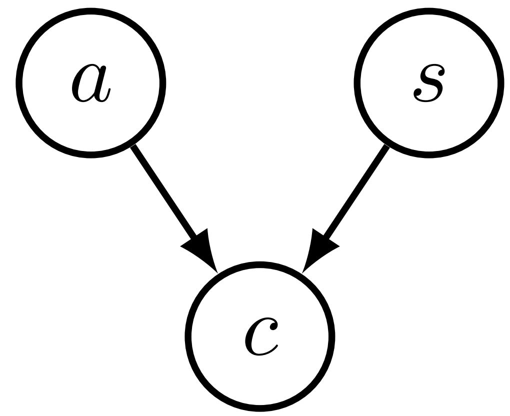
EM Overview

- **E-step:**
  - For each data instance $\mathbf{d_i}$ and each family $X,\ parents(X)$ compute $p(X, parents(X) \mid \mathbf{d_i}, \theta^t)$
  - Compute for each $(x, \mathbf{u}) \in dom(X, parents(X))$
  - $q_t(x, \mathbf{u}) = \sum_i p(x, \mathbf{u} \mid \mathbf{d_i}, \theta^t)$
- **M-Step:** Perform maximum likelihood estimation with respect to the "soft completed" data:

$$p^{t+1}(x \mid \mathbf{u}) = \theta^{t+1}_{x|\mathbf{u}} = \frac{q_t(x, \mathbf{u})}{q_t(\mathbf{u})}$$

Exercise

- Learn the parameters $\theta_1 - \theta_6$ (described below) underlying al CPTs using EM
- Start with priors estimates of 0.5 for all parameters



a	s	c
?	1	1
1	0	0
0	?	1
0	?	0
1	1	1
?	0	0
1	?	?

$$
\begin{aligned}
\theta_1 &= p(s=1) \\
\theta_2 &= p(a=1) \\
\theta_3 &= p(c=1|s=1, a=1) \\
\theta_4 &= p(c=1|s=1, a=0) \\
\theta_5 &= p(c=1|s=0, a=1) \\
\theta_6 &= p(c=1|s=0, a=0)
\end{aligned}
$$

- **One detailed E-Step:**

$$p(a \mid s, c) = \frac{p(a, s, c)}{p(s, c)} = \frac{p(s)p(a)p(c \mid a, s)}{\sum_a p(s)p(a)p(c \mid a, s)}$$

$$= \frac{p(s)p(a)p(c \mid a, s)}{p(s)p(a=1)p(c \mid a=1,s) + p(s)p(a=0)p(c \mid a=0,s)}$$

$$p(a=1 \mid s=1, c=1) =$$

$$= \frac{p(s=1)p(a=1)p(c=1 \mid a=1,s=1)}{p(s=1)p(a=1)p(c=1 \mid a=1,s=1) + p(s=1)p(a=0)p(c=1 \mid a=1,s=1)}$$

$$= \frac{\theta_1 \cdot \theta_2 \cdot \theta_3}{\theta_1 \cdot \theta_2 \cdot \theta_3 + \theta_1 \cdot (1-\theta_2) \cdot \theta_4} = \frac{0.5 \times 0.5 \times 0.5}{0.5 \times 0.5 \times 0.5 + 0.5 \times 0.5 \times 0.5}$$

E-step: Using current estimate of parameters calculate probability of each data instance

a	s	C
?	1	1
1	0	0
0	?	1
0	?	0
1	1	1
?	0	0
1	?	?

Incomplete data

	a	s	c		
d1	1	1	1	$p(a=1 \mid s=1, c=1, \theta^t)$	$=0.5$
	0	1	1	$p(a=0 \mid s=1, c=1, \theta^t)$	$=0.5$
d2	1	0	0	1	= 1
d3	0	1	1	$p(s=1 \mid a=0, c=1, \theta^t)$	$=0.5$
	0	0	1	$p(s=0 \mid a=0, c=1, \theta^t)$	$=0.5$
d4	0	1	0	$p(s=1 \mid a=0, c=0, \theta^t)$	$=0.5$
	0	0	0	$p(s=0 \mid a=0, c=0, \theta^t)$	$=0.5$
d5	1	1	1	1	= 1
d6	1	0	0	$p(a=1 \mid s=0, c=0, \theta^t)$	$=0.5$
	0	0	0	$p(a=0 \mid s=0, c=0, \theta^t)$	$=0.5$
d7	1	0	1	$p(s=0, c=1 \mid a=1, \theta^t)$	$=0.25$
	1	0	0	$p(s=0, c=0 \mid a=1, \theta^t)$	$=0.25$
	1	1	1	$p(s=1, c=1 \mid a=1, \theta^t)$	$=0.25$
	1	1	0	$p(s=1, c=0 \mid a=1, \theta^t)$	$=0.25$

$$q(a = 1) = \sum_{i=1}^{7} p(a = 1 \mid \mathbf{d_i}, \theta^t) = 0.5 + 1 + 0 + 0 + 1 + 0.5 + (2 \times 0.25) = 4$$

$$q(s = 1) = \sum_{i=1}^{7} p(s = 1 \mid \mathbf{d_i}, \theta^t) = 1 + 0 + 0.5 + 0.5 + 1 + 0 + (4 \times 0.25) = 3.5$$

$$q(c = 1, s = 1, a = 1) = \sum_{i=1}^{7} p(c = 1, s = 1, a = 1 \mid \mathbf{d_i}, \theta^t) = 0.5 + 0 + 0 + 0 + 1 + 0 + 0.25 = 1.75$$

$$q(c = 1, s = 1, a = 0) = \sum_{i=7}^{7} p(c = 1, s = 1, a = 0 \mid \mathbf{d_i}, \theta^t) = 0.5 + 0 + 0.5 + 0 + 0 + 0 + 0 = 1$$

$$q(c = 1, s = 0, a = 1) = \sum_{i=1}^{7} p(c = 1, s = 0, a = 1 \mid \mathbf{d_i}, \theta^t) = 0 + 0 + 0 + 0 + 0 + 0 + 0.25 = 0.25$$

$$q(c = 1, s = 0, a = 0) = \sum_{i=1}^{7} p(c = 1, s = 0, a = 0) \mid \mathbf{d_i}, \theta^t) = 0 + 0 + 0.5 + 0 + 0 + 0 + 0 = 0.5$$

Compute the “soft counts”

$q(c = 0, \_, \_)$ on your own and compare with result on next page

M-step: Perform maximum likelihood estimation with respect to the "soft completed" data:

$$\theta_a^{t+1} = \theta_1^{t+1} = \frac{q(a=1)}{q(a=1)+q(a=0)} = \frac{q(a=1)}{7} = 0.571$$

$$\theta_s^{t+1} = \theta_2^{t+1} = \frac{q(s=1)}{q(s=1)+q(s=0)} = \frac{q(s=1)}{7} = 0.5$$

$$\theta_{c=1|s=1,a=1}^{t+1} = \theta_3^{t+1} = \frac{q(c=1,a=1,s=1)}{q(c=1,a=1,s=1)+q(c=0,a=1,s=1)} = \frac{1.75}{1.75+0.25} = 0.875$$

$$\theta_{c=1|s=1,a=0}^{t+1} = \theta_4^{t+1} = \frac{q(c=1,s=1,a=0)}{q(c=1,s=1,a=0)+q(c=0,s=1,a=0)} = \frac{1}{1+0.5} = 0.666$$

$$\theta_{c=1|s=0,a=1}^{t+1} = \theta_5^{t+1} = \frac{q(c=1,s=0,a=1)}{q(c=1,s=0,a=1)+q(c=0,s=0,a=1)} = \frac{0.25}{0.25+(1+0.5+0.25)} = 0.125$$

$$\theta_{c=1|s=0,a=0}^{t+1} = \theta_6^{t+1} = \frac{q(c=1,s=0,a=0)}{q(c=1,s=0,a=0)+q(c=0,s=0,a=0)} = \frac{0.5}{0.5+(0.5+0.5)} = 0.333$$

M-step: Perform maximum likelihood estimation with respect to the “soft completed” data:

Using the updated parameters $\theta^{t+|1}$, compute the next iteration.

$$\theta_a^{t+1} = \theta_1^{t+1} = \frac{q(a=1)}{q(a=1)+q(a=0)} = \frac{q(a=1)}{7} = 0.571$$

$$\theta_s^{t+1} = \theta_2^{t+1} = \frac{q(s=1)}{q(s=1)+q(s=0)} = \frac{q(s=1)}{7} = 0.5$$

$$\theta_{c=1|s=1,a=1}^{t+1} = \theta_3^{t+1} = \frac{q(c=1,a=1,s=1)}{q(c=1,a=1,s=1)+q(c=0,a=1,s=1)} = \frac{1.75}{1.75+0.25} = 0.875$$

$$\theta_{c=1|s=1,a=0}^{t+1} = \theta_4^{t+1} = \frac{q(c=1,s=1,a=0)}{q(c=1,s=1,a=0)+q(c=0,s=1,a=0)} = \frac{1}{1+0.5} = 0.666$$

$$\theta_{c=1|s=0,a=1}^{t+1} = \theta_5^{t+1} = \frac{q(c=1,s=0,a=1)}{q(c=1,s=0,a=1)+q(c=0,s=0,a=1)} = \frac{0.25}{0.25+(1+0.5+0.25)} = 0.125$$

$$\theta_{c=1|s=0,a=0}^{t+1} = \theta_6^{t+1} = \frac{q(c=1,s=0,a=0)}{q(c=1,s=0,a=0)+q(c=0,s=0,a=0)} = \frac{0.5}{0.5+(0.5+0.5)} = 0.333$$

	a	s	c		
d1	1	1	1	$p(a=1 \mid s=1, c=1, \theta^t)$	$=0.56$
	0	1	1	$p(a=0 \mid s=1, c=1, \theta^t)$	$=0.44$
d2	1	0	0	1	=
d3	0	1	1	$p(s=1 \mid a=0, c=1, \theta^t)$	=
	0	0	1	$p(s=0 \mid a=0, c=1, \theta^t)$	=
d4	0	1	0	$p(s=1 \mid a=0, c=0, \theta^t)$	=
	0	0	0	$p(s=0 \mid a=0, c=0, \theta^t)$	=
d5	1	1	1	1	=
d6	1	0	0	$p(a=1 \mid s=0, c=0, \theta^t)$	=
	0	0	0	$p(a=0 \mid s=0, c=0, \theta^t)$	=
d7	1	0	1	$p(s=0, c=1 \mid a=1, \theta^t)$	=
	1	0	0	$p(s=0, c=0 \mid a=1, \theta^t)$	=
	1	1	1	$p(s=1, c=1 \mid a=1, \theta^t)$	=
	1	1	0	$p(s=1, c=0 \mid a=1, \theta^t)$	=