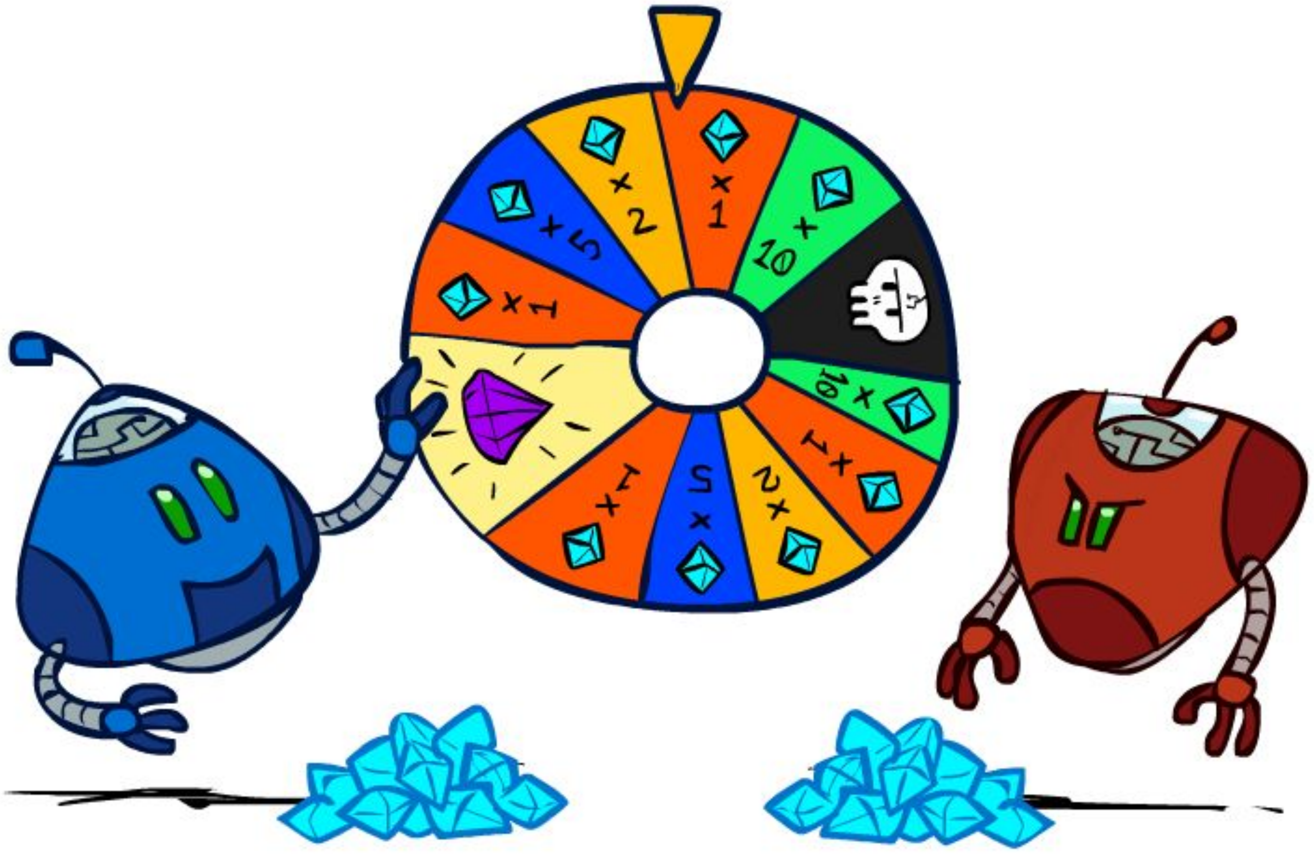
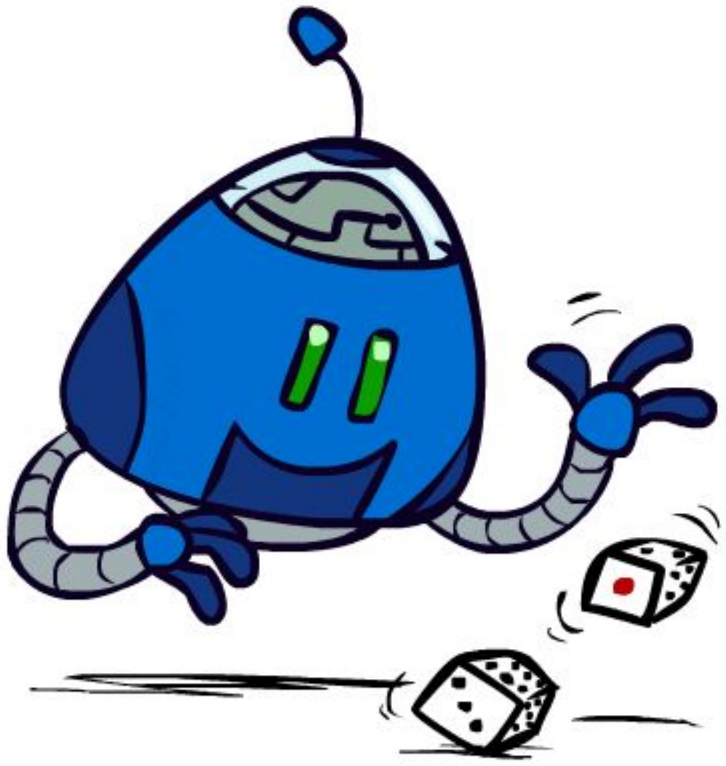
Incerteza e Utilidades

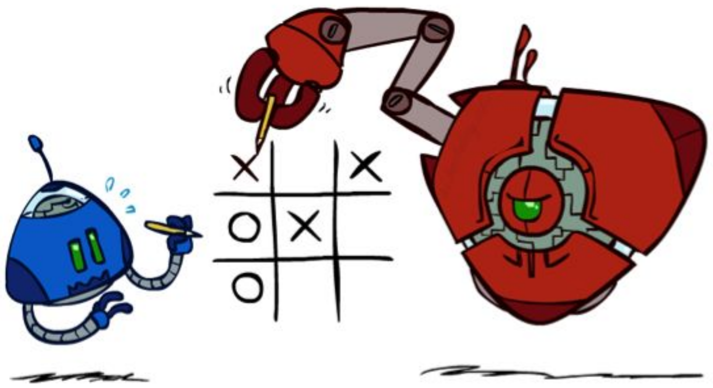


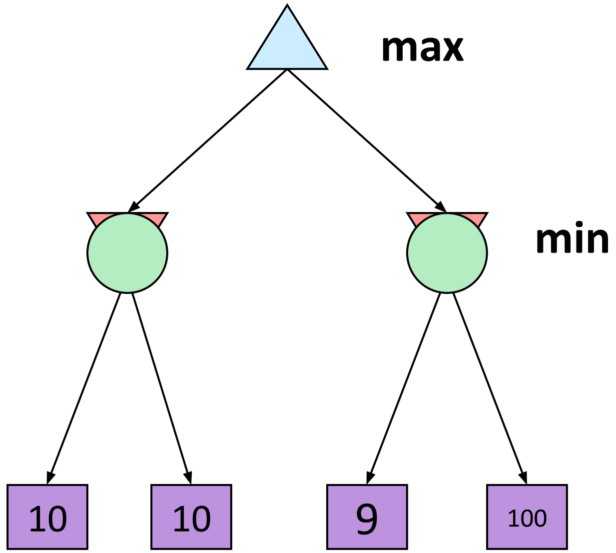
[baseado em Dan Klein and Pieter Abbeel for CS188 Intro to AI at UC Berkeley.   http://ai.berkeley.edu.]

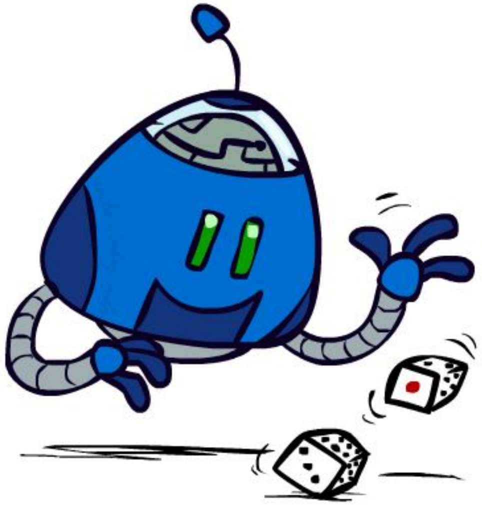
Jogos Estocásticos



Pior caso vs. Caso médio



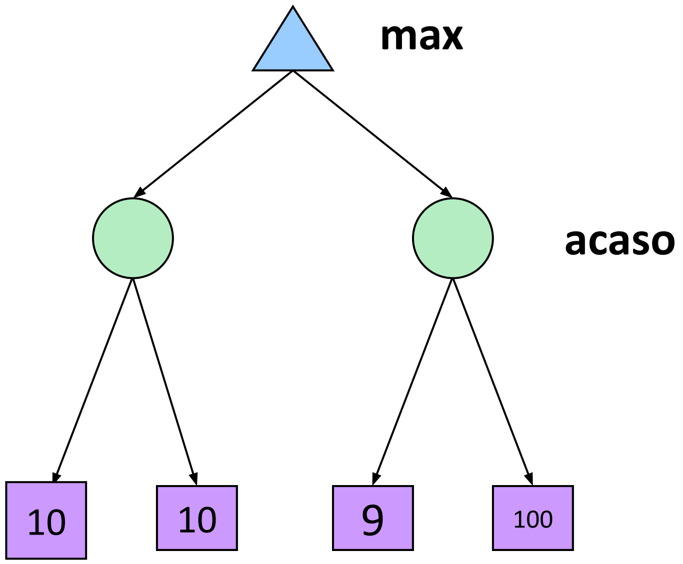




Resultado depende de sorte
Adiciona-se o nível do acaso à árvore de busca.

Expectimax Search

Quando não saberíamos o resultado de uma ação?
Rolagem de dados, pneu de um robô pode escorregar.



Valores na árvore refletem o caso médio (expectimax) e não o pior caso (minimax).

Busca Expectimax:

1. MAX funciona da mesma forma.
2. Nós acaso são como MIN, mas o resultado é incerto.
3. Calcula-se a utilidade esperada dos nós.

Implementação Expectimax

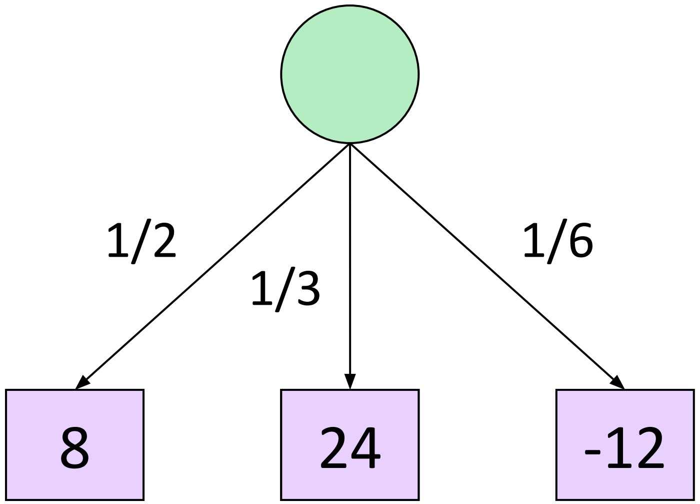
```
def valor(estado):
    se estado é terminal: retorne a utilidade do estado
    se agente é MAX: retorne max-valor(estado)
    se agente é EXP: retorne exp-valor(estado)
```

```
def max-valor(estado):
    inicialize v = -∞
    para cada sucessor do estado:
        v = max(v, valor(sucessor))
    retorne v
```

```
def exp-valor(estado):
    inicialize v = 0
    para cada sucessor do estado:
        p = probabilidade(sucessor)
        v += p * valor(sucessor)
    retorne v
```

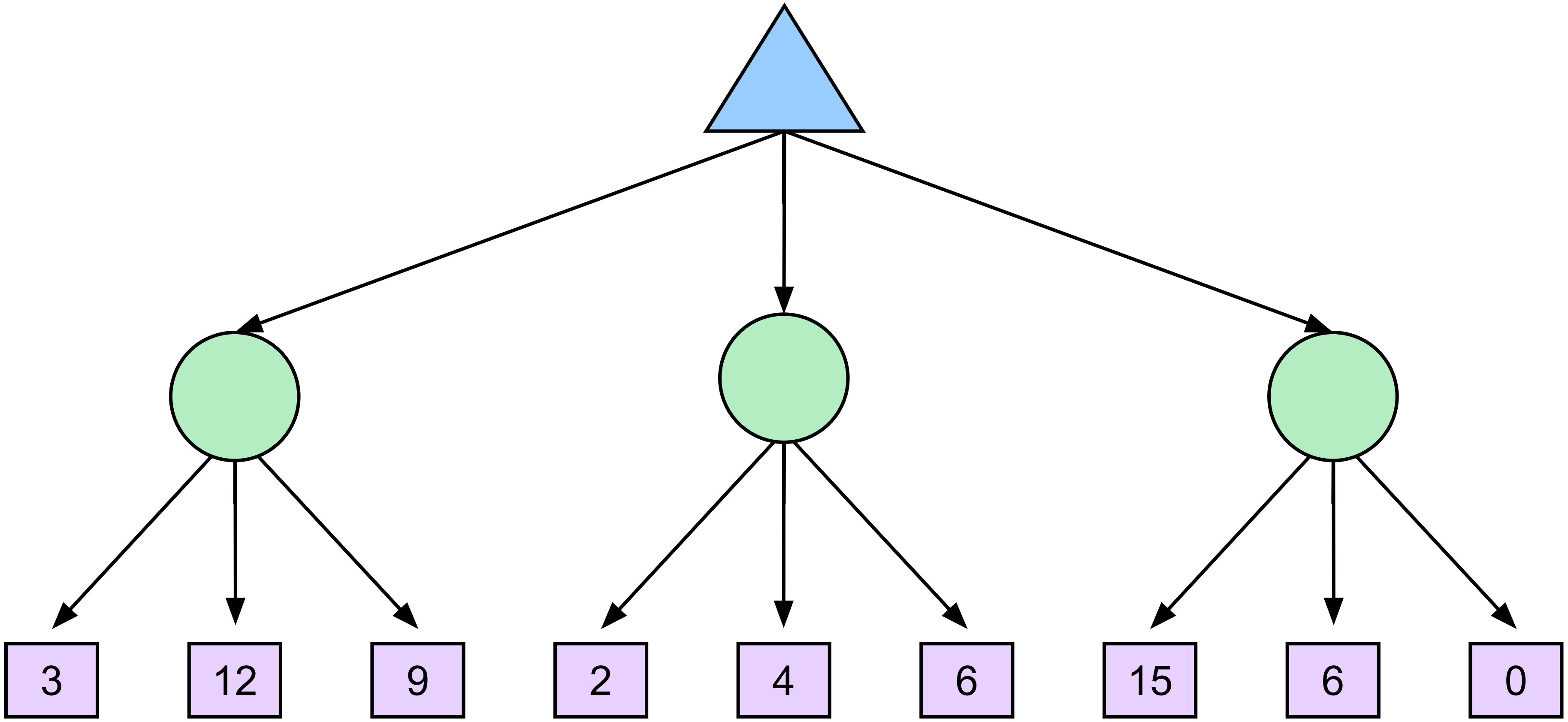
Implementação Expectimax

```
def exp-valor(estado):  inicialize v = 0
  para cada sucessor do estado:  p =
    probabilidade(sucessor)  v += p *
    valor(sucessor)
retorne v
```

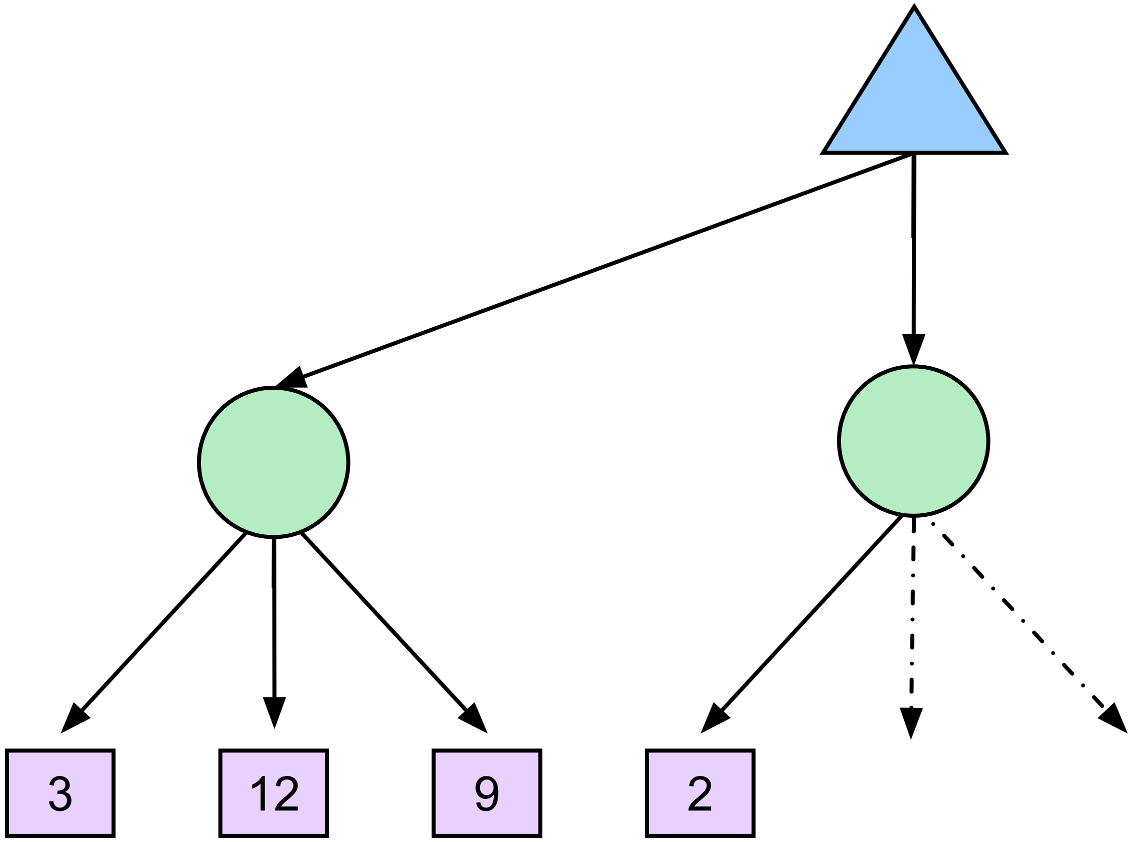


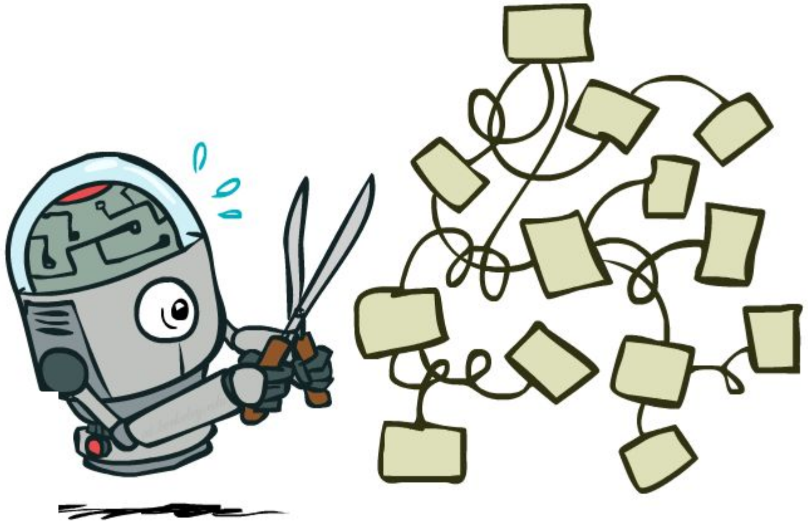
$$v = (1/2)\ (8) + (1/3)\ (24) + (1/6)\ (-12) = 10$$

Expectimax Exemplo



Poda Expectimax?

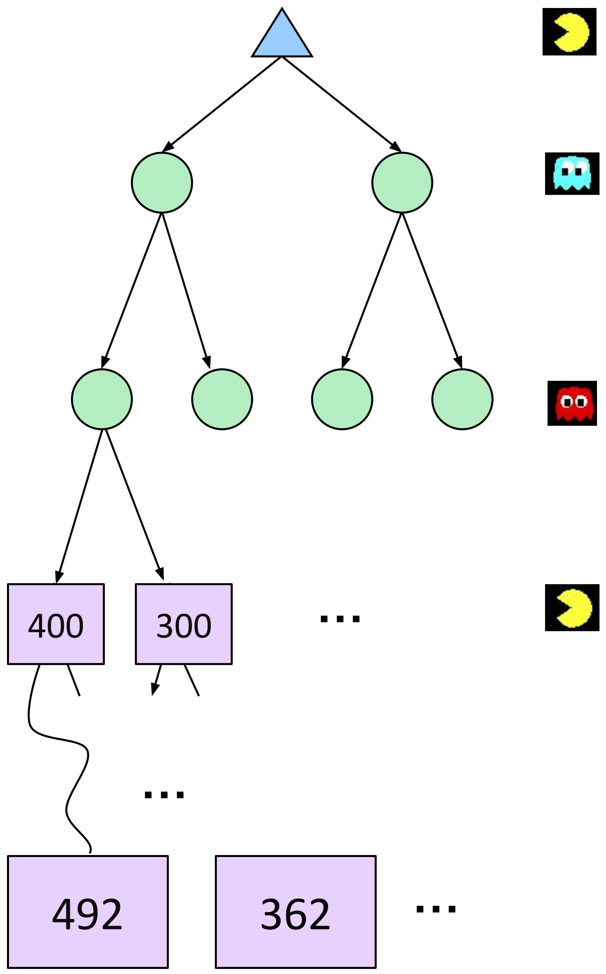




Depth-Limited Expectimax

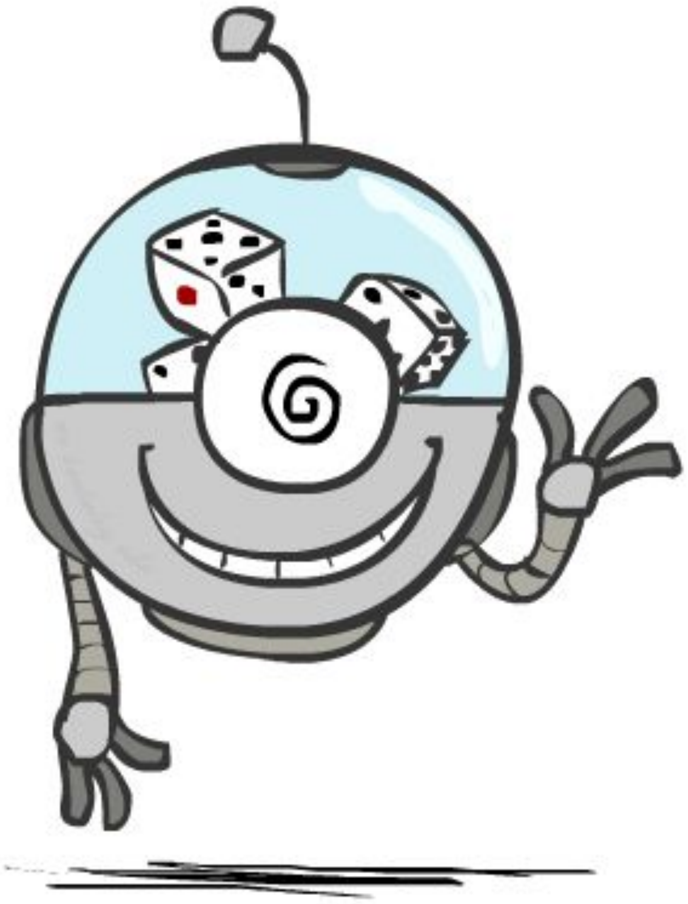
- Utiliza-se aprofundamento iterativo e truncamento de busca.
- Utiliza-se função de avaliação ou simulações Monte Carlo.

… estima o valor real de expectimax (poderia demorar muito para calcular)



Designa-se por ***método de Monte Carlo*** *(MMC) qualquer método de uma classe de métodos estatísticos que se baseiam em amostragens aleatórias massivas para obter resultados numéricos. Em suma, utilizam a aleatoriedade de dados para gerar um resultado para problemas que a priori são determinísticos.* ***Wikipedia***

Probabilidades



Relembrando: Probabilidades

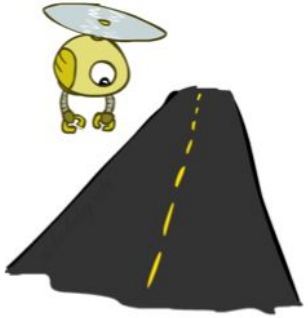
- Uma variável aleatória representa um evento cujo resultado é desconhecido
- Uma distribuição de probabilidade é um atribuição de pesos aos resultados

- Exemplo: Engarrafamento em uma estrada
  - Variável aleatória: T = engarrafamento
  - Resultados: T {nenhum, leve, pesado}
  - Distribuição: P(T=nenhum) = 0.25, P(T=leve) = 0.50, P(T=pesado) = 0.25

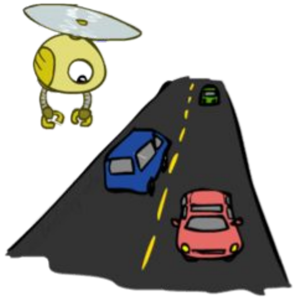
- Algumas leis da probabilidade :
  - Probabilidades são sempre não negativas
  - Probabilidades sobre todo os resultados possíveis soma 1

- A medida que obtemos mais evidências a probabilidades podem mudar:
  - P(T=pesado) = 0.25, P(T=leve | Hora=8:00) = 0.60

Probabilidade condicional



0.25

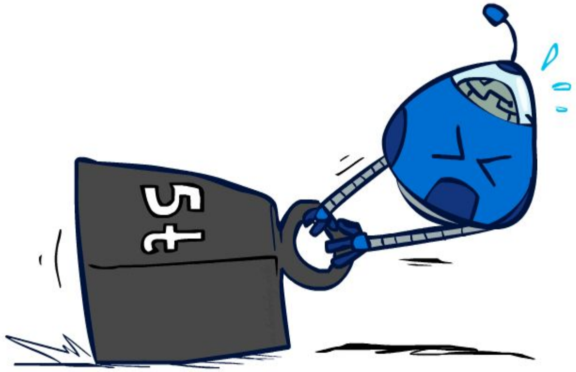


0.50

0.25

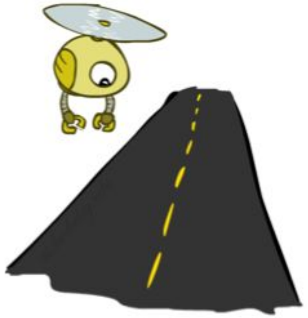
Relembrando: Expectativa

- O valor esperado de uma função de uma variável aleatória é a média, ponderada pela distribuição de probabilidade sobre os resultados



- Exemplo: Quanto tempo para chegar ao aeroporto?

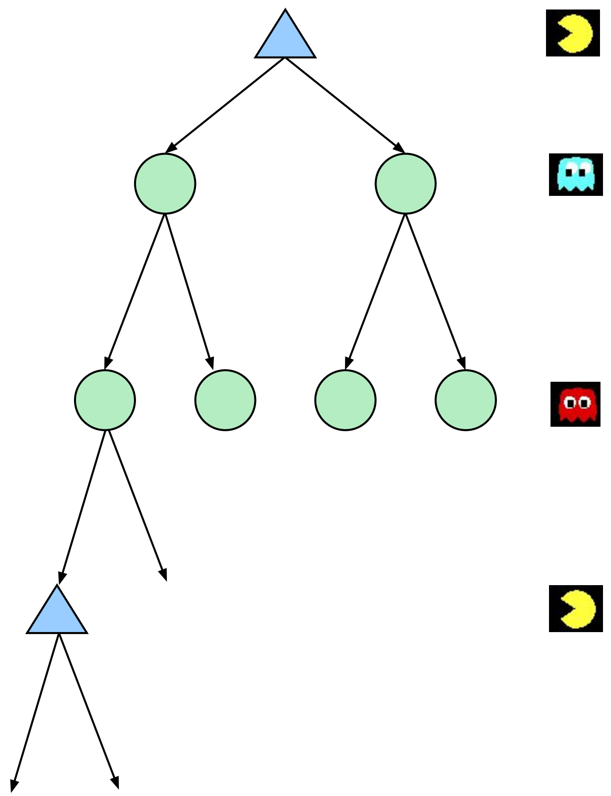
Tempo:	20 min		30 min		60 min	
	x	+	x	+	x	⇒ 35 min
Probabilidade:	0.25		0.50		0.25	



Quais probabilidades usar?

- Na busca expectimax, temos um modelo probabilístico de como o oponente (ou ambiente) se comportará em qualquer estado
  - O modelo pode ser uma distribuição uniforme simples (lançar um dado)
  - O modelo pode ser sofisticado e exigir uma grande quantidade de computação
  - Temos um nó de chance para qualquer resultado fora de nosso controle: oponente ou ambiente
  - O modelo pode dizer que as ações adversárias são prováveis!

- Por enquanto, suponha que cada nó de chance venha magicamente com probabilidades que especificam a distribuição sobre seus resultados



Ter uma crença probabilística sobre a ação de outro agente não significa que o agente está jogando moedas!

Os perigos do otimismo e do pessimismo

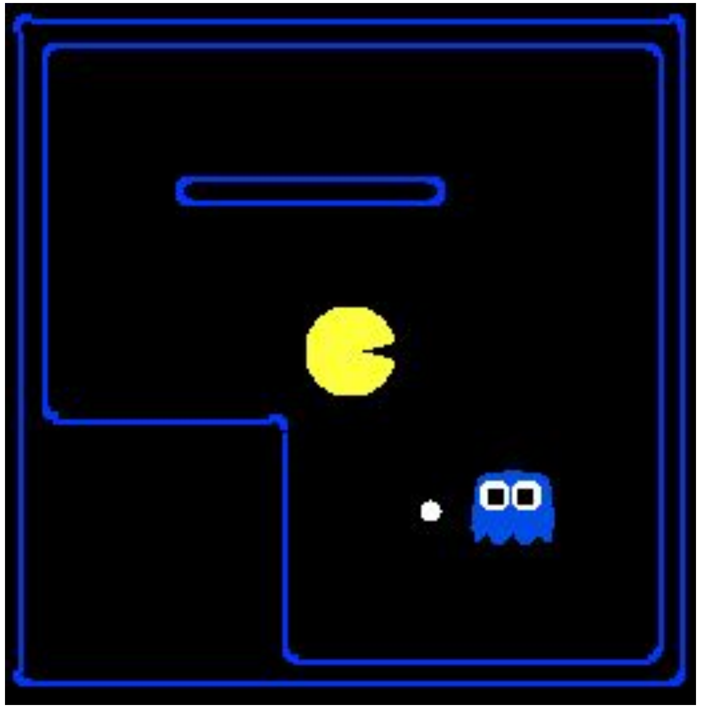
Otimismo Perigoso

Assumindo a chance quando o mundo é perigoso

Pessimismo perigoso

Supondo o pior caso, quando não é provável

Suposições vs. Realidade

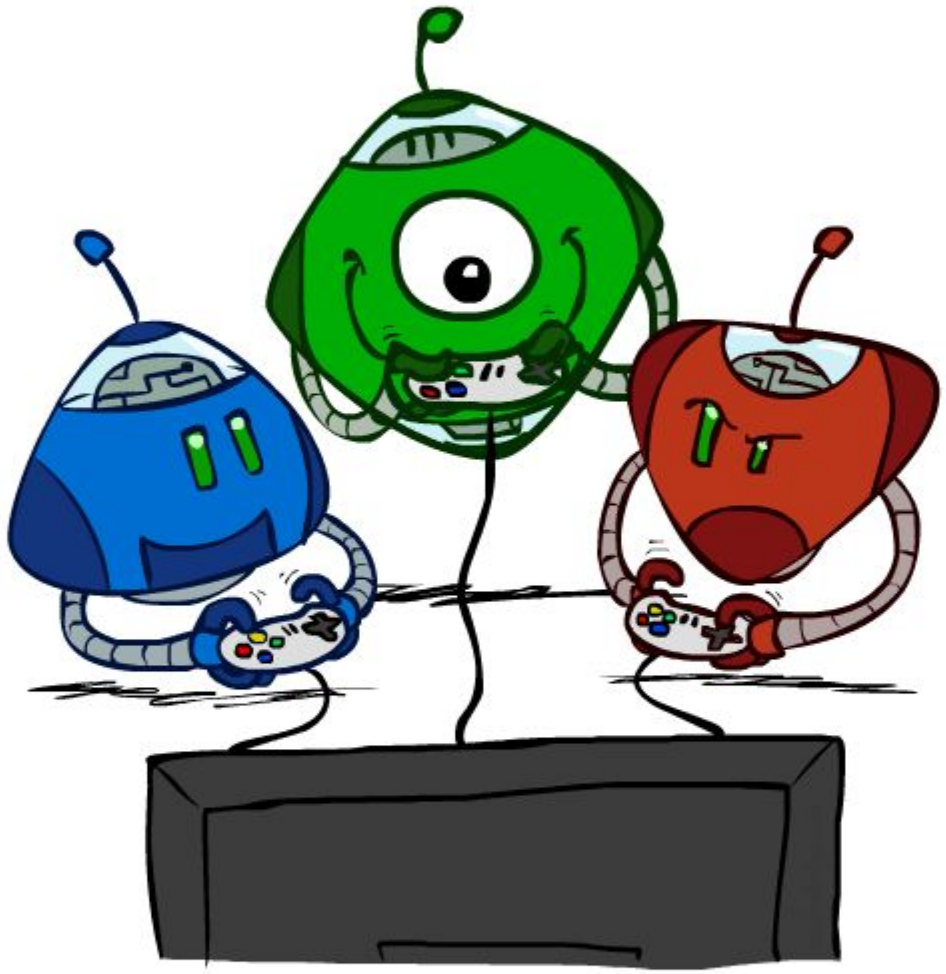


	Fantasma Adversário	Fantasma Aleatório
Minimax Pacman	vence 5/5 Pont. média: 483	vence 5/5 Pont. média: 493
Expectimax Pacman	vence 1/5 Pont. média: -303	vence 5/5 Pont. média: 503

Resultados de 5 jogos

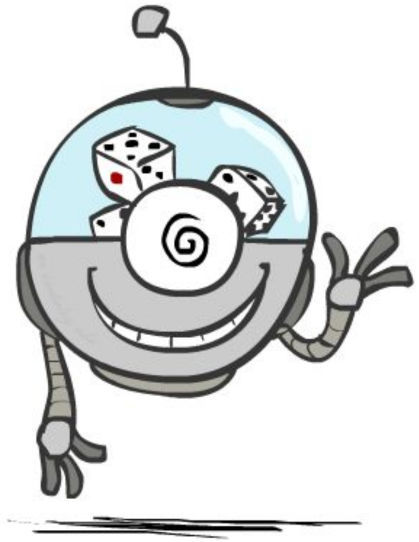
Pacman usou busca em profundidade 4 com uma função de avaliação que evita problemas
O Ghost usou busca em profundidade 2 com uma função de avaliação que busca Pacman

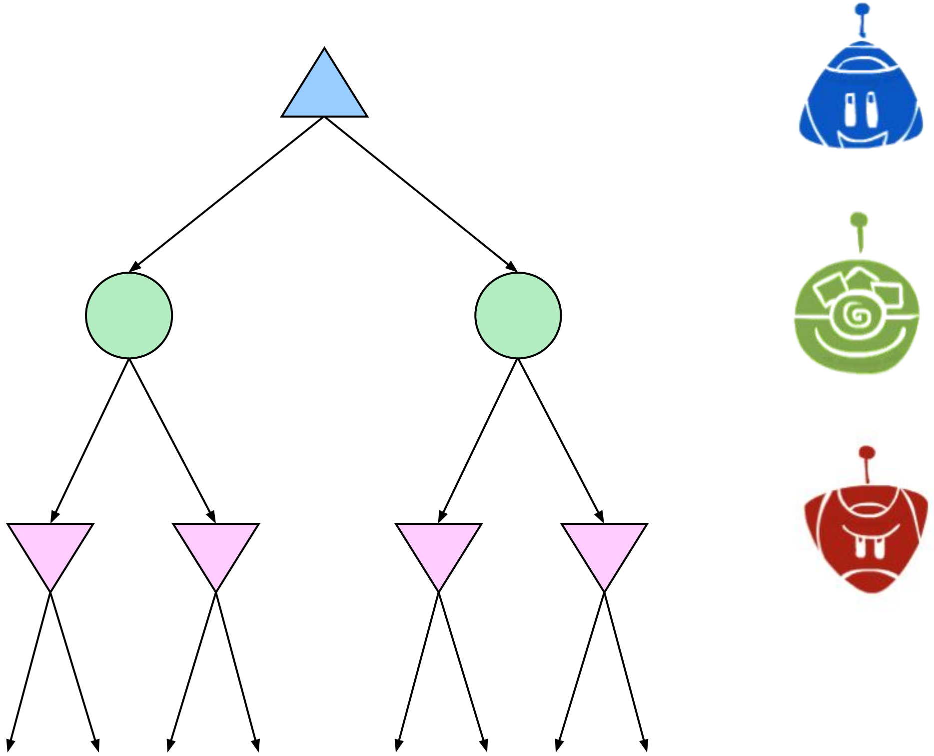
Outros tipos de jogos



Tipos de camadas mistas

- E.g. Gamão
- Expectiminimax
  - O ambiente é um jogador "agente aleatório" extra que se move após cada agente mín/máx.
  - Cada nó calcula a combinação apropriada de seus filhos



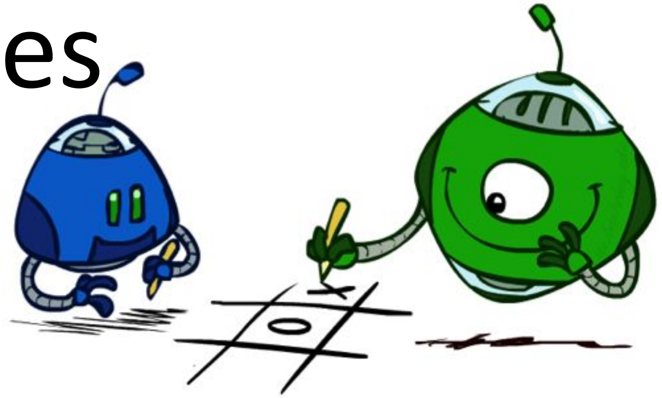


Exemplo: Gamão

- Lançamentos de dados aumentam b: 21 lançamentos possíveis com 2 dados
  - Gamão ≈ 20 movimentos legais
  - Profundidade 2 = $20 \times (21 \times 20)^3 = 1.2 \times 10^9$
- À medida que a profundidade aumenta, a probabilidade de alcançar um determinado nó de pesquisa diminui
  - Portanto, a utilidade da pesquisa é reduzida
  - Portanto, limitar a profundidade é menos prejudicial
  - Porém podar é mais complicado…
- História da IA: TDGammon usa pesquisa de profundidade 2 + função de avaliação muito boa + aprendizado de reforço:
- nível de jogo de campeão mundial
- 1º campeão mundial de IA em qualquer jogo!

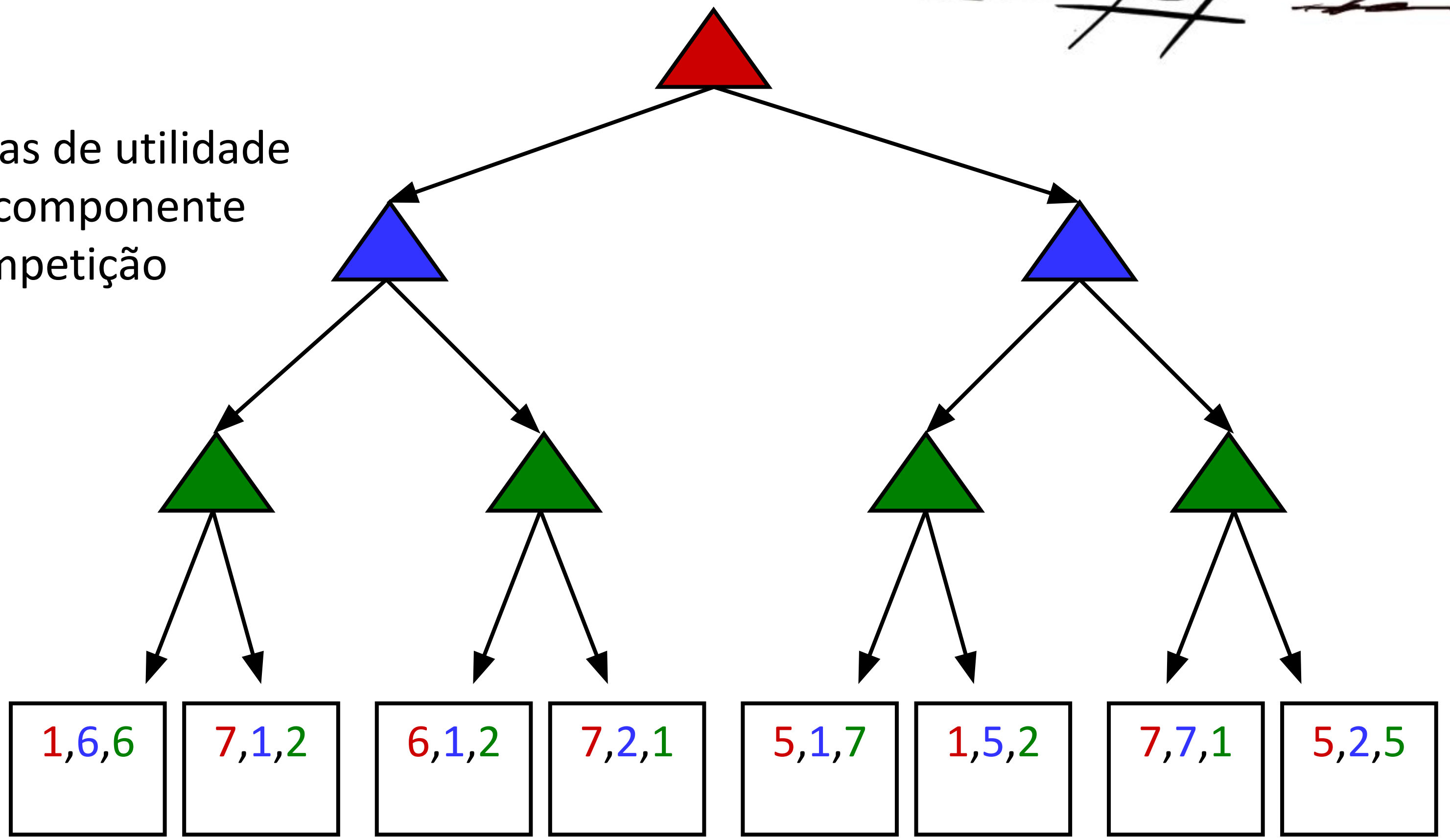
Image: Wikipedia

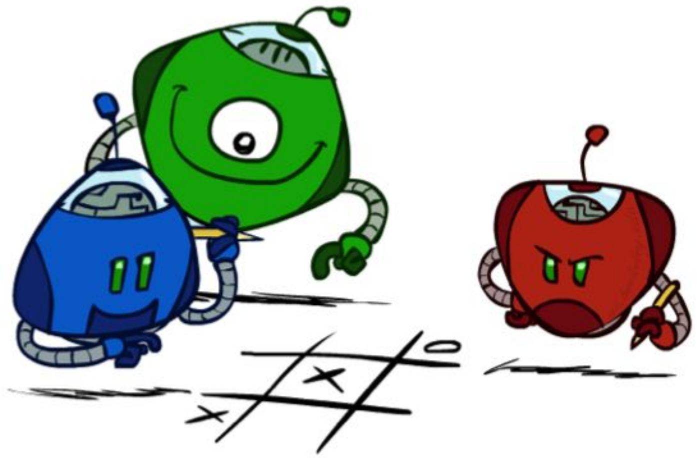
Utilidades com multiagentes



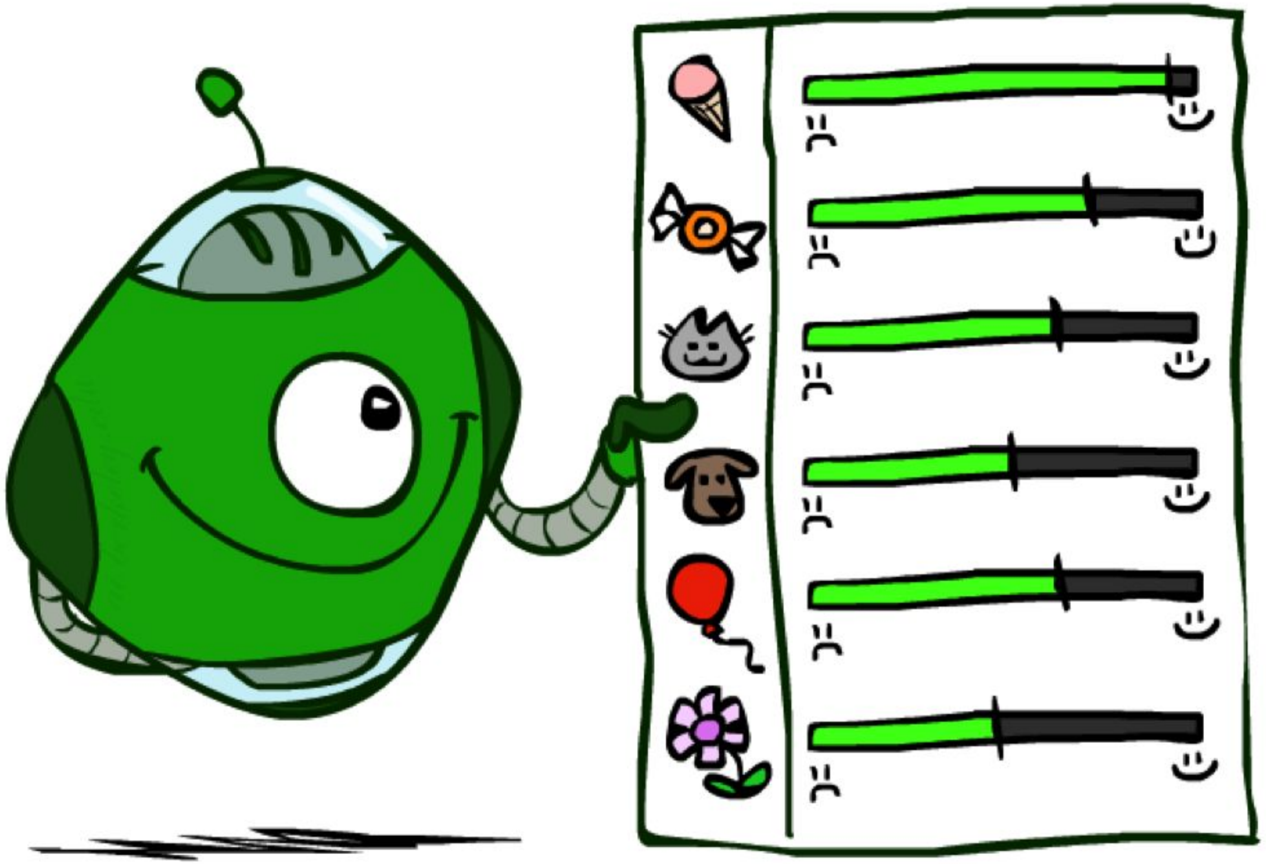
- E se o jogo não for de soma zero ou tiver vários jogadores?

- Generalização do minimax:
  - Terminais têm tuplas de utilidades
  - Os valores dos nós também são tuplas de utilidade
  - Cada jogador maximiza seu próprio componente
  - Pode dar origem a cooperação e competição dinamicamente…





Utilidades



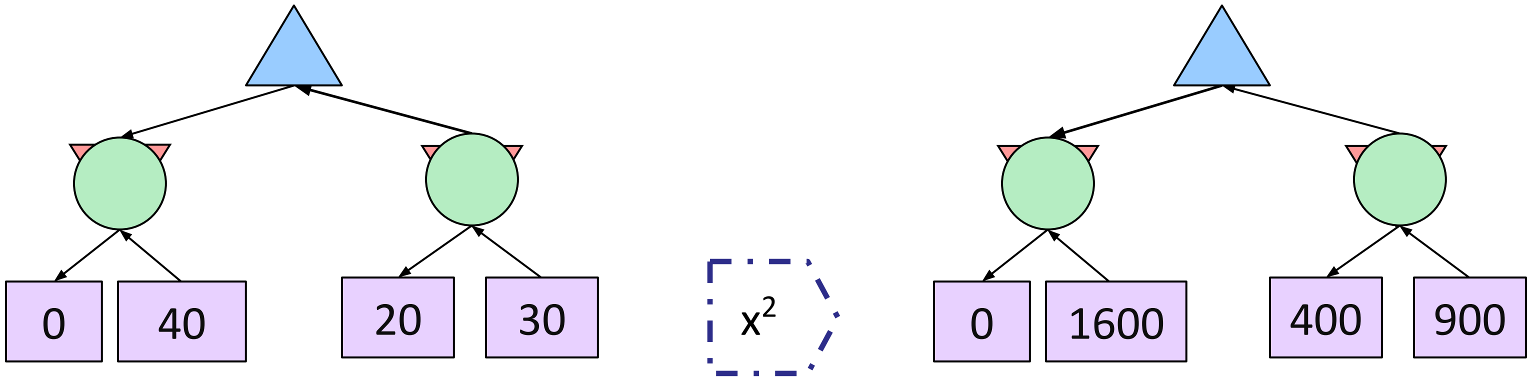
Utilidade máxima esperada

- Quando devemos fazer a média das utilidades ou optar por minimax? Depende da situação.

- Princípio da utilidade máxima esperada:
  - Um agente racional deve escolher a ação que maximiza sua utilidade esperada, dado seu conhecimento

- Questões:
  - De onde vêm as utilidades?
  - Como sabemos que esses utilidades existem?
  - Como sabemos que a média faz sentido?
  - E se o nosso comportamento (preferências) não puder ser descrito pelos utilidades?

Quais utilidades usar?



- Para o raciocínio minimax, alterar a escala da função de avaliação não importa
  - Queremos apenas melhores estados com avaliações mais altas (obter a ordem certa)
  - Chamamos isso de insensibilidade às transformações monotônicas

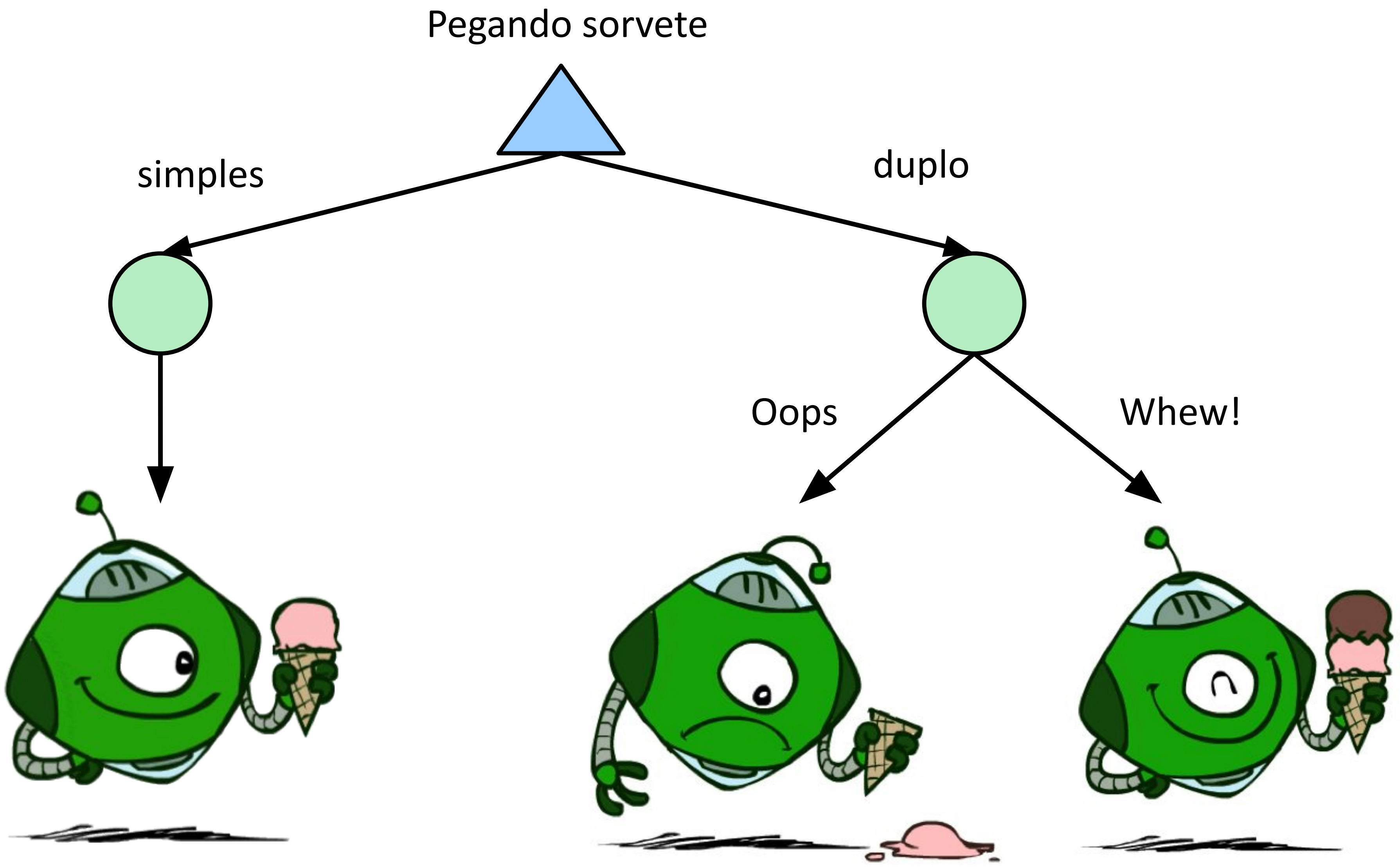
Utilidades

- Utilitários são funções de resultados (estados do mundo) a números reais que descrevem as preferências de um agente

- De onde vêm as utilidades?
  - Em um jogo, pode ser simples (+1/-1)
  - Utilidades resumem o objetivo do agentes
  - Teorema: quaisquer preferências "racionais" podem ser resumidas como uma função de utilidade

- Codificamos as utilidades e deixamos que o comportamento emergir
  - Por que não deixamos os agentes escolherem utilidade?
  - Porque não codificamos o comportamento?

Utilidades: Resultados incertos



Preferências

- Um agente deve ter preferências entre:
  - Prêmios: A, B, etc.
  - Loterias: situações com prêmios incertos

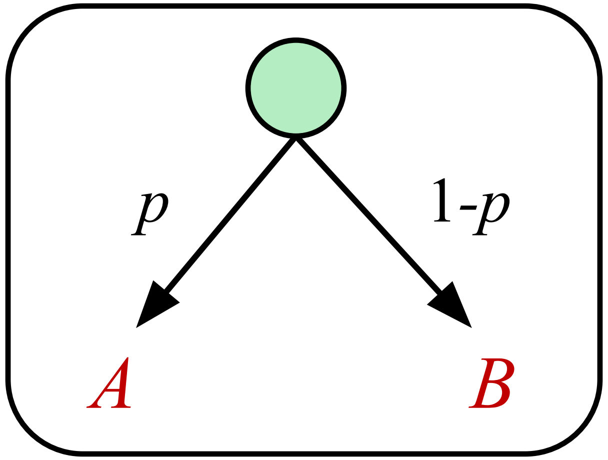
$$L = [p, A;\ (1-p), B]$$

- Notação:
  - Preferência: $A \succ B$
  - Indiferença: $A \sim B$

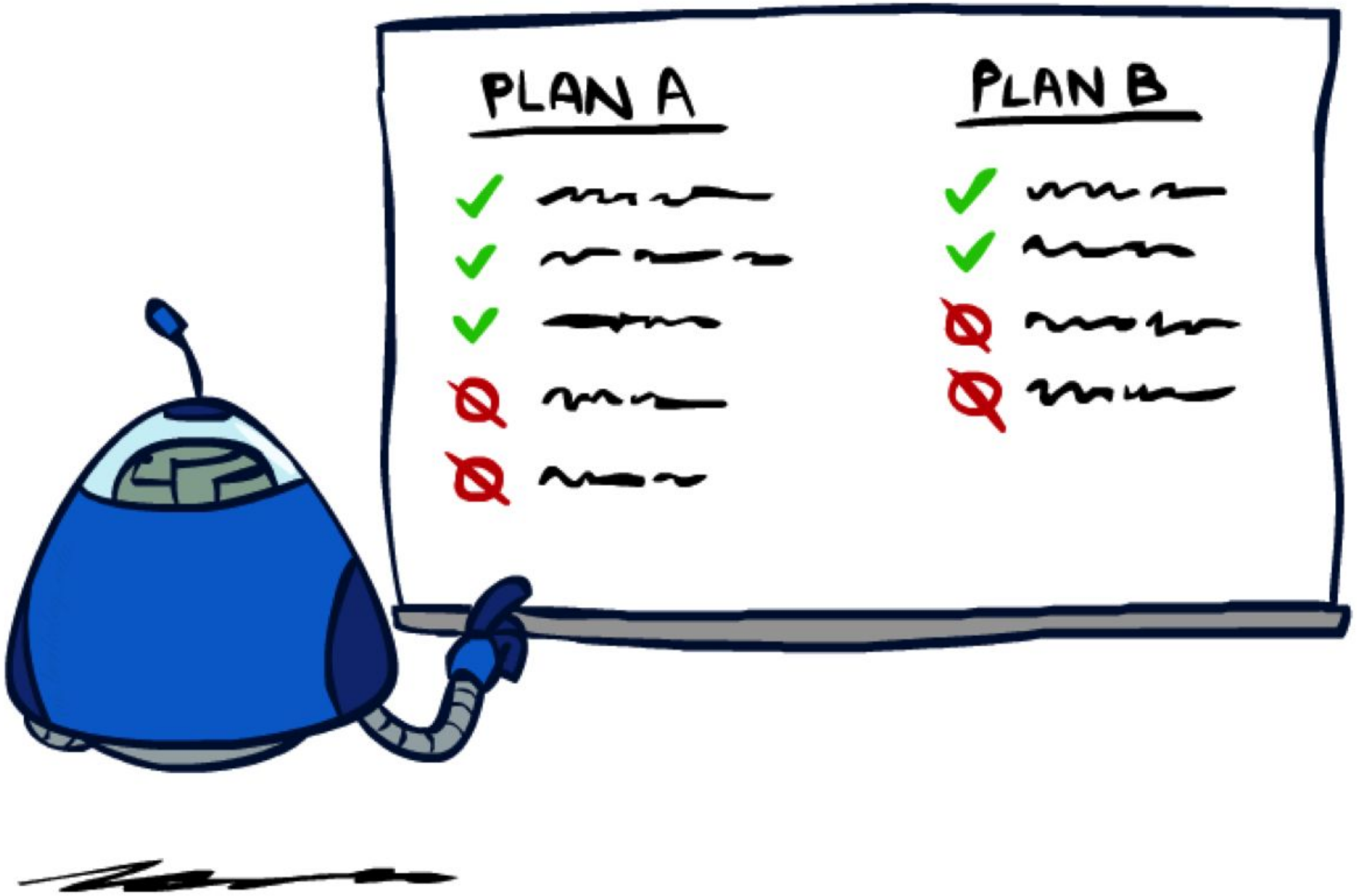
Um prêmio

A

Uma loteria



Racionalidade



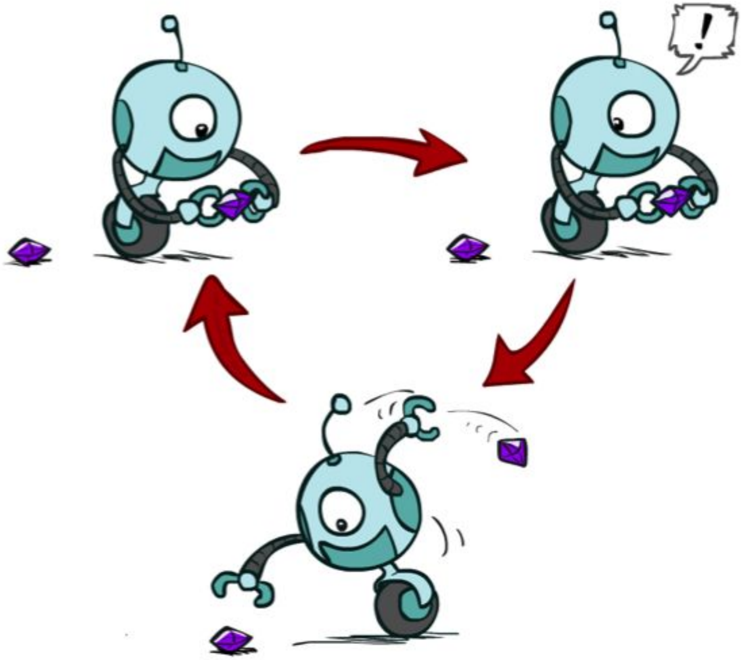
Preferências Racionais

- Queremos algumas restrições nas preferências antes de chamá-las de racionais, tais como:

Axioma de Transitividade: $(A \succ B) \wedge (B \succ C) \Rightarrow (A \succ C)$

Por exemplo: um agente com preferências intransitivas pode ser induzido a dar todo o seu dinheiro

- Se B > C, então um agente com C pagaria (digamos) 1 centavo para obter B
- Se A > B, então um agente com B pagaria (digamos) 1 centavo para obter A
- Se C > A, então um agente com A pagaria (digamos) 1 centavo para obter C



Preferências Racionais

Os Axiomas da Racionalidade

Orderability

$$(A \succ B) \vee (B \succ A) \vee (A \sim B)$$

Transitivity

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

Continuity

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

Substitutability

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

Monotonicity

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

Teorema: Preferências racionais implicam comportamento descritível como maximização da utilidade esperada

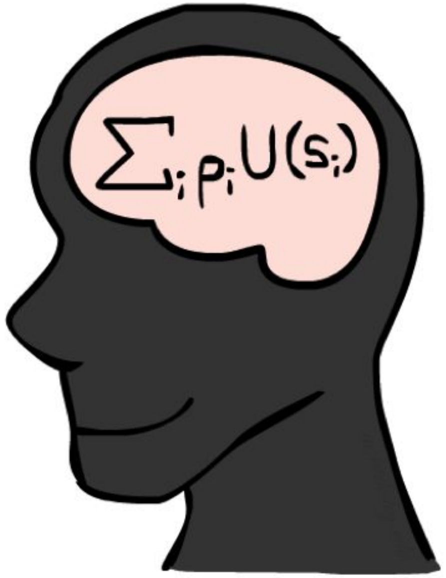
Princípio MUE

- Teorema [Ramsey, 1931; von Neumann & Morgenstern, 1944]

Dadas as preferências que satisfaçam essas restrições, existe uma função de valor real U tal que:

$$U(A) \geq U(B) \;\Leftrightarrow\; A \succeq B$$

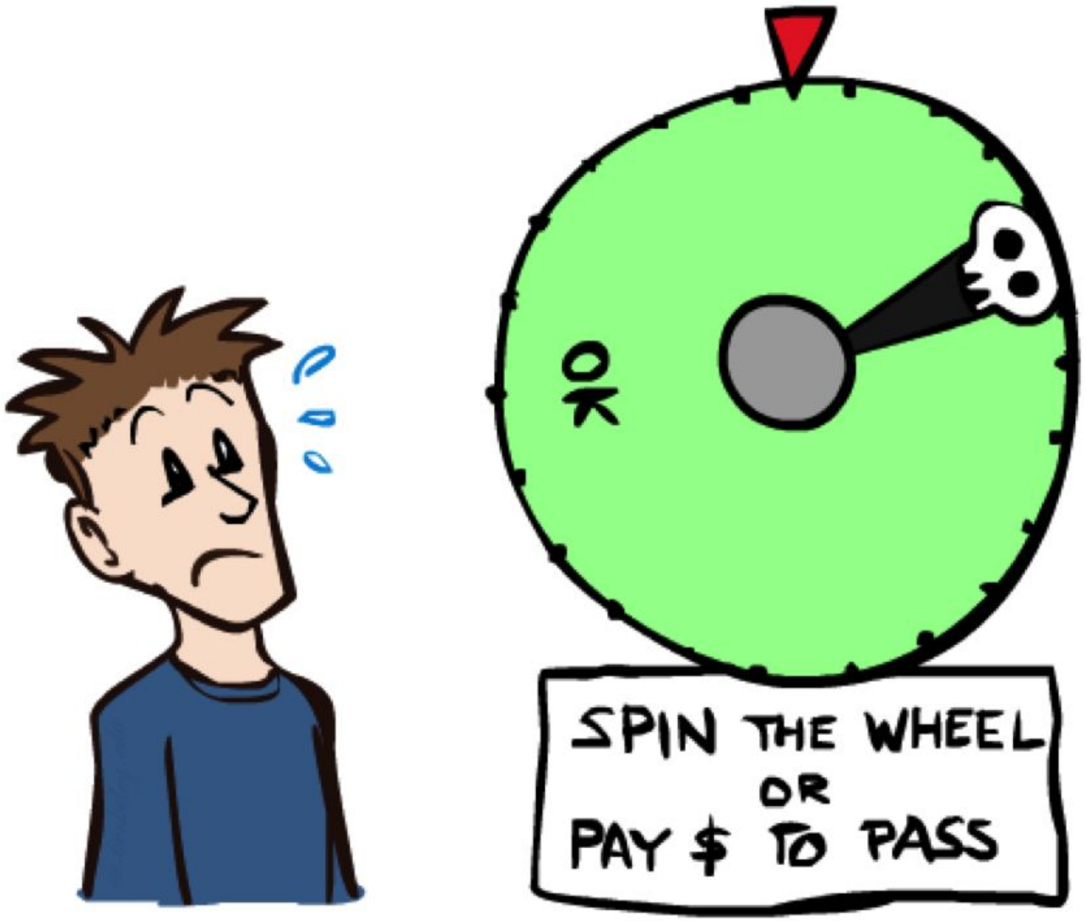
$$U([p_1, S_1; \;\ldots\; ; \; p_n, S_n]) = \sum_i p_i U(S_i)$$



- Ou seja os valores atribuídos por U preservam as preferências de prêmios e loterias!

- Princípio da máxima utilidade esperada (MUE):
  - Escolha a ação que maximiza a utilidade esperada
  - Nota: um agente pode ser totalmente racional (consistente com MUE) sem nunca representar ou manipular utilidades e probabilidades
  - Por exemplo, uma tabela de pesquisa para jogo da velha perfeito; um aspirador de pó reflexivo

Utilidades Humanas

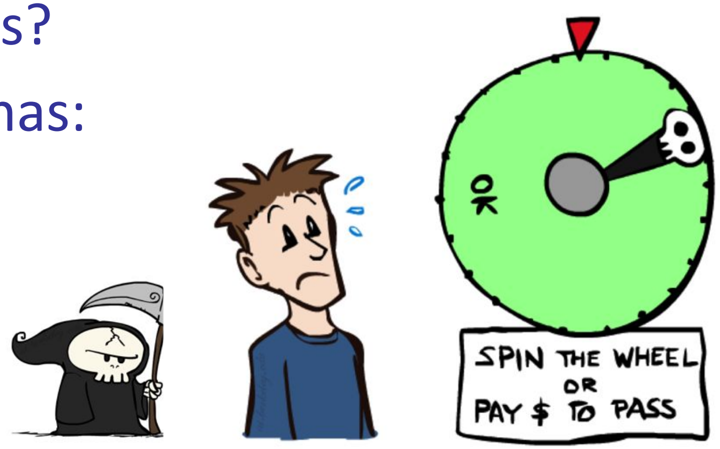


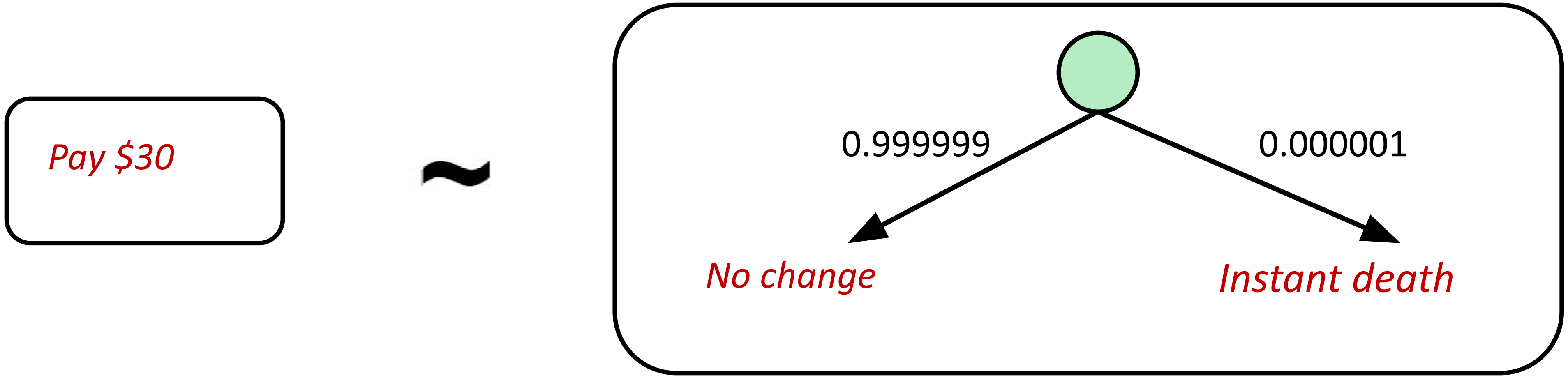
Escalas de Utilidade

- Utilidades normalizadas: $u_+ = 1.0$, $u_- = 0.0$
- Micromortes: chance de um milionésimo de morte, útil para pagar para reduzir os riscos do produto, etc.
- QALYs: anos de vida ajustados pela qualidade, úteis para decisões médicas envolvendo risco substancial

Utilidades Humanas

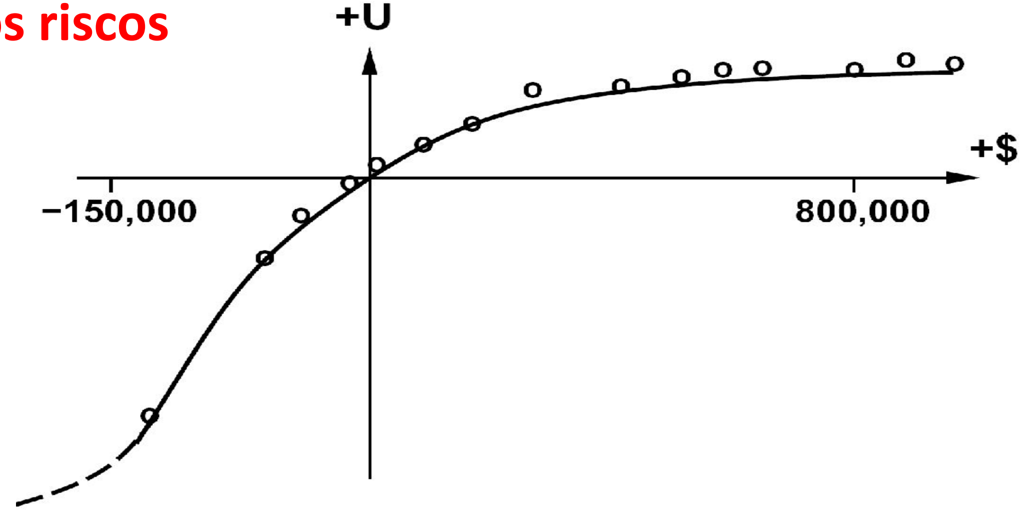
- Os utilitários mapeiam estados para números reais. Quais números?
- Abordagem padrão para avaliação (elicitação) de utilidades humanas:
  - Compare um prêmio A com uma loteria padrão L_p entre:
    - "Melhor prêmio possível" u_+ com probabilidade p
    - "Pior catástrofe possível" u_- com probabilidade 1-p
  - Ajuste a probabilidade da loteria p até a indiferença: A ~ L_p
  - O p resultante é um utilidade em [0,1]





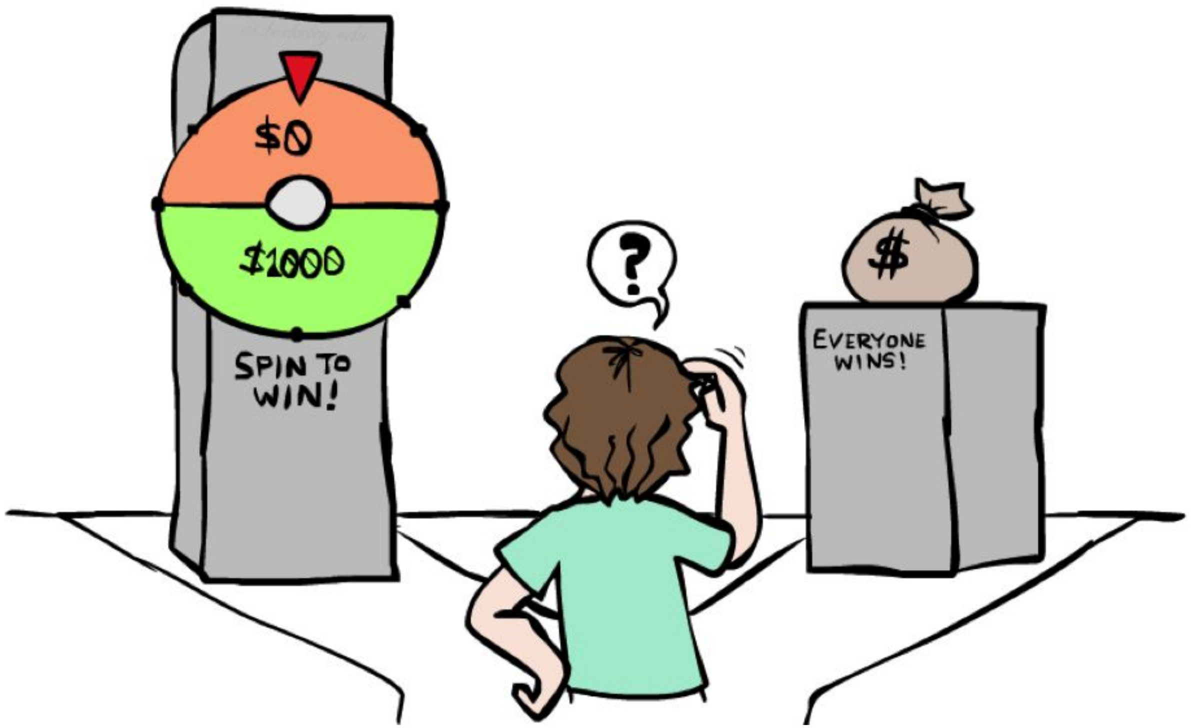
Dinheiro

- O dinheiro **não** se comporta como uma função de utilidade, mas podemos falar sobre a utilidade de ter dinheiro (ou estar em dívida)
- Dada a loteria L = [p, $X; (1-p), $Y]
  - O **valor monetário esperado** EMV(L) is p*X + (1-p)*Y
  - U(L) = p*U($X) + (1-p)*U($Y)
  - Tipicamente, U(L) < U(EMV(L))
  - Nesse sentido, as pessoas são **avessas ao risco**
  - Quando estão afundadas em dívidas, as pessoas estão **propensas aos riscos**



Exemplo: seguro

- Considere a loteria [0.5, $1000; 0.5, $0]
  - Qual é o valor monetário esperado? ($500)
  - Qual é a sua certeza equivalente?
    - Valor monetário aceitável em vez de loteria
    - $ 400 para a maioria das pessoas
  - A diferença de $ 100 é o prêmio do seguro
    - Existe uma indústria de seguros porque as pessoas pagam para reduzir seus riscos
    - Se todos fossem neutros ao risco, nenhum seguro seria necessário!
  - É ganha-ganha: você prefere os $ 400 e a seguradora prefere a loteria (sua curva de utilidade é plana e eles têm muitas loterias)



Exemplo: Racionalidade Humana?

- Exemplo famoso de Allais (1953)

  - A: [0.8, \$4k;   0.2, \$0]
  - B: [1.0, \$3k;   0.0, \$0]

  - C: [0.2, \$4k;   0.8, \$0]
  - D: [0.25, \$3k;   0.75, \$0]

- A maioria das pessoas prefere B > A, C > D

- Porém se U(\$0) = 0, então
  - B > A $\Rightarrow$ U(\$3k) > 0.8 U(\$4k)
  - C > D $\Rightarrow$ 0.8 U(\$4k) > U(\$3k)