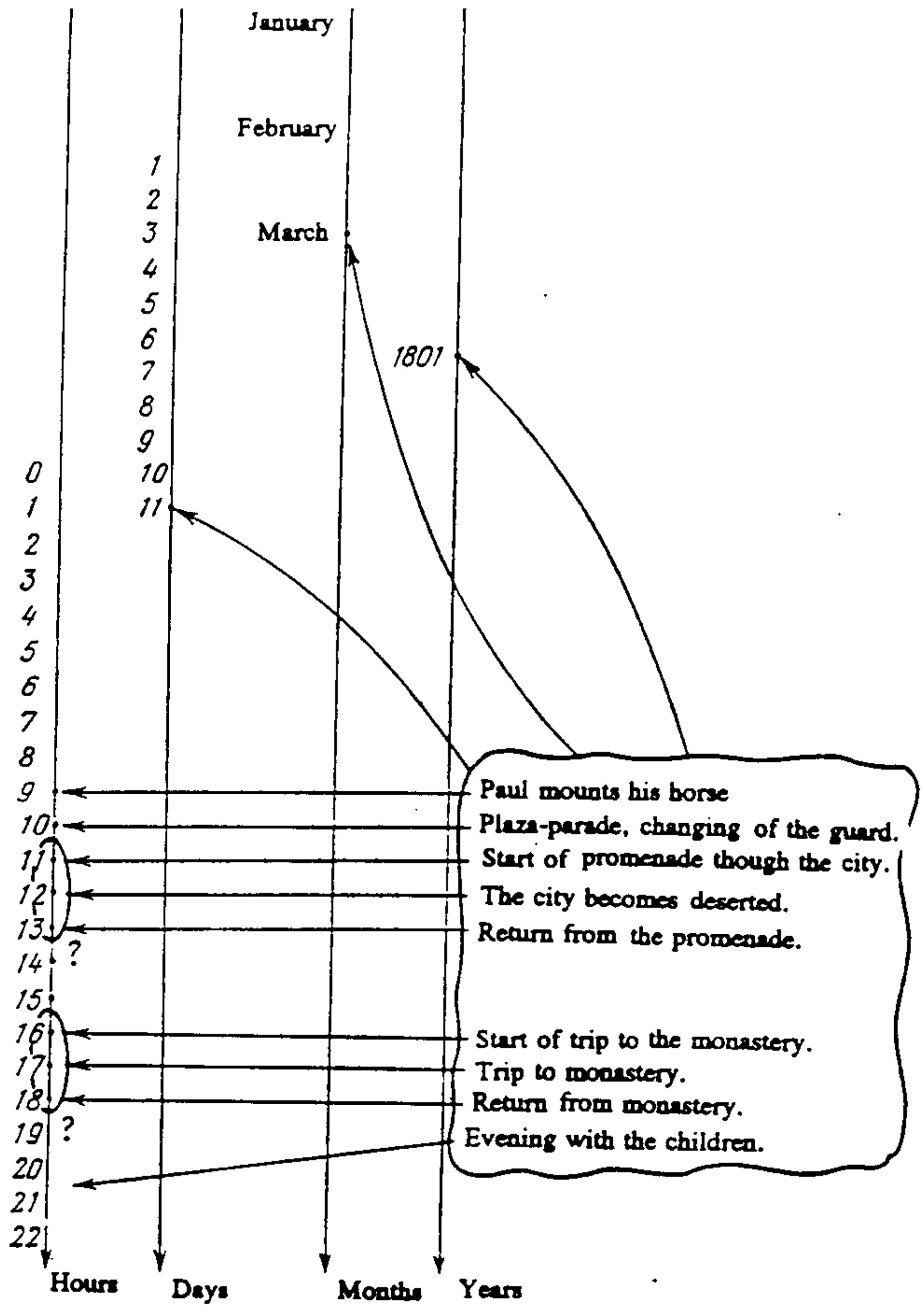


Figure 3.15

For control systems it is also important to introduce the function $\zeta(e_i,e_j)$, with the aid of which it is possible is determine the distance between events e_i and e_j as a metric scale, and also the usual arithmetic operations on distances. This permits the control system to obtain, for example, an answer to questions of the type "How much earlier did event e_i occur than event e_j?", "How much time elapsed between events e_i and e_j?", "If

e_i occurred m units after e_j and e_j was n units earlier than e_g, how much time did elapse between the events e_i and e_g?"

It is easily seen that the base system of relations chosen by us for point events is not the only possible one and not the minimal. Having the function ζ it is possible to assume that events e_i and e_j are in the relationship r_{21}, if $\zeta(e_i, e_j) = 0$, and are in relationship r_{26}, if this function assumes the value of 1. Relationship r_{22} and r_{23} can be expressed via each other by using the relationship of conjunction and negation: $(e_i r_{22} e_j) = \overline{(e_i r_{23} e_j)}\,\overline{(e_i r_2 e_j)}$. But the desire for minimalty of the base selection of relations is not always justified. The smaller the selection of acceptable relation for the system is, the more ponderous will be the representations of situations in it. True, a decrease in the number of relations may lead to a decrease of the number of rules of deduction in the model of the input, but the length of the output will still increase on the average. For this reason, neither here nor further on, will we solve the problem of minimization of the base set of relations in one or other pseudophysical logic.

In topological scales the relation τ is not realized. Instead of it there appears the relation $\bar{\tau}$, analogical to an indistinct relation of modeling. The relations r_{22}, r_{23}, and r_{26} are retained. They are specified by means of a special indistinct quantifier of the type *not long before that*, *soon*, *right after that*, *etc*. Hence, these relations are also tied to functions of affliction and indistinct modeling relation. Lastly, the relation r_{21} may be specified either as an exact or once again as an indistinct modeling relation (for instance with the aid of a quantifier of the form *almost simultaneously*).

In working with topological scales, use in also made of the idea of a universal scale, which was discussed in §2.11. All events, for which knowledge on the frequency histograms of these events can be introduced into the control system, are transferred to the universal scale.

Let us now pass on to interval events. We will introduce for them eight relations, the semantics of which is illustrated in Figure 3.16. These relations have the following names: r_{22} - *to be earlier*, r_{23} - *to be later*, r_{24} - *to begin simultaneously*, r_{25} - *to end simultaneously*, r_{26} - *to be adjacent in time on the left*, r_{27} - *to overlap in time*, r_{28} - *to coincide in time*, r_{29} - *to be inside timewise*. (The names of all relations are taken from Table 2.1.)

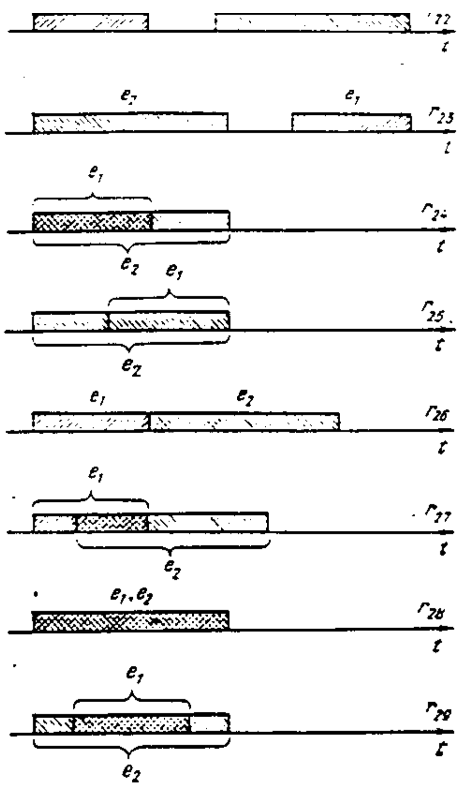


Figure 3.16

If for all interval events, are given the markers of their beginning and end or the length of the event and one of its end markers, then the relations between interval events may be substituted by a combination of relations among markers, i.e. among point events. If, for example, the events e_1 and e_2 are in the relation r_{25}, then the following relations hold for their markers:

$$((\mu_\kappa^1 r_{21} \mu_\kappa^2) \vee (\mu_\kappa^1 r_{22} \mu_\kappa^2) \vee (\mu_\kappa^1 r_{23} \mu_\kappa^2) \vee (\mu_\kappa^1 r_{26} \mu_\kappa^2) \vee (\mu_\kappa^2 r_{26} \mu_\kappa^1)) \vee (\mu_\kappa^2 r_{21} \mu_\kappa^1).$$

In other words, for markers of the beginning of events e_1 and e_2 arbitrary relations may be realized, while for markers of the ends of these events the relation r_{21} is always fulfilled. Similarly, it is possible to write relations between markers of events e_1 and e_2, when different relations for interval events are realized between them.

However, unfortunately, this is not always possible. Very often the texts contain incomplete information about the moments of beginning and ending of interval events. In that case the relations between them are actually given on a topological scale.

Example 3.13. Let there exist the following test: "Baron Münchhausen, as he was in a night cap and long night shirt, slipped his feet into his shoes and in the next second sat already in the coach and hurried the field-courier, although the latter was already desperately hurrying himself, jumped down from time to time from the driver's seat from impatience, and giving the horses an example, overtook them several times. An annoying delay occurred when they had already approached Postdam. The horses got so heated up, that there was no possibility of stopping them. It was necessary to drive around Potsdam thirteen times before, finally it was possible to hold down the thirty-six fire-breathing dragons, - this took at least three minutes. The horses traces, bits, and bellybands, were smoldering while the shaftbows and draftbars caught on fire, and they burned up before the firemen got there".* The cited text can be in different ways transformed into a tempered structure. This depends on the method of separating the events out of it. Assume that the separation of events is such that we have the following events the upper index of which shows whether they pertain to Münchhausen, the field-courier, the horses or the firemen.

e_1^M - Münchhausen slips feet into shoes, e_2^M - Münchhausen gets into the coach and sits in it; e_3^M - Münchhausen hurries the field-courier; e_4^M - Münchhausen leaves the coach; e_1^Φ - the field courier gets on the drivers seat; e_2^Φ - the field courier jumps down from the drivers seat and for some time runs along-side the horses; e_3^Φ - the field courier tries to stop the horses; e_1^z - the horses are rushing along; e_2^z - the horses ride around Potsdam; e_3^z - the horses stop; e_4^z - the horses traces, bits and bellybands are smoldering, while the shaftbows and draftbars burn; e_1^z - the firemen arrive.

The corresponding structures are shown in Figure 3.17. For better clarity of this figure, the events that are related to different personages are shown on different time axes. Different events are highlighted by lines, shading or filled with dots. Among the individual axes of time there exists a time correspondence - events lying along the same vertical line, occur simultaneously. On each of the axes are superposed those events that occur

---

*Globa, A.P. Tales for Day and Night, Moscow, Sorremebbik Publishers, 1976, p. 72

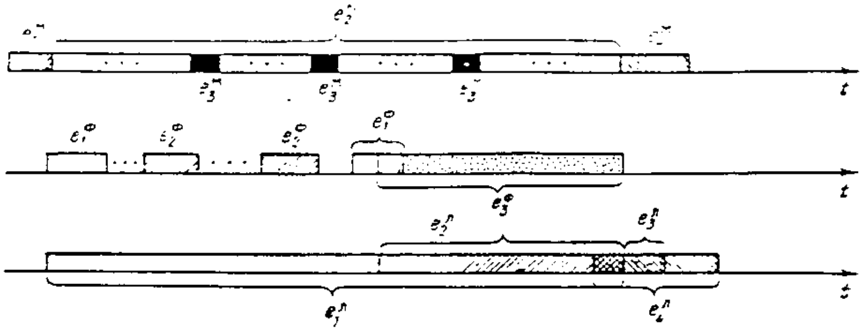


Figure 3.17

simultaneously for a given personage. Let us give some additional explanations. We do not know from the text, when exactly occur the time moments during which Baron Münchhausen hurries the field courier. Apparently this happens more than once. For that reason inside the event e_2^M there appear a certain number of events e_3^M. In a real system it can occur in a random manner. the lengths of the events for Münchhausen are not defined in the analyzed text. But it is clear, that the length of e_2^M considerable surpasses the lengths of the other events, while events e_1^M and e_3^M are close to each other in length. The text says nothing at all about event e_4^M. It could even be omitted on Figure 3.17. But if it had not occurred, then the Baron would have probably been burnt up, which contradicts the later text of Globa.

For the field-courier the number of events e_1^Φ and e_2^Φ are not defined in the text. It is only clear that the total number of events e_2^Φ must be one more than the number of events e_2^Φ. In addition, e_1^Φ and e_2^Φ must alternate and the sequence begins and ends with event e_1^Φ. The last event e_1^Φ, coincides in time with the beginning of the attempt to stop the horses. The length of the last event in the text is determined by the words "For this, it took at least three minutes". It is possible to assume, that events e_1^Φ and e_2^Φ have smaller duration, although for e_2^Φ this may turn out to be wrong, if the field-courier is able to run for a long time at the rate of the horses.

Finally, for the horses up to event e_3^z their running continues part of which ("at least three minutes") coincides with the event e_2^z. It is not clear from the text when their harness begins to burn, hence the beginning of the event e_4^z is not determined. It is clear, more or less, that the catching on fire must have occurred while they still ran. As concerns the firemen the relative placing of the event of their arrival to the pace of occurrence, there are hardly any difference of opinion.

We note that in passing from the text to the temperal structure we have used not only those data which are directly contained in the text, but also some additional information following from the scenarios related to some or other actions. A similar situation arises very often. It is precisely the scenarios which play a large role in forming the temperal structure and supplementation of description. We shall speak about this below in greater detail. For chain events we will not introduce a different system of relations. We will reduce them to a succession of interval and point events.

We have completed the construction of a model of concepts of temperal logic. This model consists of a selection of scales, relations of projections τ or $\bar{\tau}_i$ of events onto scales, of a system of relations in time among events, connecting the relation for interval events with point events via markers, and function ζ for different scales together with a selection of necessary arithmetic procedures.

Now it is possible to pass on to the model, derivation of new statements about time on the basis of those derivation rules, which we shall put into our logic and data stored in the control system.

All derivation rules can be divided into two groups: *syntactic* and *semantic*. In the first group are related the derivation rules coming from the algebraic properties of those relations to which they are applied. Rules of the second group come from a certain external semantic interpretation of these relations and scenarios used in deriving new facts and relations.

The total number of relations used by us in the temperal logic of relations is equal to eight. Rules of the first group for them are constructed as follows. A certain relation is examined. Its algebraic properties are clarified and the corresponding rules are written out. For example, examining the relation r_{22}, we convince ourselves that it is anti-reflexive, anti-symmetric and transitive. This allows us to write the following three statements:

$$(\overline{er_{22}e});(e_1r_{22}e_2) \rightarrow (\overline{e_2r_{22}e_1});(e_2r_{22}e_1);(e_1r_{22}e_2) \rightarrow (e_2r_{22}e_3).$$

Among these statements of fundamental interest is the third, which has the form of a classical rule of derivation in the form of a production. The second statement also has a productional form, but its utilization for obtaining new facts is doubtful. Rather, it may turn out to be useful in the checking of the system memory for noncontradiction. Another example may be a rule of the type $(e_1 r_{25} e_2)(e_2 r_{26} e_3) \Rightarrow (e_1 r_{26} e_3)$. Its meaning reduces to the fact that if the events e_1 and e_2 end simultaneously and event e_2 adjoins directly on the left of the event e_3, then the event e_1 also adjoins directly on the left the event e_3. Let us illustrate the operation of the rules of the first group with an example.

Example 3.14. Let us return to the situation described in the text cited in Example 3.13. Let us extract from that text a selection of kernel constructions (triads of the type [azb]). which is directly contained in it. Analyzing the text in succession, we obtain the following selection: $(e_1^{\text{M}} r_{26} e_2^{\text{M}})$; $(e_3^{\text{M}} r_{29} e_2^{\text{M}})$; $(e_2^{\Phi} r_{29} e_2^{\text{M}})$; $(e_1^{\text{I}} r_{22} e_2^{\text{I}})$; $(e_3^{\Phi} r_{21} e_2^{\text{I}})$; $(e_2^{\text{I}} r_{22} e_3^{\text{I}})$; $(e_1^{\Pi} r_{23} e_4^{\text{I}})$. Let us see, what additional information can be obtained from this selection. Comparing the first triad with the second and third, we can deduce the new information, using the rule of derivation $(e_1 r_{26} e_2)(e_3 r_{29} e_2) \Rightarrow (e_1 r_{22} e_3)$.

Applying it, we find two more triads: $(e_1^{\text{M}} r_{22} e_3^{\text{M}})$ И $(e_1^{\text{M}} r_{22} e_1^{\Phi})$. From the fourth and fifth triad, with the aid of the respective rule (we do not write them out, so as not to encumber the presentation) is obtained the triad $(e_1^{\text{I}} r_{22} e_3^{\Phi})$. The fourth and sixth triads generate $(e_1^{\text{I}} r_{22} e_3^{\text{I}})$ while the fifth and sixth generate $(e_3^{\Phi} r_{22} e_3^{\text{I}})$. Finally the last two triads generate a new fact of the form $(e_4^{\text{I}} r_{22} e_1^{\Pi})$.

It is of interest to compare the obtained result with the tempered structure shown in Figure 3.17. It becomes clear, that the derivation rules of the first group provide insufficient information for a supplementation. The events e_4^{M} and e_1^{Φ} did not occur at all, and, hence, there did not occur any temperal relations connecting them with the other events. It remains completely undefined, how the events e_2^{I} and e_4^{I} are interrelated, and also the events e_3^{I} and e_4^{I}. All this information cannot be obtained with the aid of the derivation rules of the first group. It originates only upon addressing the knowledge stored in the form of scenarios in the memory of the system. The derivation rules operating with these scenarios, belong to the second group of rules. We shall illustrate their operation on two examples of supplementation of descriptions.

Example 3.15. Consider the text already utilized in Example 3.3: "The aircraft performed the landing. The passengers are waiting for the luggage". In an attempt to construct a temporal structure of this text, using only the text itself, we experience a complete failure. We select two events: e_1 - the aircraft has landed; e_2 - the passengers are waiting for luggage. But we cannot establish any temporal relation between them. However, if the system memory stores a scenario, like the one shown in Figure 3.2, then, with its aid it is easy to establish the needed relation, namely $(e_1 r_{22} e_2)$.

Example 3.16. Assume that the input of the system receives the text "On Friday at the ball Vel'chkovskiy publicly insulted Zarenin, and Zarenin, not delaying a second, challenged him to a duel. After the ball, Zarenin called at his old friend General Murashev and asked him to be his second. A written challenge was composed, and on Saturday morning it was handed to Vel'chkovskiy. Now Zarenin and Murashev were writing a protocol about refining the duel, for the term of the reply, even with weighty reasons, had already expired".

After analyzing this text the system is able to separate out the following events; e_1 - insult of Zarenin, e_2 - challenging Vel'chkovskiy to a duel, e_3 - request by Zarenin, directed to Murashev, and the latter's agreement, e_4 - composing the written challenge, e_5 - handing in the challenge, e_6 - composing the protocol on the refusal. From the text it is possible to establish the following triads: $(e_1 r_{26} e_2)$; $(e_2 r_{22} e_3)$; $(e_3 r_{22} e_4)$; $(e_4 r_{22} e_5)$; $(e_5 r_{22} e_6)$; $((e_1 \tau f)$; $(e_2 \tau f)$; $(e_5 \tau s)$.

Here f and s denote respectively the days of the week: Friday and Saturday. Applying the general rules, related to the transitivity of the relations r_{26} and r_{22}, the system can derive further relations of the type $(e_1 r_{22} e_4)$. But from the text it can not derive any triads which would allow to map, with the aid of τ, onto the scale of days of the week the events e_3, e_4, and e_6. This could be done only in the case when the system memory contains a scenario describing the accepted order of conducting duels. If such a scenario is available[1] then it is clear that the event e_6 could not have occurred earlier than Monday morning (m), since the reply to the written challenge to the duel should have been received from the offender not later than 24 hours after handing him the written challenge. The "weighty reasons"

---

[1] For example, the dueling codex, described in the book: Durasov, Dueling Codex, City of St. Peter, 1908.

cited in the text permitted to postpone the answer by another 24 hours. After that, if the offender's reply was not forthcoming, a protocol of refusal from the duel, demeaning him, was composed. Therefore, e_6 could have occurred only on Monday, which allows us to derive the triad $(e_6 \tau m)$. The text states that Zarenin called on Murashov after the ball. This means, that most likely the event e_3 occurred on Friday, since in the XIX century balls ended by midnight, as a rule. Hence, with a great degree of probability, the events e_3 and e_4 occurred on Friday, which allows the system to derive two more triads: $(e_3 \tau f)$ and $(e_4 \tau f)$.

There remains for us to discuss the problems related to output on topological scales. For mapping events on such scales, methods like these described in §2.11 are normally used. First of all, a list of quantifiers is elected, which could form a certain ordering scale of diffused moments of time.

As an example, a list for temporal diffused quantifiers can serve the sequence: *very long ago, long ago, not long ago, quite recently, just now, now, soon after that, very soon, soon, in the near future, in the future*. It is more or less clear, that such a sequence is able to generate a certain ordinary scale. In order that the letter would serve purposes of deviation and supplementation of descriptions, it is necessary to pass on to a more exact scaling. For this is used the apparatus of functions of application, discussed earlier in the same §2.11. For each element of the stated listing is constructed, with the aid of an experiment with people, a function of application of this element to sector [0,1]. Next is selected such a level of cut-off α, which could form a good scale on the given segment. As was said earlier (see §2.11), this results in segment [0,1] being subdivided into non-intersecting intervals, each of which corresponds, with a guaranteed level of the application function, equal to $0<\alpha<1$, to a certain element from the list that forms the verbal scale. This process is indicated on Figure 2.17 for diffused frequency quantifiers. For temporal quantifiers it is realized in a similar fashion.

Upon obtaining a topological scale it turns out to be possible to construct a system of derivation rules on it. We shall not specially pause on this in temporal logic, since in the next paragraph we will describe the entire procedure in sufficient detail. Transferring it to the case of the more simple temporal logic causes no difficulties.

§3.6 Spectral logic**

As in the case of temporal logic, we will begin the construction of spatial static logic with a model of representations. Noting the sufficiently complete presentation of the technique of construction of pseudophysical logics on the example of temporal logic, we shall limit ourselves with a more fragmenting presentation for the spatial logic. Of greatest interest in spatial logic is that part of it, which is connected with the obtaining of the distances between objects. It is that part of the logic that we will consider here. Our presentation should supplement the general concept of these problems which face developers of pseudophysical logics. Other divisions, related to spatial static logic (operation with metric scales and mapping of statements on them, work with relations of mutual disposition of objects, relations of their orientation in space, and much other), are factually constructed similar to how it was done for temporal logic.

Let's begin with the formation of a list of diffuse quantifiers, which can be used to estimate distances on the topological scale of distances. Of course, this list can be more or less extensive. As an example we'll consider a list of 25 quantifiers, listed in the left column of Table 3.1. In it are shown results of experiments conducted on the basis of the given list of distance quantifiers, for a large group of test subjects. In the columns are indicated the numbers of test subjects who put a card with the respective quantifier into the position whose number is shown at the top.

What conclusions can be made from the data of Table 3.1? Several conclusions can be made. The test subjects distinguish poorly among certain of each others quantifiers: *right-up-to* and *very-very near*, *rather near* and *quite near*, *near*, *nearby*, *next to*, *beside*, *near to*, and others. Two quantifiers: *not far* and *not near* are not understood the same way by the test subjects. In constructing a good ordering scale this must be taken into account. Therefore, instead of the original list of quantifiers only those should be retained for constructing the scale which are shown in the left column of Table 3.2.

TABLE 3.1

Lexemes and Combinations of Lexemes	0	1	2	3	4	5	6	7	8	9	10
1. Up-against	27										
2. Very-very close	20	7									
3. Very close	2	21	4								
4. Fairly close	1	13	13								
5. Quite close	2	20	4	1							
6. Right next to		15	7	4	1						
7. Near		4	21	2							
8. Nearby		3	18	6							
9. Next to		3	19	3	2						
10. Beside		2	15	8	2						
11. Near to		5	18	2	2						
12. Not near				3	11		9	4			
13. Not very near				4	12		8	3			
14. Not far			1	2	11		10	3			
15. Not far, but not very near					12		12	3			
16. Not far, not near						27					
17. Not very far, but also not near					9		14	3	1		
18. Not very far				3	8		14	2			
19. Distant							2	12	10	2	1
20. Quite far									4	20	3
21. Rather far									6	18	3
22. Very-very far										7	20
23. Very far									4	21	2
24. Far								1	19	7	
25. At infinity											27

Here has already been carried out a mapping of the results from Table 3.1 into values of the application function in accordance with the procedures described in 2.11. At the top, with a spacing of 0.1, are listed the values of the variable from segment [0,1], on which are defined seven functions of affilation (the lines of the Table 3.2) for all remaining quantifiers of the list.

Table 3.2

Lexemes and Combinations of Lexemes	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
1. Up against	0.966	0.033									
2. Very near	0.2	0.666		0.133							
3. Near		0.133		0.733	0.133						
4. Not far, not near						1					
5. Far							0,1	0.666	0.133		
6. Very far								0.066	0.266	0.666	
7. Very very far											1

From the table it is clearly seen, that the remaining seven quantifiers form a good ordering scale for distance relations. Within the frames of this paragraph we will consider only the two-dimensional spatial logic of distances. Transferring the results given below to the case of three or more dimensions taken place in a purely technical manner. For an interpretation of quantifiers of distances we will have to consider also a list of diffuse quantifiers that estimate the dimensions of objects.

The connection between estimates of distances on a topological scale and estimates of sizes of objects, between which the distances are estimated, are rather well traced with the aid of examples. Let us examine the two statements: "The man is located far from the city" and "The man is located far from the automobile". Clearly, in the first case *far* estimates a greater distance, than in the second. Two further statements - "The book is located near the table" and "The forest is located near the village" - also characterize the dependence of the estimate of a factual distance, conveyed by the quantifier *near*, on the sizes of the objects discussed here.

Let us introduce a list of size quantifiers, which forms an ordinary scale for size estimates. This scale is obtained as a result of the same kind of experiment, which for distances is reflected in Table 3.1 and Table 3.2. The list of size quantifiers has the following form: *very-very small*, *very small*, *average*, *large*, *very large* and *very-very large*. Here man fits into the class of average objects, which makes him a sort of modulus, with respect to which is made the division of surrounding objects into classes. To the class of very small objects belong, for example, a pin or a button. A map or a book may represent small objects, while a city, depending on its size fits either into the class of very large or very-very large objects.

Let us now conduct a psychological experiment, in the course of which the test subjects must estimate from some adopted ten-point estimating system the degree of closeness of pairs of quantifiers, one of which pertains to sizes and the other to distances. The result of one such experiment is given in Table 3.3.

Table 3.3

Distance	Size						
	Very-very small	Very small	Small	Average	Large	Very large	Very-very large
Up against	0.966	0.033					
Very near	0.2	0.666	0.133	0.133			
Near		0.333	0.733	0.266			
Not far, not near				1.0			
Far					0.933	0.066	
Very far					0.266	0.733	
Very very far						0.066	0.933

From an analysis of this table there visibly follows the existence of a connection between estimates of distances and of sizes on topological scales. Between the lists of some or other quantifiers there exists a definite correspondence. It can be expressed in the form of a certain hypothetical statement.

Hypothesis 3.1. For estimating distances between two objects it is possible to use a third object placed between them, touching both. The size of this object defines uniquely the distance between the original objects in accord with Table 3.4.

Table 3.4

Size	Distance
1. Very-very small	Up against
2. Very small	Very near
3. Small	Near
4. Average	Not far, not near
5. Large	Far
6. Very Large	Very far
7. Very-very large	Very-very far

Let us henceforth designate by a some object named i and of size q. We will specify the sizes by ordinal numbers in accord with the numeration of sizes in Table 3.4. By R_j we shall designate the ratio of size on the topological scale. Indexing coincides with the ordinal numbers of the respective ratio in the same Table 3.4. The notation ($a_1^{q_2} R_j a_2^{q_2}$) means, that the shortest distance between the boundaries of objects a_1 and a_2 with sizes q_1 and q_2 is estimated as R_j. On Figure 3.18 is shown a graphic representation of this situation.

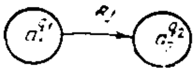


Figure 3.18

Let us state two further hypotheses about distance estimation on a topological scale. Consider Figure 3.19. On this Figure object $a_1^{q_2}$ remains unchanged, but instead of object $a_2^{q_2}$ are substituted objects of different sizes, but in such a way that the true distance between the boundaries of the objects does not change.

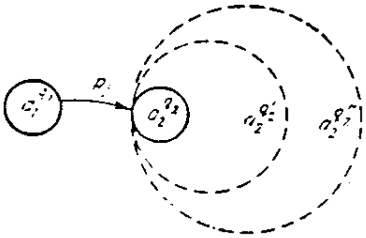


Figure 3.19

Hypothesis 3.2. Distance estimation on a topological scale for the triad ($a_1^{q_2} R_j a_2^{q_2}$) does not change upon varying the sizes of the object a_2 while retaining the time distance between the boundaries of the objects.

As did Hypothesis 3.1, so the given hypothesis may be accepted or not accepted, since its veracity cannot be proved strictly. Its only justification is that (as for other hypotheses), with its aid results are obtained which do not contradict human intuition.

Note, that upon varying the sizes of the object a_2, from which is estimated the distance (it is better to say "from the point of view" of which), the estimate of distances even with retention of the true distance can vary. If, for instance, some man is located close to a house, then for a pinhead this distance may turn out to be extremely great.

From this follows one more hypothesis, that tells of the nonsymmetry of distance estimates.

Hypothesis 3.3. With different sizes of objects a_1 and a_2 in triads $(a_1^{q_1} R_j a_2^{q_2})$ and $(a_2^{q_2} R_k a_1^{q_1})$ the values of R_j and R_k may differ from each other.

The nontransitivity of R_j is obvious. Clearly, if a_1 is near a_2, a_2 near a_3, etc., then between a_1 and the last object in such a sequence there may be set up practically any relation from *near* to *very-very far*. This depends on the number of elements in the sequence considered.

Let us form the last hypothesis, as the basis of which can serve the results of numerous psychological experiments, and also man's intuition.

Hypothesis 3.4. For any three objects of the same size, placed tightly up to one another on the same straight line, the distance between the end objects in a trial is estimated by the quantifier *very near*.

All the formulated hypotheses (if they are adopted, of course) are used in constructing the base matrix of distance estimates on a topological scale (Table 3.5).

Table 3.5

Up-against	om ? om	m ? m	s ? s	b ? b	ob ? ob	oob ? oob
1. om	obl	obl	obl	obl	obl	obl
2. m	bl	obl	obl	obl	obl	obl
3. s	nd, nbl	bl	obl	obl	obl	obl
4. b	d	nd, nbl	bl	obl	obl	obl
5. ob	od	d	nd, nbl	bl	obl	obl
6. oob	ood	od	d	nd, nbl	bl	obl

In this matrix are utilized the following designations: om - very small, m - small, s - average, b - larger, ob - very large, oob - very-very large, obl - very near, bl - near, nd, nbl - not far, not near, d - far, od - very far, ood - very-very far. The notation q?q, where q - is an indication of a certain size, corresponds to the fact that instead of the in between, object it is possible to insert tightly at the two end objects of the indirected size a third object whose size is given in the left column of the matrix. Here the distance between the end objects can be estimated by means of distance estimation shown in the corresponding cell of the matrix. For example, if between two objects of average size a small object is placed tightly next to them, then the distance between the end objects will be very near, while upon placing tightly between average sized objects of a very large object there results a distance between the original objects, which is characterized by the quantifier *not far*, *not near*.

Example 3.17. The distance between two saucers has been estimated as *very far*. What would have been the estimate of the same distance, if instead of saucers we had people? Let us assume that saucers are small objects. In column m?m of Table 3.5 we locate the quantifier od and find, that it corresponds to placing tightly between the saucers of a very-very large (oob) object. Now, in the line corresponding to this object at the intersection with column s?s (we assume, as stated earlier, that a man belongs to the class of average-sized objects) we find the estimate of distance in the form of quantifier *far*. Thus, if in the place of saucers people are substituted, then between them will be a distance, characterized by the quantifier *far*.

Example 3.18. Let two persons be located on the same straight line on both sides of a modern city building, having the shape of a closed square or a rectangle close to a square (in the ideal - a cylindrical form, like some experimental houses in Yerevan or Bratislava, for example). And we wish to estimate the distance between these two people. Since man belongs to the class of average sized objects, while a house apparently should be considered in the class of a large objects, then at the intersection of column s?s and the line for large objects in Table 3.5 we find the distance quantifier that is of interest to us. As it follows from the base matrix, this quantifier is the quantifier *near*.

Now it is possible to pass on to the discussion of the rules of derivation characteristic for a topological distance scale in a static spatial logic. Consider first the situation depicted on Figure 3.20. On this figure

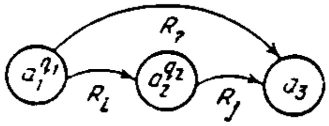


Figure 3.20

are three objects of given sizes (the size of object a_3 plays no role, as follows from our hypothesis 3.2), located on a single straight line. Estimates of distances can be made between objects $a_1^{q_1}$ and $a_2^{q_2}$, and also between objects $a_2^{q_2}$ and a_3. It is required to find the estimate of the distance between objects $a_1^{q_1}$ and a_3. In other words, it is required to construct a system of rules in the form of such productions:

$$(a_1^{q_1} R_i a_2^{q_2})\ (a_2^{q_2} R_j a_3) \Rightarrow (a_1^{q_1} R_? a3),$$

with the aid of which it would be possible to specify a complete system of derivations about distances on a topological scale for objects situated on a single straight line. Let us examine in succession four possible cases.

1. Objects a_1 and a_2 have the same sizes, R_i and R_j coincide. Then the derivation rule is written thus:

$$(a_1^q R a_2^q)\ (a_2^q R a_3) \Rightarrow (a_1^q R_? a_3).$$

In this case $R_?$ is determined with the aid of the operation $R_? = R + 1$. The meaning of this operation is explained with the aid of Table 3.6.

Table 3.6

R	$R \dotplus 1$	$R \dot{-} 1$
Up-against	Very near	Up-against
Very near	Near	Up-against
Near	Not far, not near	Very near
Not far, not near	Far	Near
Far	Very far	Not far, not near
Very far	Very-very far	Far
Very-very far	Very-very far	Very far

Example 3.19. Let a_1 and a_2 be small objects and the estimate of the distance between them, and also between a_2 and a_3 be *near*. Then the estimate of the distance between a_1 and a_2 is *not far, not near*.

2. Objects a_1 and a_2 have the same sizes, but R_i and R_j are different. In this case the derivation rules have the form

$$(a_1^q R_i a_2^q)\ (a_2^q R_j a_3) \Rightarrow (a_1^q R_? a_3).$$

The value of $R_?$ is determined on the basis of the following relation:

$$R = \begin{cases} \max(R_i, R_j), & if \quad |i - j| > 1, \\ \max(R_i, R_j) + 1, & if \quad |i - j| = 1. \end{cases}$$

If $|i-j| = 0$, then this case reduces to the preceding one.

Example 3.20. Let, for small a_1 and a_2, the distance from a_1 to a_2 be estimated by the quantifier *near*, and from a_2 to a_3 - by the quantifier *far*. Then the distance from a_1 to a_3 will be estimated by the quantifier *very far*, since in our case $|i-j| = |3-5| = 2$.

3. Estimates of R_i and R_j coincide, but the sizes of objects a_1 and a_2 differ. In this case the derivation rules have the form

$$(a_1^{q_2} R a_2^{q_1}) \ (a_2^{q_2} R a_3) \Rightarrow (a_1^{q_1} R_? a_3).$$

In this case it is first necessary to "equalize" the sizes of a_1 and a_2. For this purpose we will use hypothesis 3.2 and the base matrix, reflected in Table 3.5. In the first triad we ascribe to a_2 the size q_1, which does not affect quantifier R. In the second triad we ascribe the same size to a_3, which does not affect R, provided the size of a_2 is not changed. But since we do not change it, it is necessary to find a new value of the quantifier. That's what Table 3.5 serves for. If the new value R^1 in the second parenthesis retains the same value of R, then the case reduces to the first of the earlier examined cases. If, however, this does not happen, then our case will reduce to the second of the earlier considered cases.

Example 3.21. Let the distance from a cup (a small object) to the house (a large object) be estimated by the quantifier *near*, and the distance from the house to some third object a_3 - by the quantifier *not far, not near*. To pass from the class of large objects to small objects we use the base matrix (Table 3.5). Find the column b?b in this matrix and look in it for the quantifier *not far, not near*. To it corresponds the last line of the table, i.e., an object belonging to the class of the very-very large ones. Placing such an object between small objects (i.e., replacing the large a_2 by a small one), we find in column m?m, at its intersection with the last line the value of R^1. As it is seen from Table 3.5, this value is *very far*. Since it does not coincide with the quantifier *near*, this is the second case. Noting that $|i-j| = |3-6| = 3$ we obtain finally, that the distance from the cup to object a_3 must be estimated by the quantifier *very-very far*.

4. Let neither the sizes of a_1 and a_2 coincide, nor the quantifiers R_i and R_j. Then it is possible, using the base matrix, to first equalize the sizes of the objects, and then obtain the same situation as in the preceding case.

The cited rules of derivation which work well for the case of three objects situated on a single straight line, satisfy in this case the human intuition, but on applying to the case of an arbitrary number of objects on a single straight line can lead to undesirable effects.

Example 3.22. Let there be a row of trees planted along a road. Between adjacent trees the distance is estimated by the quantifier *near*. If a_1, a_2, ..., a_n are the trees, then from a_1 to a_2 it is *near*, from a_1 to a_3, on the basis of the first case of derivation rules it is *not far, not near*, from a_1 to a_4 on the basis of the same rule, it is *far*, and for distances from a_1 to a_5 and all subsequent trees in the row the relation *far* will be retained. This happens as result of absorption by one quantifier of the others, which are more than one unit away from it.

At the present time, to eliminate the effect of absorption, there are no methods that are effective to any degree. One particular one, unfortunately not always applicable, is the method of partitioning, at each step, of the objects into pairs, inside of which the quantifiers differ by unity in the index, and the rules of derivation are applied specifically to such pairs. If there are no such quantifiers, then use is made of pairs having the same distances between objects. This approach is applicable, if all pairwise distances differs in the quantifier indices by more than unity.

Example 3.23. Under conditions of Example 3.22, after having found the distance from a_1 to a_3, it makes sense to find the distance from a_3 to a_5, which will be estimated in the result as *not far, not near*. Applying now the derivation rule to objects a_1, a_3 and a_3, a_5, we obtain, that the distance from a_1 to a_5 is estimated by the quantifier *far*. Similarly, the distance from a_5 to a_9 is also estimated by the quantifier *far*. If now the derivation rule is used for objects a_1, a_5, and a_9, then the distance from a_1 to a_9 will be already estimated by the quantifier *very far*. Continuing the procedure in the same manner, it is possible to estimate distances by indexwise more and more increasing quantifiers (in the assumption, that beyond the quantifier *very-very far* there takes place an infinite sequence of the m: *very-very-very far*, *very-very-very-very far*, and so forth).

Let us now pass on to the construction of derivation rules for the spatial logic of distances on a plane. The base matrix (Table 3.5) provides the possibility of transforming any arbitrary, sizewise, objects into very small objects, which subsequently we shall call point-objects for brevity. Hence, in constructing derivation rules on a plane we will consider only point-objects.

In Figure 3.21 are shown three objects a_1, a_2, and a_3, positioned arbitrarily on a plane. Let the distances R_i and R_j be known to us (i.e., we

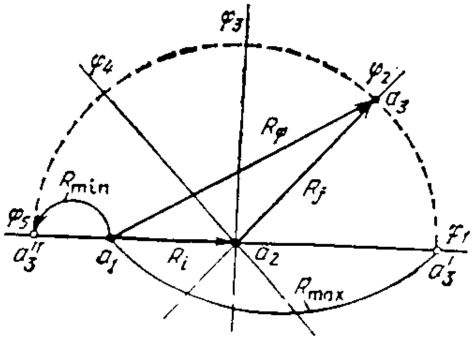


Figure 3.21

know their verbal estimate), and it is required to find distance $R\varphi$. Let us construct two auxiliary objects a_3' and a_3'' such as is shown on the figure. It is clear that for the indices of the quantifiers the irregularities $R_{min} \leq R\varphi \leq R_{max}$ are fulfilled. For finding R_{min} and R_{max} the following spatial operations are used:

$$R_{\min} = \begin{cases} \text{Up-against} & \text{if} \quad i = j, \\ R_{i-1}, & \text{if} \quad j = i+1, \\ R_{j-1}, & \text{if} \quad i = j+1, \\ R_i, & \text{if} \quad i - j \geqslant 2, \\ R_j, & \text{if} \quad j - i \geqslant 2; \end{cases}$$

$$R_{\max} = \begin{cases} R_i \dotplus 1, & \text{if} \quad i = j, \\ R_i \dotplus 1, & \text{if} \quad i = j+1, \\ R_j \dotplus 1, & \text{if} \quad j = i+1, \\ R_i, & \text{if} \quad i - j \geqslant 2, \\ R_j, & \text{if} \quad j - i \geqslant 2. \end{cases}$$

For simplicity let us limit ourselves not by any angular directions φ on the plane, but by some number of fixed directions. With a sufficiently large number of them it is possible to obtain an arc close as desired approximation to the model with continuous value of φ. For clarity let us take a small number of directions as shown on Figure 3.21. In other words, let us assume that object a_3 can be located only along these five fixed directions. We construct a spatial table (Table 3.7), which determines, from the location of object a_3 and the computed values of R_{min} and R_{max}, the value of R_φ. On increasing the number of directions the character of the table does not change.

Table 3.7

$s \backslash \varphi$	φ_1	φ_2	φ_3	φ_4	φ_5
0	R_{max}	R_{max}	R_{max}	R_{max}	R_{max}
1	R_{max}	R_{φ_1}	$R_{\varphi_2} \dot{-} 1$	R_{φ_3}	R_{min}
2	R_{max}	$R_{\varphi_1} \dot{-} 1$	R_{φ_2}	$R_{\varphi_3} \dot{-} 1$	R_{min}
3	R_{max}	R_{φ_1}	$R_{\varphi_2} \dot{-} 1$	$R_{\varphi_3} \dot{-} 1$	R_{min}
4	R_{max}	$R_{\varphi_1} \dot{-} 1$	$R_{\varphi_2} \dot{-} 1$	$R_{\varphi_3} \dot{-} 1$	R_{min}
5	R_{max}	$(R_{\varphi_1} \dot{-} 1) \dot{-} 1$	$R_{\varphi_2} \dot{-} 1$	$R_{\varphi_3} \dot{-} 1$	R_{min}
6	R_{max}	$(R_{\varphi_1} \dot{-} 1) \dot{-} 1$	$(R_{\varphi_2} \dot{-} 1) \dot{-} 1$	$R_{\varphi_3} \dot{-} 1$	R_{min}
7	R_{max}	$(R_{\varphi_1} \dot{-} 1) \dot{-} 1$	$(R_{\varphi_2} \dot{-} 1) \dot{-} 1$	$R_{\varphi_3} \dot{-} 1$	R_{min}

Here s indicates the difference of indices for the quantifiers, corresponding to R_{max} and R_{min}. For the case s = 1 it is assumed, that R_{min} is not the quantifier *up-against*. If it is so, then $R_{\varphi 1}=R_{\varphi 2}=R_{\varphi 3}=R_{\varphi 4}=$*very near*. If s=0, $R_{max}=R_{min}$. The operation of spatial subtraction, utilized in this table, is defined in the right column of Table 3.6.

For full completion of the logic on a plane it is further necessary to construct rules of derivation of directions for the given direction of two objects with respect to a third one. In other words, it is further necessary to introduce derivation rules of the following type:

$$(a_1^{q_1}\varphi_m a_2^{q_2})\ (a_2^{q_2}\varphi_n a_3^{q_3}) \Rightarrow (a_1^{q_1}\varphi_? a_3^{q_3}).$$

Let us cite examples of such rules, limiting ourselves to only four fixed directions *in front of, on the left, behind, on the right*. These rules are collected in Table 3.8.

Table 3.8

	In front of	On the left	Behind	On the right
In front	In front	In front, if $R_j \leqslant R_i$ On the left, if $R_j > R_i$	In front, if $R_j \leqslant R_i$ Behind, if $R_j > R_i$	In front, if $R_j \leqslant R_i$ On the right, if $R_j > R_i$
On the left	On the left, if $R_j \leqslant R_i$ In front, if $R_j > R_i$	On the left	On the left, if $R_j \leqslant R_i$ Behind, if $R_j > R_i$	On the left, if $R_j \leqslant R_i$ On the right, if $R_j > R_i$
Behind	Behind, if $R_j \leqslant R_i$ In front, if $R_j > R_i$	Behind, if $R_j \leqslant R_i$ On the left, if $R_j > R_i$	Behind	Behind, if $R_j \leqslant R_i$ On the right, if $R_j > R_i$
On the right	On the right, if $R_j \leqslant R_i$ In front, if $R_j > R_i$	On the right, if $R_j \leqslant R_i$ On the left, if $R_j > R_i$	On the right, if $R_j \leqslant R_i$ Behind, if $R_j > R_i$	On the right

Example 3.24. Let there be a forest far in front of the house, and near on the right of the forest is a windmill. Let us determine the direction from the house to the windmill. In our case $R_j > R_i$ and, according to Table 3.8, we observe that the windmill is in front of the house.

On increasing the number of permissible directions the table for determining the direction grows rapidly, and also increases the number of conditions related to the relation of R_i and R_j for selecting one or another different answer. But the gist of the matter does not change because of this.

As in temporal logic, in spatial logic it is not always possible to obtain a supplementation of a description, using only those rules of deviation which is described here. Often occurs the necessity of using additional information, stored in the memory of the system in the form of scenarios or in some other form.

Example 3.25. In Example 3.4 we discussed a scenario, in which is reflected the construction of a vacuum cleaner of a number of parts. Let the following description of a situation be received at the system input: "In the room, next to the table, a vacuum cleaner is operating". If a spatial structure of the situation is constructed directly from this description, it will have the following form: $(ar_{38}b)(aR_3c)$, where a is the vacuum cleaner, b the room, c the table, r_{38} is to be inside, R_3 is near. Nothing else can be obtained directly from the initial description. But using the scenario, shown on Figure 3.3, it is possible to supplement this description, for instance by the fact $(dr_{38}b)$ or the fact (dR_3c), where d is the cord of the vacuum cleaner.

§3.7. Causal Logic*

Above it was already said, that at the present time there does not exist anywhere, a nearly completed fragment of this logic, which has to operate with relationships that connect among each other the causes and the effects. Generally speaking, if causal logic is treated sufficiently broadly, it can be expanded almost to the theory of derivation, in which are examined not only the traditional deductive systems of derivation, but also derivations of the deductive type or derivations of the type "from the particular to the particular" (traductive). But we will not occupy ourselves with such an expansion, but shall consider in the framework of this paragraph a few fragments of causal logic in a narrow sense, which already now can be applied in control systems of the semiatic type.

First of all it is necessary to note, that until recently, even in the philosophical plan, there has not been conducted a sufficient analysis of the essence of the relation cause - effect. Most authors are inclined to thinking that there are at least five basic types of such relations.

1. Energy reason. The reason for variation of w, observed in some phenomenon or process Π_2, is the transfer of some energy v from a phenomenon or process Π_1. In this case it is possible to say that v is the cause of w, and w is the effect of v.

2. The cause for change of w in Π_2 is not itself the action of v. It rather plays the role of the "last impulse" after which in Π_2 there begins to develop a process that leads to w. However even in this case it is

possible to consider v as the cause (prime cause) for w, and w as the effect of v.

3. The cause of change w is the information v, in which is contained the directive for Π_2 on the character of the required change. In such a case it is also possible to think that between this directive v and the change w there exists the relation *cause - effect*.

4. The change w occurs because in Π_2 there are two subprocesses, or systems, which interact with each other and generate the change w itself. Such causes of changes can occur also as secondary ones after the appearance of the primary reasons, indicated in the preceding numbered remarks. We note, that the interaction inside a process can determine the development of the process itself, which expresses itself in a chain of its changes w_1, w_2, $\ldots w_n \ldots$

5. Acting as a cause can be some "fundamental" law, according to which any process tends toward certain steady states (for instance, any body that is under the effect of a gravitational field, tends "to fall").

There exists a rather stable concept of the cause of a certain phenomenon, as being the only one, the necessary and the sufficient cause. Even in erecting methods of derivation on the basis of induction (we shall encounter them in Chapter4), where this erroneous belief will play the role of a certain very important assumption, in the realization of which one or another method is effectively realized. In actuality the connection of causes and effects can turn out to be considerably finer. From this standpoint it is possible to give the following classification of the cause-effect relations.

1. Necessary and sufficient cause. The cause v in its actualization always brings about the effect w. And vice-versa, the appearance of w always witnesses of the preceding actualization of v. This is a rather rare case for a real situation. As its example may serve the connection between the loss of a support by a certain body and its fall.

2. Sufficient reason. This is the most frequently encountered case in actual practice, when the cause v always causes the effect w, but v does not always follow from the appearance of w. As an example of such a situation may serve the connection between the nonfullfillment of a plan and a lack of a bonus. But from an absence of a bonus there cannot follow the conclusion or the nonfullfillment of a plan.

3. Additional co-causes. Neithre v_1 nor v_2 are causes of w. Only their joint action causes w. A case is possible of not a pair of additional co-causes, but a certain multiplicity of them. For instance, if v_1 is the presence of pneumococcus and w is the infection with pnenmonia, then for an occurrence of w there is also necessary, in addition to v_2, the actualization of some co-cause, like chilling, a cold, or a tramma.

4. Necessary co-causes. In a list of co-causes, each of which do not lead to w, there may be co-causes the presence of which are necessary for causing w. For the preceding example such a necessary co-cause is the presence of the pneumococcus.

5. Possible co-causes. Causes for a bad accident may be an irregularity of some means of transportation, violation of road traffic rules, state of drunkenness of the driver, etc. None of these does necessarily lead to an accident. But an increase of multiplicity of similar simultaneously actualized co-causes increases the chances for an occurrance of w.

Into our classification of causes could also be included various pseudocauses, that often played a decision role at the intial stages of thought development. As an example of such a pseudocause can be the rule "after that, - means because of that".

In considering not a single cause-effect relation, but a chain of such relations it may turn out, that the action of some or other sequences of causes cumulates or alternates the effect, related to the necessity of appearance of w. In general, in cause-effect networks there may occur various fine effects, some of which we shall discuss immediately below; when we analyze scenarios related to causal relations.

Let us return to Example 3.2 and its related figs 3.1 and 3.6. A fundamental property of cause-effect relations is their antisymmetry and transitiveness, which allows for any phenomenon, which can play the role of an effect, to construct a tree of causes and co-causes capable of generating it. How can this tree be used? If for example, in a text entering the input of a system, there is explicitly contained the information, that an outsider supplier has upset the plan of furnishing the billets, than the text can be supplemented by the facts that there are not billets and the machine tool is idle (see Figure 3.1). The presence of a cause-effect scenario thus allows, in the presence of information about any apex of it, to complete the chains of facts that lead from it to effects. The only fine point here is contained in the necessity of accounting for the times of onset of events. If the fact "the supplier upset the plan of supplying billets" appeared at time moment t, then the facts "There are no billets" and "the machine tool is idle", which are same-moment facts, can themselves occur not at moment t, but at some moment t+τ, since the company may have a certain reserve of billets. And the size of τ may be sufficient so that facts-effects would not occur at all, if during time τ the reserve is not used up and the supplier corrects his negligence.

It is precisely here that are joined the temporal and the causal logics. In the future there may be realized not one sequence of events, but an entire fan of such sequences. And in a direction from causes to effects this must be taken into account. Hence cause-effect scenarios must be equipped by additional information (weights) about temporal delays and the onset of effects after immediate causes. With these weights it is possible to mark the arcs shown in this scenario, which gives th possibility to make derivations in caused logic more abundant.

Example 3.26. On Figure 3.22 is shown the same scenario, as in Figure 3.1. To it was only added the information on the time of occurrance of effects after the onset of their immediate causes.

Let there be, at moment t, information in the system that the machine tool is idle, and that at moment t - τ the information was received that the supplier ceased to supply the billets. If τ is smaller than 48 hours, then the conclusion that there are no billets and the machine tool is idle has not yet been realized.

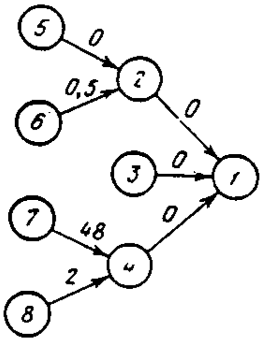


Figure 3.22

Derivation from causes to effects always uses the property of transitiveness. If there is information on two events, one of which may be considered as the cause, and the other as the effect, then all intermediate events are easily derived. Derivation from effects to causes however is more complex, for it is not single valued. In forming such rules of causal logic a great role is played by all the information available to the system for the given moment of time. If for example, the system knows that the machine tool is idle, it is not yet possible to reestablish the cause. But if there also is additional information that the machine tool is in good working order, and billets are available, then from the fact of idleness of the machine tool there will immediately follow the fact, that the worker is absent from his work-place.

In constructing cause-effect scenarios it is often possible to point out a number of relations that connect the causes among each other. For instance, it may happen, that some causes cannot exist at the same time. Thus, the effect "The tape recorder is not reproducing the recording" may occur because of two facts: "There is no tape in the recorder" and "the tape broke". These two facts-causes are mutually related. They cannot exist simultaneously. Therefore in deriving the cause for the effect "The tape recorder is not reproducing the recording" and the presence of additional information on the realization of the fact "The tape has been taken out of the recorder" the system must make the single valued deduction that the cause of the fact of interest to us consists in the absence of the tape in the recorder.

This, shown in Figure 3.6 is rather simple. But the scenario cited in Figure 3.8 contains much additional information, which can be fully used in the process of obtaining derivations both from causes to effects, as well as from effects to causes.

Apart from an indication of a direct connection of causes and effects, there is also contained in this scenario the information on those acting factors, which give rise to or stimulate the causes. Together with that, in this scenario, may be pointed out also inhibiting factors, which hinder the development of causes. And finally, in the scenario, may be pointed out certain observed regularities in the appearance or non-appearance of facts, which allows us to establish the causes and predict the development of the process from the causes to the effects.

Thus, in causal logic, there are more questions and unsolved problems than answers to them. And one of the most difficult questions, is the problem of obtaining diffuse derivations in causal logic, when the presence of causes is known with only a certain estimate of reliability, and sometimes there is even no complete certainty, that it is precisely the given facts that serve as causes of the fact that interests us. It is precisely this problem which plays a large role in all systems of control of objects of a complex nature, that we still occupy ourselves in the remaining part of the present paragraph, and in doing so we will assume that the hypothesis has already been worked out that $v_1, v_2, \ldots, v_m$ can be the causes of w. The methods of forming such hypotheses will be discussed in Chapter 4. We, however, will not consider how there are constructed special derivation rules that allow us to accomplish the transition from $v_1, v_2, \ldots, v_m$ to w, and still concentrate our attention on that case when all causes lead directly to w. For simplicity of presentation we still limit ourselves to the case when among v_2 there exist no connections which would forbid the occurrence of one cause in the presence of another. In the literature list for the present chapter sources are cited, in which this limitation is removed. And lastly, we shall assume, that w occurs, if all the v_i have been realized. The case when it is sufficient to have even just one cause v_i to realize w, does not introduce anything principally new and has also been analyzed in the studies given in the literature list.

In logical systems of deviation, wide use is made of schemes of the form

$$\frac{A_1, A_2, \ldots, A_n; \{A_1, A_2, \ldots, A_n\} \Rightarrow B}{B}.$$

Their meaning is interpreted in the following way: if $A_1, A_2, \ldots, A_n$ have already been derived, and from the totality of the derived formulas of $A_1, A_2, \ldots, A_n$ there follows the derivability of B, then B is considered as having been derived. This scheme generalizes the famous rule "modus ponens" which belongs to classical logic:

$$\frac{A, A \rightarrow B}{B}.$$

Here, the → symbol is treated as the usual implication.

Assume now that our knowledge, that the A_i has actually been derived, is incomplete. In other words, instead of that portion of the derivation scheme which is above the line, we have

$$\leftthreetimes_1 A_1,\ \leftthreetimes_2 A_2,\ \ldots,\ \leftthreetimes_n A_n;\ \{A_1,\ A_2,\ \ldots,\ A_n\} \Rightarrow B.$$

Here $\leftthreetimes_i$ are certain quantifiers, for example, quantifiers of the frequency type *almost never, very seldom, not often, not seldom, often, very often, almost always*[1] In this notation the causes, that give rise to the derivability of B, are given with some degree of assurance. What can be said in these conditions about the derivability of B?

Also our knowledge may be incomplete not on the causes A_i, but about the very fact of derivability from these causes. In that case the upper part of the scheme of the deductive derivation has the following form:

$$A_1,\ A_2,\ \ldots,\ A_n;\ \leftthreetimes[\{A_1,\ A_2,\ \ldots,\ A_n\} \Rightarrow B].$$

As before, the quantifier $\leftthreetimes$ can be treated here in different ways. As long at it would reflect in some manner the degree of assurance to the effect that the derivation from $\{A_i\}$ is realized in B. Of course, a case is also possible when there takes place the one and the other uncertainty.

The described situation occurs very often in practice. In conditions of situational control it turns out to be more often the rule than the exception. Knowledge about the object of control and of methods of controlling it is collected gradually, and the necessity of using *indistinct schemes of derivation*, examples of which we have just examined, remains on almost the entire period of functioning of the control system that is constructed on the basis of semiatic models.

Various methods of operation with such indistinct rules of derivation are possible. We will illustrate this with several approaches, and for greater clarity of presentation we shall analyze only the case, when the derivation scheme has such a form of the production $\leftthreetimes_A A, \leftthreetimes[A \Rightarrow B]$, where $\leftthreetimes_A$ and $\leftthreetimes$ are quantifiers of a certain type that reflect the doubt in the truth of the premise A and of the production itself. We are interested in the value of the quantifier $\leftthreetimes_B$ with which we have to enter the fact B into the memory of

---

[1] Other quantifiers are also possible. For instance, *almost certain, with some degree of certainty, fifty-fifty, doubtful, etc.*

the system or to use it in a subsequent derivation or in making a final decision.

Consider first the case, when instead of the quantifiers there are certain probabilities reflecting our knowledge about the influence of A on the occurrence of B and the appearance of A itself. Such probabilities may be either of the apriori or aposterion type, accumulated after some experience of system operation. In that case it is possible to propose two schemes of absorbing a probability estimate for B, which more or less correspond to man's intuition:

$$\frac{\begin{array}{l} P_A \geqslant a \\ P_{A \Rightarrow B} \geqslant b \end{array}}{P_B \geqslant \max(0,\ a+b-1)}; \quad \frac{\begin{array}{l} P_A \geqslant a \\ P_{A \Rightarrow B} \geqslant b \end{array}}{P_n \geqslant ab}.$$

In the relationships cited only the estimates from below are indicated. By how much they may be exceeded, remains unclear. Some information about this can be extracted from the estimated for the rule "modus tollens" also well known from the classical logic:

$$\frac{B,\ A \rightarrow B}{A}$$

The probabilistic analog of this rule has the form

$$\frac{\begin{array}{l} P_{A \Rightarrow B} \geqslant a \\ P_B \leqslant b \end{array}}{P_A \leqslant \min(1,\ 1-a+b)}; \quad \frac{\begin{array}{l} P_{A \Rightarrow B} \geqslant a \\ P_n \leqslant b \end{array}}{P_A \leqslant \min(1,\ b/a),\ a \neq 0}.$$

Letters P in these schemes designate everywhere the corresponding probabilities.

However, as was said already, in semiatic control systems, and in particular, in situational system of control, accumulation of statistical data proceeds too slowly, and apriori knowledge on probabilistic estimates is absent. Therefore, it is of interest to form an estimate of the certainty of B, remaining within the framework of utilizing the quantifiers that estimate the certainties of A, A ⇒ B and B.

Let us specify as such quantifiers the frequency quantifiers, which we have just cited adding to their list two further quantifiers *never* and *always*, which bound it on the left and on the right. It is for these quantifiers that we shall construct the rules of estimation of B.

Example 3.27. Consider the following scheme of reasoning:

During mooring, the ship "A. Pushkin" looses 20-30 minutes of time once in a while; if more than 20 minutes are lost, the loading plan must be made over very often. Let us denote by A the statement "The mooring of the ship 'A. Pushkin' leads to losses of 20-30 minutes of time", denote by $\nsucc_A$ the quantifier *seldom*, by B the statement "The loading plan is being made over", and by $\nsucc$ the quantifier *very often*. Then the text above can be represented by the scheme $\nsucc_A A; \nsucc [A \Rightarrow B]$. What can be said about the estimate of the truth of B, if it is a known that at the given moment the ship "A. Pushkin" is getting ready to moor? A more or less experienced dispatcher will, in the presence of such information, most likely get ready to the procedure of changing the plan of loading.

In §2.11 and in the preceding paragraph we had already stated that the quantifiers could be juxtaposed by functions of affiliation or by intervals on the segment [0,1]. In particular, in Figures 2.16 and 2.17 were shown these intervals and the form of the application functions just exactly for the case of frequency quantifiers. Then the problem can be set up in the following manner.

From the known quantifiers $\nsucc_A$ and $\nsucc$ we construct (or use the earlier constructed) functions of application (or intervals) on the segment [0,1] and we look for the mapping of a pair of these functions (or intervals) onto the function of application (or interval) for the quantifier $\nsucc_B$. Thus, we are interested in the form of the function F which specifies the relation $\mu_b = F\ (\mu_a\ \mu)$. Here μ_a, μ, and μ_B are affiliation functions corresponding to $\nsucc_A, \nsucc$, and $\nsucc_B$. Before determining the form of function F, let us formulate certain sufficiently obvious properties, which it must satisfy. If we select as quantifiers the quantifiers *always*, then the upper part of the derivation scheme will have the form A; $A \Rightarrow B$. In other words, the derivation rule will become a "modus ponens". This means that the quantifier $\nsucc_B$ also must become the quantifier *always*. For functions of application this is equivalent to the requirement of fulfilling of the equality $F(\mu_A^1, \mu^1) = \mu_B^1$, where μ_A^1, μ^1, μ_B^1 are functions of application corresponding to the quantifier *always*. For the case when $\nsucc_A$ is the quantifier *always*, and $\nsucc$ - *never*, it is obvious, that $\nsucc_B$ must also assume the meaning *never*. In other words, there must be fulfilled the equality $F(\mu_A^1,\mu^\circ)=\mu_B^\circ$, where μ° and μ_B° are application functions corresponding to the quantifier *never*.

The requirements formulated by us for the application functions μ° and μ^1 limit the form of these functions. They can be specified in only two ways:

$$1.\ \begin{cases} \mu^1(x)=0, & x\in[0,1), \\ \mu^1(x)=1, & x=1, \\ \mu^0(x)=0, & x\in(0,1], \\ \mu^0(x)=1, & x=0. \end{cases} \qquad 2.\ \begin{cases} \mu^1(x)=x, & x\in[0,1], \\ \mu^0(x)=1-x, & x\in[0,1]. \end{cases}$$

Let us consider one of the possible ways of determining F, under which are satisfied the requirements formulated by us[1]. This method reduces to obtaining μ_B from μ_A by the superposition $\mu_B(x)=\mu_A[\mu(x)]$. This operation is correct, since for the application functions the region of definition and the region of values in the segment [0,1]. It is easy to verify, that with any of the two possible methods of defining the application functions for the quantifiers *always* and *never* our two requirements are fulfilled in the case of superposition.

Let now μ_A and μ be arbitrary functions corresponding to certain values of the frequency quantifiers. In Figure 3.23 are shown two single-hump curves satisfying these functions. The result of their superposition is a two-hump curve shown by a dashed line. What is its meaning? If, for instance, μ_A is *seldom*, and μ is *often*, then the maxima of their superposition characterize those values from the segment [0,1] which are *seldom* encountered in the diffuse subset of the affiliation function *often*. If however, for instance, μ_A and μ correspond to the quantifier *often*, then the maxima

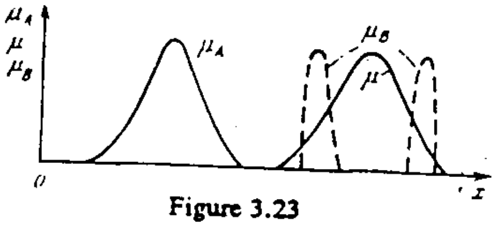


Figure 3.23

---

[1] It was proposed by I. V. Yezhkova.

correspond to those values, which are *often* encountered in the fuzzy subset of the application function *often*. Inasmuch as $\succ\!\!|$ characterizes the set of possible values of the frequency of occurrence of the fact B, specified by all possible values of the frequency of occurrence of the fact A, the superposition selects from this set only those values of the frequency of occurrence of fact B whose application with this set is characterized by the quantifier $\succ\!\!|_A$. In other words, the estimate specified by the quantifier $\succ\!\!|_B$ selects from the set of values specified by quantifier $\succ\!\!|$, those values which were encountered in the set specified by $\succ\!\!|$, with the frequency characterized by $\succ\!\!|_A$. Such an interpretation of the superposition operation is sufficiently "human", which may serve as an argument in favor of adopting such a method of defining the function F.

The only drawback of the suggested approach for defining function F is that the superposition generates affiliation functions the number of humps in which may increase on repeated application of this operation. But, as we have shown already in §2.11 and later in §3.6, the application functions corresponding to the various quantifiers used by people, are either monotonic on the segment [0,1], or are unimodal. The appearance of multihump functions renders a verbal interpretation of the results sufficiently complex. In natural language it is simply impossible to find lexemes or combinations of lexemes corresponding to them. However, this is not so bad, if all the results, related to the estimate of $\succ\!\!|_B$ during a multistep derivation by indistinct derivation schemes, will be only intrasystemic. In delivering the final result to the user of the system or for making decisions on control, it is possible to use a special procedure of "approximation" of multihump functions by the same words, by which are characterized the initial frequency quantifiers with some verbal additions (for instance, *near and often*, or between seldom and not seldom, not often).

Example 3.28. Let us return to Example 3.27. In Figure 3.24 is shown, by a dashed curve, the result of a superposition of two application

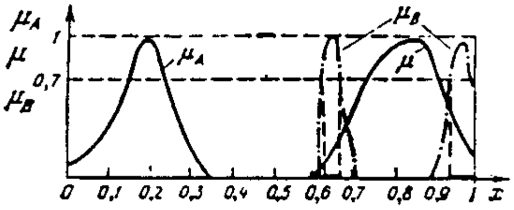


Figure 3.24

functions that correspond to those quantifiers, that are given in that example. With the aid of a level of cut-off with a value of 0.7 are obtained indistinct intervals on the abscissa axis. Now we can say, that the dispatcher must expect a change of the plan of loading with a certainty not less than 0.7 with values inside the intervals (0.62; 0.67) and (0.95: 1).

If, however, a level of cut-off equal to unity is taken, then the dispatcher can expect the change in plan in the intervals (0.6; 0.7) and (0.9; 1) with those indices of certainty that are given by function μ_B. In one and in the other case we find, that with availability of the fact (information) on the beginning of mooring of the ship "A. Pushkin" an experienced dispatcher who begins to get ready (at least morally) to a change in the loading plan, acts with reason.

We have considered only the simplest case of the rule of derivation in the case of indistinct frequency quantors. In the literature on this problem, given in the commentary to this chapter, are formulated the requirements for constructing $\nsucc_B$ in the case of n premises A_i, from which follows the statement B (premises mutually dependent or independent in the sense of influencing the appearance of each other), and also the obtainment of estimates of type $\nsucc B_1B_2$ or $\nsucc B_1 \vee B_2$, when from a group of premises is deduced the conjunction of two facts or their disfunction. A consideration of such sufficiently cumbersome constructions goes outside the framework of the present book.

Another method of determining function F, different from the method of superposition, consists in noting that upon application of some one quantifier to another quantifier there occurs a certain monotoic transformation of the original function of affiliation, which amounts to an extension and a shift of the maximum of the function in one or another direction.

The advantage of such a method of determining F is that during monotonic transformation the form of an application function does not change cardinally. Its unimodality or monotony is retained, and the transfer from a new form of the function to a verbal estimate, corresponding to some quantifier, occurs much more simply, than from multihump functions that result from superposition operation.

Example 3.25. In Figure 3.25 are shown two results, obtained by superposition and by a shift with extension, for the case when $\nmid_A$ and $\nmid$ correspond to the quantifier *often*. The difference appears to be that the superposition separates out in the affiliation function *often* those values which are frequently encountered. In the case of a shift and extension, however, we can interpret the result as the appearance of a new quantifier with the value *often-often*, which, if desired, could be approximated, for instance by the value *very often*.

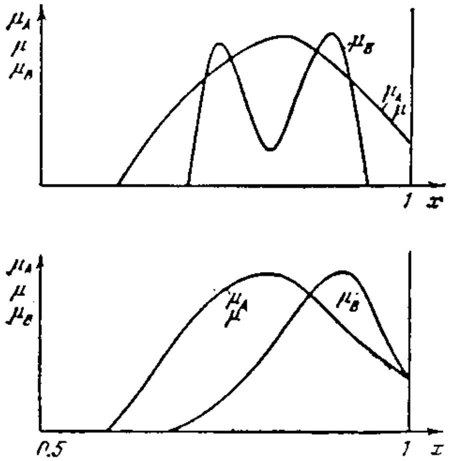


Figure 3.25

A drawback of the second approach for determining the form of function F is that neither a shift of the affiliation function for $\nmid$, nor the character of its extension that are specified by the affiliation further corresponding to $\nmid_A$, have not been defined in any way. Needed are some special considerations about these parameters. But at the present time such considerations are lacking.

In the conclusion of the present paragraph let us examine a certain new method of obtaining indistinct deviation (or conclusions), successfully developed recently by S. V. Chesnokov. The results cited below, with the possible exception of their interpretation, have been taken from the studies of that author, as given in the literature list for the present chaper.

On Figure 3.26 is shown a certain situation, issuing from which it is possible to assert, that some a_i are b, where i=1, 2, ..., n. Consider a notation of the form $a_i \mapsto b$, which we shall call a *determination*. Let us introduce two important concepts. The magnitude of the relative probability $P(b/a_i)$ we shall call the *Exactness of the determination*. The magnitude of the relative probability $P(a_i/b)$ we shall call the *Completness of the determination* $a_i \mapsto b$. While the exactness of the determination estimates the degree of certainty that the appearance of a certain element from A leads to the appearance of b, the completness of the determination estimates the truth of the statement that the appearance of b witnesses the presence of at least one element of A. Henceforth, instead of the multitude of causes in the left side of a determination, we shall consider only one cause, which we shall designate by a.

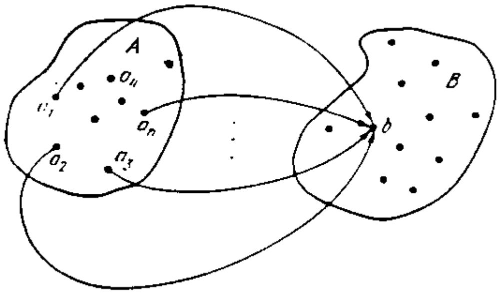


Figure 3.26

Example 3.30. Let, for example, the Table 3.9 be obtained on the basis of some observations.

Table 3.9

	a Is present	a Is not present
b Is present	60	590
b Is not present	140	210

Each of its cells, in correspondence with the definition of exactness and completness of determination, denoted by leters I and C, respectively, allows us to determine these characteristics of determination. For our example we have

$$I\,(a \longmapsto b) = P\,(b/a) = \frac{140}{140+60} = 0{,}7,$$
$$C\,(a \longmapsto b) = P\,(a/b) = \frac{140}{140+210} = 0{,}4.$$

For the inverse determination b $\mapsto$ a the exactness and the completness are obtained by repositioning of the values of exactness and completness for the direct determination. In other words I(b6$\mapsto$ a) = 0,4, a C(b$\mapsto$ a) = 0,7.

Example 3.31. Let a-denote the presence of limestone in a given place, c- the presence of Karst funnels and caves in the same place, b- the presence of underground streams and lakes. Negation of these propositional variables corresponds to the absence of the respective phenomenon. Let us consider Tables 3.10-3.12.

Table 3.10

	a Is present	a Is absent
b Is present	100	100
b Is absent	100	100

Table 3.11

	ac Is present	$\overline{ac}$ Is present	$\overline{ac}$ Is present	$\overline{ac}$ Is present
b Is present	100	40	0	60
b Is absent	0	60	100	40

Table 3.12

	c Is present	c Is absent
b Is present	140	60
b Is absent	60	140

What do these tables show? Table 3.10 shows, that on the basis of some statistic, reflected in it, it is possible to state that in half of the cases in places where there is limestone, there are underground streams and lakes. Table 3.11 tells us that in the presence of the Karst phenomena in the limestone there is always a certainty of the existence of underground streams and lakes in that location. Even in the case, when the Karst phenomena (more precisely, their analogs) occur not in a pure limestone, then in 40 cases out of 400 there exist both the underground streams and the lakes. But in a Karst stratum in which no Karst phenomena are found, it is useless to look for underground streams and lakes. Now, if they are absent in places where instead of limestone other rock formations are observed, then in 40 cases out of 400 it is also possible to give a positive answer on the presence here of underground streams and lakes.

As it follows from these tables, $I(a \mapsto b) = 0{,}5$, a $I(ac \mapsto b) = 1$. Thus, the presence of Karst phenomena in the case of limestone formations makes more exact the determination with the aid of which is established the presence of underground streams or lakes in the given place. If one considers $ac \mapsto b$, then, as it follows from Table 3.11, the exactness diminishes down to zero. Hence, an additional conjunctive premise in a determination can both increase its exactness, as well as decrease it. In a similar manner it is possible to analyze the case when in passing from a to b, a disjunction is used instead of a conjunction, i.e., a transition is made to a determination of the form $(a \vee c) \mapsto b$.

Intuitively it is clear, that the determinations are closely connected with the indistinct deviations, when the quantifiers of frequency are estimated by statistical probabilities. This is actually so. We shall illustrate their connection on an example of generalizaton of the syllogism concept, which is a classical object in logic. As it is known, Aristoteles proposed 19 basic modi of syllogisms, which always give correct results. In addition, there are also five other modi, which represent weakened basic modi. For instance, for the modus (all a are b; all b are e)⇒(all a are c) the weakened modus has the form (all a are b); (all b are c)⇒(some a are c). The other modi of the four possible figures of Aristoteles, besides these 24, are not true in a strict sense. Determinations allow us to expand the multitude of modi, operating with which becomes possible by introducing into them frequency quantifiers.

Let us have the determinations $a \mapsto c$. These determinations are characterized by their exactnesses and completness. Let us denote them for simplicity of notation in the following manner: $I(a \mapsto b) = i \in L^1$; $C(a \mapsto b) = k \in M^1$; $I(b \mapsto c) = l \in L^2$; $C(b \mapsto c) = m \in M^2$. Here L^1, M^1, L^2 and M^2 are some concrete values of intervals from the segment [0,1]. We are interested, what can be said, with known i, k, l and m about the exactness and completeness of the determination $c \mapsto a$. These characteristics of the determination $c \mapsto a$ we will denote, respectively, as r and s, and the intervals in which these value lie, as L^3 and M^3.

If we return to the consideration of Aristoteles' syllogisms and take that modus which we had just looked at, then for it $L^1 = L^2 = L^3 = 1$. About M^1, M^2, and M^3 nothing can be said. They could be any intervals in the half-interval [0,1].

The characteristics of the three determinations, introduced by us, which determine the syllogism, are calculated by standard formulas on the basis of Table 3.13. In accord with this table the quantity i, for example, is equal to

$$i = \frac{x_1 + x_3}{x_1 + x_2 + x_5 + x_7}$$

Analogically are determined also the other characteristics of the determinations that form the syllogism.

Table 3.13

a	x_1	x_2	x_3	x_4
$\bar{a}$	x_5	x_6	x_7	x_8
	b	$\bar{b}$	b	$\bar{b}$
	c		$\bar{c}$	

Let us pose ourselves the problem of searching for syllogisms, in which would participate frequency quantifiers, and let the latter correspond to relationships related to the exactness and completeness of the determinations which enter into the syllogism.

Example 3.32. As an example of such asyllogism let us examine the reasoning:

Among bus drivers there are few violators of street traffic
Among women there are few violators of traffic
Almost all bus drivers are women.

Apparently, readers will not agree with the truth of the obtained deduction. But if it is assumed, that these reasonings refer to war-time, then in a situation, where the class of bus drivers is almost included into the class of women, such an assertion could turn out to be true. This means that the whole matter consists in how are interrelated these classes, which is determined by the characteristics of determination introduced by us and forming the modus of the syllogism.

The problem of constructing syllogisms satisfying some earlier prescribed requirements that take into account their validity, can be formulated in a completely strict form. Let us return to Table 3.13. We shall consider the x_i as unknowns. Let us write out a selection of limitations, which these unknowns must satisfy.

The first limitation is trivial:

$$(1)\ x_1 + x_2 + x_3 + x_4 + x_5 + x_6 + x_7 + x_8 = 1.$$

The next six limitations, which might be called limitations on the "context" of the syllogism, follows from the fact that the frequencies with which are encountered a, b, and c, must lie within certain prescribed limits:

$$(2,3)\ 0 < g_a \le x_1 + x_2 + x_3 + x_4 + \le h_a \le 1,$$
$$(4,5)\ 0 < g_b \le x_1 + x_3 + x_5 + x_7 + \le h_b \le 1,$$
$$(6,7)\ 0 < g_c \le x_1 + x_2 + x_5 + x_6 + \le h_c \le 1.$$

And finally, there are eight more limitations related to the requirements of satisfying certain limitations of exactness and completeness of the determinations entering into the premises of the syllogism.

$$(8, 9) \quad i_0 \leqslant \frac{x_1 + x_3}{x_1 + x_2 + x_3 + x_4} \leqslant i_1,$$

$$(10, 11) \quad k_0 \leqslant \frac{x_1 + x_3}{x_1 + x_3 + x_5 + x_7} \leqslant k_1,$$

$$(12, 13) \quad l_0 \leqslant \frac{x_1 + x_5}{x_1 + x_3 + x_5 + x_7} \leqslant l_1,$$

$$(14, 15) \quad m_0 \leqslant \frac{x_1 + x_5}{x_1 + x_2 + x_5 + x_6} \leqslant m_1.$$

The limitations (8,9) and (10,11) specify the conditions imposed on the exactness and completeness of the determination a |→ b, while the remaining limitations specify the conditions on the exactness and completeness of the determination b |→ c.

The fourteen parameters in these limitations (upper and lower boundaries in the inequalities) provide rather wide possibilities. Specifying them determines the form of those frequency quantifiers, which are used in the premises of the syllogism. We however, are interested in the values of the four limitations for exactness and completeness of determination c |→ a, which forms the conclusion of the syllogism

$$r_0 \leq \frac{x_1 + x_2}{x_1 + x_2 + x_5 + x_6} \leq r_1,$$

$$S_0 \leq \frac{x_1 + x_2}{x_1 + x_2 + x_3 + x_4} \leq S_1.$$

Then, with given limitations and such x_i, that all 15 limitations are fulfilled, it is possible to state, that the exactness and completeness of the determinations c |→ a, which we can denote by z and s, satisfy the limitations $r_0 < r \leq r_1$ and $S_0 \leq S \leq S_1$. The solution of the formulated problem can be reduced to a problem of piece-wise linear programming. The result of its solution will be the finding of such regions in the x_i space, in which will be realized the derivation by means of a syllogism with presented fixed values of the quantifiers. To do this it is only still necessary to juxtappose to the verbal estimates of these quantifiers certain segments or intervals or the segment [0,1].

Our formulation of the problems can be simplified. Since from limitations 2, 4, and 6 there follows a strict positiveness of the demominators in limitations 8 - 15, all limitations can be rewritten in a linear form. Moreover, if it is noted, that P(uv) = P(u/v)P(v) and change over from the relative frequencies to the frequencies P(ab), P(ac) and P(cb), then also the goal functions in an problem can be expressed in a linear form (for this, it will be necessary to change the meaning of the limitations. Then the problem of finding a generalized (determinational) syllogism will be reduced to a standard problem of linear programming.

For a qualitative analysis of what results from solving such a problem, let us consider the case, when we are interested only in the exactness of the respective determinations. Furthermore, we assume that there exist the following uniform limitations on the exactness of the determinations (they are denoted by corresponding relative probablities) and on the context of the syllogism:

$$\begin{gathered} x \leq P(b/a) \leq 1, \\ x \leq P(c/b) \leq 1, \\ 0 < v \leq P(a) \leq 1, \\ 0 < v \leq P(b) \leq 1, \\ 0 < v \leq P(c) \leq 1, \end{gathered}$$

where x and v are the so-far unknown limitations on the frequencies.

With the inequalities and the equality (1) it is possible to reduce the original problem to a problem of the form: find x_1, x_2, x_3, x_4, x_5, x_6, x_7, and x_8, which satisfy the limitation (1) and the system of limitations

$$\begin{cases} (1-\varkappa)x_1 + x_2 + (1-\varkappa)x_3 + x_4 \geqslant 0, \\ (1-\varkappa)x_1 + x_3 + (1-\varkappa)x_5 + x_7 \geqslant 0, \\ x_1 + x_2 + x_3 + x_4 \geqslant v, \\ x_1 + x_3 + x_5 + x_7 \geqslant v, \\ x_1 + x_2 + x_5 + x_6 \geqslant v. \end{cases}$$

Furthermore, all x_i must be equal to or greater than zero (this condition together with relation (1) guarantees the fulfillment of the upper limits for all five probabilities) while P(c/a) must lie within given limits [8,1].

If the solution of the posed problem is formed, with x and v as some parameters, it will have a form that is different for the three regions into which is divided a unit square, specified by the relations $0 \leq x \leq 1$ and $0 < V \leq 1$:

$$\delta = \begin{cases} 0 \text{ for region } 1, \\ 2 - 1/v \text{ for region } 2, \\ 1 - (1 - x)(x + 1/v) \text{ for region } 3. \end{cases}$$

The numbered regions are shown on Figure 3.27.

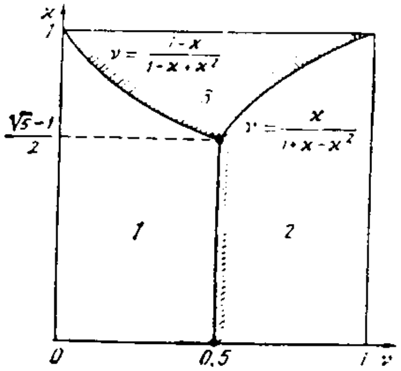


Figure 3.27

Their boundaries are defined by the following relationship:

$$\text{Region } 1 \quad \begin{cases} v \leqslant 0.5 \\ v \leqslant \frac{1-x}{1-x+x^2}; \end{cases}$$

$$\text{Region } 2 \quad \begin{cases} v > 0.5 \\ v \geqslant \frac{x}{1+x-x^2}; \end{cases}$$

$$\text{Region } 3 \quad \begin{cases} v < \frac{x}{1+x-x^2} \\ v > \frac{1-x}{1-x+x^2}. \end{cases}$$

Let us examine these regions[1]. In Region 1 is fulfilled the inequality $0 \leq P(c/a) \leq 1$. It is tautological, since it is fulfilled always and does in no way depend on the exactness of the determinations that form the premises. For this it is sufficient to have the fact of existence of the frequency P(c/a) itself, which always takes place in our square. This means that Region 1 does not offer us anything interesting. In Region 2 is fulfilled the inequality 2, $1/V$ $P(c/a) \leq 1$. Let us examine only those properties which we denoted by a and c, and about b we will not have any information. Then the inequality for P(c/a) in the given region will still take place only because for P(a) and P(c) the magnitude of V exceeds 0.5. Thus, Region 2, although not tautological as is Region 1, does not contain syllogisms proper. Statements

---

[1] Hatching indicates an open boundary.

from this region might be called, for example, trivial truths. For Region 3 is fulfilled the inequality

$$1 - (1-x)(x+1/v) \leq P(c/a) \leq 1.$$

The presence of x in this inequality relates the exactness of the concluding determination with the exactnesses of determinations entering into the premises. This is the region of generalized syllogisms. If, for instance, $x=1$, then $P(b/a)=1$, $P(c/b) = 1$ and $P(c/a) = 1$. We have obtained the well-known aristotelian syllogism, commonly called *Bazbaza*. It is interesting, that for $x=1$ the values of v turn out to be insignificant, i.e. the character of the context of the syllogism. This means, that the condition $v=2$, which is always understood in the classical interpretation of Aristoteles' syllogisms, is not obligatory for the modi of the Bazbaza type. If one renounces the equality of limitations of v in the context, then it is possible to obtain also other expansion of the syllogisms. We note, that the value of v is easily related with the quantifiers of the estimate of frequency of the type *often*, *very often*, *almost always*. And if the right limitations in the context are further modi different and not equal to unity, then the context will determine the frequency quantifiers with which the premises of the syllogism enter into the scheme of derivation.

Example 3.33. Let $0<x<\epsilon$ and $0<v\leq\epsilon$, where ϵ is a small number. Then, according to Aristoteles, such parameters will yield false syllogisms. On Figure 3.28 are shown regions, constructed in the square (x,v) similar to

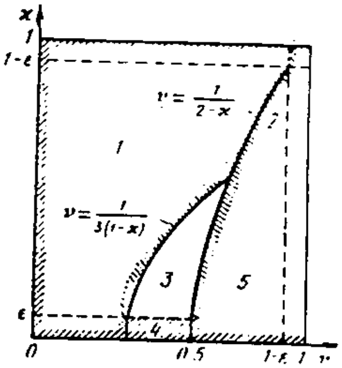


Figure 3.28

the ones given in Figure 3.27. Region 1 is the region of falsehood of the syllogism. Region 2, shown by a heavy line, is the region of trivial truth, similar to Region 2 of Figure 3.27. In Regions 3 and 4 the syllogism exists, while in Region 5 it does not. Open boundaries are indicated by hatching.

Here is an example of a syllogism contained in the region of existence of syllogism, shown in Figure 3.28.

Among fanciers of cigars women are seldom encountered
<u>Among women political workers are seldom found</u>
Among cigar fanciers political workers are not seldom encountered

In conclusion, let us pause on one more class of derivation rules of the indistinct type. Such rules, apparently, do not have a direct relationship specifically to causal logic, but by their ideology they are loosely connected to the rules examined by us here. This group of rules is connected to the *reasonings by analogy*, which play an enormous role for man and find wider application in semiformal explanation of the technology of control and of decision making methods in conditions of alternative selection.

First of all arises the question, what specifically should be considered an analogy. Let us examine the following example[1].

Example 3.34. Let there be two texts describing two situations. The first text: "Romeo loves Juliet. Juliet loves Romeo. Romeo is a man. He is an Italian. Juliet is a woman. She is beautiful. She is unmarried". The second text: "Tristan loves Isolade. Tristan is a Breton. Isolade is a woman. She is beautiful. She is married to King Mark". On Figure 3.29 are shown graphs corresponding to the cited texts.

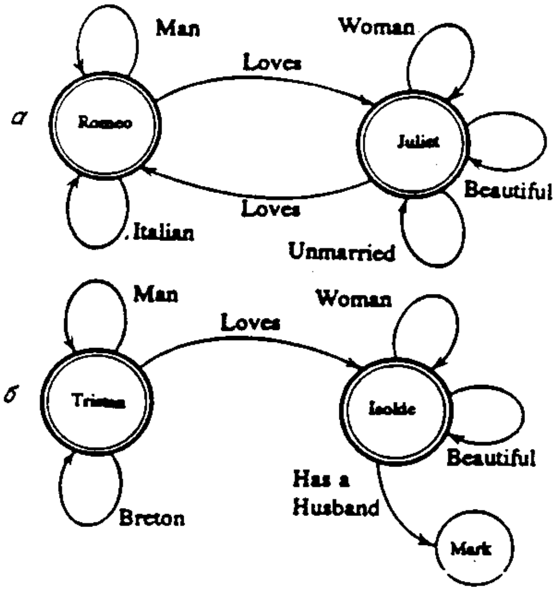


Figure 3.29

---

[1] This example was examined in the report of A. Farreni and A. Prade at the European Conference on Artificial Intelligence, at Orsee in 1982.

From the texts themselves and from Figure 3.29 it follows (from the standpoint of human intuition), that there is much in common between the two situations. But in what does this manifest itself? What is important in the descriptions for establishing an analogy, and what is not so important? Is it important, that Romeo and Tristan are men? Does their nationality play a role? Is the distinction ponderable that Isolade has a husband while Juliet does not? Answers to these questions depend on how will be defined the concept of analogy itself.

At the present moment, all investigators agree, that one of the models which could be used in exposing the essence of the reasoning by analogy and constructing corresponding derivation rules, may be the predicate of analogy introduced already by Leibnitz. This predicate is described by a diagram very similar to the diagram of modeling mentioned in §2.11:

$$\begin{tikzcd}
A \arrow[r, Rightarrow] \arrow[d, squiggly, "V"'] \arrow[dr, phantom, "T"] & A' \\
B \arrow[r, Rightarrow, "Q"'] & B' \arrow[u, squiggly, "W"']
\end{tikzcd}$$

Let us examine a class of objects X, in which is realized the derivation from object A to object A', and another class of objects Y, where is realized the derivation from object B to object B', and finally let them be established some correspondence between the objects entering into X and Y (for instance let there be a mutually singe-valued correspondence between X and Y). Then, if according to this correspondence A is mapped in B, then the requirement of commutativity of the diagram must be fulfilled. In other words, there must exist a mapping W, which is so designed, that B' is mapped into the same element of A' which is derived from A by means of T, if B' is obtained by derivation of Q from B. Let us recall, that for the diagram of modeling also the requirement of commutativity was necessary. And this coincidence is not accidental. Both the diagram cited by us, illustrating the essence of analogy according to Leibnitz, as well as the diagram characterizing the relation of modeling, are in essence alike. The introduction of an indistinct modeling relation is one of the possible methods of introducing analogies. But they can be introduced even without the apparatus of the affiliation functions, staying within the framework of the usual algebraic systems. This possibility is illustrated by the following simple example[1].

Example 3.35. On Figure 3.30 are shown three elements A, A' and B, which enter into the diagram for the predicate of analogy. The question arises, what picture should be in the place of B' to preserve the commatativity of the diagram. For answering it, it is necessary to construct the mappings of V and W. But before doing it, it is necessary to describe the pictures themselves, having selected a suitable language. For us it is natural to choose as such language the language of situational control, described is Chapter 2. Borrowing from Table 2.1 the notation for the relationships, we will introduce the following designation: a_1 - small unpainted circle, a_2 - larger circle, a_3 - rectangle, b_1 - cross, b_2 - triangle, b_3 - small black circle, r_{40} - *to be alone*, r_{41} - *to be below*, r_{43} - *to be inside*.

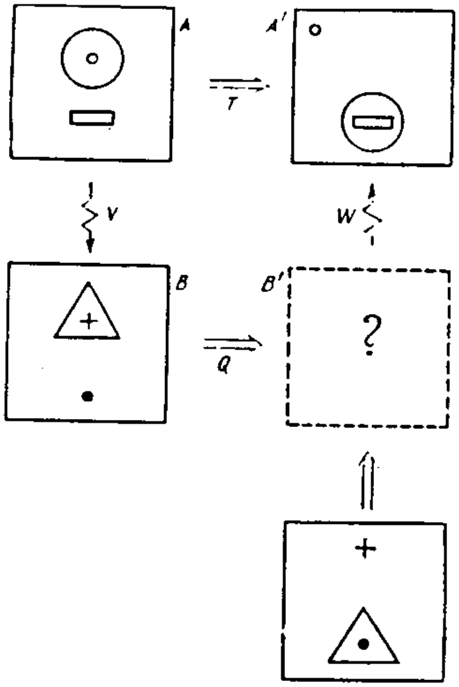


Figure 3.30

---

[1] This example is cited in the report of D. Pëtcke, represented in the materials of the same European Conference on Artificial Intelligence, that took place at Orsee in 1982.

Let us describe in these terms the pictures corresponding to A, A' and B.

$$A: (a_2 r_{40} a_3)(a_1 r_{43} a_2);$$
$$A': (a_3 r_{43} a_2)(a_2 r_{41} a_1);$$
$$B: (b_2 r_{40} b_3)(b_1 r_{43} b_2).$$

Now let us introduce the correspondence V; we specify it as the isomorphism: $a_1 \leftrightarrow b_1$, $a_2 \leftrightarrow b_2$, $a_3 \leftrightarrow b_3$. Since we have established a relationship of isomorphism, then W is single valuedly determined together with V. Then, for passing from description A' to description B' it is sufficient to make the needed substitution that follows from the given isomorphism. Having done so, we obtain B': $(b_3 r_{43} b_2)(b_2 r_{41} b_1)$. To this description corresponds the picture shown in the right lower corner of Figure 3.30. It is it, then, that must be placed as B' into the original drawing.

Let us make a few important remarks. The procedure of establishing correspondence of V and W depends on the selected method of description of A, A' and B. Imagine, for instance, that A' is described not as we had done it, but in the following form: A': $(a_1 r_{40} a_2)(a_3 r_{43} a_2)$. Then a comparative analysis of the descriptions does not give the possibility to find any kind of homomorphism among the given descriptions. As a second remark let us point out that for a description it is necessary to choose such relationships for which would exist derivation rules which translate combinations of one kind of relation into combinations of other relations. In one case it was possible due to the avoidabilty of a derivation rule related to the spatial logic of mutual disposition of objects, which has the form: $(a_2 r_{40} a_3)(a_1 r_{43} a_2) \Rightarrow (a_1 r_{40} a_3)$.

And the final remark. If T and V are established and are relationships of isomorphism, then Q and W always exist are also relationships of isomorphism.

An analysis of the cited example and the remarks made after it show that to-date, great difficulties still exist on the path of practical utilization of rules that reflect reasonings by analogy. This concerns especially the selection of a suitable language for describing of the sets of objects, that enter into the analogy diagram, and the development of powerful means of transforming descriptions, one into another. In Chapter 4 we shall once more return to the rules of reasoning by analogy in connection with the description of a special method that allows to generate hypothesis about the presence or absence of cause-effect relations among phenomena and to extend these hypotheses to formerly unknown cases.

Concluding the paragraph, devoted to causal logic, one must note, that compared to temporal and spatial logic, these logics have so far been developed very superficially. Great efforts are still needed for their completion. Particularly little has been done in the area of analysis of mutual influence of derivations in causal logic or derivations in temporal and spatial logics (on Figure 3.1) these connections are marked by strength arrows, passing between the respective logics, and also in the area of influence of the usual logics or the making of decision in the logics of action. About this last influence we shall yet speak in following chapters, where action logic will be discussed.

§ 3.8 Incompleteness and "Absurdity" of Knowledge

The problem of contradictiveness of knowledge, stored in a control system, was already touched upon by us in §1.6. During the operation of a system of complementation of knowledge, especially during utilization of derivation rules, which do not have an absolute character of the truth of the obtained statements, the appearance in the memory of the control system of contradictory descriptions of the object of control, processes occurring in it, or the very procedure of control is not quite so unbelievable. Therefore we will discuss here some of the reasons of the appearance of "absurd" knowledge and methods of its elimination.

The source of annoyances, which may occur in the system memory, is concealed in the incompleteness of information stored in the knowledge base. If some information is absent in the base, then it is completely unclear, whether the absent fact is true (and, for example, contradictory to some fact which is already in the memory) or is false. Because of incomplete falling-out of the memory, the absence of some statement in it says nothing about its truth or falsehood, possibility or impossibility of its occurrence in the memory. In the well-known logic system Nav'ya-n'yaya of Middle-Age India are distinguished four kinds of absence: absence due to non-appearance of it to the present moment of time, absence of something because by the present moment of time it has already disappeared, absence due to an impossibility of the appearance of it ever, neither in the past, nor at the present, and absence because at the present moment there is present something with which the appearance of that which interests us is incompatible. All four forms of absence in the Nav'ya-nyaya are strictly distinguished, and operation with them takes place in a different manner. Something similar is observed also in models bases of knowledge, both in semiatic control systems, as well as in other intellectual systems, based in their operation on a base of knowledge. For this reason we will attempt here to discuss some fragments of the not-yet-created theory of operation with incomplete bases of knowledge.

We begin with the observation that the incompleteness, which we are talking about, is many sided. There exist several forms of incompleteness, which effect differently the possibilities of occurrence of contradictory information in the system memory. Firstly, the values of the data, beginning with the numerical ones, may not be defined exactly. Instead of an exact value of q, about which it is known, that it assumes a certain numerical value, the memory of the system contains only an indication to a certain interval, within which this value occurs. Let us have, for example, a q, the value of which is fixed somewhere in the interval (0, 10) and a g, the value of which is fixed somewhere in the interval (5, 12). And let it be known to us, that $g<q$. Then the information stored in the system memory, is in a certain sense contradictory, since for values of g from the interval (10, 12) the condition $g<q$ will be violated. To remove the contradiction it is possible to improve the accuracy of the information for g and to consider, that g is assigned a value from the interval (5, 10), while q is also assigned values from the interval (5, 10). With these interval values it is possible to perform all necessary operations. At the present time, the arithmetic which operates with interval concepts, is well developed. Two such systems are described in studies mentioned in the commentary to the present chapter.

It is possible to assume, that concrete values which actually correspond to q, are not just fixed somewhere on the interval (a, b), but that there exists some additional information on the selection of a true value of q on that interval. This information can have the character of a probalistic distribution for finding a true value, or can be given, for example, in the form of a function of affiliation of the type we introduced in §2.11 and have already used several times. In such cases, there can also occur contradictory situations, related, for example, to the incompatibility of such statements: "With a probability greater than 0.99, the value of q is found in the interval (0, 1)" and "with a probability greater than 0.95, the value of q is formed in the interval (10, 60)". The arithmetic of such interval quantifiers is very weakly developed and as yet awaits its investigation.

In the same manner it is possible to determine an undetermined set, an undetermined variable, an undetermined predicate and an undetermined relation. Somewhat more difficult is introduced the concept of the undetermined function. This entire selection of non-objects requires, for operating with it, the creation of a special mathematical apparatus. Only in that case, appears the hope of mastering the theory of knowledge bases, in which undetermined objects are represented. In the framework of this theory there should, in particular, be developed methods of improving the exactness of interval boundaries and of descriptions of sets, predicates, relations and functions, which would allow the system to keep optimally eliminating in the context of its memory the contradiction possible in future caused by incoming new information.

Another form of incorrectness, threatening the "peace" of the contents of a base of knowledge, is the so-called overdetermination. With overdetermination, the identical interval constants, the variables and other non-objects are ascribed, simultaneously and with equal justification mutually contradictory values. For instance, a certain propositional variable is simultaneously ascribed the value of truth and falsehood. For working with such over-determined (absurd) objects there are also constructed special procedures. As an example of such a procedure may serve a four-symbol logic, in which to each propositional variable is ascribed one of the four values: truth, falsehood, total indeterminacy, and truth-falsehood (absurd). In recent years studies in the area of logic of this type are actively being developed, and there is hope, that in the near future the necessary apparatus will be created.

In conclusion of the present paragraph let us recall that which was already spoken of in §1.6. In a number of cases mutually contradictory statements occur in the system memory because, for every sufficient, life-like (not purely mathematical) concept it is always possible to examine it under different aspects or in different possible universes. And this contradiction is eliminated by a fixation of that aspect or universe, to which the given object is related.□

CHAPTER 4

GENERALIZATION AND CLASSIFICATION OF SITUATIONS

Interview with a fireman in a conservatory:

- *You have been working here for 20 years and now, probably, know well the difference between a base fiddle and a violin?*
- *Of course. The base fiddle burns considerably longer.*

From musical folklore

4.1 The Essence of the Problem

Let us return for a moment to Figure 1.7 on which we cited the general scheme characteristic for systems of situational control; its central part was the CLASSIFIER. With its aid is solved the fundamental problem--the obtainment of classes of situations each of which corresponds uniquely or with definite priorities to some or other decisions on control. From this becomes clear the important role of the process of generalization of descriptions and of their classification. At first sight it may appear that such a problem is actively solved in other sciences, for example in the theory of pattern recognition or in cluster analysis. This is so and not so. Certainly, many of the methods developed in the given divisions of discrete mathematics are used (and not unsuccessfully) here also. But the cardinal distinction of the problem of generalization and classification in situational control and generally in semiotic models consists in that, aside from the problem of forming generalized concepts and classification by a set of given criteria, it is also required to solve the problem of determining the pragmatically important criteria, which as a rule is substituted in the theory of pattern recognition by a search of informative criteria.

The distinction in these approaches can be illustrated by the citation* *...it is possible to see in a logarithmic slide rule an object for pounding in nails. With a logarithmic slide rule it is possible to pound in a nail, it (the slide rule) possesses a length and a certain mass. But that is the very same variant in which you see a tenth or a twentieth function of an object, which at the same time with the slide rule it is possible to calculate and not just drive nails.* It is possible to formulate many different concepts, which will be based on well defined logical considerations, but they will not play any role in solving problems of control. Why, it is not for nothing that among specialists working on the diagnostics of psychic diseases there exists such a definition: *He reasons absolutely logically, but incorrectly.* What does that mean? Well known is the test by means of which is checked the ability for *correct* classification. Out of the images of objects depicted on cards it is required to form certain groups--classes. In this problem are checked those bases, which a man will use in forming concepts. A classical example is the one in which the tester unifies into a single class the images of a man, a tree, and a pencil, and into another (class) the images of a closet, a water drain pipe, and an inflatable balloon. To the question about the reasons for such a classification, the tester replies that he selected the information criterion--the presence of an empty space inside the object. Such a classification is, of course, logically faultless. But is it the one which we expected a priori when starting the tests? Apparently not.

And so, the first peculiarity (characteristic) of problems of forming concepts and of classification in situational control is a search of *pragmatic criteria* of classification, capable of ensuring the finding of such generalized descriptions of situations, which would allow one to successfully solve the problems of finding solutions on the control of an object.

In the traditional theory of pattern recognition and in cluster analysis the concept of a criterion plays the critical role. It is precisely the criteria that act as the parameters on the basis of which takes place the selection of generalized concepts and is constructed one or another classification. In languages of situation description, typical for all methods of control based on semiotic models, a huge role is played not by

---

*Semenov, Yu., Seventeen Moments of Spring, Baku, 1981, p. 17.

criteria, but by structures of the relations among the objects. In many cases, of course, the complexes of relations could be considered as a peculiar kind of criteria. But this is hampered by two considerations. First of all, let us note the sharp increase of the number of criteria, because the number of possible complexes of relations increases sharply with the growth of the number of relations and increases of their type multiplicity, while the majority of their combinations does not have any significance for the solution of the problem of generalization. Furthermore, criteria are ascribed to definite objects, and the utilization of complexes of relations as criteria will lead to the necessity of introducing an enormous number of extraneous objects. The concepts connected with these objects will also be extraneous, for *the concept which we are not able to develop into a judgment does not have any logical significance for us*.* And, finally, often it is practically impossible to articulate out that criterion which is used for forming a concept. As examples of such pseudoconcepts (within the framework of the criterial approach) may serve: control handle, nose of a ship, queue for services, etc.

The second peculiarity (characteristic) of the problem or concept formation and of generalization of situations in the considered region is the availability of the procedures of generalization based on the *structure of relations* present in a description of situations.

And, lastly, the third peculiarity (characteristic) of the described procedures that are characteristic for all procedures operating on knowledge, is the possibility of working with names assigned to individual concepts and situations. The names, they, of course, are not criteria. In any case, they are criteria of a special kind. And their use during classification distinguishes the described approach from the traditional procedures. For man and his model of the world, the names, apparently, play a special role. As did write V. Katayev, *I have noticed that man is three times more tormented by the appearance of an object if he doesn't know its name. To give a name to surrounding things--may be, this alone is what distinguishes man from another creature. But I do not possess such a store of words as to name the millions of creatures, concepts, and things surrounding me. This is tormenting. But*

---

*Asmus, V. F., Logic, Moscow, 1947, p. 70.

*greater still is the torment, apparently, experienced by a thing devoid of a name: its existence is not fully valued. Hosts of unnamed objects are tormenting themselves around me, and in turn they torment me by the frightening realization that I am not God. Unnamed things and concepts stand in the glass cabinets of eternity like new gilded figures of not yet reincarnated Buddhas in the dark annexes of a temple.. They are waiting for their incarnation, and that can occur only then, when a completely new concept will appear in the world which will require a plastic expression.**

These three characteristics of the problem of generalization of concepts and of classification force the development of specific procedures of concept forming and of their classifications oriented toward the solution of problems of situational control and of other methods of control of similar types. The general setting up of these problems in one case has the following form (the designations were taken from §1.4).

On a set of concrete situations $\{Q_i\}$ (it is required) to find such a partitioning of them into classes, under which each class Q^i would have within the framework model of control a certain *sensible* interpretation of the process of control of a situation. On a set of complete situations $\{S_i\}$ it is required to assign such a set of classes $\{S_i\}$ that each of them would allow a *sensible* interpretation for the procedure of search of a solution on the control of the object. In particular, classification S^i should be matched on a certain basis with the classification on the set of influences (controls) $\{U_k\}$.

Let us make some classifications in connection with such a formulation of problems. As we see, the problem splits into two stages: a generalization of current situations and a generalization of complete situations. Such staging is not compulsory. But in the majority of practically interesting cases it is justified. Current situations are related only with the object of control itself, with the peculiarities of processes occurring in it. Therefore, generalized descriptions of the given object conform to the classes Q^i. They are based on criteria and structural relations characteristic of the object of control and important for the compression of information on the object and

---

*Katayev, V., The Herb of Oblivion, Moscow, Childrens Literature, 1967, p. 31.

processes occurring in it. Classes S^i on the other hand are formed mainly on the basis of those criteria and structural relations which are related to the process of control of the object; they utilize both a priori knowledge about the object and methods of its control, as well as the real pre-history collected in the process of functioning of the control system.

In setting up the problem of generalization for current situations, we spoke of a search for the partitioning of the set $\{Q_i\}$. In other words, the classes Q^i satisfied two requirements, obligatory for the partitionings: $Q^i \cap Q^j = \phi$ if $i \neq j$ and $\bigcup_i Q^i = Q$, where Q is the total set of all current situations. One or another solution of this problem is always possible. In solving the problem of generalization, however, and of classification of complete situations, it does not succeed, as a rule, to find a partitioning of their set into classes. This is related to the fact that for searching the classes S^i a very incomplete information is used which is obtained from technologists--administrators who possess on many questions of control their own opinion that does not coincide with the opinions of other experts. For this reason, instead of partitioning $\{S_i\}$ into classes in this case, a solution is made of the problem of finding of an envelope of the set S. Here $S^i \cap S^j$ are not required to be empty when $i \neq j$, but $\bigcup_i S^i = S$.

And one last explanation. In the process of generalizing situations S_i and appearance of classes of situations, there occurs a natural hierarchy, defined by the entry of some classes into others. Since the pragmatic classification is based on the desire to so form the classes as to ease the search for a solution on control of an object, the hierarchy of the classes of complete situations must be matched with the hierarchy of effects (controls) in the following manner: if $S^i \subset S^j$ and for S^i are characteristic, the controls from the set U^i, while for S^j--from the set U^j, then an entry of $U^i \subset U^j$ must take place.

Below we shall indicate the paths of solution of the posed problems. In the immediately following paragraph, however, we shall communicate some data of their psychology of thinking, which turn out to be important in solving problems of formation of concepts and of classification in systems of semiotic control.

4.2 Generalization of Concepts and Classification in Man

Here we sufficiently cursorily and concisely shall touch on data stored in the area of psychology of thinking and important for the solution of problems discussed in the present chapter.

First of all, attention is attracted by contemporary concepts on the manner in which concepts occur in a child, and the manner in which it separates out of the flow of personal emotions, one or another of their complexes which later appears for it as a single unified whole, as a concept or as a type situation. If these concepts are summed up, they will reduce to the following. In the initial period (so-called period of mumbling) the child, not yet mastering the word, and at the same time the ability to decompose the feelings and the concepts, perceives concrete situations in their undismembered form. Here the visual, the acoustical, and other sensations in it (the child) become closely connected with the given situation. For this reason, at the first stage of mastering the tongue, it is the audio concepts that begin to play the role of a name of a certain object (for example, beep-beep to designate an automobile or kit-kit for a cat). A characteristic example of the importance of *intonational mumbling* to understand a situation can be the numerous recent multifilms for children, the heroes of which do not speak any words, but only mumble intonationally *in English, in Polish, or in Russian*. And in spite of that, such wordless mumbling is quite unambiguously interpreted by the film viewers. During this period, the child unites into single complexes the facts and the phenomena that randomly occurred in the situation. For it (the child) the role of concepts is played by such concrete situations in which the determining as well as the random factors possess equal rights.

At the next stage of formation of its model of the ambient world, the child will use that, which psychologists call *pseudoconcepts*. These formations, in distinction from complexes, are no longer random. The main role in them is played by the associative connections among separate phenomena and facts, supported by that activity, which the child realizes in its surrounding medium. During this period it already masters morphology and obtains the possibility to separate among themselves the objects, the processes, and the phenomena (from beep-beep it forms beep-beeper). A little later occurs the desire to classify the accumulated knowledge. In this period, when the child still cannot operate with abstract concepts, but uses only concepts-complexes consisting of a finite number of concrete representatives included in the given concept, the child's classification is concrete. From this originates such classifying *definitions* as *the dog--it bit me right here* or *factory--that's where daddy works*. But at this age (4-6 years old) there already goes on an active development of the ability to separate from the surrounding world the concepts-classes. The concrete yields the place to the general. And if earlier to the question *Who drives the train*, the child replied *The engineers*, and to the question *Who breaks the toys?--Pete*, then now to the second question he is quite capable of giving the answer *The breaker*.

If one traces how the child structures its surrounding reality, then the following stages can be separated: non-separated gestalts of the situations, the separation, and the naming of some elements of such situations (assigning to them names according to some component of a given situation), the assignment of an independent name to a situation, the assignment of names to individual objects and facts, the formation of concepts-complexes, the naming of true concepts and the formation (as a rule, not formulated clearly) of classification functions that ensure the referencing of some or other single essences under some or other concepts, the identification (separation) of concepts as independent attributes connected to concepts the situational classification of concepts, and theological classification of concepts.

Much of the above requires classifications. We will give them gradually, using for this purpose not only observation or the development of childrens abilities toward classification and concept formation, but also corresponding observations on adult people. First of all, we shall again focus attention on the fact, that in forming true concepts (which grownup people operate with too), the basis which serve for singling them out, appear genetically before the creation of a clear realization of criteria of objects, processes and phenomena viewed as special, separate essences. For this reason in a grownup man, as in a child, enormous difficulties are caused by the problem of determining, via the criterial structure, of those concepts which have for him a self-evident character (so-called daily-life concepts in distinction from scientific concepts, that are introduced via strict scientific definitions).

In one of his popular books the well known specialist in the field of language and culture A. M. Kondratov cites the results of an experiment conducted with students of upper grades and with adult test subjects. It was necessary to give a definition of the concept "door". It turned out that the problem was exceedingly difficult. These are samples of the answers received in the course of the experiment.

–Door... On it you can hang a tablet, cover it with leatherette, make a nice handle.

–Well, door... That is when you go in the apartment. It opens. Also you can lock it, so as not to get burglarized.

–Door, it is wooden. Although not always – there are also glass doors, in hotels for example. It turns, the door. Although there are also slide-apart ones – in a compartment of a train.

On reading such "definitions" of an object, well known to everyone, a feeling of embarrassment for the person overcomes one. But one should not hurry with conclusions about the level of development of those who gave such answers. Try yourself to give definitions of such simple concepts as "window", "table", "spoon", and many others[1]. And here is how the concept "door" is defined in the "short explanatory dictionary of the Russian language for foreigners": "Door – an opening in a wall for entry and exit, and also that with which this opening is closed". Now the difficulty of the problem is evident. For even the explanatory dictionary gives an obviously inexact definition (or interpretation) of the concept "door". According to it, for example, a curtain which closes the "opening in a wall for entry and exit", is also a door.

And so, the majority of concepts, with which people operate in daily life, do not have for them a strongly distinct criterial structure. What then serves for unifying the concrete essences of some class, to which a special name is given? The answer to the stated question is apparently not difficult. In the overwhelming majority of cases this unification occurs according to the principle of an equal pragmatic function realized by all essences entering

---

[1] From hearsay, a well-known Leningrad fantasy fiction writer is credited with the following remarkable definition of a spoon. "A spoon, by military terminology, is subdivided into a gulper, a holder, and a connecting plank".

into the given class. In the answers about the door, attempts at manifestation of this function were encountered several times among the test subjects. In the interpretive dictionary an attempt is also made to utilize, for the determination of a door, its functional purpose. That is why definitions are found so often of the type "chair—that is on what [people] sit", "spoon—that is what cereal and liquid food is eaten with", and so forth.

However, along with that are encountered concepts, which have been formed by a different principle. It is interesting, that the classification possibilities of natural languages reflect precisely a different possibility of singling out concepts. The subdivision by gender, characteristic for a number of languages, is based on the singling out in matter of the complex criterion – sex. True, we extend gender to inanimate objects also, although to speak about the ungrammatical basis of relegating a table to the male gender and a stool – to the feminine, is hardly worth mention. But for animated objects this criterion highlights completely definite classes of essences. And in the Papua language, asmat grammar is used to distinguish six classes: objects narrow and tall, resting on a small base (people, trees); objects equally tall and wide (house, dog); wide and low, resting on a large base (felled tree trunks, smoke); swimming (fish, boat); flying (bird, butterfly); lying on top so that, in order to see them it is necessary to lift oneself on something (low, lying beneath the roof of a hut on a shelf and not visible from the floor of the hut). And for each such class there are individual methods of designation with the aid of grammatical forms.

Language allows us to also introduce abstract concepts which are articulated by direct reference to a classification criterion. Such, for example, is the concept *blue*, that separates out of the entire surroundings everything that has a blue color.

The classical experiments of Luriya, carried out during the 1930's in obscure villages of Central Asia (especially when the test subjects were women) have highlighted a number of peculiarities of concept formation which become little noticeable in persons subjected to the influence of science and of being taught scientific thinking. The main peculiarity, noticed by the investigation, is the dominance of distinction over similarity. The hypothesis that, historically first originated the operation of distinction and highlighting of distinction, and only afterwards the operation of unification and highlighting of similarity, was put forth already by A. Binet. He based his thoughts on the idea that the fact of distinction is directly observed in objects or phenomena being compared, while the fact of similarity is the result of carrying out a certain logical, abstract operation. The hypothesis of Binet was corroborated in the tests of Luria. The test subjects, for example, could not in any way formulate the class "domestic animals", when they were shown images on cards of domestic animals, plants, wild animals, objects of household use. And even to a direct question of the experimenter "What is there in common between a chicken and a dog?", followed the answer "Nothing in common. The dog has four legs, and the chicken two". Like children, the natives of outlying villages of the Pamir during those years gave thus the "definitions" of concepts: "Tree—that is an apple tree, a karagach tree, a poplar", "An autobus moves itself by the power of fire and a man moves it... If you don't pour oil in there and no people are there, it will not move". It is most characteristic, that many things which were proposed to be determined in the experiment, were not determined because "it is clear anyhow". In reply to a request to give a definition to the concept "mountains", the answer received was "If you have never seen any mountains, then how could it be explained to you?" Similarly, to the question how could it be possible to explain to a person who had never seen an autobus, what that is, the reply was obtained "I'll tell like this: autobuses are moving, they have four feet, chairs in front to sit, a roof for shade and a motor... But in general, I'll say: if you get in one, you'll see what it is".

The cited experiments once again underscore the idea that is important for us, that it is not the criterial structure at the center of attention of specialists on pattern recognition, but the situation position of one or another fact or object that determine its interpretation in a person that has no contact with scientific definitions and theories.

Let's now go on to the problem of classification. Here too the scientific style of reasoning has played a considerable role in our evaluation of this procedure. Beginning with Linneus who gave a brilliant structure of classification of plants, the principle of genus-species classifications has prevailed in science. Such tree-like classifications have become a general occurrence. The obtainment of a classification of this type appears to be tempting and absolutely necessary. The role of the relations of the type *part-whole, genus-species, set-subset, element-class*, etc. However, the same psychological experiments indicate, that the classification natural to man is founded on different principles. In laying out the cards the test subjects of A. R. Luriya had never put in a single group the images of a sheep and a wolf, because "the wolf will bite a sheep to death, and that is not good", but did put into a single group the images of a sheep, a kettle, a knife, a table cloth, saying "There will be a feast, we'll spread a table cloth, kill the sheep, cook a nice pilaf". Thus, the situative classification plays in man's thinking the dominant role, while the categorical classification (in particular genus-species)—plays a much lesser one.

Let's note, that among categorical classifications the use of directly observable criteria is also found most infrequently. Many classifying systems are based on pragmatic criteria which we have already mentioned. So for example were formed the concepts *implements of work* or *plants*. A special class of concepts are the social concepts: *family, school class, the numismaticists of a town*, etc. Classifications based on them are also very characteristic of man.

Let's pause on one other property of human thinking important for constructing classification systems. It is known that in forming concepts a considerable role is played by associations, that originate in all of us during work with linguistic and extra-linguistic texts. Let's mention the results related to the semantic space of Charles Osgood. The idea, which this scientist and his numerous pupils and followers had developed for many years, consisted in the assumption that all concepts with which man operates, occupy a definite position in the special subjective semantic space, the metric of which reflects the associative nearness of some or other concepts. This space itself is constructed in the following manner. For a natural language are selected antonymous pairs of the type: large-small, sharp-blunt, fast-slow, kind-wicked, etc. For each such pair (usually there are 300-400 of them in a language) a scale is set up, in which these words themselves occupy the end positions, while the middle of the scale is occupied by a neutral statement (for example not big, not little, etc.). In addition, the scale also has a number of divisions, which may even be left undesignated by any words. For a given list of words, the test subjects must mark on each of the scales the position of each word. Let's underscore that as a rule, the criterion which characterizes the scale, is in no way connected with the concept expressed by the given word. Still, from the requirement of the experiment, the test subject must place the word from the list into some division of each of the scales. A statistical treatment of the results of such an experiment is most interesting. It turns out that definite statistically meaningful regularities of word disposition on the scales are observed. After suitable treatment (for instance by the method of separation of factors of dispersional analysis) we obtain the possibility to separate out three main factors, that consume in total, almost all of the dispersion. These factors Osgood called: *scale of activity*. It is they that form the axes of the semantic space[1] On them are projected practically all the initial scales. Thus, the scales good-wicked, useful-useless, good-bad, etc. are projected onto the scale of estimates; the scales big-little, heavy-light, strong-weak, etc.—onto the power scale, and scales of the type: sharp-blunt, fast-slow, quiet-restless, etc.—onto the activity scale.

Now each concept may be considered as a point of semantic space. The most interesting result of Osgood for us, is that during the use of cluster analysis in that space we obtain, in individual clusters, concepts related to each other either by situative or by associative connections. Examples of clusters which are separated in the semantic space of Osgood are (knight, nose, tournament, glove, kiss) or (father, mother, son, daughter, dog).

Thus we arrive at a conclusion, very important for us, that in daily life man forms concepts and organizes systems of classification, as a rule, on the basis of those situations which he encounters in his daily activity and on the basis of those pragmatic criteria, the consideration of which is important for its successful realization. This fact must be taken into account in constructing classification systems in control systems.

And one more, final remark. In conducting psychological experiments for "instant response, in the course of which the subject must "without hesitation" reply by some word to the word spoken by the experimenter, there was discovered a stable fact of a probabilistic preference of the model

---

[1] There exist many critical remarks on the subject of Osgood; method and results, obtained by means of it. But the main ideas, about which we are speaking, are apparently outside this criticism.

answer. The character of the answer depended on the social group to which the subject belonged, but within the limits of this group the answer to many words is correctly foretold with a probability close to unity. If for example, the subjects are asked to reply rapidly to a query of the type "name a poet", "Name a fruit", "Name an object related to furniture", then in the overwhelming number of cases we will receive as replies "Pushkin, apple, and table". This occurs because, in the system of classification of real things of the surrounding world the man also possesses a certain evaluation of the typicalness of one or another situation, of one or another context, in which the given concept is encountered. It is like a ready answer to a situation, in which it is required to make an instantaneous decision, when there is no time for an analysis of the problem. Apparently it is exactly this ability (and also an a-priori playing-through of absurd situations, about which we spoke in §3.8) that allows man to rather successfully react toward encountered unforeseen and unfamiliar, from previous experience, situations. In any case, such tactics turns out to be successful in many cases.

Let's note that the idea of scaling, which follows from arguments cited by us, is in many ways similar to the conclusions and statements presented in §2.4, devoted to general properties of pseudophysical logics.

4.3 GENERALIZATION BY CRITERIA

Let us begin with the simplest and best studied case of generalization by criteria. The general formulation of the problem has the following form. There is a set of objects $A=\{a_i\}$ and a set of criteria $M=\{\pi_i\}$, each of which may take on some value from a corresponding set of values of criteria $\{M^i\}$. Here the upper index shows that criterion to which is related this set of values. All the sets, except A, are assumed to be finite. A "generalized object" is a certain subset of set A. It if is possible to construct a function $\phi(\pi_1, \pi_2, . . . \pi_R)$ such that it would determine the belonging or the non-belonging of any a_i to the subset of interest to us, then this function is called the "decision function" (in its place may be formulated a certain "decision rule" which utilizes values of this function of a certain procedure, which is difficult to express in the form of an analytical functional notation, but we shall not concentrate our attention on this as yet). In the simplest case the function ϕ may be a Boolean function, if for the determination of the belonging of an element a_i to the subset $A^i \subseteq A$ it is necessary to answer the question which criteria or values of the criteria it possesses or does not possess.

What possibilities do we have for constructing a decision function?

Example 4.1. On Figure 4.1 is shown a histogram of the distribution of values of a certain criterion among a set of objects. What

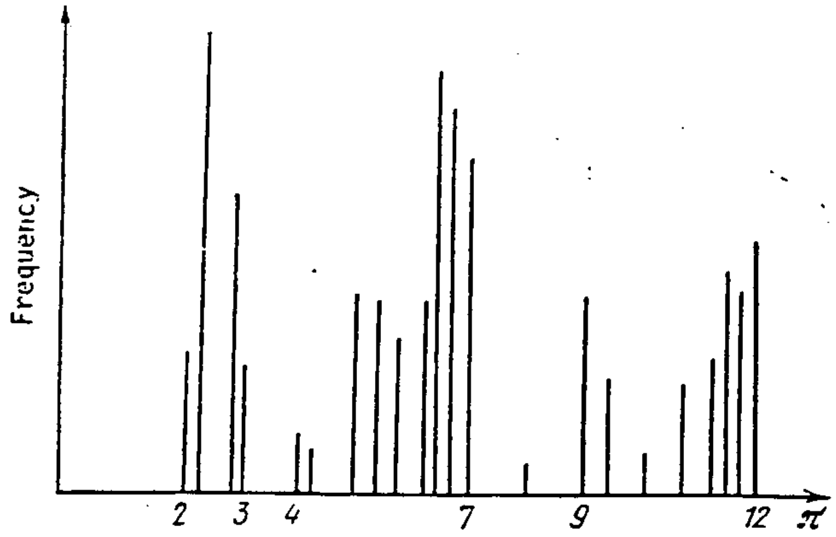


Figure 4.1

may be the basis for its partition into classes? Apparently, the only method would be such a partition in which a single class contains the objects with a sufficiently frequently occurring value of a criterion, and each class is separated from another by "gaps". On Figure 4.1 by this method it is possibl to separate three classes; corresponding to these are three generalized concepts, which can be given new names. To define the concepts it is possibl to use the assignment of sections from which are taken the values of the criterion: [2,3] [4,7], and [9,12].

The method illustrated in example 4.1, was in its time called "falling apart into piles" by M.M. Bongard. In methods of this type it is assumed, that the structure of the criteria had been selected so successfully, that in the critical space the objects forming a certain concept, are grouped in a compact manner. They sort of "concentrate" about a certain "nucleus", which characterizes the most significant, for the given criterion, combinations of their values.

On this principle were built rather many methods of generalization of concepts, which would be called "methods of partitioning in the space of criteria". In the simplest case, the situation shown in Fig. 4.1, is generalized onto a space of arbitrary dimensionality, and methods are constructed of separating the largest accumulations of objects, for which the distances between criteria are considerably less than the distances between individual accumulations. On this idea is based the majority of methods developed within the framework of cluster analysis.

As a generalization of and an approach to the formation of concepts serves the idea of using a special function for the separation of "nuclei" from the given set of objects. This function can be constructed by different methods. For example, it may assume, on objects entering into the nucleus, values significantly surpassing the values of the same function in a certain vicinity of the criterion space surrounding a nucleus. These functions sort of specify the potential distribution in the criterion space, which is why the method itself of working with such functions has received the application "method of potential functions".

However such methods, based on the utilization of only the information about the criterion space and some of its properties often studied in the theory of pattern recognition, give almost nothing for the solution of the problem of concept formulation which occurs in the theory of situational control and similar approaches to the solution of the problem of an expedient control of an object. And this is related to the fact, that in such conditions we always possess certain a priori information on concrete specimen that enter into the concept being formed. In other words, we know beforehand that certain objects a_i should enter into the concept being formed ("positive examples"), while other a_i should not enter into it ("negative examples"). This leads to a formulation of the problem of generalization by criteria on the basis of an "educated selection", consisting of positive and negative examples. Precisely these problems we are going to study below in the present chapter.

There exist a larger number of methods occupying a bordering position between the methods of the type of nuclei separation and methods based on the enlistment of additional essential information on the object of control and on methods of organizing its control. Let's describe briefly the most widespread method of such a borderline type which is usually called the "method of hyperplanes".

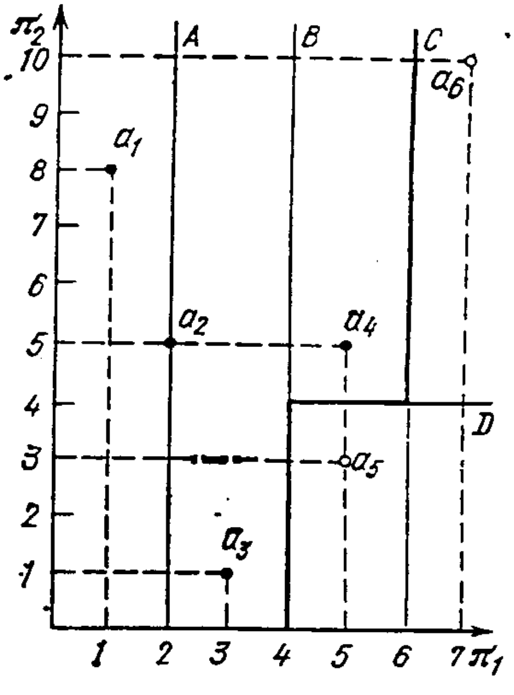


Figure 4.2

Let's illustrate it on the example of two criteria ascribed to objects a_i. This will allow us to illustrate the idea of this method geometrically, which assures the required visual clarity. In realizing the method of hyperplanes for an arbitrary number of criteria it is necessary to carry out a direct transposition of all procedures from a two-dimensional case to a n-dimensional. And so, let us have a certain set of objects $A=\{a_i\}$, the elements of which possess two criteria π_1 and π_2, assuming values from a certain sets M^1 and M^2. For simplicity we will assume that M^1 and M^2 are definite segments on the real numbers axis. Then all objects a_i in the criterion space are contained inside a certain rectangle. One of its vertices we shall take as the origin of coordinates. Let us consider the situation shown on Fig 4.2. On it are shown objects available in the educated selection. The black circles (dots) correspond to the position examples, the white circles -- to the negative ones. In this manner, in our concrete example there are four positive examples and two negative. We shall form the general concept in turn. We will take one of the positive examples (let it be a_1) and draw a straight line which would separate from a certain negative example (let that be a_5). For constructing the separating straight line $M\pi_1 + \ell\pi_2 + q=o$, we shall used the following technique. Let us choose q in such a way that the points corresponding to a_1 and a_5 will lie on different sides of our straight line. Substituting into the equation of the straight line the coordinates of points corresponding to a_1 and a_5, we obtain $m + 8\ell + q=o$ and $5m + 3\ell\ q=o$. Since we have two equations for determining three unknowns, we can construct the separating straight line in a different manner. Setting, for example, $\ell=o$ and taking $m=1$ and $q=-2$, we obtain the separating straight line in the form: $\pi_1-2=o$, which is denoted by letter A on Fig 4.2. However, this straight line is bad by the fact that it separates only one point of the formed set, and the complexity of the remaining problem is almost the same as that of the original one. Much more sensible would be to take as the separating line the straight line $\pi_1-4=o$. On Fig 4.2 this line is denoted by letter B. This line separates from objects, not belonging to the formed set, all the necessary points, except a_1. Now our task consists in separating this object from a_5 and a_6. For this, it is possible, for instance, to utilize straight lines C and D on Fig 4.2. The heavy outline highlights the obtained boundary between the formed set of objects (black dots) and the other objects. To these whether an arbitrary object $a^*(\pi_1^*, \pi_2^*)$ belongs to the formed set, it is necessary to substitute its coordinates into the equations of straight lines B, C, and D and to check the signs of deviation. If the signs of deviation comprise a combination (-, +, -), then the object in question does belong to the formed concept; if the combination is different, then the object does not belong to our concept. Such a partitioning does not succeed always. On Fig 4.3 is shown a situation, in which the positive and negative examples are grouped in such a way that upon drawing the dividing lines, regions occur that do not contain either positive or negative examples from the educated selection. These regions are denoted on the drawing by question marks. Concerning objects from them we cannot say anything about their belonging or non-belonging to the constructed concept. Therefore, in methods based on the separation of criteria in space, use is made of special techniques for increasing the efficiency of separation (for example, hyperplanes are passed halfway between two separated objects perpendicular to the segment connecting them) and for decreasing the number of regions in which the objects are not classified. In the literature list for this chapter these techniques are described in greatest detail.

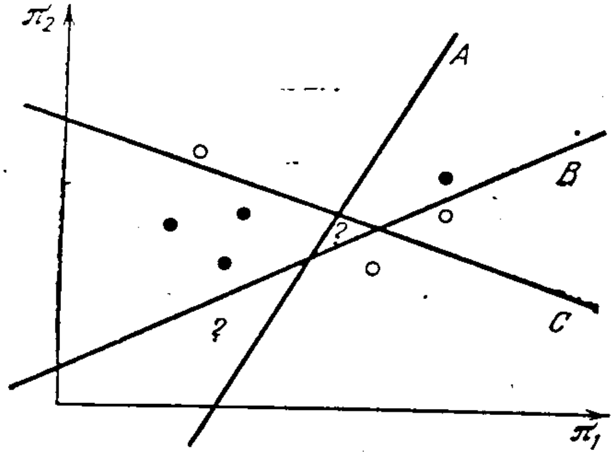


Figure 4.3

A modification of the method of separating hyperplanes are methods in which the separation takes place by means of curvilinear surfaces. However, due to the difficulties of forming successful separating surfaces of the second and higher orders, these methods have actually not found application in practice.

There is also a group of methods in which instead of a single separation in the criterial space are used metric functions of that space. For instance, the grouping of objects into a new concept occurs in a such a way, that the distances between pairs of objects that fit into a single concept, exceed considerably the distance from them to objects being outside of it (separation by potential function).

The shortcoming of the described method and of these methods which, similar to the method of hyperplanes, use only the hypothesis on the formation in the space of selected criteria of a compact set of objects fitting into a single generalized concept, consists in the fact that the languages in which is given both the description of objects a_i (i.e., the language of selection of selected criteria and of their values) as well as the description of the formed concept itself, are different. In our concrete method the concept is given by an enumeration of the separating hyperplanes with an indication of the signs of deviation from them. Such a method of describing the concepts does not permit carrying out the generalization procedure at the succeeding step, i.e., to obtain a concept of the second level of generality, and hence does not provide a basis for constructing a classification of concepts.

The indicated shortcoming, in spite of the simplicity of concept forming procedures by the method of separation in the criterial space, renders these methods of little use for problems of situational control. Precisely for that reason in systems of situational control, as a rule, use is made of other principles of concept formation and classification organization.

First of all, it is necessary to point out a group of methods, at the base of which lies the Bongard-suggested scheme of forming the decision rule for estimating the belonging of objects to the concept being formed. This scheme is arranged as follows. Certain logic functions of the criteria which may characterize the objects are given. For realization of such a possibility it is assumed that the criteria take on binary values (of the type yes-no). Such transition from arbitrary criteria to binary ones is theoretically always possible. For this it is sufficient to assume, that every acceptable value of criterion π_i acts as an independent criterion or, that in the set of its values have been introduced some predicates which transform their values into binary variable ones. The selection of these logic functions must be sufficiently rich as to encompass as many as possible relations among the criteria. In Bongard's method, depending on the specifics of the problem, the selection of predicates is made each time from scratch. If it is necessary to distinguish the arithmetic expressions, then acting as such predicates would be predicates related to arithmetic; if the matter concerns a comparison of certain programs and the formation of classes of programs, then acting as predicates would be predicate-procedures which evaluate the truth or the falseness of statements concerning, for example, the length of the program, the required volume of memory for its placement or its time of execution. If on the other hand, the object region in which generalization is being carried out, is given by problems related to situational control, then the predicates which transform the selections of criterial into binary variables, are fully determined by the semantics of that problem area in which use is made of situational control. It is important only to select a sufficiently complete assortment of such predicates, while their application to objects may be organized even with the aid of a random procedure.

When all objects had been juxtaposed by selections of binary criteria, the second stage commences. In it, by random means, are generated functions of the obtained binary arguments which also assume binary values. The number of arguments in these functions may either be fixed or vary from 1 to m, where m is the number of different binary criteria resulting from applying predicates to the original set of criteria. Each such function is considered as a potential elementary separation rule. To insure separation, it is necessary that for all objects contained in the number of negative examples in the educated selection this function would become zero, while for all objects contained in the number of positive examples it would assume the value of 1 at least once (but the more, the better is the separation rule). Then a disjunction of such elementary separation rules may be considered a final separation rule under the condition of a correct classification of all objects that form the education selection.

Also possible are other methods of organizing a general decision rule and also a selection of elementary separation rules to include into it. But the gist of the method does not change because of that.

Let us illustrate the execution of the above procedure first on a very simple example, in which the initial set of criteria is such that they are already binary. In that case the first stage, dealing with the application of unstable predicates to the initial criteria, is absent.

Example 4.2. On Figure 4.4 is shown an educated selection consisting of five positive and four negative examples. Acting as criteria are four independent criteria, which it is convenient for us for the sake of simplicity to denote by a, b, c, and d. All criteria assume binary values, and these values are assigned to them in the following manner: a=1 (hat is present), a=0 (hat is absent), b=1 (eyes are round), b=0 (eyes are slits), c=1 (nose is wide towards bottom), c=0 (nose is narrow towards bottom), d=1 (mouth is "happy"), d=0 (mouth is "sad"). Bongard's method consists of a multiple repetition of the procedures of random formation of an elementary separation rule in the form of a certain logical function, a checking of this rule for effectiveness (efficiency) of separation on the educated selection, and

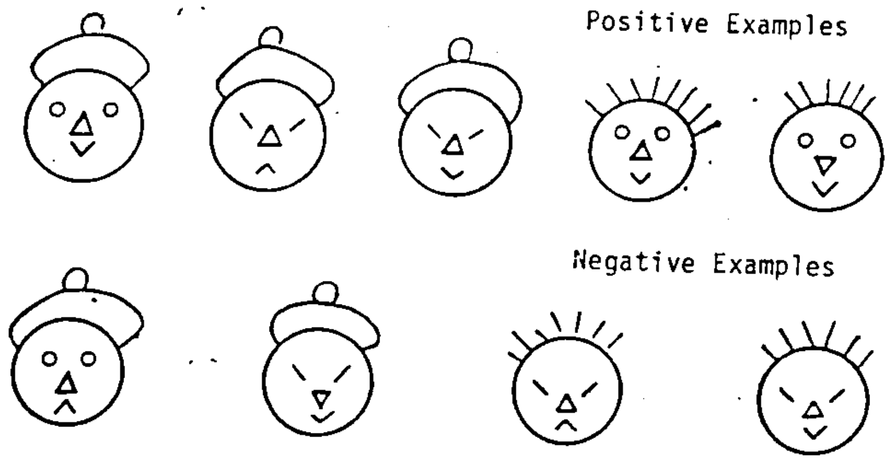


Figure 4.4

amassment of such effective separation rules, which, after their unification through a disjunction, will yield a global separation rule.

Let, as a result of operation of some random mechanism, there be worked out the following functions of the criteria: bd, $c \vee \bar{a}$, $a\bar{b}c$, abcd. What can be said about the effectiveness of the generated rules? Is it worth keeping them for later? The function bd corresponds to the selection of criteria round eyes and "happy" mouth. In the negative examples of the educated selection the corresponding function assumes in all examples the value of 0. For positive examples the function bd assumes the value of 1 in three cases out of five. This allows us to evaluate the elementary separation rule bd-1 as a very effective one. Corresponding to function $c \vee \bar{a}$ is "nose wide toward bottom" or "no hat". On the set of negative examples this function becomes 1 three times, while for the set of positive examples -- all five times. However the rule $c \vee \bar{a}=1$ does not ensure a high effectiveness of separation and it is not retained for further use. The function $a\bar{b}c$ requires that the personage should have a hat, slit-eyes, and a nose winding downwards. Among the positive examples there are two such personages. Among the negative examples such personages are absent. This means that the elementary separation rule $a\bar{b}c=1$ is effective (efficient). Finally, the function abcd becomes 1 only for those personages which have a hat, round eyes, nose widening towards the bottom, and a "happy" mouth. Among the positive examples in the educated selection there is one such personage, while among the negative examples there are none at all. This means that the separation rule abcd-1 may be retained for subsequent work. Thus, we have decided to retain three rules of separation. If it is assumed that the procedure of searching for elementary rules of separation is completed, then a global rule may be formed which will have the form $bdva\bar{b}cVabcd=1$. To be able to assume this rule as final, it should absolutely correctly classify the personages entering into the educated selection. This requirement is fulfilled. But the rule itself turned out to be superfluous. The elementary separation rule abcd-1 is obviously superfluous, because the rule $bdVa\bar{b}c=1$ already classifies correctly all personages entering the educated selection. Formally this could have been discovered by carrying transformations known from the theory of function of logic algebra: $bdVa\bar{b}cVabcd=bd(1VAC)Va\bar{b}c=bdVa\bar{b}c$.

To the questions on how correctly was constructed the separation rule it is possible to reply only on the basis of an examining selection that does not intersect with the educated selection. If, in the examining selection, the separation rule works effectively, this means we have constructed it successfully. In the reverse case this selection is considered as a complementary portion of the educated selection and the procedure of search for a decision rule is repeated anew.

In conclusion we will describe one other method of forming concepts on the basis of criterial structure, which is distinct in its idea from those previously considered. This method carries the name "method of growing pyramidal nets" (RPS-method). For us it is interesting not only as another method of generalization by criteria, but also because it can be utilized in the planning of multi-step decisions. And, in a corresponding chapter of this book we shall talk about it.

A growing pyramidal network is a dynamic object, its configuration changes in the process of functioning. It consists of apexes of two types, shown in Fig 4.5 by small squares and circles. The apexes of the first type are called "receptors", while the apexes of the second type -- "associative elements"[*] Each receptor corresponds to a certain fixed value of a definite criterion. Thus, the number of receptors in an RPS is equal to the product of the number of criteria by the total number of values assumed by it. Any element of an RPS can be in two states -- excited and unexcited. For all elements of an RPS there is fixed a certain definite number of cycles t_1 characterizing a period during which the excited elements retain their excitation and transmit it through their output areas. For excitation of receptors a special external signal is required, and the associative elements transit into excited state, if on all their inputs during a definite number of cycles t_2 the exciting signal is retained. Thus, although the transmission of excitation along the arcs of RPS occurs instantaneously, each associative element acts also as a delay the magnitude of which depends on the length of the path travelled by the excitation signal. In any case the magnitude of t_2 must be small compared to t_1.

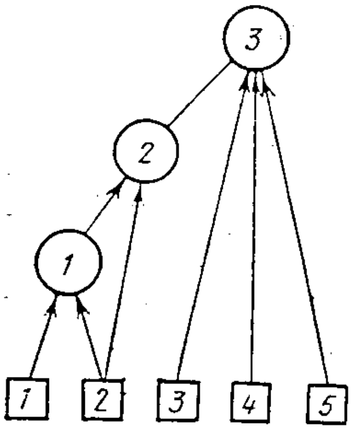


Figure 4.5

A selection of excited receptors at a certain moment of time is generated as a result of supplying to the input of RPS, at a given moment of description, of a certain object via the values of its criteria. The interval between changes of descriptions of objects, supplied to the input of RPS, is always greater than the interval during which RPS is excited. This means, that upon supplying any description to the receptors the RPS is in an unexcited state.

If there exists some associative element, then a subset of this element is formed by all those elements from which there are paths leading to the inputs of the given element. The subset is defined completely correctly if it is noted that in the RPS are forbidden connections which might lead to the origination of leaps and cycles; in other words, no associative element may be connected by its output with the input to itself or inputs to any

---

[*] By its functioning this element is an element of type I, but we will retain its name as given to it by the authors of the RPS-method. Usually, an associative element is understood to be an element that realizes a certain threshold function.

elements that form its subset. The zeroth layer of a subset of a given element formed by those elements the outputs of which are directly connected to the inlets of the given element.

Analogically it is possible to define also a superset of the given associative element as a set of those elements to which there exists a path leading from the output of that element. Those elements, for which this inlet constitutes directly an input, form a zeroth layer of the superset.

The dynamics of change (growth) of RPS is given by means of special rules.

1. Assume that there existed during excitation of a certain combination of receptors in the pre-existing RPS, certain associative elements, in which the zeroth layer of a subset, there became excited not fewer than two elements. For all these elements, the following takes place. The outputs of all excited elements in the zeroth layer of the subset become the inputs of a new associative element that is part of RPS; the output of the new element then becomes the input of the original associative element; all indirect connections existing in it before with the excited elements of the zeroth level of the subset are abolished.

The execution of this rule is illustrated in Fig 4.6. Those receptors that became excited are blackened. On Fig 4.6a is shown the RPS before the rule was applied, while on Fig 4.6b after its application. Afterwards, when in compliance with rule 1, all new elements had been introduced, a check is made on the conditions for applying rule 2, and if they are satisfied, then this rule is realized.

2. Let, during the excitation of a certain combination of receptors, take place the excitation of a definite part of elements in the RPS. If the set of excited elements, in which there remained unexcited all elements entering into the zeroth level of their superset, consists of more than one element, then a new element is introduced whose inputs are the outputs of all mentioned excited elements.

The execution of this rule is illustrated by the situation shown in Fig 4.7. On Fig 4.7a are shown two excited elements: the associative element 2 and the receptor 5, for which in the zeroth layer of their superset there are no excited elements. Then a new associative element 4 is introduced, and the RPS takes on the form shown in Fig 4.7b.

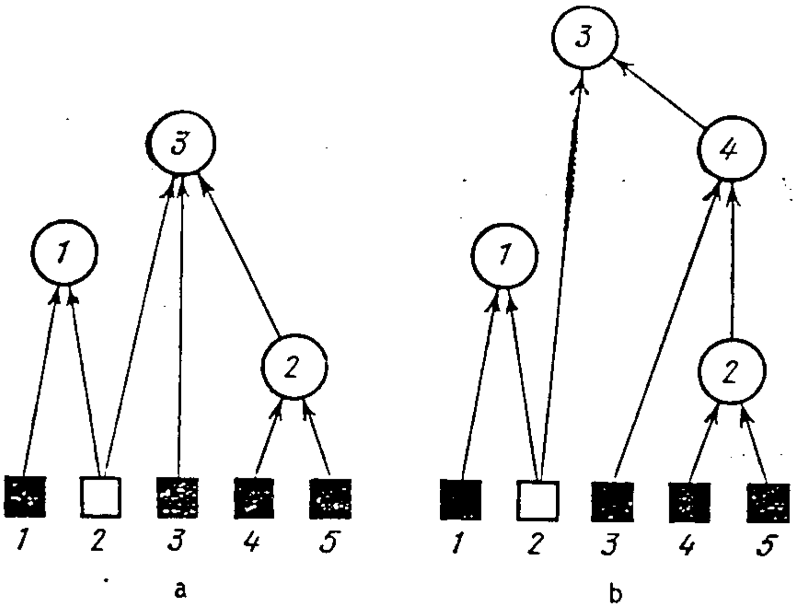


Figure 4.6

Let us now pass on to the procedure of forming concepts with the aid of RPS. In the educated selection with binary coded values of the criteria, the receptor of RPS corresponding to the given value of a certain criterion, is excited if this value is actualized, and does not become excited in the reverse case. To the input of RPS are supplied, in succession, all the values of the criteria of the objects from the educated selection, both from the group of positive examples as well as from the group of negative examples. Upon supplying each value there occurs a reforming of the RPS in accordance with the described rules.

To each associative element we will ascribe two parameters m and l, the values of which may vary in the process of RPS operation. Parameter m is the summary number of excitations of an associative element when a succession of examples is fed to the receptors of RPS. Parameter l characterizes the number of receptors entering into the subset of a given associative element. In the process of transforming RPS, special elements are separated out of the set of its associative elements; these special elements are named the positive and the negative control elements for the given concept.

The formation of concepts occurs by carrying out three special procedures.

1. If the receptors are fed a description of an object which represents a positive example from the educated selection, and if in the RPS no positive control element of the concept has been separated, then there exists an associative element for which m and l are maximal. Such an element becomes a positive control element of the concept. If, in a group of elements with an equal maximal value of m there are several elements with an equal maximal value of l, then any one of them is taken as a positive control element.

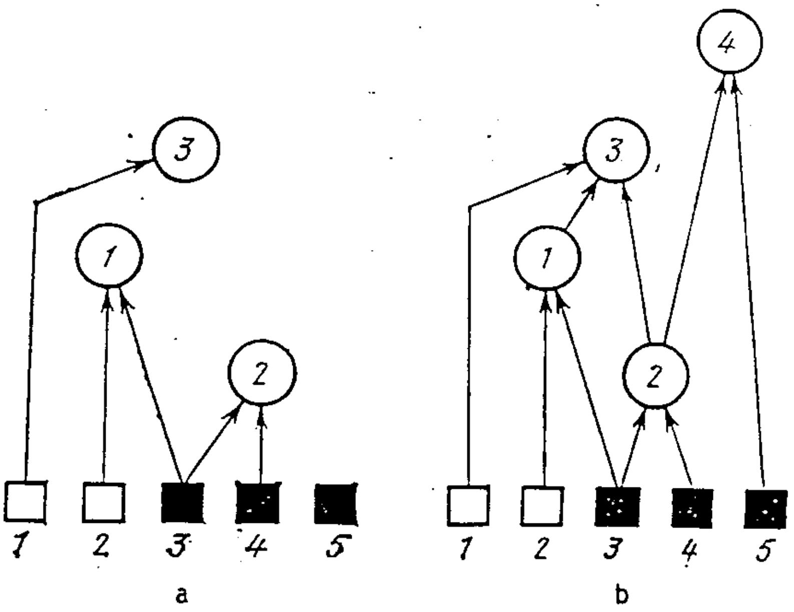


Figure 4.7

2. If the receptors are supplied with a description of an object which represents a negative example from the educated selection, and if in the RPS there are positive control elements of the concept which do not contain in their superset any other excited control elements, then in each of these supersets elements are selected as negative control elements that have minimal value of l. If there are several such elements, any one of them is chosen.

3. If the receptors are fed a description of an object which is a positive example from the educated selection, and if in the RPS there are negative control elements not containing in their supersets other excited control elements, then in each of these supersets elements are selected as positive control elements that have a minimal value of l. If such elements are several, then any one of them is chosen.

The carrying out of all three procedures is illustrated by situations shown in Fig 4.8. On Fig 4.8a is shown an RPS in which near all

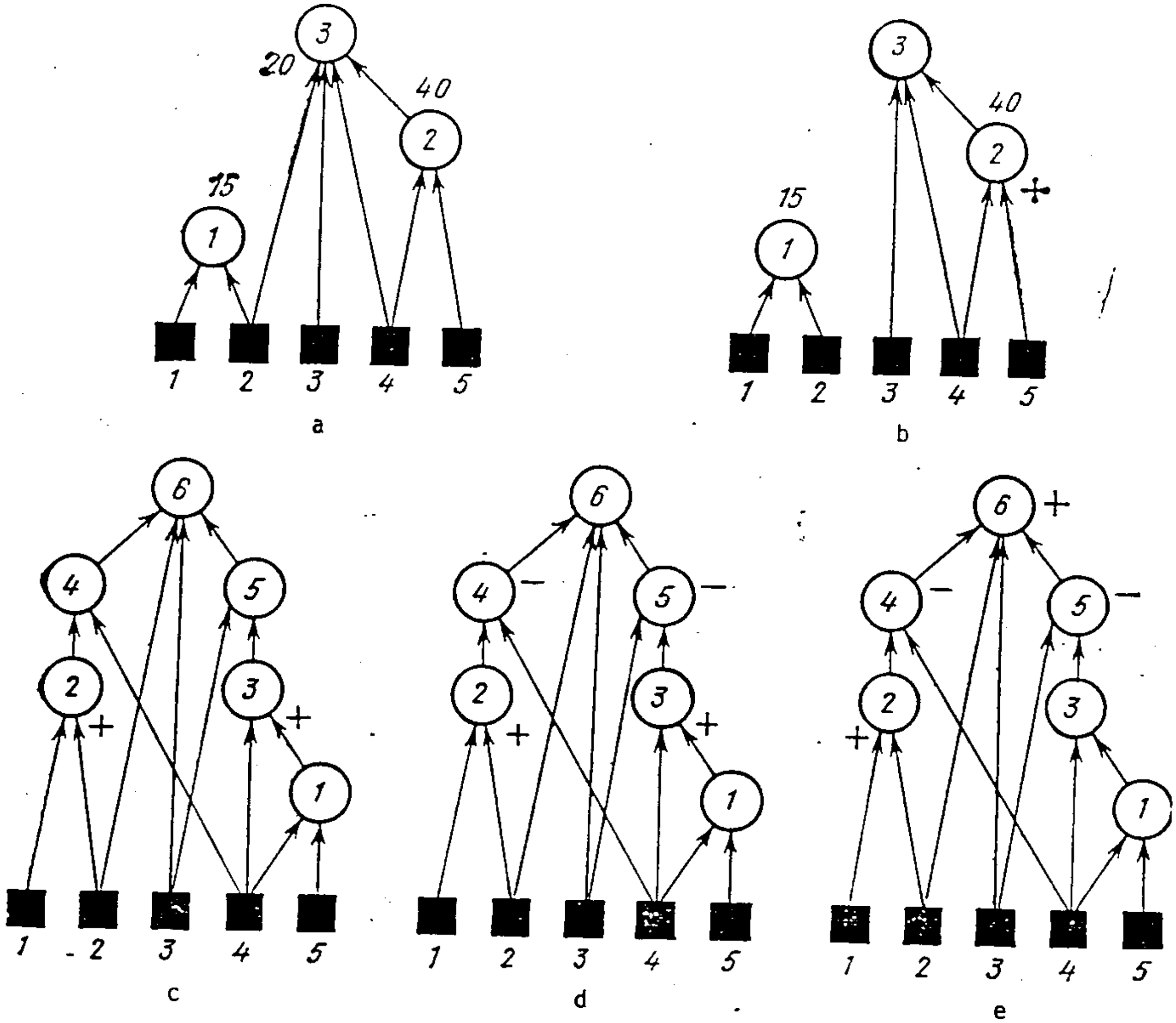


Figure 4.8

the associative elements a value of m is marked as collected before this cycle. To the input of RPS a positive example is fed. In accordance with the first of the procedures Fig 4.8b shows element 2 which in the subsequent cycle will play the role of a positive control element of the concept. On Fig 4.8b is shown an RPS to the input of which a negative example is supplied. On Fig 4.8c is shown the result of applying the second procedure. If now for the RPS, cited in this example, the receptors receive a description of a positive example, then, by execution of the third procedure, the RPS shown in Fig 4.8e will result.

The control elements play an important role in the process of forming the concepts. By means of positive control elements are separated those combinations of receptors (values of the criteria), which were encountered with a frequency m in the educated selection on positive examples. By means of the negative control elements are separated the combinations of values of the criteria of objects entering into the group of negative examples of the educated selection. But not for all such objects, but only for those in which the combination of the criterial values excites the positive control elements, which corresponds to the presence, in a negative example, of such a combination of criterial values which prior to that was identified as an important one (controlling one) in the formation of the rule of relating the objects to the concept being formed.

During the function of RPS, the educated selection is used many times. As soon as in the RPS, in accordance with one of the above described procedures, there occurs the appearance of a new control element, a new survey is started of the examples from the educated selection, beginning with the first one. The education is considered completed if the descriptions of all the objects entering into the education selection, upon their being fed to the receptors of RPS, will not cause an appearance of new control elements.

Once the concept has been formed, the RPS offers the possibility of using the following rule of relating the objects to the formed concept. An object a_i will enter into the concept if, upon feeding its description to the receptors of the RPS, it (the RPS) does not contain excited positive control elements that would not contain in their superset at least one excited negative control element. If none of the conditions formulated above is carried out, then no final judgments are made with respect to an object a_i.

It can be shown that the three procedures of RPS tuning, which are used in the process of working with the educated selection, do always during a finite number of steps lead to a completion of forming a concept. In so doing, all positive and negative examples from the educated selection are identified by the RPS correctly.

The above described method of forming concepts possesses several merits. All concepts formed with its aid, admit of a pithy interpretation for they are specified via a logic function into which the values of the original criteria enter in explicit form. The method lends itself well to automation when working on electronic computing machines. After forming the ultimate RPS, that part of it which does not contain control elements and does not affect their excitation, can be dispensed with, which in many instances allows to use the computer memory very economically.

For problems of the type characteristic for systems of situational control, a situation is very often encountered, in which, because of incompleteness of data on the object of control and methods of control, the educated selection can be contradictory. In this case the same descriptions can enter both into the group of positive examples and the group of negative examples. This important for us case can be realized in the RPS method with the aid of a certain modification of the procedures of its construction.

Because of the contradictiveness of the educated selection the decision or the inclusion or non-inclusion of object a_i into the formed class must be non-determinate. This non-determinacy must reflect those frequencies with which in the educated selection the object a_i is related to the set of the positive and the set of negative examples.

In the RPS the formed pyramids may now be of three types: M1, corresponding to a positive object, i.e., an object formed on the basis of those examples of the educated selection which enter only into the class of positive examples; M2, corresponding to a negative object, i.e., an object formed on the basis of those examples of the educated selection which enter only into the class of negative examples; M3, corresponding to a bivalent object, formed on the basis of examples from the educated selection which enter simultaneously into both classes of examples.

Let us describe now the procedures characteristic for that case. The procedure 1 does not change. Procedures 2 and 3 now assume the following form.

2. If the receptor is supplied a description of an object of the educated selection relating to the negative examples, the pyramid is not different from the pyramid of type M3, and among the excited positive and bivalent control elements there are such that in their supersets there are not other excited control elements, then in each of these supersets is selected an

element with the smallest value of l (or any one of them if there are several), which become the negative control element of the concept.

3. If to the receptors is supplied a description of an object of the educated selection from among the positive examples, and the pyramid is different from pyramid M3, and if among the excited negative or bivalent control elements there are such that in their supersets there are no other excited control elements of concepts, then in each of those supersets is selected an element with the smallest value of l (or any of these if there are several), which becomes the positive control element.

In addition, new procedures are added that are related to the appearance of bivalent concepts.

4. If to the receptors is supplied a description of an object of the educated selection which belongs to positive or negative examples, and the pyramid is of type M3, and among the excited control elements there are such that do not contain any other excited control concepts (except maybe the very apex of the pyramid), the at the intersection of these supersets is located an element with a minimal value of l (or any of them, if there are several), which becomes the bivalent control element.

5. If, upon acceptance of a certain object into the apex element of a pyramid formed by excited elements, a criterion M3 is entered (this happens when this apex corresponded to M1, and a negative example was received, or else when the apex corresponded to M2 and a positive example was received), then this apex becomes the bivalent control element.

In working with bivalent concepts there occurs the possibility to separate in the RPS individual pyramids which correspond to positive subconcepts, negative subconcepts and bivalent subconcepts. These pyramids can be inserted one into another in an arbitrary manner. For bivalent elements it is necessary to memorize a pair of parameters m+ and m- instead of a single m. They reflect the number of positive and negative examples of the educated selection in which the given element became excited.

The rules for relating objects a_i to the formed concept or for excluding from it remain the same. but to the decision rule another two cases are added. Object a_i is included in the formed concept, if the bivalent control element is excited, does not have in its supersets any other excited control elements, and if the equality is fulfilled:

$$\psi = \frac{h^{+} \sum_{j=1}^{s} m^{+} (a_i)_j}{h^{-} \sum_{j=1}^{s} m^{-} (a_i)_j} > 1.$$

Here h^+ and h^- are the prices (penalties?) for non-recognition of the corresponding object as being positive or negative, while s is the bivalent excited control elements, in the supersets of which there are no other excited control elements. If $\psi<1$, then the element a_i is not included in the formed concept.

It can be proved that with growth of the volume of the educated selection the decision rule ψ will in the limit give the well known rule of Bayes which minimizes the risk of error in classifying objects.

In concluding the consideration of RPS we note, that in practice cases are often encountered, when it is necessary to form concepts that have a dynamic nature. For such concepts the values of criteria defining them occur with a definite regularity in time. For forming similar concepts, apparently, it is possible to use an RPS with slightly modified conditions of excitation of the associative elements. Namely, we will assume that an element becomes excited if the signals at its inputs appear in a definite time sequence. For this purpose it is possible to set up at the set of inputs of such elements, a relation of strict order, for example by numbering them.

On Fig 4.9 is shown a certain RPS. If the receptors of this network are excited in succession, then the excitation of the associative elements will depend on the sequence of excitation of the receptors. For example, element 3 will become excited only in the case if at first to its input will arrive the excitation from element 1, next from receptor 4, and after that from element 2. This will occur if the sequence of excitation of the receptors will have the form 1-2-3-4-5 or 1-2-4-3-5. With other excitation sequences of the receptors, element 3 will not become excited.

In the language of situational control the criterial structures of the objects in their descriptions are given with the relations r_7 to possess a criterion, r_8 - criterion-value, r_9 - criterion-measure, and r_{10} - measure-value. These relations allow to describe all necessary knowledge about criteria and their values. Upon carrying out the generalization by criteria it is possible to reform the initial descriptions into descriptions that operate with the newly formed concepts. Procedures of generalization by criteria are possible which are oriented directly toward the language of situational control. There we won't talk about them, because later we shall consider this problem from more general positions.

Now we will proceed to another kind of generalization that takes into account not only the criteria inherent in objects but also these relations which exist among these objects.

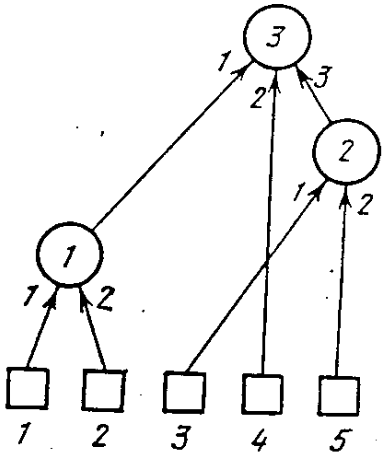


Figure 4.9

4.4 Generalization by Structures*

This type of generalization for situational control is most important. Because in problems of situational control, for describing situations, use is made of the language of relations. It is precisely in the structure of the latter that most frequently is hidden that which unifies into a single class the descriptions of situations. The problem of generalization by structures is tightly connected with recent actively developing methods of searching for regularities based on the apparatus of logic, and not on the traditional apparatus of mathematical statistics and theory of probability for this class of problems.

With availability of properly collected data on the process of interest to the researcher or to the control system, and satisfying known statistical requirements, it is possible to use many good methods allowing us to describe all the regularities that can be derived from these statistics. But, as was mentioned more than once before, the trouble with objects of control based on semiotic models, is either a total lack of observations which would be statistically reliable, or the presence of only indirect observations, from which the needed information can be obtained only by conducting additional logical operations.

These circumstances have led to the result that during the last decade, studies in the area of constructing regularities have been based only on a small number of corroborating or disproving examples. By their statements such problems are similar, on the one hand, to traditional problems of pattern recognition on educated selections, and on the other, to problems occurring in inductive logic.

At the beginning of this paragraph we will describe a procedure of advancing theses and construction of regularities, the procedure being based on schemes of inductive logic suggested already in the last century by S. Mill. These schemes, called by him *methods of inductive arguments*, have the following appearance.

---

* Parts of the book marked by single asterisks represent the "second level" of difficulty of the material discussed.

1. Method of Similarity. Consider the following derivation:

$$\begin{array}{r} a,\ b,\ c \Rightarrow d, \\ a,\ e,\ f \Rightarrow d, \\ a,\ g,\ h \Rightarrow d, \\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot \\ a,\ l,\ m \Rightarrow d, \\ \hline a \Rightarrow d \end{array}.$$

The number of transfers here is not determined. In any case they are sufficient that a person realizing the inductive derivation would be convinced that it is precisely transfer (a) as result of which is realized consequence (d). Of course this derivation is not indisputable. In actuality it may turn out that (d) originates not because of (a) at all. Reasons for (d) may be, for example, (b, e, g,..., ℓ). But the conclusion that it is precisely (a) that is the cause of (d), appears to be much better founded, than any other assumptions.

2. Method of Distinction. It is characterized by the following scheme of derivation (the sign $\overline{\rightarrow}$ is interpreted as "is not derived"):

$$\begin{array}{r} a,\ b,\ c \Rightarrow d, \\ a,\ e,\ f \Rightarrow d, \\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot \\ a,\ l,\ m \Rightarrow d, \\ b,\ c \overline{\Rightarrow} d, \\ e,\ f \overline{\Rightarrow} d, \\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot \\ l,\ m \overline{\Rightarrow} d \\ \hline a \Rightarrow d \end{array}.$$

This method does include in itself the preceding one, and strengthens it. In a more single variant, the scheme of the method of distinctions may contain only a few lines that are used in the method of similarity. As an example, only one such line.

3. Method of "Residues". In this method the following scheme of derivation is used:

$$\frac{\begin{array}{ll} a,\ b,\ c & \Rightarrow d,\ e,\ f, \\ a & \Rightarrow d, \\ b & \Rightarrow e, \end{array}}{c \Rightarrow f}.$$

As with the preceding schemes, this scheme cannot be considered absolutely reliable. Actually, (f) may originate not because of (c), but may be defined by the joint appearance of (a) and (c).

4. Method of Accompanying Changes. Let there be two sets A and B. Then the scheme of derivation can be written thus:

$$\begin{array}{ll} a_1,\ c \Rightarrow b_1, & a_i \in A, \quad i = 1,\ 2,\ \ldots,\ m, \\ a_2,\ c \Rightarrow b_2, & b_i \in B, \\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot & c \,\overline{\in}\, A, \\ \underline{a_m,\ c \Rightarrow b_m,} & c \,\overline{\in}\, B. \\ \quad A \Rightarrow B & \end{array}$$

With the aid of this method is made the usual inductive step. From corroborating examples, we pass on to the generalized statement that pertains to all objects of a certain class.

Let us now describe a special method of forming hypotheses on the basis of availability in the system, selections of examples, that corroborate and negate the formed hypothesis *(positive and negative examples)*. This method was called by its authors, the JSM-method[1].

Let us introduce two sets – the set of causes $C = \{c_1,\ c_2,\ \ldots,\ c_n,\ \ldots\}$ and the set of consequences $A = \{a_1,\ a_2,\ \ldots,\ a_n,\ \ldots\}$. We will refer to as the hypothesis an expression of the type $J_n^{+}(c_i \Rightarrow^{+} a_\ell)$, where J_n^{+} – special quantifier of the estimate of validity (reliability) of the hypothesis that c_i is the cause of a_ℓ. Simultaneously with the positive hypothesis $J_n^{+}\{c_i \Rightarrow^{+} a_\ell)$ we will consider negative hypotheses $J_n^{-}\{c_i \Rightarrow^{-} a_\ell)$, in which it is assumed, with reliability J_n^{-}, that c_i is not the cause of a_ℓ. These two hypotheses are

---

[1] This name is formed from the initials of John Stuart Mill.

formed in the JSM method independently. For this reason the estimates J^+_m and J^-_m can be, for example, simultaneously high, or simultaneously low, or else one of them can be high and the other low. In the JSM-method the estimates of reliability take on rational numerical values. For convenience of interpretation of J_m we will consider that a certain hypothesis has an estimate of reliability equal to m. We will omit the upper index in the estimates of reliability, in as much as for the positive and the negative hypotheses the estimates are taken from the same set.[2]

Values of J^+_m and J^-_m form two rectangular matrices J^+ and J^-. At the intersection of the i-th line and the j-th column of the first matrix stands the estimate of reliability of the hypothesis $c_i \Rightarrow^+ a_l$, and at the same place in the second matrix – the estimate of reliability of the hypothesis $(c_i \Rightarrow^- a_l)$.

The main procedure of the JSM-method consists in finding, on the basis of work with positive and negative examples, of new hypotheses, i.e., elements of indicated matrices J^+ and J^-, whose dimensionality increases in the process of operation of the method and also during recalculation of elements of these matrices on the basis of analysis of verification or negation of the corresponding hypothesis. Moreover, in the JSM-method there are values for carrying over of the hypotheses formed by the rules of the first group onto other "analogical" cases. Thus, rules in the JSM-method can be of two types: rules of distinguishing of regularities constructed according to the type of inductive rules of Mill, and rules of propagation of formed regularities onto those cases for which so far, such examples have not been found. It is important to underline, that both the rules of the first type as well as those of the second type constitute rules of not an exact, but of a plausible deduction.

Let I take on values $\{0, 1/(n-1), 2/(n-1), \ldots (n-2)/(n-1), 1\}$. Thereby the set $\Omega = \{w_{ij}\} = \{c_i \Rightarrow a_j\}$ is broken up into (n+1) subsets. Into subset Ω_0 enter those elements of Ω, for which the estimate of reliability is equal to 0. With respect to these hypotheses it is known exactly, that they are not true. Into subset Ω_1 enter absolutely reliable hypotheses, because their

---

[2] Let us note that in the JSM-method the estimates of reliability are rather crude (they are expressed by rational numbers). It is of course possible to use, instead of rational numbers, indistinct intervals or functions of affiliation.

estimate of reliability is equal to 1. Into subsets $\Omega i/_{(n-1)}$, where i = 2, 3, ... n-2, enter correspondingly the hypotheses, for which the degree of reliability is equal to the number equal to the index of the subset. The subset with the index 1/(n-1) is a special one. Hypotheses that enter into it are considered to be uninvestigated and the estimate $I_{1/(n-1)}$ corresponds to the value *incompletely defined*. Let us use the deduction rule of the first type. It could have generated new positive or negative hypotheses, to which are ascribed estimates of reliability on the basis of some rule. Moreover we find whether a number of earlier formed hypotheses (positive and negative) have been verified or not verified. For these hypotheses the estimates of reliability are recalculated. If the rule has verified certain hypotheses, then the recalculation is carried out in the following manner. The sets Ω_0 and Ω_1 are retained. To all verified hypotheses that enter into subset $\Omega i/_{(n-1)}$, were i = 2, 3, ..., n-2, are ascribed estimates (i+1)/n. With the aid of rules of analogy (see below) a part of the elements of $\Omega_{1/(n-1)}$ receives some estimates of reliability. Those elements from subset $\Omega_{1/(n-1)}$, which have not received such estimates, go over to subset $\Omega_{1/n}$ and are considered as incompletely defined at the next step.

If certain hypotheses were not verified, then their estimate of reliability should be decreased. The recalculation occurs in the following manner. Subsets Ω_0 and Ω_1 are retained. To all unverified hypotheses entering into sets $\Omega_{i/(n-1)}$, where i = 2, 3, ..., n-2, are assigned estimates (i-1)/n. With the aid of rules of analogy a part of the estimates is carried over to the hypotheses which entered into the set $\Omega_{1/(n-1)}$. Those of them, for which no new information was produced, will form the set $\Omega_{1/n}$.

The procedure of the JSM-method is concluded in two cases: if no new hypotheses are generated and if no change occurs of estimates of reliability of earlier found hypotheses.

Note, that within the frames of logic at each step of the JSM-method there takes place a transition from the logic of one base to the logic of another base (the base of the logic increases). The smallest base of logic with which the JSM-method begins its operation is the tertiary logic. In that case the values 1 and 0 are treated as an absolute estimate of the truthfulness of a hypothesis (it is ascribed to positive examples) and an absolute estimate of the falsehood of a hypothesis (it is ascribed to negative examples), while the value of 1/2 is ascribed to all those (if they are already there) hypotheses (or examples), on which the need information is lacking.

The last thing that we still need to consider, is the arrangement of the rules of generation of new hypotheses and of ascribing to them a certain estimate of reliability, and the rules of analogy.

As initial objects we will assume that the situation descriptions are in some relational language, for example the language of situational control. Let S^* be a fragment of a situation description. If it is known that for some situation S it is required to make a decision U, while fragment S^* enters into the description of situation S, then the hypothesis $S^* \Rightarrow^+ U$ has a right to existence. If, on the basis of an analysis of positive examples, we are convinced, that fragment S^* exists in a majority of descriptions of situations, for which the decision U is made, then the reliability of the hypothesis $S^* \Rightarrow^+ U$ must increase. It increases particularly strongly if this fragment is absent in the negative examples, that correspond to cases of making other decisions on control, different from U. If however, it is not so, then it can be assumed that making other decisions, different from U, in the presence of the fragment S^*, affects a certain fragment S^{**} that plays the role of a "brake" for S^*. An ideology of this type can propagate onto the brake of the second and higher levels (the brake of a brake, etc.). The brakes generate negative hypotheses $S^{**} \Rightarrow^+ U$.

Thus, in the subject area considered by us and with the selected language of description of objects and of hypotheses, the rules of elaboration of hypotheses are connected with the finding of fragments in the description of situations, which may turn out to be the reason for elaborating some or other decisions.

The general scheme for the *rule of similarity*, utilized in the JSM-method, has the following form:

$$\frac{J_{1(n-1)}(c_i \Rightarrow^+ a_j);\ J_1 M(c_i, a_j)}{J_{2n}(c_i \Rightarrow^+ a_j)};$$
$$\frac{J_{i(n-1)}(c_i \Rightarrow^+ a_j);\ J_1 M(c_i, a_j)}{J_{(i+1)n}(c_i \Rightarrow^+ a_j)}.$$

Here $M(c_i, a_j)$ is the predicate of similarity, and its estimate J_1 shows that its truthfulness is absolute. The predicate of similarity itself is defined in the following manner:

$$M(c_i, a_j) = \exists_k \tilde{M}(c_i, a_j, k).$$

The parametric predicate of similarity $\tilde{M}(c_i, a_j, k)$ becomes true at a definite value of parameter k, which determines the required number of examples used for manifestation of similarity. This parametric predicate of similarity is determined in the following manner:

$$\begin{aligned}
&\tilde{M}(c_i, a_j, k) = \exists_{x_1, y_1}\exists_{x_2, y_2}\ldots\exists_{x_k, y_k}\\
&\left(\left(\mathop{\&}_{l=1}^{k} J_b(x_l \Rightarrow y_l)\right)\left(\left(\bigcap_{l=1}^{k} x_l = c_l\right)(\overline{c_l = \varnothing})\right)\right)\\
&\forall_{l,q}[((l \neq q)(1 \leqslant l,\ q \leqslant k)) \to (\overline{(x_l = x_q)(y_l = y_q)})]\\
&\forall_{z,u}\Big[((J_b(z \Rightarrow u)(c_l \subseteq z)) \to ((a_j \equiv u)(\overline{a_j = \varnothing}))\\
&\left(\left(\bigvee_{l=1}^{k}(z = x_l)\right)\right)\Big] \quad (k \geqslant 2).
\end{aligned}$$

For those who are unaccustomed to the symbolism of mathematical logic, we shall give the necessary explanations, related to the interpretation of this predicate. The first part of the expression on the right of the equal sign states that from previous observations the relation $x_l \to y_l$ has been realized k times. Here J_b indicates that this relation did actually exist, because the index b highlights on the set of values of the logical variable some one value selected for fixation of true values (into these selected values may not enter, as in our case for example, the value 0, treated as falsehood, and the value 1/(n-1), which we treat as an indeterminacy). The parenthesis in which there is a sign of crossing, corresponds to the statement that n all observed cases of the relation $x_l \to y_l$, there was something common in its left parts, some fragment of the description c_i and this fragment is not empty. The expression in the first square parentheses is interpreted as a statement that among the observed k cases of interrelations of x_l with y_l of interest to us, there were no completely identical ones. Finally, the last square parenthesis highlights the very statement or the presence of connection between c_i and a_j, obtained on the basis of observations. Condition is the condition of total exhaustion of the entire set of examples.

If to the predicate described by us we add one more condition, attached to $\tilde{M}(c_i, a_j, k)$ c by means of an operation of conjunction,

$$\forall_{z_1, u_1} ((J_{н} (z_1 \Rightarrow u_1)(a_j \subseteqq u_1)) \rightarrow (\overline{c_l \subseteqq z_1})),$$

then we obtain a predicate that reflects a strong positive similarity between c_i and a_j. In this notation J_H means that $Z_1 \rightarrow u_1$ has an unseparated (unhighlighted) estimate.

The method of similarity was proposed by Mill in the following verbal formulation: "If two or more cases of the phenomenon subject to study have in common only one circumstance, then this circumstance, in which alone all these cases agree, is the cause (or consequence) of the given phenomenon".[1]

The method of strong positive similarity, which we had obtained from the simple method of similarity, is absent in Mill's work.

The second fundamental method, which follows from the schemes of deductive reasonings of Mill, is the *method of distinction*. The author himself had thus formulated its essence: "If the case, in which the studied phenomenon occurs, and the case in which it does not occur, are similar in all circumstances, except for one, which occurs only in the first case, then this circumstance in which alone they differ, is the consequence or the cause, or the necessary part of the cause".[2]

The scheme of the derivation rule by the method of distinction is the same as for the derivation by the method of similarity. The difference is only that instead of the predicate $\tilde{M}(c_i, a_j, k)$ use is made of the predicate of distinction $\tilde{N}$, which has the following form:

$$\begin{aligned}
&\bar{N}(c_l, a_j, k) = \exists_{x_1, y_1} \exists_{x_2, y_2} \ldots \exists_{x_k, y_k} \\
&\left(\left(\mathop{\&}_{l=1}^{k} J_b (x_l \Rightarrow y_l)\right)\left(\left(\bigcap_{l=1}^{k} x_l = c_l\right)(\overline{c_l = \varnothing})\right)\right) \\
&\forall_{l, q} [((l \neq q)(1 \leqslant l,\ q \leqslant k)) \rightarrow (\overline{x_l = x_q)(y_l = y_q})] \\
&\forall_{z, u} [((J_b (z \Rightarrow u)(c_l \subseteqq z)) \rightarrow
\end{aligned}$$

---

[1] Mill, J. S., System of Syllogistic and Inductive Logic, Moscow, Knizhnoye Delo publishers, 1900, page 313.

[2] Mill, J. S., System of Syllogistic and Inductive Logic, Moscow, Knizhnoye delo publishers, 1900, page 314.

$$\begin{aligned}&\rightarrow ((a_j \subseteq u)\,\overline{(a_j = \varnothing)}\left(\bigvee_{l=1}^{k} (z = x_l))\right)\\&(z_2 = v_2 - c_i)\left(\bigvee_{l=1}^{k} (v_2 = x_l)\right) J_b\,(z_2 \Rightarrow u_2)]\\&\forall_{z_3, u_3, v_3}\,[((z_3 = v_3 - c_l)\left(\bigvee_{l=1}^{k} (v_3 = x_l)\right)\\&J_b\,(z_3 \Rightarrow u_3)) \rightarrow (a_j \subseteq u_3)] \quad (k \geqslant 2).\end{aligned}$$

In the notation of this predicate its initial part coincides with the predicate of similarity. An explanation is needed only for the last part of the notation, starting with the parenthesis $(z_2 = v_2 - c_i)$. The minus sign, used here and after, is a sign of the theoretics-set operation generating the difference of two corresponding sets. This entire part of the notation formalizes the statement that, upon exclusion from consideration of the common $x\ell \rightarrow y\ell$ contained in the studied examples, the correlation between $x\ell$ and $u\ell$ disappears.

Since the predicate of distinction $\tilde{N}(c_i, a_j, k)$ contains the predicate of similarity $\tilde{M}(c_i, a_j, k)$, the method of distinction turns out to be more powerful than the method of similarity. Facts established by the method of distinction are better founded than those that were formed on the basis of manifestation of similarity alone. Similar to the predicate of strong positive similarity, it is possible to construct the predicate of strong positive distinction.

Finally, it is also possible to introduce methods of similarity and distinction not for a set of positive examples, but for a set of negative examples. The form of the corresponding rules of deduction and of predicates of negative similarity and distinction both in the weak as well as the strong form, has the appearance similar to the notations which we had already cited. Hence we shall not encumber the reader. We will note only the explicit connection of techniques used in the JSM-method, with the techniques we have already encountered in using pyrannidal growing networks.

As was already mentioned, the JSM-method consists not only of already analyzed rules of search of regularities, but also of rules of extension of the found results by analogy onto other cases. Two types of derivation by analogy are possible: with the aid of *negative analogy*. The rule of deduction by the method of positive analogy has the form

$$\frac{J_1 A^+ (c_i, a_j)}{J_{b'} (c_i \Rightarrow a_j)} \cdot$$

Here A^* – predicate of positive analogy, J_1 means that the truth of this predicate has been established, and b′ characterizes the estimate from a set of separated true vales, into which unity does not enter, i.e., the derivation on the connection $c_i \rightarrow a_j$ does not have an absolute character. The predicate A^* is stated in the following manner:

$$\begin{gathered} A^+ (c_i, a_j) = \exists_{z,u} (J_{b'} (z \Rightarrow u) J_\tau (c_i \Rightarrow a_j) \\ (z \subseteq c_i)(u \subseteq a_j)) \forall_{x,y} [((z \subseteq x)((y \subseteq u) \vee \\ \vee (u \subseteq y))) \rightarrow J_{u'} ((x - z) \Rightarrow y)]. \end{gathered}$$

Let us make a few classifications. As initial transfers in the method of positive analogy appears the statement that, with a certain reliability, characterized by an estimate of reliability from b′, there exists a relation $z \rightarrow u$, while nothing is known for the present about the relation $c_i \rightarrow a_j$ (the value of τ is the value of the under-certainty, which for example, could be 1/(n-1), as in our case). It is also known, that z enters into c_i as a substructure, and u is a substructure for a_j. Then it is asserted, that upon exclusion of z from the transfer the derivation (x-z $\Rightarrow$ will be estimated as unreliable. The set of non-singled out estimates of truthfulness H^1 can be defined by three different methods, which gives three versions of the predicate of positive analogy and, therefore, three versions of the rule of derivation by the positive analogy. In the first version a set consisting only of the value 0 is taken as H^1. In the second version this set is expanded to all the non-singled out true values, except value τ. In the third version τ is also included in H^1. Clearly, these three versions are ordered by their strength. A derivation made, for example, by the first of these, involves the other two derivations also.

The predicate of negative analogy, which enters into the rule of derivation by the negative analogy (the appearance of what has remained coincides with the rule of positive analogy) is written in the following manner:

$$A^{-}(c_{i},\ a_{j})=\exists_{z,u}(J_{H}(z\Rightarrow u)J_{\tau}(c_{i}\Rightarrow a_{j})$$
$$(z\subseteq c_{i})(u\subseteq a_{j}))\forall_{x,y}[((z\subseteq x)((y\subseteq u)\vee$$
$$\vee(u\subseteq y)))\longrightarrow J_{b^{-}}((x-z)\Rightarrow y)].$$

As is seen from the notation of the predicate of negative analogy, externally it is almost indistinguishable from the predicate $A^{\cdot}$. Except that instead of the set b' the set H, which consists of all non-singled out elements, except 0 and, was used, and instead of set H^{1} – set $b^{\cdot}$ was used. If, as $b^{\cdot}$, 1 is used, we will obtain the first, strongest version of derivation by the negative analogy; upon identifying $b^{\cdot}$ with the set of all singled-out elements the second version is obtained, and upon further addition of element τ – the third version.

In the conclusion of the rule of derivation by analogy itself, using predicate A^{-}, J_{H} appears instead of J_{b}, where H is the set of non-singled out estimates of reliability, into which values 0 and τ do not enter.

We have introduced, following the authors of the JSM-method, some rules of derivation which form the hypotheses and some derivation rules, that entered, by analogy, the stored information onto pairs $(c_{i},\ a_{j})$, for which the hypothesis has not been derived directly. In as much as the order of application of the rules of the first group is not fixed, it is quite possible, that at a certain step, contradictory information will occur in the memory of the system. For instance, with regard to some hypothesis there will exist two statements, one of which "votes" for the existence of relation of c_{i} with c_{j}, while the other denies it. However it can be shown, that non-contradictioness can be achieved by preventing a combination of pairs of rules of derivation of the first group, about which it is known that at least one of them is strong, if simultaneous use is made of the rules by similarity and by distinction. It is possible not to worry about contradictions at all, especially since with regard to combinations of rules of derivation by analogy with rules of the first and second types the problem of generating contradictory statements has not been studied. A practical way out from the situation when a contradiction occurs, is the exclusion of this pair (c_i, c_j) from the set of elaborated hypotheses and by ascribing to the value of estimate τ.

In a practical application of the JSM-method two important questions arise. How to estimate the truthfulness of one or another hypothesis obtained? How to organize the sequence of application of rules of derivation of new hypotheses and rules of analogy for attaining the most effective hypotheses under conditions of the assumed estimate of their truthfulness? The answer to these questions is most important. But, unfortunately, at the present time it has not yet been obtained. The authors of the JSM-method propose, as an estimate of the reliability of hypotheses, a vector estimate of the type $\nu = \langle n, m, \ell \rangle$, where n-number of steps of the JSM-method which is needed for derivation of the given hypothesis, m- a characteristic of the set of rules, utilized in the given derivation, ℓ - a characteristic depending on which exactly sets Ω_i were used in the derivation (i.e., what were the indices of J in the predicates used in the derivation). In the estimate ν were taken into consideration, apparently, all required data capable of affecting the reliability of the ultimate hypothesis. But the trouble consists in the fact, that at present there are no considerations about procedures that allow to calculate m and ℓ, and also no methods at all (which by the way is the common trouble of all problems with vector optimization) of comparing among each other the estimates of ν in the region of Pareto sets.

For solving problems related to the second question, it is possible to use some local considerations obtained by the authors of the JSM-method. Let us denote by R_f^g a certain rule of derivation, so that R_j^+ will be a positive rule, while R_j – a negative rule, R_{cc}, R_c, $R_{a(i)}$ – will be respectively, the rules of strong and weak similarity and the rule of analogy of the version i(i = 1,2,3). Their superposition means a successive application of the rules of derivation. The relations, proven within the framework of the JSM-method, that connect among themselves the rules of derivation, have the form

$$\begin{aligned}
&(1)\quad R_c^+(R_c^+(\Omega_\ell)) = R_c^+(\Omega_\ell),\\
&(2)\quad R_{cc}^+(R_{cc}^+(\Omega_\ell)) = R_{cc}^+(\Omega_\ell),\\
&(3)\quad R_c^-(R_c^-(\Omega_\ell)) = R_c^-(\Omega_\ell),\\
&(4)\quad R_{cc}^-(R_{cc}^-(\Omega_\ell)) = R_{cc}^-(\Omega_\ell).
\end{aligned}$$

These four properties reflect the circumstance that a successive application of two similar methods of similarity does not lead to the appearance of new information. Therefore such sequences of rules of derivation are not rational.

In addition, use is made of six more relations

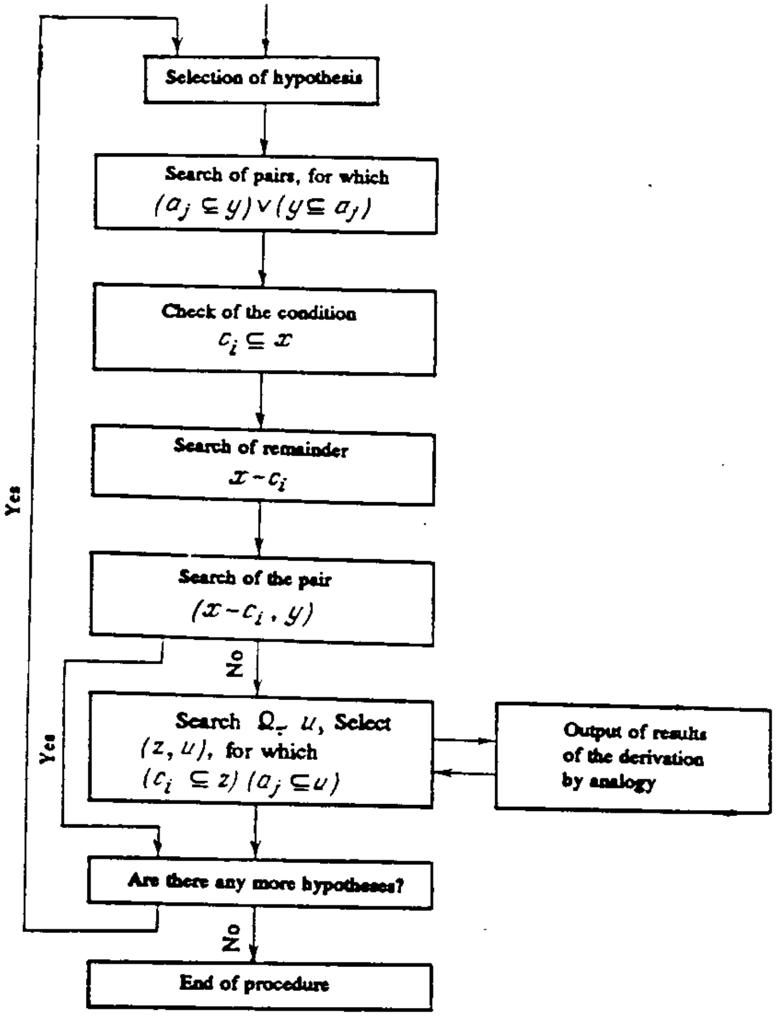
$$\exists_{\Omega_i}((R^{q}_{a(l)}(R^{q}_{p}(\Omega_l)) \neq R^{q}_{p}(\Omega_l))), \tag{5}$$
$$R^{q}_{p}(\Omega_l) \subseteq R^{q}_{a(l)}(R^{q}_{p}(\Omega_l)), \tag{6}$$
$$\exists_{\Omega_i}((R^{+}_{h}(R^{+}_{h}(\Omega_l)) \neq R^{+}_{h}(\Omega_l))), \tag{7}$$
$$R^{+}_{h}(\Omega_l) \subseteq R^{+}_{h}\ (R^{+}_{h}(\Omega_l)), \tag{8}$$
$$R^{+}_{h}(\Omega_l) \subseteq R^{+}_{a(l)}(R^{+}_{h}(\Omega_l)), \tag{9}$$
$$\exists_{\Omega_i}(R^{+}_{h}(\Omega_l) \neq R^{+}_{a(l)}(R^{+}_{h}(\Omega_l)). \tag{10}$$

In these relations the index p means a strong or a weak similarity, and the index h – a strong or a weak distinction (difference). The latter relations show that an application of the rules of analogy leads to an expansion of the set of hypotheses, which, in this manner, are considered further defined. Use of the cited relations among rules of derivation will apparently be able to aid in the future in constructing a more effective procedure of finding hypotheses than the one carried out in the presently existing programmed realization of the JSM-method. That which is available now requires a repeated examination of aggregates of k pairs, search of intersection in the left and the right elements of these pairs and singling out the descriptions entering into these intersections. Non-empty intersections will correspond to c_i and a_j and will form the hypothesis $c_i \rightarrow a_j$. In methods based on distinction it is required, in addition to this procedure, to find a counter-example that arises in the absence of the left parts of hypothesis c_i in the intersection. To remove the factorial complexity according to k, which occurs during direct search of all possible intersections of descriptions by k pieces (the magnitude of k varies from 2 to n, where n-number of descriptions stored in the original set), it is possible to use various methods of decreasing the number of realizations of intersection of the same description with others. But we will not dwell on this. In realizing rules of analogy a resorting is also needed, but not by k elements as for the rules of forming hypotheses, but only by k=2. But even in that case with a large volume of the initial set and the set of generated hypotheses, a direct sorting by pairs to single and conditions of applicability of one or another method of analogy does also have a polynomial complexity of a high order. For this reason, here too are needed special techniques which decrease the number of compared pairs of descriptions.

Although the essence of the procedure, realized on electronic computers is sufficiently clear from the form of notation in the predicates of similarity, distinction and analogy, let us illustrate one of these procedures by a block-scheme of the program shown in Figure 4.10. It illustrates the block-scheme for the rule of positive analogy of version 3.

Figure 4.11 shows, in a very schematic form, the general block-scheme of a program that operates in conjunction with the JSM-method.

Before concluding the account about the JSM-method, let us make one important remark, which has great significance for the use of the method of structural generalization in situational control systems. As it follows from the descriptions of the main predicates that enter into the rules and the procedures of the JSM-method, this method works successfully only in the presence of a non-empty intersection of c_i with c_j. In other words, it is assumed, that the reason which generates relations of the type $x_1 \rightarrow y_1$, $x_2 \rightarrow y_2$, ..., $x_n \rightarrow y_n$, is unique. In reality such an assumption is too rigid. It can often lead to the appearance of an empty intersection because the reason (cause) is not unique (not the only one). In other cases it may lead to continued conclusions with regard to the true reasons. For instance, while generalizing in this manner two statements "Today the weather is fine, I'm on leave – I'll go to the vacation cottage" and "Today the weather is fine, it's Sunday – I'll go to the vacation cottage", the system will come to the conclusion, that the reason for the trip to the vacation cottage is the presence of good weather, i.e., it will consider as reliable the hypothesis: "Today the weather is good – I'll go to the vacation cottage". In actual reality the reason may be a certain logic function of the fragments of descriptions entering into x_j, i.e., instead of the single hypothesis $c_i \rightarrow a_j$ there has been introduced a hypothesis of the type $f(c_i^1, c_i^2, \ldots., c_i^n) \rightarrow a_j$. The form of this function can be arbitrary. For example, a function of the type $f = c_2^1 \overline{c_2^2}$ tells that the relation $f \rightarrow a_j$ takes place only then when in the description there is present c_2^1 and is absent the fragment c_i^2. Thus, c_i^2 here plays the role of a brake which forbids the appearance of consequence a_j even in the presence of cause c_2^1.



Selection of hypothesis
Search of pairs, for which $(a_j \subseteq y) \vee (y \subseteq a_j)$
Check of the condition $c_i \subseteq x$
Search of remainder $x - c_i$
Search of the pair $(x - c_i, y)$
No
Yes
Search $\Omega_{\div}$ u, Select (z, u), for which $(c_i \subseteq z)(a_j \subseteq u)$
Output of results of the derivation by analogy
Are there any more hypotheses?
Yes
No
End of procedure

Figure 4.10

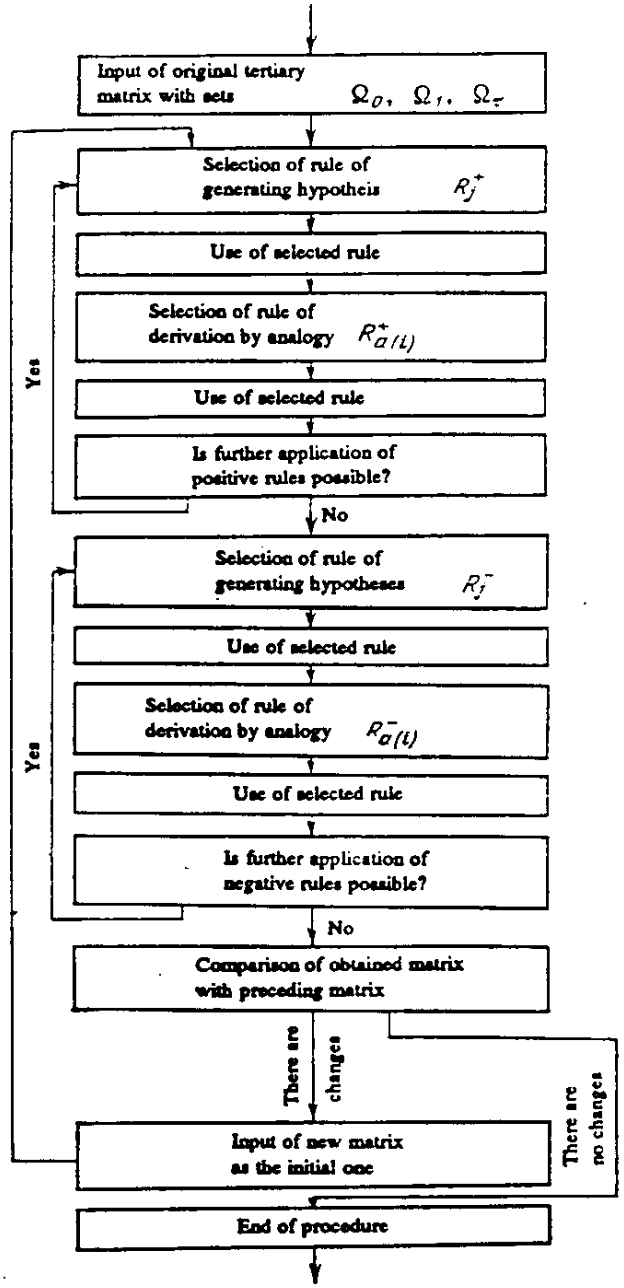

Input of original tertiary matrix with sets Ω_0, Ω_1, Ω_τ
Selection of rule of generating hypotheis R_j^+
Use of selected rule
Selection of rule of derivation by analogy $R^+_{a(i)}$
Use of selected rule
Is further application of positive rules possible?
Yes
No
Selection of rule of generating hypotheses R_j^-
Use of selected rule
Selection of rule of derivation by analogy $R^-_{a(i)}$
Use of selected rule
Is further application of negative rules possible?
Yes
No
Comparison of obtained matrix with preceding matrix
There are changes
There are no changes
Input of new matrix as the initial one
End of procedure

Figure 4.11

In the case, however, when $f = c_i^1(c_i^2 \vee \overline{c_i^3})$ the fragment c_i^2 may be considered as a brake for the brake c_i^3, since the presence of c_i^2 even in the presence of c_i^3 creates conditions for the occurrence of the relation $f \rightarrow a_j$.

The transition from the principle of a single cause to a consideration of an aggregate of causes together with the brakes and the brakes on brakes sharply widens the possibilities of the method of formation of hypotheses. In this case it is possible to consider the generalized rules of derivation of the type of rules of positive and negative similarities and distinctions and the modified rules of derivation by analogy. At the present time, in the JSM-method, there are no workable operating procedures of this type, but their creation does not appear to be excessively complex. Interesting and as yet unsolved is also the problem of establishing correspondence between the rules of derivation by analogy, used in our discussed method of generalization by structures, and the derivation by analogy methods described in §3.7.

In addition to the approach toward generalization of concepts by structure, which is based on ideas, expressed already by F. Bacon and S. Mill, it is possible to construct also other procedures, not based in explicit form on the idea of induction. Let us consider in this regard a number of examples.

Example 4.3. In Figure 4.12 is shown a certain structure, consisting

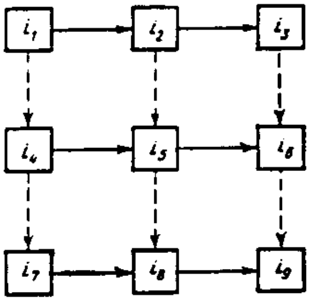


Figure 4.12

of single-type objects with names i_1 i_2, ..., i_9. It can be considered, for example, as a game field for the tic-tac-toe game familiar to everyone since childhood. To the solid arrows on that picture there corresponds the relation r_{32} – *to be on the right*, while the dashed arrows indicate the relation 2_{41} –

to be below. The entire structure shown can be described in the language of situational control in the following manner:

$$(((api_2)\,r_{32}\,(api_1))\,((api_3)\,r_{32}\,(api_2))),$$
$$(((api_5)\,r_{32}\,(api_4))\,((api_6)\,r_{32}\,(api_5))),$$
$$(((api_8)\,r_{32}\,(api_7))\,((api_9)\,r_{32}\,(api_8))),$$
$$(((api_4)\,r_{41}\,(api_1))\,((api_7)\,r_{41}\,(api_4))),$$
$$(((api_5)\,r_{41}\,(api_2))\,((api_8)\,r_{41}\,(api_5))),$$
$$(((api_6)\,r_{41}\,(api_3))\,((api_9)\,r_{41}\,(api_6))).$$

Triple parentheses in this notation single out same-type structures. The first three and, respectively, the last three structures differ from each other only by the names used in the monatomic formulas. Let us establish a relation of isomorphism among the first three fragments: $i_2 \leftrightarrow i_5 \leftrightarrow i_8$; $i_1 \leftrightarrow i_4 \leftrightarrow i_7$; $i_3 \leftrightarrow i_6 \leftrightarrow i_9$. If, for the isomorphic triads, we introduce the designation ξ_2, ξ_1, and ξ_3, then all three first fragments will become equal and will take on the appearance $(\xi_2 r_{32} \xi_1)(\xi_3 r_{32} \xi_2)$. This description may be considered as a new concept – a horizontal row of three elements. It can be very easily generalized to the case when there are n elements in the row. The description of this concept is written in the following manner:

$$(\xi_2 r_{32} \xi_1)\,(\xi_3 r_{32} \xi_2) \ldots (\xi_n r_{32} \xi_{n-1}).$$

Let us now consider the three last fragments in the original notation. Let us introduce the following isomorphism for names: $i_4 \leftrightarrow i_5 \leftrightarrow i_6$; $i_1 \leftrightarrow i_2 \leftrightarrow i_3$; $i_7 \leftrightarrow i_8 \leftrightarrow i_9$. If, for the isomorphic names, we introduce the designations η_2, η_1, and η_3, then all three last fragments will take on the appearance $(\eta_2 r_{41} \eta_1)(\eta_3 r_{41} \eta_2)$. This description may be considered a new concept – a vertical row of three elements. It can also be easily generalized to the case when there are n elements in the vertical row:

$$(\eta_2 r_{41} \eta_1)\,(\eta_3 r_{41} \eta_2) \ldots (\eta_n r_{41} \eta_{n-1}).$$

Let us now establish the isomorphism $\xi_1 \leftrightarrow \eta_1$; $\xi_2 \leftrightarrow \eta_2$; $\xi_3 \leftrightarrow \eta_3$; $r_{32} \leftrightarrow r_{41}$ and designate the relation, characterizing the last isomorphism, by ω. Then we will obtain the notation, in which ν_i is the designations of the isomorphic ξ_i and η_i: $(\nu_2 \omega \nu_1)(\nu_3 \omega \nu_2)$. The meaning of this concept is a row of three elements. For arbitrary n it is a row of n elements of the type:

$$(\nu_2 \omega \nu_1)\,(\nu_3 \omega \nu_2) \ldots (\nu_n \omega \nu_{n-1}).$$

For a more complete determination of the last concept it is necessary to include, in the isomorphism among the relations, the relation *to be on the left* and *to be above*.

Let us make some remarks in connection with the just cited example. With the aid of the isomorphism operation we have introduced into the description certain new elements; these may be considered as classes, into which enter all elements that are isomorphic among each other. Such classes are customarily called *classes of equivalence*. In the course of forming the final description we have successively increased the power (strength) of these classes. The resulting hierarchy is shown in Figure 4.13. It is a typical genus-species hierarchy which was mentioned in §4.2.

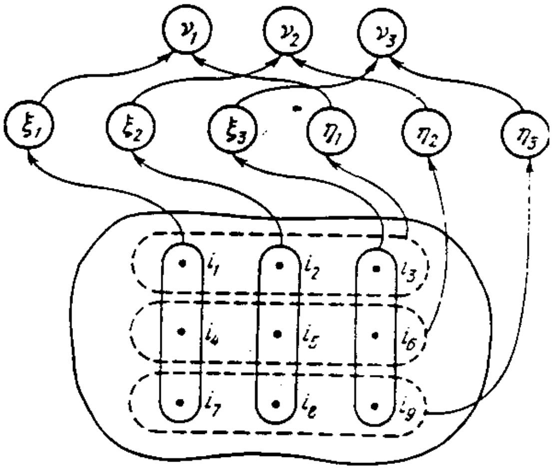


Figure 4.13

Example 4.4. Let there be given the following description of a situation in the language of situational control:

$$((api_1)\, r_{74} d)\, ((api_2)\, r_{34}\, (api_1)),$$
$$((api_3)\, r_{34}\, (api_2)) \ldots (api_{40})\, r_{34}\, (api_{39})),$$
$$((api_1)\, r_7\, (vr_6 0))\, ((api_2)\, r_7\, (vr_6 0)) \ldots$$
$$((api_{40})\, r_7\, (vr_6 0)).$$

Here the following designations are made: a-automobile, $i_j (j = 1,2,\ldots,40)$ – various national numerical symbols, d–auto refueling, v–velocity, r_{76}–*to have a criterion*, r_{34} – *to be behind*, r_7 – *to have a criterion*, r_8 – criterion – meaning. After clarification of the designations the situation described above, becomes clear. At a certain refueling place there is a queue of 40 cars with real national numerical symbols. The auto with number i_2 is being refueled at the given amount.

Let us gradually carry out the generalization of our description. First will introduce the free numbers $\xi_1, \xi_2, \ldots, \xi_{40}$, which will correspond to any pairwise different national numerical signs. Automobile $(\alpha\rho\xi_1)$ has a certain, exactly unspecified number. Moreover, let us not fix the number of automobiles in the queue. For that we will introduce, in addition to the free names, also a free ordinal number δ. Values of δ may be any elements of the national series. After that we find

$$
\begin{gathered}
((a\rho\xi_1)\, r_{76} d)\, ((a\rho\xi_2)\, r_{34}\, (a\rho\xi_1)) \ldots \\
((a\rho\xi_\delta)\, r_{34}\, (a\rho\xi_{\delta-1}))\, ((a\rho\xi_1)\, r_7\, (vr_8 0)) \ldots \\
((a\rho\xi_\delta)\, r_7\, (vr_8 0)).
\end{gathered}
$$

This description may be considered as the structure of the concept "a queue of automobiles for refueling". If now we exclude from the description the first parenthesis, then the remaining part may be considered as the structure of the concept "a standing column of automobiles". If however instead of the value of velocity 0 there is present in the description a certain non-zero free value of λ, then the resulting structure may be interpreted as a "moving column of automobiles". If, finally, instead of (a) there is substituted into the description the free variable η, interpreted on the set of objects capable of motion, we then find

$$
\begin{gathered}
((\eta\rho\xi_2)\, r_{34}\, (\eta\rho\xi_1)) \ldots ((\eta\rho\xi_\delta)\, r_{34}\, (\eta\rho\xi_{\delta-1})) \\
((\eta\rho\xi_1)\, r_7\, (vr_8 \lambda)) \ldots ((\eta\rho\xi_\delta)\, r_7\, (vr_8 \lambda)).
\end{gathered}
$$

This description reflects the structure of the concept "a moving column of objects". If the information on velocity is eliminated from it, the monatomic triplets with names $(\eta\rho\xi_1)$ are substituted by ν_i, and the ratio r_{34} is substituted by w, we will obtain

$$(v_2 \omega v_1)\,(v_3 \omega v_2) \ldots (v_\delta \omega v_{\delta-1}).$$

If this description is compared with the one we obtained in the preceding example, then, upon identifying δ with n we realize their complete coincidence. And indeed, the column of immovable objects may be completely treated as a row of n elements.

Example 4.5. Let us consider the following description of a situation:

$$((\beta_1 r_2 f)\, r_{59} c)$$
$$((\beta_2 r_2 f)\, r_{59} c) \ldots ((\beta_\delta r_2 f)\, r_{59} c).$$

Here β_j – abbreviated notation of a monatomic description of the type (gpi_j), where g – machine tool, i_j – individual numbers of the machine tools, f – a milling machine, c – the state *to be free*, r_2 – *to be an element of a class*, r_{59} – *to be in a state of*. Let us recall once more, that between all parentheses there exists the relation *to be simultaneous*. In this manner, here is described the situation, that at a certain moment of time δ milling machines are free. Let us generalize this description with the aid of a quantifier Ӽ, interpreted as *some of the*. The new description will have the form ((Ӽg)r_2f)r_{59}c). The meaning of this notation is the same as the initial one, because the value of δ is undefined. If, however, for example, we have a priori information on the number of the milling machines in the given plant and the value of δ is concretely indicated, then, depending on δ and the total number of milling machines, the quantifier Ӽ may take on the values: *only a small part of*, *about half of*, *almost all*, and so forth.

Let us discuss another two small examples.

Example 4.6. Let us consider the description:

$$((hpi_1)\, r_7\, (cr_8 b_1))\, ((hpi_2)\, r_7\, (cr_8 b_2))\, ((hpi_3)\, r_7\, (cr_8 b_3)).$$

where h – crane, i_1, i_2, i_3 – individual numbers of cranes, c – color, b_1, b_2, b_3 – concrete values of colors. The relations have the same meaning as in Example 4.4. Let us replace this description by another, in which the modifier m has the meaning *colored*: $((mh)\rho i_2)((mh)\rho i_2)\ ((mh)\rho i_3)$. Its meaning consists in the assertion on the existence of three cranes painted in certain colors.

Example 4.7. Let a – hammer, b – saw, c – pliers, d – chisel, e – plane, r_{47} – *to be an instrument for*, p_1 – pounding nails into wood, p_2 – cutting wood into pieces, p_3 – pulling nails out of the wood, p_4 – making depressions in wood, p_5 – smoothing out the surface of wood. Let us consider a description in which semi-colors standing between parentheses, correspond to an absence of the relation *to be simultaneous* among them $(ar_{47}\rho_1)$; $(br_{47}\rho_2)$; $(cr_{47}\rho_3)$; $(dr_{47}\rho_4)$; $(er_{47}\rho_5)$. Let us introduce the concept z – instrument and the concept p –process of working wood. The description $(zr_{47}p)$, which replaces the preceding one, can be interpreted, as the concept wood-working tool.

What do the examples cited by us tell? Most likely, that methods of generalization by structure are rather numerous. In Example 4.3 objects were generalized that participated in the description, their names and relations. And this generalization occurred by means of establishing a certain isomorphism. In Example 4.4 in addition to generalization by names, use was also made of generalization by the value of the criterion v, and also by means of introducing the indefinite quantitative index δ. In Example 4.6 the generalizing factor was the introduction of the modifier *colored* instead of a concrete indication of color. Finally, in the last example was used a generalization by processes with a generalization by objects.

For each kind of generalization it is necessary to have special means. However, for many of them there exists a common model which is related with a model of representing descriptors in the form of a *semantic graph*. Each semantic graph (SG) is a weighted multigraph, and the weights can be ascribed to both the apexes of SG as well as its arcs. Let us consider the SG's, in which five types of apexes are used. The apexes of the first type are named the *objective* ones. The weight of objective apexes has the form $<n,\tau>$. Here n – number of single-type objects, ascribed to the given apex or a certain quantifier estimating that number, τ – is the type of object (a class to which belong the objects corresponding to the given apex). Apexes of the second type are referred to as *criterial* ones. The weight of such apexes has the form $<(\pi_1, \Pi_1), (\pi_2, \Pi_2) \ldots (\pi_k, \Pi_k)>$, where π_i – names of criteria and Π_i – values of criteria. Apexes of the third type are referred to as the *predicative* ones. For their weight use is made of the name of some predicative P_i with indication of its locality. The *functional* apexes comprise the fourth type of apexes in the SG. As their weight is used that functional symbol which defines the essence of that apex. For the formational symbol its locality is indicated. Finally, the last type of apex in the SG are the *name* apexes; their weight is expressed through some symbols from the set of names.

Let us now describe the rules of drawing areas in an SG. Let us begin with the criterial apexes. Into them may enter arcs marked with weights, which correspond to the names of criteria present in the weight of the given apex. From a criterial apex there issues always only one arc, which leads to some object apex. For predicative apexes the number of incoming areas must correspond to the locality of the given predicate. These arcs issue from the object apexes and have weights determining the position of corresponding values in the predicate structure. They may be considered as definite indications on the roles which they play in the predicate expression. Outgoing arcs for these vertices may go towards the object or the functional apexes. For incoming arcs of functional apexes may serve the arcs coming from the object, the predicate and the functional apexes. The weights of the arcs determine either the values of the corresponding arguments or the possibility or impossibility of carrying out one or another succession of operations depending on the values issued by the predicate apexes. But for each functional apex there are known beforehand to exist at least one path, leading from some object apex to another object apex. Finally, the name apexes do not have any ingoing arcs, while the only outgoing arc is connected with some object apex. For object apexes the sole limitation on incoming arcs is that there may be only one arc coming from a criterial apex.

Among SG and descriptions in the language of situational control there exists a direct connection. Let us illustrate it by an example.

Example 4.8. On Figure 4.14 is shown an SG, which corresponds to the description cited in Example 4.4. On it the object apexes are marked with double circles, the predicative – by ordinary circles, the name – by small black circles, the criterial – by squares. The roles R' and R" are respectively the receiver and the sender. The rules Q' and Q" determine, respectively those, that stand or the left and the right in the relations τ_{34}. On Figure 4.14, b is shown one of the stages of generalization, when there occurred the generalization of all automobiles with numbers from 3 to 6 into a single object apex. Such generalization raises many questions related to the

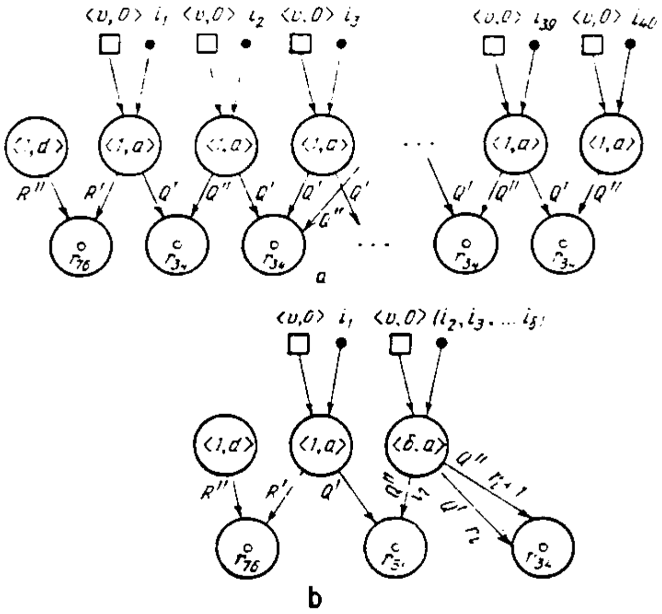


Figure 4.14

single-valued interpretation of those knowledges that are contained in the generalized SG, in particular, the knowledges or the selection of arguments for the predicate apex, connected with the generalized object apex, knowledges or the ascribing of names to objects entering into the generalized apex, etc. For an answer to such questions it is necessary to first formulate more precisely the concept of a generalized apex of one or another type.

Let us start with the object apexes. Let us say, that the object apex V_1 is *more general* than the object apex V_2, if the following conditions are fulfilled.

1. The component of weight n for V_1 is larger than for V_2. This corresponds to the case that the object, juxtaposed to apex V_1, contains more single-type elements, than the same kind of object juxtaposed to apex V_2. On Figure 4.14b the apex with weight <δ,a> plays in particular precisely that role with respect to the apex with weight <1,a>.

2. Weight component τ for V_1 includes in it as a subset the component τ for V_2. As an example of such generalization may serve the transition from the apex, for which τ – class of automobiles, to the apex, for which τ – class with the name "means of transport". This generalization is, in its essence,

with the name "means of transport". This generalization is, in its essence, the ascribing to the given object apex, of c different name, although the name juxtaposed to it by the name apex, may be retained.

3. The criterial apex ascribed to V_1, includes in it the entire selection of values of the criteria in the criterial apex, ascribed to V_2, with the proviso that all ranges of values of criteria for V_1 are not smaller than the ranges of values of criteria for V_2, and contains in the roster of criteria in its criterial apex at least one more criterion. If however the selections of criteria for V_1 and V_2 are identical, then the region of values for at least one criterion for V^1 includes in itself the region of values of this criterion for V_2.

These three conditions reflect the different methods of generalization, which we have considered in Examples 4.3–4.7. Thus, the first condition is a generalization according to a quantifier. In particular, in writing in a single apex of an indeterminate number of objects (as an example of this serves SG on Figure 4.14b) it is possible to utilize a qualitative quantifier of quantity instead of the quantitative value of the weight component n. The second condition is a generalization by name, since the τ are connected by some hierarchical relation of the genus-species type. In §2.9 we had already spoken about hierarchies of the type of AKO-connection, which is exactly what specifies conditions for generalization carried out by this principle. Finally, the third condition is the condition for generalization by modifiers, illustrated by us in Example 4.6.

Let us now consider apexes of other types. For name apexes V_1 and V_2 the apex V_1 will be the more general one, if the name represented by V_2, is included in the name represented by V_1. The entry of names should be understood in two ways. On the one hand, a set of names may correspond to V_1, while V_2 is a subset of names (in particular, one name), that enters into the set of names V_1. An example of such a situation is demonstrated to us by Figure 4.14; thus, in Figure 4.14b the respective generalized object apex is more general, than any of the name apexes on Figure 4.14a which are ascribed to those object apexes which later entered into the structure of the generalized apex in Figure 4.14b. On the other hand, there may exist a single hierarchy of names. As an example may serve the men's and women's names, characteristic for each language and in particular Russian. With their aid strength (power) of which there may even be no information. Another example – is the w the children name class, of which the elements in the Russian language are obtained by utilizing diminutive suffixes on many men's and women's names. Analogically there exists a definite tradition in the selection of football comments, of steamships, mountains, cats, dogs, and so forth.

For criterial apexes the generalization is already described by Condition 3 for object apexes. For predicate apexes the generalization is attained by a transition from one weight of the apex to another. The predicate apex V_1 is more general than V_2, if from the truth of the predicate ascribed to V_1 there always follows the truth of the predicate ascribed to V_2.

Finally the functional apex V_1 is a more general one than the functional apex V_2, if during realization of the function corresponding to V_1, there is also realized the function that corresponds to V_2.

Let us assume that SG_1 is a *generalization* of SG_2, if there is in SG_1 at least one apex which is a generalization of one or several apexes of SG_2, while in SG_2 there are no such apexes. Let us also say that in that case SG_2 is *superimposable* on SG_1.

Now it is possible to connect the above described apparatus of generalization based on the representation of information by semantic graphs, forming a definite subclass of semantic networks, with the apparatus developed in the JSM-method. Let there exist a certain educated selection, containing both positive as well as negative examples. In the set of positive examples, characterizing the formed concept, let there enter descriptions in the form of semantic graphs SG_1, SG_2, ..., SG_k. Then the problem of generalization can be formulated as follows. To find such an aggregate of semantic graphs SG^1, SG^2, ..., SG^*, which would be superimposable on SG_1, SG_2, ..., SG_k but would not be superimposable on the semantic graphs that enter into the set of counterexamples. In the simplest case, equivalent to the case of existence of a single reason for the hypothesis $c_i \rightarrow a_j$ in the JSM-method, it is required to find one SG^* which would be superimposable on any of the semantic graphs, representing a class of positive examples of an educated selection, and would not be superimposable on any graph, characterizing a class of negative examples of the educated selection. In a more general case, when the set of covering SG for all positive examples contains more than one element, there occurs the situation similar to that which in the JSM-method results in the appearance of generalized rules. Not wanting to encumber the reader with an exposition of corresponding procedures of search for SG* or a certain disjunctive conception of covering SG, which in many ways are analogical to the procedures of the JSM-method, we refer the reader to the studies of N. P. Viktorova, given in the roster of literature for the present chapter. In these studies is described the entire spectrum of necessary procedures of work with semantic graphs.

§4.5. "Layer Cake"

We have examined methods which allowed us to form generalized descriptions of situations entering into the educated selection. In the process of functioning of a system of situational control the work on the formation of classes of situations and on the more precise definition of earlier formed classes taken place continuously, since the educated selection may not exhaust the entire wealth of possible situations occurring on the object of control.

Generalization may take place in many stages, and therefore the initial descriptions of situations and their generalized descriptions form a hierarchical structure, in each layer of which are situated descriptions, obtained from the initial ones with the aid of some or other procedures of generalization. If the initial descriptions are taken as the zeroth level, then on the first level will be located descriptions obtained directly from descriptions of situations located at the zeroth level. To the second level will go descriptions which will originate from application of procedures of generalization to descriptions of the first level, and so forth. Thus will originate the "layer cake" shown in Figure 4.15. The situations at all levels of generalization correspond to certain decisions on control. In the ideal case, at the very top level of the system of classification are generated descriptions, to each of which corresponds a definite decision on control. In the more prevalent cases with an insufficient educated selection a final separation of classes may not even occur. It never occurs in an undeterminate process of forming generalized situations.

If the decisions on control themselves possess a certain structure, then, in forming the "layer cake" it is necessary to juxtapose the levels of control in the structure of the chosen control decision with the level of classifier. To achieve this, it is possible to use the following technique. Assume that in a certain situation it is necessary to make a decision U, which itself is decomposed into three independent decisions on control U_1, U_2, and U_3, related to decision U in the following manner. First, the decision U_1 is taken, and then, from the result of its realization, it is necessary to select either decision U_2, or else the decision U_3. In classifying situations this means, that there exists a certain class of situations, to which corresponds decision U_2, another class of situations, to which corresponds decision U_3, and finally, a class of situations to which corresponds decision U. Precisely U, and not U_1, because U_1 may define a completely different class of situations for which the single-step decision U_1 is final.

This leads to the necessity of dispersing of the descriptions of generalized situations for U and for U_2 and U_3 to different levels of generalization, because for U_2 and U_3 the teaching of forming generalized situations must occur only on the set of situations related to decision U. Moreover, at the same level as for U, there may exist classes of situations for which the single-step decisions U_2 and U_3 are final.

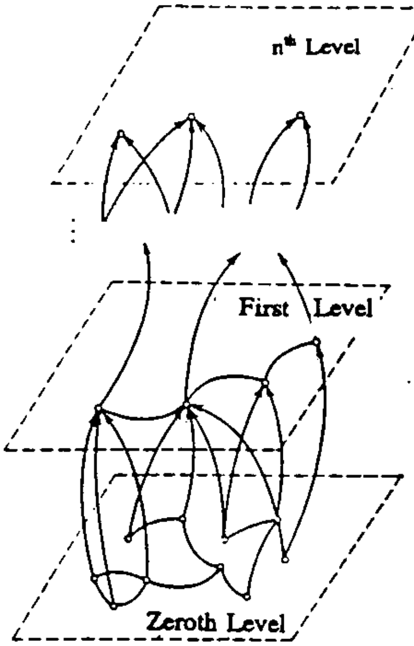


Figure 4.15

If there exists a control (a decision) U, which consists of n successive controls, then n levels of classification are produced in the "layer cake".

The hierarchic nature of control allows to form generalized descriptions during considerably lesser number of steps, than in the general case. Equally useful are also hierarchical descriptions of situations on the object of control, which is determined by the structure of the object itself. Hierarchical descriptions allow us each time to single out the general (common?) parts of the compared descriptions (it is precisely this procedure that is central in all methods of generalization in situational control) during a lesser number of operations, because the graph of the description naturally falls apart into graphs of lesser dimensionality.

In many problems, important is the theory of situational control, all these hierarchies have the special form of the I-graph. The general form of an I-graph is shown on Figure 4.16. Such hierarchical graphs possess the following properties.

1. An I-graph consists of a certain number of subgraphs G_i, which do not intersect with each other. On Figure 4.16 are shown three such subgraphs.

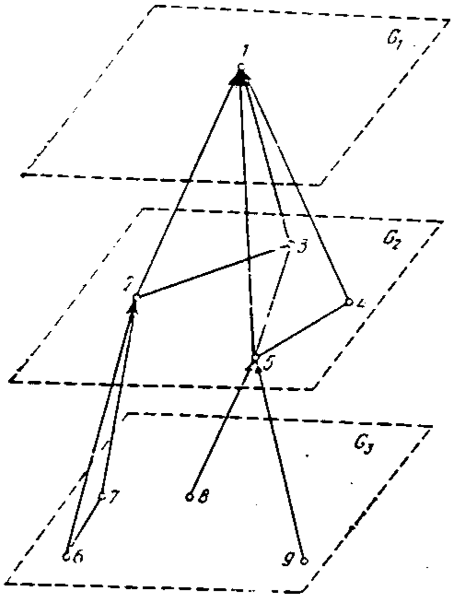


Figure 4.16

2. Subgraphs G_i can be ordered by levels under the condition of fulfilling the following condition i any apex of subgraph G_i is connected by an arc with some one apex of G_{i-1}. Subgraph G_1 always consists of a single apex with which is connected one or several apexes of subgraph G_2. The direction of the arcs, connecting the apexes of different subgraphs, is always such, that the arc is directed from the subgraph with a higher number to the subgraph with a lower number.

3. If, in a certain subgraph G_i, two apexes are connected by an arc, then they are connected to the same apex of subgraph G_{i-1}.

Example 4.9. On Figure 4.17 is shown the structure of a semantic graph, describing the knowledges in a certain region of a power system, where the following designations are adopted: П – user, CT – station, CB – communication line. Designations on the third level identify concrete representatives of classes shown at the second level of this I-graph.

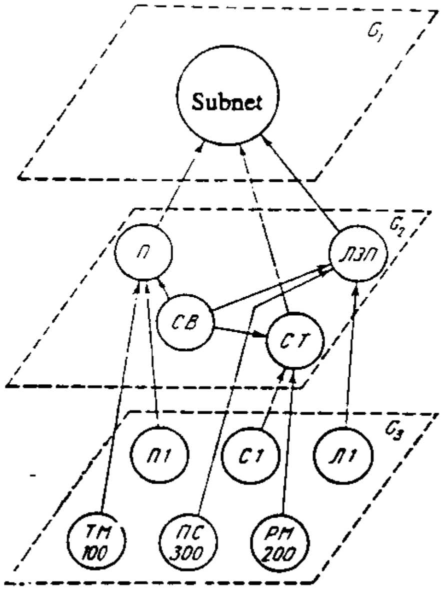


Figure 4.17

For I-graphs the solution of the problem of search of common fragments is semantic graphs requires considerably less time, than for SG's of arbitrary structure. Therefore it is desirable to try to represent the description of situations at an object of control and of common knowledges on the structure of control in the form of I-graphs.

In constructing a "layer cake" which comprises the essence of the Classifier, the following sequence of procedures is carried out. First the initial description of the current situation is complemented with all additional data through the operation of procedures of description complementation, described in Chapter 3. Next, the obtained description are subjected to generalization procedures. The obtained results, relating to the first level of the layer cake, are considered again as an initial description. To them are again applied methods of complementation, and the next sequential step of generalization takes place. This continues until such time, until at some sequential level the obtained fragments can no longer be treated by either complementation procedures, nor generalization procedures that would not lead to contradictory control decisions.

Once the Classifier has been formed, its operation consists of the following. If, at the entry to the system, a certain concrete situation, it is enriched by the operation of procedures of complementation of situation descriptions and enters to the zeroth level of the "layer cake". With the aid of vertical connections it is generalized to the highest possible level. If at this level there is a matching solution (decision?) on control, it passes from the Classifier to the Correlator. If however, because of impossibility of further enrichment, there is no solution matching the given level, then the Classifier goes over into a stage of instruction.

The hierarchical structure formed in the Classifier during the process of instruction, reflects that situative classification, which was discussed in §4.2.

§4.6. Some data on Classifying Systems**

Here we shall endeavor to communicate certain data on systems of classification, which may turn out to be useful in the construction of Classifiers of systems of situational control. In the most general form the system of classification is determined in the following manner. Let there be given a certain set M and, defined on it, a system of subsets $M_1, M_2, \ldots, M_n$. This system forms either a partitioning of M, or its envelope. In the first case, for any pair of subsets M_i, M_j, we have $M_i \cap M_j = \blacklozenge$, but $\cup_i M_i = M$. Subsets M_i in this case are called *classes of equivalence*. In the second case, there is at least one pair of such subsets M_i and M_j, so that at $i \neq j$ $M_i \cap M_j \neq \blacklozenge$ but $\cup_i M_i = M$. Subsets M_i in that case are called *classes of tolerance*.

Subsets M_i can in their turn be represented by systems of their subsets, which can either give a partitioning of M_i or its envelope. And so forth. Thus arises a classification of situations in which is simultaneously reflected both the genus-species classifications and that situative classification, which is defined by a selection of some or other subsets M_i, which form the classification structure.

A natural question arises – how to compare among each other the various classification for one and the same set? Of course, their comparison is largely determined by the objective of the classification. For example, for situational control it is necessary to have a classification which would provide the quickest and single-valued answers about the recommendation of some or other decisions or control. Therefore the classification will be best if the sets of the upper level of the "layer cake" produce the partitioning of the set of initial situations in such a manner, that to each class of equivalence of the given upper level these corresponds one decision on control. And among all such classification, apparently, the best one will be that, which contains a minimal number of levels, i.e., is the "quickest" in searching for an answer. But it has already been repeatedly noted, that it is practically impossible to construct classifications of this type. The majority of classifications will contain classes of tolerance rather than classes of equivalence. In that case the question on selection of the best classification becomes not quite so trivial.

Let us examine the problem on the quality of description of some subset M_i by criteria characteristic the set M. As criteria here and later we will understand to be both the criteria in the usual sense, as well as structural criteria, i.e., fragments of a structure specified or objects included in M (for example, representing such objects may be descriptions of situations in the language of situational control). If it was a matter of characterizing M_i by some selection of the best criteria, then it would be possible to introduce the criterion π, which is 1, if the object belongs to M_i, and is 0 – if it does not belong to M_i. If M_i is a class of equivalence for M, then the theorem is justified, of which we do not give the proof. It states, that the criterion π, if it does not constitute a criterion for M, can be expressed as a binary logic function of these criteria, which are utilized in M for organizing the classification structure. A drawback of this theorem is that it only declares the existence of this binary logic function, but does not give the opportunity to obtain its construction. The second drawback – is the requirement that M_i belong to the equivalence class, which limits the classification structures to only such, in which all M_i are classes of equivalence. Let use be made for the classification of the set M, of a selection of the n criteria π_i, i = 1, 2, ..., n. Then, corresponding to M_i, there is a certain conjunction of part of the π_i. If the classification is such, that among M_i there exists sets, which correspond to any of such possible conjunctions, then the solution (decision) that follows from the cited theorem, will be the only possible one. On Figure 4.18 is shown such a case of classification for two criteria defining the objects of the set M.

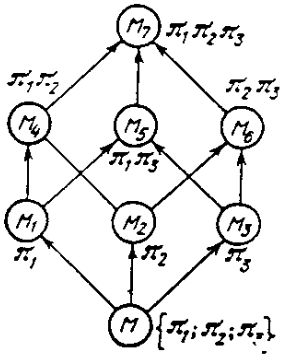


Figure 4.18

However, in the practically important cases there are rather many criteria, used for classification, and by far not all of their combinations are encountered in those subsets into which it is possible to decompose M. And beside that, classes of tolerance may be encountered in classifications.

For this reason, let us somewhat change the formulation of the problem. We shall search for such sets M_i of which the selection of characteristics is "sufficiently similar" to a certain system of criteria defined on M for the purpose of classification. We note that the classes, shown in Figure 4.18 on the same level, amy be considered pairwise "sufficiently similar", since they differ from one another by only one criterion.

Let us call an *α-etalon* for set M_i such a selection of criteria, that for any object x from M_i the following unequality holds:

$$|\pi^{M_i} \cap \pi^{x}| \geqslant \alpha |\pi^{M_i}|.$$

Here π^Y is a set of criteria which describe Y, $|Y|$ denotes the number of elements of set Y. The meaning of the α-etalon is that with its help there is singled out such a selection of criteria, that for any element included in M_i, the relative part of its criteria, common with all other criteria of set M_i, is not less than α. If there exists such an object $x \in M_i$, so that all of its criteria are included in the composition of the α-etalon, then it is called an *α-etalon object*. It is obvious that $0 \leq \alpha \leq 1$. For $\alpha = 1$ all objects included in M_i, possess certain common criteria. As an 1-etalon object in that case it is possible to take any object from M_i. If $0 < \alpha < 1$ the existence of an etalon object is not mandatory, since the criteria ascribed to various M_i^j from M_i, may not even intersect so that the condition for the existence of an α-etalon object is fulfilled. In that case the α-etalon given an idea of a certain "ideal object". Such a situation often arises in real life. When you participate in an exchange of your apartment, then even with the condition of an exact enumeration of all the criteria, by which the apartment must be characterized, an ideal apartment is hardly bound to exist that would satisfy all the requirements posed by us. But the fact that the exchange still takes place, speaks of the fact that for the procedure of classification of apartments and a selection, in conformance with this classification, of a variant suitable to us, a real existence of an etalon object is not mandatory at all.

Now it is possible to define the concept of a *measure of commonality* for set M_i. Let us say that M_i has a measure of commonality ϵ, if ϵ is the maximal value for which there exists an α-etalon for this set. Such an

ϵ-etalon is also called a *total etalon* for set M_i. For the same set M_i there may exist several total etalons. A selection of criteria which enters into some α-etalon, will be called *maximal*, if any expansion of the selection leads to its ceasing to be an α-etalon. Let now $\ell_\varepsilon = \max\limits_i |\pi_j^\varepsilon|$ be the maximal number of criteria in the total etalons (Π_j^ε is the number of criteria included in the j-th total etalon for M_i). In addition, let ℓ_m be the maximal number of criteria, that are used in M_i for the characteristic of some object included in this set. Then we will call a *measure of similarity* for set M_i the quantity $B_{Mi} = \ell_\varepsilon : \ell_{Mi}$. It is not difficult to see that $B_M i$ varies from 0 to 1. For $B_M i = 0$ there are no criteria in the total etalons, i.e., the total etalons do not exist. All objects in M_i are described by different criteria. For $B_M i = 1$ all objects in M_i are described by the same selections of criteria. Thus, the greater $B_M i$, the more similarity there is among the elements in M_i.

It is possible to consider as the best one that classification, for which the quantity $\min B_M i$ is greater. Before stating some other considerations related to the quality of classifications, let us show how it is possible to utilize the concept of α-etalons in the solution of problems of classification and the formation of concepts of a higher level.

Let there be two sets M_i and M_j, characterized by etalons α_i and α_j, which distinguish these sets. These etalons are called the *etalons of distinction*. For them the following statement is true: for any $x \epsilon M_i$ the expression $|\Pi^{m_i} \bigcap \Pi^x| < a_j |\Pi^{m_i}|$ applies, while for any $y \epsilon m_j$ the expression $|\Pi^{m_i} \bigcap \Pi^y| < a_j |\Pi^{m_i}|$ holds. Then it is possible to propose the following conditions of affiliation of a certain element z to the set M_i or M_j.

1. If, upon addition of z to M_i, are preserved the α_i-etalons of that set (let us note that an addition of new elements to a set can only decrease the similarity, but not increase it) and remains true the inequality $|\Pi^{m_i} \bigcap \Pi^z| < a_j |\Pi^{m_i}|$, then object z is relegated to set M_i.

3. If conditions 1 and 2 are not satisfied (their simultaneous fulfillment, as it is easy to see, is impossible), then object z cannot be included in these sets. For it there must exist another set M_k in M, or else it represents a new set, which must be singled out in M.

It is possible to use total etalons also in selecting characteristic criteria, defining the belonging of objects to some earlier formed generalized concept, to which corresponds the set M_i. If M_i has a single total etalon and the measure of commonality for the given set is sufficiently large, then this etalon may be considered as a interior for the given set, i.e., to consider, that "satisfying" the etalon is a reason of relegating the examined object to the given set.

Let us point out also a number of characteristics, that can be utilized in estimating the quality of a classification. First of all, these estimates turn out to be related to the effectiveness of solution of these problems, because of which the classification is being carried out. For us, such problems ar the problems of finding solutions on control for the situation, the description of which at that moment has entered the control system. Let us consider the set of possible solutions (decisions) on control $U = \{U_1, U_2, \ldots, U_n\}$. Each check, related to an analysis of descriptions of situation S, allows us to carry out in set U a partitioning, in which there is singled out from U a subset U' of solutions, which can be applied in the given situation. If ν' is used to denote the number of elements entering in U', then, the smaller is ν', the more precise recommendations on control we obtain at the given step. Ideally, ν' should be equal to unity. In that case the solution (decision) on control will be single-valued. A gradual more exact definition of U' is a certain process of checks connected with the analysis of U. Each such check possesses a certain value η', expressed, for example, by the number of operations used for it. If one pictures the entire procedure of the transition from U to a single-element U' (or a U' with the least number of elements possible for the given method of classification) in the form of a tree-like graph, then we obtain a graph, as shown in Figure 4.19. On it, the apexes correspond to those classes, which are singled out in the "layer cake", and the arcs—to the checks of the descriptions of S that allow to transfer from one generalized description to another. These arcs are loaded, as was already stated, by weights η', while the weights ν' are juxtaposed to the apexes.

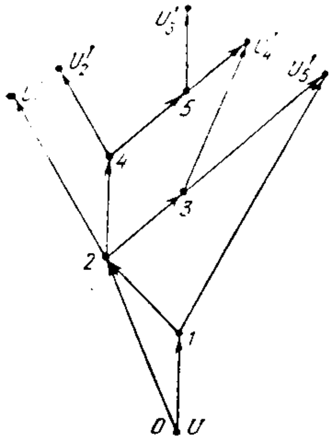


Figure 4.19

Structures of this type in the theory of diagnostics have received the name of *questionnaires*. To obtain the answer that interests us – a set of recommendations on control, we, as though, pose questions, to which in our case correspond the checks of the description of S. To each question an arc corresponds on the graph. The parameter ν' in the theory of questionnaires is called the *base of the question*, while parameter η' – the *value of the question*.

The introduced parameters allow us to pose the problem on optimization of the classification graph, and since its form is directly related to the structure of the "layer cake", also the problem or optimization of the structure formed by the Classifier. It is easily seen, that the average value (price) of bypassing the questionnaire depends on the order of questions, and so it (the value or price) can be varied by ascribing certain priorities to some or other questions. If for example, the value (price) of a question, leading from a root apex to apex 2, does not exceed the summary value (price) of questions leading from this root apex to apex 1, and then to apex 2, then for moving toward apex 2, it is more expedient to pose this single question. But the move to apex 1 allows, with a definite description of S, to immediately arrive by the next questions to the final apex U_5^1, which may turn out to be more advantageous, than moving to apex 2. This dilemma can be resolved only in that case if, besides the parameters N' and η', the system has some additional information about the frequency of relating S, during solving problems on control, to one or another class U_j'. But, as we had severely emphasized, such information is absent. In its place it is possible to utilize he information related to the structure of the questionnaire, but introducing the concept of *apex weight*. As weights of the end apexes remain to be the N_j' which correspond to the number of elements, entering into U_j', while as the weights of the intermediate apexes remain to be the sums of those end apexes, which are attainable with the aid of the questions from the given intermediate apex. In this manner, the weight of the root apex of the questionnaire is the number of all the possible solutions (decisions) on control. It is more convenient to utilize the relation weight of the apexes, which is equal to the local one from the division of the weight of the given apex by the sum of weights of all apexes forming the questionnaire.

If, for example, for the questionnaire shown in Figure 4.19, it is assumed that the weights of the end apexes, in the order as they are indexed, are equal respectfully to 1, 2, 2, 1, 1, then the weight of the intermediate apexes is the order of their numbers, beginning with the last one, will be equal to: 3, 5, 2, 7, 7, 7, while the relative weights will be obtained by dividing these values by 38.

The questionnaire parameters introduced by us can be utilized under such a transformation of their structure, during which there decreases the price (value) of their envelope (bypass), defined as $\xi = \sum_{i} \eta(i) w(i)$. Here $\eta(i)$ is the price (value) of the path leading from the root apex to apex $_{2}^{'}$. If this path is the only one, then $\eta(i)$ is the sum of values (η') of the questions forming it. If there are several paths, then the path with minimum value is selected. The parameter w(i) is the sum of the weights of all apexes that can be reached from apex i. The questionnaire is called *optimal*, if ξ is minimal. The minimization of ξ is achieved by transforming the questionnaire by means of four group rules.

1. The magnitude of ξ can be decreased, if it is possible to so redistribute the values of the paths toward two apexes of the questionnaire (neither of the apexes of the questionnaire is the predecessor or the successor of the other) or of the weight of two apexes, so that the apex with

greater weight would have a lesser value of the path leading to it, or vice versa.

2. The magnitude of ξ can be decreased, it is possible to rearrange the values or the weights of two questions in such a way, that the questions with the high weight would have a lessor value, or vice versa.

3. The magnitude of ξ can be decreased, if there is possible such a rearrangement of the bases and path values for apexes corresponding to two questions, so that the path values of paths toward the successors of the apex with the smaller base would turn out to be greater than the values of the paths toward the successors of the apex with the greater base, and vice versa.

4. The magnitude of ξ can be decreased by redistributing the weights of the successors of the apexes, from which are equal the lengths of the paths leading to them from the root apex. This rearrangement should be such, that the successors of the apexes arranged in the non-increasing order of their bases, could be ascribed weights in a non-increasing order. Next, it is necessary to apply the first rule for a pair of sub-questionnaires entering into the questionnaire obtained in the preceding step. This pair of sub-questionnaires is selected in such a way, that for the first of them the rest is the question with the smaller base, and for the second - the successor with the greatest weight.

Applying the rules above, it is possible to arrive at the optimal questionnaire. Each rule is connected with the requirement of rearranging the structure of the "layer cake". The drawback of the entire procedure of -- optimization is its non-directivity. The attainment of the best result is insured by a certain succession of application of the transformation rules, but the solution of this best succession is not formalizable, which makes the task of optimization of the questionnaire a combinational one in its complexity. It is precisely this that apparently explains that the problem of optimization of classifications and of the structure of the "layer cake", remains to this day a property of the theory of classification and not of the practice of building of effective systems of generalization and classification of descriptions in situational control.

§4.7. Additional Remarks

In working with individual structural descriptions in the form of semantic graphs or formulas of the language of situational control there frequently arises the problem of carrying out theoretical-set operations on these descriptions. In other words, it would be desirable to have the possibility to obtain a unification, a crossing-arc, and a difference for such objects. Figure 4.20 illustrates the carrying-out of operations of unification and intersections for networks that map the recordings in the language of situational control. These operations are carried out in the usual way and do not cause any difficulties. The operation of determining the difference is not that simple. If we have the situation S, shown in Figure 4.21a, and from its description is situation S_1, shown in the same figure by dotted lines, then the questions arises, what to consider as complement to S_1 in S? If the complement is taken to be only that description which is obtained after removal of S_1 with all linkages that existed at the apexes in situation S_1 with the remaining apexes (S_2), then such a complement, upon unification with S_2 will not yield the original description S, because all the linkages between S_1 and S_2 are lost. Therefore, in forming of a complement, it is necessary to introduce fictitious apexes, equal in S_1 and S_2, and such, that upon unification these fictitious apexes are superimposed on each other and vanish, and only the arcs are retained of which the apexes were the markers. On Figure 4.21b, these additional apexes are chairs-hatched.

The introduced operations allow, in particular, to shorten the time for obtaining generalized descriptions for the cases $S_1 \cup S_2$, $S_1 \cap S_2$ and $S_1 \backslash S_2$ had already been found. If one limits oneself to the case on conjunction generalization, to which in the CROSS-method corresponded the idea of a single cause, then the generalized representation of unification, the intersections and the differences will be observed by means of corresponding operations on the generalized representations of those descriptions, which enter into the operation, provided these generalizations were obtained for the class of positive examples of the teaching selection. If we are interested in negative concepts formed on the basis of a class of teaching examples, having a negative character, then, upon unification, intersection and taking differences of positive generalized descriptions for S_1 and S_2, it is necessary to take correspondingly the intersection and unification, illustrated in Figure 4.21b, of generalized descriptions of negative concepts that correspond to the considered positive concepts.

If we are interested in disjunctive generalized descriptions, i.e., descriptions in which common fragments of the descriptions obtained in the course of forming of the generalized description, are unified with each other by the symbol OR, then the obtainment of generalized descriptions for $S_1 \cup S_2$, $S_1 \cap S_2$ and $S_1 \setminus S_2$ is somewhat complicated. For finding $S_1 \cup S_2$ it is necessary to carry out the following operations.

1. Introduce the set $W = D_1 \cup D_2$, where D_1 and D_2 are respectively the generalized descriptions for classes S_1 and S_2 obtained by sets of positive examples as result of procedures described in §4.4.

2. Introduce the set H. In it are included such semantic graphs S^*, which satisfy the following properties. For each S^* in $\mathfrak{g}_1$ and $\mathfrak{g}_2$, which represent generalized descriptions of classes of negative examples from a teaching selection corresponding to classes S' and S'', that the following three conditions are fulfilled.

2a. S^* is isomorphously superimposable on S' under the condition that the types of compatible apexes coincide and that between the weights of these apexes there exist relations or inclusion, investigated in §4.4.

2b. S^* is isomorphously superimposable on s'' under the same conditions. If in addition, S' itself is isomorphously superimposable on S'' under the same conditions, then S'' is included in H.

3. To the obtained sets W and H is applied the method of finding of the generalized description; it can be any of the methods which are described in the same §4.4.

If it is necessary to find the generalized description for $S_1 \cap S^2$ in the presence of disjunctive generalizations, then the procedure of finding them reduces to the procedure of deriving generalizations for $\bar{S}_1 \cup \bar{S}_2$, where $\bar{S}_1$ and $\bar{S}_2$ are respectively the classes of negative concepts for S_1 and S_2. It is possible to construct a procedure of finding of the generalized disjunctive description for $S_1 \setminus S_2$ also. In that case the sequence of necessary steps has the form:

1. Introduce the set $W = \mathfrak{g}_1 \cup D_2$.

2. Introduce the set H, which is constructed as result of the following procedure. A list L is prepared, into which at the initial moment are placed elements from D_1. For the sequential element there is taken from this list the sequential element from D_2, and for the obtained pair is checked the condition of isomorphous superposition of S' onto S", where S' is CΓ from the list D_1 and S" from D_2, provided the conditions are fulfilled formed in item 2a for th procedure of finding generalized disjunction descriptions for unification. If this condition is not fulfilled, then the next element from D_2 is taken. If all the elements from D_2 have already been used up, then the next element from list L is taken, while S' is deleted from L and with it are again compared in succession all the elements from D_2. If however, for some pair S'and S" the condition of isomorphous superposition is fulfilled, then for this pair is formed a new description in which the apex V_{ig} corresponds to a pair of vertices V_i and V_j from S' and S", for which the superposition has taken place, and the weight of this new vertex will be the difference of weights, which corresponded to V_i and J_j. In the semantic graph S', instead of vertex V_i use is made of vertex V_{ij}, and the other vertices are retained. This is done for all pairs of vertices, which leads to the formation of N new descriptions, where N is the number of vertices in S' (in S" there are equally many vertices). All these new descriptions are placed into the list Γ, while S' is deleted from it. If, after the appearance of the first isomorphous superposition for S' all the elements from D_2 are already compared with it, then S' becomes an element of H. The process terminates when the list Γ becomes empty.

The introduction of theoretical-set operations that are applied to the descriptions, turns out to be useful also in problems related to classification. If S^1 and S^2 are classes of equivalence or classes of tolerance, then $S^1 \cup {}^2$ and $S^1 \cup S^2$ are also such classes.

Finally, the operations on descriptions defined by us are most useful in representing information in modern data bases of the relational or network type. In such data bases the operations of intersection, complementation, unification, and difference of situational descriptions are basic and are carried out by the means built into those data bases, i.e. they "cost nothing".

Let us note some other properties of the classes of tolerance, useful in constructing classifiers, for which a partitioning into classes of equivalence with regard to control turns out to be impossible because of a lack of information in the teaching sequence.

Let there be a set of descriptions of situations S and a set of criteria, by which these situations are classified. Two descriptions S_1 and S_2 turn out to be related by the relations of tolerance, if they have some combination of common criteria. If it is specified (fixed) then the maximum subset S', in which any two descriptions are related by the relation of tolerance via the selected group of criteria, forms a *class of tolerance*. Since the relation of tolerance is reflexive, any element S, if it does not resemble any other interior from the selected combination, does still form a class of tolerance with itself, i.e. a class consisting of only that one element. Looking at all the combinations of classifying criteria and forming systems of classes of tolerance by these combinations, we will obtain the space of tolerance. An identical description in this space can enter into several classes of tolerance. In the space of tolerance it is possible to introduce a certain likeness of metric, if one attaches to the classifying criteria some weights that reflect the measure of similarity. In that case, a measure of similarity of descriptions that fell into one class of tolerance can be weighted sum of criteria characterizing this class. It is possible to prove, that in any space of tolerance there exists a set of classes of tolerance which forms a basis. A property of a basis is its mimimality. For any two tolerance descriptions there will be found in a basis a class, into which they both fit. But upon removal of any of the tolerance classes that enter into the basis, the property of completeness vanishes. There will always be found such a pair of descriptions which is tolerant with respect to some criteria combination, for which there will not be formed a class of tolerance in the set of classes obtained from the basis by tossing out some class. The transition to the basis allows to minimize the description in the "layer cake", but does not give the opportunity to effectively utilize the measure of similarity.

And finally, the last remark. In forming generalized descriptions with small volumes of teaching selections an occurrence of "fallarios??" is possible, related with the fact that the decision roles used in classification do not reflect the classification by the technologist - controller. This makes interesting the study of what functions of the criteria does man use in his daily practice when he forms concepts that are new to him. In experiments it was shown[1] that there is a quite definite priority of these functions for a person who has, at least, a school education. The most simple classifying functions for him are the conjunctions of criteria or the disjunctions of individual criteria. With functions containing four - five criteria man works easily; with increase of the number of classifying criteria above that the complexity of the problem of classification begins to increase for him. Much more complex for man are those situations where it is necessary to apply functions in which are simultaneously used the operation of conjunction and disjunction. But still more difficult, practically non-encountered in the practice of human classification, are the cases when the classifying function includes in itself not only the fact of the presence of some criterion but also the fact of the absence of a certain criterion. As an example of such a "non-human" function of classification may sense the function of the form: $y = x_1\bar{x}_2 V \bar{x}_1 x_2$, where x_1 and x_2 are certain criteria.

---

[1] These experiments were carried out by Ye.V. Romanava, N.V. Chudova with students of technical education institutes and with scientific coworkers.

CHAPTER 5

FORMATION OF SOLUTIONS ON CONTROL

If it were so, then that would be alright,
But if alright then that is how it would be.
But because it is not so, then it is not so.
Such is the logic of things.

Lewis Carroll

5.1 Goals, Situations and Plans

In the present chapter we shall look at procedures used during semiotic control for forming a sequence of solutions with the aid of which it is possible to convert a current situation into a certain goal one. In the general scheme of situational control this problem is solved jointly by the Correlator, the Classifier, and the Extrapolator, shown in Figure 1.7. Systems that assure the solution of the given problem are customarily called *planning systems* or *planners*. Planners first form a plan, then check its executability and effectiveness (efficiency), select among those formed the best plan, begin its execution and if necessary correct the plan upon entry of additional information from the object of control and the environment.

Usually two types of planning are distinguished: *planning by states* and *planning by subtasks*. In the first case the concept of state is introduced which is composed from the state of the object of control, the state of the surrounding medium, and the state of the control system. The construction of the plan takes place in the statespace in such a manner that each one time solution on control converts the whole system from one state to another in the statespace. The plan is represented in this case by a certain trajectory in the statespace.

In the second case it is assumed, that there exists a certain selection of moduli, capable of solving definite problems. The planning process consists in the search of such a decomposition of the original problem of which the elements would turn out to be ready-made moduli. Then the plan is a tree-like structure, moving along which it is possible to gradually obtain a solution of the original problem from solutions of more local problems.

In spite of a certain difference in the formulations of these two problems of planning, they can be described by a certain common model, shown in Figure 5.1. On it is shown a network consisting of 15 apexes. The cross-hatched apex 1 is the initial one. It symbolizes either a current situation or a state in the statespace, or the initial problem that must be solved. The apexes 13, 14 and 15 (on the figure they are cross-hatched) symbolize either those goal situations or the final states in the statespace into which it is required to convert the object of control and the control system, or else ready-made moduli for the solution of the problem, which are already available in the system. During planning in the statespace it is necessary to find a path that leads from apex 1 to one of the cross-hatched apexes. In setting up an optimization problem it is possible, for instance, to look for the shortest path of this type. Or, if the transitions along the arcs require a certain expenditure of resources ascribed to these arcs, it is possible to pose the problem of attaining a final situation with a minimal expenditure of resources or to pose more complex problems: to find a path with a minimal expenditure of resources, but such, that the number of steps (time on the way) would be not greater than specified, or, to find a path, minimal in the number of steps, but not minimal by resources, although satisfying the upper limitations on resources.

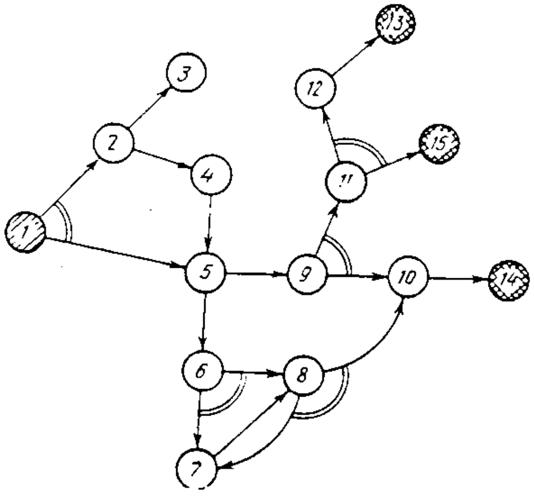


Figure 5.1

During planning by problems it is necessary to go through the entire spread of paths leading from the initial apex labelled 1 to modular cross-hatched apexes that assure the solution of the initial problem.

In this manner, while searching for a plan in the statespace, all branchings in the apexes are considered as alternative. It is necessary to select one(any one) continuation of motion. Whereas while working in the problem-space, all arcs issuing from a given apex must be gone through, if they are interpreted as ratios of the type *part-whole*. Naturally, even then it is possible to consider not the only one decomposition of the initial problem into subproblems, but an alternative decomposition of it. In that case are needed special directives on the character of the arcs issuing from the apexes: are they alternative (OR-arcs) or are they mandatory for joint passage through (AND-arcs). A network with both types of arcs is commonly called an AND-OR network or an AND-OR graph. On Figure 5.1 small double arclets are used which unify among each other the AND-arcs. This means that the initial problem, corresponding to Apex 1, is decomposed into subproblems 2 and 5. Problem 2 reduces to problem 3 or to problem 4, which reduces simply to problem 5. Problem 5 can be reduced either to problem 6 or to problem 9. Problem 9 decomposes into problems 10 and 13. Problem 10 reduces to problem 14 that is already solved in the system, while problem 13 is decomposed into already solved problems 12 and 15.

Thus, in the case of a space of subtasks the plan of solution of problem 1 has the following form: using the already solved problem 14, problem 10 is solved. Using the solved problems 10, 12, 15, problem 9 is solved. Using the solved problem 9, problem 5 is solved. Using the solved problem 4, problem 2 is solved. Using the solved problems 2 and 5, problem 1 is solved. For the case of the statespace on the same network, the plan of attaining the goal-state might be as follows: 1-5-9-10-14. On this example is seen the difference of the above-considered problems of planning. Therefore we shall speak about the problem of planning in general, without indicating its type.

Before considering the problems of planning themselves, it is useful to introduce certain classes of networks, of the type shown in Figure 5.1. Firstly, let us divide such networks into *closed* ones and *open* ones. A network is closed, if the number of its apexes and arcs is fixed and does not change in the process of planning. Such networks are the most simple, but unfortunately, occur rarely in practice. The possibility of constructing a closed network for systems of situational control would signify that the number of complete situations is finite and is enumerated from the start. Then, while analyzing the network based on these situations, it would be possible to search for a plan of transforming a current situation into the one that is of interest to us. Then the arcs could be interpreted as controlling actions under the influence of which the transition is accomplished. Clearly, such a finite model is not realistic. Our knowledge about the object of control and the surrounding medium is usually incomplete and therefore, does not permit us to assume, that the network for planning has been given to us a priori.

In open networks new apexes may appear because of the occurrence of new situations. New arcs may also occur if in the process of functioning of the control system we are able to find new links among the control decisions and the changes of situations.

It is possible to introduce one other characteristic of the planning network which subdivides such networks into *deterministic* and *nondeterministic* ones. In deterministic networks the selected solution, corresponding to some arc, realizes exactly the transition to a new apex, indicated in the network. In nondeterministic networks this is absent. Upon realization of some single-step solution the transition to one or another state is determined either by a certain distribution of probabilities, or is estimated qualitatively with the aid of frequency quantifiers or distribution functions. Such a situation is most typical for control systems of complex objects, when the technology of control depends on the level of training of the controller or on the completeness of the knowledge about control, stored in the memory of the control system.

The classification of the planning networks cited by us is also true for the planning in the problem-space. Openness in such a system is related, for example, to the introduction of new methods of decomposition and a conversion of some problems to type moduli, the solution of which is known, i.e., to the expansion of the library of standard moduli. Nondeterminacy may be treated, for example, as an inexact reduction of the decomposed problem to the totality of these or other subproblems. But, while for planning in the statespace the introduced classes of networks are most important for planning methods, for planning in the space of subproblems such a subdivision is less significant.

It is possible, finally, to divide the problems of planning into *single-level* and *multilevel* ones. Multilevel planning occurs then, when a totality of plans is built on the basis of information on the statespace or the space of subproblems of a different level of detailing. Before constructing an exact and detailed plan, it is always desirable to have a certain assurance that the final goals of planning are attainable. If they are not attainable, then it is desirable not to waste time on planning. It can turn out to be very great. This is where the multilevel plans can be of use. First, at the very highest level, the possibility of planning is established. If it is possible, then a variant of the most crude plan is constructed, the steps of which must be defined more precisely at lower levels, operating with more detailed descriptions of situations or problems. If the initial crude plan passes all levels of planning, then we obtain at the very lowest level the final plan. If however, it is found at some level that the transition planned high-up from crude generalized information cannot be realized, then a return upwards is made where another variant of the plan is selected (if it exists).

In systems of situational control, hierarchical planning occurs in a natural manner because of the hierarchy of the Classifier. The *layer cake* in the Classifier demands, for the possibility of working with it, a similar *layer cake* in the Correlator. Only in that case will all the information stored in the multi-level system of generalized descriptions be successfully utilized. We will describe below how this is done.

5.2 Strategy of Planning

We shall describe here some universal strategies of planning, which were utilized in different planners, both for systems of situational control as well as for solving problems on automation of constructing programs for ECM* or for finding ofa goal-oriented succession of actions for robots. For questions of planning all these problems have much in common, which allows us to consider the problem of constructing a planner in a certain detachment from the pithy formulation of a problem.

---

* Electronic Computing Machines

1. Method of the Direct Wave. It represents the simplest of all methods of planning. Its essence boils down to the following. At the initial moment all apexes are excited which correspond to the initial conditions of the problem (for example the description of the initial situation). Next, if the planning is carried out in the statespace, all the arcs issuing from the excited apexes are excited. If however, planning is carried out in the problem space, then all arcs issuing from excited apexes are excited, but the arcs themselves are marked with special markings. Namely, all arcs related to one alternative variant are tagged with the number of that variant. For example, for the planning network shown in Figure 5.1, upon excitation of apex 1, both arcs are excited. They are labeled with the number of the alternative 1. Next the excitation continues to propagate along the planning network. In the case of planning in the statespace the propagation of excitation will either terminate at the goal vertices, or will cease to propagate, if it is impossible to excite new vertices. For planning in the problem space, the planning will cease upon reaching the same conditions. However, while for the statespace the attainment of even one goal apex serves as a signal about the solution of a problem, for the problem space it is still necessary to check whether all the paths are excited that are related to even one alternative and lead from the initial apexes to the goal one. The searched-for plan is formulated by a reverse check of the excited paths leading from the goal apex, which had been excited, to the initial apexes. In the case of statespace this usually is the path (any excited path, unless an optimization problem is posed), while for the problem space - it is the totality of paths that corresponds to a certain alternative.

Example 5.1. Propagation of excitation by the method of the direct wave for the case of planning in the problem space is shown in Figure 5.2. The planning network is the same as on Figure 5.1. The number of the alternative variant is formed by a gradual increase from left to right. The goal apexes have been excited by alternatives with numbers 121 and 122. For alternative 121 all apexes are excited which are related to it, and hence for this alternative the plan is the sequence of carrying out problems 13, 12, 15, 11, 14, 10, 9, 5, 4, 2, 1. For alternative 122 are also excited all apexes that correspond to it. The plan by this alternative has the form of the following sequence: 14, 10, 8, 7, 6, 5, 4, 2, 1. If for any one alternative, apexes would become excited that happen to be *hanging*, i.e., from which there is no further path toward the cross-hatched apexes, then such an alternative would not correspond to the plan of the solution of the problem. On Figure 5.2 this case is related to alternative 11.

The weakness of the direct wave method is that the excitation propagates along the entire network, which may require a large expenditure of time.

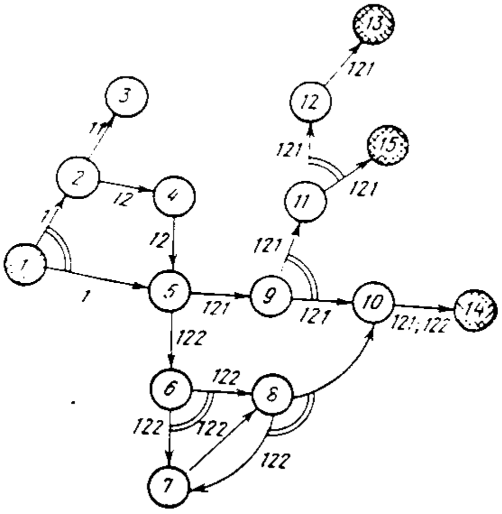


Figure 5.2

2. **Method of the Reverse Wave.** This method is completely similar to the preceding one. Except that excitation begins to propagate from those goal apexes which it is required to attain. Cessation of excitation propagation occurs upon attaining the initial apexes. In planning by states, this occurs upon the first such attainment, and in planning by problems, this occurs upon attainment of initial apexes by all branches of any one alternative.

3. **Method of Counter-Waves.** For accelerating the search process for the path between the initial and the goal apexes it is possible to simultaneously utilize the methods of direct and the reverse waves. Here the process of excitation in the statespace terminates, when there occurs the first encounter of the excited fronts, while in the case of the problem space such an encounter must occur for all the branches of some one alternative. The weakness of the method of counterwaves may manifest itself in the case, when the fronts of the waves diverge, and the problem is posed of determining

the moment of cessation of the excitation process. This situation is shown on Figure 5.3.

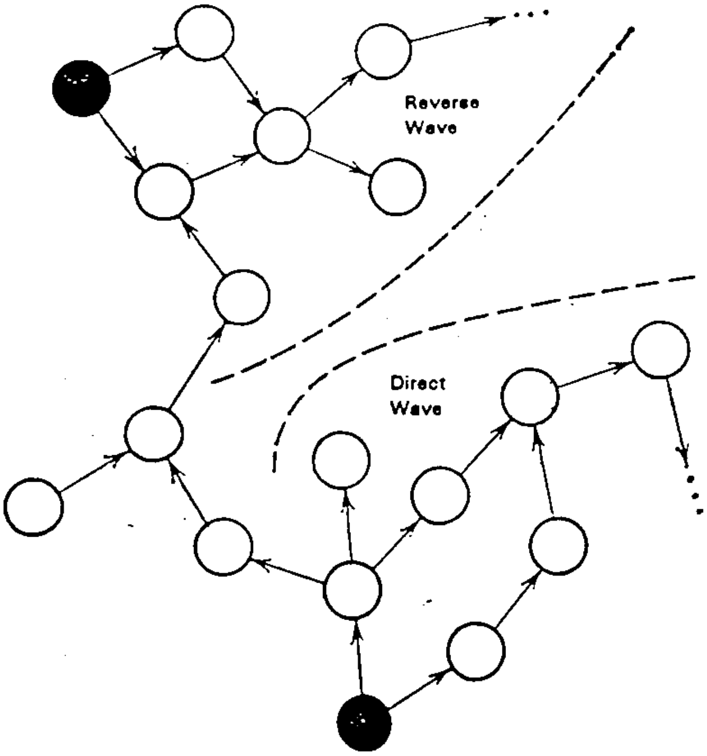


Figure 5.3

4. Method of Local Improvements. This method has been for many years the basis of many planning systems. Its idea consists in that a certain metric is introduced into the planning network. In each apex of the planning network, in the presence of alternative selection of the continuation of motion along the network, it turns out to be possible to estimate, how effective one or another selection leads closer to the goal apex. If such an estimate were true, then the motion from the initial apex to the goal one would not cause any difficulties. It would be sufficient, in each apex with alternative selection, to move along the arc that had the best estimate. Such a solution is possible only in small deterministic closed planning networks. However, for such networks the problem of planning practically vanishes. For any initial and any final apex it is possible to find ahead of time the best path and to write it into the system memory. Upon excitation of the corresponding pair of apexes or of a set of such pairs, the answer is obtained in one step of addressing the information kept in the system memory.

For the case of open networks, however, or for closed networks whose dimensionality does not permit enumeration of all the paths available in the system, there occurs the problem of the type of a searching for a path in a labyrinth* with availability of local information on approach to the goal apex. This idea was used for the first time in the well known program *General Solver of Problems* (GSP). The method of local improvements is based on the introduction of a special metric in the situation space. With its aid it turns out to be possible to measure the distances between situations, to estimate the measure of their nearness. And, in particular, to estimate the measure of nearness of the situation corresponding to the apex of the network in which we find ourselves, to the goal situations attainable from that apex. For each transformation permitted in the planning network (i.e., transition from one situation to another along an arc characterizing this transformation) are indicated those changes in the degree of similarity with the goal situations, which occur after the transformation. And at each step of the method of local improvements is selected that transformation, which maximally decreases the differences of the obtained situation from the goal ones. This method is applicable in planning both in the statespace as well as in the problem space. It can also be used as a basis of establishing the equivalence of descriptions of different situations in the sense of their transformality, one into another, with the utilization of a given set of operators.

Example 5.2. In the program LOGICIAN-THEORETICIAN, that served as prototype of the program GSP, is introduced a system of differences and a system of operators for their removal for transforming the descriptions made in the language of formulas of sentential calculus. Altogether there have been introduced eight distinctions (differences) among these formulas.

1. Distinction P_1. In the tested equivalence of two expressions of sentential calculus $\mathfrak{w}_1 = \mathfrak{w}_2$ (i.e., while trying to establish the equivalence of these two descriptions) there is a distinction between $\mathfrak{w}_1$ and $\mathfrak{w}_2$ in that either $\mathfrak{w}_1$ or $\mathfrak{w}_2$ contains a propositional variable which is absent in the other compared expression. Such a distinction exists for example, in the

---

* In §1.3 such problems were called *labyrinth* ones. There too, were discussed the psychological assumptions that formed the bases of the strategy of local improvements. An finally, there too was characterized the General Solver of Problems (GSP) about which we are yet to speak. Systems based on the modeling principle we shall discuss a little later.

descriptions a and $a\underline{b}$. It is necessary to select a transformation which adds this variable.

2. **Distinction P_2.** It is analogical to P_1, but it is necessary to select a transformation which removes the excess variable.

3. **Distinction P_3.** In one of the formulas ω_1 or ω_2 a certain propositional variable has l entries more than the other. Such distinction is observed, for instance, between the descriptions: $a \vee \bar{a}b$ and ab. It is necessary to select a transformation which decreases the number of entries of this variable.

4. **Distinction P_3.** It is analogical to P_3, but it is necessary to select a transformation which increases the number of entries of this variable into that description where there are fewer entries.

5. **Distinction P_5.** A distinction in the outermost operation contained in the formula. It is observed, for instance, in the descriptions $(a \vee b)(\bar{a} \vee c)$ and $ab \vee \bar{a}c$.

6. **Distinction P_6.** One of the descriptions contains a common sign of negation over the entire description, while the other description does not have it. As an example may serve the descriptions $a \vee b$ and $\overline{a \vee b}$.

7. **Distinction P_7.** A distinction in the character of grouping of the propositional variables in ω_1 and ω_2, as for example, in the descriptions $a \vee (b \vee c)$ and $(a \vee b) \vee c$.

8. **Distinction P_8.** A distinction in the position of subformulas with respect to the external operation. As an example of such a distinction may serve the distinction in the descriptions $ab \to c$ and $c \to ab$.

The distinctions between descriptions introduced by us, can, of course, be combined among each other. But it is important that the above-cited system of distinctions forms a complete base system. Any distinction in two descriptions in the language of sentential calculus is expressed by a certain combination of these base distinctions.

Let us enumerate now those transformation operators serving for removal of distinctions, which were used in the program LOGICIAN-THEORETICIAN. Altogether, ten groups of such operators are available. Operators entering into the first eight groups are symmetrical, they can be applied in any direction. Operators of the last two groups are unsymmetrical, and their use is possible only in the direction shown by the arrow. In the descriptions are adopted the following standard designations:

$$O_1 \begin{cases} a \vee b \equiv b \vee a, \\ ab \equiv ba; \end{cases} \qquad O_2 \begin{cases} \bar{a} \rightarrow b \equiv \bar{b} \rightarrow a, \\ a \rightarrow \bar{b} \equiv b \rightarrow \bar{a}; \end{cases}$$

$$O_3 \begin{cases} a \vee a \equiv a, \\ aa \equiv a; \end{cases} \qquad O_4 \begin{cases} a \vee (b \vee c) \equiv (a \vee b) \vee c, \\ a(bc) \equiv (ab)c; \end{cases}$$

$$O_5 \begin{cases} a \vee b \equiv \overline{\bar{a}\bar{b}}, \\ ab \equiv \overline{\bar{a} \vee \bar{b}}; \end{cases} \qquad O_6 \begin{cases} a \rightarrow b \equiv \bar{a} \vee b, \\ a \rightarrow b \equiv \overline{a\bar{b}}; \end{cases}$$

$$O_7 \begin{cases} a \vee bc \equiv (a \vee b)(a \vee c), \\ a(b \vee c) \equiv ab \vee ac; \end{cases} \qquad O_8 \begin{cases} a \equiv a \vee b\bar{b}, \\ a \equiv a(b \vee \bar{b}); \end{cases}$$

$$O_9 \begin{cases} ab \Rightarrow a, \\ ab \Rightarrow b; \end{cases} \qquad O_{10} \quad a \Rightarrow a \vee b.$$

The transformations of the first eight operator groups do not alter the truth of the expressions of sentential calculus. The last two groups of transformations, serving for introduction and removal of propositional variables are such, that from the truth of the expression to the left of the arrow there follows the truth of the expression to the right of it, but not the reverse.

After introduction of the selection of distinctions and selection of operators for their removal, a special table is set up, in which are listed in the order of priorities assigned to them the operators that remove these or those distinctions. Among the distinctions themselves are also set up priorities according to the importance of one or another of them. In our example the numeration of distinctions defines these priorities, while for the operators there are no priorities. The table itself that relates the distinctions and the operators that remove them, is shown below (Table 5.1). Crosses on it mark the positions in which the application of the respective operator removes the existing distinction.

Table 5.1

	O_1	O_2	O_3	O_4	O_5	O_6	O_7	O_8	O_9	O_{10}
P_1								+		+
P_2								+	+	
P_3			+				+			
P_4			+				+			
P_5					+	+	+			
P_6		+			+	+				
P_7				+			+			
P_8	+	+								

As an example of utilization of the program LOGICIAN-THEORETICIAN let us consider the problem of establishing the equivalence of the descriptions $(a+\bar{b})\bar{c}$ and $\overline{a\bar{b}\vee c}$. For definiteness let us transform the second description, attempting to reduce it to the first one. The main observable distinction consists in the presence of a mutual negation over the formula, i.e., the distinction P_6. For its removal we use the operator O_5 and as a result we obtain the description $(\overline{a\bar{b}})\bar{c}$. Comparing the obtained expression with the untransformable original one, we find a distinction in the fact that a subformula appears that contains a mutual negation, i.e., the distinction P_6 is present again. We apply to this subformula the same operator O_5 and obtain $(\bar{a}\vee\bar{b})\bar{c}$. Now the distinction P_5 is observed to exist for the same subformula, because the more external operation on the compared descriptions coincides (conjunction). To remove the distinction we apply the operator O_6. As a result we obtain $(a+\overline{b)c}$. The coincidence of the descriptions proves the equivalence of both initial descriptions.

A problem for the program LOGICIAN-THEORETICIAN, and in general for all planning methods, based on the idea of local improvements, is the determination of the nonequivalence of the compared descriptions. In that case, at a certain state of transformations, it is necessary to terminate them as being non-prospective. This may not occur automatically either, inasmuch as the operators, as it happened in the example cited above, may yield cyclically repeated descriptions. This situation is similar to one in which, while wandering about in a labyrinth, the people entrapped in it repeat many times the same path. As is known, a way out by marking the traversed path (the thread of Ariadne in the well-known Greek myth about Theseus who searches for a way out in the labyrinth of the Minotaur). For the planning algorithm this means, that it is necessary to remember the pre-history and to execute the operation of return (back-tracking) upon finding a repetition of the transformation in the form of a certain cycle.

In the program LOGICIAN-THEORETICIAN the table of distinctions and of operators of their removal was given in advance. However there exist programs which formulate the distinctions in the description obtained at a given stage, and also in the goal-description*.

5. **Method of Constructing the Proof.** In this approach the descriptions of the initial and the goal situations are viewed as formulas of a certain calculus. All elements entering into the initial descriptions are treated as axioms of this calculus, while all elements forming the goal description, --as formulas which need to be derived from the given selection of axioms with the aid of an available system of rules of deduction. Usually, the calculus of the predicates of the first degree is used as such a calculus. The rules of derivation are divided into two types: *rules of derivation of a universal character*, used in the traditional calculus of predicates, and the rules of derivation that are true only within the framework of that problem region which describes the given object of control, the peculiarities of its functioning, and the specifics of controlling it. The rules of the second type are customarily called the *heuristic rules of derivation*. As procedure

---

* Such a program is, for example, the program FDS of Quinlan and Hunt, which is discussed in the Commentary for the present chapter.

for the search of derivation use is made of some sufficiently powerful procedure of the type of the method of resolutions or of the reverse method of derivation of Maslov. These procedures will be described below.

Usually the technology of planning by the method of proof construction has the following form. First, an attempt is made to find the derivation of the goal description from the initial one. If such derivation is found, then the plan has been constructed. If it cannot be done, as determined on the basis of some principle of stopping, similar to the one indicated for the program LOGICIAN-THEORETICIAN, then a certain heuristic rule of derivation is applied to the initial description, and the process of proof is repeated now with a new initial description. Methods of this type are presently the fundamental ones in constructing the planners in the most diverse areas.

6. Method of Forestalling Planning. This strategy of planning is finding ever-increased application in systems of semiotic control and in planning the actions of intellectual robots. In its basis lies the idea of multilevel planning, about which we spoke already. Before realizing the final planning, for example by searching for a derivation (which may not exist, too), a *generalized planning* is carried out. It differs from ordinary planning by the utilization for the construction of plans not of the complete descriptions of possible paths of attaining goals, but of truncated descriptions or of descriptions obtained from original ones by aggregation, by merging of individual apexes into generalizes apexes. If generalized planning allows construction of a plan, then a transition is made to a precise level of planning, where this generalized plan serves as an initial approximation to the precise plan.

Let us make one other important observation. Two types of planning are possible: the *monotonic* (linear) and the *nonmonotonic* (non-linear). During planning of the first type we approach, at each step of planning, the goal apex, so to speak continually locally improving our position in the network. Procedures of planning of the second type do not possess this property. At individual steps they kind of permit a worsening of the plan and the next apex may turn out to be more remote from the goal apex one than the one previously found. But then later it turns out to be possible to sharply improve the plan, i.e., to approach closer to the goal apex. In our life experience we will find numerous examples when such nonlinear plans turned out to be the only possible means for reaching the goal. Local non-success (sometimes created consciously) allowed them to quickly move towards the designated goal.

In the conclusion of the present paragraph we will note, that the planner can utilize not one fixed strategy, but can have at his disposal a selection of planning strategies, which can assure a more effective functioning of it.

5.3 Utilization of the Method of Resolution During Planning**

The *method of resolution* was proposed in the mid-1960s. It belongs to that group of methods of searching for a derivation which is usually called *methods of refutation*. This appelation is caused by the fact that in these methods, instead of searching for a derivation of a certain postulate F, is being searched a proof of the non-derivability of the negation of the given assertion F. The equivalence of both problems follows from the closed nature of the calculus of predicates of the first order, in which is used the method of resolution, i.e., from the rule of the eliminated third in such calculations the truth of F follows from the falseness of $\bar{F}$.

For applying the method of resolution it is necessary to translate the description of the statement that interests us into the language of logic of predicates of the first order and to represent it in a certain standard canonical form. Such representation is usually called a *standardized* one. It does not contain quantifiers (they are replaced by Scolem functions known from mathematical logic) and represent a conjunction of disjunctions of predicates. As arguments in the predicates may be used constants, variables, and Scolem functions. Standardization is carried out by means of using the system of equivalent transformations of the calculus of predicates of the first order.

Example 5.3. Let there be the description

$$\forall x \exists \tau [[(xr_{52}b) V (xr_{53}b)] (tp\tau)]$$

Here x - some man whose name is not fixed, b - a letter, t - time, r_{52} - to be the sender, r_{53} to be the receiver. The meaning of this description is that each man at some moment of time τ (there will always be at least one such moment) is either a sender of the letter, or its receiver. Let us depart from this notation in the language of situational control to a notation in the language of predicates. Expressions in round parentheses we will replace by corresponding predicate, whose interpretation is defined by the interpretation of the relations within the parentheses. We then obtain

$$\forall x \exists \tau (P_1 (x,b,\tau) V P_2 (x,b,\tau))$$

In the predicate τ is entered as a variable. The quantifier of existence in this description is in the zone of action of the quantifier of generality. Therefore, in eliminating the quantifier we introduce the function of Scolem f_2, while the quantifier of generality is, as always, eliminated without any change in notation. After that our description takes on the form

$$P_1 (x,b,f_1 (x)) VP_2 (x,b,f_2 (x)) .$$

Example 5.4. Consider the predicate notation

$$(P_1 (x) \rightarrow P_2 (x)) (P_3 (x,y) \rightarrow P_4 (y)) .$$

Using equivalence relation of predicate logic, we transform it so as to eliminate from the notation the implication operations

$$(\overline{P_1}\ (x)\ \ VP_2\ (x)\ )\ \ (\overline{P_3}\ (x,y)\ \ VP_4\ (y)\ )\ .$$

Both representations obtained in the last two examples are already standardized. For the resolution method it is also customary to omit the signs of conjunction and to write instead of the conjunction form the set of *disjuncts* that form it. In Example 5.3 there is only one such disjunct, while in example 5.4 there are two. The set of disjuncts for a formula F we will denote by S_F. For negating the formula F it is sufficient to show, that at least one of its disjuncts is underivable. Per force of the conjunctive relation of the disjuncts here, the formula F itself will be underivable. The meaning of all above-said is that there arises the possibility to replace the check of the contradiction of some formula F, which may be quite complicated, by a successive check for contradiction of rather structurally simple formulas-disjuncts. But the method of resolution allows to make even this process, which with a large number of disjuncts can turn out to be cumbersome, more uniform and effective.

In what consists the basic difficulty of checking the falseness or the truth of a predicative expression? In that it is necessary to check the truth of such expressions during combinations of variables, taken from regions of predicate definitions containing such variables. These regions themselves can be infinite. Hence the prospect of direct checking of all possible interpretations of predicate meanings and of their combinations appears to be most sad. Actually, the situation is not so bad. It has been shown that there exists such a region of definition of all predicates entering into a description called the *universum of Erbran* which possesses an extremely important characteristic for us, namely: if in the universum of Erbran S_F is contradictory, then F is derivable.

The Erbran universum is constructed with the aid of a special procedure. If in S_F there are l constants, then a zero set H_0 is formed that consists of these constants. If there are no constants in S_F, then some arbitrary constant is adopted as the zero set of the Erbran base. Next, the sets of the Erbran base are constructed inductively. The H_{m+1} set is obtained by a unification of H_m and these propositional variables $f(w_1, w_2, \ldots\ldots w_n)$, for which w_2 are the elements of the set H_m.

Example 5.5. Consider the predicative description, obtained in example 5.3. If an S_F is constructed for it, consisting of one disjunct, then it is seen that there is no constant in this disjunct. Let us introduce an arbitrary constant, x = a. Then the Erbran universum for this disjunct takes on the form

$$H_0 = \{a\},$$
$$H_1 = \{a,\ f_1(a),\ f_2(a)\},$$
$$H_2 = \{a,\ f_1(a),\ f_2(a),\ f_1(f_1(a)),\ f_1(f_2(a)),\ f_2(f_1(a)),\ f_2(f_2(a)\},$$
$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$

If constant expressions from the Erbran universum are substituted into all disjuncts from S_F instead of the variables, then an *Erbran base* is formed as a result.

Example 5.6. For S_F from example 5.3 and with consideration of the result of example 5.5, we obtain the following Erbran base:

$$P_1(a,\ b,\ f_1(a)),\ P_2(a,\ b,\ f_2(a)),$$
$$P_1(f_1(a),\ b,\ f_1(f_1(a))),\ P_2(f_1(a),\ b,\ f_2(f_1(a))),$$
$$P_1(f_2(a),\ b,\ f_1(f_2(a))),\ P_2(f_2(a),\ b,\ f_2(f_2(a))),$$
$$P_1(f_1(f_1(a)),\ b,\ f_1(f_1(f_1(a)))),\ \ldots$$

If, instead of the variables, some element from the universum of Erbran is substituted into the disjoints, then a certain interpretation S_F is obtained. In our example, in explicit form, are written out three interpretations and is given one expression selected to a fourth interpretation. In this manner, the Erbran base contains all interpretations of the disjuncts of S_F on the Erbran universum. The number of these interpretations can be finite if the Erbran universum is finite, or can be infinite, as in our example.

To each Erbran base can be single-valuedly juxtaposed to a certain semantic tree. Its construction we shall illustrate on an example.

Example 5.7. For the Erbran base built in the preceding example, the semantic tree has the form shown on Figure 5.4.

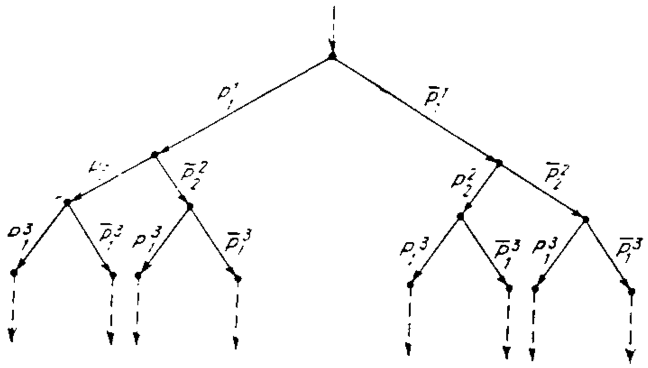


Figure 5.4

On it P_1^j and P_2^k designate the predicates P_1 and P_2 with ordinal numbers j, k, assigned to them by the order of recording in example 5.6. Thus, the predicate P_1^3 has the form

$$P_1(f_1(a),\ b,\ f_1(f_1(a))).$$

In the semantic tree the paths, leading from the root apex downward, correspond to interpretations S_F on the universum of Erbran. The complete tree contains all such interpretations.

Example 5.8. Let $S_F = \{P_1(x),\ P_2(a) \vee P_3(y),\ \bar{P}_1(x) \vee \bar{P}_2(x),\ \bar{P}_1(x) \vee \bar{P}_3(a)\}$. For this set of disjuncts the Erbran universum is very simple. It consists of only one constant , $H_s = \{a\}$. Then the Erbran base is also finite and has the form

$$I = \{P_1(a),\ P_2(a) \vee P_3(a),\ \bar{P}_1(a) \vee \bar{P}_2(a),\ \bar{P}_1(a) \vee \bar{P}_3(a)\}.$$

Or, omitting for ease of uniting the a's

$$I = \{P_1,\ P_2 \vee P_3,\ \bar{P}_1 \vee \bar{P}_2,\ \bar{P}_1 \vee \bar{P}_3\}.$$

To this Erbran base corresponds the semantic tree shown in Figure 5.5.

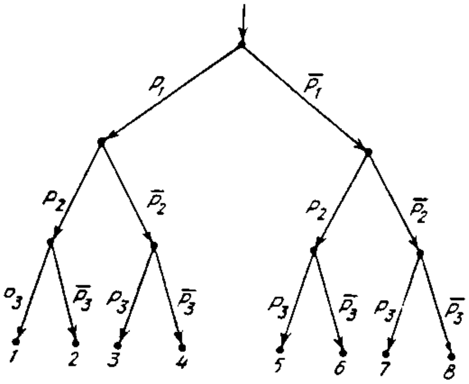


Figure 5.5

Let us consider those eight interpretations which this tree specifies for S_F. For interpretations with numbers 5-8, the set of disjuncts is unsatisfiable already because, with a false P_1, the first disjunct of S_F will be false regardless of interpretations P_2 and P_3. For this reason, the right portion of the semantic tree, shown on Figure 5.5, needs to be analyzed no further (i.e., no further motion is needed downward along the branches of this tree from this apex), which allows to sharply decrease the examination and checking of the realizability for the set of allowable interpretations S_F. thus, it is only necessary to check out those interpretations in which P_1 assumes a true value. With this condition, the first disjunct of S_F is real. For trueness of the third and fourth disjuncts in S_F, it is necessary that P_2 and P_3 should be false simultaneously. But, fulfillment of that condition leads to the falseness of the second disjunct in S_F, whence it follows that all eight allowable interpretations are such that not one of them does not lead to the satisfying of S_F. Therefore, F is non-derivable, while $\bar{F}$ is a derivable formula.

The apexes of the semantic tree in which is first discovered the uselessness of a given interpretation for derivation of F (satisfying F_F) are usually called *unfavorable*. If all arcs issuing from some apex of the tree end up in unfavorable apexes, then it is said that for the given apex the tree is *closed*. If it is closed for the root apex, then this witnesses the non-derivability of F.

The method of resolution coexists in the availability of a procedure of a gradual ascent up the semantic tree, a result of which the unfavorable apexes move up among the levels of the tree. The ascent continues until such time as this is possible. If in this manner it becomes possible to reach the root apex (i.e., to make the entire initial semantic tree closed), then this will prove the non-derivability of F.

This procedure is based on the search of resolvents. A *resolvent* of two disjuncts is called that part of a description which remains if from these two disjuncts a complementary pair is removed. A *complementary pair* is two expressions in which are utilized identical predicates, except that in one expression they enter without negation, and in the other with negation. Arguments of the predicate play no role. As an example of a complementary pair may serve the following two disjuncts:

$$P_1(x) \vee P_2(y) P_3(z); \; P_4(x,y) \vee \overline{P}_2(f(x)) P_3(z).$$

The resolvent of this complementary pair has the form $P_1(x) \vee P_4(x,y)$. What good is a resolvent? If two disjuncts containing a complementary pair are such that they lead to a non-satisfying of S_F, then their resolvent possesses the same property. It is this that is taken into account in the process of proving the non-derivability of F. If a pair of disjuncts does not contain a complementary pair, then the resolvent can be absent. If a pair of disjuncts is such that they both completely form a complementary pair, then their resolvent is equal to zero. The possibility of ignoring, during a search for complementary pairs, their arguments, is based on the fact that a change of variables is the predicate, if it is false in a certain range of values, cannot make it true upon changing the variable and keeping or decreasing this range of values.

Now it becomes possible to describe the essence of the procedures of the method of resolution. Among a set of disjuncts are searched such two disjuncts which contain a complementary pair. To the set S_F is added a resolvent of these two disjuncts. The process is repeated until in the set S_F there appears an *empty* disjunct (i.e., a zero resolvent). This will witness the termination of the process of proving the non-derivability of F.

Example 5.9. Let us consider the set $S_F = \{\bar{P}_1(x) \vee P_2(x), P_1(f(z)) \vee P_2(x); P_2(y) \vee P_3(f(y)), \bar{P}_3(x)\}$. For two of the disjuncts the resolvent is $P_2(x)$. Together with the third disjunct this resolvent generates a new resolvent, $P_3(f(y))$. finally, the resolvent of this secondary resolvent together with the last disjunct gives the zeroth disjunct. Now we can state that the F in our example is non-derivable.

We have presented only the essence of the method of resolution. While carrying out the procedure of consecutive search for resolvents, various unpleasantnesses are possible, when the next-in-line resolvent can't be generated or when the process becomes cyclic. Moreover, with a large number of initial disjuncts (and they may even be infinitely many) special measures are required for accelerating the process of search for an empty disjunct. And, finally, as in all methods of refutation, for the method of resolution there occurs the problem of cessation of search for the zero disjunct, when in reality it is impossible to obtain a refutation because of the derivability of F in the set S_F. There exists a large literature devoted to all these problems, but the discussion of such questions, important for increasing the effectiveness of the work of the planners based on the realization of the given method of derivation search, goes beyond the limits of our book. Those interested in a detailed study of these problems may familiarize themselves with the literature given in the commentary to the present chapter.

In conclusion of the present paragraph, let us note some peculiarities of systems of decision planning based on derivation in a formal system. Firstly, the plan itself, that is formed with the aid of a logical derivation, is realized in a real time and space. And the characteristics of this real environment must be considered during planning. the next example illustrates those unpleasantnesses which may occur if they (the characteristics) are ignored.

Assume that if a certain moment of planning, while analyzing a given situation, it turned out to be possible to work out two mutually independent solutions: U_1 - *to move a crane 200 meters forward along the dock*, and U_2 - *to move the boom of the crane 90° to the left.* However, from the derivation of these two solutions there is no way follows the conclusion of the possibility of their simultaneous realization, although in ordinary deduction systems there follows, from the derivability of U_1 and U_2, the derivability of their conjunction U_1 & U_2. *To move the crane 200 meters forward along the dock, while simultaneously rotating its boom 90° to the left.* In a real situation this may lead to a disaster. It may happen, that along the path of motion of the crane, 50 m from the point where it was, there is another crane, whose body can be touched by the boom of our crane during turning. In other words, planning demands the consideration of the derivability of one or another solution with ties to the dynamics of change of situation resulting from realization of another demand solution (decision).

There is yet another remark, selected to function if plans based on the idea of local improvement of motion toward the goal. In many problems, the goal can be reached only in the case, when at a certain stage of solution it is necessary to kind-of back off from the goal, to make one's situation *worse than it was*, which would allow at the next step to get considerably closer to the final goal. The theory of such non-monotonic plans is still in an embryonic condition and not anything like general approaches have been developed for their formation.

The method of resolution, of course, is not the only method of deduction derivation that can be used in systems of semiotic control. There exist also other planning methods based on logical derivation. Commentaries to the present chapter discuss them in particular. But in all similar methods there is one common defect. It consists in that it is necessary to have available a sufficiently complete description of the problem area to serve as initial axioms with the aid of which derivations will be made. And this causes great difficulties and presumably leads to uncorrectible errors due to the incompleteness of the initial description. Besides, the earlier noted difficulties related to the dynamic nature of the problem areas, also limit the possibilities of using purely deductive systems of forming solutions.

Below, we shall consider that path along which have proceeded systems of planning solutions in practical control systems.

5.4 Functional Models for Planning

In searching for a solution related to the control of some object, it is natural to use those data which are known to the control system, both on the object of control itself as well as on the processes occurring in it. Those data, as was already mentioned before, are kept in the memory of the control systems in the form of a data base. Data on each current situation must be filled with concrete data that characterize it. The base of knowledge is sort of immersed in the base of data, it is *fed* by the data. And depending on the filling by concrete data, the planner works out some or other solutions, expressed in plans of influencing the object of control. On Figure 5.6, all said is represented in the form of a scheme, that has now become standard, of planning the search of solutions. On this figure are not shown the connections that the data base has with the object of control, and the connections of the planner with other subsystems of the control system.

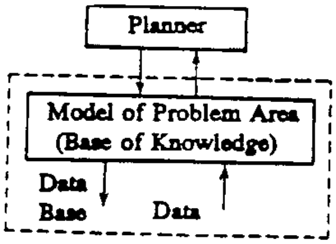


Figure 5.6

To illustrate how the model of the problem area is arranged and how the planner can utilize it in searching a solution to attain the posed control goal, let us consider one class of such models frequently occurring in practice. These are the so-called *functional models*. Every functional model represents a network without orientation including in itself apexes of two types: *descriptors* and *specifiers*. In graphic representation of functional models, the descriptors are commonly designated by circles and the specifiers by rectangles. Each specifier is juxtaposed to a certain functional expression of the type $f(x_1, x_2, ..., x_n) = 0$, where it is assumed that f allows an explicit solution with respect to at least one of the arguments, x_i. In other words, for at least one x_i, this equation can be rewritten in the form: $x_i = f'(x_1, ..., x_{i-1}, x_{i+1}, ..., x_n)$. Such an operation is called an *allowable solution of the specifier.* It is assumed that for one specifier there may exist from 1 to n allowable solutions. With the specifier $f(x_1, x_2, ... x_n) = 0$ are connected n ribs (edges) of the functional model each of which is connected with one of the descriptors of sets $\{x_1, x_2, ..., x_n\}$.Two specifiers in the functional model are connected with each other by a certain descriptor if the latter defines an argument common to both descriptors.

Example 5.10. On Figure 5.7 is shown a fragment of a functional model which is related to the subject area, *Planimetry of a Triangle.* On the figure the elements of the triangle are designated as it is customary in a school course: α, β, and γ are the angles of the triangle; a, b, and c--its sides (the selection between the designations of angles and of sides is illustrated by the sketch in the upper part of the figure); R and r are, respectively, the radii of the circumscribed and the inscribed circles of the triangle; p--the semiparameter of the triangle; and s--its area. Inside each specifier is written the relation connecting various descriptors. The maximal number of possible solutions of each specifier corresponds to the number of edges (ribs) connected with it. For instance, for the specifier with the relation $s - rp = 0$, such possible and allowable solutions are three: $s = rp$; $r = s/p$, and $p = s/r$.

Functional models similar to the one just considered in example 5.10 allow the description of an entire complexus of procedures related to the determination of some descriptors through others. For this, it is sufficient to only specify the initial set of descriptors and the goal descriptors and to find, with the aid of the planner, such allowable solutions for the specifiers which would form an oriented path, leading from the initial descriptors to the goal ones. The search for and a path (at least one of the possible ones) constitute the fundamental problem of the planner.

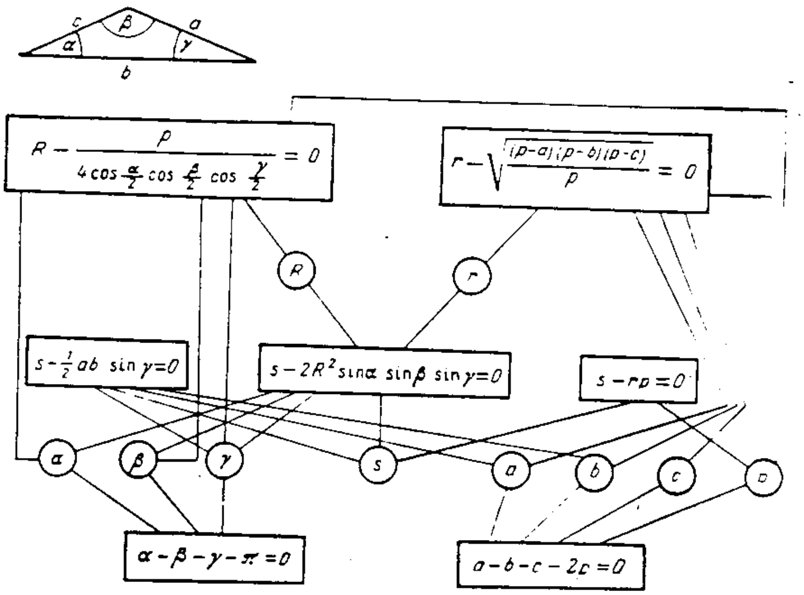


Figure 5.7

Example 5.11. Assume that the task *to find the area of a triangle, if its sides a, b, and c are given* has been inputted to the entry of the planning system. According to the text of the task, the planner excites in the functional model the descriptors that correspond to a, b, and c and begins the search for a path leading to the goal descriptor S. The search proceeds in accordance with one of the strategies about which we spoke in section §5.2. Assume, for definitiveness, that this shall be the strategy of the direct wave. Then the process of search of the plan will proceed as is shown in Figure 5.8. At the first step, the planner becomes convinced that the allowable solution is provided only by the specifier marked by number 1 on the figure. A solution for it gives the possibility to find the descriptor P, if values for a, b, and c are known. Next, the solution for the specifier marked by the number 2 becomes possible, thus allowing a determination of the value of descriptor 2. Finally, at the third stage of planning is resolved the specifier marked by the number 3, which leads to the determination of the value of descriptor S. Since the descriptor S is the goal one, the planner concludes his work.

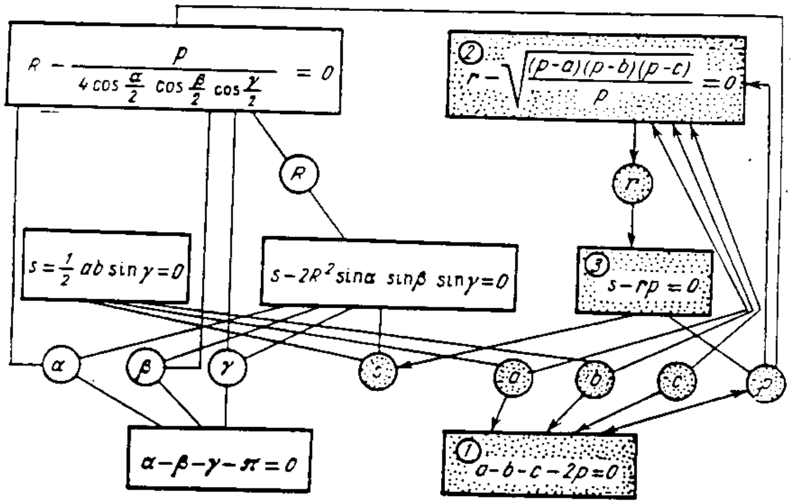


Figure 5.8

After having found the oriented path, leading from the initial descriptors to the goal ones, it is necessary to formulate a computing program. This is possible if in the memory of the control system there is an earlier prepared selection of standard program moduli, each of which corresponds to one or another allowable solutions for a defined specifier. With availability of a selection of such moduli, the path found by the planner indicates the sequence of carrying out the moduli one after another.

The described procedure of planning solves only the first part of the problem of solution search--it formulates a program, after which it is necessary to carry out the filling of the program with concrete data. At that stage it also becomes necessary to turn to the base of knowledge about the problem area, for it is only in it that the regularities are contained that allow judgment on the allowability of one or another combination of data. For example, if we utilize the functional model described in examples 5.10 and 5.11, then upon input at the entry of the planning system of the problem *to find the area of a triangle, if its sides are given as a = 100, b = 1, and c = 20*, a program will be constructed corresponding to the path found in example 5.11. Next, it will begin to be carried out for concrete values of the sides of the triangle, specified in the task. But, the execution of this program will not be concluded successfully. An emergency stoppage will occur, because during the calculation of the radius of the inscribed circle, in accordance with the specifier marked by the number 2 on Figure 5.8, a negative expression will be obtained under the radical sign.

In order to exclude such cases, regularities are allowed for ahead of time in the model of the problem area; the carrying out of these must be obligatory in work with concrete data. In our case, in the model of the problem area, should be stored the following regularities: *The sum of the lengths of two sides of a triangle is always greater than the length of the third side; the difference of the lengths of two sides of a triangle is always (less) than the length of the third side; the sum of the angles of a triangle is always equal to 180°*, and so forth. These preliminary checks should expose all cases of unallowability of carrying out programs with such fixed initial data, which violate the fundamental regularities of the problem area.

The totality of the deductive conclusion assured by the planner on a functional model, the description of the model itself and its related program moduli and regularities of the given problem area along with procedures for checking them form what is now customary to call *intellectual package of applied programs*. It is namely as such a package that the correlator manifests itself in structural control. Its main component is a solution of logico-transformational rules:

$$S_1;\ S_2 \underset{D}{\Longrightarrow} S_3.$$

Here, S_1 is a description of a fragment of the current situation the presence of which defines the applicability of the logico-transformational rule; S_2 is a description of the fragment being transformed; and S_3 is the resultant description of the new fragment of description. If S_1, S_2, and S_3 are considered as descriptors, and $\underset{D}{\Longrightarrow}$ as a certain specifier, then it is easy to establish a correspondence between the functional models and the selection of logico-transformational rules. Therefore, it can be assumed that traditional methods for situational control have anticipated the idea of working with functional models.

In real problems are often found cases, when transitions between states in the state space or between subproblems in the subproblem space are undetermined, which reflect the incompleteness of our knowledge or the possibilities of such transitions. In this case, the arcs of the network, or which planning is carried out, are weighted, for example, by values of a function of affiliation. A general evaluation of the formal found plan is calculated from the estimates of these values or different arcs by various methods. In particular, a general evaluation may coincide with a minimal evaluation encountered on the given path, or may be the arithmetic mean of these evaluations.

5.5 Scenarios and Logics of Action*

In searching for decisions in systems of situational control are used not only functional models but also other methods of finding solutions. In §5.4 we had already mentioned logico-transformational rules, laying in the base of many real systems of situational control. The use of similar rules assumes that all necessary information for deriving a solution is stored in a specially organized base of knowledge (see Figure 5.6). This organization consists in breaking up the entire base of knowledge into separate, possibly overlapping, regions called *spheres*. Each sphere contains knowledges among which exists a close semantic, pragmatic, or situational connection. A definite name is assigned to a sphere. Each fact, stored in the memory of the control system, contains information on the names of those spheres of which it is a part. Moreover, the fact has its own, belonging only to it, name. Finally, for the parts are given lists of selections by which it is tied to the other parts, and of the types (names) of these selections.

Depending on the concrete utilized base of data, into which is immersed the base of knowledge, the formalization of the fact may be different. For example, it can be described in the form of a *role frame*.

Example 5.12. In creating a control system for loading and unloading railroad cars at a railroad station, it might be necessary to describe in the form of a fact the knowledge on the technology of unloading the cars. This knowledge can be represented by the following role frame:

{<UNLOADING><PLACE><Value of slot><TYPE OF CAR><Value of slot>
<TYPE OF CARGO><Value of slot><PRIORITY><Value of slot>
<TIME OF ARRIVAL><Value of slot><RECEIVER><Value of slot>}.

Such a role frame may have an inlaid structure. In other words, individual slots corresponding to roles defined by their names, may themselves be the names of other role frames. For instance, the Slot PLACE can itself be specified in the form of the following frame:

{<PLACE FROM WHERE><Value of slot><WHERE TO><Value of slot>
<TYPE OF CAR><Value of slot><MEANS><Value of slot>}.

In its turn, the slot TYPE OF CAR can also have the form of a certain role frame.

{<TYPE OF CAR><MEANS OF LOADING><Value of slot><MEANS OF UNLOADING>
<Value of slot><CONDITIONS><Value of slot>}.

In filling the slots of the cited role frames with concrete data from the data base, concrete frames are formed that result from those prototype frames. For example,

{<UNLOADING PLACE><Platform No. 2><TYPE OF CAR><Type Δ>
<TYPE OF CARGO><Sand><PRIORITY><First><TIME OF ARRIVAL>
<25 December 1985><RECEIVER>< Brick factory No. 68?};

{<PLACE FROM WHERE><Platform No. 1><WHERE TO><Platform No. 2>
<TYPE OF CAR><Type Δ><MEANS><Diesel locomotive 02652>};

{<TYPE OF CAR><MEANS OF LOADING><Any><MEANS OF UNLOADING>
<Any><CONDITIONS><Length of platform not less than 50 m,
height of platform not more than 0.5 m>};

{<SAND><MEANS OF LOADING-UNLOADING><Transporters, type 1,4,7;
crane type 6; manual loading-unloading><LOCATION OF MECHANISMS>

<Transporters type 1, 4 on platform No. 2; transporter type 7 and crane type 6 on platform No. 7>}.

What do these descriptions provide? They contain in them the information that on 25 December 1985 a railroad car type 4 with sand which arrived for use at brick factory No. 68, is situated near platform No. 1. However, at that platform there are no mechanical means suitable for unloading sand. therefore, with the aid of diesel locomotive 02652, the car can be brought up to platform No. 2, at which there are available transporters of types 1 and 4 suitable for unloading sand. In carrying out this decision (solution) it is further necessary to check the fulfillment of a condition related to the type of the car and the length of the platform No. 2 and its height. If these conditions are fulfilled, then a solution of the problem has been found. If they are not fulfilled, then it is necessary to find another variant of unloading related, for example, to the moving of the car with sand to platform No. 7.

As is seen from the cited example, two types of descriptions are stored in the system's memory: intentional and extentional. The intentional descriptions are descriptions of pure knowledges about the object of control and of the processes occurring in it. A role frame in which the roles are only named but not filled represents an example of such an intentional description. In filling out the roles, however, there originate a number of extentional descriptions--concrete facts which in their totality give a description of a current situation.

Both intentional as well as extentional descriptions may be connected among each other by various selections. In our example, for greater clearness, these selections,as well as the names of spheres into which these facts enter, were not pointed out. Most important for intentional descriptions are the classifying relations, characteristics for the classifier (i.e., the selections of the type *class-element of class*. These relations describe the structure of the object of control and the structure of the processes of decision-making on the control of the given object. On the other hand, the relations on the set of extentional descriptions are situative. They connect individual concrete facts among each other into a single situation.

Connecting together descriptions of two types, it is possible to obtain a scenario for decision-making.

Example 5.13. On Figure 5.9 is shown a fragment of a scenario for control of water usage and water consumption in a certain region. On the first level of this scenario, seven classes of situations are singled out, during the occurrence of which control is required. These classes are interpreted in the following manner: 0 - a change in the operation of the system of water usage is required because of 1.1 - state of weather conditions, 1.2 - agricultural water use effects, 1.3 - water transport effects, 1.4 - daily communal use effects, 1.5 - effects of hydroelectric structures, 1.6 - industrial effects, 1.7 - effects of the tourism industry. Next is analyzed only that part of the scenario which is related to the necessity of control as a result of requirements posed to the control system by the daily communal usage. The remaining scenario apexes on Figure 5.9 are interpreted in the following manner: 2.1 - increase in quantity of water used, 2.2 - increase in quantity of effluent water and deterioration of its

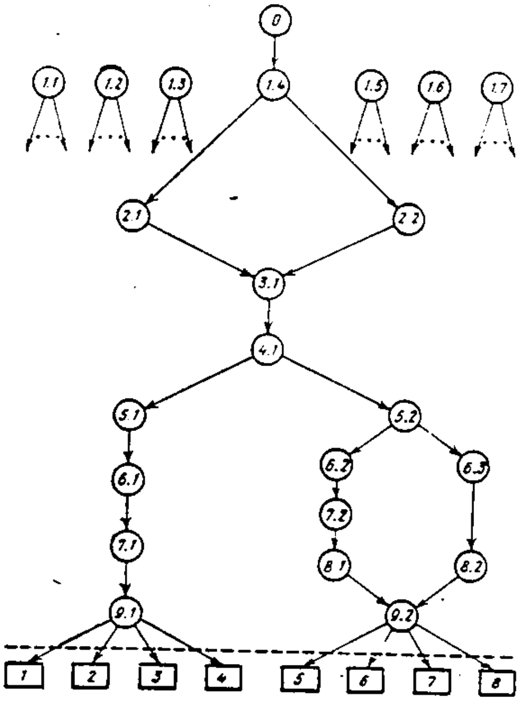


Figure 5.9

quality, 3.1 - analysis of reasons of balance violation, 4.1 - prognosis of further changes, 5.1 - changes have a temporary character, 5.2 - changes have a permanent character, 6.1 - modeling of the effect of changes on the state of the environment, 6.2 - determination of the leaking quantity of water, 6.3 - modeling of the effect of change of quality of run-offs or the state of the environment, 7.1 - evaluation of the modeling results, 7.2 - analysis of the effect of water economy balance on the perspective (outlook), 8.1 - estimate of possibility of increasing water usage, 8.2 - evaluation of modeling results, 9.1 and 9.2 - determination of solutions (decisions) on the basis of the carried-out analysis.

Next follow the solutions (decisions) themselves, indicated on Figure 5.9 by numbers: 1 - Introduction of a water economy regime, 2 - increase of load or water economy facilities, 3 - putting into action of reserve powers, 4 - construction of water conduits, 5 - modernization of the equipment of available water purification facilities, 6 - construction of new facilities for purification, 7 - redistribution of run-offs and their further utilization, 8 - increase of load or purification facilities and utilization of reserve powers.

Thus, the scenario contains in itself the most heterogeneous elements related to the checking of conditions, requiring the intervention of a control system, analysis of the current situation, forecasting of its development, and the adoption proper of one or another solution (decision). But the scenario examined in our example 5.13 is only the upper level of the decision-making system. At the lower levels the solution must be defined concretely. Concretization may require several steps, including procedures of analysis, invitational modeling (extrapolation), or optimization. In a number of systems of situational control created in our country, systems of decision-making have three levels of concretization. At the uppermost level there occurs the elicitation of the conflict situation that requires making decisions on control. At the second level are divided into parts all allowable solutions. An extrapolation is carried out for them and, if possible, an optimization. Finally, at the third level the found solution scheme (intentional description of the solution scenario of the type of the role frame from example 5.12 or of a logico-transformational rule) is filled with a concrete content from a data base and is transformed into an extensional description of the scenario of decision-making, which is realized command by command. The transition from the second level to the third is attended with a possible nonsingle-valuedness of the concrete scheme of realization of an intentional description because of the nonsingle-valuedness of the working sequence of logico-transformational rules. To prevent these conflicts, special metarules are introduced with the aid of which are selected allowable sequences of application of logico-transformational rules.

Let us consider one more approach related to the making of multistep decisions, when the logic of actions is utilized on the apparatus for making decisions. Let us consider a fragment of such a logic related to a partial orderliness in time of actions that constitute in their totality the realization of a certain solution. Let d_i be elementary actions, for which the control system has standard techniques of realization; D_j - complex actions consisting of chains of actions, d_i; A_k - subjects capable of carrying out actions d_i from a certain totality of such actions. Here the role of subjects can be played by both humans as well as various possible mechanisms and devices. By T_k we shall designate the time intervals during which subject A_k acts. Finally, O_m are objects that experience the effect. Let us now introduce the main basic triplets characteristic for the logic of actions in time. In labeling the relations entering into them we will make use of the relations from Table 2.1 and Table 2.2, namely:

$d_i r_6 D_j$ - the elementary act d_i is a part of act D_j,
$A_k r_{76} d_i$ - subject A_k carries out act d_i,
$A_k r_{76} D_j$ - subject A_k carries out act D_j,
$d_i r_1 T_k$ - act d_i is realized in interval T_k,
$D_j r_1 T_k$ - act D_j is realized in interval T_k,
$A_k r_1 T_k$ - subject A_k acts in time interval T_k,
$O_m r_1 T_k$ - object O_m experiences effect over interval T_k,
$d_i r'_{26} d_l$ - acts d_i and d_l can be carried out in arbitrary order and without being adjacent,
$d_i r_{26} d_l$ - acts d_i and d_l must be carried out adjacent to each other and in the order indicated,

$D_j r'_{26} D_m$ - acts D_j and D_m can be carried out adjacently in any order,

$D_j r_{26} D_m$ - acts D_j and D_m must be carried out adjacently and in the order indicated,

$d_i r_{21} d_l$ - acts d_i and d_l are carried out simultaneously,

$D_j r_{21} D_m$ - acts D_j and D_m are carried out simultaneously,

$d_i r_{71} O_m$ - act D_i is carried out on object O_m.

Let us also mention examples of rules of deduction used in such logic of actions.

1. Each subject may carry out a certain action (d_i or D_j) only over that time interval on which it can do so. For actions d_i, this rule assumes the form

$$(A_k r_1 T^*)(A_k r_{76} d_i)(d_i r_1 T^{**}) \Longrightarrow (T^{**} r_{29} T^*).$$

2. The time during which a subject does an act on some object, and the time during which an object experiences the effect of this act (d_i or D_j) must be equal. For action d_i this rule takes on the form

$$(d_i r_1 T^*)(A_k r_{76} d_i)(d_i r_{71} O_m)(O_m r_1 T^{**}) \Longrightarrow (T^* r_{21} T^{**}).$$

3. Elementary acts that enter into a complex action must have intervals of realization that are contained in the interval of realization of the complex action. For act d_i this rule has the form

$$(d_i r_8 D_j)(d_i r_1 T^*)(D_j r_1 T^{**}) \Longrightarrow (T^* r_{29} T^{**}).$$

4. For immediately adjacent actions (d_i or D_j) the intervals of their realizations should be in relation of adjacency. For d_i this rule has the form

$$(d_i r_{26} d_l)(d_i r_1 T^*)(d_l r_1 T^{**}) ==> (T^* r_{26} T^{**}).$$

5. If a certain action (d_i or D_j) is completed before another action (d_l or D_m) then the same relation exists also between the intervals on which they are realized. For d_i this rule has the form

$$(d_i r_{22} d_l)(d_i r_1 T^*)(d_l r_1 T^{**}) ==> (T^* r_{22} T^{**}).$$

6. If two actions (d_i or D_j) are carried out simultaneously, then their intervals of execution coincide. For d_i this rule has the form

$$(d_i r_{21} d_l)(d_i r_1 T^*)(d_l r_1 T^{**}) ==> (T^* r_{21} T^{**}).$$

Similarly, it is possible to construct a fragment of the logic of actions related to the spatial realization of actions. For this, it is necessary to select from Table 2.1 and Table 2.2 the corresponding spatial relationships.

Such fragments of the logic of actions can be utilized in constructing complex solutions from elementary ones and for checking the conditions of their combination (matching) in time and space.

5.6 Extrapolation of Solutions

The complexity of the objects, for which normally are used methods of situational control, requires the making of one or another decisions during the search of a solution or the control of possible consequences. The problem of imitational modeling is not related to any particular problem tied in specifically with situational control. There exist many methods of such modeling which allow extrapolation of the dynamics of the states of an object and of the processes occurring in it. Special languages of imitational modeling have been developed and program packages have been created, allowing realization and imitation.

However, in utilizing methods of the semiotic type and of those languages of situation descriptions that are used in them, special requirements arise toward modeling systems. These requirements are tied to the fact that it would be desirable to obtain in the ideal a forecast of the development of events at the level of descriptions of those situations which may occur in the future. In other words, it is desired to obtain an extrapolation in the form of an inverted tree, as shown in Figure 5.10. Its root corresponds to a situation (description of a situation) at the object of a given moment of time. If the solution is planned as $D^* = d_1d_2d_3$, then the subsequent levels of the tree show those situations into which the object can get a result of realization of that particular solution. The branching of the tree corresponds to that uncertainty with which it is possible to conceive the process of development of events. Around each situation located at the end branches of the tree, estimates g_i are written that characterize the possibility of such a consequence (they are indicated by shaded circles). These estimates may have an export, a probability, or a diffused character.

If, in the initial situation, in addition to solution D*, it is possible to use certain other solutions, then an imitational process is constructed for all of them, generating its own tree of the same type as in Figure 5.10. Next, following some decision rule, are evaluated the estimates g_i obtained as a result of modeling and that solution D* is selected for which the decision rule gives the best result.

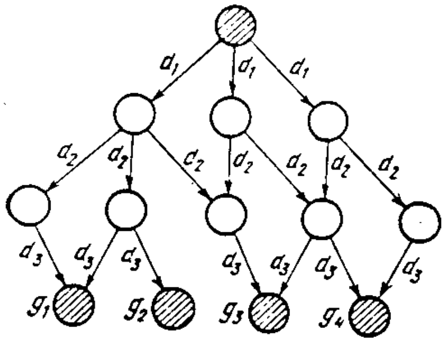


Figure 5.10

An important characteristic of the above cited method during modeling is that each time we have a description of the obtained situation, and hence can classify it by means of the classifier and to evaluate its conflictness or nonconflictness for controlling the object. This leads to imitational modeling as though on two levels: at the level of the dynamics of the controlled object itself and at the level of description of operations forming on it with consideration of possibilities (for instance, resources) of the control system itself. It is in this two-step manner that the values of estimates g_i are formed.

The following peculiarity of imitational processes in systems of decision-making on during control, in the base of which lie logico-linguistic models, consists in that of interest are not only the statistical data accumulated in the process of modeling (for instance, average times of expectation or lengths of queues of objects waiting for servicing), but also the dynamics of one concrete situation, and here are varied not only (and not how much) some numerical parameters, but also structural descriptions (for example, relations realized among separate elements entering into the description of the situation). All this leads to the necessity in constructing an extrapolator for systems of situational control to utilize possibilities contained not only in traditional languages for imitational modeling, but also in languages characteristic for describing and transforming the descriptions of situations.

And, finally, one more peculiarity of the extrapolator--its operation must be closely tied with the operation of other blocks of the system of situational control and first of all with the operation of the classifier and of the system of selection of solutions. this indicates that the imitational process in systems of situational control proceeds in such a way that motion along a tree of the type shown in Figure 5.10 is being carried out with consideration of not only the object itself, but also with involvement of procedures of situation classification that come up in the course of modeling, and with utilization of procedures of solution forming based on the results of this classification.

In §2.7 we introduced the concept of objects of control in the form of a discrete situational network (DSN). Such a model allows a very simple demonstration of the essence of the process of imitational modeling in the extrapolator of the system of situational control.

Let us recall that a DSN includes in itself elements of four types: sources, risks, passive solvers, and active solvers. Along a DSN are moving, during discrete moments of time, objects, each of which is specified by a selection of characteristics. Characteristics can be of two types: *static*, characterizing the essence of the given object, and *dynamic*, related to temporal estimates, ascribed to the object during a given concrete process that is being realized or a DSN. Each element entering into the DSN can be represented by some deterministic or probabilistic automaton. These automata can even be schematically realized, but in systems of situational control it is preferred to have their programmatically realized models.

Example 5.14. Let us consider the DSN shown in Figure 5.11. As we see, the DSN has one source, the functioning of which generates objects of two types: $stock_1$ and $stock_2$. The objects enter into the passive solver and then into a solver in which are combined the passive and the active solvers AR+R.

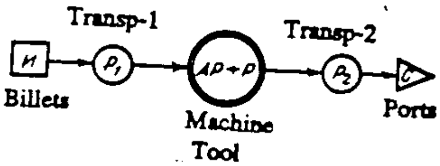


Figure 5.11

Finally, after that the objects get into one more passive solver and flow off. The first and the last passive solvers imitate the transporters, Transp-1 and Transp-2, feed the stocks to the solver AR+R, which imitates a machine tool, and pass the ready parts to the bunker, shown on the DSN as the sink C. The processing of the object $stock_1$ and $stock_2$ on C machine tool requires a certain time, which alters the dynamic characterization of these objects. The working on the machine tool itself changes their static characteristics.

In imitating the process of motion of the stock and their working on the machine tool we shall introduce the following processes.

1. Setting Up. Corresponding to this process is the logico-transformational rule (LTR) of the following form: *If* the machine tool is free and the $stock_1$ is on transp-1, *then* it is necessary to place the object $stock_i$ on the machine tool after taking it off the transp-1, and then switch the machine tool into the occupied position.

2. Switching On. The corresponding LTR has the form *if* the machine tool is occupied, *then* switch it to the switched-on state. Note that the unoccupied position has an indefinite, in time, interval of actualization. The occupied state, however, depending on the type of the object $stock_i$ (i = 1, 2 for our example) is actualized for a definite period of time (for example, for i = 1 for 10 minutes, and for i = 2 for 45 minutes).

3. Feeding the Stock. The corresponding LTR has the form: *If* on transp-1 there are not objects of the type $stock_i$, *then* transfer the control to the source and wait for the appearance of the first object of $stock_i$, after which the functioning of the source ceases until the vacating of transp-1.

Such conditions of operation of the LTR for the stock feed process may, of course, not be fulfilled in practice. The source may function independently of the presence or absence of stock on the transporter which supplies the parts to the machine tool.

4. Switching Off. The corresponding LTR has the form: *If* the counter of the astronomical time, for the occupied position, has reached a limit corresponding to the object $stock_i$, *then* switch the machine tool into the free position and place the ready part onto the transporter trans-2.

5. Feeding the Part. The corresponding LTR has the form: *If* the part is on transp-2, *then* switch it on and, after a specified astronomical time, switch it off.

This rule corresponds to the transposition of the part from the machine tool to the exit bunker (sink of the DSN).

The above-described five LTR contain in them the conditions of applicability, the fulfillment of which permits changes in the situation, described in these LTR. The conditions of applicability are introduced by the marker *If*, and the operators of transformation--by the marker *then*.

The above-cited very simple example shows that the description of a current situation and the description of the LTR for a slightly complex real object will be extremely cumbersome, while the work itself on the formulation of the needed list of LTR represents a very uneasy task. In real objects there will be moving along the DSN many very multitype objects, and the LTR for their processing in active solvers will be numerous and interconnected by the conditions of their applicability.

Example 5.16. For illustrating the above, we will cite as an example Table 5.2, in which are enumerated types and numbers of objects moving along a DSN and forming its active solvers, for the Kalinnigred Sea fishing port.

Here we do not separate the objects and the active solvers, because depending on desire they can be transformed one into the other. an analysis of this table shows how complex may turn out to be the LTR system which takes into account all kinds of objects, resources, and process-synchronizing parameters. An estimate of the number of LTR, characteristic for modeling systems of such objects, lies within the bounds of 200-400. The number of decisions adopted in processing ships in a port of the sea-fishing port type is also very large. In practice, each such decision (solution) is a multistep one. The following fundamental operations are parts of it: loading, unloading, trans-shipment, towing; dislocational operations; positioning by the first board, positioning by the second board, pulling over, manipulating, mooring, remooring; filling operations: fuel, oil, fuel-oil, water, settling, conveying; auxiliary operations: carbonation, deactivation, disinfection, surface conditioning, washing and drying, packaging, ship repair, cleaning, fire prevention, diver operations, passenger operations. For cargoes, there is also an entire spectrum of operations. A number of operations (for example, loading or towing) themselves consisting of a number of simpler operations. These operations are interconnected by the scenarios that determine the necessary and allowable sequences of execution of individual operations. As an example may serve the scenario shown on Figure 5.12.

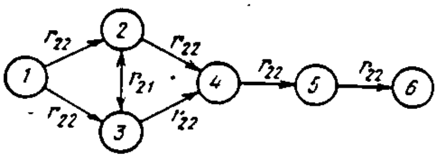


FIGURE 5.12

Table 5.2

Objects and Active Solvers	Number Encountered in the Framework of a Single Situation
Pier Cranes	2
Gantry Cranes	36
Moorings	
Industrial Harbor	8
Lumber Harbor	7
Oil Base	2
Integrated Works	1
Gasification or Aeration	1
Electric Loaders	115
Automotive Loaders	28
Ships of the Harbor Service	45
Refrigerator	3
Covered Storage	13
Open Storage	30
Railroad Track	7
Workers	2459
Brigades	32
Ships of the Kaliningrad Registry	more than 400
Types of Ship Winches	25
Types of Twin Decks	126
Types of Cargo Spaces	63
Types of Railroad Cars	17
Assortment of Cargoes	about 500
Types of Automotive Vehicles	7

Here, as always, the relation 2_{21} is the relation *simultaneously*, while the relation 2_{22}--*to be earlier*. Numbers denote the following operations: 1 - transmittal of the cargo plan, the date of arrival, 2 - preparation for the processing of the ship, 3 - transference of the ship into the port, 4 - familiarization of arrival, 5 - processing of the ship, 6 - formalization of the cargo documents and of departure. Of course, each apex of the scenario is itself the name of some scenario, and such a *nesting* of scenarios can be sufficiently deep. As parameters that change the form of the scenarios or their sequence, act the concrete characteristics of the types of ships, moorings, and other elements enumerated in Table 5.2.

Besides solutions which are connected with the consideration of the totality of scenarios defining the technological requirements toward the realization of solutions and entering, as a rule, into conditions of applicability of some or other LTR, there are also imitated and adopted solutions determined only by those control criteria which the control system itself is guided by. These criteria, as a rule, are contradictory, because they touch upon the interests of different administrations. The plan of unloading and of cargo shipment is always born as a result of a certain compromise, evaluated in the process of imitation of the consequences of the adopted decisions. Behind the simplicity of the final solution, as illustrated for example by Figure 5.13, there is concealed the realization of a very cumbersome procedure of imitation and decision-making.

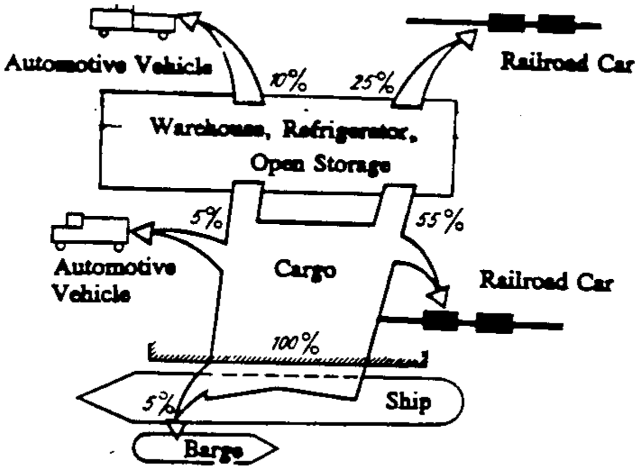


FIGURE 5.13

5.7 Peculiarities of Reasoning in a Human

In the process of preliminary filling of the memory of a control system of semiotic type and in the process of teaching, when the technologist-controller forms for the system the fundamental strategies of control, and also the teaching samples, he involuntarily introduces into this information those peculiarities of his thinking which are inherent in him as a human.

In the present paragraph we will try to enumerate the methods of thinking readily utilized by man in the course of his reasonings and not fitting into the traditional framework of logic. The purpose of the present section is to alert the specialist on control, who frequently makes light of that level of strictness which in reality characterizes the position and the pronouncements of the expert-controller. A second purpose of what is expounded below--is to offer a selection of some recommendations on the construction of systems which mistake human methods of reasoning. The necessity in this is called for not specifically by problems of imitation of human abilities by artificial systems, but by the fact that in systems of communication between a human and an intellectual system it is necessary to create a level of mutual understanding and trust occurring during the communication of two people.

Subsequent exposition will have a concise (recapitulative) character. A more extensive consideration of some or other problems connected with the methods of reasoning of man, can be found in books and articles discussed in the commentaries to the present chapter. The order of discussion of the peculiarities of reasoning and of the conclusions made by man on the basis of these considerations, reflects that significance which can today be ascribed to them in solving management problems.

1. Scaleability of Deductions. In §4.2 we spoke much about scales forming the space in which is realized the classification of objects and phenomena of the surrounding world in a human. Understandably, his reasonings and conclusions made by him cannot ignore these scales. However, the scaling of the conclusions occurs, as a rule, sort of outside their verbal expression, remain *outside the sequence*, although it is recognized in a necessity. But sometimes such scaling manifests itself in an explicit form. Let us have two scales: a *scale of causes* and a *scale of consequences*. On each scale is defined a certain relation of order. Let us designate by $\alpha_1 < \alpha_2$ and by $\beta_1 < \beta_2$ that consequences β_1 is *less than* cause β_2. Assume it to be known that the conclusion $\alpha_1 ==> \beta_1$ is justified. What can be said about the conclusion $\alpha_2 ==> \beta_2$, for which it is known that $\alpha_1 < \alpha_2$? We are interested in the case when from the information given it is possible to make either the conclusion $\beta_1 < \beta_2$ or the conclusion $\beta_2 < \beta_1$. In the first case we will say that the scale of causes and the scale of consequences are *coordinated*, and in the second that they are *inversely coordinated*. It appears that for man these two cases of coordination of the scales of causes and consequences occur practically always.

Let us consider the following example. Let scale α be the scale of criminal offenses. Then the relation *less than* may be interpreted as *to be lighter from the standpoint of inflicted damage*. Let scale β be the scale of punishments determined for crimes. Clearly, the scales α and β are coordinated. For a graver crime, a graver punishment is prescribed. If scale β is taken as a scale on which is marked the probability of correction of the criminal, then it is possible to assume that it is inversely coordinated with scale α.

2. Incorrect Use of Causes. In conclusions of the type $\alpha ==> \beta$ in man's reasonings, cases are quite often formed when their realization is modified by certain conditions. In other words, the conclusion has the form $\gamma:\alpha ==> \beta$, where the predicate γ characterizes the possibility of the conclusion $\alpha ==> \beta$. However, very often γ is *forgotten* and the conclusion $\alpha ==> \beta$ is made *according to tradition*. That is the layer of knowledge which it is customary to call *parental*. Such knowledge is absorbed by man uncritically without an evaluation of its reliability. In primitive reasoning, in myths, in rituals, such conclusions are used everywhere. And in our days' expressions like *It is customary to consider that...* or *In accordance with the established tradition...*, and so forth, demonstrate that some method of reasoning. But frequently such conclusions turn out to be false because of the falseness of the predicate γ in a given situation.

Another example of an incorrect use of causes is the case when the *childish* layer of knowledge participates in a conclusion. Conclusions based on such knowledge, as a rule, base themselves on reasonings of the type *I desire β but in order for β to be, α must be.* Let α be then.* Next follows the conclusion $\alpha \Longrightarrow \beta$, although α may not even exist.

Finally, very often as cause β is selected not that true reason which causes β, but some quasi-reason that lies on the surface. This is the most frequently encountered human error. It is precisely because of it that Napoleon thought that steamships could not float: They must sink, because the iron from which they are made sinks in water.

3. Personal Disinterest in Consequences. Very often in the conclusion $\alpha \Longrightarrow \beta$ man is rather indifferent as to just what it is that is being concluded from α. Such a situation is in a certain cause juxtaposed to reasonings based on the childish layer of knowledge. Often in reply to the question *What follows from this?* or *Why is this done?*, there follows the answer *I don't know*. Rationalism that is usually ascribed to human thinking and to human reasonings is in fact not encountered so often. For this reason many conclusions β, which follow from α, have actually a chance and nonpragmatic character.

4. Traductivity of Conclusions. In human practice the conclusion is most often a traductive one (i.e., a conclusion from particular to particular). In child reasoning, the percentage of such conclusions is close to 100. As their examples may serve reasonings of the type *This uncle is bad; he has a rough voice*. or *This thing is red; that means it tastes good*. But conclusions of this type are also found very often in reasonings of grownup people.

5. Subjective Utilization of Quantifiers. Quantifiers are always subjective and reflect the personal experiences of a person using them in reasonings. We spoke about this already when discussing quantifiers. Here, we will just point out a somewhat different aspect of this subjectivity, which plays an important role in evaluating the information obtained from experts-managers. In their utilization of quantified reasonings it is always

---

* Conclusions of this type belong to the mathematical folklore.

necessary to remember that the evaluation of a quantifier is enormously affected by the personal gain aspect of the user. While, for instance, many people purchase tickets of a money-goods lottery or a sports-lotto, there would remain few takers if their conditions were altered as follows: With that probability, close to zero, with which the grand prize is won, the winner of that ticket would be subjected to a very severe penalty, while in all cases corresponding to loss, a small win were paid. Because, when the consequences for a person may be severe, the probability of the risk increases subjectively. Roughly speaking, with increase of possible unpleasantnesses the a priori estimates of the subjective probability of expectation of this event grows. The traducive conclusions with and quantifiers is a common phenomenon.

6. **Scenario Conclusion.** In human practice are often found conclusions based on certain temporal series, causal scenarios, or scenarios of other types. An example of such conclusions can serve *The electric train is approaching a station, the platform is deserted, many people are crowded by the car doors. This means that there will be many people on the platform. The lecturer entered the auditorium and started to write formulas on the board. Hence, the students are writing in the abstracts that which is written on the board.* We note that in such reasonings great importance is played by role structures about which is spoken in §2.2. In particular, it is precisely the knowledge of the role structure that explains such a conclusion: *The man was chopping the wood with an axe; the axe was sharp. Hence, it was easy to chop.*

7. **Paradoxicality and Absurdity of Reasonings.** This property of human reasonings is less significant for us than the others. Hence, we confine ourselves only with examples that show that a human in a number of cases commits a violation of even fundamental laws of logic. Most often this occurs in proverbs and sayings. Thus, in the well-known saying *He didn't burn it, but scorched it* or *Not half of forty, but twenty* is violated the logical law of identity. In the sayings *He would have been a clever man had he not been a fool* or *Ulya is full when she doesn't want to eat* from the standpoint of logic shouldn't be contained any information, for it is asserted in them

that A = A. For us are especially interesting the quasiconclusions based on the fact that with a false premise it is possible to describe the implication any consequence. *If you find a man on a cow, then a mane on a cow, then a mane will have horns.* An interesting modification of such conclusions is the saying *The gelding is bay, but there is no fur on it* and the beautiful text of D. Harms constructed on the same idea *There was a redheaded man who didn't have any eyes and ears. He didn't have any hair either, so that he was called redheaded conditionally.*

D.A. Pospelov, Situational Control, Theory and Practice, Nauka Publishers, Moscow, 1986, 288 pp. PSL #45364

APPENDIX

HISTORICAL SKETCH OF THE DEVELOPMENT OF SITUATIONAL CONTROL

(pp. 254-258 of the book)

Cart after cart, the wagon train is huge,
And it's a wonder to behold
That which in one day, just one day
Became extinct and antiquated
They cart like rotten wooden logs
Like junk, like rusty metal scrap,
Erroneous representations
and piles of imaginary axioms.

L. Martynov

In the early 1960's bionics was in vogue. Its main purpose -- to use that which nature has realized in plants and in animals for creating technical devices -- seemed very attractive. Conferences and symposia on bionics were convened. Biologists, engineers, mathematicians and cyberneticists learned to listen to and to understand each other. Formation of specialists went on who would be capable to work at the juncture of very diversified scientific disciplines.

And against this background there naturally arose the question of cooperation of cyberneticists and psychologists. To utilize the knowledge accumulated by the psychologists in the creation of intellectual systems was a no less interesting goal when compared with the purpose of bionics. In analogy with bionics the new direction was labeled psychonics. This appelation turned out to be not very successful and did not take root. But a seminar with that name functioned for about 10 years. It was organized at the Moscow Energetics Institute. Its main membership, the composition of which varied with the passage of time, were specialists working at MEI and the Institute of General and Pedagogical Psychology. The constant leader of this seminar was the author of this book.

It was precisely at that seminar where for the first time were formulated the principles of the modeling method of problem solution by man. The renunciation of the theory of labyrinth reasoning prevailing in psychology, which we talked about in Chapter 1 of our book, the transition to a theory of thinking where was dominant the procedure of constructing a labyrinth that lead often a search to a solution, was repeatedly declared in the addresses of V.N. Pushkin and his students. Ideas of structurization of the initial description of a problem and the interconnections of this structure with a structure of the aimed-at situation were understandable to cyberneticists. Instead of heuristic procedures of sorting, imitating a search in the labyrinth of possibilities, the new approach required the creation of procedures based on work with structurized descriptions. And this in turn required the creation of new models of representation of the objects of control and an elaboration of special languages for describing the situations being formed at the object of control and in the systems of controlling it. These two problems received their solution in the first dissertations defended in the area of situational control in 1967. In the work of Zh. Zhelezov was developed a theory of discrete situational networks which served as a good model of objects of control for many subsequent developments on situational control. Discrete situational networks were described in Chapter 2 of the present book. Another graduate student of the author of the present book, Yu.I. Klykov, basing himself on the then known language of RX-codes (brief data on it are given in Chapter 2 of this book), had developed a special model language, named by him later as the "language of syntagmatic chains". This language is described in Chapter 2. For many years it became the basic language of describing situations and decision making in situational control.

The results of 1967 laid a firm foundation for the concept of model control, actively developed by V.N. Pushkin and D.A. Pospelov. The appearance in 1972 of their joint monograph "Thinking and Automata" completed the initial stage of the development of the new approach.

The term "situational control" itself did not become formed suddenly. In the studies of the initial period the new approach to control was first called the modeling approach. Next, the term "situational model" was formed. And only with the appearance in 1971 of the article of D.A. Pospelov "Principles of Situational Control" (Izvestiya Akademii Nauka SSSR, series Technical Cybernetics, No. 2) did this term displace all the others. Simultaneously it became clear that the principles of situational control (described above) are not necessary and universal for the methods utilizing the logic-linguistic models. This served as an impetus for the origination of a new term: "semiotic control" (sometimes referred to as "semiotic modeling").

For the first time this term was used in the title of a seminar organized at the Moscow House of Scientific-technical Propaganda in 1971. The seminar, the leaders of which were L.T. Kuzin and D.A. Pospelov, had the title "Semiotic Methods of Control in Large Systems". In the introductory report of the seminar organizers "Problems of Semiotic Control" for the first time was formulated the statement that the use of logical and linguistic means in problems of control, which are considered in situational control, and problems arising in other fields, have much in common. In other words, the ideas used in situational control have actually a more general character.

A result of such a view on logical-linguistic models was the origination of two new scientific directions, different from situational control, but utilizing in their studies many ideas first originated in the midst of specialists in the area of situational control. One of these directions is the "characterizational control", actively developed by V.A. Gorbatov and his students. They created some principally new methods of solving traditional problems of synthesis of discrete control systems, based on the principle of "semantic equivalency", structured descriptions of schemes and demands for their functioning.

At the Moscow Engineering Physics Institute, under the guidance of L.T. Kuzin, began to be actively developed ideas of control in ACS* of different nature, based on logical-linguistic models different from situational control. Such studies naturally lead to the solution of problems

---

* Automated Control Systems

of building intellectual banks and languages of representation of knowledge and their manipulation.

By 1971 there was formed a body of investigators who saw in the method of situational control a means of overcoming those difficulties which they encountered in designing systems of control of complex objects. This body was not too big, for the unfamiliarity of the new approach based on other ideas, sharply different from traditional methods of automatic control, created a psychological barrier, which can scarcely be sensed now. But at that time there were many more critics than followers. In 1970, L.S. Zagadskaya defended a dissertation in which the new method served as a basis for the system for controlling cargo handling operation in a sea port. This program little resembled the simple programs of 1967 and 1968, in which the method of situational control was used for directing the movement of ships in sluice-lock portions of canals. In the program system of L.S. Zagadskaya there were already present all the basic blocks of situational control: analyzer, classifier, correlator and extrapolator.

There began to form local territorial schools in the area of situational control. In Odessa the school was headed by L.S. Zagadskaya, in Groznyy - by L.A. Afonin, in Kaliningrad - by V.F. Ponomarev. Somewhat later appeared groups of specialists on situational control also in other cities: Riga (under the direction of Ya.A. Gel'fandbeyn), Ustinov (under the direction of A.Yu. Leviatov), Baku (under the direction of R.E. Gadzhiyev). These bodies were aimed at utilizing the method of situational control for solving real problems posed before them. In Odessa and Kaliningrad this was the problem of operational dispatcher control of loading/unloading operations in a seaport, in Groznyy - a problem of operational control of moving around among bore holes of special installations for carrying out plugging operations and of boring installations among points of boring, in Riga was being solved a complex of problems related to ACS in civil action, in Ustinov - operational control of production areas at instrument building facilities, in Baku - operational diagnostics of illnesses and prescribed therapy. Such a wide spectrum of applications naturally required in each concrete case to develop and modify those methods and models of situational control which were born in the initial body of investigators.

The participants of these studies were full of enthusiasm and optimism. It appeared to them that the principles of situational control would be realized in the systems created by them easily and effectively. And it is not their fault that the matter turned out to be considerably more complex than they thought. The then existing native ECM* in their technical possibilities and mathematical support did in no way come up to the requirements which were posed to them by methods of situational control. This did not become obvious at once, either. It was necessary to experience through practice what it was that was concealed behind this new approach which they were using. For this reason the first program systems remained purely experimental.

In the same 1971, so important for situational control, the resolution was adopted on conducting All-Union symposia on situational control of large systems. The first such symposium took place in Odessa in 1972. From that time on Odessa became for a long time the traditional place of conducting symposia; the latest (sixth by count) occurred in 1981.

In the plenary report, made at the first symposium by Yu.I. Klykov, D.A. Pospelov and V.N. Pushkin, which was titled "Why is Situational Control Needed?" its authors tried for the first time to clearly indicate the place of situational control among other methods of control of complex objects. Replying to the question posed in the title of the report, the authors formulated the thesis that situational control, in distinction from other methods of control, solves not only the problem of control on the basis of a certain already known model of the controlled object, but also builds that very same model. The construction of the model of the object of control, the construction of the procedure of controlling it, and the search for goal-oriented solutions on control are inseparably connected in the method of situational control. And it is precisely this characteristic of it that allows to hope for its applicability in those cases, when the complexity of the object of control does not permit to build its formal mathematical model and to pose the problem of control in the traditional spirit.

After the first symposium followed two more, in 1973 and 1974. At the second symposium the central problem discussed was the problem of

---

* Electronic Computing Machines

construction of a multi-level Classifier. The idea of a "layered cake" that originated back in the model of a gyromat (a device capable of tuning its structure to the structure of the problem being solved), first proposed at a seminar on psychonics in the mid-1960s, received at that symposium support and approval on the part of its participants. There were proposed several principles of formulating such a Classifier, described in the pages of the present book.

At the third symposium at the center of attention were problems related to the creation of mathematical support for systems of situational control. Precisely this question was then the most urgent, because on its solution depended the very possibility of applying the ideas of situational control in control systems of real objects. The participants of the discussions on these problems pointed out the necessity of using languages capable of working with data having a complex internal structure and oriented toward such progressive for the time ECM as the BESM-6 and YeSRYAD. Not a few suggestions were voiced, many of which were realized considerably later in languages of representation of knowledge of the type of relational and frame languages.

The necessity in expanding the sphere of application of the method of situational control, the transmission to other organizations of the methods and programs created by that time, posed the problem of development of special typical techniques. This work, actively conducted in the early 1970s, was crowned with success. There were created the first typical techniques for the design of situational models of control (L.S. Zagadskaya, O.V. Sokolova, 1973), formation of the model of describing the object of control in the language of syntagmatic chains (V.F. Ponomarev, A.V. Kolesnikov, 1973), the development of program support for models of situational control (O.V. Sokolova, L.S. Zagadskaya, Yu.I. Klykov, 1973). These first techniques, in spite of their imperfection, played a definite role in the development of the method and its extension to practical program models.

In 1974 appeared the first collection of articles on situational control, after which followed three more collections (1975, 1977, and 1980).

All the collections were published by the Scientific Council on the Complex Problem "Cybernetics" under the Presidium of the ASUSSR*.

For coordination of all studies on situational control and utilization of it for solving practical problems the project "Situation" was established, which was organizationally included in the Scientific Council on the problem "Artificial Intelligence" of the Committee on System Analysis under the Presidium of the ASUSSR. Starting in 1975, the working council of this project has coordinated all studies in the area of situational control, carried out in the USSR.

In 1974, appeared the first monograph on situational control (Klykov, Yu.I. Situational Control of Large Systems - M. Energiya, 1974), which played along with collections on situational control an important role in the popularization and dissemination of the method.

In 1975 the method of situational control in its main parts was already formed, there appeared a stable, generally accepted terminology, was created a complex of basic methods and procedures. Beginning from that year the study groups, in which was being used the new approach, began to transition from experimental systems and model problems toward the construction of real systems of control, based on the ideas of situational control. This required the coordination of efforts of a number of study groups and an exchange of the experience accumulated during the creation of real systems. Thus was born the idea on conducting conferences, devoted to the "technology" of systems of situational control. Altogether were conducted five such conferences (Groznyy, 1975, Ustinov, 1977, Kaliningrad, 1979, Kuybyshev, 1982, Kaunas, 1982).

The number of organizations interested in applying the method of situational control began to grow rapidly. For popularizing the method are being carried out study schools, at which are read lectures on all basic problems, connected with situational control and of its use in developing various real systems (Moscow, 1974, Baku, 1975, Tashkent, 1977, Moscow, 1979, Kuybyshev, 1982). With the same purpose the publishing house Znaniye in 1975 published the popular book of D.A. Pospelov "Large Systems, Situational Control".

---

* Academy of Sciences, USSR

By that time it became clear that aside from traditional problems of dispatching and operational control of technical and organizational systems, the situational control could be used even in those cases, when traditional approaches do not yield a solution because of the dimensionality of the solved problem. In those cases, instead of an exact solution, the obtainment of which becomes practically impossible, the method of situational control allows to obtain approximate solutions, acceptable in their quality. The first examples of such problems were the problem on the distribution of programs in complexes of ECM (D. Boyev, 1974) and the problem on the cutting of the graph with weighted edges into components with a minimization of the summary weight of the edges that fell on the section line (N. Georgiyeva, 1977). Both problems are well known in discrete mathematics. For them were suggested many methods of solutions. But because of the time wasted for the search for their solutions, these methods are not useful in operational systems, whose problems include a segmentation of the programs and a distribution of the segments onto parallel working ECM. The use of the methods of situational control allowed to create procedures for the solution of the indicated problems executable in the real time scale.

In the mid-1970s Yu.I. Klykov began to actively use the method of situational control during the modeling of the process of teaching humans. With time under his guidance there originated a school, in which there began to develop methods of modeling the learning activity of man and methods of controlling it.

A characteristic feature of the end of the 1970s became the essential expansion of the understanding of situational control. Factually situational control began to be considered from unified positions of semiotic modeling and of control. The reason for this was the impetuous development in the theory of artificial intelligence of that area which carries the name "representation of knowledge". Situational control had anticipated by ten years the development of that area, being first to begin work with structured information. In situational control were created the first models of representation of knowledge and of languages of representation and of manipulation of knowledge. For this reason new trends in the area of artificial intelligence and the active development of the theory of semantic networks and frame concepts were adopted by study groups which had collected a large experience in the area of situational control "on the run" and allowed to adopt the new ideas and to use them in their development. Equally painless occurred the transition to programming languages of the new type (LISP and its developments FRL). In this plan, the specialists who mastered the principles of situational control, found themselves at the forefronts of studies in artificial intelligence.

Beginning in 1980, it is practically impossible to separate the development of the situational control proper from the development of semiotic controls which organically included in itself its historical predecessor. And most symbolic sound the titles of the plenary reports made at the VI World Symposium on Situational Control of Large Systems (Odessa, 1981) "Situational Analysis" (L.S. Zagadskaya), "Situational Programming" (V.S. Lozovskiy), "Scheme of Outputs in Semiotic Models" (D.A. Pospelov).

The logic of science development is harsh and does not submit to anyone's will. Situational control has played the role of a catalyst, having evoked in life new ideas and principles of constructing models of the reality and activity of man, that have so widely developed in the wide flow of studies on artificial intelligence. But one wants to believe that in the theory of semiotic models that is being born before our eyes, the results obtained in the depths of situational control, shall occupy their proper place.

SITUATIONAL CONTROL
Theory and Practice

BIBLIOGRAPHY

Chapter 1.

1.1 Pospelov D.A., *Logico-Linguistic Models in Control Systems,* Moscow: Ehnergiya Publishers, 1981, 231 pp.

1.2 Zakharov V.N., Pospelov D.A., Khazatskiy V.Ye., *Control Systems: Task, Design, Realization,* Moscow: Ehnergiya Publishers, 1977, 423 pp.

1.3 Moiseyev N.N., *A Mathematician Asks Questions...* Moscow: Znanie Publishers, 1974, 190 pp.

1.4 Kini R.L., Raiffa H., *Decision Making with Many Criteria: Preferences and Substitutions,* Moscow: Radio and Svyaz' Publishers, 1981, 560 pp.

1.5 *Problems of Management and Planning-By-Objective,* ed. G.S. Pospelov, Moscow: Nauka Publishers, 1981, 460 pp.

1.6 Blauberg I.V., Yudin E.G., *The Composition and Nature of the Systems Approach,* Moscow: Nauka Publishers, 1973, 270 pp.

1.7 Tsetlin M.L, *Studies on the Theory of Automata and on the Modeling of Biological Systems,* Moscow: Nauka Publishers, 1969, 316 pp.

1.8 Varshavskiy V. I., *The Collective Behavior of Robots,* Moscow: Nauka Publishers, 1973, 407 pp.

1.9 Woodwords, R., "The Solution of Problems by Animals," in the book: *Selections on General Psychology: The Psychology of Thinking,* Moscow: Moscow State University Press, 1981, pp. 228-234.

1.10 Pervozvanskiy A.A., *Search,* Moscow: Nauka Publishers, 1970, 262 pp.

1.11 Slagle J., *Artificial Intelligence,* Moscow: Mir Publishers, 1973, 319 pp.

1.12 Nilsson N., *Artificial Intelligence,* Moscow: Mir Publishers, 1970, 270 pp.

1.13 Reitman U., *Knowledge and Thinking: Modeling at the Level of Information Processes,* Moscow: Mir Publishers, 1978, 400 pp.

1.14 Newell A., Simon H., "GPS - A Program That Simulates Human Thought," in the book: *Computers and Thinking*, Moscow: Mir Publishers, 1967, pp. 283-301.

1.15 Newell A., Shaw J., Simon H., "Empirical Study of Machines 'Logistician - Theoretician'; An Example of A Study of Heuristics," in the book: *Computers and Thinking*, Moscow: Mir Publishers, 1967, pp. 113-144.

1.16 Newell A., Shaw J., Simon H., "A Program for Playing Chess and Problems of Complexity," in the book: *Computers and Thinking*, Moscow: Mir Publishers, 1967, pp. 33-71.

1.17 Pospelov D.A., Pushkin V.N., Sadovskiy V.N., "Toward a Definition of the Subject of Heuristics," in the book: *Problems of Heuristics*, Moscow: Vyshaya Shkola Publishers, 1969, pp. 5-19.

1.18 Pushkin V.N., *The Definition of Thinking in Large Systems*, Moscow and Leningrad: Ehnergiya Publishers, 1965, 375 pp.

1.19 Pushkin V.N., *Heuristics - The Science of Creative Thinking*, Moscow: Political Literature Publishing House, 1967, 270 pp.

1.20 Pushkin V.N., *Psychology and Cybernetics*, Moscow: Pedagogika Publishers, 1971, 230 pp.

1.21 Pospelov D.A., Pushkin V.N., *Thinking and Robots*, Moscow: Sovetskoye Radio, 1972, 222 pp.

1.22 Pushkin V.N., *The Psychological Possibilities of Man*, Moscow: Znanie Publishers, 1972, 63 pp.

1.23 Pushkin V.N., "The Psychological Bases of Constructing Teaching Systems," in the book: *Questions of Cybernetics. Man-Machine Teaching Systems*, Moscow: The Scientific Council for Complex Problems of 'Cybernetics' on the Presidium of the Academy of Sciences of the USSR, 1979, pp. 8-39.

1.24 Pushkin V.N., "Principles of Mathematical Modeling of Knowledge and Ways of Constructing Artificial Intelligence Systems,"in the book: *Questions of Cybernetics. Situation Management. Theory and Practice.*, Moscow: Scientific Council for Complex Problems of 'Cybernetics' on the Presidium of the Academy of Sciences USSR, 1980, pp. 5-33.

1.25 Pospelov D.A., "Principles of Situational Control",. Bulletin of the Academy of Sciences USSR, Technical Cybernetics, 1971, No. 2, pp. 10-17.

1.26 Klykov Yu.I., Pospelov D.A,, "Situational Control in ASU," *Controlling Machines and Systems*, 1972, No. I, pp. 27-34.

1.27 Klykov Yu. I., Pospelov D.A., Pushkin V.N., "Why is Situational Control Necessary?" in the book: *Questions of Cybernetics. Situational Control. Theory and Practice.* Moscow: Scientific Council for Complex Problems of 'Cybernetics' on the Presidium of the Academy of Sciences USSR, 1974, pp. 7-15.

1.28 Pospelov D.A., *Large Systems. Situational Control,* Moscow: Znanie Publishers, 1975, 62 pp.

1.29 Klykov Yu.I., *Situational Control with Large Systems*, Moscow: Ehhergiya Publishers, 1974, 134 pp.

Chapter 2.

2.1 Stepanov Yu.S., *Subjects, Predicates, Prepositions*, Moscow: Nauka Publishers, 1981, 369 pp.

2.2 Sil'dmyaeh I.Ya., "Knowledge, Meaning, and Thought," in the book: *Scientific Proceedings of Tartu State University,* No. 551, Tartu: 1981, pp. 121-127.

2.3 Sil'dmyaeh I.Ya., "On Cognitology," in the book: *Scientific Proceedings of Tartu State University*, No. 594, Dialogue Systems and the Representation of Knowledge, Tartu, 1981, pp. 126-133.

2.4 Fillmore C., "The Case for Case," in the book: *What's New in Foreign Linguistics*," Issue 10, *Linguistic Semantics*, Moscow: Progress Publishers, 1981, pp. 369-495, "The Case for Case is Reopened", ibid, pp. 496-530.

2.5 Martynov V.V., *Semiological Bases of Data Processing*, Minsk: Nauka i Tekhnika Publishers, 1974, 191 pp.

2.6 Martynov V.V., *Universal Semantic Code*, Minsk: Nauka i Tekhnika Publishers, 1977, 190 pp.

2.7 Martynov V.V., "Man-Machine Methods of Solving Problems in the USK [Punched Card Reader] System," in the book: *Questions of Cybernetics. Man-Machine Training Systems*, Moscow: Scientific Council for Complex Problems of 'Cybernetics' on the Presidium AN USSR, 1979, pp. 49-55.

2.8 Apryesyan Yu.D., *Experimental Research of the Semantics of the Russian Verb*, Moscow: Nauka Publishers, 1967, 251 pp.

2.9 Revzin I.I., *Modern Structural Linguistics: Problems and Methods*, Moscow: Nauka Publishers, 1977, 283 pp.

2.10 Boguslavskiy I.M., "Semantics of the Particle Только [Tol'ko]" in the book: *Semiotics and Data Processing*, No. 14, Moscow: VINITI, 1980, pp. 134-158.

2.11 Pshenichiaya L.Eh., Skorokhod'ko Eh.F., "Units of Informational Language in Comparison with Units of Natural Language," in the book: *Transactions of the III All-Union Conference on Information Retrieval Systems and Automated Processing of Scientific-Technical Information*, Vol. 2, Moscow: VINITI, 1967, pp. 187-191.

2.12 Skorokhod'ko Eh.F. and others, *Information Retrieval System BIT*, Kiev: Naukova Dumka Publishers, 1968, 120 pp.

2.13 Klykov Yu.I., "Model Language," in the book: *Transactions of the III All-Union Conference on Information Retrieval Systems and Automated Processing of Scientific-Technical Information*, Vol. 2, Moscow: VINITI, 1967, pp. 360-371.

2.14 Klykov, Yu.I., "The Language of Situational Control," in the book: *Semiotic Methods of Control in Large Systems*, Moscow: Dom Nauchno-Tekhnicheskoy Propaganda, 1971, pp. 76-81.

2.15 Kuznetsov I.P., *A Mechanism of Processing Semantic Information*, Moscow: Nauka Publishers, 1978, 179 pp.

2.16 Nisenbaum L.B., "Several Models Generalizing the Understanding of the Language of Situational Control," in the book: *Questions of Cybernetics. Problems of Artificial Intelligence*, Moscow: Scientific Council for Complex Problems of 'Cybernetics' on the Presidium of the Academy of Sciences USSR, 1980, pp. 131-136.

2.17 Uyemov A.I. *Objects, Characteristics and Relationships*, Moscow: Publishing House of the Academy of Sciences, USSR, 1963, 184 pp.

2.18 Uyemov A.I. *Logical Fundamentals of a Method of Modeling*, Moscow: Mysl' Publishers, 1971, 311 pp.

2.19 Ponomarev B.F., Kolesnikov A.V., Kirikov I.A., "One Approach in Semiotic Modeling of the State of Transport Systems," in the book: *Questions of Cybernetics. Situation Control. Theory and Practice*, Moscow: Scientific Council for the Complex Problem of 'Cybernetics' on the Presidium of the Academy of Sciences USSR, 1980, pp. 109-130.

2.20 Zhelezov Zh.Y., Pospelov D.A. "Discrete Situational Criteria - Subclass for Homomorphs of a System", in the book: *Automation and Computer Technology*, 1969, No. 2, pp. 5-12.

2.21 Zhelezov Zh.Y., Pospelov D.A. "One Class of Large Systems", Bulletin of the Academy of Sciences USSR: Technical Cybernetics, 1970, No. 2, pp. 49-53.

2.22 Lazarev V.G., Pilyl' Ye.I. *Synthesis of Controlling Robots*, Moscow: Ehnergiya Publishers, 1978, 408 pp.

2.23 Pospelov D.A. *Probabilistic Automats*, Moscow: Ehnergiya Publishers, 1970, 87 pp.

2.24 Bibin G.B. "Use of the Automatic Proof of Theorems Method in Situational Control," in the book: *Theoretical Problems of Situational Control*, Moscow: Scientific Council for Complex Problems of "Cybernetics" on the Presidium of the Academy of Sciences USSR, 1975, pp. 28-34.

2.25 Aleksandrov V.V., Bulkin G.A., Polyakov A.O. *Automated Processing of Information in the Language of Predicates*, Moscow: Nauka Publishers, 1982, 103 pp.

2.26 Osipov G.S. "Two Problems of the Theory of Semiotic Models of Control, I. Concepts of Semiotic Models, Bulletin of the Academy of Sciences USSR: Technical Cybernetics, 1981, No. 6, pp. 100-110.

2.27 Osipov G.S. "Two Problems of the Theory of Semiotic Models of Control," *II. Semantic Analysis*, Bulletin of the Academy of Sciences USSR: Technical Cybernetics 1982, No. 1, pp. 131-137.

2.28 Paducheva Ye.V. "Semantic Analysis of Natural Language with Translation into Languages of Mathematical Logic," *Proceedings of the III All-Union Conference for Information Search Systems and Automatic Processing of Scientific-Technical Information*, Vol. 2, Moscow: VINITI, 1967, pp. 156-169.

2.29 Popov Eh.V., Rodionov Ye.V. "The Algorithm of Synthesis of Clauses of Circumscribed Russian Language from a Semantic Network," Bulletin of the Academy of Sciences: Technical Cybernetics, 1978, Nol 6, pp. 44-51.

2.30 Bryabrin V.M., Senin G.V. "Analysis of Natural Language in a Narrow Context," in the book: *Questions of Cybernetics. Problems of Artificial Intelligence*, Moscow: Scientific Council for the Complex Problem of "Cybernetics" on the Presidium of the Academy of Sciences, 1980, pp. 111-117.

2.31 Paducheva Ye.V. "Problems of Logical Analysis of Place Names," in the book: *Semiotics and Data Processing,"* No. 15, Moscow: VINITI, 1980, pp. 125-153.

2.32 Popov Eh.V. *Interfacing with the Computer in Natural Language*, Moscow: Nauka Publishers, 1982, 360 pp.

2.33 *Dialogue Systems in Automated Control Systems*, ed. D.A. Pospelov, Moscow: Ehergoatomizdat, 1983, 204 pp.

2.34 *Dialogue in Automated Systems*, Moscow: Moscow House of Scientific-Technical Propaganda, 1981, 160 pp.

2.35 Sil'dmyaeh I.Ya. "Fact-Text-Knowledge," in the book: *Scientific Proceedings of Tartu State University*, No. 621, Tartu, 1983, pp. 118-134.

2.36 Fayn V.S. "New Aspects of Understanding Machine Information," Bulletin of the Academy of Sciences USSR: Technical Cybernetics, 1982, No.5, pp. 100-112.

2.37 Fayn V.S. "Subjective Aspect of Information Exchange in a Dialogue," Moscow: Institute of Problems of Information Transmissions, Academy of Sciences, 1983, reprint No. 1, 33 p.

2.38 Plesnivich G.S. "Concept of Knowledge in Associative Networks," Bulletin of the Academy of Sciences USSR: Technical Cybernetics, 1982, No. 5, pp. 6-23.

2.39 Lozovskiy, V.S. "Extended Database on the Bases of Semantic Networks," Bulletin of the Academy of Sciences USSR: Technical Cybernetics, 1982, No.5, pp. 23-42.

2.40 Kleshchev A.S., Chernyakhovskaya, M.Yu. "Systems of Presentation of Problem-oriented Knowledge," Bulletin of the Academy of Sciences USSR: Technical Cybernetics, 1982, No. 5, pp. 43-63.

2.41 Aleksandrov V.V., Bulkin G.A., Polyakov, A.O. *Automated Processing of Information in the Language of Predicates*, Moscow: Nauka Publishing, 1982, 101 pp.

2.42 Skrehgg G. "Semantics Networks as Models of Memory," in the book: *New in Foreign Linguistics*, No. 12, Moscow: Raduga Publishers, 1983, pp. 228-271.

2.43 Preobrazhenskiy A.B. "The Linguistic Process for Realizing Information with Databases," Bulletin of the Academy of Sciences USSR: Technical Cybernetics, 1982, No. 5, pp. 121-129.

2.44 Kord'e M.O. "Use of Knowledge in an Objective Area in a Work Control System in Natural Language: An Experiment with Culinary Receipts," in the book: *Linguistic Processes and the Concept of Knowledge*, Novosibirsk: Computer Center, Siberian Division of the Academy of Sciences USSR, 1981, pp. 5-21.

2.45 *Russian Semantic Dictionary*, Moscow: Nauka Publishers, 1982, 564 pp.

2.46 *Semantic Types of Predicates*, Moscow, Nauka Publishers, 1982, 365 pp.

2.47 Zadhe, L. *Concept of Linguistic Variable and its Use in Making Approximate Decisions*, Moscow: Mir Publishers, 1976, 165 pp.

2.48 Varosyan S.O., Pospelov D.A. "Non-Metric Space Logic", Bulletin of the Academy of Sciences USSR, Technical Cybernetics, pp. 86-99.

2.49 Yeshkova I.V., Pospelov D.A. "Decision Making With Fuzzy Sets," I. Universal Base, Bulletin of the Academy of Sciences USSR: Technical Cybernetics, 1977, No. 6, pp. 3-11.

2.50 Averkin A.N. "Fuzzy Relations of Modeling in Systems of Planning of the Behavior of a Robot," *Artificial Intelligence in Information Control Systems of Robots*, October 18-22, Somolenice, 1982, pp. 6-10.

Chapter 3.

3.1 Debrenn M.M., Narin'yani A.S. "The Speech Act as an Element of Dialogue Interaction," *Scientific Proceedings of Tartu State University No. 654*, Tartu, 1983, pp. 39-48.

3.2 Vinograd T. "One Approach to the Study of Conversation," *Interaction with a Computer in Natural Language*, Novosibirsk, Computer Center, Siberian Division of the Academy of Sciences USSR, 1978, pp. 11-47.

3.3 Vinograd T. *Programming Natural Language*, Moscow: Mir Publishers, 1976, 292 pp.

3.4 Chyernyak Yu. "Deductions and Knowledge," *New in Foreign Linguistics No. 12*: Applied Linguistics, Moscow: Raduga Publishers, 1983, pp. 171-207; 272-317.

3.5 Shenk R. *Processing Conceptual Information*, Moscow: Ehnergiya Publishers, 1980, 358 pp.

3.6 Chernyakhovskaya M.Yu. *The Concept of Knowledge in Expert Systems of Medical Diagnostics*, Vladivostok: Institute of Automatics and Processes of Control with the Computer Center of the Far Eastern Center of the Academy of Sciences, 1983, 212 pp.

3.7 Pospelov D.A. "Elements of Principles of Time Relationships," *Situational Control. Theory and Practice*, Moscow: Sovetskoye Radio Publishers, 1975, pp. 15-21.

3.8 Kandrashina Ye.B., Ochakovskaya O.N., Golubeva L.A. "Experimental Question - Answer System, Including a Simple Model of Time with Elements of Logical Deduction," in the book: *Interaction with a Computer in Natural Language*, Novosibirsk, Computer Center, Siberian Division, Academy of Sciences USSR, 1978, pp. 207-222.

3.9 Dembovskaya V.N. "The Logic of Time for Formation of Coherent Text," *Questions of Cybernetics. Problems of Artificial Intelligence*, Moscow: Nauka Publishers, 1980, pp. 144-158.

3.10 Litvintseva L.V, Pospelov D.A. "Time in Robot and Dialogue Systems," in the book: *Questions of Cybernetics. Problems of Artificial Intelligence*, Moscow: Nauka Publishers, 1980, pp. 62-70.

3.11 Kandrashina Ye.Yu. "Time in the Concept of Knowledge: Correlations between Individual Events," preprint, Vostok Project, No.7, Novosibirsk, 1983, 25 pp. and No. 8, Novosibirsk, 1983, 32 pp.

3.12 Yezhkova I.V., Pospelov D.A. "Decision Making on Fuzzy Basis. Rules of Deduction," Bulletin of the Academy of Sciences USSR: Technical Cybernetics, 1978, No. 2, pp 5-11.

3.13 Yezhkova I.V. "Generalization of Schemes of Decision Making Using Fuzzy Logic," in the book: *Artificial Intelligence Database Software Systems*, Moscow: Moscow House of Scientific-Technical Propaganda, 1978, pp. 102-111

3.14 Prade H. "Data Bases with Fuzzy Information and Approximate Reasoning in Expert Systems," Transactions of the International Symposium on Artificial Intelligence, IFAF, Leningrad, 4-6 October 1983, 20 pp.

3.15 Chesnokov S.V. *Determinational Analysis of Socio-Economic Data*, Moscow: Nauka Publishers, 1982, 168 pp.

3.16 Chesnokov S.V. "Syllogism in Determinational Analysis," Bulletin of the Academy Sciences USSR: Technical Cybernetics, 1984, No. 5, pp. 55-83.

3.17 Zakrpevskiy A.D. "Toward Formalization of Polysyllogistics," in the book: *Logical Conclusion*, Moscow: Nauka Publishers, 1979, pp. 300-309.

3.18 Poya D. *Mathematics and Correct-Like Reasoning*, Moscow: IL [Institute of Linguistics], 1957, 535 pp.

3.19 Narin'yani A.S. "Fuzzy Sets - A New Type of Data for Representing Knowledge," preprint, Computer Center, Siberian Division, Academy of Sciences USSR, No. 232, Novosibirsk, 1980, 26 pp.

3.20 Narin'yani A.S. "Means of Modeling Fuzzy Data in Computers for Representing Knowledge," in the book: *Representing Knowledge and Modeling in the Process of Comprehension*, Novosibirsk, Computer Center, Siberian Division, Academy of Sciences USSR, 1980, pp. 153-162.

3.21 Razlogova Ye.Eh. "Logical Relationships between Meaning and its Components," Scientific-Technical Information, Series 2, 1981, No.1, pp. 14-20.

3.22 Shokin Yu.I. *Interval Analyses*, Novosibirsk: Nauka Publishers, 1981, 111 pp.

3.23 Narin'yani A.S. "Uncertain Models and Operations with Undefined Values," preprint, Computer Center, Siberian Division, Academy of Sciences USSR, No. 400, Novosibirsk, 1982, 33 pp.

3.24 Ingolls D.G.Kh. *Introduction to the Indian Logic of Nav'ya-n'yaya*, Moscow: Nauka Publishers, 1975, 237 pp.

3.25 Parkhomenko P.P. *Technical Diagnostics*, Moscow: Znanie Publishers, 1969, 64 pp.

3.26 Novorusskiy V.V. *Finite-Dimensional Automatic Systems of Control*, Novosibirsk: Nauka Publishers, 1982, 269 pp.

3.27 Kovrigin O.V., Smolyaninov N.D., Chmyr' A.Ya. Experimental Medical Diagnostic Systems, Bulletin of the Academy of Sciences USSR: Technical Cybernetics, 1982, No. 5, pp. 199-216.

3.28 Wright von G.H. *A Treatise on Induction and Probability*, Patterson N.Y.: Littlefield, Adams and Co., 1960, 310 pp.

Chapter 4.

4.1 Leont'yev A.A. *Linguistics and Psychology*, Moscow: Nauka Publishers, 1966, 80 pp.

4.2 Vygotskiy L.S. *A Collection of Articles*, Vol. 2, Moscow: Pedagogika, 1982, 502 pp.

4.3 Zubkova, T.I. "One Direction of Development of a Child's Speech," in the book: *Structural and Applied Linguistics*, Leningrad: LGU [Leningrad State University] Press, 1983, No. 2, pp. 53-57.

4.4 Negnevitskaya Ye.I., Shakhnarovich A.M. *Language and Children*, Moscow: Nauka Publishers, 1981, 110 pp.

4.5 Piazhe Zh. *Selected Psychological Works*, Moscow: Prosveshchenie Publishers, 1969, 659 pp.

4.6 Luriya A.P. *The Historical Development of Cognitive Processes*, Moscow: Nauka Publishers, 1974, 176 pp.

4.7 Luriya A.P. *Language and Sense*, Moscow: Moscow State University Press, 1979, 319 pp.

4.8 Koul M., Skribner S. *Culture and Thinking*, Moscow: Progress Publishers, 1977, 261 pp.

4.9 Losskiy N.O. *Logic*, Ch. 1, Berlin: Obelisk Publishers, 1923, 166 pp.

4.10 Voyshvilo Ye.K. *Concept*, Moscow: Moscow State University Press, 285 pp.

4.11 Hunt E., Marin J., Stone F. *Modeling the Process of Conception on a Computer*, Moscow: Mir Publishers, 1970, 301 pp.

4.12 Hunt E. *Artificial Intelligence*, Moscow: Mir Publishers, 1978, 558 pp.

4.13 Arkad'yev A.G., Braverman Eh.M. *Teaching Machines Classification of Objects*, Moscow: Nauka Publishers, 1971, 110 pp.

4.14 Ayzerman M.A., Braverman Eh.M., Rozonoer L.I. *Methods of Potential Functions in the Theory of Teaching Machines*, Moscow: Nauka Publishers, 1970, 384 pp.

4.15 Aleksandrov V.V., Gorskiy N.D. *Algorithms and Programs of the Structural Method of Processing Data*, Leningrad: Nauka Publishers, 1983, 207 pp.

4.16 *Classification and Cluster*, Moscow: Mir Publishers, 1980, 389 pp.

4.17 Bongard M.M. *The Problem of Recognition*, Moscow: Nauka Publishers, 1967, 320 pp.

4.18 Gladun V.P. "Forming Ideas by Way of Instructing Growing Networks," Cybernetics, 1970, No. 2, pp. 99-104.

4.19 Gladun V.P., Vashchenko N.D. "Methods of Forming Ideas by Computer," Cybernetics, 1975, No. 2, pp. 107-112.

4.20 Gladun V.P., Yefimyenko O.M., Rabinovich Z.L. "Isolation of Characteristic Combinations of Attributes of Objects by Computer," Controlling Machines and Systems, 1972, No. 2, pp. 34-37.

4.21 Kamilov M.M., Zhuravlev Yu.I., and others, *Algorithms for Computing Values and Their Use*, Tashkent: Fan Publishers, 1974, 120 pp.

4.22 Gladun V.P. *Heuristic Search in Complex Environments*, Kiev: Naukova Dumka, 1977, 166 pp.

4.23 Starchenko A.A., Kalaydzhleva M.A. *Comprehension Devices of Thesauruses and Operations by Computer*, in the book: Logico-Methodological Studies, Moscow: Moscow State University Press, 1980, pp. 109-129.

4.24 Reznikov, L.O. *Idea and Word*, Leningrad: Leningrad State University Press, 1958, 124 pp.

4.25 Finn V.K., Zabezhaylo M.I. "Logical Problems of Automatic Forming of Hypothesis," in the book: *Theses of Reports of the IX All-Union Symposium on Cybernetics*, Sukhumi, 10-15 November 1981, Vol. 2, pp. 54-61.

4.26 Finn V.K. "Machine Oriented Formalization of Correct-like Reasoning in the Style of F. Behkon - D.S. Mill", in the book: *Semiotics and Informatics*, No. 20, Moscow: VINITI, 1983, pp. 35-101.

4.27 Zabezhaylo M.I., Finn V.K., Kozlova S.P., and others, "One Method of Automatic Forming of a Hypothesis and Its Program Realization," *Scientific-Technical Information*, Series 2, 1982, No. 4, pp. 20-26.

4.28 Gayek P., Gavranek T. *Automatic Formation of Hypotheses*, Moscow: Nauka Publishers, 1984, 277 pp.

4.29 Plotkin G.D. "A Note on Inductive Generalization," *Machine Intelligence*, 1970, Vol. 5, pp. 153-164.

4.30 Plotkin G.D. "A Further Note on Inductive Generalization," *Machine Intelligence*, 1971, Vol. 6, pp. 101-124.

4.31 Zhukov L.G. "One Algorithm of Inductive Formation of Ideas," in the book: *Questions of Cybernetics. Theory and Practice of Situational Control*, Moscow: Sovetskoye Radio Publishers, 1977, pp. 167-171.

4.32 Vargun V.N. "Deductive and Non-deductive Systems of Decision Making," in the book: *Questions of Cybernetics. Situational Control. Theory and Practice*, Moscow: Sovetskoye Radio Publishers, 1980, p. 45-59.

4.33 Vargun V.N., Viktorova, V.P. "Questions of Structural Generalization and Classification in Systems of Decision Making," Bulletin of the Academy of Sciences USSR: Technical Cybernetics, 1982, No.5, pp. 64-73.

4.34 Viktorova N.P. "Two Approaches to the Solution of Tasks of Classification of Woody Structures," *Scientific-Technical Information*, Series 2, 1981, No.9, pp. 17-21.

4.35 Viktorova N.P. "Development of Algorithms of Structural Generalization," in the book: *Works of MEhI*, 1981, No. 533, pp. 49-53.

4.36 Golender V.Ye., Rozenblit A.B. *Computing Methods of Development of Drugs*, Riga: Zinatne Publishers, 1978, 231 pp.

4.37 Maginskas Yu.A. "One Approach to Analysis of Causality," in the book: *Problems of Logic*, Moscow: Academy of Sciences USSR Press, 1963, pp. 144-150.

4.38 Loš I. "The Foundation of a Methodological Analysis of Mill's Methods," in the book: *25 Years of Logical Methodology in Poland*, Warsaw: 1977, pp. 291-325.

4.39 Dzhini K. *Logic in Statistics*, Moscow: Statistics, 1973, 128 pp.

4.40 Shreyder Yu.A. "Mathematical Model of a Theory of Classification," *Scientific-Technical Information*, Series 2, 1968, No. 10, pp. 7-24.

4.41 Naydenova K.A. "Relational Model of Analysis of Experimental Data," Bulletin of the Academy of Sciences USSR: Technical Cybernetics, 1982, No. 4, pp. 103-119.

4.42 Shreyder Yu.A. *Equality, Similarity, Order*, Moscow: Nauka Publishers, 1971, 254 pp.

4.43 Sharov A.A. "Comprehension of Properties and Theory of Classification," in the book: *Semiotics and Informatics*, No. 11, Moscow: VINITI, 1979, pp. 37-51.

4.44 Pogosyan Eh.M. "Toward a Theory of Automatic Synthesis of Ideas," in the book: *Semiotics and Informatics*, No. 8, Moscow: VINITI, 1977, pp. 125-152.

4.45 Yefimov Ye.I. "Automatic Formation of Ideas, Described by Completed Formulas of Highly-Specialized Calculation of Predicates," in the book: *Semiotics and Informatics*, No. 13, Moscow: VINITI, 1979, pp. 86-108.

4.46 Zarenkov N.A. *Lectures on the Theory of Systematics*, Moscow: Moscow State University Press, 1976, 140 pp.

4.47 Panova N.S., Shreyder Yu.A. "Symbolic Nature of Classification," *Scientific-Technical Information*, Series 2, 1974, No. 12, pp. 3-10.

4.48 Vilenkin N.Ya., Shreyder Yu.A. "Majority Extent and Quantifier 'of the Majority'," in the book: *Semiotics and Information*, No. 8, Moscow: VINITI, 1977, pp. 45-82.

Chapter 5.

5.1 Quinlan J., Hunt E.A. "A Formal Deductive Problem Solving System, J. ASM, 1968, Vol. 15, No. 4, pp. 625-646.

5.2 Fikes R. "A Mechanism of Searching Output for Models which Describe Conditions," *Transactions IV, Combined International Conference on Artificial Intelligence, No. 3, Methods of Searching for Solutions, Heuristic Methods*, Moscow: Scientific Council for the Complex Problem of 'Cybernetics' on the Presidium of the Academy of Sciences USSR, 1975, pp. 206-226.

5.3 Fikes R., Nilsson N. "STRIPES: A New Approach to the Application of the Theorem Proving Method to Problem Solving," in the book: *Integrated Operations*, Moscow: Mir Publishers, 1973, pp. 382-403.

5.4 Yefimov Ye.I. *Solvers of Intellectual Tasks*. Moscow: Nauka Publishers, 1982, 320 pp.

5.5 Robinson J. "Machine Oriented Logic Based on the Principle of Resolution," in the book: *Cybernetic Collection, New Series*, No. 7, Moscow: Mir Publishers, 1970, pp. 194-218.

5.6 Chen Ch., Lee R. *Mathematical Logic and Automatic Theorem Proving*, Moscow: Nauka Publishers, 1983, 358 pp.

5.7 Maslov S.Yu. "The Inverse Method of Determining Derivability in Classical Computing of Predicates," *Reports of the Academy of Sciences USSR*, 1964, Vol. 159, No. 1, pp. 17-20.

5.8 Maslov S.Yu. "Searching Output As A Model of the Heuristic Process," *Cybernetics*, 1972, No. 5, pp. 74-78.

5.9 Maslov S.Yu. "Theory of Searching Output," *Cybernetics*, 1975, No. 4, pp. 134-144.

5.10 Sakerdouti Eh. "The Non-linear Nature of Plans," in the book: *Transactions of the IV Combined International Joint Conference on Artificial Intelligence*, No. 4: Instruction, Debugging and Automatic Programming, Moscow: Scientific Council for the Complex Problem of 'Cybernetics' on the Presidium of the Academy of Sciences USSR, 1975, pp. 99-133.

5.11 Tyugu Eh.Kh. "Solving Problems on Computer Models," Journal of Mathematical and Material-Physical Sciences, 1970, No. 3, pp. 716-733.

5.12 Kakhro M.I., Kal'ya A.P., Tyugu Eh.Kh. *Instrumental System of Programming the Unified System of Computers* (Integrated Software Package), Moscow: Finance and Statistics Publishers, 1981, 157 pp.

5.13 Mints G.Ye., Tyugu Eh.Kh. "Structural Synthesis and Non-classical Logic," in the book: *Theses Reports*, Tallin, 1983, pp. 52-60.

5.14 Dikovskiy A.Ya. "Evaluation of the Effectiveness of Algorithms Connected with Computer Models," in the book: *Theses Reports*, Tallin, 1983, pp. 42-51.

5.15 Tyugu Eh.Kh. "Computer Frames and Structural Synthesis of Programs," Bulletin of the Academy of Sciences USSR: Technical Cybernetics, 1982, No. 6, pp. 176-182.

5.16 Averkin A.N., Gvida J., Pospelov D.A., Somal'vinko M. "Generalized Strategies in the Solution of Problems," Bulletin of the Academy of Sciences USSR: Technical Cybernetics, 1978, No. 5, pp. 71-83.

5.17 Borisov A.N., Aleksyeev A.V., and others, *Models of Decision Making on the Basis of A Linguistic Variable*, Riga: Zinitne, 1982, 256 pp.

5.18 Kolesnikov A.V., Ponomarev V.F., and others, "Systems of Situational Control: Preliminary Publication, Moscow: Scientific Council for the Complex Problem of 'Cybernetics' on the Presidium of the Academy of Sciences USSR, 1978, 43 pp.

5.19 Vishnyakov N.A., Gel'fandbeyn Ya.A., Mamontov A.I. "Base-line Graph of a Topological Model of Operational Control of Aircraft Repair," in the book: *Questions of Theory and Practice of Creating an ASU [automated control system] in Civil Aviation*, Transactions of GosNII GA [State Scientific Research Institute of Civil Aviation], No. 143, Moscow: 1976, pp. 25-34.

5.20 Vishnyakov N.A., Gel'fandbeyn Ya.A. "Mechanism for Resolving Conflicting Situations on the Bases of a Solution Development Model," in the book: *Questions of Theory and Practice of Creating an ASU in Civil Aviation*, Transactions of the State Scientific Research Institute of Civil Aviation, No. 143, Moscow: 1976, pp. 35-43.

5.21 Vishnyakov N.A., Zenkovich Yu.A. "Application of the Principles of Situation Control for Solving Tasks of Operational Control of Civil Aviation Activity," in the book: *Questions of Theory and Practice of Creating an ASU in Civil Aviation*, Transactions of the State Scientific Research Institute of Civil Aviation, No. 173, Moscow: 1979, pp. 201-207.

5.22 Vishnyakov N.A., Gel'fandbeyn Ya.A., Tsvetkov G.S. "Functional-Structural Approach to Modeling an Object of Control in Tasks of Automated Control of Activity." Preliminary Publication. Scientific Council for the Complex Problem of 'Cybernetics' on the Presidium of the Academy of Sciences USSR, Moscow, 1978, 29 pp.

5.23 Zagadskaya L.S., Sokolova O.V. "Methodology of Designing Situational Models of Control," preprint, Scientific Council for the Complex Problem of 'Cybernetics' on the Presidium of the Academy of Sciences USSR, Moscow: 1973, 40 pp.

5.24 Zholondz' V.Ya. "The Use of the Mechanism of Time and Function in the Situational Control of Discrete Objects," in the book: *Questions of Cybernetics. Situational Control. Theory and Practice*, Moscow: Scientific Council for the Complex Problem of 'Cybernetics' on the Presidium of the Academy of Sciences USSR, 1980, pp. 60-69.

5.25 Zagadskaya L.S., Lozovskiy V.S., Sokol'nikov A.I., Goryachuk V.F. "Realization of the Basic Processes in a System of Situational Control," in the book: *Questions of Cybernetics. Situational Control. Theory and Practice*, Moscow: Scientific Council for the Complex Problem of 'Cybernetics' on the Presidium of the Academy of Sciences USSR, 1980, pp. 94-108.

5.26 Bakayev A.A., Kostina N.M., Yarovitskiy N.V. *Automatic Models of Economic Systems*, Kiev: Naukova Dumka, 1970, 191 pp.

5.27 Pastin Yu.Ya., Pozhidayev V.G. "Correlation of Methods of Modeling Dispatch Control in a Fishing Port," in the book: *Automated System of Control in the Fishing Industry*, Kaliningrad: Atlantic Scientific Research Institute of Fishing Industry and Oceanography, Transactions No. 55, 1974, pp. 40-46.

5.28 Gel'fandbeyn Ya.A., Kolesnikov A.V., Rudinskiy I.D. "Semiotic-Statistical Model of the Behavior of a Transportation Center," in the book: *Situational Control and Semiotic Modeling*, Moscow: Scientific Council for the Complex Problem of 'Cybernetics' on the Presidium of the Academy of Sciences USSR, 1983, pp. 40-49.

5.29 Kirikov I.A. "The Use of PPP Modeling of the State and Behavior of Large Systems of MOSOP in the ASU of a Fishing Port," in the book: *Economic Problems of the Development of Soviet Fishing in the Atlantic*, Kaliningrad: Atlantic Scientific Research Institute of the Fishing Industry and Oceanography, 1981, pp. 82-84.

5.30 Pospelov D.A. "Human Reasoning in Intellectual Systems," in the book: *Logic of Reasoning and its Modeling*, Moscow: Scientific Council for the Complex Problem of 'Cybernetics' on the Academy of Sciences USSR, 1983, pp. 5-37.

5.31 Kilks F. *Awakening Thinking. The Sources of Human Intellect*, Moscow: Progress Publishers, 1983, 302 pp.