Departamento de Ciência de Computadores - FCUP
Segundo Teste de Inteligência Artificial / Sistemas Inteligentes
(Duração: 50 minutos)

Data: 24 de Maio de 2018

1) Um algoritmo de geração de planos apresenta três momentos de decisão:

(A) escolha da próxima pré-condição a ser resolvida

Todas devem ser resolvidas, portanto a escolha poderia ser aleatória. Porém uma heurística que resolva primeiro pré-condições sem variáveis pode também ser uma opção.

(B) escolha da ação cuja pós-condição resolve a pré-condição escolhida

Uma boa heurística poderia ser a escolha da ação que introduz o menos número de pré-condições. Para alguns tipos de problemas, esta heurística poderia causar um número grande de retrocessos (backtracking), portanto, combinada com esta heurística poderíamos escolher a ação que introduz um número maior de **pós-condições** para que se tenha uma probabilidade maior de resolução das pré-condições.

(C) escolha da pós-condição que vai resolver a pré-condição escolhida

Escolha da pós-condição que resolve a pré-condição com a unificação de uma constante invés de uma outra variável.

Para cada um destes momentos, diga quais seriam as heurísticas mais apropriadas. Justifique sua resposta.

2) Assuma duas árvores de decisão: uma com ramos balanceados e outra com ramos não balanceados. Qual delas teria sido gerada a partir de variáveis com maior entropia? Justifique.
Um valor menor de entropia discrimina melhor instâncias de cada classe e separa-as em grupos distintos. Desta forma, a árvore com ramos **não balanceados** apresenta atributos com entropia menos do que uma árvore com ramos balanceados. Se um nó-atributo não discriminar bem instâncias (entropia mais alta), teremos que colocar novos atributos como sub-árvores deste nó tornando-a balanceada.

3) Considere o conjunto de treino mostrado na Tabela 1 para um problema de classificação binária.

Tabela 1: Tabela para a pergunta 3

Instância	a_1	a_2	a_3	Classe
1	T	T	2.0	$+$
2	T	T	5.0	$+$
3	T	F	7.0	$-$
4	F	T	3.0	$+$
5	F	T	8.0	$-$
6	T	F	2.0	$-$
7	T	F	9.0	$-$
8	F	F	6.0	$+$
9	F	F	4.0	$-$

a) Indique a entropia inicial (não precisa fazer cálculos) dos dados desta tabela, onde a entropia é calculada como:

$$-\frac{p}{p+n}log_2(\frac{p}{p+n}) - \frac{n}{p+n}log_2(\frac{n}{p+n})$$

$$-\frac{4}{9}log_2(\frac{4}{9}) - \frac{5}{9}log_2(\frac{5}{9})$$

A entropia inicial de um conjunto é dada pela variável de classe, antes de considerarmos a informação adicionada por qualquer outra variável do conjunto.

b) Mostre os contadores necessários para o cálculo da entropia das variáveis a_1 e a_2 em relação à variável de classe (Nota: não esqueça que a entropia de uma variável em relação à classe é um somatório das contribuições de cada valor daquela variável em relação à classe).

a1:	+	-	Total
T	2	3	5
F	2	2	4

a2:	+	-	Total
T	3	1	4
F	1	4	5

c) Para a_3, que é um atributo numérico, qual é o melhor ponto (ou pontos) de divisão (split point)?

O melhor ponto de quebra que minimiza a entropia de a3 é entre 6.0 e 7.0. Também aceitei outro intervalo entre 2.0 e 3.0 ou entre 3.0 e 4.0.

4) Dada a rede Bayesiana da Figura 1 e dadas as evidências para as variáveis “bad weather” (B), “traffic jam” (T) and “sirens” (S) qual é a expressão que melhor representa o cálculo da probabilidade da variável “accident” (A): (Justifique.)

(A) $P(A) = P(T \mid R \wedge B \wedge A) \times P(S \mid A) \times P(A \mid B) \times P(R) \times P(B)$

(B) $P(A) = P(A \mid B) \times P(B)$

(C) $P(A) = P(B) \times P(T \mid R \wedge B \wedge A) \times P(S \mid A)$

Todas as variáveis envolvidas (A, B, S e T) estão dentro do Markov Blanket, portanto a resposta mais correta é a (a), que corresponde justamente à expressão desta rede. Não esquecer que como não conhecemos R temos que desenvolver a expressão para R e (not R). Minha intenção com esta pergunta era ter uma rede em que algumas variáveis estariam fora do Markov blanket e perceber se sabem como simplificar a expressão. Não é o caso desta rede.

5) Use a linguagem Prolog para codificar o conhecimento apresentado nas frases seguintes:

(A) Cavalos, vacas e porcos são mamíferos

mam(X) :- cavalo(X); vaca(X); porco(X).

(B) A prole de um cavalo é um cavalo

cavalo(X) :- cavalo(Y), prole(Y,X).

(C) Barba Azul é um cavalo

cavalo(barbazul).

(D) Barba Azul é o pai de Carlos

pai(barbazul,carlos).

(E) Todo mamífero tem pais

pai(X,Y) :- mam(Y).

Também aceitei: tempPais(Y) :- mam(Y).

Mostre como Prolog chegaria à conclusão de que Barba Azul é um mamífero.
A prova sai da consulta ?- mam(barbazul). Esta regra unifica com a primeira que diz mam(X) :- cavalo(X), unificando X com o valor barbazul. A seguir, Prolog tenta provar cavalo(X/barbazul). Como existe um facto que diz cavalo(barbazul), a prova termina aqui e Prolog devolve: true (ou yes).

6) Suponha que um paciente tem um sintoma (S) que pode ser causado por duas doenças diferentes (A e B). É conhecido que uma variação do gene G tem uma associação muito forte na manifestação da doença A. Uma rede Bayesiana e suas tabelas de probabilidade representando este cenário é mostrada na Figura 2. Calcule a probabilidade $P(+g, +a, +b, +s)$.
Basta seguir a expressão da rede: $P(G, A, B, S) = P(S \mid A, B) \times P(A \mid G) \times P(G) \times P(B)$ e retirar os valores das tabelas, onde A=+a, B=+b, S=+s e G=+g.

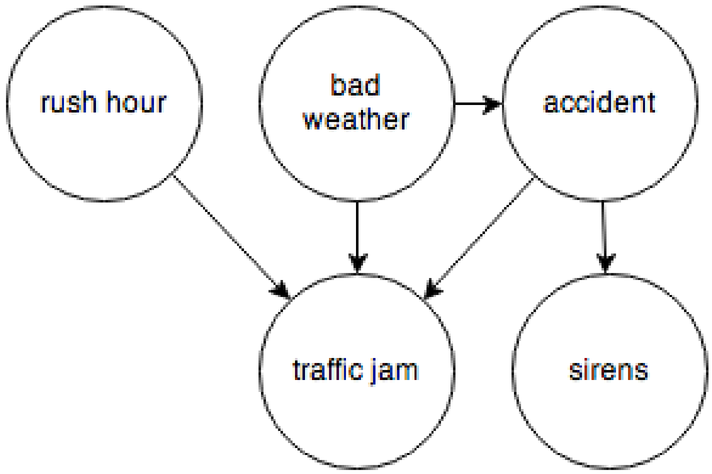
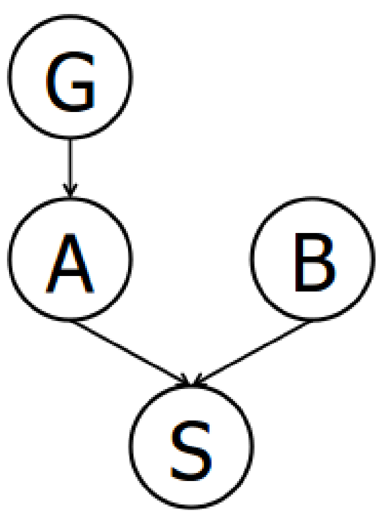


Figura 1: Rede Bayesiana para a pergunta 4

$\mathbb{P}(G)$	
$+g$	0.1
$-g$	0.9

$\mathbb{P}(A \mid G)$		
$+g$	$+a$	1.0
$+g$	$-a$	0.0
$-g$	$+a$	0.1
$-g$	$-a$	0.9



$\mathbb{P}(B)$	
$+b$	0.4
$-b$	0.6

$\mathbb{P}(S \mid A, B)$			
$+a$	$+b$	$+s$	1.0
$+a$	$+b$	$-s$	0.0
$+a$	$-b$	$+s$	0.9
$+a$	$-b$	$-s$	0.1
$-a$	$+b$	$+s$	0.8
$-a$	$+b$	$-s$	0.2
$-a$	$-b$	$+s$	0.1
$-a$	$-b$	$-s$	0.9

Figura 2: Rede Bayesiana para a pergunta 6