

# Inteligência Artificial

---

## Representação do conhecimento(RC)

Exemplos de formalismos para RC:

- lógica de primeira ordem(LPO);
- regras;
- redes semânticas(RS);
- ontologias.

### Lógica de Primeira Ordem(LPO)

- Uma dada RC tem associado um equilíbrio que é preciso gerir:
  - mais expressividade implica normalmente maior complexidade.
- O formalismo mais poderoso que podemos usar é a LPO;
- Problemas com LPO:
  - difícil de usar;
  - pouco prático de implementar.

### Lógica proposicional(LP)

- Antes de chegarmos à LPO, temos a LP;
- A LP é um ramo da Lógica que trata o estudo das proposições;
- Uma proposição pode ser construída a partir de outras com o auxílio de **conetivas lógicas**:
  - negação;
  - conjunção;
  - disjunção;
  - implicação;
  - etc..
- Tanto as premissas como a conclusão são exemplos de proposições;
- A conclusão é obtida usando uma regra de inferência;

- 
- A LPO amplia as capacidades da LP com a inclusão de:
    - quantificadores;
    - funções;
    - predicados;
    - variáveis;
    - etc..
  - Enquanto que a LP trabalha apenas com factos, a LPO acrescenta objetos e as suas relações;

### LPO: categorias

- O conceito de **categoria** é fundamental para a representação de conhecimento. Permite agrupar vários objetos dentro de uma só entidade;
- Na LPO temos 2 formas de representar categorias: **predicados** e **objetos**;

- Podemos representar a função membro com o símbolo  $\in$ ;
- Para dizer que uma categoria é sub-classe de outra usa-se o símbolo  $\subset$ ;

## Regras

- Uma outra forma de representar conhecimento é a partir de regras do tipo **if-else**;
- As regras são tipicamente usadas nos sistemas periciais;

## Sistemas periciais(SP)

- Usados em domínios muito específicos onde existem normalmente peritos humanos;
- As regras usadas são do tipo: **if x then y**;
- O **x** é a **condição** e o **y** o **consequente**;
- Uma regra dispara se a condição for verdadeira;
- Estão divididos em dois sub-sistemas:
  - a base de conhecimento(BC);
  - o motor de inferência.
- A BC guarda factos e regras;
- o SP pode fazer **dedução**: cada consequente é um novo facto;
- Ou fazer **reação**: cada consequente é uma ação;
- **Vantagens**:
  - o formato das regras é intuitivo e permite explicar as conclusões obtidas;
  - fácil de manter: não é necessário escrever código para alterar as regras;
  - fácil e rápido fazer protótipos.
- **Desvantagens**:
  - Obtenção de conhecimento: o tempo dos peritos é valioso;
  - Desempenho: corriam originalmente em sistemas interpretados(lisp).
- Os Sps devem ser usados quando:
  - faça sentido economicamente;
  - peritos humanos não estejam sempre disponíveis;
  - o problema requiera raciocínio simbólico.

## Redes Semânticas(RS)

- Uma RS é um **grafo** que representa relações semânticas entre conceitos;
- Existem muitas variantes de RS mas todas permitem representa **objetos**, **categorias** de objetos e as **relações** entre eles;
- As RS tipicamente não têm todo o poder expressivo da LPO. Mas têm a vantagem de serem simples e o processo de inferência transparente;
- As **relações** são binárias apenas, pois estão codificadas em arestas;
- Alguns problemas com as RS:
  - são pesadas do ponto de vista computacional pois para responder a questões é necessário fazer a travessia de rede;
  - falta alguma capacidade de expressão:
    - quantificadores;
    - negação;
    - entre outras.

## Ontologias

- Uma **ontologia** é uma representação de entidades e das suas relações;
- As ontologias são escritas usando linguagens próprias para a descrição de ontologias;
- Uma ontologia pode ser **visualizada** num grafo;

### Componentes duma ontologia

- Indivíduos: instâncias ou objetos;
- Classes: conjunto, coleções conceitos, tuplos de objetos ou de coisas;
- Atributos: aspetos, características, propriedades ou parâmetros de objetos ou classes;
- Relações: formas segundo as quais indivíduos e classes se relacionam;
- Restrições: Descrições formais que devem ser verdade para que uma dada afirmação seja aceite;
- Regras: afirmações com a forma **if-then** que descrevem uma inferência lógica;
- Axiomas: afirmações (inclui regras) que contêm toda a teoria descrita pela ontologia;
- Eventos: alterações de atributos ou relações.

### Ontologia vs Rede Semântica

- A RS é uma **notação gráfica** usada para representar conhecimento com os nodos e arestas de um grafo;
- Uma **ontologia** é a representação de conceitos dentro de um domínio e das suas relações, de forma explícita e formal, que **pode ser visualizada como um grafo**, mas que existe sem qualquer relação com um grafo.

### Semantic Web(SW)

- Mais recentemente, a SW surge integrando RC e raciocínio recorrendo a linguagens baseadas em XML;
- A SW é uma **extensão da web** que possibilita a partilha de dados entre aplicações, empresas e comunidades;
- Isto é conseguido **adicionando meta-dados** às páginas *web*, em formatos que são legíveis por máquinas;
- Isto permite que agentes/ motores de pesquisa, consigam um acesso aos dados mais fácil e mais rico, resultando num maior número de tarefas realizáveis;
- A *Resource Description Framework*(RDF) permite representar conhecimento;
- A *Web Ontology Language*(OWL) adiciona semântica e permite o uso de sistemas de raciocínio automáticos, como os classificadores;
- A ideia é adicionar uma camada de significado(semântica) sobre a *internet*;
- A abordagem básica usa palavras contidas nas páginas para construir índices para depois permitir a pesquisa com os motores de busca;
- Com a SW, são criadas **ontologias** de conceitos.

### Base de Conhecimento(BC) vs Base de Dados(BD)

- Uma base de dados é muito boa para guardar informação **quando sabemos quais as características**(atributos) que queremos guardar relativas aos dados;
- Em IA, **não sabemos** muitas vezes todas as possíveis características que irão ser medidas: conforme o agente vai recolhendo conhecimento do mundo este tem que ser guardado.

## Tipos de Raciocínio

- **Inferência:** chamamos inferência ao processo de derivação de novo conhecimento a partir de conhecimento já existente. Na IA o componente do sistema que faz inferência é o **motor de inferência**;
- **Dedutivo:** nova informação é deduzida a partir de informação com relação lógica;
- **Indutivo:** partir de um conjunto de observações e generalizar;
- **Abdutivo:** é uma forma de dedução que permite inferência plausível;
- **Analogia:** fazer analogias entre duas situações;
- **Senso-comum:** raciocínio informal que usa regras aprendidas pela experiência(heurísticas);
- **Não-monótono:** usado quando os factos podem mudar;

## Incerteza

### Teoria da decisão

- Para que o agente possa tomar a sua decisão deve então levar em conta a **probabilidade** de uma dada ação o levar a um estado que deseja e a **utilidade** o estado que resulta dessa ação:
  - \$Teoria da decisão = Teoria das probabilidades + Teoria da Utilidade\$
- **Um agente é racional** se e só se escolhe a ação que tem maior **utilidade esperada**:
  - $UE(a) = \sum_{e \in E(a)} P(e)U(e)$  onde \$e\$ é um estado, \$U(e)\$ é a utilidade do estado \$e\$, \$E(a)\$ é o conjunto dos estados que podem resultar da ação \$a\$ e \$P(e)\$ é a probabilidade da ação \$a\$ resultar no estado \$e\$;
- Note-se que a **utilidade** é característica de um estado mas a **utilidade esperada** é característica duma ação;

### Incerteza e decisões racionais

- Atribuímos probabilidades e utilidade ao estado resultante das ações;
- A escolha final é uma ação cujo estado resultante não tem a maior utilidade, mas a ação tem a maior utilidade esperada.

### Conceitos básicos de probabilidade

- Consideremos uma experiência com espaço amostral \$S\$. Para cada acontecimento \$A\$ assumimos que se define um número chamado a **probabilidade do acontecimento** \$A\$, \$P(A)\$, que obedece às seguintes condições:

1.  $0 \leq P(A) \leq 1$ ;
2.  $P(S) = 1$ ;
3. Para qualquer sequência de eventos \$A\_1, A\_2, \dots\$, que sejam **mutuamente exclusivos** (\$A\_i \cap A\_j = \emptyset\$ quando \$i \neq j\$) temos

$$P(\cup_{n=1}^{\infty} A_n) = \sum_{n=1}^{\infty} P(A_n)$$

- Para achar a **probabilidade da união de dois acontecimentos** \$P(A \cup B)\$ em geral, mesmo quando não forem mutuamente exclusivos:

- Consideramos  $P(A) + P(B)$  que é a soma da probabilidade de todos os resultados em  $A$  com a probabilidade de todos os resultados em  $B$ ;
- Resultados que estejam em ambos os acontecimentos serão contados duas vezes, mas só aparecem uma vez em  $P(A \cup B)$ , logo temos  $P(A \cup B) = P(A) + P(B) - P(A \cap B)$
- **Probabilidade condicional:** queremos saber qual a probabilidade de um acontecimento  $A$  sabendo que outro,  $B$ , já aconteceu:
 
$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$
 para  $P(B) > 0$
- O nosso agente vai estar interessado em usar probabilidades condicionais pois quer ir juntando toda a informação que vai recolhendo sobre o ambiente para poder tomar decisões.

## Inferência

- É possível **chegar a conclusões** tirando partido da teoria das probabilidades;
- A **base de conhecimento** será a distribuição conjunta dos acontecimentos que estão envolvidos no mundo que estamos a considerar;
- A soma das probabilidades tem que dar 1;
- Se se quiser apenas a probabilidade relativa a uma única variável, soma-se os valores de todos os eventos em que ela está envolvida e obtem-se a **probabilidade marginal**:  $P(A) = \sum_{B \in Z} P(A, B)$  onde  $Z$  representa o conjunto de variáveis envolvidas no problema;
- Temos uma relação semelhante para o caso em que as probabilidades são condicionais em vez de conjuntas (aplicar a penúltima equação a última equação):  $P(A) = \sum_{B \in Z} P(A|B)P(B)$
- Achar a probabilidade condicional de alguma variável dada informação sobre outras:
  - Usa-se  $P(A|B) = \frac{P(A \cap B)}{P(B)}$  e depois avalia-se a partir da distribuição conjunta;
 
$$P(A|B) = \alpha P(A \cap B)$$
 com  $\alpha = \frac{1}{P(B)}$
  - Só tem que se garantir que  $\alpha$  seja tal que a soma destas duas probabilidades dê 1;
  - Isto permite achar os valores destas probabilidades mesmo desconhecendo o valor de  $P(B)$ ;
  - Pode-se então escrever:  $P(X|Y) = \alpha P(X, Y)$

## Regra de Bayes(RB)

$$P(A, B) = P(A|B)P(B) \quad P(B, A) = P(B|A)P(A)$$

Como  $P(A, B) = P(B, A)$  vem  $P(A|B)P(B) = P(B|A)P(A)$

Isto permite escrever:  $P(A|B) = \frac{P(B|A)P(A)}{P(B)}$

- Pode-se usar a RB para descobrir a **probabilidade de uma dada causa estar por trás de um efeito** observado com:  $P(B|A) = \frac{P(A|B)P(A)}{P(B)}$
- Se não sabemos a causa entre várias possíveis, podemos achar as probabilidades condicionais acima para as várias causas e ficamos a saber qual é a mais provável geradora do efeito observado;
- A probabilidade  $P(B|A)$  é na **direção da causa**: tem-se uma causa( $A$ ) e procura-se um efeito( $B$ );
- A probabilidade  $P(A|B)$  é na **direção da diagnóstico**: tem-se um efeito( $B$ ) e procura-se uma causa( $A$ );

## Independência condicional

- A definição para **independência condicional** de duas variáveis  $X$  e  $Y$  dada a variável  $Z$  é  $P(X,Y|Z) = P(X|Z)P(Y|Z)$
- Vantagens:
  - Isto permite que os sistemas de inferência que lidam com  $n$  variáveis escalem com  $O(n)$  em vez de  $O(2^n)$ , o que é muito importante;
  - É mais fácil ter informação relativa a independência condicional que a independência absoluta.