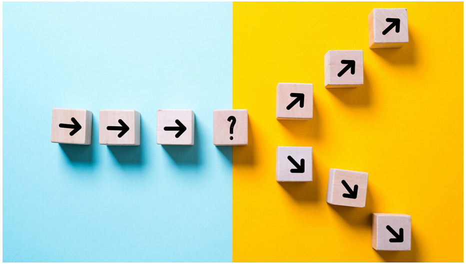
Parte III: IA e Incerteza.

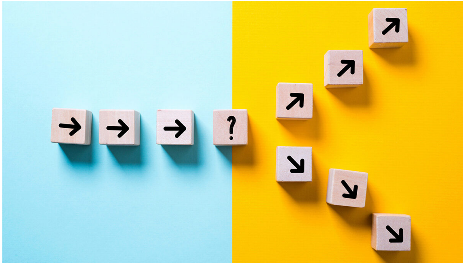
Prof. Fabiano Araujo Soares, Dr. / FGA 0221 - Inteligência Artificial

Universidade de Brasília

2025



Como um agente pode tomar decisões corretas mesmo quando não conhece todos os fatos?



Como um agente pode tomar decisões corretas mesmo quando não conhece todos os fatos?

Inteligência precisa lidar com decisões sob informações imperfeitas: ignorar a incerteza é ignorar o mundo real.

Motivação: Por Que Lidar com Incerteza?

- Agentes no mundo real precisam lidar com a incerteza, seja devido à observabilidade parcial, não determinismo, limitações ou adversidades.
- Até o momento, vimos agentes lógicos lidarem com a incerteza acompanhando todos os estados de mundo possíveis em que possam estar.
- Essa abordagem funciona em problemas simples, mas possui desvantagens:
  - O agente deve considerar todas as explicações possíveis (mesmo as mais improváveis);
  - Plano de contingencia que lide com todas as eventualidades pode ser computacionalmente inviável;
  - Às vezes, não há garanta de alcance da meta, mas o agente deve agir mesmo assim.
- Logo o agente precisa de alguma estratégia para lidar com planos que não são garantidos.

Exemplo de Agente lidando com Incerteza.

Imagine que um agente de Smart Taxi precisa atender um cliente que terá um voo as 21h e por isso precisa chegar ao aeroporto às 20h, no entanto o agente sabe dos seguintes fatos:

- O tempo médio para chegar ao aeroporto é de 30 min em transito médio, 45 min em transito pesado e 20 min em transito leve;
- Entre as 19h e às 21h o transito é médio, entre as 17h e às 19h o transito é pesado;
- Quando chove o transito é sempre pesado;
- Existe 50% de chances de chuva no dia do voo;
- Existe uma chance de 5% de haver acidente na rota, e caso isso ocorra, o tempo para chegada ao aeroporto aumenta em 10 min por acidente, independente do tráfego.
- Existem uma série de outras condições que podem ocorrer mas não são previsíveis.

Exemplo de Agente lidando com Incerteza.

Imagine que um agente de Smart Taxi precisa atender um cliente que terá um voo as 21h e por isso precisa chegar ao aeroporto às 20h, no entanto o agente sabe dos seguintes fatos:

- O tempo médio para chegar ao aeroporto é de 30 min em transito médio, 45 min em transito pesado e 20 min em transito leve;
- Entre as 19h e às 21h o transito é médio, entre as 17h e às 19h o transito é pesado;
- Quando chove o transito é sempre pesado;
- Existe 50% de chances de chuva no dia do voo;
- Existe uma chance de 5% de haver acidente na rota, e caso isso ocorra, o tempo para chegada ao aeroporto aumenta em 10 min por acidente, independente do tráfego.
- Existem uma série de outras condições que podem ocorrer mas não são previsíveis.

Qual é o melhor plano para chegar com segurança no aeroporto?

Incerteza e lógica proposicional

- Vamos considerar um exemplo de raciocínio incerto: diagnosticar a dor de dente de um paciente odontológico;
- O diagnóstico - seja para medicina, conserto de automóveis ou qualquer outra coisa - quase sempre envolve incerteza.
- Vamos tentar escrever regras para diagnóstico odontológico usando lógica proposicional:
  - Dor de dente $\implies$ Cárie (Errado! Nem toda dor de dente está relacionado à cáries);
  - Dor de dente $\implies$ Cárie $\vee$ Gengivite $\vee$ Abcesso $\vee$ ... (lista grande);
  - Cárie $\implies$ Dor de dente (Errado! Nem toda cárie está relacionada a dor);
- A única maneira de corrigir a regra é torná-lo logicamente exaustiva: aumentar o lado esquerdo com todas as qualificações necessárias para que uma cárie cause dor de dente.

Como lidar com incerteza?

- Tentar usar a lógica para lidar com um domínio como o diagnóstico médico não é prático por três razões:
  - **Complexidade**: É trabalhoso listar o conjunto completo de antecedentes ou consequentes necessário para garantir uma regra sem exceções e é muito difícil usar essas regras.
  - **Ignorância teórica**: A ciência médica não tem uma teoria completa para o seu domínio.
  - **Ignorância prática**: Mesmo que conheçamos todas as regras, podemos estar incertos sobre um determinado paciente porque nem todos os testes necessários foram ou podem ser executados.
- A conexão entre dores de dente e cáries não é uma consequência lógica estrita em ambas as direções.
- Isso é típico do domínio médico, bem como em vários outros domínios de julgamento: direito, negócios, design, reparo de automóveis, jardinagem, namoro e assim por diante.

O que fazer?

- Quando o conhecimento do agente não é certo $\Longrightarrow$ **Probabilidade**.
- **compromissos ontológicos**: Iguais na lógica e na teoria das probabilidades — o mundo é composto de fatos que são válidos ou não em qualquer caso particular;
- **compromissos epistemológicos**: são diferentes — um agente lógico acredita que cada sentença é verdadeira ou falsa, um agente probabilístico possui um grau numérico de crença entre 0 (para sentenças que certamente são falsas) e 1 (certamente verdadeira).

Teoria da Probabilidade

- **A teoria da probabilidade**: Lida com a incerteza que vem de nossa falta de tempo, ignorância ou limitação.
- Podemos não ter certeza sobre o que aflige um paciente em particular, mas acreditamos que há, digamos, 80% de chance — isto é, uma probabilidade de 0,8 - que o paciente que reclama de dor de dente tenha uma cárie.

Utilidade

- **Retornando ao exemplo do smart taxi**: existem várias soluções possíveis para o cliente, no entanto, ele certamente possui preferências;
- Usaremos a **teoria da utilidade** para representar preferências e raciocinar quantitativamente a respeito dos possíveis resultados;
- A teoria da utilidade diz que todo estado (ou sequência de estados) tem um **grau de utilidade**, para um agente e que o agente preferirá estados com maior utilidade;
- A utilidade de um estado é relativa a um agente;
- Uma **função de utilidade** pode levar em conta qualquer conjunto de preferências — peculiares ou típicas, nobres ou perversas.

- As preferências, expressas pelas utilidades, são combinadas com as probabilidades para formar a **teoria de decisão**:

Teoria da Decisão

$$Decisão = Probabilidade + Utilidade$$

- **Teoria da decisão**: Um agente é racional se e somente se escolhe a ação que produz a **maior utilidade esperada** calculada pela média de todos os resultados possíveis da ação;
- Isso é chamado de princípio da utilidade máxima esperada (MEU, do inglês *maximum expected utility*);

function DT-AGENT(*percept*) **returns** an *action*
  persistent: *belief_state*, probabilistic beliefs about the current state of the world
    action, the agent's action

  update *belief_state* based on *action* and *percept*
  calculate outcome probabilities for actions,
    given action descriptions and current *belief_state*
  select *action* with highest expected utility
    given probabilities of outcomes and utility information
  return *action*

Fundamentos de Probabilidade

- Em teoria de probabilidade, o conjunto de todos os estados de mundos possíveis é chamado de **espaço amostral**;
- Os estados de mundos possíveis são **mutuamente exclusivas** e **exaustivas**;
- A letra grega Ω é usada para representar o espaço amostral e a letra ω é usada para representar um elemento do espaço amostral;
- O axioma básico de probabilidade diz que todo estado de mundo possível possui uma probabilidade entre 0 e 1 e a probabilidade do conjunto de estados de mundo possíveis é 1:

$$0 \leq P(\omega) \leq 1 \text{ para todo } \omega \text{ e } \sum_{\omega \in \Omega} P(\omega) = 1.$$

Noção básica de probabilidade

- Por exemplo, imagine um lance de dois dados de 6 lados justos, os resultados possíveis seriam: (1,1), (1,2), (1,3),... (6,6);
- Considerando os dados justos, a probabilidade de cada resultado possível é 1/36;
- Afirmações e consultas probabilísticas geralmente são sobre conjuntos de realizações;
- **Por exemplo**: qual a probabilidade de dois dados somarem 11, ou a probabilidade de que os dados rolem o mesmo número, etc;
- Na teoria da probabilidade, esses conjuntos são chamados de eventos.

Comparação com a Lógica Proposicional

- Em **lógica** esses conjuntos são chamados de **proposições** e é basicamente sinônimo do termo **eventos** em probabilidade, mas as proposições são escritas com linguagem formal.
- A **probabilidade** associada à uma **proposição** é definida como sendo a soma das probabilidades dos mundos em que ela é válida:

$$P(\phi) = \sum_{\omega \in \phi} P(\omega)$$

onde ϕ é uma proposição.

Noção básica de probabilidade

- Vamos considerar dois casos para o lançamento de dois dados:
  - Obter uma soma de 11 em dados justos:

    $P(Total = 11) = P(5,6) + P(6,5) = \frac{1}{36} + \frac{1}{36} = \frac{1}{18}.$
  - Os dados favorecem números iguais 25% das vezes (Dados Tendencisos):

    $P(Duplos) = \frac{1}{4}$
- No segundo caso, sabemos a probabilidade de duplos sem saber se os dados preferem duplos 6 a duplos 2, etc. Assim como em asserções lógicas, essa asserção restringe o modelo de probabilidade subjacente sem determiná-lo completamente.

Noção básica de probabilidade

- Probabilidades como $P(Total = 11)$ ou $P(Duplos)$ são chamadas de probabilidades **incondicionais** ou probabilidades *a priori*.
- Na maioria das vezes, temos informações já reveladas que chamamos de **evidência**.
- Por exemplo, digamos que lancemos primeiro um dado e tiramos 5, qual a probabilidade de termos um duplo?
- Nesse caso, não estamos interessados na probabilidade incondicional de rolar duplos, mas a probabilidade **condicional** ou *a posteriori*.

- Essa probabilidade é escrita como: $P(Duplos \mid Dado_1 = 5)$
- As probabilidades *a posteriori* são definidas em termos de probabilidades *a priori* da seguinte forma: para quaisquer proposições a e b, temos:

$$P(a \mid b) = \frac{P(a \wedge b)}{P(b)}, \text{ para qualquer } P(b) > 0$$

- Por exemplo:

$$P(Duplos \mid Dado_1 = 5) = \frac{P(Duplos \wedge Dado_1 = 5)}{P(Dado_1 = 5)}.$$

- A definição de **probabilidade condicional** (ou *a posteriori*), pode ser escrita em um forma chamada de **regra do produto**:

Regra do produto

$$P(a \wedge b) = P(a \mid b)P(b)$$

Noção básica de probabilidade

- Variáveis em teoria de probabilidade são chamadas de **variáveis aleatórias** e seus nomes iniciam com uma letra maiúscula (exemplo: $Dado_1$);
- Toda **variável aleatória** mapeia do **espaço amostral** para uma faixa de valores — o conjunto de valores que essa variável pode assumir (exemplo para o lance de dois dados: $\{2, ..., 12\}$, Para o lance do primeiro dado $Dado_1 = \{1, ..., 6\}$;
- Nomes de valores são escritos com letras minúsculas, por exemplo:

$$\sum_x P(X = x)$$

- Variáveis Booleanas tem a faixa verdadeiro, falso ou $0, 1$;
- Proposições da forma $A = verdadeiro$ é abreviada na forma a enquanto $A = falso$ é abreviado como $\neg a$.

Noção básica de probabilidade

- Os intervalos podem ser conjuntos de símbolos arbitrários como $Idade = \{juvenil, adolescente, adulto\}$ ou $Tempo = \{sol, chuva, nublado, neve\}$;
- Variáveis podem ter uma faixa infinita discreta ou contínua;
- Inequalidades são permitidas como $NumeroDeAtomosNoUniverso \geq 10^{70}$;
- Finalmente, podemos combinar esses tipos de proposições elementares usando os conectivos da lógica proposicional, por exemplo: "A probabilidade do paciente possuir uma cárie dado que ele é um adolescente e não possui dor de dente é 0,1":

$$P(Cárie \mid \neg DorDeDente \wedge Jovem) = 0,1.$$

Noção básica de probabilidade

- Em probabilidade, também é comum utilizar virgula para representar conjunção, então podemos reescrever a proposição anterior como

$$P(Cárie \mid \neg DorDeDente, Jovem) = 0,1.$$

- Algumas vezes, precisamos falar de todos os valores de uma variável aleatória, nesse caso, podemos escrever:

$$P(Tempo = Sol) = 0,6$$

$$P(Tempo = Chuva) = 0,1$$

$$P(Tempo = Nublado) = 0,29$$

$$P(Tempo = Neve) = 0,01$$

- ou:

$$P(Tempo) = \langle 0.6, 0.1, 0.29, 0.01 \rangle$$

Noção básica de probabilidade

- Aqui o **P** em negrito indica que o resultado é um vetor de números onde assumimos uma ordenação predefinida $\langle sol, chuva, nublado, neve \rangle$ na faixa de Tempo.
- Dizemos que o **P** define uma **distribuição de probabilidade** para a variável aleatória Tempo.
- A notação **P** também é usada para distribuições condicionais: $\mathbf{P}(X \mid Y)$ fornece os valores de $\mathbf{P}(X = x_i \mid Y = y_j)$ para cada par i, j possível.

Noção básica de probabilidade

- Para variáveis aleatórias contínuas utilizamos a **função densidade de probabilidade** (PDF do inglês *probability density function*) com a notação:

$$P(TemperaturaNoturna = x) = Uniform(x; 18^\circ C; 26^\circ C)$$

- Que significa que a temperatura noturna varia uniformemente entre 18 e 26 graus célsius.
- Escrevemos a densidade de probabilidade para uma variável aleatória discreta X no valor x como $P(X = x)$ ou apenas $P(x)$.

Noção básica de probabilidade

- Além de distribuições em variáveis únicas, precisamos de notação para distribuições em múltiplas variáveis. As vírgulas são usadas para isso.
- Por exemplo, se desejarmos a distribuição conjunta de probabilidade entre tempo e cárie teríamos:

$$\mathbf{P}(Tempo, Cárie) = \mathbf{P}(Tempo \mid Cárie)\mathbf{P}(Cárie)$$

- ao invés de

$$P(T = sol \wedge C = Verdadeiro)P(T = sol \mid C = Verdadeiro)P(C = Verdadeiro)$$

$$P(T = chuva \wedge C = Verdadeiro)P(T = chuva \mid C = Verdadeiro)P(C = Verdadeiro)$$

$$P(T = nublado \wedge C = Verdadeiro)P(T = nublado \mid C = Verdadeiro)P(C = Verdadeiro)$$

$$\vdots$$

Noção básica de probabilidade

- Essa notação também pode ser usada para valores de variáveis, como $P(sol, cárie)$;
- Para completar a semântica, precisamos dizer o que são os mundos e como determinar se uma proposição é válida em um mundo:
  - Um mundo possível é definido como uma atribuição de valores a todas as variáveis aleatórias em consideração.
- Um modelo de probabilidade é completamente determinado pela distribuição conjunta de todas as variáveis aleatórias - as chamadas **distribuição de probabilidade conjunta completa**.

- Relação entre uma proposição e sua negação:

$$\begin{aligned}
P(\neg a) &= \sum_{\omega \in \neg a} P(\omega) \\
&= \sum_{\omega \in \neg a} P(\omega) + \sum_{\omega \in a} P(\omega) - \sum_{\omega \in a} P(\omega) \\
&= \sum_{\omega \in \Omega} P(\omega) - \sum_{\omega \in a} P(\omega) \\
&= 1 - P(a)
\end{aligned}$$

- Probabilidade para uma disjunção:

$$P(a \vee b) = P(a) + P(b) - P(a \wedge b)$$

Noção básica de probabilidade

- Vamos usar um exemplo para discutir distribuição conjunta:
  - **Cavity** = Cárie = {verdadeiro, falso};
  - **Toothache** = dor de dente {verdadeiro, falso};
  - **Catch** = Sonda prende no dente = {verdadeiro, falso}.
- A distribuição de probabilidade conjunta para essas variáveis aleatórias é:

	$toothache$		$\neg toothache$	
	$catch$	$\neg catch$	$catch$	$\neg catch$
$cavity$	0.108	0.012	0.072	0.008
$\neg cavity$	0.016	0.064	0.144	0.576

Noção básica de probabilidade

- A soma das probabilidades conjuntas é 1;
- Por exemplo, probabilidade para *cavity* $\vee$ *toothache* é:

$$P(cavity \vee toothache) = 0.108 + 0.012 + 0.072 + 0.008 + 0.016 + 0.064 = 0.28.$$

- Somar os valores da primeira linha da tabela nos dá a probabilidade incondicional de cárie:

$$P(cavity) = 0.108 + 0.012 + 0.072 + 0.008 = 0.2.$$

Noção básica de probabilidade

- Esse processo é chamado de marginalização, ou somatório - porque somamos a probabilidade para cada valor possível das outras variáveis, retirando-as da equação.
- Para qualquer variável **Y** e **Z**, temos:

$$P(Y) = \sum_{z} P(Y, Z = z)$$

- Para o exemplo odontológico, teríamos:

$$\begin{aligned}
P(Cavity) =& P(Cavity, toothache, catch) + P(Cavity, toothache, \neg catch) + \\
& P(Cavity, \neg toothache, catch) + P(Cavity, \neg toothache, \neg catch) \\
P(Cavity) =& \langle 0.108, 0.016 \rangle + \langle 0.012, 0.064 \rangle + \langle 0. - 72, 0.144 \rangle + \langle 0.008, 0.576 \rangle \\
P(Cavity) =& \langle 0.2, 0.8 \rangle
\end{aligned}$$

Noção básica de probabilidade

- Podemos substituir $P(Y, z)$ na equação de marginalização por $P(Y \mid z)P(z)$ para obter a **regra de condicionamento**:

Regra de condicionamento

$$P(Y) = \sum_z P(Y \mid z)P(z)$$

- Na maioria dos casos, estamos interessados em calcular probabilidades condicionais de algumas variáveis, dadas evidências sobre outras.
- Por exemplo, podemos calcular a probabilidade de uma cárie, dada evidência de uma dor de dente, como:

$$P(cavity \mid toothache) = \frac{P(cavity \wedge toothache)}{P(toothache)} = \frac{0.108 + 0.012}{0.108 + 0.012 + 0.016 + 0.064} = 0.6$$

Noção básica de probabilidade

- Podemos verificar essa probabilidade calculando a probabilidade de não termos uma cárie e termos dor de dente:

$$
\begin{aligned}
P(\neg cavity \mid toothache) &= \frac{P(\neg cavity \wedge toothache)}{P(toothache)} \\
&= \frac{0.016 + 0.064}{0.108 + 0.012 + 0.016 + 0.064} = 0.4
\end{aligned}
$$

- Usaremos α como representação de uma constante para normalização de probabilidade:

$$
\begin{aligned}
P(cavity \mid toothache) &= \alpha P(cavity, toothache) \\
&= \alpha[P(cavity, toothache, catch) + P(cavity, toothache, \neg catch)] \\
&= \alpha[\langle 0.108, 0.016\rangle + \langle 0.012, 0.064\rangle] = \alpha\langle 0.12, 0.08\rangle = \langle 0.6, 0.4\rangle
\end{aligned}
$$

- $\alpha = 0.108 + 0.012 + 0.016 + 0.064 = 0.2$

Aplicações

- Diagnóstico em medicina
- Robótica com sensoriamento (precisão limitada pela tecnologia)
- Sistemas de recomendação
- Análise financeira e previsão de risco

Desafios e Limitações

- Modelagem adequada do problema e dependências
- Cálculo eficiente em redes grandes
- Obtenção de dados confiáveis

Resumo

- Agentes inteligentes precisam tratar incertezas para agir no mundo real.
- Probabilidade e inferência são ferramentas cruciais.
- Abordagens modernas incluem redes Bayesianas, Markov, etc.

Para reflexão

- Que estratégias a IA pode usar para aprender probabilidades a partir de dados?
- Quais limitações atuais precisamos superar para diagnosticar doenças raras de modo automatizado?
- Quando a heurística pode ser preferível ao modelo probabilístico completo?

Referências

- Russell, S. J., & Norvig, P. (2021). *Artificial Intelligence: A Modern Approach.* 4ª ed.
- Slides oficiais: *aima.cs.berkeley.edu*
- Outros: materiais complementares em IA.

Obrigado!

E-mail: fabianosoares@unb.br

LinkedIn: https://www.linkedin.com/in/fabiano-soares-06b6a821a/

Site do curso: https://www.fabianosoares.eng.br/fga0221-inteligência-artificial