Inteligência Artificial

Aula 6
Profª Bianca Zadrozny
http://www.ic.uff.br/~bianca/ia-pos

Tomada de decisões simples

Capítulo 16 – Russell & Norvig
Seções 16.1 a 16.5

Combinação de crenças e desejos sob incerteza

- "*Para julgar o que se deve fazer para obter um bem ou evitar um mal, é necessário considerar não apenas o bem ou o mal em si, mas também a probabilidade de ele acontece, e ainda visualizar geometricamente a proporção que esses itens têm em conjunto*"
  (Port-Royal Logic, Arnauld, 1662)

Combinação de crenças e desejos sob incerteza

- Função de Utilidade: Atribui um número para expressar a **desejabilidade** de um estado para o agente.
  - As utilidades são combinadas com probabilidades dos estados para tomada de decisão.
  - Notação: U(S)

Utilidade Esperada

- Ação tem i resultados possíveis $Resultado_i(A)$
- Probabilidade de cada resultado: $P(Resultado_i(A)|Fazer(A),E)$, onde E resume a evidência disponível do agente sobre o mundo.
- A utilidade esperada da ação A dada a evidência atual E é:
  $EU(A|E) = \sum_i P(Resultado_i(A)|Fazer(A),E)U(Resultado_i(A))$

Princípio de Utilidade Máxima Esperada (UME)

- Um agente racional deve escolher uma ação que **maximize** a utilidade esperada do agente.
- Pode ser usado diretamente para tomada de decisões simples (i.e., escolher uma única ação).
- Para tomada de decisões complexas (sequências de ações), precisamos enumerar todas as sequências e escolher a sequencia com máxima utilidade esperada.
  - Próximo capítulo, técnicas para fazer isso de forma eficiente.

A Base da Teoria da Utilidade

- Por que maximizar a utilidade média é tão especial?
- Por que não tentar minimizar a pior perda possível?
- Como sabemos que existe uma função de utilidade que captura as preferências do agente?
- Podemos demonstrar que o princípio de UME pode ser derivado a partir de **restrições** sobre as **preferências** que um agente racional pode ter.

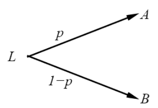
Preferências

- Notação
  - $A \succ B$: 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

onde A e B são estados.

Loterias

- Loteria: distribuição de probabilidade sobre um conjunto de resultados (prêmios)



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

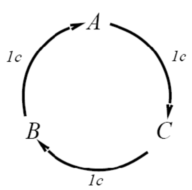
- Cada resultado de uma loteria pode ser um estado ou outra loteria.

Restrições para as Preferências Racionais: Axiomas da Utilidade

- Ordenabilidade: $(A \succ B) \vee (B \succ A) \vee (A \sim B)$
- Transitividade: $(A \succ B) \wedge (B \succ C) \implies (A \succ C)$
- Continuidade: $A \succ B \succ C \implies \exists p [p, A;\ 1-p, C] \sim B$
- Substitutibilidade:
  $$A \sim B \implies [p, A;\ 1-p, C] \sim [p, B; 1-p, C]$$
- Monotonicidade:
  $$A \succ B \implies (p \geq q \Leftrightarrow [p, A;\ 1-p, B] \succeq [q, A;\ 1-q, B])$$

Violar restrições leva a irracionalidade

- Exemplo: se um agente viola a restrição de transitividade, seria possível induzí-lo a gastar todo o seu dinheiro.
  - Se B$\succ$C, então se ele tem C daria 1 centavo (por exemplo), para trocar C por B.
  - Se A$\succ$B, então ele daria 1 centavo (por exemplo), para trocar B por A.
  - Se C$\succ$A, então ele daria 1 centavo (por exemplo), para trocar A por C.
    - Volta ao estado inicial com 3 centavos a menos.



Princípio de Utilidade e Princípio de Utilidade Máxima Esperada

- Teorema (Ramsey, 1931; von Neumann e Morgenstern, 1944): Dadas preferências satisfazendo as restrições, então existe uma função de valores reais U que opera sobre estados tal que

$$U(A) \geq U(B) \quad \Leftrightarrow \quad A \succeq B$$
$$U([p_1, S_1;\ \ldots\ ;\ p_n, S_n]) \quad = \quad \sum_i p_i U(S_i)$$

Princípio de Utilidade e Princípio de Utilidade Máxima Esperada

- Teorema (Ramsey, 1931; von Neumann e Morgenstern, 1944): Dadas preferências satisfazendo as restrições, então existe uma função de valores reais $\boldsymbol{U}$ que opera sobre estados tal que

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

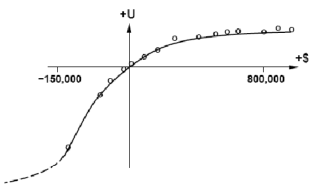
$$U([p_1, S_1; \ \ldots \ ; \ p_n, S_n]) \quad = \quad \sum_i p_i U(S_i)$$

Função de Utilidade

- Sabemos que existe uma função de utilidade mapeando estados a números reais, mas como encontrá-la?
- Uma maneira é usar uma “Loteria padrão”
  - Comparar o estado S com uma loteria L_p que tem probabilidade p de retornar a utilidade máxima possível $\overline{u}$ e probabilidade 1 - p de retornar a utilidade mínima possível $\underline{u}$.
  - Ajustar p até que $S \sim L_p$
  - Aí teremos $U(S) = p(\overline{u} - \underline{u})$

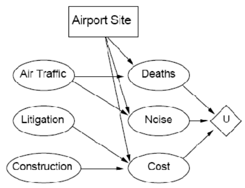
Utilidade do Dinheiro

- Não é uma função linear: conforme aumenta a quantidade de dinheiro, a taxa de crescimento da utilidade diminui.
  - Pessoas ficam mais avessas a risco conforme aumentam os valores
    - Preferimos ganhar R$500.000 to que ter a chance de ganhar de R$1.000.000 com probabilidade 0.6.
  - Do lado da dívida a tendência é ser favorável ao risco.



Redes de Decisão

- Adicionar **nós de ação** (retângulos) e **nós de utilidade** (losangos) a redes bayesianas.
  - Nós de ação têm que ser pais e funcionam como evidências (variáveis observadas).
  - Nós de utilidade têm como pais todos os nós que afetam a utilidade do agente.



Redes de Decisão

- Algoritmo:
  1. Para cada valor do nó de ação:
     - Calcular probabilidades dos pais do nó de utilidade dadas a ação e as outras variáveis de evidência (usar um algoritmo de inferência em redes bayesianas).
     - Calcular valor esperado da utilidade usando as probabilidades obtidas.
  2. Retornar ação com maior valor esperado de utilidade.

Exemplo: Redes de Decisão

Umbrella = leave

$$\mathrm{EU}(\mathrm{leave}) = \sum_w P(w) U(\mathrm{leave}, w)$$

$$= 0.7 \cdot 100 + 0.3 \cdot 0 = 70$$

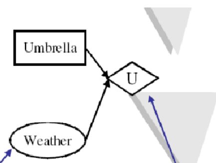
Umbrella = take

$$\mathrm{EU}(\mathrm{take}) = \sum_w P(w) U(\mathrm{take}, w)$$

$$= 0.7 \cdot 20 + 0.3 \cdot 70 = 35$$

Optimal = leave

$$\mathrm{MEU}(\varnothing) = \max_a \mathrm{EU}(a) = 70$$



W	P(W)
sun	0.7
rain	0.3

A	W	U(A,W)
leave	sun	100
leave	rain	0
take	sun	20
take	rain	70

Slide 7
Hal Daume III

CS 5300: Decisions + Information

Exemplo: Redes de Decisão

Umbrella = leave

$$\mathrm{EU}(\mathrm{leave}|\mathrm{bad}) = \sum_w P(w|\mathrm{bad})U(\mathrm{bad}, w)$$

$$= 0.34 \cdot 100 + 0.66 \cdot 0 = 34$$

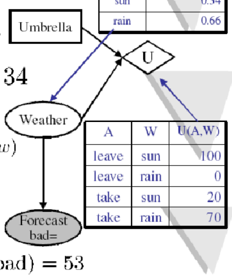
Umbrella = take

$$\mathrm{EU}(\mathrm{take}|\mathrm{bad}) = \sum_w P(w|\mathrm{bad})U(\mathrm{take}, w)$$

$$= 0.34 \cdot 20 + 0.66 \cdot 70 = 53$$

Optimal = take

$$\mathrm{MEU}(F = \mathrm{bad}) = \max_a \mathrm{EU}(a|\mathrm{bad}) = 53$$



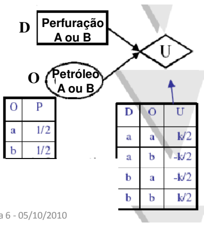
W	P(W|F=bad)
sun	0.34
rain	0.66

A	W	U(A,W)
leave	sun	100
leave	rain	0
take	sun	20
take	rain	70

Slide 8
Hal Daume III
CS 5300: Decisions + Information

Valor da Informação

- Calcular valor de adquirir cada possível evidência
  - Pode ser feito diretamente a partir da rede de decisão
- Exemplo:



O	P
a	1/2
b	1/2

D	O	U
a	a	k/2
a	b	-k/2
b	a	-k/2
b	b	k/2

Exemplo: Valor da Informação

- Exemplo: comprar direito de perfuração de petróleo
  - Duas regiões A e B, uma tem petróleo, valendo k.
  - Probabilidade a priori em cada região é 0.5
  - Preço atual de cada região é k/2
  - UME = 0 (qualquer ação maximiza essa utilidade)
- Solução: calcular valor da informação = ganho esperado de UME para cada evidência obtida.
  - Suponha que podemos verificar com um teste O se a região A contém ou não óleo. Quando valeria à pena pagar por esse teste?
    - Depois da verificação o ganho esperado em utilidade é k/2.
    - Logo VPI(O) = k/2 (VPI = valor da informação perfeita)

Valor da Informação

- Valor esperado da ação de utilidade máxima dada a evidência atual

$$\mathrm{MEU}(e) = \max_a \sum_s P(s|e)\, U(s,a)$$

- Valor esperado da ação de utilidade máxima da a evidência atual + nova evidência

$$\mathrm{MEU}(e, e') = \max_a \sum_s P(s|e, e')\, U(s,a)$$

- Como E' é desconhecida, tenho que calcular ganho esperado considerando todos os valores E'=e'

$$\mathrm{VPI}_e(E') = \sum_{e'} P(e'|e)\left(\mathrm{MEU}(e, e') - \mathrm{MEU}(e)\right)$$

THE UNIVERSITY OF UTAH

Hal Daumé III (hal@cs.utah.edu)

VPI Example

MEU with no evidence

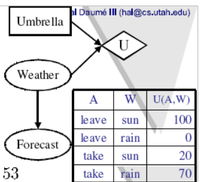
$$\mathrm{MEU}(\varnothing) = \max_a \mathrm{EU}(a) = 70$$

MEU if forecast is bad

$$\mathrm{MEU}(F = \mathrm{bad}) = \max_a \mathrm{EU}(a|\mathrm{bad}) = 53$$

MEU if forecast is good

$$\mathrm{MEU}(F = \mathrm{good}) = \max_a \mathrm{EU}(a|\mathrm{good}) = 95$$



A	W	U(A,W)
leave	sun	100
leave	rain	0
take	sun	20
take	rain	70

Forecast distribution

F	P(F)
good	0.59
bad	0.41

$$0.59 \cdot (95 - 70) + 0.41 \cdot (53 - 70)$$

$$0.59 \cdot (+25) + 0.41 \cdot (-17) = +22$$

$$\mathrm{VPI}_e(E') = \sum_{e'} P(e'|e)\left(\mathrm{MEU}(e, e') - \mathrm{MEU}(e)\right)$$

Slide 11
CS 5300: Decisions + Information