Inteligência Artificial

Sistemas Especialistas
Tratamento de Incertezas

Prof. Paulo Martins Engel

Conceitos de Incerteza

- A incerteza pode ser considerada como a falta de informação adequada para tomar uma decisão.
- Lidar com incertezas é uma importante capacidade dos especialistas humanos.
- **Conhecimento incerto**: freqüentemente o perito possui somente conhecimento heurístico sobre alguns aspectos do domínio.

  Ex: ele sabe que um certo conjunto de fatos *provavelmente leva* a uma certa conclusão.
- **Dado incerto**: mesmo quando se está seguro do conhecimento sobre certo aspecto do domínio, pode haver *incerteza nos dados* que descrevem o ambiente externo.

  Ex.: tentando inferir a causa de um fenômeno observado, pode-se depender de resultados *questionáveis* de testes.
- **Informação incompleta**: freqüentemente é necessário tomar decisões com base em *informação incompleta*.

  Ex.: o médico pode ter que iniciar o tratamento de um paciente *antes* mesmo de ter *todos* os resultados dos exames de laboratório que ele gostaria.

Exemplo: conhecimento incerto

- O problema de o pedal do freio estar baixo ou ir todo ao fundo tem como (*algumas*) causas possíveis:
  1. As lonas do freio estão gastas (FC = 0,8)
  2. O nível do líquido de freio está baixo (FC = 0,6)
  3. O cilindro está defeituoso (FC = 0,5)
- A incerteza neste caso é tipicamente tratada por uma abordagem numérica, baseada em probabilidades, quer na forma subjetiva, quer na forma tradicional (baseada em experimentos de freqüência).
- Obs.: o pedal pode estar baixo porque a mola quebrou!

•Regra *causal* consistente:

(pouco útil para diagnóstico)

se
as lonas do freio estão gastas ou
o nível do líquido de freio está baixo
ou o cilindro está defeituoso
então
o pedal do freio está baixo

•Regra *heurística* (não é consistente):

(útil para diagnóstico)

se
o pedal do freio está baixo
então
as lonas do freio estão gastas ou
o nível do líquido de freio está baixo
ou o cilindro está defeituoso

Raciocínio abdutivo

- Normalmente as regras úteis são de natureza heurística, *inconsistentes*:

  se
  o motor não pega e
  as luzes não acendem
  então
  o problema é bateria ou cabo
- Seria possível, apesar de improvável, que a bateria ou os cabos estivessem em ordem, mas que o carro simplesmente tivesse o motor de partida com problemas e os faróis queimados.
- Raciocínio abdutivo: A partir de $\mathbf{P} \rightarrow \mathbf{Q}$ e $\mathbf{Q}$ é possível inferir $\mathbf{P}$.
- A abdução é uma regra inconsistente, a conclusão não é necessariamente verdadeira quando as premissas são verdadeiras.
- As regras heurísticas podem expressar uma política explícita para a crença.
- P. ex., a regra $\mathbf{P} \rightarrow \mathbf{Q}$ (0,9) expressa a crença: "se você acredita que **P** seja verdadeiro, então você acredita que **Q** irá ocorrer em 90% dos casos.

Dados incertos

- Pode-se usar medidas de certeza para refletir a nossa crença na *qualidade* dos dados.
- No exemplo anterior, a proposição: **as luzes acendem (0,2)** pode indicar o resultado da verificação da intensidade com que acendem os faróis: acendem, mas são fracos e pouco visíveis.
- Crenças e dados imperfeitos podem ser propagados através das regras a fim de gerar conclusões.
- Os formalismos baseados em lógica podem ser estendidos para tratar *abdução*.
- O raciocínio não-monotônico aborda o problema de crenças que podem ser modificadas.
- Ele trata a incerteza fazendo as suposições mais razoáveis com base na informação incerta, procedendo como se estas informações fossem verdadeiras.
- Se mais tarde a crença mudar, é necessário reexaminar as conclusões derivadas daquela crença.
- Os *algoritmos de manutenção da verdade* são empregados para manter a base de conhecimento consistente.

Raciocínio não-monotônico

- O raciocínio convencional (lógica de predicados) é baseado em três suposições:
- *Os predicados devem ser suficientes em relação ao domínio de aplicação*:
  Toda informação necessária para resolver o problema deve ser representada.
- *A base de informação deve ser consistente*:
  As porções de conhecimento não se contradizem.
- *Através do uso de regras de inferência, a informação conhecida cresce monotonamente*.
- Se qualquer destas suposições não for satisfeita, a abordagem baseada em lógica convencional não funcionará.
- Sistemas não-monotônicos lidam com estas questões.
- Questão importante sobre escassez de conhecimento: como tratar a falta de conhecimento sobre um predicado *p*?
- A falta de conhecimento significa que *não temos certeza que p é verdadeiro* ou significa que *temos certeza que não p é verdadeiro*?
- PROLOG usa *hipótese de mundo fechado*:

Tudo que não pode ser provado verdadeiro é falso!

Raciocínio não-monotônico

- Uma abordagem para a falta de conhecimento é fazer suposições explícitas sobre a verdade.
- Raciocina-se com base como o mundo normalmente funciona:
- A maioria dos pássaros voa.
- A lógica pode ser estendida por *operadores modais*: **a_menos_que**

  p(x) a_menos_que q(x) → r(x)

  Podemos inferir **r(x)** se **p(x)** for verdadeiro e não acreditamos que **q(x)** seja verdadeiro,

  Se, mais tarde, constatamos que **q(x)** é verdadeiro, então **r(x)** deve ser *retratado*.
- O operador modal **a_menos_que** trata da *crença* em vez da verdade: “ou desconhecido ou presumidamente falso”.

Problemas Gerais

- Suponha que x é A com grau de confiança x_1 e que a Base de Conhecimento contenha a regra:
- R_1: Se x é A Então y é B (FC = r_1)

1. Como extrair do especialista valores consistentes para r_1?
2. Como computar ou obter de outro modo o valor de x_1?
3. Como r_1 e x_1 devem ser representados: um escalar, um intervalo, uma expressão lingüística, uma distribuição,…
4. No caso mais geral, quando a premissa é composta por múltiplas cláusulas, como agregar o grau de certeza total da premissa?
5. Como definir a função que agrega os graus de certeza da premissa com o da regra?
6. Se múltiplas regras determinam a mesma conclusão, qual a função que determina o grau de certeza final da conclusão?

Raciocínio baseado em Probabilidades Condicionais

- A inferência probabilística usa a informação disponível sobre os valores de algumas variáveis para obter a probabilidade para valores de outras variáveis.
- A probabilidade condicional $P(\mathrm{B}|\mathrm{A})$ é a probabilidade de B ocorrer dado que A tenha ocorrido. Ela é calculada a partir da probabilidade conjunta $P(\mathrm{A,B})$:

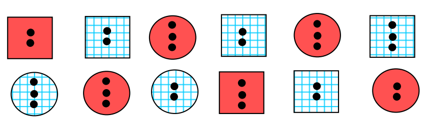
$$P(\mathrm{B}|\mathrm{A}) = \frac{P(\mathrm{A,B})}{P(\mathrm{A})}$$

- Esta probabilidade pode ser interpretada como um *fator de confiança* que se pode inferir a partir dos dados na regra: $P(\mathrm{A}) \rightarrow P(\mathrm{B})$
- $P(\mathrm{A})$ e $P(\mathrm{B})$ são as *probabilidades marginais* de A e B, respectivamente. Elas são também chamadas de probabilidades a priori destes valores de variáveis.

O Teorema de Bayes

- O Teorema de Bayes mostra uma maneira de se calcular a probabilidade de um evento em particular, a partir de um conjunto de observações.
- O evento pode ser um determinado valor de um atributo especial, que seria o rótulo de uma classe *i*, ou hipótese H_i.
- Inicialmente, se determina a probabilidade a priori de cada hipótese H_k, $P(\mathrm{H}_k)$ presente no domínio. (H_1, H_2, ..., H_i, ..., H_n)
- A seguir, a partir de conjuntos de amostras contendo objetos pertencentes a cada uma das hipóteses possíveis, se determinam as probabilidades com que ocorrem os atributos observáveis (E_1, E_2, ..., E_j, ..., E_l), as $P(\mathrm{E}_j|\mathrm{H}_k)$.
- O TB afirma que, para um objeto do domínio no qual ocorra p. ex. E_j, a probabilidade de que este objeto satisfaça a hipótese H_i é dada por $P(\mathrm{H}_i|\mathrm{E}_j)$:

$$P(\mathrm{H}_i|\mathrm{E}_j) = \frac{P(\mathrm{E}_j|\mathrm{H}_i)\ .\ P(\mathrm{H}_i)}{\sum_{k=1}^{n} P(\mathrm{E}_j|\mathrm{H}_k)\ .\ P(\mathrm{H}_k)}$$



- Determinar pelo Teorema de Bayes, a probabilidade de um objeto que tenha dois furos ser azul, $P(\mathrm{A}|\mathrm{D})$.

$$P(A \mid D) = \frac{P(D \mid A)P(A)}{P(D \mid A)P(A) + P(D \mid V)P(V)}$$

Probabilidades a priori, $P(\mathrm{A})$ e $P(\mathrm{V})$:

$$P(\mathrm{A}) = 0{,}5 \qquad P(\mathrm{V}) = 0{,}5$$

Verossimilhanças (em relação a D) das classes (A e V), $P(\mathrm{D}|\mathrm{A})$ e $P(\mathrm{D}|\mathrm{V})$:

$$P(\mathrm{D}|\mathrm{A}) = 4/6 = 0{,}667 \qquad P(\mathrm{D}|\mathrm{V}) = 2/6 = 0{,}333$$

$$P(A \mid D) = \frac{0{,}667 \times 0{,}5}{0{,}667 \times 0{,}5 + 0{,}333 \times 0{,}5} = 0{,}667$$

Exemplo de SE com raciocínio probabilístico

- O PROSPECTOR é um exemplo de SE que utiliza raciocínio probabilístico para localizar depósitos de minérios:
- Sabe-se a probabilidade a priori de se encontrar cada minério.
- Sabe-se a probabilidade de que se existir um certo minério, determinadas características físicas serão encontradas (verossimilhança).
- Então, utiliza-se o TB para calcular a probabilidade de que os vários minérios estejam presentes, a partir das evidências (dados) que foram coletados.

Dificuldades com o método de Bayes

- O uso do TB pode depender de um número grande de variáveis.
- Em medicina, pode-se usar o TB para determinar a probabilidade de uma doença específica dado um conjunto de sintomas.
- Deve-se estimar a probabilidade a priori de um paciente ter uma determinada doença.
- Deve-se estimar a probabilidade condicional dos diversos sintomas em cada (todas!) doença.
- É normalmente impossível obter valores consistentes e completos para todas estas probabilidades para a população em geral.
- Como $P(\mathsf{H}) + P(\neg\mathsf{H}) = 1 \therefore P(\mathsf{H}) = 1 - P(\neg\mathsf{H})$
- Similarmente, $P(\mathsf{H|E}) = 1 - P(\neg\mathsf{H|E})$
- Os médicos que trabalharam no MYCIN eram extremamente relutantes em formular seu conhecimento nesta forma.

Fatores de Certeza (FC)

- Um FC é um valor numérico que expressa quanto se acredita que, com base num conjunto de evidências, pode-se aceitar uma dada conclusão.
- O FC pode variar de 1 (crença total) a –1 (descrença total).
- O FC é uma quantificação subjetiva baseada no julgamento e intuição do perito.
- Exemplo: regra para diagnosticar problema numa lâmpada de cabeceira:

SE
    a lâmpada de cabeceira está ligada na tomada e
    a tomada tem corrente e
    o interruptor está ligado
    (e a lâmpada não acende)

ENTÃO
    há uma evidência sugestiva (0,8) que a lâmpada esteja queimada.

Medidas de Crença

- O fator de certeza depende de medidas de crença e descrença:

$$FC(\mathsf{H,E}) = MC(\mathsf{H, E}) - MD(\mathsf{H, E})$$

$FC(\mathsf{H, E})$ é o *fator de certeza* na hipótese H baseado na evidência E

$MC(\mathsf{H, E})$ é a *medida de crença* em H baseado na evidência E.

$MD(\mathsf{H, E})$ é a *medida de descrença* em H baseado em E.

- Por definição, $0 \le MC \le 1$; $0 \le MD \le 1$.
- Num sistema que usa fatores de certeza, as regras devem ser estruturadas de forma que cada regra ou aumenta a crença numa determinada conclusão, ou aumenta a descrença na conclusão.

Medidas de Crença e Probabilidades

- As medidas de crença e descrença estão relacionadas com probabilidades:

$$MC(\mathsf{H,E}) = \begin{cases} 1 & \text{Se } P(\mathsf{H}) = 1 \\ \dfrac{\max[P(\mathsf{H|E}), P(\mathsf{H})] - P(\mathsf{H})}{\max[1,0] - P(\mathsf{H})} & \text{Caso contrário} \end{cases}$$

$$MD(\mathsf{H,E}) = \begin{cases} 1 & \text{Se } P(\mathsf{H}) = 0 \\ \dfrac{\min[P(\mathsf{H|E}), P(\mathsf{H})] - P(\mathsf{H})}{\min[1,0] - P(\mathsf{H})} & \text{Caso contrário} \end{cases}$$

- Note que ou $MC = 0$ ou $MD = 0$.
- $FC(\mathsf{H,E}) = MC(\mathsf{H, E}) - MD(\mathsf{H, E})$
- Se FC for positivo, o sistema acredita que a hipótese é verdadeira; se for negativo, o sistema acredita que a hipótese é falsa.

Combinando Fatores de Certeza

- Combinando várias regras sobre a **mesma hipótese** (*conclusão*):
  As medidas de crença e descrença de uma hipótese dadas duas regras R_1 e R_2 são calculadas pela seguinte combinação de medidas individuais:

$$MC(H,R_1 \wedge R_2) = \begin{cases} 0 & \text{Se } MD(H,R_1 \wedge R_2) = 1 \\ MC(H,R_1) + MC(H,R_2) \,.\, (1 - MC(H,R_1)) & \text{Caso contrário} \end{cases}$$

$$MD(H,R_1 \wedge R_2) = \begin{cases} 0 & \text{Se } MC(H,R_1 \wedge R_2) = 1 \\ MD(H,R_1) + MD(H,R_2) \,.\, \underbrace{(1 - MD(H,R_1))} & \text{Caso contrário} \end{cases}$$

Pode ser visto como a parcela que ainda não temos certeza, ou que R_1 não garante

- Dadas 4 regras que sugerem a conclusão *C*, ache o *FC* acumulado para *C*:

Regra	*FC*
R_1	0,8
R_2	0,3
R_3	– 0,2
R_4	0,7

Para a regra R_1: $MC = 0{,}8$ e $MD = 0$.

Considerando a regra R_2: $MC = 0{,}8 + 0{,}3\,(1 – 0{,}8) = 0{,}86$
$MD = 0 + 0\,(1 – 0) = 0$

Considerando a regra R_3: $MC = 0{,}86 + 0\,(1 – 0{,}86) = 0{,}86$
$MD = 0 + 0{,}2\,(1 – 0) = 0{,}2$

Considerando a regra R_4: $MC = 0{,}86 + 0{,}7\,(1 – 0{,}86) = 0{,}958$
$MD = 0{,}2 + 0\,(1 – 0{,}2) = 0{,}2$

$FC_{acum} = 0{,}958 – 0{,}2 = 0{,}758$.

Conjunção de conclusões

$$MC(C_1 \wedge C_2, e) = \min[MC(C_1, e) , MC(C_2, e)]$$

$$MD(C_1 \wedge C_2, e) = \min[MD(C_1, e) , MD(C_2, e)]$$

onde C_1 é a conclusão 1, C_2 é a conclusão 2 , e é a evidência disponível

Disjunção de conclusões

$$MC(C_1 \vee C_2, e) = \max[MC(C_1, e) , MC(C_2, e)]$$

$$MD(C_1 \vee C_2, e) = \max[MD(C_1, e) , MD(C_2, e)]$$

Exemplo de conjunção/disjunção de conclusões

Dadas as seguintes conclusões de um diagnóstico de um problema num carro:

C_1 : o problema exige imediata atenção (0,8)

C_2 : há um problema no sistema elétrico (0,6)

C_3 : há um curto-circuito no sistema elétrico (0,4)

C_4 : há um problema no distribuidor (0,2)

- Encontre o *FC* de que há um problema no sistema elétrico (**e**) que exige imediata atenção **e** o problema é um curto-circuito **ou** um problema no distribuidor

$$MC(C_1 \wedge C_2) = \min[0{,}6 , 0{,}8] = 0{,}6$$

$$MC(C_3 \vee C_4) = \max[0{,}4 , 0{,}2] = 0{,}4$$

$$MC(C_1 \wedge C_2 \wedge (C_3 \vee C_4)) = \min[0{,}6 , 0{,}4] = 0{,}4$$