

# Aspectos Humanos da Tomada de Decisão nas Arquiteturas Cognitivas

Wilson Roberto Loreto Rodrigues

FEEC - Faculdade de Engenharia Elétrica e Computação, UNICAMP  
Cidade Universitária Zeferino Vaz, Distrito Barão Geraldo, Campinas-SP, Brasil

## Resumo

A teoria da decisão é um tópico explorado há muitos anos, em diversas áreas do conhecimento. Com a evolução das arquiteturas cognitivas e o crescimento das discussões sobre a importância das emoções, a decisão humana passou a ser importante tema para pesquisa, abrindo novas fronteiras e desafios. Esse artigo apresenta algumas arquiteturas cognitivas que apresentam aspectos humanos da decisão, além de demonstrar a importância da presença das emoções nos sistemas que visam alcançar uma capacidade de decisão semelhante àquela apresentada em seres humanos.

## Palavras-chave

tomada de decisão; arbitração; seleção de ação; emoções; arquiteturas cognitivas

## Introdução

A tomada de decisão é um aspecto inerente da cognição humana, e fundamental em qualquer humano. Decidir significa escolher uma ou mais alternativas, dentre as diversas alternativas capazes de resolver um problema. A tomada de decisão humana envolve diversos aspectos, e pode requerer processamento complexo em criaturas artificiais.

Diversas arquiteturas cognitivas apresentam módulos de decisão, e demonstram o funcionamento desses módulos de decisão. Anderson et al (2006) apresentam o ACT-R, cujo processo de decisão é representado pelas produções, nos mecanismos de associação, seleção e execução (Figura 1). Dentro da decisão do ACT-R está a resolução de conflitos, utilizada quando mais de uma proposição podem ser executadas.

Na arquitetura SOAR, Lehman et al (2006), apresentam um módulo de decisão separado, ligado à memória de trabalho (Figura 2). O SOAR apresenta um módulo de resolução de impasses, análogo ao módulo de solução de conflitos do ACT-R

Já Imbert (2005) sugere uma arquitetura em que cada um de seus módulos é influenciado por emoções (Figura 3). A tomada de decisão ocorreria após uma interação entre os módulos sociais e deliberativos, em que não só os estados

internos, mas também as regras sociais iriam ter influência.

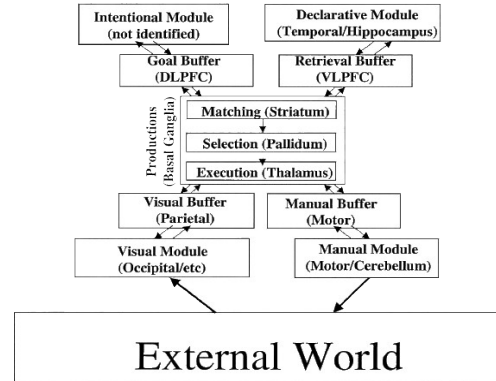


Figura 1 - ACT-R

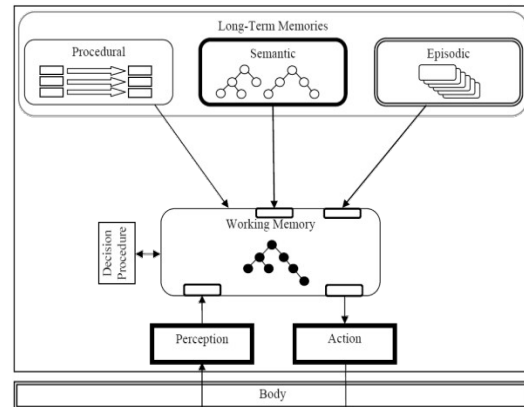


Figura 2 - SOAR

Alguns aspectos que tornam a tomada de decisão humana peculiar serão apresentados a seguir.

**Ambientes dinâmicos:** A tomada de decisão humana muitas vezes ocorre em situações que dependem fortemente do tempo, e em ambientes que variam de acordo com o tempo e de acordo com as próprias ações tomadas. Ambientes dinâmicos trazem restrições importantes aos sistemas cognitivos.

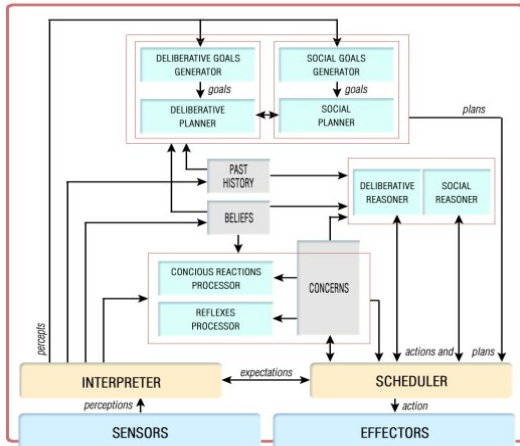


Figura 3 – COGNITIVA

### Aversão ao risco e aversão a ambigüidades:

A mente humana escolhe naturalmente por opções menos arriscadas ou sem ambigüidades. Isso foi demonstrado pelos paradoxos de Ellsberg e Allais. O paradoxo de Ellsberg apresenta um problema em que uma urna é apresentada ao participante, com bolas vermelhas, pretas e amarelas. A urna possui 30 bolas vermelhas e uma quantidade desconhecida de bolas pretas e amarelas.

O participante deve então, escolher uma das apostas:

- A) Ganhar \$100 se sacar uma bola vermelha
- B) Ganhar \$100 se sacar uma bola preta

Da mesma maneira, o participante deve escolher entre:

- C) Ganhar \$100 se sacar uma bola vermelha ou amarela
- D) Ganhar \$100 se sacar uma bola preta ou amarela

A maioria dos participantes opta pela opção A sobre a opção B. De forma curiosa, os participantes que votaram em A sobre B, optam por D sobre C. Isso ocorre devido ao fato de A e D apresentarem probabilidades certas (30/90) e (60/90). Mesmo que a segunda escolha aparente contrariar a primeira (se o participante opta por A, ele acredita que as bolas pretas estão em menor quantidade que as amarelas, e, portanto, deveria apostar em C), as escolhas que parecem ser “garantidas”, ou sem ambigüidades, prevalecem.

O paradoxo de Allais apresenta um problema em que o participante escolhe uma entre duas proposições:

- A) 1/3 de chance de ganhar \$300 ou 2/3 de chance de não ganhar nada

- B) Um ganho certo de \$100

Da mesma forma, o participante escolhe entre:

- C) 1/3 de chance de perder \$300 ou 2/3 de chance de não perder nada
- D) Uma perda certa de \$100

A maioria dos participantes escolhe A ante B, e, quando se troca o ganho pela perda, escolhe C entre D, apresentando uma aversão ao risco.

**Variabilidade:** O comportamento dos seres humanos pode apresentar variações em situações semelhantes. Quando expostos a situações semelhantes, diferentes seres humanos podem agir de diferentes maneiras. Da mesma forma, um mesmo ser humano, ao ser exposto mais de uma vez a uma situação, pode agir de formas diferentes.

**Emoções:** As emoções possuem papel fundamental nas decisões que se toma no dia-a-dia, e não podem ser descartadas quando se tenta reproduzir um a tomada de decisão humana.

Considerando esses aspectos da decisão humana, buscou-se apresentar estudos que tentam de alguma maneira reproduzi-los em arquiteturas cognitivas.

### Estudos envolvendo aspectos humanos da tomada de decisão

#### 1. Decision making without paradoxes (Belavkin 2006)

Este estudo pretende demonstrar a capacidade de se tomar decisões que se comparem àquelas tomadas por seres humanos quando expostos a situações de risco, mais especificamente quando colocados diante dos paradoxos de Ellsberg e Allais.

Os autores propuseram duas formas de modelar o problema, utilizando a arquitetura ACT-R, sendo a primeira sub-simbólica e a segunda simbólica.

Na forma sub-simbólica, os autores fizeram uma alteração no módulo de resolução de conflitos da arquitetura, que normalmente se baseia em distribuição gaussiana.

O ACT-R utiliza a seguinte função para utilidade:  $U_i = PG_i - C_i + noise(\sigma^2)$ , em que:

- $U_i$  é a utilidade esperada;

- $P_i$  é a probabilidade de se atingir o objetivo
- $G$  é o valor objetivo
- $C_i$  é o custo associado ao valor objetivo
- $noise(\sigma^2)$  é a função de variação, que permite a escolha entre alternativas semelhantes

A implementação substituiu essa função por uma função de distribuição de probabilidade (PDF) cumulativa.

Foi possível obter resultados próximos àqueles obtidos em estudos com seres humanos. Entretanto, o próprio autor discorre sobre o realismo desse tipo de solução, que se deve mais ao aprendizado da arquitetura.

Na representação simbólica, os autores atribuíram valores às seguintes propriedades:

- $U^s$  - ganho
- $P^s$  - probabilidade de ganho
- $U^f$  - perda
- $P^f$  - probabilidade de perda

As comparações entre as proposições recebiam valores qualitativos do tipo

- $a > b$  – preferência de a sobre b
- $a < b$  – preferência de b sobre a
- $a \sim b$  – indiferença entre a e b

Resumindo as proposições, teríamos:

Prop	A	B	C	D
$U^s$	\$300	>	\$100	\$0 ~ \$0
$P^s$	1/3	<	1	1/3 > 0
$U^f$	\$0	~	\$0	-\$300 < -\$100
$P^f$	2/3	<	0	2/3 > 1
União		<		>

Os resultados obtidos foram semelhantes aos encontrados em pesquisas com seres humanos, e demonstra uma maior aproximação da realidade, uma vez que considera a apresentação do problema.

## 2. Variability in human behavior (Wray et al, 2003)

Esse artigo apresenta uma implementação utilizando a arquitetura SOAR. Utilizou-se o SOAR para modelar agentes inteligentes em situações de combate urbano, para aplicações de treinamento de militares.

A variabilidade é importante nesse caso para que se tenha companheiros e adversários mais realistas, menos previsíveis (Wray et al 2005), e executar situações de combate menos repetitivas. É útil também para se trabalhar com situações em que erros são cometidos, uma vez que soldados humanos cometem erros.

Os autores categorizaram variabilidade com relação à fonte e com relação ao tipo. Com relação à fonte, pode ser (Figura 4):

- **Diferenças mentais:** nível de treinamento, educação, inteligência, capacidade de memória, etc.;

- **Diferenças físicas:** acuidade visual, forma física, saúde, etc.

Quanto ao tipo, a variabilidade pode ser:

- **Intra-sujeito:** um mesmo agente exposto a situações semelhantes agindo de formas diferentes;

- **Entre sujeitos:** agentes diferentes expostos à mesma situação agindo de formas diferentes.

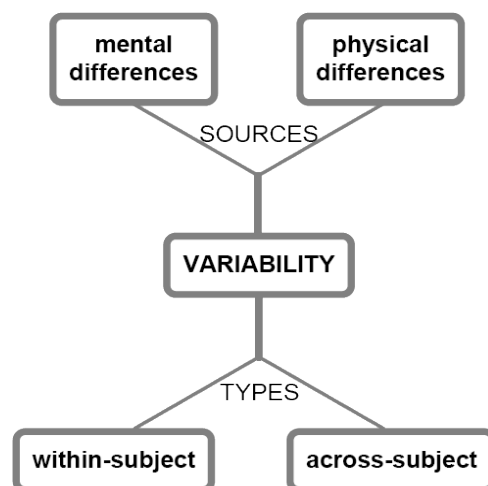


Figura 4 - Variabilidade

Utilizando essas categorizações, os autores produziram agentes, preocupando-se em não criar variabilidade arbitrária, mas sim uma variabilidade controlada.

A forma de se codificar isso foi através de uma alteração na estrutura padrão do SOAR.

Originalmente, o SOAR escolhe randomicamente entre proposições com probabilidades equivalentes. Entretanto, isso não é realista para a situação de combate militar, em que dificilmente um combatente adotaria o uso de pistolas caso rifles automáticos também estivessem disponíveis.

A alteração feita incluindo-se um modificador numérico às proposições. Esse valor numérico é um peso, que vai implicar em uma distribuição não uniforme, na escolha entre alternativas equivalentes.

Os autores afirmam a importância também, da variabilidade entre sujeitos, que não chegou a ser desenvolvida no estudo apresentado.

A proposta para resolver a variabilidade entre sujeitos é apresentada apenas na fase de design, e compreende o levantamento de valores fixos, gerados de forma randômica. Esses valores representariam um “perfil de variabilidade”, que pode representar diversos aspectos físicos e mentais dos agentes.

### 3. SORTS (Wintermute 2007)

Este artigo apresenta o uso do SOAR como agente inteligente para jogar jogos de estratégia em tempo real (RTS). O agente foi criado através da integração do SOAR ao *game engine* ORTS (Figura 5).

O objetivo dessa experiência é criar um jogador eletrônico que se comporte assim como um jogador humano se comportaria. Da mesma forma que a mente humana, o SOAR é primariamente simbólico, e manipulação de grande quantidade de dados é computacionalmente cara. Dessa forma, obtém-se uma arquitetura capaz de pensar de forma mais próxima à humana.

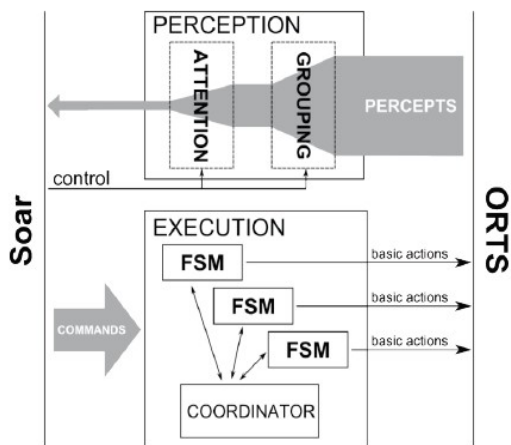


Figura 5 - SORTS

Os RTS permitem que se execute múltiplas tarefas de forma simultânea, e essas tarefas muitas vezes podem ser quebradas em sub-tarefas.

O SORTS utiliza como forma de seleção das ações a classificação hierárquica das tarefas, e a execução dessas tarefas de forma a respeitar essa hierarquia.

O exemplo apresentado no artigo sugere duas tarefas: construir uma base e combater um inimigo.

O sistema envia as tropas, e enquanto espera a chegada das tropas ao local de combate, inicia a construção da base. Ao se dar início ao combate, a tarefa de combate ganha prioridade e a construção é interrompida. A construção não será retomada até que o combate finalize.

Essa abordagem difere de outros sistemas, que tentam trocar de tarefas a maior quantidade de vezes possível, pois, tanto para o SORTS como para o jogador humano, trocar de tarefas implica em um *overhead* alto.

O SORTS foi colocado para participar de 3 competições.

**Jogo 1:** reunir a maior quantidade possível de recursos minerais em um intervalo fixo de tempo. A maior dificuldade foi a de coordenar personagens em locais próximos, com tarefas semelhantes, de forma a se evitar colisões nos caminhos.

Ainda assim, a implementação ganhou essa competição.

**Jogo 2:** destruir as bases adversárias, iniciando o jogo com 5 bases e 50 tanques. Foram utilizadas heurísticas para agrupamento de tanques e políticas para seleção de alvos. A dificuldade novamente esteve na movimentação, e acabou implicando na derrota.

**Jogo 3:** um RTS de verdade. Através da utilização de subdivisão de tarefas, priorização em múltiplas tarefas e agrupamento de personagens, o jogador eletrônico desenvolvido nesse artigo foi capaz de superar o único adversário inscrito em aproximadamente 60% das vezes.

Com o SORTS, os autores foram capazes de desenvolver um sistema capaz de jogar um RTS de uma forma muito mais parecida com a forma com que um humano joga.

Mostrou-se também que os jogos RTS podem ser considerados uma importante fonte de pesquisa para aspectos comuns das ciências cognitivas.

## O uso das emoções na tomada de decisão

As emoções estão intimamente ligadas ao processo decisório (Sloman et al. 1981). As emoções influenciam, balizam e agilizam o processo decisório de um ser humano. De forma análoga, para pensarmos em agentes computacionais que tenham capacidade decisória similar a de um ser humano, devemos considerar a existência de emoções.

Sistemas com emoções podem demonstrar comportamento mais flexível, com “erros” típicos de humanos e animais (Cañamero, 1997).

Uma interessante proposta para o uso de emoções na tomada de decisão foi feita por Roesener et al, 2006. Em seu artigo, ele apresenta uma arquitetura para arbitração de comportamentos orientada por emoções. O argumento principal é a aparente necessidade de desenvolvimento de agentes robóticos capazes de lidar com idosos e deficientes de forma autônoma. Esses robôs devem ser capazes de apresentar senso-comum semelhante ao de um humano, e tomar decisões considerando o ambiente dinâmico em que estão inseridos.

Os autores apresentam um modelo psicanalítico para emoções, no qual baseiam a arquitetura (Figura 6).

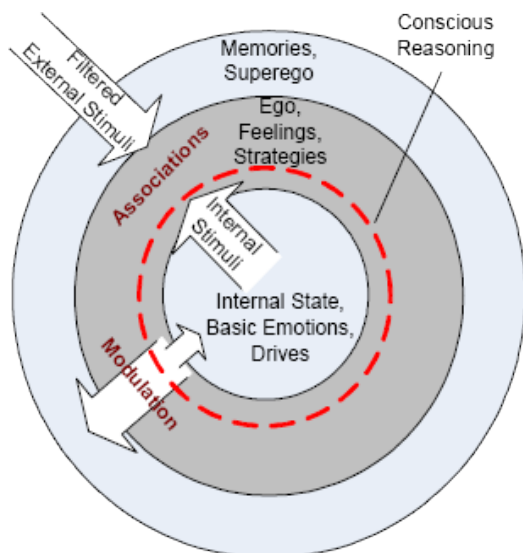


Figura 6 - Modelo psicanalítico de emoções

Esse modelo apresenta 3 níveis: estado interno, em que se enquadram as emoções básicas (medo, alegria, raiva); ego, em que se enquadram os sentimentos mais complexos; e superego, em que se armazenam as regras de convívio social. Os estímulos externos são responsáveis pelo mapeamento do ambiente e reconhecimento de padrões. Os estímulos

internos apresentam as necessidades internas do corpo, o estado interno, e as sensações boas *versus* ruins.

Inspirada na classificação de Damásio, 2004, é proposta uma arquitetura (Figura 7) que apresenta dois níveis de decisão. A decisão de baixo nível (pré-decisão) considera os drives e emoções básicas, e produz respostas rápidas. A decisão de alto nível (chamada simplesmente de decisão) considera o super-ego e as emoções mais complexas, além de ter uma forte interação com a memória episódica.

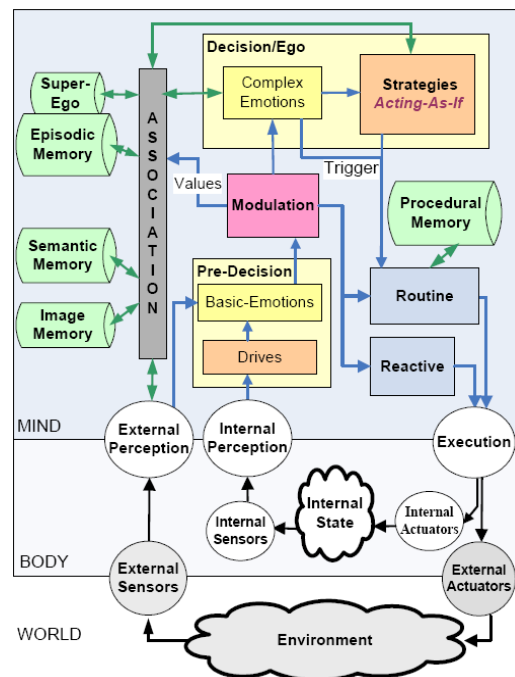


Figura 7 - Arquitetura de arbitração emocional

## Discussão

Foram apresentados alguns exemplos de sistemas que tentam reproduzir aspectos humanos da tomada de decisão. No entanto, esses sistemas ainda apresentam características aleatórias, e acabam apresentando semelhanças com a decisão humana mais em seus resultados do que de fato na forma como o processamento ocorre.

No campo das emoções, ainda são poucos os sistemas capazes de utilizar emoções para balizar a tomada de decisão.

A arquitetura sugerida por Roesener et al (2006) aparenta uma capacidade promissora de filtrar dados, agir conforme experiências anteriores, de uma forma mais humana, porém requer estudos mais profundos para provar sua eficácia.

Por fim, devemos tomar cuidado com a aplicação de decisão semelhante à de seres humanos em agentes cognitivos. Deve-se analisar friamente a necessidade de se reproduzir os aspectos humanos da tomada de decisão, uma vez que a criação desses sistemas pode exigir recursos financeiros desnecessários, além de introduzir a possibilidade das falhas e imperfeições inerentes ao raciocínio humano.

### **Conclusão**

Este artigo aponta, através do estudo de alguns casos, para a possibilidade de termos, no futuro, sistemas de cognição artificial capazes de tomar decisões semelhantes às decisões humanas.

Há, no entanto, um longo caminho a ser percorrido, que envolve um maior entendimento do mecanismo de decisão, especialmente no que diz respeito à interação com emoções. Existe também a necessidade de um consenso na classificação de emoções.

Com tudo isso, pode-se concluir que a área da decisão humana, embora estudada há muito tempo, mantém-se como uma área de estudos aberta, especialmente para os cientistas cognitivos.

### **Bibliografia**

Anderson, J.R., Bothell, D., Byrne, M. D., Douglass, S., Lebiere, C. and Qin, Y. An integrated theory of the mind. *Psychol Rev*, 111(4):1036–1060, October 2004.

Lehman, Laird, and Rosenbloom, 2006 A Gentle Introduction to Soar: 2006 update

Imbert, R., and de Antonio, A. 2005. An Emotional Architecture for Virtual Characters. In *Virtual Storytelling: Using Virtual Reality Technologies for Storytelling*, ed. G. Subsol, 63-72. Berlin: Springer.

Busemeyer, J. R. (2001). Dynamic decision making. In *International encyclopedia of the social and behavioral sciences: Methodology, mathematics, and computer science*. Amsterdam: Pergamon.

Belavkin, R. V. (2006). Towards a theory of decision-making without paradoxes. In *Proceedings of the Seventh International Conference on Cognitive Modeling* (pp. 38-43). Trieste, Italy.

Wray, R.E., Laird, J. E. (2003) "Variability in Human Behavior Modeling for Military Simulations", *Behavior Representation in Modeling and Simulation Conference*. Scottsdale, AZ.

Wray, R.E., Laird, J.E., Nuxoll, A., Stokes, D., Kerfoot, A. Synthetic adversaries for urban combat training. *AI Magazine*, 26(3):82--92, 2005.

Wintermute, S., Xu, J., and Laird, J.E. SORTS: A Human-Level Approach to Real-Time Strategy AI. *Proceedings of the Third Artificial Intelligence and Interactive Digital Entertainment Conference (AIIDE-07)*, Stanford, California

Wintermute, S., Xu, J., Irizarry, J. 2007. SORTS: Integrating Soar with a Real-Time Strategy Game (Report No. CCA-TR-2007-01). Ann Arbor, MI: Center for Cognitive Architecture, University of Michigan.

Cañamero, D. (1997) A Hormonal Model of Emotions for Behavior Control. *Proceedings of the 4th European Conference on Artificial Life (ECAL'97)*.

Damásio, António R., *O Erro de Descartes. Emoção, Razão e Cérebro Humano*, 12ª edição, Lisboa, Publicações Europa-América, 1995, (col. Forum da Ciência, 29).

Roesener, C. Lorenz, B. Vock, K. Fodor, G. (2006) Emotional Behavior Arbitration for Automation and Robotic Systems - *Industrial Informatics, 2006 IEEE International Conference on*