

# 1 Representação de Conhecimento

## 1.1 Introdução

Embora não tenha sido formalmente reconhecida como tal pelos primeiros pesquisadores de Inteligência Artificial, a *Representação de Conhecimento* surgiu lentamente como uma área de estudos distinta a partir de meados da década de 60. Nesta época, começaram a surgir várias abordagens diferentes para representação de conhecimento. Estas abordagens desenvolveram-se posteriormente e originaram a maioria dos paradigmas conhecidos hoje em dia.

As abordagens mais importantes atualmente são conhecidas como *redes semânticas*, *lógica de primeira ordem*, *frames* e *regras de produção*. Cada um destes paradigmas subdivide-se segundo diversas variantes, havendo inclusive certas composições de mais de um dos paradigmas principais.

## 1.2 Representação de Conhecimento

Numa primeira observação dos paradigmas existentes, um esquema de representação do conhecimento pode ser descrito como "um conjunto de convenções sintáticas e semânticas que tornam possível a descrição das coisas" (1).

Mas J. Delgrande (2) procura definir a representação de conhecimento de forma mais acurada. Segundo ele, por *conhecimento* entende-se uma crença verdadeira justificada. Por *representação* entende-se uma codificação em uma estrutura de dados. Então, representação de conhecimento significa a codificação de crenças justificadas em estruturas de dados apropriadas.

## 1.3 Base de Conhecimento

A estrutura de dados utilizada para representar conhecimento é chamada de base de conhecimento.

Uma *base de conhecimento* pode ser definida como um par ordenado  $(BC_0, \phi_L)$ , onde  $BC_0$  é uma coleção de afirmações em uma linguagem  $L$ , e  $\phi_L$  é uma regra que permite estabelecer quais afirmações seguem (ou não) da base de conhecimento.

Por exemplo, se:

$$BC_0 = \{ \text{estudante(João)}, \text{supervisor(João, Maria)} \}$$

e  $\phi_L$  é a relação de derivabilidade em  $L$  (onde  $L$  é a lógica de predicados de primeira ordem), então uma afirmação  $a$  só poderá pertencer à base de conhecimentos (BC) se a puder ser derivada de  $BC_0$  por  $\phi_L$ , ou seja:

$$a \in BC \text{ sss } BC_0 \phi_L a$$

Assim, a base de conhecimento  $BC$  é definida como um conjunto associado ao fechamento de  $BC_0$  pela operação  $\phi_L$ .

Há, então, duas maneiras de interpretar uma  $BC$ , ou bem como um conjunto  $BC$  de afirmações verdadeiras, ou também como um mecanismo  $(BC_0, \phi_L)$  usado para construir este conjunto. A primeira interpretação pode ser classificada como extensional e a segunda como intencional.

Afirmções em uma base de conhecimento podem ser associadas a um valor verdade relativo a um mundo ou domínio de discurso. A atribuição de valores verdade a afirmações é feita em termos de uma função semântica  $v$ :

Uma função semântica  $v$  associa a cada afirmação um valor verdade:

$$v : \text{Afirmação} \rightarrow \{ V, F \}$$

Um método padrão para atribuição de valores verdade é o de A. Tarski, o qual trata a interpretação  $I$  da  $BC$  como uma tripla  $\langle D, R, U \rangle$ , onde  $D$  é o conjunto de indivíduos no domínio de discurso, e  $R$  e  $U$  são respectivamente as relações e funções entre indivíduos que ocorrem no mundo. Assim,  $I$  estabelece um mapeamento das afirmações no conjunto  $\{ V, F \}$ . Uma interpretação  $I$  é um modelo de uma  $BC$  se e somente se todas as sentenças na  $BC$  são verdadeiras na interpretação  $I$ .

### 1.3.1 Critérios para Análise de Linguagens de Representação de Conhecimento

É possível admitir que um determinado formalismo pode representar tudo o que um sistema pode vir a conhecer. A questão principal, então, é saber quão facilmente as coisas podem ser representadas num ou noutro formalismo. Jackson, citado em (1) propõe três critérios para análise de linguagens de representação de conhecimento:

a) *Adequação Lógica*. É a sua capacidade ou potencial expressivo.

b) *Poder Heurístico*. É a facilidade para interpretar as estruturas de dados com vistas a resolver problemas.

c) *Conveniência Notacional*. É a naturalidade da expressão do conhecimento e a facilidade para sua leitura ou escrita para entender o que está expresso à margem do mecanismo de interpretação da máquina.

### 1.3.2 Teorias Semânticas de Conhecimento

Bases de conhecimento devem se referir a alguma coisa, e a formalização desta referência é uma *teoria semântica*.

J. Delgrande (2) considera três teorias semânticas: a dos *mundos possíveis*, a *sintática* e a *situacional*.

Para exemplificar a teoria de mundos possíveis, considere-se  $BC_0 = \{ \text{professor}(\text{João}) \}$ . Esta  $BC$  pode ser vista como aquilo que se sabe sobre o mundo. Ela cria restrições sobre o mundo, isto é, neste mundo, João deve ser um professor. Também há irrestrições: se Pedro é uma pessoa, ele pode ser ou não um professor no mundo real. Um sistema acredita em uma sentença  $a$  somente se ela é verdadeira em todos os mundos possíveis, de acordo com a  $BC$ , isto é, se  $a$  é verdadeira em todas as interpretações que são modelos da  $BC$ .

A abordagem sintática faz a estrutura do modelo conter, ou ser isomorfa a, um conjunto explícito de sentenças. Dada uma  $BC$ , tudo o que se pode conhecer está na  $BC$ . Porém se  $a \vee b$  é verdadeiro, não se pode afirmar que  $b \vee a$  também seja se  $b \vee a$  não estiver explicitamente na  $BC_0$ . Por isso, deve-se especificar quais as crenças que seguem de um dado conjunto.

A abordagem situacional diz que a  $BC$  é relevante para algumas informações e irrelevante para outras. Esta abordagem permite distinguir crenças implícitas e explícitas.

## 1.4 Natureza do Conhecimento

### 1.4.1 Crença, Hipótese e Asserção

O tratamento de noções como crença ou hipótese de uma afirmação corresponde ao seu status epistêmico. O *status epistêmico* das afirmações é especialmente importante nos assim chamados sistemas de manutenção de verdades. Nestes sistemas, certas afirmações são colocadas a título de simples hipóteses. Se for comprovado que essas hipóteses não são verdadeiras, então todas as afirmações que se apoiavam nelas deixam de ser consideradas como verdadeiras.

A confiança total na asserção representada por uma afirmação corresponde ao *status assercional* da afirmação. Neste caso, tem-se uma crença que não pode ser desmentida; pois foi estabelecido que ela é verdadeira a título de axioma.

A noção de crença é definida inicialmente da seguinte maneira: dado um sistema de base de conhecimento ou agente  $A = (BC_0, \phi_L)$ . O agente  $A$  acredita na afirmação  $a$  apenas se  $a$  aparece na  $BC_0$  de  $A$ . Ou seja:

$$a \in \text{crença}(A) \text{ sss } BC_0 \phi_L a$$

Sistemas lógicos de primeira ordem podem ser aumentados com um operador modal  $K$ , onde  $Ka$  pode ser lido como "acredita-se que  $a$  seja verdadeiro". Assim, a afirmação  $Ka \rightarrow a$  significará que "se é acreditado que  $a$  é verdadeiro, então  $a$  realmente é verdadeiro". Se este axioma for verdadeiro numa  $BC$ , então todas as afirmações contidas na  $BC$  são de fato verdadeiras, e o status assercional é válido para todo o conhecimento da  $BC$ . Caso contrário, o conhecimento contido na  $BC$  corresponde a noção de hipótese. Uma *hipótese* é uma crença da qual não se conhece o valor verdade.

### 1.4.2 Semântica

Quando se procura extrair conhecimento de sentenças em linguagem natural, depara-se frequentemente com problemas de ordem semântica. Por exemplo, a afirmação "elefantes são cinza" tem pelo menos três interpretações:

- a) "todos os elefantes são cinza";
- b) "tipicamente elefantes são cinza";
- c) "elefantes são cinza, entretanto existem exceções".

Estas possibilidades levam a três abordagens diferentes para especificar o significado de uma afirmação. No caso (a), existe a idéia de que um termo pode ser exatamente identificado com uma

coleção de propriedades. Esta abordagem tem sido chamada de *teoria tradicional do significado*, e é expressa diretamente na lógica clássica de predicados de primeira ordem.

Mas termos como "elefante" são de tipo natural e as características normais de identificação destes elementos podem dar lugar a exceções, e são explicadas por mecanismos de raciocínio mais profundos.

No caso (b) um termo é identificado com um membro típico. Esta é a essência da *teoria de protótipos*.

A terceira leitura (c) mantém a sentença geral enquanto admite exceções. Esta abordagem leva à *teoria da herança não-estrita* e requer uma formalização específica.

### 1.4.3 Incompleteza

Uma base de conhecimentos não pode ser completa, no sentido de que ela não pode ser grande o suficiente para conter todo o conhecimento sobre o mundo. Uma *BC* só pode ser hipoteticamente completa para domínios muito restritos do conhecimento. Mas mesmo assim, não se pode admitir que o conhecimento sobre esse domínio cresça, pois se tal fato suceder, a *BC* se tornará novamente incompleta.

Assume-se então que aquilo que não está representado na *BC* é falso de acordo com o conhecimento existente. Essa abordagem é chamada de *hipótese do mundo fechado* (em inglês usa-se a sigla *CWA* para *Closed World Assumption*). Pela hipótese do mundo fechado, se uma pergunta a *BC* não obtém resposta afirmativa, ela não é respondida como desconhecida mas como falsa.

O teorema da incompleteza de Gödel estabelece que há limites inerentes à completeza de uma base de conhecimento quando um esquema de representação de conhecimento é suficientemente poderoso.

Expressar incompleteza envolve dizer que alguma coisa tem uma propriedade sem identificar a coisa ou dizer que todas as coisas numa classe tem uma propriedade sem dizer o que está na classe e permitir a possibilidade de duas expressões não idênticas nomearem o mesmo objeto. Para expressar, por exemplo, que existem estudantes desconhecidos podemos usar

$$(\exists x) \text{estudante}(x) \wedge \neg K\text{estudante}(x)$$

A incompleteza das bases de conhecimento leva, como consequência à não-monotonicidade.

### 1.4.4 Não-Monotonicidade

Se *BC* e *BC'* são bases de conhecimento e  $\phi$  é a relação de derivabilidade da lógica de primeira ordem, e o conjunto de axiomas de *BC* está contido no conjunto de axiomas de *BC'*, ou seja:

$$BC = (BC_0, \phi) \quad \text{e} \quad BC' = (BC_0 \cup A, \phi)$$

então  $BC \subseteq BC'$ .

Esta propriedade torna a lógica de primeira ordem monotônica. A monotonicidade gera problemas por não tratar alguns casos:

**a) Atribuições por falta.** "A menos que você saiba o contrário, assuma que um elefante é cinza";

**b) Incompleteza do conhecimento.** Pela hipótese do mundo fechado: "a menos que você saiba outra coisa, assuma que um objeto não é um estudante".

**c) Herança default.** Considerando o caso prototípico da descrição de aves em que estabelecemos que elas voam, isto pode ser falso tanto para indivíduos (Piu-piu) ou classes (pingüim). Isto pode ser entendido como uma regra da forma: "a menos que você saiba o contrário, assuma que uma ave voa.

A lógica default é uma formalização para não-monotonicidade. Ela inclui a lógica de primeira ordem e tem regras de inferência da forma:

$$pessoa(x): Mmora(x, Florianópolis) \rightarrow mora(x, Florianópolis)$$

Isto é, se alguém é uma pessoa e pode ser consistentemente assumido que ela mora em Florianópolis (isto é, não pode ser derivado da *BC* que ela não mora em Florianópolis) então assuma que ela mora em Florianópolis.

Uma teoria default pode ter mais de um conjunto possível de teoremas dependendo da ordem de aplicação das regras de inferência default.

Outra abordagem ao raciocínio não-monotônico é a noção de circunscrição de J. McCarthy. Intuitivamente, a idéia de circunscrição pode ser pensada como uma regra de conjectura, a qual permite pular para certas conclusões quando houver falta de conhecimento. Isto é obtido estabelecendo que todos os objetos que demonstram ter a propriedade *p*, dado algum conjunto de fatos, são de fato os únicos objetos que satisfazem *p*.

Outra abordagem a não-monotonicidade atribuída a McDermot & Doyle e depois re-sistematizada por Moore, utiliza lógica de primeira ordem aumentada com um operador sentencial *M* cuja interpretação informal é "é consistente". Assim:

$$(\forall x) ave(x) \wedge Mvoa(x) \rightarrow voa(x)$$

#### 1.4.5 Inconsistência

Grandes bases de conhecimento são inerentemente inconsistentes. Do ponto de vista de representação de conhecimento, trabalhar com inconsistência envolve dois temas: assimilação da informação inconsistente e a acomodação da informação inconsistente (3).

Uma solução para trabalhar com inconsistência é restringir a ordem em que as inferências são tentadas. Assim, olha-se primeiramente as propriedades da classe antes das da superclasse. O problema é que então a semântica desta relação de herança não é clara. É necessário estabelecer

claramente quais propriedades são herdadas por alguma classe a partir de suas superclasses, quando as superclasses determinam propriedades inconsistentes.

Outro problema com lógica de primeira ordem são os paradoxos da implicação como:

$$A \rightarrow (B \rightarrow A) \text{ e } (A \wedge \neg A) \rightarrow B$$

Uma maneira de eliminar este problema é modificar o conjunto de axiomas e revisar a noção de prova tal que a prova de  $B$  pelas hipóteses  $A_1, \dots, A_n$  é bem formada somente se ela realmente usa cada hipótese em algum passo. Chamamos a esta abordagem de *lógica com implicação relevante*.

Uma nova proposta para tratamento de inconsistência (Borgida & Imielinsky) trata a  $BC$  como membro de um comitê, cada membro com um ponto de vista particular. A derivabilidade com respeito a  $BC$  é determinada por uma regra de decisão do comitê. Exemplos de regras alternativas são:

**a) Regra cética:**  $BC \models a$  se para todo ponto de vista  $PV$ ,  $PV \models a$ .

**b) Regra crédula:**  $BC \models a$  se existe ao menos um ponto de vista  $PV$ , tal que  $PV \models a$  e não existe nenhum ponto de vista  $PV'$ , tal que  $PV' \models \neg a$ .

Outra abordagem (Belnap) é a *lógica de quatro valores* (verdadeiro, falso, desconhecido e inconsistente). Uma maneira de resolver o problema da acomodação é tratar fórmulas "suspeitas" como hipóteses. É assumido que fatos numa  $BC$  são de três tipos diferentes: fórmulas atômicas básicas como *estudante(João)*, afirmações gerais hipotetizadas como "Elefantes hipoteticamente são mamíferos" e sentenças arbitrárias que se presume estejam além da refutação. Dada esta hipótese, três pontos são explorados:

**a)** Como gerar e manter a consistência entre as afirmações gerais hipotetizadas dadas fórmulas atômicas básicas e outras afirmações;

**b)** Como prescrever formalmente o conjunto de afirmações gerais que podem ser hipotetizadas e

**c)** Como este sistema de formulação de hipóteses pode interagir com um sistema para raciocinar dedutivamente com hipóteses e conhecimento.

Outra abordagem de acomodação (Borgida) toma a visão de que leis gerais são úteis e devem estar disponíveis no esquema de representação de conhecimento com um mecanismo para acomodação de exceções. Por exemplo, "aves voam"; se "pingüins são aves e não voam então "aves voam ou a ave é um pingüim".

#### 1.4.6 Inexatidão

Uma  $BC$  pode conter informações que não sejam verdadeiras no mundo real. O mínimo que se pode fazer com a inexatidão é falar sobre ela. Assim, uma informação como: "uma entidade conhecida como estudante na  $BC$  é, de fato, um estudante no domínio de discurso" pode ser representado pela expressão:

$$(\forall x) Kestudante(x) \rightarrow estudante(x)$$

Ainda, "a idade de uma pessoa registrada na *BC* pode estar imprecisa por no máximo dois anos" pode ser representada por:

$$(\forall x)(\forall y)[Kidade(x) = y \rightarrow |idadeReal(x) - y| < 2]$$

Se for assumido que  $Ka \rightarrow a$ , então a *BC* inexata também será inconsistente. McCarthy procura resolver o problema da inexatidão relacionando o conceito (p. ex. "João") à entidade denotada pelo conceito (a pessoa chamada João). Ele usa uma função de denotação  $denot("João")=João$ .

#### 1.4.7 Relatividade

Uma característica importante do conhecimento é que ele é *relativo* a um agente. Agentes diferentes podem ter pontos de vista inconsistentes sobre o mundo.

#### 1.4.8 Incerteza

O grau de certeza de uma afirmação pode ser dado por uma função *incert*. Assim, se *a* é mais certa do que *b* então:

$$incert(a) < incert(b)$$

Seja *P* um conjunto de proposições, *Λ* a proposição inconsistente, *V* a proposição sempre verdadeira e *R* a medida de probabilidade. Então:

$$R : P \alpha [0,1]$$

$$R(\Lambda) = 0$$

$$R(V) = 1$$

$$R(p \vee q) = P(p) + P(q) \text{ se } p \wedge q = \Lambda$$

#### 1.4.9 Imprecisão

O conteúdo de uma proposição pode ser impreciso. Por exemplo, "João nasceu em 1956", "João é muito jovem", etc. Um meio de trabalhar com imprecisão é usando conjuntos difusos (4). São conjuntos definidos por uma função  $\mu$  que varia sobre o intervalo  $[0,1]$ . Uma proposição da forma "*X* é *A*" (p. ex. "Jorge é calvo") descreve *X* como pertencente ao conjunto difuso  $S_A$ . Por exemplo, o conjunto difuso  $S_{CALVO}$  pode ter uma função *m* tal que  $\mu_{CALVO}(Jorge)=0.9$  e  $\mu_{CALVO}(Maria)=0.05$ .

### 1.5 Adequação das Bases de Conhecimento

Um sistema de representação de conhecimento deve se adequar a tarefa que ele se propõe a fazer. Bench-Capon (5) destaca três tipos de adequação para representação de conhecimento de (parte do) mundo:

**a) Adequação metafísica:** existe se não existe nenhuma contradição entre os fatos que se quer representar e a representação real deles. Assim, por exemplo, uma representação do mundo como um conjunto de partículas que não interagem seria metafisicamente inadequada pois no

mundo real se sabe que partículas interagem. Por outro lado uma representação do mundo como um conjunto de partículas que interagem pode ser metafisicamente adequada.

**b) Adequação epistemológica:** trata da habilidade de expressar conhecimento na representação. Para ser epistemologicamente adequada, uma representação deve ser capaz de ser usada para expressar os fatos que se conhece sobre o mundo. Assim, nenhuma das representações do item a) são epistemologicamente adequadas para representar fatos como "pássaros voam" ou "Tomé tem setenta anos".

**c) Adequação heurística:** é a menos compreendida. Ela ocorre se for possível expressar na representação os processos de resolução de problemas que são necessários para tratar o problema.

## 1.6 Paradigmas de Representação de Conhecimento

Vários autores, dentre eles (5) dividem os paradigmas de representação de conhecimento em quatro grandes grupos:

- a)** Regras de produção.
- b)** Redes semânticas.
- c)** Frames.
- d)** Lógica de primeira ordem.

Regras de produção tem a forma geral "Se ocorrer a situação *A* então execute a ação *B*". Um conjunto destas regras forma o mecanismo de inferência que atua sobre uma determinada base de conhecimento. A independência de cada regra de conhecimento acarreta falta de uma estrutura global para o mecanismo de inferência. Isto faz com que regras de produção não sejam consideradas como um bom paradigma para tratar problemas que se afastem muito do trivial.

Objetos estruturados como redes semânticas e frames tem sua estrutura inspirada nas estruturas de grafos e registros.

Assume-se que todos esses formalismos possam ser modelados pela lógica. A questão que retorna então é a de se saber qual formalismo é mais adequado a qual situação.



### 1.6.1 Filosofias de Representação de Conhecimento

Há duas linhas principais de pesquisa em representação de conhecimento. Uma, representada por McCarthy procura desenvolver a lógica para que esta se aproxime tanto quanto possível da capacidade humana de raciocinar e representar conhecimento. A outra linha, cujo principal defensor é Minsky, rejeita a lógica como processo de representação de conhecimento e procura imitar as estruturas da mente humana.

Minsky faz as seguintes críticas à lógica como ferramenta de representação de conhecimento:

- a) Existe dificuldade para estabelecer os axiomas adequados para um problema dado;
- b) Não se sabe a priori que axiomas são relevantes;
- c) A lógica é monotônica;
- d) A separação entre axiomas e dedução torna pouco prático incluir conhecimento para classificar as proposições e para controlar o processo dedutivo;
- e) O processo de inferência gera explosão combinatorial;
- f) Exige-se consistência.

Porém, Hayes tenta provar que os frames de Minsky são apenas uma variante notacional para a lógica de primeira ordem. Em suas conclusões afirma que uma das poucas sugestões positivas dos frames é a capacidade de raciocínio reflexivo (auto-reflexão).