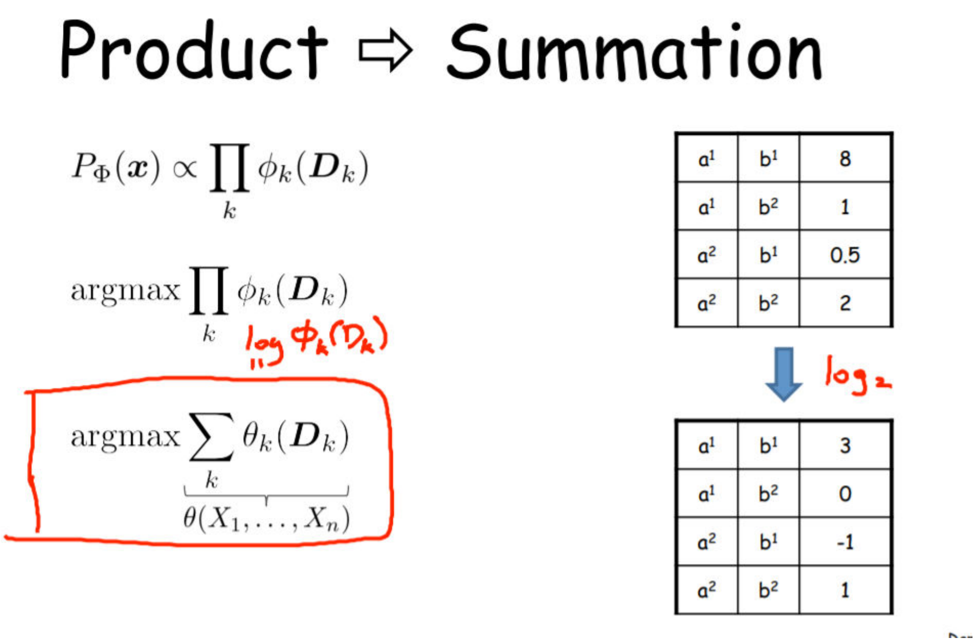
Semana 3

Paso de mensaje de MAP

Máxima inferencia exacta

Se trata de encontrar los argumentos que maximicen

la probabilidad, en lugar de trabajar con multiplicaciones se aplica el logaritmo



Product ⇨ Summation

$$P_{\Phi}(\boldsymbol{x}) \propto \prod_k \phi_k(\boldsymbol{D}_k)$$

$$\operatorname{argmax} \prod_k \phi_k(\boldsymbol{D}_k)$$

$\log \phi_k(D_k)$

$$\boxed{\operatorname{argmax} \underbrace{\sum_k \theta_k(\boldsymbol{D}_k)}_{\theta(X_1,\ldots,X_n)}}$$

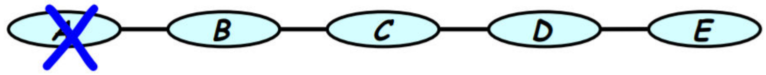
a^1	b^1	8
a^1	b^2	1
a^2	b^1	0.5
a^2	b^2	2

$\log_2$

a^1	b^1	3
a^1	b^2	0
a^2	b^1	-1
a^2	b^2	1

Es similar a eliminación de variables

Max-Sum Elimination in Chains



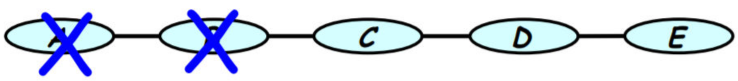
$\theta(A,B,C,D,E)$

$$\max_D \max_C \max_B \max_A \left(\theta_1(A,B) + \theta_2(B,C) + \theta_3(C,D) + \theta_4(D,E)\right)$$

$$\max_D \max_C \max_B \left(\theta_2(B,C) + \theta_3(C,D) + \theta_4(D,E) + \max_A \theta_1(A,B)\right)$$

$$\max_D \max_C \max_B \left(\theta_2(B,C) + \theta_3(C,D) + \theta_4(D,E) + \lambda_1(B)\right)$$

Max-Sum Elimination in Chains

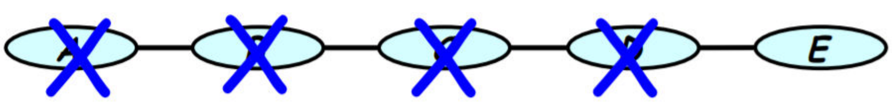


$$\max_D \max_C \max_B \left(\theta_2(B,C) + \theta_3(C,D) + \theta_4(D,E) + \lambda_1(B)\right)$$

$$\max_D \max_C \left(\theta_3(C,D) + \theta_4(D,E) + \max_B \left(\theta_2(B,C) + \lambda_1(B)\right)\right)$$

$$\max_D \max_C \left(\theta_3(C,D) + \theta_4(D,E) + \lambda_2(C)\right)$$

Max-Sum Elimination in Chains



$$\max_D \max_C \left(\theta_3(C,D) + \theta_4(D,E) + \lambda_2(C)\right)$$

$$\max_D \left(\theta_4(D,E) + \lambda_3(D)\right)$$

$\lambda_4(E)$ $\qquad$ $\lambda_4(e) = \max_{a,b,c,d} \theta(a,b,c,d,e)$

best value that I can get if we mandate $E=e$

se escoge a el mayor

Factor Summation

a^1	b^1	3
a^1	b^2	0
a^2	b^1	-1
a^2	b^2	1

b^1	c^1	4
b^1	c^2	1.5
b^2	c^1	0.2
b^2	c^2	2

a^1	b^1	c^1	3+4=7
a^1	b^1	c^2	3+1.5=4.5
a^1	b^2	c^1	0+0.2=0.2
a^1	b^2	c^2	0+2=2
a^2	b^1	c^1	-1+4=3
a^2	b^1	c^2	-1+1.5=0.5
a^2	b^2	c^1	1+0.2=1.2
a^2	b^2	c^2	1+2=3

Daphne Koller

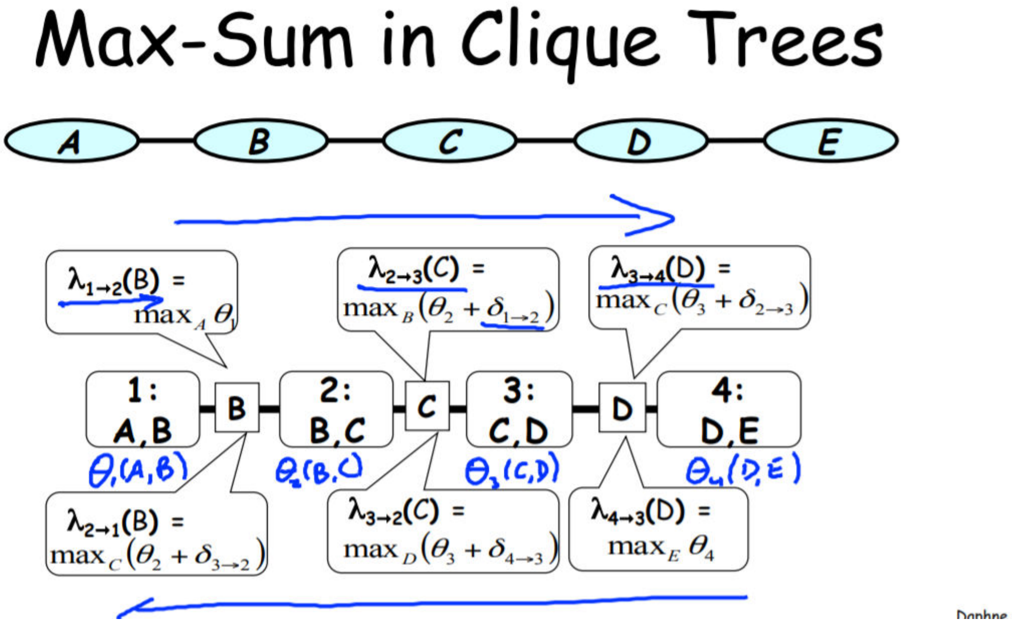
Factor Maximization

max - marginalization

a^1	b^1	c^1	7
a^1	b^1	c^2	4.5
a^1	b^2	c^1	0.2
a^1	b^2	c^2	2
a^2	b^1	c^1	3
a^2	b^1	c^2	0.5
a^2	b^2	c^1	1.2
a^2	b^2	c^2	3

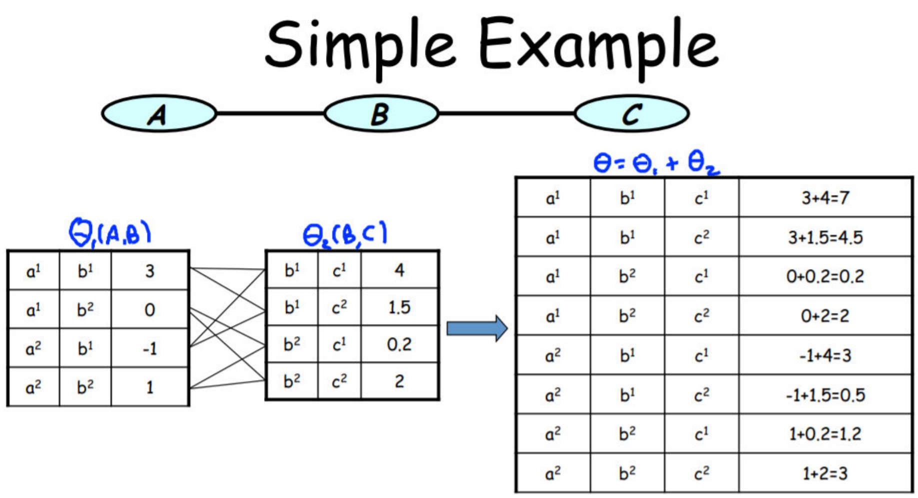
a^1	c^1	7
a^1	c^2	4.5
a^2	c^1	3
a^2	c^2	3

Se puede hacer como propagacion de creencias

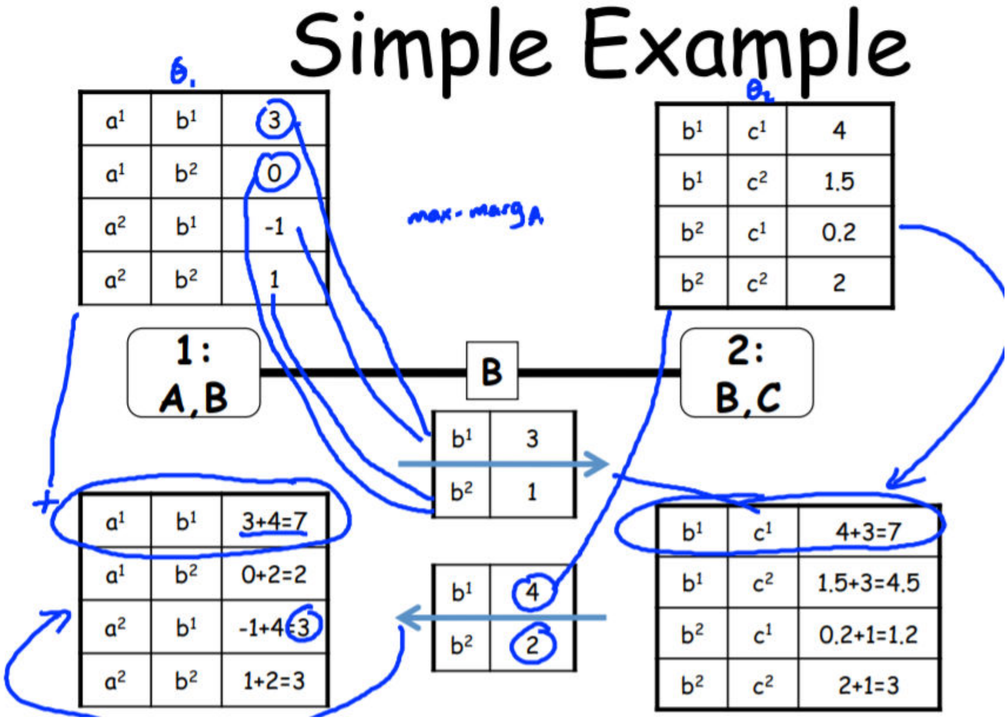


Convergencia del paso de mensaje

- Una vez que C_i recibe un mensaje final de todos los vecinos, excepto C_j, entonces $\lambda_{i\to j}$ también es final (nunca cambiará)
- Los mensajes de las hojas son inmediatamente final



los argumentos optimos son a^1, b^1, c^1, si lo hacemos por propagacion de creencias



vemos que en el paso de mensajes, vemos que los argumentos optimos son los mismos

Max-suma BP en la convergencia

• Las creencias en cada camarilla son max-marginals.

$$\beta_i(C_i) = \theta_i(C_i) + \sum_k \lambda_{k\to i} \quad \text{incoming msgs}$$

$$\beta_i(\boldsymbol{C}_i) = \max_{\boldsymbol{W}_i} \theta(\boldsymbol{C}_i, \boldsymbol{W}_i) \qquad \boldsymbol{W}_i = \{X_1, \ldots, X_n\} - \boldsymbol{C}_i$$

Calibración: las camarillas están de acuerdo con las variables compartidas

$$\max_{\boldsymbol{C}_i - \boldsymbol{S}_{i,j}} \beta_i(\boldsymbol{C}_i) = \max_{\boldsymbol{C}_j - \boldsymbol{S}_{i,j}} \beta_j(\boldsymbol{C}_j)$$

a^1	b^1	3+4=7
a^1	b^2	0+2=2
a^2	b^1	-1+4=3
a^2	b^2	1+2=3

b^1 7
b^2 3

b_1 7
b_2 3

b^1	c^1	4+3=7
b^1	c^2	1.5+3=4.5
b^2	c^1	0.2+1=1.2
b^2	c^2	2+1=3

$\beta_2 = \theta_2 + \lambda_{1\to 2}$

El maximo para b^1 es 7 y para b^2 es 3

Resumen

• El mismo algoritmo de árbol de clique que se usa para sum-producto se puede utilizar para Max-Sum

• Al igual que en el producto, la convergencia se logra después de un solo paso arriba abajo

• El resultado es un máximo marginal en cada clique C:

- - Para cada asignación c a C, la cuál es la puntuación de la mejor finalización a c

Encontrar una asignación de mapa

Decodificación de una asignación de mapa

- Fácil si la asignación de mapas es única
    - - asignación de maximización única en cada camarilla
    - - cuyo valor es el valor θ de la asignación de mapas
    - - Debido a la calibración, las elecciones en todas las camaras deben estar de acuerdo

a^1	b^1	c^1	7
a^1	b^1	c^2	4.5
a^1	b^2	c^1	0.2
a^1	b^2	c^2	2
a^2	b^1	c^1	3
a^2	b^1	c^2	0.5
a^2	b^2	c^1	1.2
a^2	b^2	c^2	3

a^1	b^1	3+4=7
a^1	b^2	0+2=2
a^2	b^1	-1+4=3
a^2	b^2	1+2=3

b^1	c^1	4+3=7
b^1	c^2	1.5+3=4.5
b^2	c^1	0.2+1=1.2
b^2	c^2	2+1=3

Daphne Koller

- Si la asignación de mapas no es única, podemos tener varias opciones en algunas camarillas
- La vinculación arbitraria puede no producir una asignación de mapa

a^1	b^1	2
a^1	b^2	1
a^2	b^1	1
a^2	b^2	2

b^1	c^1	2
b^1	c^2	1
b^2	c^1	1
b^2	c^2	2

- Dos opciones:
    - - Parámetros ligeramente perturbados para hacer que el MAP sea único.
    - - Use el procedimiento de rastreo que crea incrementalmente una asignación de mapa, una variable a la vez