(1)

- $P(X_0)$ prior · (from the algo)

Q1 Set A &
Q 4 Set B

- $P(X_1|X_0)$ transition model
- $P(E_1|X_1)$ Sensor model
- S : vector of samples of size N
- w : vector of weights of size N

Initial weight for all samples = 1

At time step 3 , $t = 3$

$$\left.\begin{aligned} A1 &= +a \\ A2 &= -a \\ A3 &= +a \end{aligned}\right\} \text{ evidence: } E$$

- The weight of the particle is given by. } as given in algorithm

$$w[i] \leftarrow P(e|x_1 = S[i])$$

(a) Here the weight will be

$$P(A_3 = a \mid \overset{W_3}{X_3 = +w}) = \underline{0.5}$$

Similarly, $P(A_3 = a \mid \underset{W_3}{X_3 = -w}) = \underline{0.9}$

(b) $+w$: 3 particles at $t = 6$

$-w$: 5 particles

$A_6 = -a$ (evidence)

- Each sample (particles) in state $+w$ will have weight $P(A_6 = -a \mid W_6 = +w) = 0.5$

  Total weight for 3 particles in state $+w$: $0.5 \times 3 = 1.5$

- Each particle in state $-w$ will have weight $P(A_6 = -a \mid w_6 = -w) = 0.1$

  total for 5 particles $0.1 \times 5 = 0.5$

Normalizing the weights to form a prob. distribution

$P(w_7 = +w) = \underline{0.75}$

$P(w_7 = -w) = \underline{0.25}$

(3) parameters: $\theta_a \ \theta_s \ \theta_c^{11} \ \theta_c^{01} \ \theta_c^{10} \ \theta_c^{00}$

Q2 set B

	$\underline{\theta_a}$	$\underline{\theta_s}$	$\underline{\theta_c^{11}}$	$\underline{\theta_c^{01}}$	$\underline{\theta_c^{10}}$	$\underline{\theta_c^{00}}$
MLE from the table	$4/7$	$4/7$	1	$1/2$	$1/2$	0

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

Bayesian parameter learning.

prior prob. is given as beta distribution

mean of the distribution is $\dfrac{\alpha}{\alpha+\beta}$

Initially $\alpha = 1$, $\beta = 1$ (uniform dist)

$$\text{beta}[\alpha, \beta](\theta) \propto \theta^{\alpha-1} (1-\theta)^{\beta-1}$$

α and β can be viewed as virtual counts.

→ Relate this to candy example. when cherry candy arrives in the observation we used to increment α and if lime candy arrives we used to increment β to get the posterior.

Also, recall how the the distribution changes as α and β change. we will select θ at the peak which is mean of the distribution $= \dfrac{\alpha}{\alpha+\beta}$.

for θ_a: the observations from the table

1 1 0 0 1 0 1

increment α

increment β and so on.

initial value of α

$$P(a=1 \mid \text{data}) = \frac{1 + \#(a=1)}{2+N} = \frac{5}{9} \longrightarrow \theta_a$$

$$P(s=1 \mid \text{data}) = \frac{1 + \#(s=1)}{2+N} = \frac{5}{9} \longrightarrow \theta_s$$

θ_c^{11}

$P(c=1 \mid a=1, s=1, \text{data})$

$$= \frac{1 + \#(c=1, a=1, s=1)}{2 + \#(c=1, a=1, s=1) + \#(c=0, a=1, s=1}$$

$$= \frac{3}{4}$$

$\theta_c^{10} = \frac{2}{3}$, $\theta_c^{01} = \frac{1}{2}$, $\theta_c^{00} = \frac{1}{3}$

Q4 Expectation maximization

Q3 Set B

E - step:

$$\hat{N}(a=1) = \sum_{j=1}^{N} P(a=1 \mid s_j, c_j)$$

from the Table we get

	$c=0$	$c=1$
$s=0$	2	1
$s=1$	1	3

In E - step, we calculate the expected counts using the given parameters (Prob.)

given: $\theta_a = \theta_s = \theta_c^{11}, \theta_c^{10}, \theta_c^{01}, \theta_c^{00} = 0.5$

$$\hat{N}(a=1) = \sum_{j=1}^{N} P(a=1 \mid s_j, c_j)$$

first observation - $s=1$, $c=1$, we have 3 such observations

$$3 \times P(a=1 \mid s=1, c=1)$$

$$= 3 \times \frac{P(c=1 \mid a=1, s=1)\, P(a=1)\, P(s=1)}{P(c=1, s=1)}$$ ← marginalise this

$$= 3 \times \frac{0.5 \times 0.5 \times 0.5}{P(c=1 \mid a=1, s=1)\, P(a=1)\, P(s=1) + P(c=1 \mid a=0, s=1)\, P(a=0)\, P(s=1)}$$

$$= 3 \times \frac{0.5 \times 0.5 \times 0.5}{(0.5)^3 + (0.5)^3} = 3 \times 0.5 = 1.5$$

similarly, we can do this for other observations.

$(s=0, c=0)$ $\quad$ $(s=1, c=0)$ $\quad$ $(s=0, c=1)$

$2 \times 0.5 = 1$ $\quad$ $1 \times 0.5 = 0.5$ $\quad$ $1 \times 0.5 = 0.5$

$$\boxed{\hat{N}(a=1) = 3.5}$$

this value of prob::
will be 0.5 for other observations as well because the initial values were all set to 0.5.

M - step: $\hat{\theta}_a = \dfrac{3.5}{7} = 0.5$