

Perspectivas em Cognição e Teoria da Informação

Nicolau Werneck, 29 de Junho de 2006

1 Introdução

Inicialmente apresentaremos uma cronologia de fatos antológicos em áreas de estudo ligadas à ciência cognitiva. Consideramos importante o *insight* obtido sobre estas descobertas quando as consideramos de um ponto de vista histórico, e não meramente técnico. Também falaremos sobre o *Perceptrons* [1], questionando a visão estabelecida.

A segunda parte mostra como algumas idéias ligadas à teoria da informação e computação podem contribuir mais ao estudo da ciência cognitiva.

2 Histórias científicas

Abordaremos aqui as histórias da computação, lógica e neuroanatomia. Estas histórias começam na antigüidade, mas começaremos na Renascença.

2.1 Computação

Um dos marcos para o início da Renascença na Europa é sem dúvida a invenção do tipo móvel, que influenciou a matemática possibilitando a divulgação do conhecimento islâmico. Isto inclui a popularização dos algarismos árabes e de noções de álgebra.

Na computação, a Renascença viu chegarem máquinas capazes de realizar cálculos sem a manipulação de pedrinhas, como nos ábacos antigos. Antes vieram sofisticados ábacos, como os ossos de Napier. O matemático escocês também re-inventou o logaritmo na Europa, fato fundamental para a quase imediata construção da régua de cálculo. Napier morreu em 1617, depois de publicar seu *Rabdologiae*.

A primeira máquina para somar números discretos foi a de Schickard. Ele comunicou a Kepler em 1623 seu uso para cálculos de astronomia.

Esta máquina foi esquecida, mas depois vieram as de Pascal (1642) e Leibniz (1673), esta capaz de multiplicar. As idéias de Leibniz foram reaproveitadas em dispositivos, como o *Aritmômetro* de Thomas (1820), primeira calculadora com sucesso comercial.

Guiado pela necessidade de calcular funções mais rápida e precisamente, e inspirado pelo motor a vapor, o cientista inglês Charles Babbage voltou-se no início do século XIX para o problema de fazer uma máquina para calcular tabelas de polinômios.

Esta máquina, a *Difference Engine*, foi concebida em 1822. Sua construção foi iniciada, mas interrompida. A idéia foi aproveitada por Scheutz, que construiu em 1843 uma máquina capaz de calcular até a 4ª ordem, com 15 dígitos. Thomas e Scheutz foram premiados na segunda Feira Mundial, em 1855 [2].

Babbage percebeu que ainda poderia mecanizar mais tarefas, e começou a conceber a *Analytical Engine*. O trabalho foi interrompido por sua morte em 1871. Ada Byron, que contribuía com Babbage, já mostrava em um artigo questionamentos sobre a máquina ter ou não capacidades cognitivas [3][4].

Este trabalho de Babbage foi quase esquecido após sua morte. Equanto isso foram feitos avanços apenas na construção de simples máquinas de calcular. O próximo grande acontecimento no mundo do cálculo automático veio com o desenvolvimento de componentes eletro-mecânicos e eletrônicos.

Em 1835 Henry inventou o relé, que foi empregado em máquinas como as de Stibitz, Zuse e Aiken no início dos anos 1940. É interessante estudar as máquinas desta época, porque as diferentes características e limitações nos fazem questionar a definição de um "computador." Também havia mais diversidade em questões que não são debatidas hoje, mas são importantes. Por exemplo, diversas máquinas utilizavam registradores decimais ou bi-quinários.

Shannon demonstrou a aplicabilidade da álgebra booleana na confecção de circuitos em 1938, o que sem dúvida facilitou estes trabalhos. Seu trabalho em Teoria da Informação veio em 1948, quando ele modelou a linguagem como uma cadeia de Markov. Este revolucionário trabalho iria mais tarde influenciar Chomsky[5], além de vários outros. Shannon também estudou genética e IA, e foi um dos organizadores da conferência em Dartmouth [6].

2.2 Lógica e Matemática

Aristóteles fundou a lógica no século IV AEC, influenciando pra sempre a ciência. O mais importante estudioso da lógica desde então foi Leibniz (1646-1716). Ele foi um dos primeiros a tentar reduzir as questões do raciocínio humano à “mera” manipulação sistematizada de símbolos.

Esta visão de Leibniz só se popularizou com a publicação em 1896 de *La Logique de Leibniz* por Couturat. Este livro influenciou Russell antes de lançar seu *Principia Mathematica*, e foi contemporâneo dos trabalhos de Peirce e Peano em lógica simbólica. Antes disso Boole (1815-1864) e De Morgan (1806-1871) já haviam iniciado a moderna lógica formal.

Este movimento levou nos anos 1920 ao programa de Hilbert, que questionava os limites da formalização. Ele também questionava uma visão pessimista em relação à ciência existente na Alemanha desde a metade do século XIX, iconificada pela frase em latim *ignoramus et ignorabimus*. Emil du Bois-Reymond defendeu esta teoria em 1874.

Em 1930 Gödel publicou seus teoremas, e em 1936 Church e Turing publicaram os seus resultados a respeito do *Entscheidungsproblem* de Hilbert. No caso de Turing, ele o reduziu ao problema da parada.

Tanto o artigo de Turing [7] quanto o primeiro de McCulloch e Pitts [8], de 1943, ajudaram a formalizar o conceito de autômata. Outros pioneiros na área foram Kleene, estudando expressões regulares, e von Neumann, que criou o autômata celular estudando sistemas auto-reprodutores.

É interessante notar ainda que as máquinas de Mealy e Moore, conhecidas por qualquer aluno da área, só foram publicadas em 1955 e 1956. Já o mapa de Veitch e Karnaugh surgiram apenas em 1953 [9].

2.3 Neuroanatomia

Ao contrário da matemática, a neuroanatomia dependeu em parte de instrumentos de medição mais poderosos, como em outras ciências naturais.

A descoberta mais antiga a que daremos atenção é a de Galvani, que demonstrou em 1780 a natureza elétrica do funcionamento dos nervos. Este estudo levou ao desenvolvimento da pilha de Volta em 1800. É interessante perceber que neste período a eletricidade estava mais ligada ao elusivo estudo da vida, do que à computação, como enxergamos hoje.

Entre 1830 e 1847 a teoria celular foi formalizada por Schleiden e Schwann. A neuroanatomia só foi se integrar a esta idéia com o trabalho de Ramón y Cajal, que foi uma das figuras mais importantes na defesa da idéia de que o sistema nervoso é composto por neurônios unitários que se tocam formando uma rede. O trabalho de Cajal se deu a partir de 1887, desencadeado pela descoberta da reação de Golgi em 1873. Ambos pesquisadores foram agraciados com o Nobel em 1906. O termo *sinapse* foi criado por Sherrington em 1897, e *neurônio* em 1891 por von Waldeyer, promotor do trabalho de Cajal na Alemanha.

Outras descobertas importantes não dependeram do estudo microscópico. Em 1861 Broca realizou a autópsia que o levou à descoberta da área do cérebro que leva o seu nome. Wernicke achou mais uma área relacionada à afasia em 1874.

Um fato curioso demonstra o estado em que a computação e a neuroanatomia estavam uma em relação à outra no fim do século XIX. Na ocasião da morte de Babbage em 1871, seu cérebro foi extraído para estudos. O mesmo ocorreu com Gauss (1855) e Einstein (1955) [10], bem como outras personalidades, mas nenhum outro caso carrega tanta ironia...

De 1848 a 1884 Emil du Bois-Reymond publicou seu trabalho em eletrofisiologia. Ele foi o primeiro a observar o potencial de ação neural, em 1848.

Foi apenas em 1939 que K. Cole, H. Curtis, Hodgkin e Huxley conseguiram fazer medições de tensão dentro de um axônio gigante de lula, amostrando correntes de ions através da membrana neuronal. O sensor de Cole era baseado em amplificação com realimentação negativa, inventada por Black em 1927. Hodgkin e Huxley partiram deste trabalho para ex-

plicar os mecanismos por trás do potencial de ação, e publicaram seus resultados em 1952.

Em 1943 McCulloch e Pitts publicaram o artigo em que demonstram que o simplório modelo de neurônio deles pode implementar funções booleanas e máquinas de estados genéricas [8]. Ambos trabalharam com Lettvin e Maturana em um famoso estudo sobre a visão de sapos em 1959 [11].

3 Estruturas computacionais

Minha opinião é que nas histórias acima os estudos “racionalistas” sempre estiveram juntos de objetivos “holistas,” e que o importante para os pesquisadores sempre foi modelar com precisão a dinâmica de fenômenos envolvendo o processamento e transmissão de informação em diferentes níveis.

Na busca por uma maneira de abandonar polarizações maniqueístas como “conexionismo” versus “serialismo,” busquei questionar o que normalmente se ouve, e ler os trabalhos originais que sempre citamos em aulas de redes neurais [8][12][1].

Curioso a respeito do que sempre se ouve a respeito do livro *Perceptrons* e a função ou-exclusivo, acabei aprendendo sobre um erro extremamente comum na idéia que as pessoas têm deste assunto. O equívoco é tão divulgado, que é raro encontrar sites na Internet que comentem o fato [13]. Já as referências mais importantes criticando a má interpretação do *Perceptrons* são a própria edição de 1988 do livro [1], e o site na Wikipédia sobre Marvin Minsky, editado *por ele próprio*, esclarecendo a questão [14].

3.1 A visão estabelecida

A história sobre como o *Perceptrons* teria jogado uma pá de cal sobre o conexionismo tem muitas versões, mas o básico é o seguinte: McCulloch e Pitts teriam desenvolvido seu modelo simplificado neurônio, que teria sido estudado por Rosenblatt, o que aplicou na prática. Por fim, Minsky e Papert teriam escrito o *Perceptrons*, e este livro apresentaria resultados tão negativos e inesperados sobre esta ferramenta matemática que o estudo das redes neurais teria sido congelado por mais de uma década, até que

nos anos 80 o conexionismo ressurgiu como uma fênix, contrariando as expectativas dos “cognitivistas” e pessoas “contrárias às redes neurais.”

Esta história possui algumas incoerências sutis que passam sempre despercebidas no ensino sobre perceptrons, talvez porque a novidade da matéria não permite um bom juízo pelos estudantes.

Sempre que se vai tentar explicar qual seria um dos tais “resultados negativos” de Minsky e Papert, cita-se o exemplo da função XOR, ou paridade. A fábula começa com a demonstração da incapacidade de um único neurônio em implementar a função XOR. A partir daí, eles teriam inferido que redes mais complexas também apresentariam limitações igualmente graves. Alguns ressaltam a incapacidade de *implementar*, enquanto outros ainda falam na dificuldade em *generalizar* a função.

Esta visão sobre o *Perceptrons*, incluindo a questão do XOR, está em muitos livros [15][16]. Outros ainda mantém a discussão em um melhor nível [17], mas sempre deixando claro que as tais conjecturas a respeito de limitações das redes seriam responsáveis por desviar o interesse para ferramentas “cognitivistas,” ou “simbólicas.” Descobertas nos anos 1980 teriam então contrariado as expectativas e inferências de Minsky e Papert, ressuscitando a área de estudos.

3.2 A crítica

Contada rapidamente, a fábula acima dá a impressão de que até os anos 1980 não se sabia implementar o XOR com neurônios de McCulloch e Pitts, e que isto só se tornou possível com o uso de “redes com múltiplas camadas,” e o “back-propagation.”

Esta idéia é absurda. Pra começar, o resultado central do artigo de 1943, é justamente o fato de que é possível implementar *qualquer função booleana* com as redes propostas. Além disso a questão de máquinas de estado foi abordada, um problema mais complexo do que um sistema sem memória.

Se redes perceptron têm como virtude poderem implementar funções lógicas genéricas, e foi isto que chamou a atenção das pessoas, como pode o fato da função paridade não ser implementável *por uma única unidade* ter arrefecido o interesse!?

O fato é que a tradicional estrutura onde as entradas são ligadas a uma “camada oculta,” e as saídas desta são processadas ainda por uma última unidade, a “camada de saída” já está presente ao menos no livro de Rosenblatt. Os dispositivos construídos pela equipe de Rosenblatt, por exemplo, capazes de reconhecimento de caracteres, não eram estas singelas unidades incapazes de implementar o XOR.

3.3 Onde erramos

Creio que o maior problema das pessoas com o *Perceptrons* é uma confusão entre *eficiência* de computação e *computabilidade*. Minsky e Papert investigaram esses dois aspectos, e resumir a obra com o “desproblema” do ou-exclusivo de duas variáveis é uma simplificação mais do que ingênua.

Minsky e Papert realmente estudam muito a função paridade no *Perceptrons* [1]. Lá é demonstrado que o XOR não pode ser implementado por uma única “linear threshold unit,” já que estas implementam somente o que eles chamam de “predicados de ordem 1.” Mas o livro deixa claro que estas unidades podem servir como funções intermediárias a serem lidas por uma unidade final, implementando predicados “de ordem N” como a paridade.

Fala-se muito em “separabilidade linear” quando se quer explicar as limitações computacionais dos *perceptrons*, mas é preciso ver que uma coisa é o hiperplano singular determinado por um único neurônio, e outra coisa são as curvas compostas por segmentos de reta que delimitam as regiões de classificação de *redes* destas unidades. Estas curvas podem aproximar qualquer curva “bem-comportada” com a resolução que se desejar.

Uma suspeita sobre esta confusão é talvez a responsável por se querer falar em predicados “fortemente separáveis linearmente” [15].

Talvez o problema do XOR seja tão famoso por parecer a epítome do que seria o “reduccionismo” sendo derrubado pelo “holismo.” Pode se entender que “a simples soma de neurônios” não implementaria o XOR, enquanto que um poder maior de computação “emerge” da rede...

Os pesquisadores da área nunca deram tanta importância a este problema quanto parece. Enquanto

esta questão do XOR é repetida *ad nauseum* em materiais introdutórios, os livros nos anos 1960 estavam preocupados em mostrar como que as redes neurais se relacionam à teoria de autômatas [18][19].

O grande achado do *Perceptrons* é que existem problemas que inerentemente não se beneficiam de estruturas neurais. Um exemplo é a própria capa do livro, que apresenta um par de figuras, uma conectada, e outra desconexa. A dificuldade de nós, leitores, em reconhecer qual é qual serve até para corroborar o resultado do livro.

Para finalizar o assunto, é importante lembrar que os próprios autores reconhecem na edição expandida de 1988 do *Perceptrons* que eles não concordam com a visão que se formou a respeito do livro.

4 Conceitos de computação na cognição

O objetivo desta última seção é mostrar como é possível aplicar alguns conceitos de computação em problemas clássicos de ciência cognitiva, para tentar se obter algum novo *insight* sobre eles.

4.1 Dicotomias

Minha esperança ao apresentar aquelas cronologias, é facilitar a visão de que sistemas de computação podem ter características muito diferentes, mas nunca escapam do limite de Turing e da álgebra.

Estou ao lado de Minsky quando ele anuncia no terceiro e quarto parágrafos do prólogo de 1988 do *Perceptrons* que as diviões entre sistemas “seriais” e “paralelos,” “localizados” e “distribuídos,” “discreto” e “contínuo” não fazem sentido [1][20]. É preciso acabar com esse debate entre arquétipos.

É verdade que alguns problemas podem ser resolvidos mais rapidamente com máquinas “paralelas,” mas não existe algoritmo que uma máquina de Turing não rode, mesmo que lentamente. A computabilidade limite de qualquer sistema de processamento de dados é sempre a mesma. Por isso não acredito que a IA forte precise *inerentemente* de uma “máquina paralela.”

A argumentação contrária sempre acaba se valendo de questões de precisão, levando à dicotomia “contínuo” versus “discreto.” Minha resposta é que nenhum sistema é perfeitamente contínuo e instantâneo na prática. Sempre existem ruídos e inércias que eventualmente se ligam à equação de Shannon para revelar as capacidades digitais dos canais de transmissão existentes no sistema.

É preciso demolir esse maniqueísmo. Precisamos criar uma visão integrativa, pensando sempre em termos de teoria da informação e estatística.

...Eu ainda tenho uma teoria de que talvez algumas destas dicotomias tenham surgido de brigas pessoais entre cientistas. O melhor exemplo seria o de Wiener e von Neumann, que disputavam ferozmente sobre o papel de cientistas em guerras [21].

4.2 Sistemas abertos, modulares, complexos e autopoieticos

Estes vários conceitos possuem em comum o uso da idéia de componente. O estudo ostensivo da complexidade se iniciou com o trabalho do Santa-Fé Institute, com a contribuição de Gell-Mann entre outros. Parte do sucesso destas idéias foi talvez devido a um melhor entendimento sobre o caos [22].

O estudo do caos foi a coqueluche dos anos 80. Um dos pioneiros dos sistemas caóticos foi Poincaré, no início do século XX. Ele também questionou o logicismo defendido por Couturat, e também questionou a termodinâmica de Boltzmann. Este questionamento de Poincaré, baseado em seu Teorema da Recorrência, acabou levando a um resultado importante no estudo de ergodicidade evidencia a dificuldade da existência de sistemas fechados.

A idéia de sistema aberto contribui muito para nossa discussão. É piada tradicional entre estudantes de IA que o objetivo é encontrar problemas que se adequem às ferramentas. Uma verdade revelada por esta piada é que os dados analisados possuem um importantíssimo papel nos sistemas.

Existem duas fontes de conhecimento para um sistema inteligente. Um é o ambiente externo, e a outra são imposições de sua estrutura interna. As informações do ambiente colocam o sistema como um

componente de sistema maior. Este ambiente possui uma dinâmica própria, e o componente sempre funciona em um contexto específico.

Estes conceitos são encontrados nos livros de Maturana, Varela e Rosen [23][24]. Acredito ser importante estudar como estas idéias afetam diferentes estruturas de processamento de informação.

As idéias de supercompiladores e formação de hierarquias de Turchin também se encaixam aqui [25].

A idéia de autopoiese ainda não foi muito explorada no contexto da teoria de autômatas [26]. Acredito ser preciso questionar quando que as estruturas “vivas” em autômatas celulares não podem ser interpretadas como “meras” auto-funções.

4.3 A sala chinesa

O problema da sala chinesa de Searle promoveu um enorme debate na ciência cognitiva [27].

Algumas análises sobre este *Gedankenexperiment* tocam na questão da eficiência de computação. A primeira objeção de Hofstadter e Dennet, por exemplo, questiona a habilidade de Searle em executar manualmente um algoritmo de IA. Outras objeções, como a de Rey, levantam a questão da estrutura do sistema. Rey demanda que uma estrutura computacional inteligente seja conexionsita.

Precisamos analisar melhor como diferentes estruturas computacionais afetam o problema.

O problema da sala chinesa parece se relacionar com esta outra questão: Podemos considerar uma tabela de referência (*look-up table*) um sistema de computação [28][29]? O livro com instruções de Searle é realmente um análogo claro de uma tabela de referência, se imaginarmos um cenário simplório.

A hesitação em responder positivamente à pergunta me parece similar ao questionamento de que o computador paralelo de Rey ou o sofisticado algoritmo “serial” de Hofstadter e Dennet sejam condições necessárias para a existência de inteligência.

Antes de questionarmos qual é o grau de complexidade de um sistema inteligente, precisamos determinar a posição na hierarquia de computabilidade desse sistema, e decidir se a inteligência precisa ser executada em uma máquina de Turing, ou em algum

sistema contínuo, ou se basta um autômato finito, ou apenas uma “mera” tabela-verdade.

A memória interna da sala chinesa desempenha um papel fundamental na questão, tanto no aspecto de memória de curto prazo quanto no aprendizado. Nunca vi esta questão ser debatida explicitamente neste contexto, e não existe estudo onde mais se analisam diferentes estruturas de memória do que no da hierarquia de computabilidade.

Voltando ao problema de substituir um sistema de IA por uma tabela de referência, podemos questionar como o desempenho seria comprometido. O maior problema é a generalização. Sistemas com estruturas restritivas, uma vez que “treinados”, podem responder a situações não vistas antes. Ao definir uma tabela nos forçamos a explicitar como faremos para extrapolar e interpolar os dados de treinamento.

Estudar este problema revela a verdadeira natureza do uso de diferentes estruturas de processamento de informação. O que se escolhe na verdade é o comportamento que o sistema terá em situações não especificadas explicitamente em seu projeto.

Estruturas como autômatas finitos ou processos de Markov podem impressionar bastante. É possível por exemplo definir um autômata finito para reproduzir a obra completa de Shakespeare [30], ou de Mozart. Basta que cada estado signifique uma de cada das letras ou notas de *toda* a obra. Precisamos desenvolver melhor os argumentos que impedem de ver um sistema desses como uma solução elegante.

O problema nestes sistemas é controlar o enorme número de parâmetros. O que diferentes estruturas de computação fazem é fornecer meios de domar estes parâmetros, fazendo com que os dados de treino à nossa disposição se tornem suficientes para obter resultados “interessantes.”

4.4 Comportamento serial

É notório que o trabalho de Karl Lashley levantou críticas ao modelo estímulo-resposta do cérebro baseado em simplórios arcos-reflexo [31]. Creio que a crítica de Lashley pode ser enriquecida se pensarmos em diferentes estruturas de computação.

Lashley questionou a idéia de que a fala decorre-

ria de um encadeamento em que a ativação de cada palavra ativa a palavra seguinte [32]. Assim é o modelo que supõe que a formação de frases depende da realimentação através da audição. Escutar a última palavra faria o locutor proferir a seguinte.

Podemos imaginar quais estruturas computacionais poderiam apresentar os erros de fala desejados. Uma seria o encadeamento de blocos como o acima, formando a frase internamente. Este sistema necessitaria de mais uma etapa de saída coletando as palavras geradas. Outra possibilidade é um sistema com uma pilha, e um único bloco de processamento gerando as palavras.

Ambos sistemas atendem os requisitos de Lashley, mas possuem diferentes estruturas e maneiras de gastar memória e unidades de processamento.

5 Conclusão

Neste artigo lancei argumentos a favor de uma postura mais abrangente com relação ao estudo da ciência cognitiva através da computação. Esta postura não é nova, mas sim uma retomada de um sentimento já manifestado por vários pesquisadores de antes dos anos 1980.

Meu desejo é seguir a citação de von Neumann no artigo de Ackley sobre máquinas de Boltzmann. Aquilo não é uma chamada ao “conexionismo,” mas muito antes uma intenção em expandir a idéia da computação para sistemas mais diversificados utilizando a teoria da informação e estatística [33].

Já sabemos que problemas grandes não cedem a sistemas simples como redes neurais ou máquinas de estados. Minsky chamou a atenção à necessidade de se estudar melhor como quebrar problemas em sub-problemas [34][20]. Esse tipo de entendimento não vai ocorrer enquanto os debates sobre IA se resumirem a inócuos conflitos maniqueístas. É preciso questionar que características os sistemas devem ter num nível hierárquico mais elevado, utilizando conceitos mais genéricos e abstratos.

Creio que seria interessante analisar as diferentes arquiteturas cognitivas utilizando conceitos de teoria da informação, procurando ver as vantagens das macroestruturas que ligam os componentes.

Referências

- [1] M. Minsky and S. Papert, *Perceptrons*, pp. 22, 28, 56. The MIT Press, 3rd ed., 1988.
- [2] S. Augarten, *Bit by Bit: An Illustrated History of Computers*. Houghton Mifflin, 1984.
- [3] C. D. G., *The Transformation of Psychology: Influences of 19th-Century Philosophy, Technology, and Natural Science*, ch. Charles Babbage, the Analytical Engine, and the Possibility of a 19th-Century Cognitive Science, pp. 133–152. American Psychological Association, 2001. Disponível em <http://www.yorku.ca/christo/papers/Babbage-CogSci.htm>.
- [4] A. Turing, “Computing machinery and intelligence,” *Mind*, vol. LIX, pp. 433–460, Outubro 1950. Disponível em <http://www.abelard.org/turpap/turpap.htm>.
- [5] D. Jurafsky and J. H. Martin, *Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition*. Prentice-Hall, 2000. Disponível em <http://www.cs.colorado.edu/~martin/slp.html>.
- [6] J. McCarthy, M. L. Minsky, N. Rochester, and C. Shannon, “A proposal for the dardmouth summer research project on artificial intelligence,” Agosto 1955. Disponível em <http://www-formal.stanford.edu/jmc/history/dartmouth/dartmouth.html>.
- [7] A. Turing, “On computable numbers, with an application to the entscheidungsproblem,” *Proceedings of the London Mathematical Society*, vol. 42, no. series 2, 1936. Disponível em <http://www.abelard.org/turpap2/tp2-ie.asp>.
- [8] W. McCulloch and W. Pitts, “A logical calculus of ideas imminent in nervous activity,” *Bulletin of Mathematical Biophysics*, vol. 5, pp. 115–133, 1943.
- [9] M. M. Mano, *Digital Design*. Prentice Hall, 3rd ed., 2001.
- [10] Random Samples, “Gauss: Just another brain,” *Science Magazine*, vol. 287, no. 5455, 2000. Disponível em <http://www.sciencemag.org/cgi/content/short/287/5455/963a>.
- [11] Lettvin, H. Maturana, W. S. McCulloch, and W. Pitts, “What the frog’s eye tells the frog’s brain,” *Proceedings of the IRE*, vol. 47, pp. 1940–1951, 1959.
- [12] F. Rosenblatt, *Principles of Neurodynamics*. Spartan, 1962.
- [13] I. S. N. Berkeley, “A revisionist history of connectionism.” Internet, 1997. Disponível em <http://www.ucslouisiana.edu/~isb9112/dept/phil341/histconn.html>.
- [14] “Marvin minsky.” Wikipédia, Junho 2006. Disponível em http://en.wikipedia.org/wiki/Marvin_minsky.
- [15] R. D. Reed and R. J. M. II, *Neural Smithing: Supervised Learning in Feedforward Artificial Neural Networks*, p. 18. The MIT Press, 1999.
- [16] S. Y. Kung, *Digital Neural Networks*, p. 102. Prentice Hall, 1993.
- [17] S. Haykin, *Neural Networks: A Comprehensive Foundation*, p. 120. Prentice Hall, 2nd ed., 1998.
- [18] M. L. Minsky, *Computation: Finite and Infinite Machines*. Prentice-Hall, 1967.

- [19] M. A. Arbib, *Brains, Machines and Mathematics*. McGraw-Hill Book Company, 1964.
- [20] M. Minsky, "Logical versus analogical or symbolic versus connectionist or neat versus scruffy," *AI Magazine*, vol. 12, no. 2, pp. 34–51, 1991.
- [21] F. Dyson, *De Eros a Gaia*, ch. 23, pp. 296–300. Editora Best Seller, 1992. Publicado em *Technology Review*, fev-mar de 1981, pg. 17-19.
- [22] R. Gallagher and T. Appenzeller, eds., *Complex Systems*. Science Magazine, April 1999. Série de artigos publicados em uma edição especial da revista, volume 284.
- [23] H. Maturana and F. Varela, *Autopoiesis and Cognition*. D. Reidel Publishing Company, 1972.
- [24] R. Rosen, *Life Itself: A Comprehensive Inquiry into the Nature, Origin, and Fabrication of Life*. Columbia University Press, 1991.
- [25] V. Turchin, *The Phenomenon of Science*. Columbia University Press, 1977. Disponível em <http://pespmc1.vub.ac.be/POSBOOK.html>.
- [26] R. Beer, "Autopoiesis and cognition in the game of life," *Artificial Life*, 2004.
- [27] V. J. M. III, *June 15, 1997*. PhD thesis, Department of Electrical Engineering, Computer Systems Laboratory, Stanford University, June 1997. Disponível em <http://csl-publications.stanford.edu/papers/CSLI-97-202nographic.pdf>.
- [28] P. S. Churchland and T. J. Sejnowski, *The Computational Brain*, ch. 3, p. 69. The MIT Press, 1994.
- [29] Z. W. Pylyshyn, *Computation and Cognition*, ch. 7, p. 214. The MIT Press, 1984.
- [30] O. S. McAckus. Internet. Disponível em <http://www.markovbible.com/>.
- [31] H. Gardner, *A Nova Ciência da Mente*, pp. 275–280. edUSP, 1985.
- [32] D. W. Glasspool, *Modelling serial order in behaviour: Studies of spelling*. PhD thesis, University College London, 1998.
- [33] D. H. Ackley, G. E. Hinton, and T. J. Sejnowski, "A learning algorithm for boltzmann machines," *Cognitive Science*, vol. 9, pp. 147–169, 1985.
- [34] M. Minsky, "Artificial intelligence," *Scientific American*, vol. 215, pp. 246–263, 1966.