

## 4 Arquitetura BDI – *Belief Desire Intention*

Entre os diversos seguimentos da Inteligência Artificial (IA), alguns são dedicados ao estudo de arquiteturas de agentes deliberativos baseadas em estados mentais. Entre estes estados mentais estão crenças, desejos e intenções. As arquiteturas que utilizam esta tríade de estados são denominadas, na literatura, como arquiteturas *Belief, Desire and Intention (BDI)*. Os conceitos sobre este assunto são abordados neste capítulo, assim como ferramentas que os utilizam.

### 4.1. Estados Mentais

As arquiteturas baseadas em estados mentais fazem parte da arquitetura de agentes deliberativos. Os estados mentais humanos possuem um vínculo com o mundo através do qual estabelecem sua existência e significância. A característica pela qual os estados mentais humanos referem-se a objetos ou situações do mundo é chamada de intencionalidade. Crença, desejo, expectativa, capacidade e intenção são exemplos de estados intencionais. O nome dado a esta abordagem que utiliza estes estados para modelagem de agentes deliberativos é abordagem mentalista (Zamberlam e Giraffa, 2001). Na **Figura 20**, está a representação esquemática de um exemplo de uma arquitetura baseada em estados mentais, extraída do trabalho de Corrêa (Juchem e Bastos, 2001, p.12).

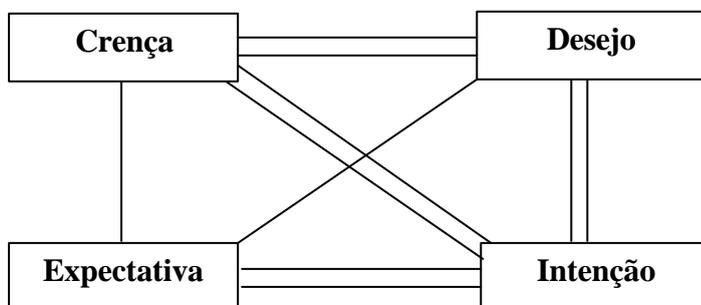


Figura 20. Arquitetura baseada em 4 estados mentais (Juchem e Bastos, 2001, p.12).

Esta intencionalidade, encontrada no estados mentais, indica uma certa direcionalidade, ou seja, uma propriedade de direcionamento do mundo para o agente e vice-versa. Como exemplo, a afirmação “a porta está fechada” é uma crença sobre a porta, do mundo para o agente e “entrar na sala” é um desejo, do agente para o mundo. Logo, intencionalidade não está ligada ao estado mental “intenção”, mas sim a todos os estados que possuem a propriedade de direcionalidade (Zamberlam e Giraffa, 2001).

Na literatura são encontradas várias denominações associadas a estes estados mentais. Os conceitos individuais associados a esses estados são descritos nas subseções a seguir.

#### **4.1.1. Crenças**

A crença contém a visão fundamental do agente com relação ao seu ambiente, ele as usa para expressar suas expectativas sobre possíveis estados futuros. As crenças de um agente podem ser vistas como o provável estado do ambiente, isto é, como um componente informativo do estado do sistema.

Um agente pode ter crenças sobre o mundo, sobre outros agentes, sobre interações com outros agentes e crenças sobre suas próprias crenças. As crenças podem ser até contraditórias.

Segundo Corrêa, (Zamberlam e Giraffa, 2001, p.6), uma crença é um estado mental intencional fundamental para as interações dos agentes com noção idêntica a de conhecimento.

#### **4.1.2. Desejos**

Diferentes dos estados prováveis representados pelas crenças, os desejos representam estados desejáveis que o sistema poderia apresentar. Para um desejo deixar de ser “desejável” e tornar-se algo mais real, falta o conhecimento inerente a ele e o contexto favorável à sua realização.

Os desejos motivam o agente a agir de forma a realizar metas, tais ações são realizadas através das intenções causadas pelos desejos.

Segundo Corrêa, (Zamberlam e Giraffa, 2001, p.7), com base na Teoria das Situações, um desejo  $D$  de um agente  $A$  é apresentado como um estado mental intencional e motivador pela seguinte situação:

$D \models \{ \langle Des, A, P, e, l_0, t_0, v \rangle \}$ , em que:

- $Des$  é a relação para a representação de desejo;
- $A$  representa o agente que possui o desejo;
- $P$  é uma proposição;
- $e$  é uma situação que informa sobre a satisfação, não satisfação, urgência, intensidade e insistência do desejo  $D$ ;
- $l_0$  é a localização espacial (local) associada à ocorrência do desejo  $D$ ;
- $t_0$  é o tempo associado à ocorrência do desejo  $D$ ;
- $v$  1, se ocorre o desejo  $D$  ao agente  $A$   
0, se não ocorre o desejo  $D$  ao agente  $A$

### 4.1.3. Intenções

As intenções dos agentes são um subconjunto dos desejos. Se um agente decide seguir uma meta específica, então esta meta torna-se uma intenção.

Em (Zamberlam e Giraffa, 2001) indica-se que a base para a definição do conceito de intenção está fortemente associada aos trabalhos filosóficos de *Michael Bratman*. Este autor claramente distingue o conceito de fazer as coisas intencionalmente, na forma de uma ação, e possuir intenção de fazê-las, ou seja, tendo como base um estado mental.

As intenções determinam o processo de raciocínio prático, pois a propriedade mais óbvia das intenções é que elas determinam as ações a serem realizadas.

Uma vez adotada uma intenção haverá um direcionamento no raciocínio prático futuro, ou seja, enquanto se tem uma intenção particular haverá consideração por ações que são consistentes para a realização desta intenção. Por exemplo, se uma pessoa tem a intenção de tornar-se um atleta, ela terá que excluir opções que dificultem a realização desta intenção como a ingestão de bebidas

alcoólicas regularmente. A escolha dessas opções deverá ser feita cuidadosamente para não surgirem sub-intenções que sejam inconsistentes.

#### 4.2. Arquitetura *Belief-Desire-Intention*

Entre as arquiteturas de estados mentais está a arquitetura *BDI*. Segundo Giraffa, (Zamberlam e Giraffa, 2001, p.5), as idéias básicas da abordagem *BDI* são: descrever o processo interno de um agente utilizando o conjunto de estados mentais: crença, desejo e intenções; e definir uma arquitetura de controle através da qual o agente possa selecionar o curso de suas ações. Além destes componentes, algumas arquiteturas *BDI* usam o conceito de planos. Planos seriam o conjunto de sub-tarefas que deve ser seguido, quando gerada uma intenção, para a realização de uma tarefa sobre o ambiente (Weiss, 1999) (ver **Figura 21**).

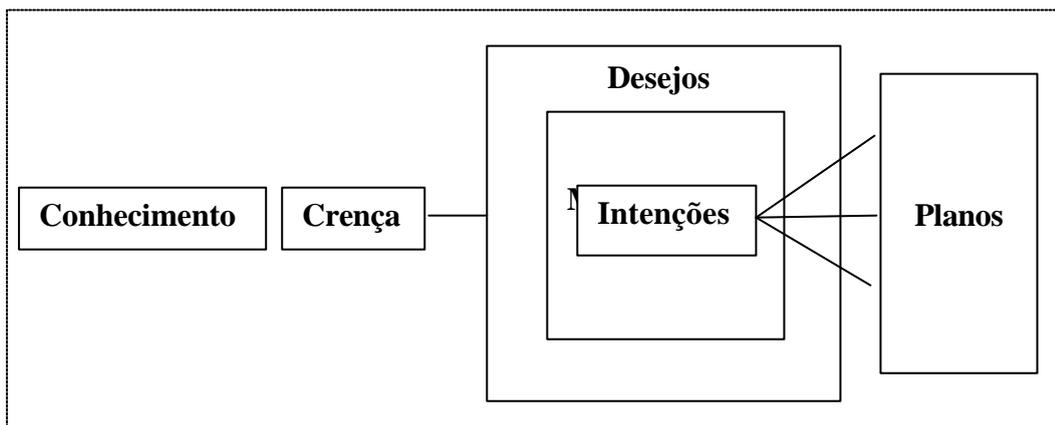


Figura 21. Esquema de uma Arquitetura BDI.

A referência relativa ao conhecimento do agente (**Figura 21**), embora não seja citada explicitamente no modelo, é válida. A razão para isso está fundamentada no fato desse modelo se preocupar com o aspecto comportamental do agente, tomando como pressuposto básico a existência de um “conhecimento” sobre o qual o agente vai trabalhar, ou seja, como realizar determinadas tarefas ou ações.

Ao projetar um agente baseando-se no modelo *BDI*, são especificados suas crenças e seus desejos, mas a escolha das intenções fica sob a responsabilidade do

próprio agente, isto é, de uma auto-análise desses estados inicialmente disponíveis.

O problema no desenvolvimento do raciocínio prático do agente é como alcançar um balanceamento entre estes diferentes conceitos, principalmente no que se refere a abandonar intenções. Às vezes é necessário que um agente deixe de considerar uma intenção, porém as reconsiderações têm um custo computacional. Este é um dos principais problemas para a utilização de agentes deliberativos em jogos. O fato de o agente ter que reavaliar sua situação, porque a intenção escolhida por ele não pôde ser realizada, consome um certo custo de processamento que pode influenciar no desempenho de outros elementos do jogo. Dependendo da complexidade de realizar as intenções de um agente em um jogo, talvez não seja possível ter várias intenções ao mesmo tempo. Uma possível sugestão seria o agente determinar uma intenção, cumprir todas as tarefas relacionadas a ela, e depois determinar outra intenção.

Em (Weiss, 1999) são definidos sete componentes importantes de uma arquitetura *BDI* (ver **Figura 22**) :

- Um conjunto de crenças atuais, que representam as informações que o agente tem sobre seu ambiente atual;
- Uma função de revisão de crenças, que a partir da entrada percebida e com as crenças atuais do agente, determina um novo conjunto de crenças;
- Uma função geradora de opções, que determina as opções disponíveis para o agente, ou seja, seus desejos, tendo como base suas crenças atuais sobre seu ambiente e suas intenções atuais;
- Um conjunto de desejos atuais, representando possíveis cursos de ações disponíveis para o agente;
- Uma função filtro, que representa o processo de deliberação do agente, e que determina as intenções dos agentes, tendo como base suas crenças, desejos e intenções atuais;
- Um conjunto de intenções atuais, representando o foco atual do agente;
- Uma função de seleção de ação, que determina uma ação para executar, tendo como base as intenções atuais.

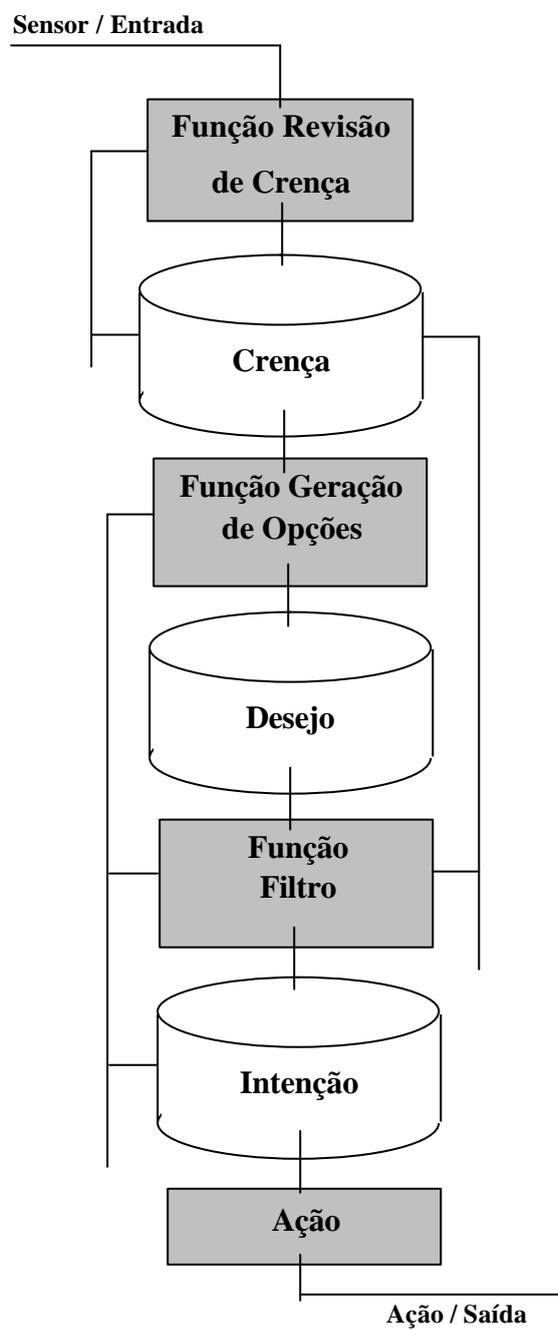


Figura 22. Componentes de uma Arquitetura BDI (Weiss, 1999).

### 4.3. Ferramentas

A arquitetura *BDI* está instanciada em diversos exemplos, entre eles destacam-se: *Procedural Reasoning System* (PRS) e *Distributed Multiagent Reasoning System* (dMars).

#### 4.3.1. *Procedural Reasoning System*

O Sistema de Raciocínio Procedimental (*Procedural Reasoning System - PRS*) é um *framework* para construção de sistemas de raciocínio em tempo real que podem realizar tarefas complexas em ambientes dinâmicos. É uma arquitetura *BDI* genérica desenvolvida pelo *Stanford Research Institute International* (Prs-Guide, 2001).

A arquitetura *PRS* é constituída pelos seguintes componentes (Prs-Guide, 2001):

- um banco de dados contendo crenças correntes e fatos sobre o mundo;
- um conjunto de objetivos que estão esperando para serem realizados;
- um conjunto de planos, chamados Atos (*Acts*), descrevendo como as seqüências de ações e testes precisam ser executadas para que os objetivos sejam realizados, ou para que haja uma reação para uma situação específica;
- intenções contendo planos que foram escolhidos para uma eventual execução;
- um interpretador que é responsável por manipular todos estes componentes, selecionando os Atos mais apropriados.

O funcionamento do *PRS* está fundamentado na especificação dos Atos que serão usados para realizar os desejos. Por exemplo, para realizar o desejo de “lavar as roupas” é necessária a seqüência de ações composta pelas ações: “colocar as roupas na máquina de lavar”, “colocar sabão na máquina de lavar”, “ativar o procedimento de lavar da máquina” e “esperar 45 minutos”. Desta forma, o PRS oferece um ambiente para que este tipo de informação, entre outras,

seja especificada e executada. Na **Figura 23** está ilustrado o *PRS* e as relações entre seus componentes.

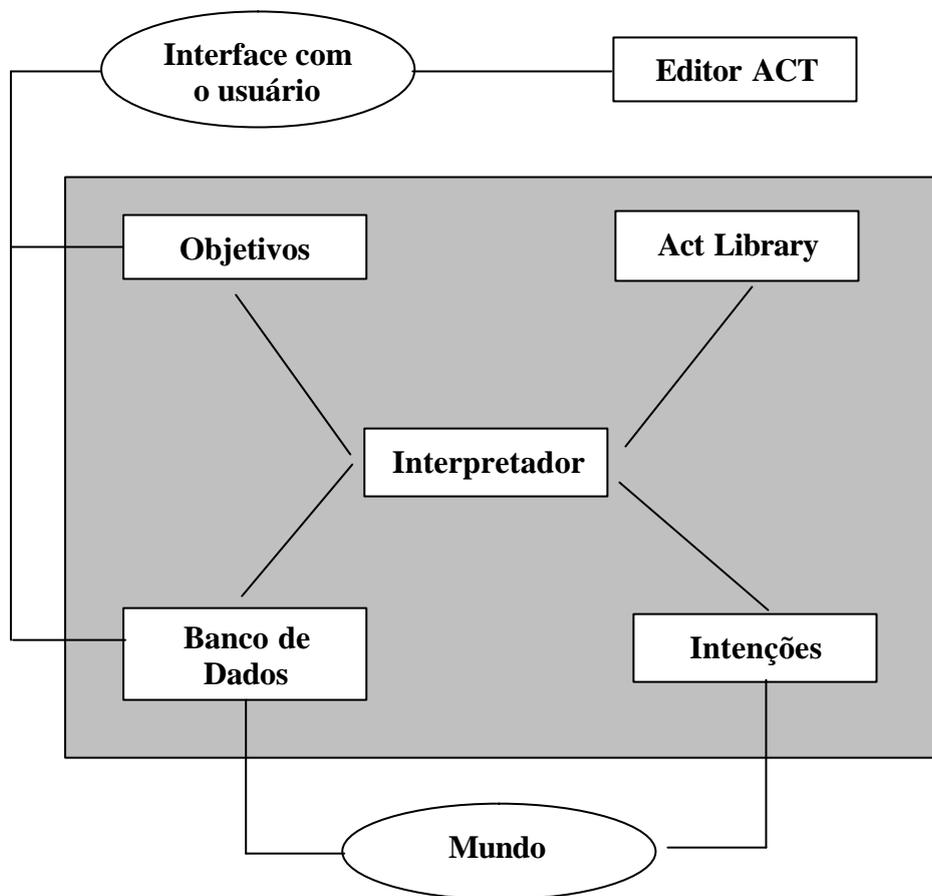


Figura 23. Componentes da Arquitetura *PRS* (Prs-Guide, 2001).

A versão original do *PRS* em vez de utilizar a representação de ações através de Atos utilizava representações baseadas em Áreas de Conhecimento (*Knowledge Areas- KAs*). O formalismo baseado em Atos apresenta algumas vantagens em relação ao *KAs*, como a utilização de uma sintaxe mais intuitiva.

#### 4.3.2. dMars

Baseada na arquitetura *PRS*, o Sistema MultiAgente de Raciocínio Distribuído (*distributed Multi-Agent Reasoning System - dMars*) foi desenvolvido em C++ e faz parte de um ambiente para desenvolvimento de agentes (Zamberlam e Giraffa, 2001). O *dMars* tem como principal foco aplicações utilizadas em domínios dinâmicos onde há conhecimento incerto e complexo. Exemplos de

aplicações estão no campo das telecomunicações, tráfego aéreo, viagens espaciais, entre outras. É composto por uma ferramenta gráfica, que fornece uma interface intuitiva para as atividades de desenvolvimento e *debugging*.

Segundo (D’Inverno *et al.*, 1998), uma arquitetura *BDI* é tipicamente formada por crenças, desejos, intenções e uma biblioteca de planos. Cada plano contém vários componentes:

- **Trigger ou Condição de Execução:** especifica as circunstâncias sob as quais o plano deve ser considerado. Normalmente são descritas em termos de eventos. Por exemplo, o plano “tomar chá” pode ser acionado pelo evento “sede”;
- **Pré-Condições:** estabelece o contexto em que o plano pode ser iniciado. Ou seja, é necessário que algumas circunstâncias sejam satisfeitas para dar início a execução do plano. No plano “tomar chá”, a execução só pode ser iniciada se a pré-condição “ter saquinhos de chá” for satisfeita;
- **Condições de Manutenção:** indicam algumas condições que devem permanecer verdadeiras enquanto o plano estiver sendo executado;
- **Corpo:** define um curso bastante complexo de ações, que podem consistir em objetivos, subobjetivos, ou ações primitivas. No plano “tomar chá”, são encontradas as seguintes ações: “ferver água”, “colocar o saquinho de chá no copo” e “adicionar água no copo”. A primeira ação “ferver água”, pode ser considerada um subobjetivo (alguma coisa que precisa ser atingida, quando o plano alcança este ponto da execução), e as demais são ações primitivas, chamadas de procedimentos.

Os agentes *dMars* monitoram o mundo, seus estados internos próprios e quaisquer eventos que são percebidos e colocados em uma fila de eventos. Assim como no *PRS*, há um interpretador (ver **Figura 24**) que é responsável por gerenciar as operações do agente. O interpretador executa o seguinte ciclo (D’Inverno *et al.*, 1998):

- observa o mundo e o estado interno do agente e atualiza a fila de eventos para indicar os eventos que foram observados;

- gera novos desejos possíveis (tarefas) a partir de planos encontrados com eventos compatíveis na lista de eventos;
- seleciona deste conjunto de planos compatíveis um evento para execução;
- coloca o evento escolhido em uma pilha nova ou já existente de intenção, dependendo se o evento é ou não é um subobjetivo;
- seleciona uma pilha de intenção, pega o plano do topo e executa o próximo passo do plano corrente: se o passo for uma ação, executa esta ação; se não, se for um sub-objetivo, envia este sub-objetivo para a fila de eventos.

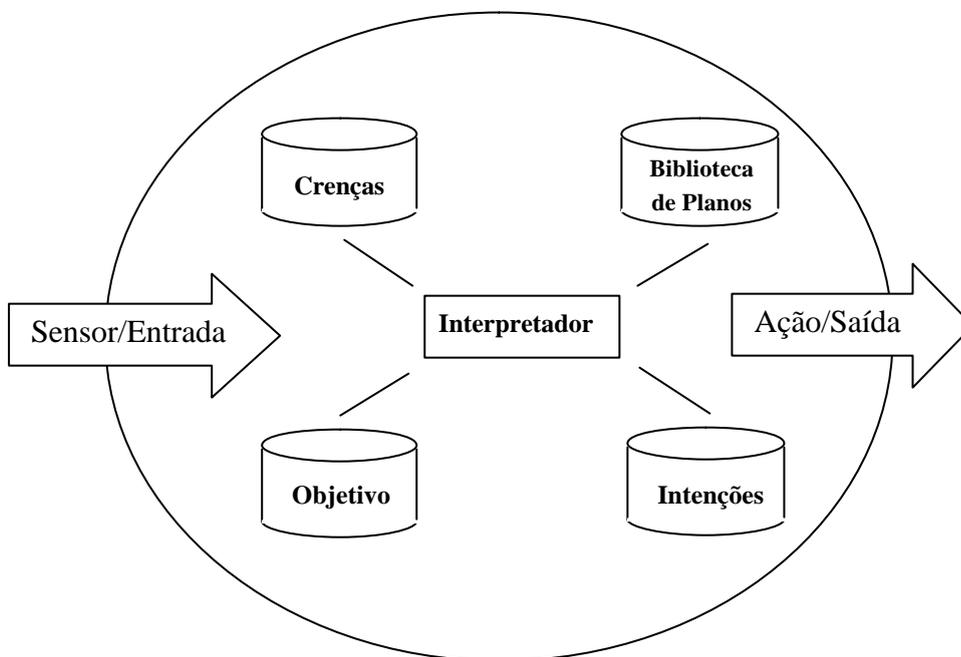


Figura 24. Componentes da Arquitetura *dMars*.

Dessa forma, a execução básica do modelo de agentes modelados em *dMars* tem início quando um plano inicia a execução e seus subobjetivos são colocados na fila de eventos, onde, um por um, acionam planos que realizam seus subobjetivos, os tornando ativos.

Os agentes *dMars* não planejam tudo, como no *PRS*, pois os planos precisam ser definidos pelo programador durante a fase de modelagem do agente. O planejamento exercido pelos agentes consiste na expansão dos subobjetivos sensíveis a contexto.