



Universidade do Minho
Escola de Engenharia
Departamento de Informática

Incerteza e Imprecisão no Raciocínio

Human-like decision making

LICENCIATURA EM ENGENHARIA INFORMÁTICA
MESTRADO integrado EM ENGENHARIA INFORMÁTICA

Inteligência Artificial
2022/23

“Everything is vague to a degree you do not realize till you have tried to make it precise.”

Bertrand Russell

(English philosopher and mathematician (1872-1970))

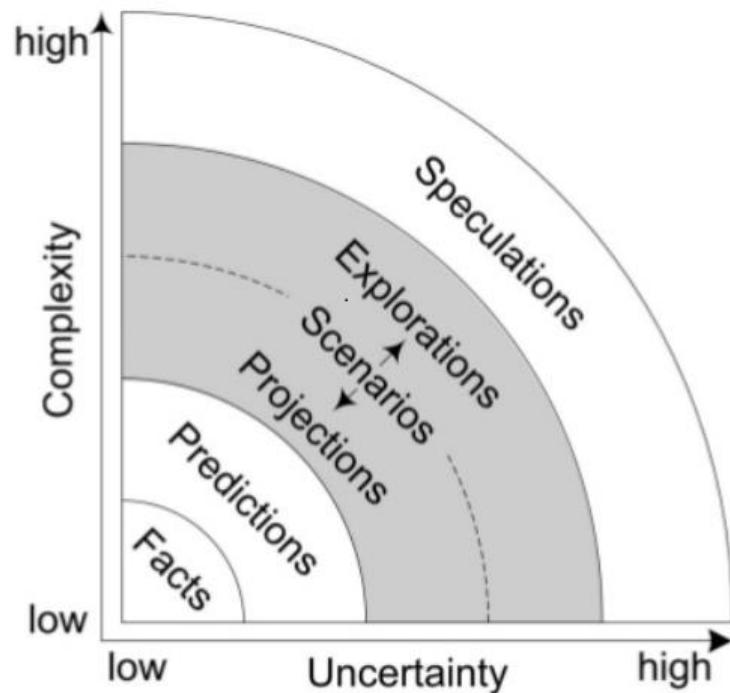
Uma entidade que conheça todo o conhecimento (os factos e regras) acerca do ambiente que o rodeia, usando uma qualquer aproximação (lógica) tem a capacidade de gerar (derivar) planos (cujo sucesso é garantido).

Mas:

- Muito raramente uma qualquer entidade têm acesso a todos os factos (verdade) acerca do seu ambiente;
- O que implica, que são obrigados a atuar constantemente numa base de incerteza/Imprecisão;

- Nota: Entidade = Pessoa ou mesmo Entidade = Agente

Making Decision



- Dados incertos;
- Conhecimento incerto (não confiável);
- Representação incerta do conhecimento;
- Processo de Inferência.

- Incerteza:
 - O predicado está bem definido, mas não se conhece o seu valor verdadeiro.
 - os dados laboratoriais de um sintoma podem atrasar-se;
 - a medicina pode ter uma teoria incompleta para algumas doenças.
 - Raciocínio Estatístico - Redes Bayesianas - que são utilizadas para tratar da incerteza usando probabilidades.
- Imprecisão:
 - O predicado é vago por si.
 - a hora em que um evento ocorreu pode ser conhecida apenas aproximadamente.
 - Conjuntos Vagos (Fuzzy) - tratam da imprecisão (possibilidade).
- Mistos
 - Em alguns domínios coexistem estes dois tipos de incerteza: a Imprecisão e a Probabilidade.

Como podemos capturar o facto de não ter certeza de que:
O Paulo tem febre?

Algumas possíveis respostas....

Lógica

Não sei se o Paulo tem febre ou não.
febre e \neg febre.

Probabilidades

A probabilidade de o paciente Paulo ter febre é 0,6.

Podemos saber isso a partir de dados estatísticos ou através de algum conhecimento geral sobre o assunto.

Métodos Qualitativos

- Hipótese de mundo fechado
- Circunscrição
- Lógica e raciocínio *Default*
- *Truth maintenance system*
- Modelos Conexionistas
- ...

Destacam-se em ambientes de inexistência de conhecimento.

Métodos Quantitativos

- Teoria das Probabilidades
- Fatores de Certeza
- Redes *Bayesianas*
- *Fuzzy Sets and Fuzzy Logic*

Destacam-se em ambientes de incerteza/imprecisão de conhecimento

- Qualquer fonte de Conhecimento pode ter associado incertezas e imprecisões;
- É necessário aumentar as capacidades de Representação de Conhecimento com medidas refletindo os graus de evidência e crenças;
- Duas abordagens diferentes:
 - Teoria das Probabilidades (aproximação *Frequency based Approach*, resulta de cálculos sobre experiências);
 - Teorias Probabilísticas Subjetivistas (resulta das crenças acumuladas).

- Teoria das Probabilidades (aproximação por Frequência).
- Método quantitativo de tratamento do conhecimento imperfeito, que representa a incerteza por um valor numérico simples e usa o teorema de *Bayes* como método de inferência.
- Representa uma Teoria estatística da evidência baseada na noção de probabilidades condicionais

Teorema de Bayes:

$$P(H_i|E) = \frac{P(E|H_i) * P(H_i)}{\sum_{n=1}^k P(E|H_n) * P(H_n)}$$

Simplificando:

Reflete a confiança no acontecimento H (hipótese) dado que se observaram uma ou mais evidências (Ei).

- O Teorema (ou Regra ou Lei) de Bayes relaciona a probabilidade condicional de um evento com a ocorrência conjunta de vários eventos:
 $p(H \cap E) = p(H|E) * p(E)$
 $p(E \cap H) = p(E|H) * p(H)$
- de onde se deriva a fórmula de Bayes relacionando probabilidades condicionais e que na sua forma mais simples se pode escrever:
 $p(H|E) = (p(E|H) * p(H)) / p(E)$
- O teorema de Bayes é uma teoria estatística da evidência baseada na noção das probabilidades condicionais

- Probabilidade incondicional

- A probabilidade $P(a)$ de um evento a é um número dentro do intervalo $[0,1]$.
 - $P(a) = 1$ se a não tem associado a si incerteza.
 - Se a e b são mutuamente exclusivos, então: $P(a \vee b) = P(a) + P(b)$.

- Probabilidade condicional

- Probabilidade condicional $P(a|b) = x$, é interpretada como:
“Dado um evento b , probabilidade do evento a é x ”.
- $P(b|a) = P(a|b)P(b)/P(a)$ - Regra de Bayes.

- Tabela de Conjunção de probabilidades

- $P(A|B) = P(A,B)/P(B)$
- Tabela $n \times m$, representada pela probabilidade de cada configuração (a_i, b_j)
- Representam todo o domínio
- Para valores booleanos teríamos 2^n entradas

- Para casos simples é possível ter tabelas de probabilidades conjuntas.
- Exemplo:

	dor	\sim dor
cavidade	0.04	0.06
\sim cavidade	0.01	0.89

A soma das probabilidades deve ser = 1

$$P(c)=0.1 \quad p(c \vee d)=0.11 \quad P(d)=0.05 \quad p(c \cap d)=0.04$$

aplicando a **fórmula de Bayes**: $p(H \cap E) = p(H|E) * p(E)$

$$P(c|d)=p(c \cap d)/p(d)=0.04/0.05=0.8$$

*Naive Bayes simplifica de $O(2^n)$
Para $O(n)$*

- Mas com muitas variáveis (n sintomas) a representação torna-se impraticável $O(2^n)$. Não se usam as probabilidades conjuntas mas sim diretamente as condicionais

Exemplo Probabilidades

“A Anabela parece ter um aperto no peito e dificuldades respiratórias. Em que medida sofre de uma bronquite ligeira? ou de um enfarte do miocárdio?”

	dor_peito	¬dor_peito
enfarte_miocárdio	0,3	0,1
¬enfarte_miocárdio	0,2	0,4

$$P(\text{dor_peito}) = 0.3 + 0.2 = 0.5$$

$$P(\text{enfarte_miocárdio} \wedge \text{dor_peito}) = 0.3$$

$$P(\text{enfarte_miocárdio}) = 0.3 + 0.1 = 0.4$$

Fonte: Eduardo Fermé, Raciocínio Incerto e Impreciso, Universidade da Madeira, 2019.

Exemplo Redes de Bayes

- E1 = 'aperto de peito' E2 = 'dificuldades respiratórias '
- H1 = 'bronquite ' H2 = 'enfarre ' H3 = 'nem bronquite nem enfarre '
- H4 = 'bronquite e enfarre '

Quantificar subjetivamente as variáveis:

- $P(H1)=0.05$ $P(E1 | H1)=0.9$ $P(E2 | H1)=0.9$
- $P(H2)=0.01$ $P(E1 | H2)=0.8$ $P(E2 | H2)=0.5$
- $P(H3)=0.93$ $P(E1 | H3)=0.1$ $P(E2 | H3)=0.1$
- $P(H4)=0.01$ $P(E1 | H4)=0.9$ $P(E2 | H4)=0.9$

Exemplo:

- Usando a regra de *Bayes* e os valores anteriores:
- Bronquite dado que apresenta aperto no peito e dificuldades respiratórias:

$$P(H_1 | E_1, E_2) = \frac{0.0405}{0.0619} = 0.65$$

- Enfarte dado que apresenta aperto no peito e dificuldades respiratórias:

$$P(H_2 | E_1, E_2) = \frac{0.004}{0.0619} = 0.06$$

- Nem bronquite, nem enfarte dado que apresenta aperto no peito e dificuldades respiratórias:

$$P(H_3 | E_1, E_2) = \frac{0.0093}{0.0619} = 0.16$$

Pode-se concluir que um paciente com um aperto no peito e dificuldades respiratórias terá maiores probabilidades (65%) de sofrer de bronquite.

Caso:

Bronquite dado que apresenta aperto no peito e dificuldades respiratórias

$$P(H_1 | E_1, E_2) = \frac{0.0405}{0.0619} = 0.65$$

$$\begin{aligned} P(H_1 | E_1, E_2) &= P(E_1 | H_1) \times P(E_2 | H_1) \times P(H_1) / \\ &\quad (P(E_1 | H_1) \times P(E_2 | H_1) \times P(H_1) + \\ &\quad P(E_1 | H_2) \times P(E_2 | H_2) \times P(H_2) + \\ &\quad P(E_1 | H_3) \times P(E_2 | H_3) \times P(H_3) + \\ &\quad P(E_1 | H_4) \times P(E_2 | H_4) \times P(H_4)) \\ &= 0,9 \times 0,9 \times 0,005 (0,0405) / \\ &\quad 0,0405 + 0,004 + 0,0093 + 0,0081 (0,0619) \\ &= 0,0405 / 0,0619 = 0,65 \end{aligned}$$

Fonte: Eduardo Fermé, Raciocínio Incerto e Impreciso, Universidade da Madeira, 2019.

Vantagens vs Desvantagens

Desvantagens

- **Se o número de variáveis aumenta, os cálculos aqui aplicados tornam-se complicados;**
- **Fica difícil compreender plenamente as relações entre as variáveis do ambiente.**
- **A aquisição de conhecimento é difícil. Os seres humanos são considerados fracos avaliadores.**

Vantagens

- **Permite criar modelos alternativos baseados neste método, tais como as redes bayesianas.**

- a) Modelo dos Fatores de Certeza
- b) Rede de Bayes
- c) Teoria de Dempster-Shaffer

O Modelo dos Fatores de Certeza foi introduzido por Buchanan, inicialmente para o Sistema Pericial MYCIN*:

- Todas as regras da BC são ou puramente causais ou puramente de Diagnóstico;
- As evidências são só introduzidas nas raízes das regras;
- As regras são quase sempre simplesmente ligadas.

Regras que tem Fatores de Certeza associados

$$FC[h|e] = MB[h,e] - MD[h,e]$$

FC: Certainty Factor; MB: Measure of Belief; MD: Measure of Disbelief

Os Fatores de Certeza podem ser combinados uma vez que:

- i) várias regras suportam separadamente uma mesma hipótese;
- ii) várias evidências conjuntas suportam uma hipótese.
- iii) o resultado de uma regra é (pode ser) a entrada em outra.

*Buchanan, B.G.; Shortliffe, E.H. (1984). Rule Based Expert Systems: The MYCIN Experiments of the Stanford Heuristic Programming Project. Reading, MA: Addison-Wesley. ISBN 978-0-201-10172-0.

Modelo dos Factores de Certeza

- É um mecanismo “informal” para quantificar o grau de crença numa certa conclusão baseado na presença de um conjunto de evidências.
- Factor de Certeza é um valor numérico que exprime quanto devemos acreditar numa conclusão dado um conjunto de evidências;

$$1 \geq FC \geq 0$$

$$0 \leq MB \leq 1$$

Measure of Disbelief $MD[H, E]$

Para a mesma evidência :

$$MB > 0 \rightarrow MD = 0$$

$$MD > 0 \rightarrow MB = 0$$

either $MB=0$ or $MD=0$

- FC é uma quantificação subjetiva;
- Ex., o diagnóstico de falhar a net.

Essa certeza que depositamos numa regra é representada da seguinte forma:

Se <evidência> então <hipótese> [$FC = k$]

Isto quer dizer que, dada a evidência, o nosso grau de crença na hipótese é dado pelo valor de FC.

Exemplo:
Incerteza em Regras de Produção

- Nos sistemas previamente apresentados a informação é do tipo verdade ou falso;
- Não são considerados valores intermédios (ex., falso, pouco provável, provável, altamente provável, verdadeiro);
- No mundo real isto é irrealista;
- Os sistemas têm de saber lidar com a incerteza (ex., o grau de risco, a probabilidade, a confiança);

- Associamos a cada proposição um grau de confiança.

Factos: facto(Proposição) :: C.

Regras: se Condição então Ação :: C

sendo C um número entre 0 e 1. (0-probabilidade 0%; 1:probabilidade 100%)

facto(enjoos)::1.

facto(vomitos)::0.5.

(se enjoos e vomitos

entao 'dor de cabeca') :: 0.75.

(se vomitos e 'dor de estomago'

entao bebedeira) :: 0.50.

(se 'dor de cabeca' ou bebedeira

entao 'problemas intestinais') :: 0.35.

(se enjoos e 'dores lombares' e 'dor de rins'

entao reumatismo) :: 0.20.

| ?- demo('dor de cabeca',C).
C = 0.375 ?

- A descrição do mundo real não implica, necessariamente, o uso de uma enorme tabela de probabilidades para se conhecer todas as combinações dos eventos possíveis;
- Um método para inferir conhecimento com incerteza;
- Uma notação gráfica simples para declaração de independência condicional e, portanto, para a especificação compacta de distribuições.

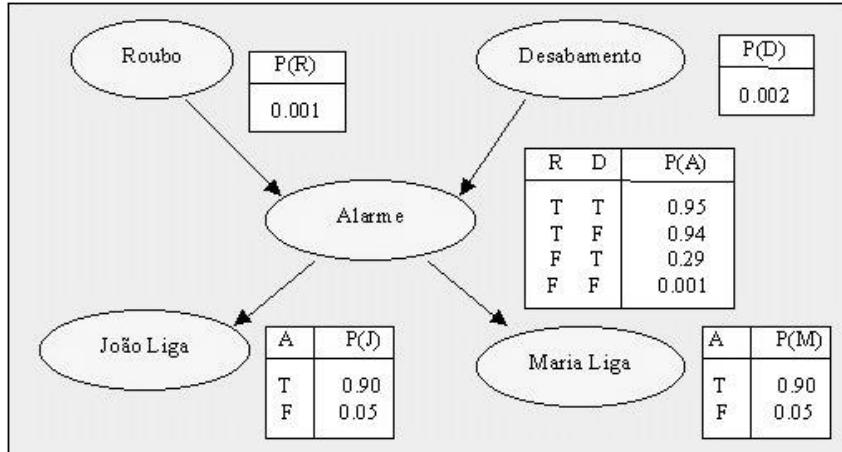
Sintaxe:

- um conjunto de nós, um por variável
- um gráfico acíclico direcionado
- uma distribuição condicional para cada nó, considerando seus pais:
 $P(X_i | \text{Pais}(X_i))$

Estou no trabalho. O meu vizinho João liga-me para dizer que meu alarme está a tocar, mas minha vizinha Maria não ligou. Às vezes, o alarme é acionado por pequenos desabamentos . Pode ser há um assalto?

- Variáveis: Roubo, Desabamento, Alarme, JoaoLiga, MariaLiga
- A topologia de rede reflete nosso conhecimento "causal":
 - um ladrão pode acionar o alarme;
 - um desabamento pode acionar o alarme;
 - o alarme pode fazer com que a Maria ligue;
 - o alarme pode fazer o João ligar.

Exemplo: AIMA



Qual a probabilidade de que o alarme toque, a Maria e o João liguem, mas não ocorra nem um roubo nem um desabamento.

$$\begin{aligned}
 & P(\text{não roubo e não Desabamento e Alarme e João e Maria}) \\
 & = P(\sim R).P(\sim D).P(A|\sim R, \sim D).P(J|A).P(M|A) \\
 & = (0.999)(0.998)(0.001) (0.9)(0.9)=0.00062
 \end{aligned}$$

- As redes *Bayesianas* são bastantes flexíveis, permitindo o cálculo de probabilidade de qualquer nó.
- As inferências podem ser de vários tipos:
 - **Inferências de diagnóstico** – do efeito para a causa. Dado que o João ligou qual a probabilidade de ocorrer roubo?
 - **Inferências causais** – da causa para o efeito. Dado que houve roubo qual a probabilidade do João ligar?
 - **Inter causais** – entre causa comuns para o mesmo efeito. Qual a probabilidade de roubo dado o alarme e desabamento.

- Teoria de Dempster-Shafer é uma técnica de modelagem da incerteza, que permite a distinção entre ignorância e incerteza, pois como são conceitos diferentes devem ser tratados de maneira diferente;
- Esta teoria baseia-se na atribuição de probabilidades a todos os subconjuntos de um universo de discurso e não apenas aos membros individuais, como na teoria clássica da probabilidade. Suportando ainda o uso de probabilidades subjetivas no raciocínio, que muitas vezes refletem melhor o raciocínio de um especialista humano.
- Nesta teoria, não é calcular a probabilidade de uma proposição, mas sim a probabilidade de que a evidência apoia a proposição.

Esta é uma medida de crença denotada por Bel (X).

- Usa um par ([Crença, plausibilidade]) atribuído a cada conjunto de proposições:

Crença (S) (ou B) de 0 (sem evidência) a 1 (certeza).

Plausibilidade (S) = $1 - \text{Bel}(\neg S)$ entre 0 e 1.

Bel (S) mostra o grau de confiança em uma proposição S

Plausibilidade ($1 - \text{Bel}(\neg S)$) indica o valor limite máximo por esse valor de confiança.

- Em vez de calcular a crença ou a probabilidade de uma proposição S;
- Calculamos a probabilidade da evidência existente para apoiar a proposição S;
- Este processo pode ajudar a decidir se devemos ou não esperar por novas evidências antes de tomar a decisão.

Ex., atirando uma moeda ao : podemos usar $p(\text{cara}) = 0,5$ e $p(\text{coroa}) = 0,5$

No entanto, se não tivermos certeza se a moeda foi alterada

Ex.,

$p(\text{alter}) = 0,1$, ENTÃO $p(\text{cara})$ e $p(\text{coroa})$ variam entre 0,45 e 0,55

Fuzzy Sets and Fuzzy Logic

- A teoria dos conjuntos difusos (Fuzzy Sets) é uma maneira de especificar quão bem um objeto satisfaz uma descrição “vaga”.
Ex., O Paulo é alto.
- Lógica difusa (Fuzzy Logic) é um método de raciocínio com expressões lógicas que descrevem a associação em conjuntos difusos.
- Os sistemas de controle difuso aplicaram a lógica difusa em muitos produtos comerciais (ex., câmaras de vídeo e outros aparelhos domésticos).

Se afirmo que “está calor”:

- Poderia induzir a convicção a alguém concluir que está 40 graus;
- No entanto, noutro a convicção poderia estar cerca de 26 graus;
- Enquanto que num terceiro, seria de que está 22 graus.

Definição de difuso (fuzzy)

- Difuso - “não claro, distinto ou preciso”.

Definição de *Fuzzy logic*

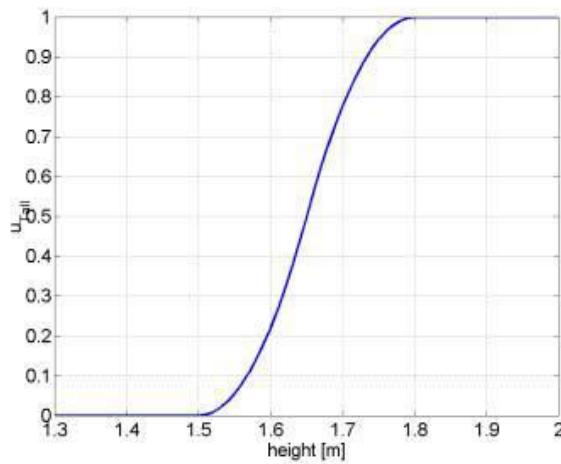
- Forma de representação do conhecimento adequada para noções que não podem ser definidas com precisão, mas que dependem de seus contextos.

Conjuntos e conceitos difusos (fuzzy) são vulgarmente usados em linguagem natural:

- Miguel é alto
- Paulo é esperto
- Alexandre está feliz
- A sala está quente

- Ex., o conjunto de alto pode ser definido como:
$$\{x \mid \text{altura } x > 1,8 \text{ m}\}$$
- Mas como podemos considerar uma pessoa com uma altura = 1,79 m?
- E com 1,78 m?
- ...
- e 1,52m?

- Num conjunto difuso, uma pessoa com 1,8 metros de altura seria considerada alta em alto grau;
- Uma pessoa com uma altura de 1,7 metros seria considerada em menor grau, etc.
- A função no entanto pode mudar para jogadores de basquetebol, mulheres, crianças etc.



- Um subconjunto (Difuso/fuzzy) F de um conjunto C pode ser definido como um conjunto de pares ordenados X, Y , em que X é um membro de C e Y pertence ao intervalo $[0,1]$.
 - O valor 0 diz que X definitivamente não pertence a F e 1 o oposto.
 - Os valores intermediários representam um grau de associação de X a F .
- O Conjunto C é o corresponde ao universo do discurso do conjunto F (Difuso/fuzzy);
- O mapeamento aqui é chamado Função de Membros em relação ao conjunto F (Difuso/fuzzy);
- A proposição 'a pertence a F ' tem um valor de verdade dado por Y no par X, Y , de modo que: $X = a$ e Y estão no intervalo $[0,1]$.

em Lógica

```
bool speed;  
get the speed  
if ( speed == 0 ) {  
// speed is slow  
}  
else {  
// speed is fast  
}
```

usando *fuzzy logic*

```
float speed;  
get the speed  
if ((speed >= 0.0)&&(speed < 0.25)) {  
// speed is slowest  
}  
else if ((speed >= 0.25)&&(speed < 0.5))  
{  
// speed is slow  
}  
else if ((speed >= 0.5)&&(speed < 0.75))  
{  
// speed is fast  
}  
else // speed >= 0.75 && speed < 1.0  
{  
// speed is fastest  
}
```

- A imperfeição, imprecisão ou incerteza, atravessa todos os cenários do mundo real;
- Naturalmente tem de ser incorporada a todos os sistemas (de informação) que tentam apresentar um modelo completo e preciso do mundo real;
- A uma dificuldade inerente ao entendimento dos vários aspectos da imprecisão e incerteza;
- Apresenta-se uma classificação em que imprecisão e incerteza são as principais classes;
- Apresenta-se resumidamente várias abordagens para modelar imprecisão e incerteza.
- Agrupados em duas grandes categorias: os modelos qualitativos (simbólicos) e os modelos quantitativos.

- Russell and Norvig, Artificial Intelligence - A Modern Approach, 3rd edition, ISBN-13: 9780136042594, 2009, section IV Uncertain knowledge and reasoning.
- P. Krause and D. Clark. Representing Uncertain Knowledge. An AI Approach. Intellect, Oxford, 1993.
- Smets P., Imperfect Information: Imprecision and Uncertainty. In: Motro A., Smets P. (eds) Uncertainty Management in Information Systems. Springer, Boston, MA, 1997.



Universidade do Minho
Escola de Engenharia
Departamento de Informática

Incerteza e Imprecisão no Raciocínio

Human-like decision making

LICENCIATURA EM ENGENHARIA INFORMÁTICA
MESTRADO integrado EM ENGENHARIA INFORMÁTICA

Inteligência Artificial
2022/23