POSGRADO EN CIENCIA E INGENIERÍA DE LA COMPUTACIÓN

Aprendizaje Automatizado

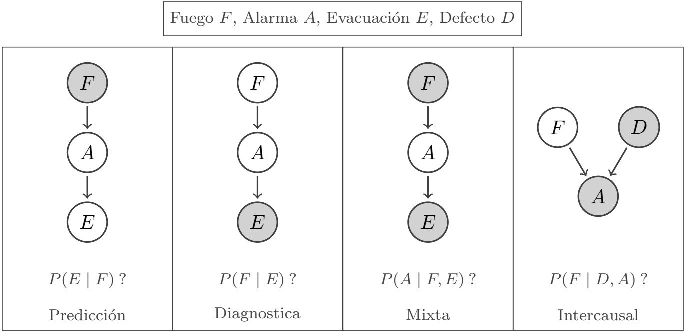
MGP: Inferencia

Prof. Gibran Fuentes Pineda

Ayud. Bere y.Ricardo

Inferencia

- Derivar una respuesta a una consulta sobre un modelo probabilístico dada cierta evidencia.



Tipos de inferencia

Métodos de inferencia

- Idea principal: paso de mensajes locales.
- Exactos:
  - fuerza bruta.
  - eliminación de variables (suma producto).
  - máximo producto.
  - árbol de uniones.
- Aproximados:
  - propagación de creencias.
  - métodos variaionales.
  - métodos de muestro (Monte-Carlo).

Marginalización

- Dada $p(A, B)$ podemos calcular $p(A)$ marginalizando B.

$$p(A) = \sum_{b \in val(B)} p(A, b)$$

$\cap$	b_1	b_2
a_1	x	y
a_2	u	v

$$p(a_1) = x + y \qquad p(a_2) = u + v$$

Fuerza bruta

- Podemos inferir sobre las variables de interés I dadas las variables de evidencia E usando:

$$p(I \mid E) = \frac{p(I, E)}{p(E)}$$

- Calculamos $p(I, E)$ marginalizando las variables restantes R:

$$p(I, E) = \sum_{r \in val(R)} p(I, E, r)$$

- Cómputo intratable!

Eliminación de variables (EV)

- Idea: eliminar las variables ocultas una a una distribuyendo las sumas sobre los productos.

  1. Obtener la distribución conjunta:

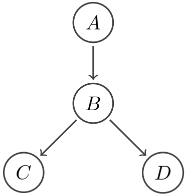
  $$p(A, B, C, D) = p(A)p(B \mid A)p(C \mid B)p(D \mid B)$$

  2. Marginalizar las variables ocultas:

  $$p(B, C) = \sum_d \sum_a p(a)p(B \mid a)p(C \mid B)p(d \mid B)$$

  3. Redistribuir

  $$p(B, C) = \sum_d p(C \mid B)p(d \mid B) \sum_a p(a)p(B \mid a)$$

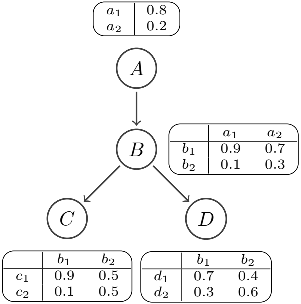


EV: ejemplo 1 (a)

- $p(B \mid c_1)$?

$$
\begin{aligned}
p(b_1, c_1) &= \sum_d p(c_1 \mid b_1) p(d \mid b_1) \sum_a p(a) p(b_1 \mid a) \\
&= \sum_d p(c_1 \mid b_1) p(d \mid b_1) (0.8 \times 0.9 + 0.2 \times 0.7) \\
&= \sum_d p(c_1 \mid b_1) p(d \mid b_1) (0.86) \\
&= (0.9 \times 0.7 + 0.9 \times 0.3)(0.86) \\
&= (0.9)(0.86) \\
&= 0.774
\end{aligned}
$$

$$p(c_1) = p(b_1, c_1) + p(b_2, c_1) = 0.774 + p(b_2, c_1)$$



EV: ejemplo 1 (b)

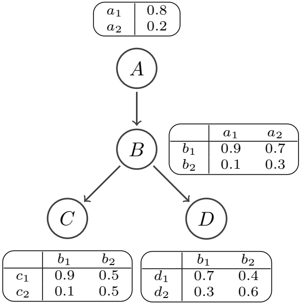
- $p(B \mid c_1)$?

$$
\begin{aligned}
p(b_2, c_1) &= \sum_d p(c_1 \mid b_2) p(d \mid b_1) \sum_a p(a) p(b_2 \mid a) \\
&= \sum_d p(c_1 \mid b_2) p(d \mid b_2)(0.8 \times 0.1 + 0.2 \times 0.3) \\
&= \sum_d p(c_1 \mid b_2) p(d \mid b_2)(0.14) \\
&= (0.5 \times 0.4 + 0.5 \times 0.6)(0.14) \\
&= (0.5)(0.14) \\
&= 0.07
\end{aligned}
$$

$$p(c_1) = p(b_1, c_1) + p(b_2, c_1) = 0.774 + 0.07 = 0.844$$

$$p(b_1 \mid c_1) = 0.774/0.844 \approx 0.917$$

$$p(b_2 \mid c_1) = 0.070/0.844 \approx 0.083$$



a_1	0.8
a_2	0.2

	a_1	a_2
b_1	0.9	0.7
b_2	0.1	0.3

	b_1	b_2
c_1	0.9	0.5
c_2	0.1	0.5

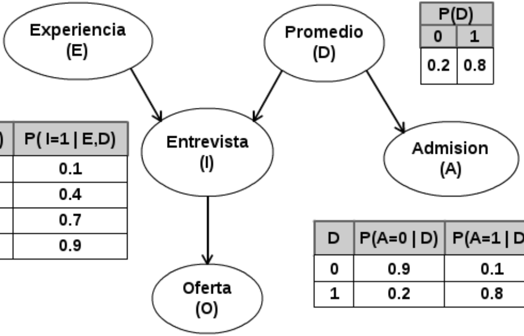
	b_1	b_2
d_1	0.7	0.4
d_2	0.3	0.6

EV: ejemplo 2

Una compañía está entrevistando candidatos para un puesto. Antes de ser invitado a entrevista, el candidato es evaluado de acuerdo a la cantidad de experiencia que tenga, así como el promedio que obtuvo en la licenciatura. Si el candidato pasa la evaluación es llamado a una entrevista. Una vez que el candidato ha sido entrevistado, la compañía puede realizar una oferta de empleo. Los candidatos además están evaluando la posibilidad de realizar estudios de posgrado, lo cual depende del promedio obtenido.

EV: ejemplo 2

P(E)	
0	1
0.7	0.3



P(D)	
0	1
0.2	0.8

E	D	P(I=0 | E,D)	P(I=1 | E,D)
0	0	0.9	0.1
0	1	0.6	0.4
1	0	0.3	0.7
1	1	0.1	0.9

D	P(A=0 | D)	P(A=1 | D)
0	0.9	0.1
1	0.2	0.8

I	P(O=0 | I)	P(O=1 | I)
0	0.9	0.1
1	0.2	0.8

EV: ejemplo 2

- $p(E \mid O = 1)$?
  1. Obtener la distribución conjunta:
  $$p(E, D, I, A, O) = p(E)p(D)p(I \mid E, D)p(A \mid D)p(O \mid I)$$
  2. Marginalizar las variables ocultas:
  $$p(E, O = 1) = \sum_d \sum_i \sum_a p(E)p(d)p(i \mid E, d)p(a \mid d)p(O = 1 \mid i)$$
  3. Redistribuir
  $$p(E, O = 1) = p(E) \sum_d p(d) \sum_i p(O = 1 \mid i)p(i \mid E, d) \sum_a p(a \mid d)$$

EV: ejemplo 2

- Eliminando A:

$$p(E, O = 1) = p(E) \sum_d p(d) \sum_i p(O = 1 \mid i) p(i \mid E, d) \sum_a p(a \mid d)$$

$$f_a(d) = \sum_a p(a \mid d)$$

D	A	$p(A \mid D)$
0	0	0.9
0	1	0.1
1	0	0.2
1	1	0.8

D	$f_a(d)$
0	$0.9 + 0.1 = 1$
1	$0.2 + 0.8 = 1$

$$p(E, O = 1) = p(E) \sum_d p(d) \sum_i p(O = 1 \mid i) p(i \mid E, d) f_a(d)$$

EV: ejemplo 2

- Eliminando I:

$$p(E, O=1) = p(E)\sum_d p(d)\sum_i p(O=1|i)p(i|E,d)f_a(d)$$

$$f_i(E,d) = \sum_i p(O=1 \mid i)p(i \mid E,d)f_a(d)$$

E	D	I	$p(O=1|i)p(i|E,d)f_a(d)$
0	0	0	$0.1 \times 0.9 \times 1 = 0.09$
0	0	1	$0.8 \times 0.1 \times 1 = 0.08$
0	1	0	$0.1 \times 0.6 \times 1 = 0.06$
0	1	1	$0.8 \times 0.4 \times 1 = 0.32$
1	0	0	$0.1 \times 0.3 \times 1 = 0.03$
1	0	1	$0.8 \times 0.7 \times 1 = 0.56$
1	1	0	$0.1 \times 0.1 \times 1 = 0.01$
1	1	1	$0.8 \times 0.9 \times 1 = 0.72$

E	D	$f_i(E,d)$
0	0	$0.09 + 0.08 = 0.17$
0	1	$0.06 + 0.32 = 0.38$
1	0	$0.03 + 0.56 = 0.59$
1	1	$0.01 + 0.72 = 0.73$

$$p(E, O=1) = p(E)\sum_d p(d)f_i(E,d)$$

EV: ejemplo 2

- Eliminando D:

$$p(E, O = 1) = p(E) \sum_d p(d) f_i(E, d)$$

$$f_d(E) = \sum_d p(d) f_i(E, d)$$

E	D	$p(d)fi(E, d)$
0	0	$0.2 \times 0.17 = 0.034$
0	1	$0.8 \times 0.38 = 0.304$
1	0	$0.2 \times 0.59 = 0.118$
1	1	$0.8 \times 0.73 = 0.584$

E	$f_d(E)$
0	$0.034 + 0.304 = 0.324$
1	$0.118 + 0.584 = 0.702$

$$p(E, O = 1) = p(E) f_d(E)$$

EV: ejemplo 2

- Calculando $p(E \mid O = 1)$:

$$p(E \mid O = 1) = p(E) f_d(E)$$

E	$p(E) f_d(E)$
0	$0.324 \times 0.7 = 0.2268$
1	$0.702 \times 0.3 = 0.2106$

$$p(E = 0 \mid O = 1) = \frac{0.2268}{0.2268 + 0.2106} \approx 0.52$$

$$p(E = 1 \mid O = 1) = \frac{0.2106}{0.2268 + 0.2106} \approx 0.48$$