



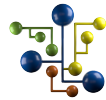
Formação Inteligência Artificial



Introdução à Inteligência Artificial



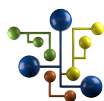
Tomada de Decisão



Teoria da decisão = teoria da probabilidade + teoria da utilidade

Princípio de **Utilidade Máxima Esperada** (UME)





função AGENTE-TD(*percepção*) **retorna** uma *ação*

variáveis estáticas: *estado_de_crença*, crenças probabilísticas sobre o estado atual do mundo
ação, a ação do agente

atualizar *estado_de_crença* com base em *ação* e *percepção*

calcular probabilidades de resultados de ações,

dadas descrições de ações e o *estado_de_crença* atual

selecionar *ação* com utilidade esperada mais alta

dadas as probabilidades de resultados e informações de utilidade

retornar *ação*



Teoria da Decisão



Teoria da Decisão

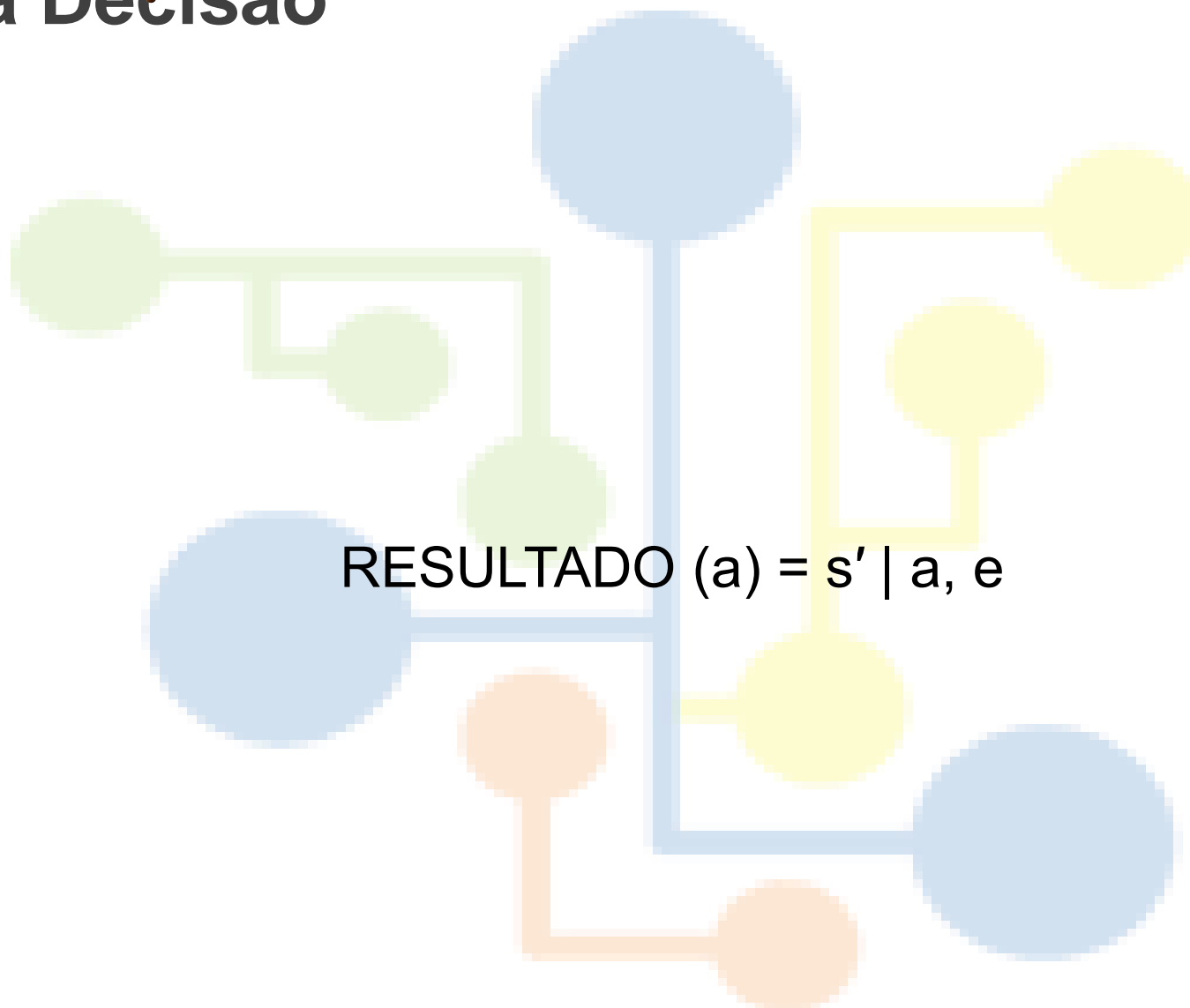
Utilizamos a notação $\text{RESULTADO}(s_0, a)$ para o estado que é o resultado determinístico de tomar a ação a no estado s_0 e neste caso estamos considerando ambientes não determinísticos parcialmente observáveis.

$\text{RESULTADO}(s_0, a)$

Estado

Ação







Teoria da Decisão

A probabilidade do resultado s' , dadas as observações de evidências:

$$P(\text{RESULTADO } (a) = s' \mid a, e)$$

onde o a no lado direito da barra significa o evento em que a ação a é executada.



Teoria da Decisão

As preferências dos agentes são aprendidas por uma função utilidade, $U(s)$, que atribui um único número utilidade para expressar a conveniência de um estado. A utilidade esperada de uma ação, dada a evidência, $UE(a | e)$, é apenas o valor da utilidade média ponderada dos resultados, pela probabilidade que o resultado ocorra:

$$UE(a|e) = \sum_{s'} P(\text{RESULTADO}(a) = s' | a, e) U(s')$$



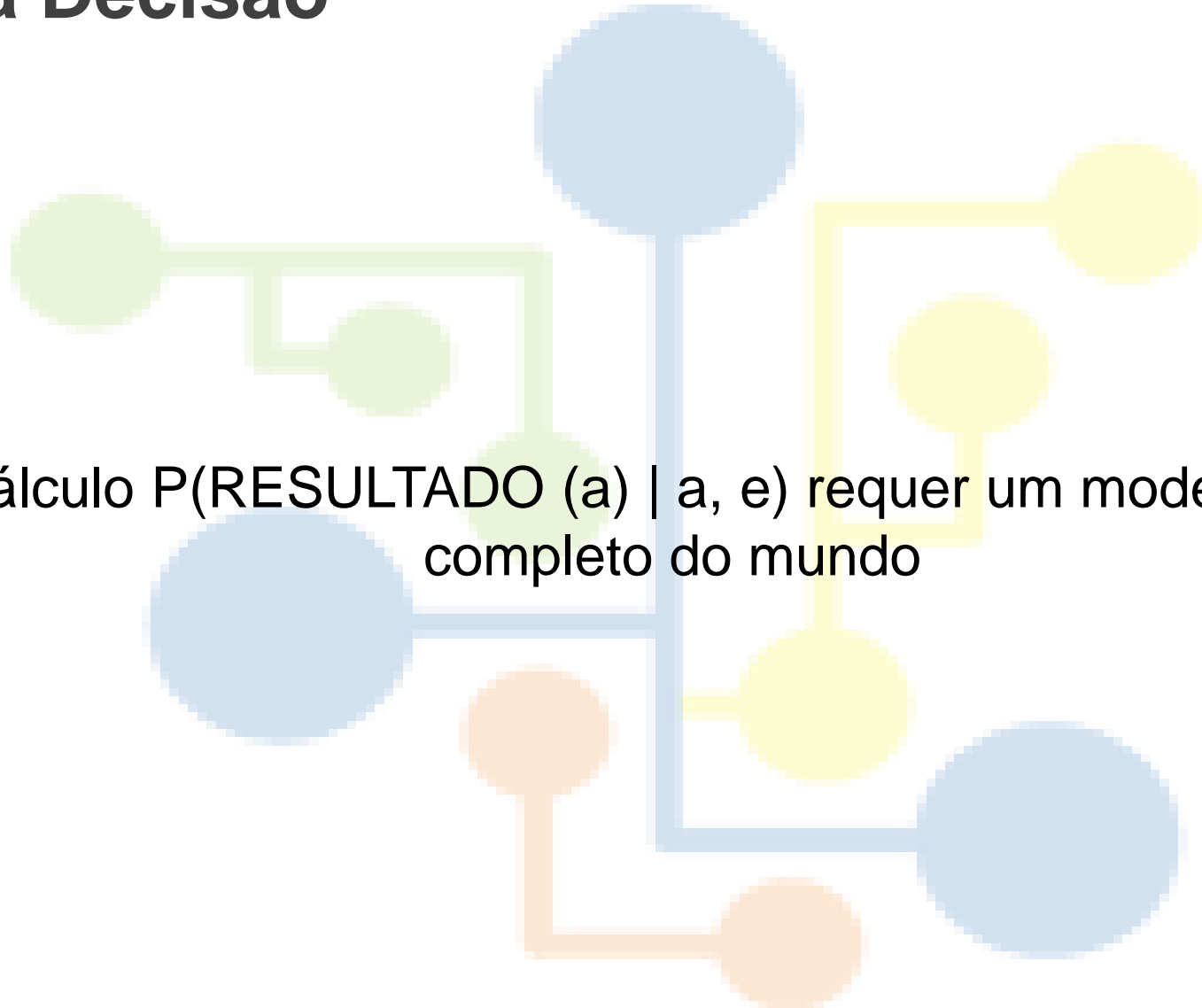


Teoria da Decisão

O princípio da utilidade máxima esperada (UME) diz que um agente racional deve escolher a ação que maximize a utilidade esperada do agente:

$$ação = \underset{a}{\operatorname{argmax}} UE(a \mid e)$$





O cálculo $P(\text{RESULTADO}(a) \mid a, e)$ requer um modelo causal completo do mundo

Teoria da Decisão



Data Science
Academy

Data Science Academy raphaelbsfontenelle@gmail.com 615c1fdde32fc361b30c9ec2



Data Science Academy



Data Science Academy

A Base da Teoria da Utilidade



A Base da Teoria da Utilidade

Intuitivamente, o princípio de utilidade máxima esperada (UME) parece um modo razoável de tomar decisões, mas não é de forma alguma evidente que ele seja o único modo racional.

Afinal, por que maximizar a utilidade média é tão especial?





A Base da Teoria da Utilidade

Utilizamos esta notação a seguir para descrever as preferências de um agente:

$A \succ B$ A é preferível a B .

$A \sim B$ O agente está indiferente entre A e B .

$A \succeq B$ O agente prefere A a B ou está indiferente entre eles.

$$L = [p_1, S_1; p_2, S_2; \dots p_n, S_n]$$





A Base da Teoria da Utilidade

Ordenabilidade: Dadas duas loterias quaisquer, um agente racional deve preferir uma à outra ou, então, classificar as duas como igualmente preferíveis. Ou seja, o agente não pode evitar a decisão.

Apenas $(A > B)$, $(B > A)$, ou $(A \sim B)$ são possíveis.





A Base da Teoria da Utilidade

Transitividade: Dadas três loterias quaisquer, se um agente preferir A a B e preferir B a C, então o agente deverá preferir A a C.

$$(A > B) \wedge (B > C) \Rightarrow (A > C)$$





A Base da Teoria da Utilidade

Continuidade: Se alguma loteria B estiver entre A e C em preferência, haverá alguma probabilidade p de que o agente racional fique indiferente entre escolher B por garantia ou escolher a loteria que produz A com probabilidade p e C com probabilidade $1 - p$.

$$A > B > C \Rightarrow \exists p [p, A; 1 - p, C] \sim B$$





A Base da Teoria da Utilidade

Substitutibilidade: Se um agente está indiferente entre duas loterias A e B, então o agente está indiferente entre duas outras loterias complexas que são a mesma loteria, exceto pelo fato de A ser substituído por B em uma delas. Isso é válido independentemente das probabilidades e do(s) outro(s) resultado(s) das loterias.

$$A \sim B \Rightarrow [p, A; 1 - p, C] \sim [p, B; 1 - p, C]$$





A Base da Teoria da Utilidade

Monotonicidade: Suponha que existam duas loterias que tenham os mesmos dois resultados, A e B. Se um agente prefere A a B, então o agente deve preferir a loteria que tem uma probabilidade mais alta para A (e vice-versa).

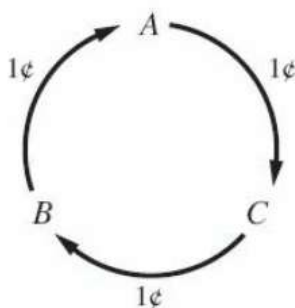
$$A \succ B \Rightarrow (p > q \Leftrightarrow [p, A; 1 - p, B] \succ [q, A; 1 - q, B])$$



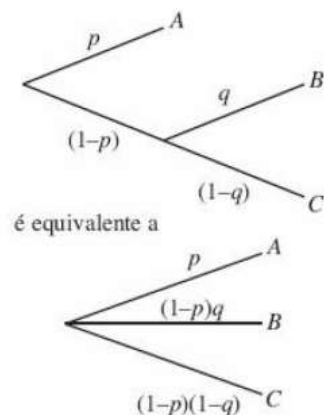


A Base da Teoria da Utilidade

Decomponibilidade: As loterias compostas podem ser reduzidas a loterias mais simples com o uso das leis da probabilidade. Isso se chama regra de “nada de diversão no jogo” porque afirma que duas loterias consecutivas podem ser compactadas em uma única loteria equivalente, como mostra a figura b abaixo:



(a)



(b)

$$[p, A; 1-p, [q, B; 1-q, C]] \sim [p, A; (1-p)q, B; (1-p)(1-q), C]$$





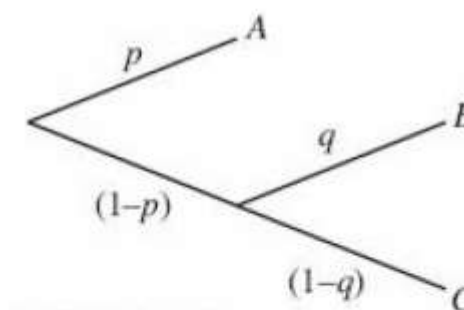
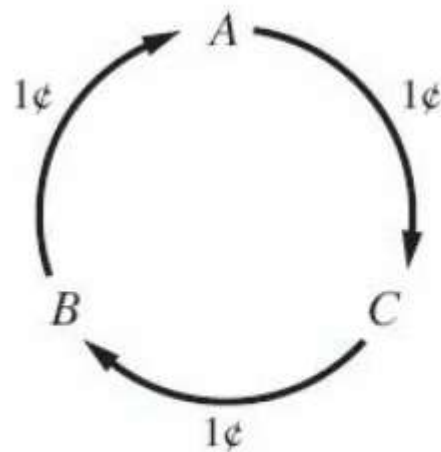
A Base da Teoria da Utilidade

Axiomas da teoria da utilidade

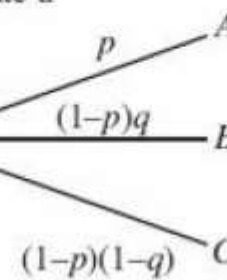




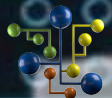
A Base da Teoria da Utilidade



é equivalente a



(b)



Data Science
Academy

Data Science Academy raphaelbsfontenelle@gmail.com 615c1fdde32fc361b30c9ec2

Redes de Decisão



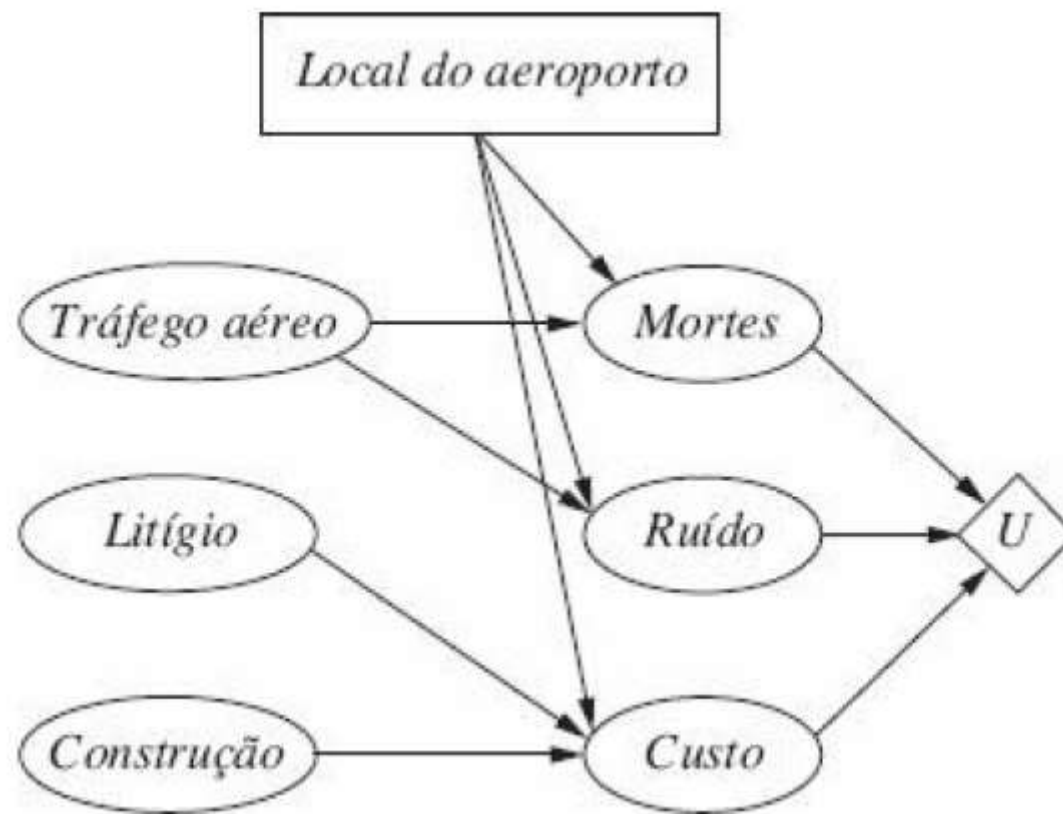
Data Science Academy

Redes de Decisão



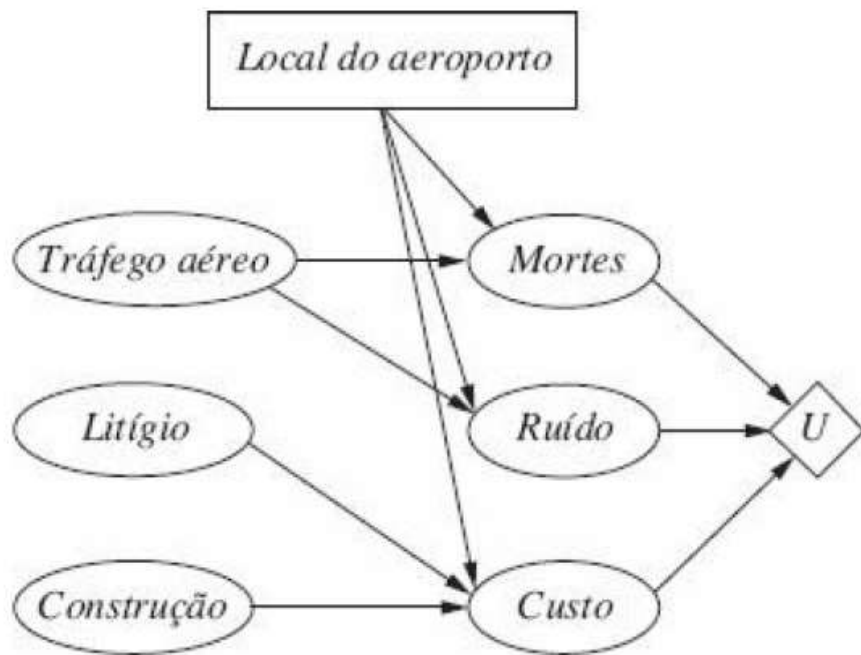
Data Science
Academy

Data Science Academy raphaelbsfontenelle@gmail.com 615c1fdde32fc361b30c9ec2





Redes de Decisão



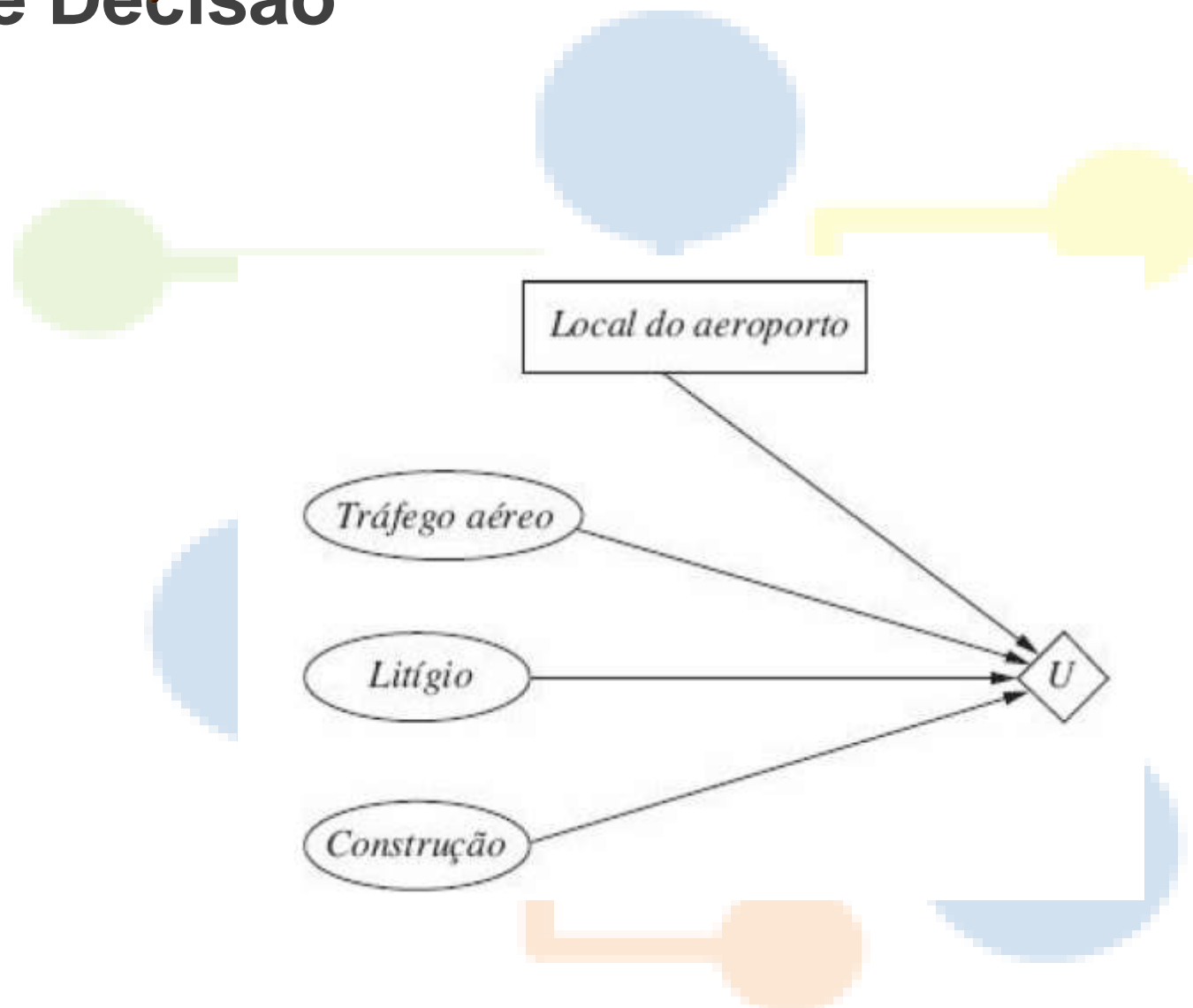
- Nós de acaso (elipses)
- Nós de decisão (retângulos)
- Nós de utilidade (losangos)

Redes de Decisão



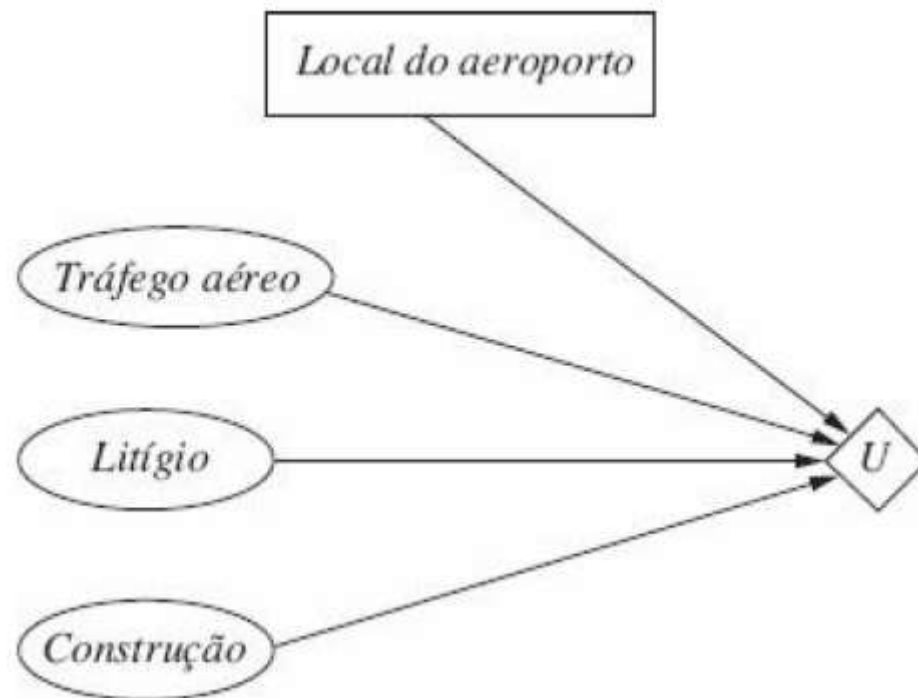
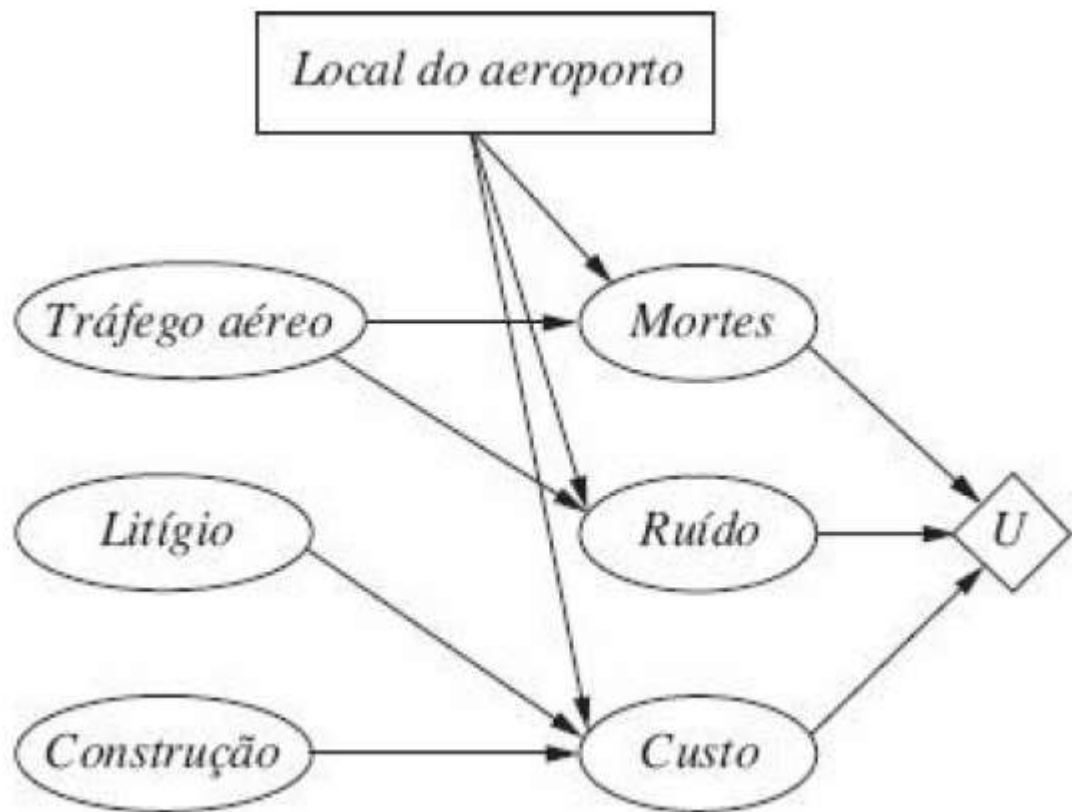
Data Science
Academy

Data Science Academy raphaelbsfontenelle@gmail.com 615c1fdde32fc361b30c9ec2





Redes de Decisão





Redes de Decisão

Avaliação de redes de decisão

1. Definir as variáveis de evidência para o estado atual.
2. Para cada valor possível do nó de decisão:
 - (a) Definir o nó de decisão com esse valor.
 - (b) Calcular as probabilidades posteriores para os nós pais do nó de utilidade, usando um algoritmo-padrão de inferência probabilística.
 - (c) Calcular a utilidade resultante para a ação.
3. Retornar a ação com a utilidade mais alta.



Agente de Coleta de Informações



Agente de Coleta de Informações

função AGENTE-DE-COLETA-DE-INFORMAÇÕES(*percepção*) **retorna** uma *ação*
variáveis estáticas: D , uma rede de decisão

integrar *percepção* a D

$j \leftarrow$ o valor que maximiza $VIP(E_j) / Custo(E_j)$

se $VIP(E_j) > Custo(E_j)$

então retornar SOLICITAR(E_j)

senão retornar a melhor ação a partir de D



Sistemas de Teoria da Decisão



Sistemas de Teoria da Decisão

Como cada vez mais processos de decisão tornam-se **automatizados**, a análise de decisão é cada vez mais utilizada para garantir que os processos automatizados se comportem como desejado.





Sistemas de Teoria da Decisão

As primeiras pesquisas de sistemas especialistas se concentravam em responder a perguntas, e não na tomada de decisões.





Sistemas de Teoria da Decisão

Um sistema que incorpora utilidades pode evitar uma das armadilhas mais comuns associadas ao processo de consulta:

confundir probabilidade e importância.

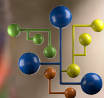




Sistemas de Teoria da Decisão

Descreveremos agora o processo de engenharia do conhecimento para sistemas especialistas de teoria da decisão. Como exemplo, vamos considerar o problema de selecionar um tratamento médico para uma espécie de doença congênita do coração, em crianças.



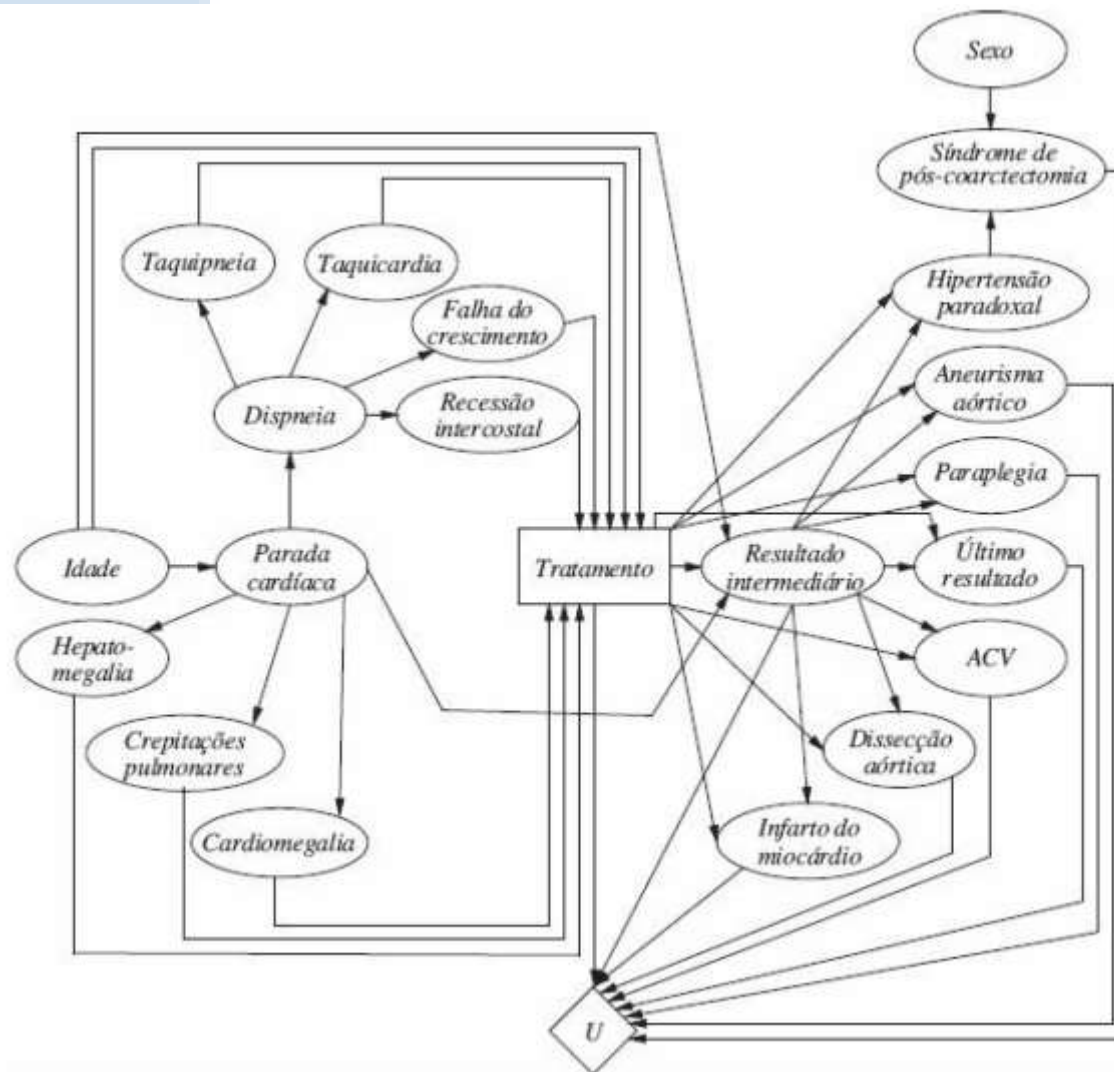


Cerca de 0,8% das crianças nascem com uma anomalia do coração, sendo a mais comum o estreitamento da aorta (uma constrição da aorta).

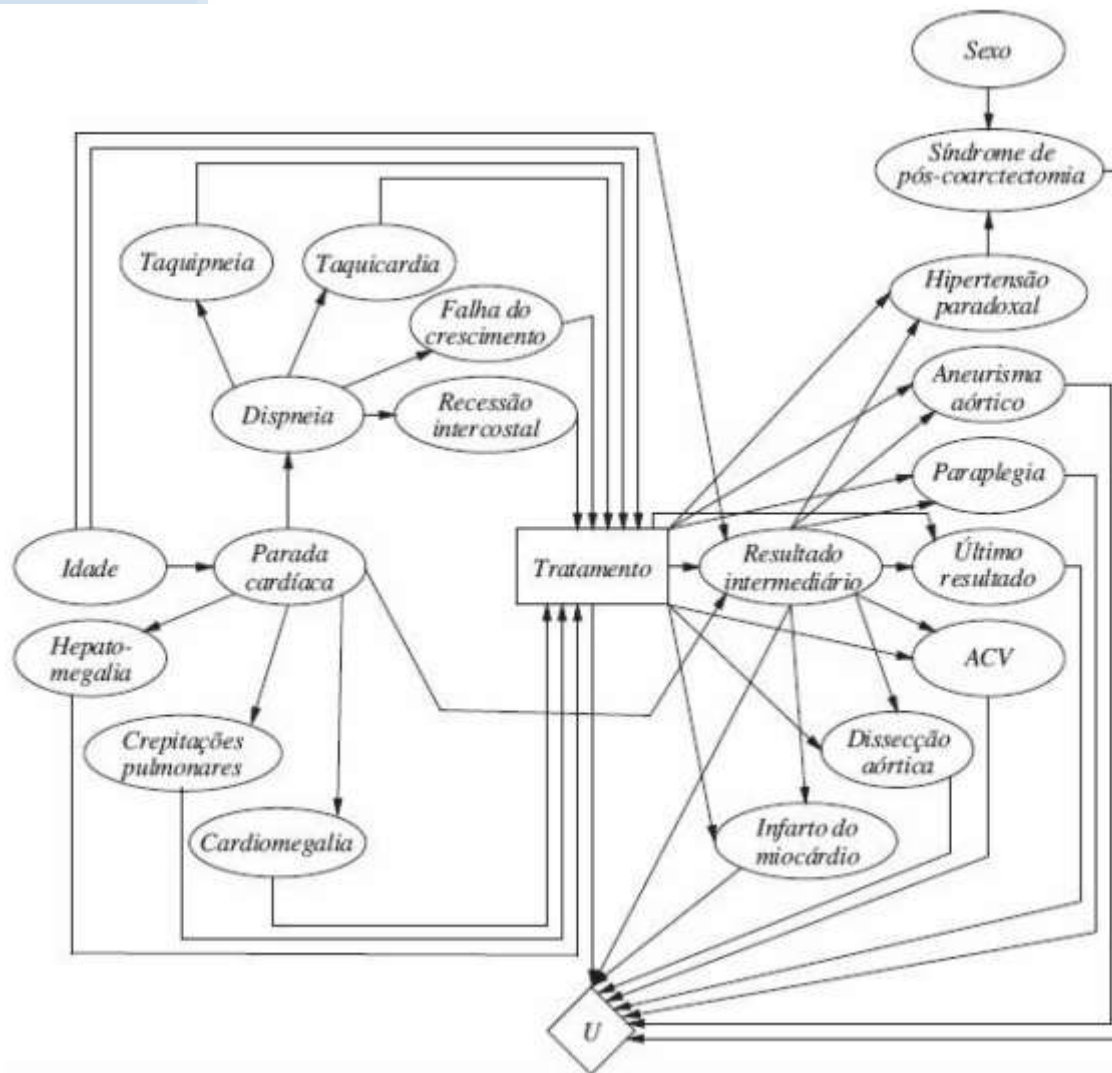


Sistemas de Teoria da Decisão

Crie um modelo causal.
Determine quais são os
sintomas possíveis, as
doenças, os tratamentos
e os resultados.



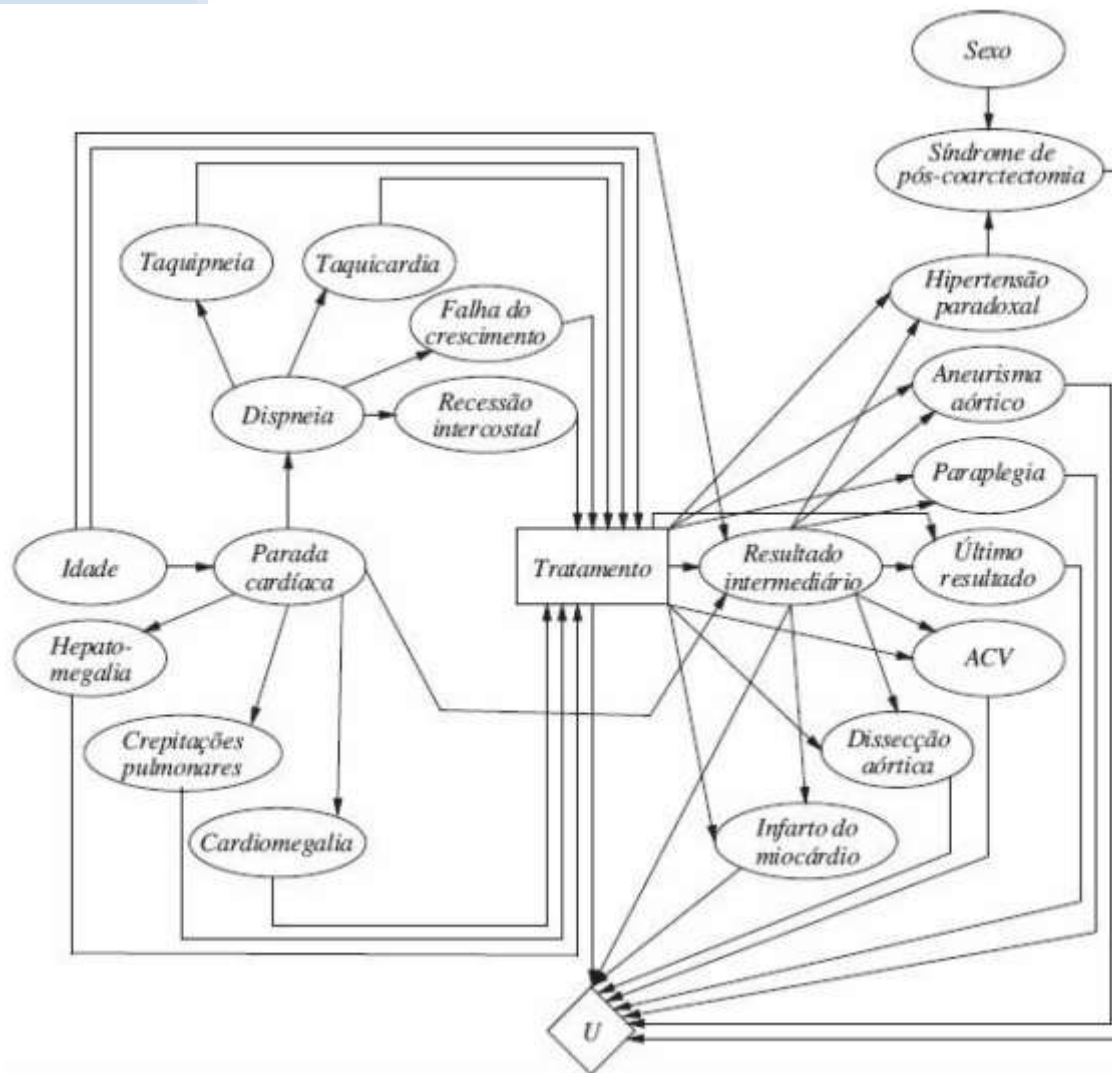
fique até chegar a
modelo de decisão
ativa.





Sistemas de Teoria da Decisão

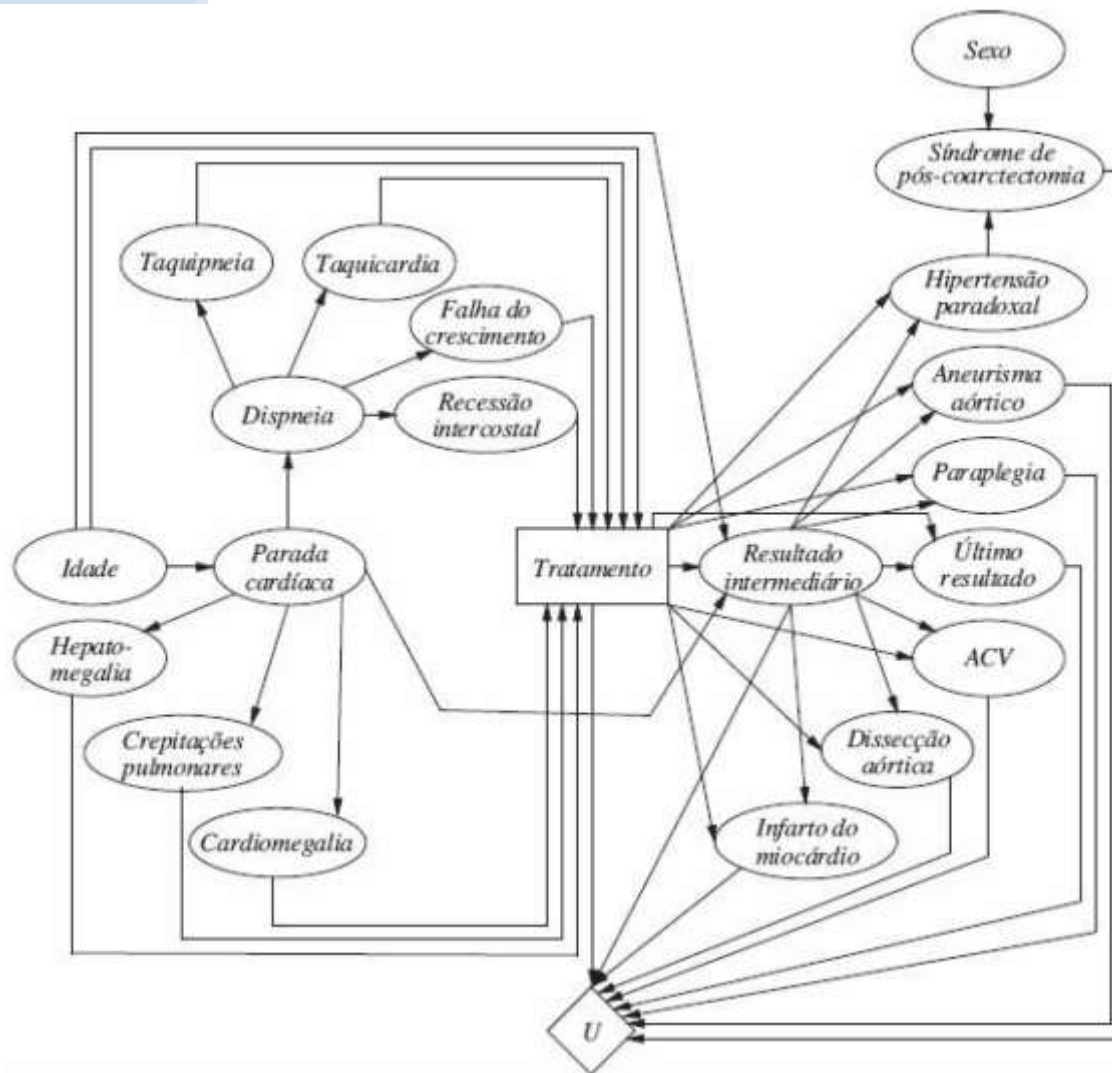
Atribua probabilidades.



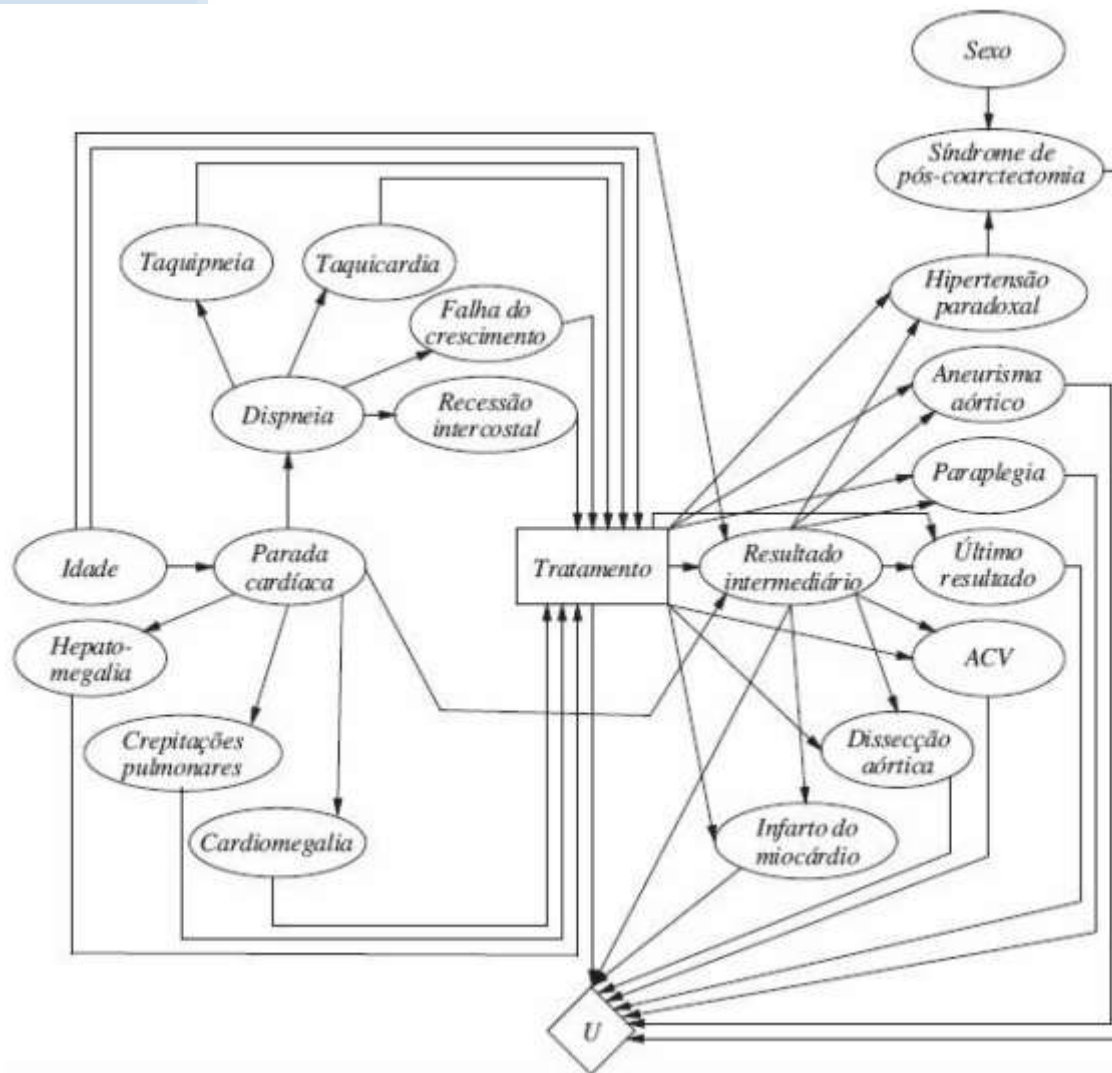


Sistemas de Teoria da Decisão

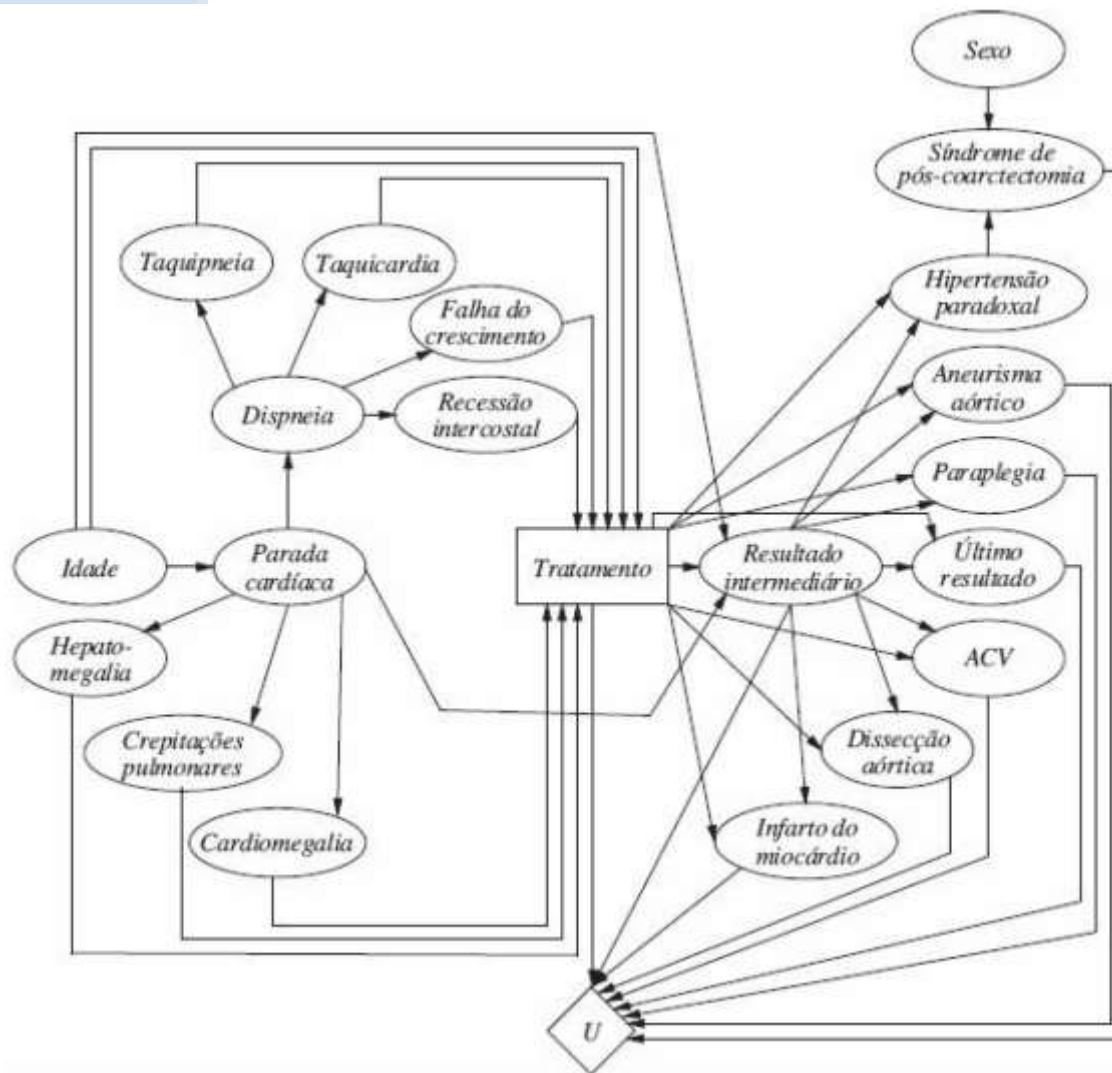
Atribua utilidades.



Verifique e refine o modelo.



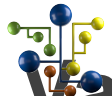
Execute a análise de sensibilidade.



Iteração de Valor

Apresentaremos agora, um algoritmo chamado de iteração de valor.

A ideia básica é calcular a utilidade de cada estado e, em seguida, usar as utilidades do estado para selecionar uma ação ótima em cada estado.

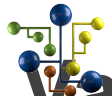


A utilidade de um estado s é dada por:

$$U(s) = R(s) + \gamma \max_{a \in A(s)} \sum_{s'} P(s' | s, a) U(s')$$

A utilidade de um estado é a recompensa imediata correspondente a esse estado mais a utilidade descontada esperada do próximo estado, assumindo que o agente escolha a ação ótima.



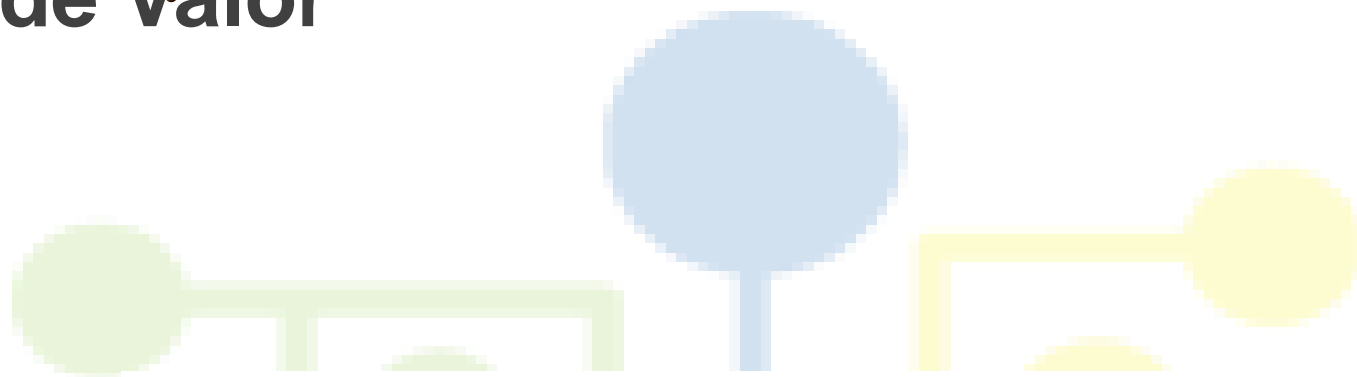
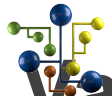


A utilidade de um estado s é dada por:

$$U(s) = R(s) + \gamma \max_{a \in A(s)} \sum_{s'} P(s' | s, a) U(s')$$

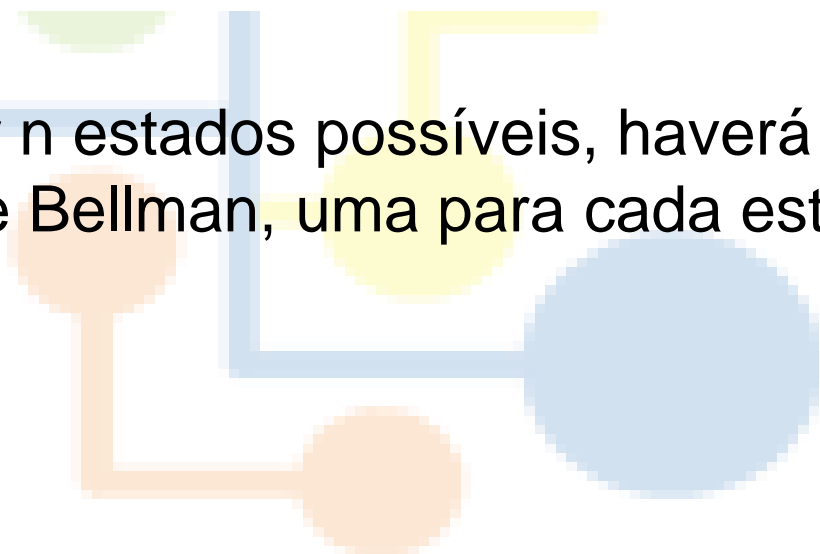
Equação de Bellman

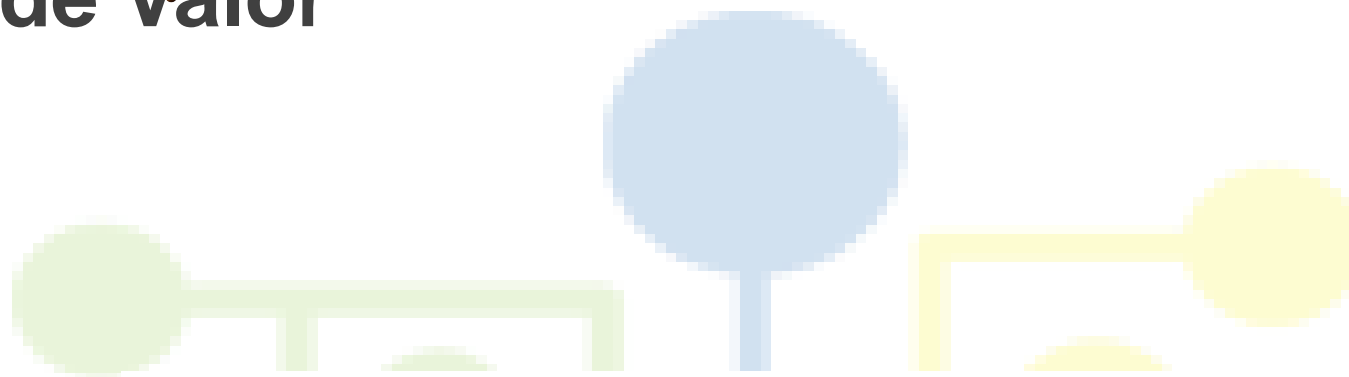
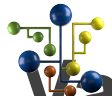




$$U_{i+1}(s) \leftarrow R(s) + \gamma \max_{a \in A(s)} \sum_{s'} P(s' | s, a) U_i(s')$$

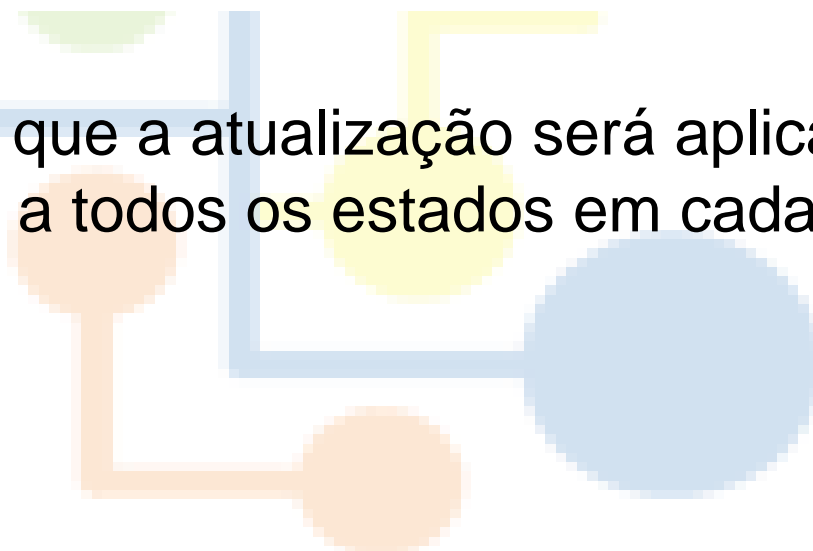
Se houver n estados possíveis, haverá n equações de Bellman, uma para cada estado

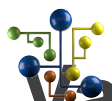




$$U_{i+1}(s) \leftarrow R(s) + \gamma \max_{a \in A(s)} \sum_{s'} P(s' | s, a) U_i(s')$$

Assumimos que a atualização será aplicada simultaneamente a todos os estados em cada iteração





Iteração de Valor

função ITERAÇÃO-DE-VALOR(mdp, ϵ) **retorna** uma função utilidade

entradas: mdp , um MDP com estados S , ações $A(s)$, modelo de transição $P(s'|s,a)$, recompensas $R(s)$, desconto γ , ϵ , o erro máximo permitido na utilidade de qualquer estado

variáveis locais: U, U' , vetores de utilidades para estados em S , inicialmente zero

δ , a mudança máxima na utilidade de qualquer estado em uma iteração

repita

$U \leftarrow U'; \delta \leftarrow 0$

para cada estado s **em** S **faça**

$$U'[s] \leftarrow R[s] + \gamma \max_{a \in A(s)} \sum_{s'} P(s' | s, a) U[s']$$

se $|U'[s] - U[s]| > \delta$ **então** $\delta \leftarrow |U'[s] - U[s]|$

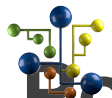
até $\delta < \epsilon(1 - \gamma)/\gamma$

retornar U





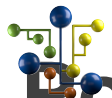
Iteração de Política



Iteração de Política

Se uma ação é claramente melhor que todas as outras, a magnitude exata das utilidades nos estados envolvidos não precisa ser exata.



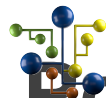


Iteração de Política

The diagram illustrates the Policy Iteration process. It features two main dark blue rounded rectangular boxes. The left box contains the text 'Avaliação de política: Dada uma política, calcular a utilidade de cada estado se a política fosse executada.' The right box contains the text 'Aperfeiçoamento de política: Calcular uma nova política utilizando a observação de um passo para frente baseado em uma utilidade.' The boxes are connected by a network of lines and circles. A light green line connects the top of the left box to a light green circle. A light blue line connects the top of the right box to a light blue circle. A yellow line connects the bottom of the right box to a yellow circle. An orange line connects the bottom of the left box to an orange circle. These circles are further interconnected by a central vertical blue line and horizontal lines, forming a complex network structure.

Avaliação de política:
Dada uma política, calcular a
utilidade de cada estado se a
política fosse executada.

Aperfeiçoamento de política:
Calcular uma nova política
utilizando a observação de um
passo para frente baseado em
uma utilidade.



Iteração de Política



função ITERAÇÃO-DE-POLÍTICA(mdp) **retorna** uma política

entradas: mdp , um MDP com estados S , ações $A(s)$, modelo de transição $P(s' | s, a)$

variáveis locais: U , um vetor de utilidades para estados em S , inicialmente zero

π , um vetor de política indexado pelo estado, inicialmente aleatório

repita

$U \leftarrow \text{AVALIAÇÃO-DE-POLÍTICA}(\pi, U, mdp)$

$inalterado? \leftarrow \text{verdadeiro}$

para cada estado s **em** S **faça**

se $\max_{a \in A(s)} \sum_{s'} P(s' | s, a) U[s'] > \sum_{s'} P(s' | s, \pi[s]) U[s']$ **então faça**

$\pi[s] \leftarrow \operatorname{argmax}_{a \in A(s)} \sum_{s'} P(s' | s, a) U[s']$

$inalterado? \leftarrow \text{falso}$

até $inalterado?$

retornar π





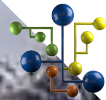
Teoria dos Jogos



Teoria dos Jogos

A Teoria dos Jogos, a qual poderia se chamar muito apropriadamente de Teoria das Decisões Interdependentes, tem como objeto de análise situações onde o resultado da ação de indivíduos, grupo de indivíduos, ou instituições, depende substancialmente das ações dos outros envolvidos.





Teoria dos Jogos

A Teoria dos Jogos fornece a linguagem para a descrição de processos de decisão conscientes e objetivos envolvendo mais do que um indivíduo.



Teoria dos Jogos



Data Science
Academy

Data Science Academy raphaelbsfontenelle@gmail.com 615c1fdu

A teoria dos jogos é usada para se estudar assuntos tais como eleições, leilões, balanço de poder, evolução genética, etc.



Teoria dos Jogos



Data Science
Academy

Data Science Academy raphaelbsfontenelle@gmail.com 615c1fdu

Em 1950, John F. Nash demonstrou o teorema minimax para grandes números de agentes (decisões com vários agentes).



Análise, projeto e implementação de mecanismos de alocação de recursos.





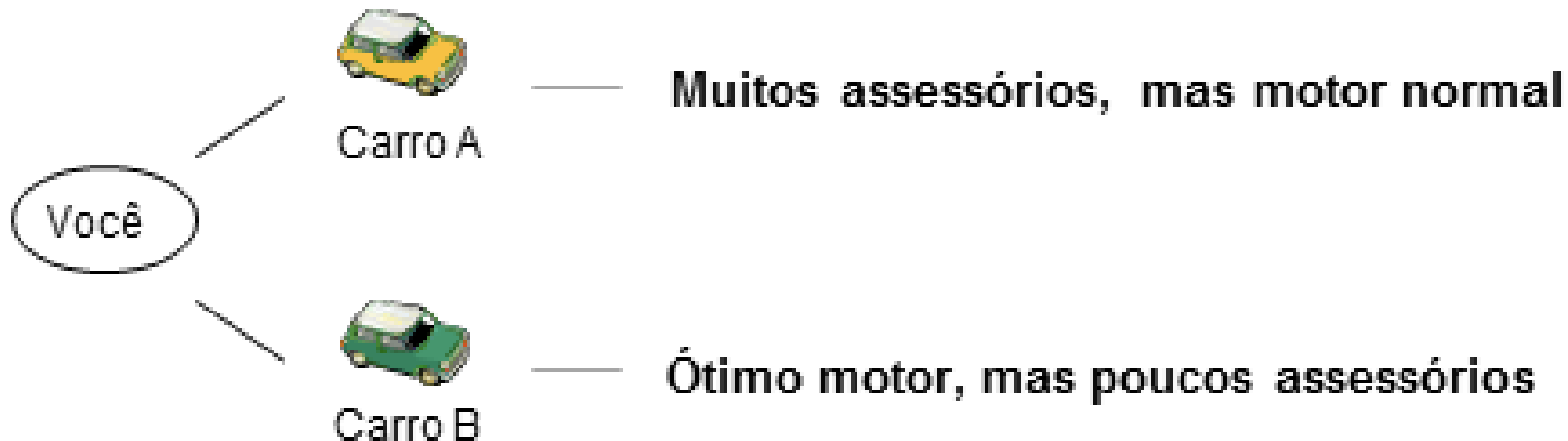
Teoria dos Jogos

Teoria dos Jogos é o estudo das tomadas de decisões entre indivíduos quando o resultado de cada um depende das decisões dos outros, numa interdependência similar a um jogo.



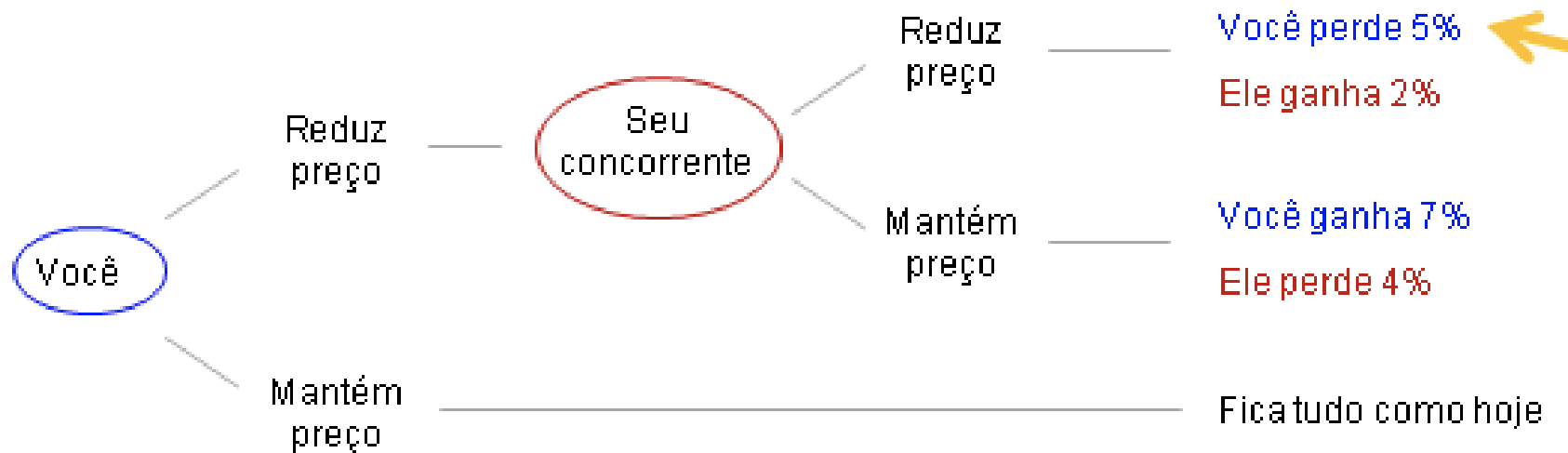


O exemplo do carro é uma **decisão isolada** - a decisão é só sua e não há interferência de outros no resultado.





A **Teoria dos Jogos** estuda cenários onde existem vários interessados em otimizar os próprios ganhos, as vezes em conflito entre si.





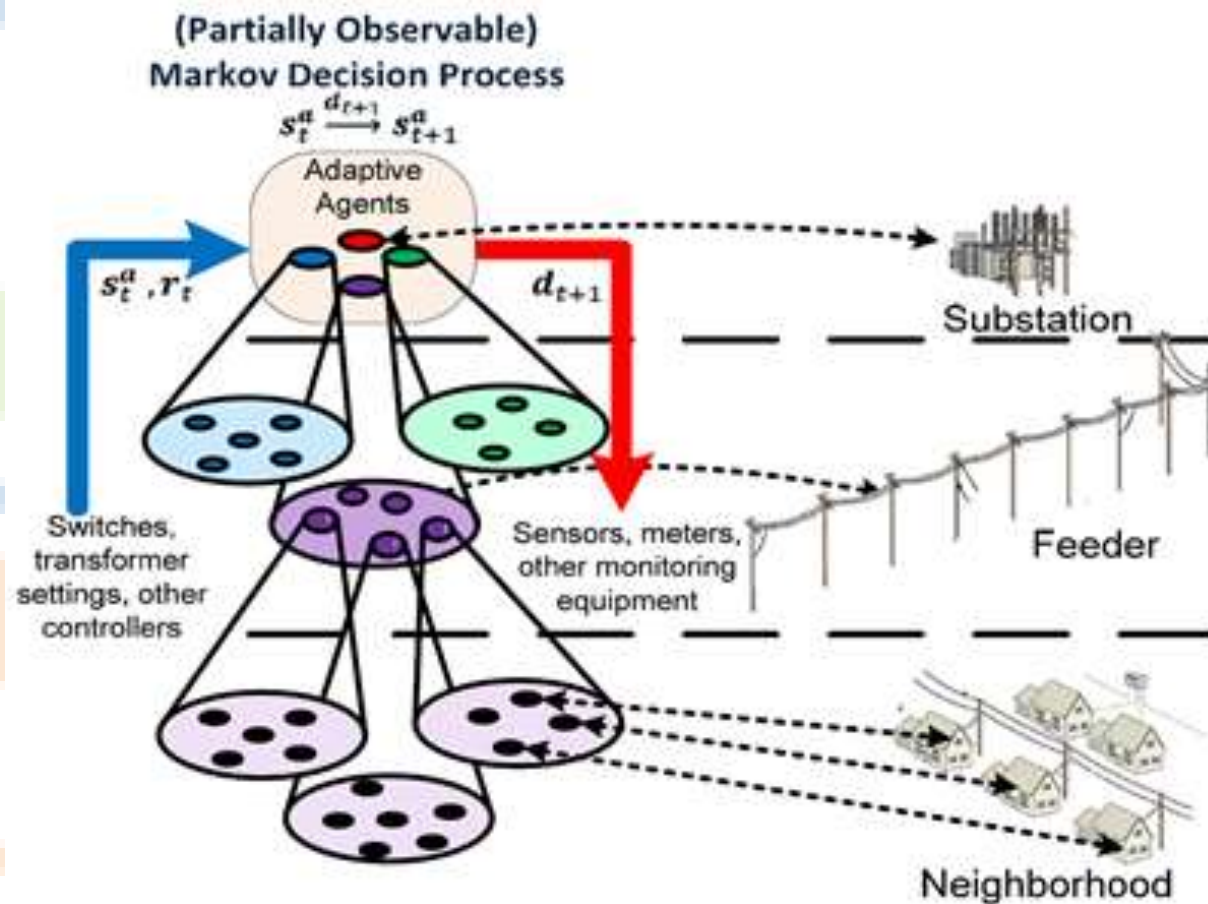
Por isso, a melhor recomendação é: antes de tomar uma decisão, coloque-se no lugar do concorrente e imagine qual seria a reação dele dadas as ações e incentivos existentes!

Decisões com Vários Agentes



Decisões com Vários Agentes

- E se a incerteza se deve a outros agentes e às decisões que eles tomam?
- E se as decisões desses agentes, por sua vez, forem influenciadas por nossas decisões?





Decisões com Vários Agentes

Projeto de agentes

The diagram features two large, dark blue rounded rectangular boxes. The left box contains the text 'Projeto de agentes' and the right box contains 'Projeto de mecanismo'. These boxes are connected by a network of lines and circles in light blue, yellow, and orange. The lines form a complex web, with some circles acting as nodes or junctions. The background is white.

Projeto de mecanismo



Decisões com Vários Agentes

Não é fundamental que as ações ocorram exatamente ao mesmo tempo; o que importa é que nenhum jogador tenha conhecimento das escolhas dos outros jogadores!





Decisões com Vários Agentes





Decisões com Vários Agentes

Na Teoria das Jogos, está em jogo o pensamento estratégico!

Prisioneiro Vermelho

Prisioneiro Azul

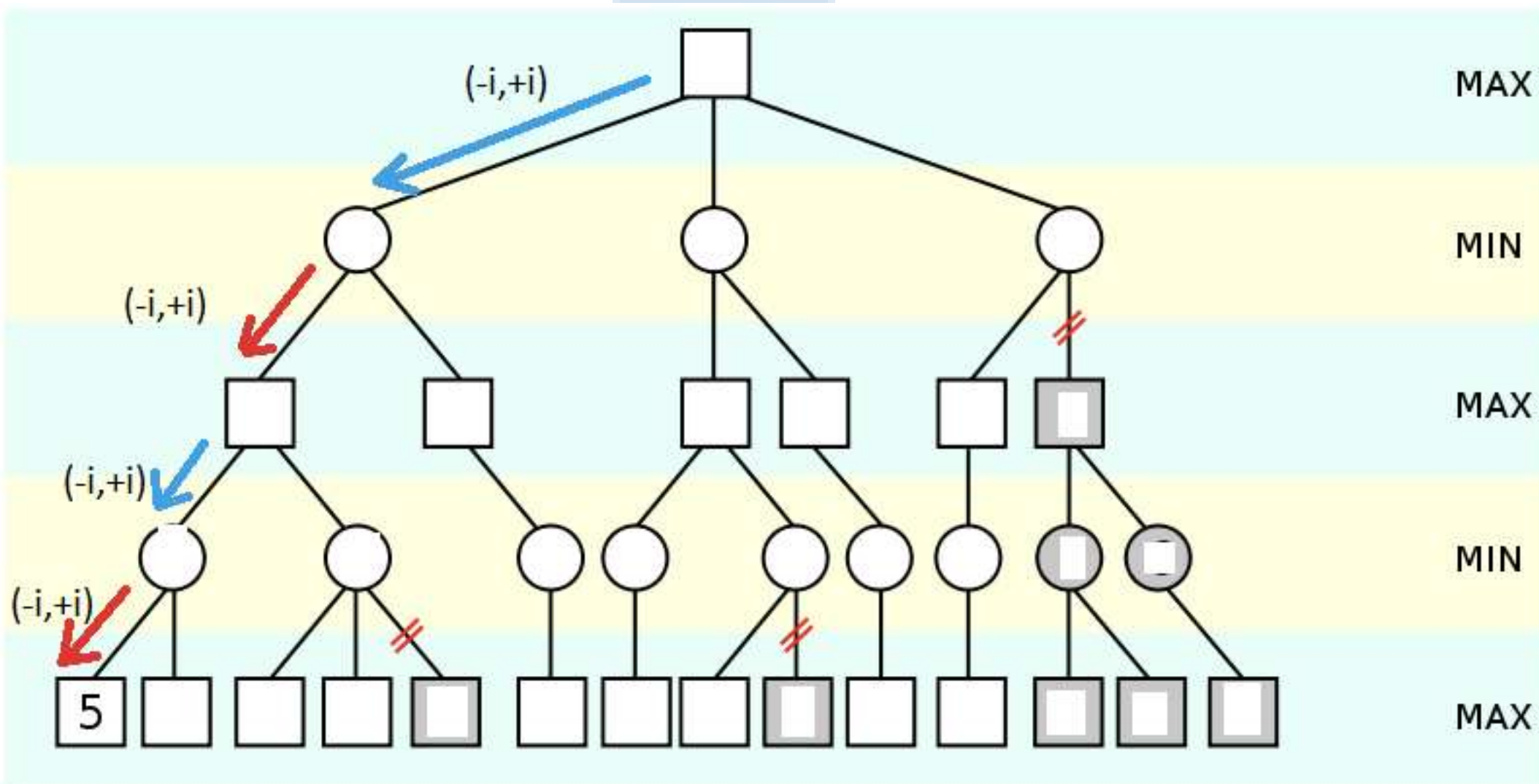
	CONFESSA	NÃO CONFESSA
CONFESSA	<p>PENA DE 5 ANOS</p> 	<p>PENA DE 10 ANOS</p> 
NÃO CONFESSA	<p>É SOLTO!</p> 	<p>PENA DE 6 MESES</p> 

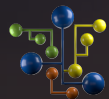
	CONFESSA	NÃO CONFESSA
CONFESSA	<p>PENA DE 5 ANOS</p> 	<p>É SOLTO!</p> 
NÃO CONFESSA	<p>PENA DE 10 ANOS</p> 	<p>PENA DE 6 MESES</p> 





Decisões com Vários Agentes





Data Science
Academy

Data Science Academy raphaelbsfontenelle@gmail.com 615c1fdde32fc361b30c9ec2

Obrigado



Data Science Academy



Data Science Academy