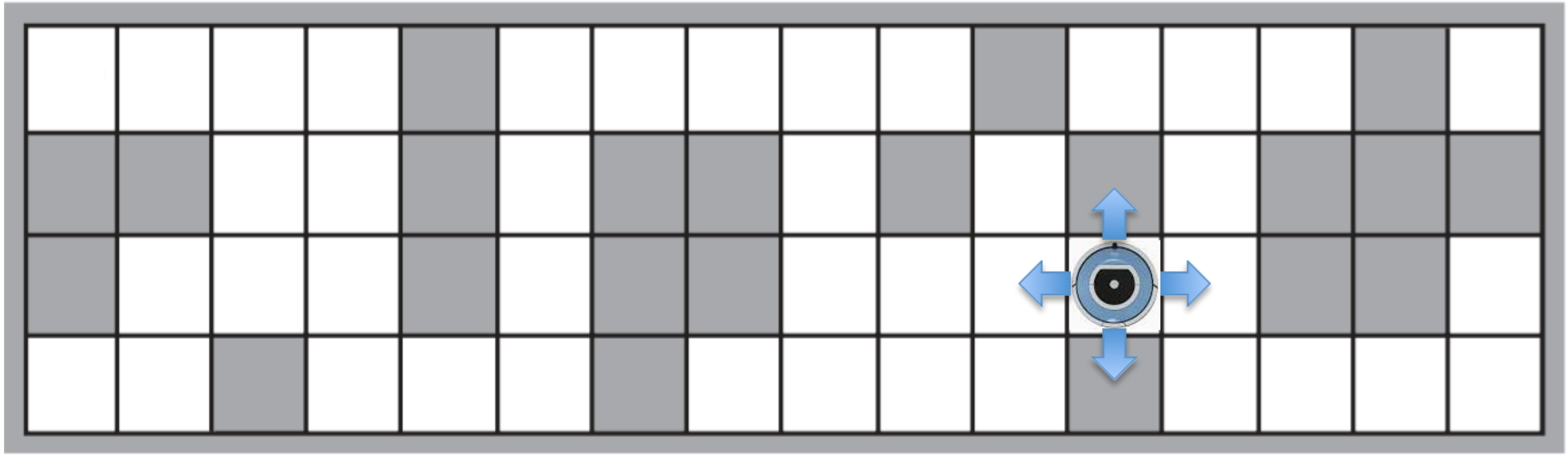
Probabilistic Reasoning Over Time

3007/7059 Artificial Intelligence

Slides by T-J. Chin, adjusted by Wei Zhang

School of Computer Science
University of Adelaide

Example: robot localisation



- A robot is moving in a 2D map with obstacles. Its sensor detects the directions are blocked by obstacles at the robot's location.
- A sequence of sensor readings is observed:

Time step	**1**	**2**	**3**	**…**	***t***	***t*+1**
Blocked directions	N S W	N S	N	…	S E W	S E

- At time step t+1, where is the robot?

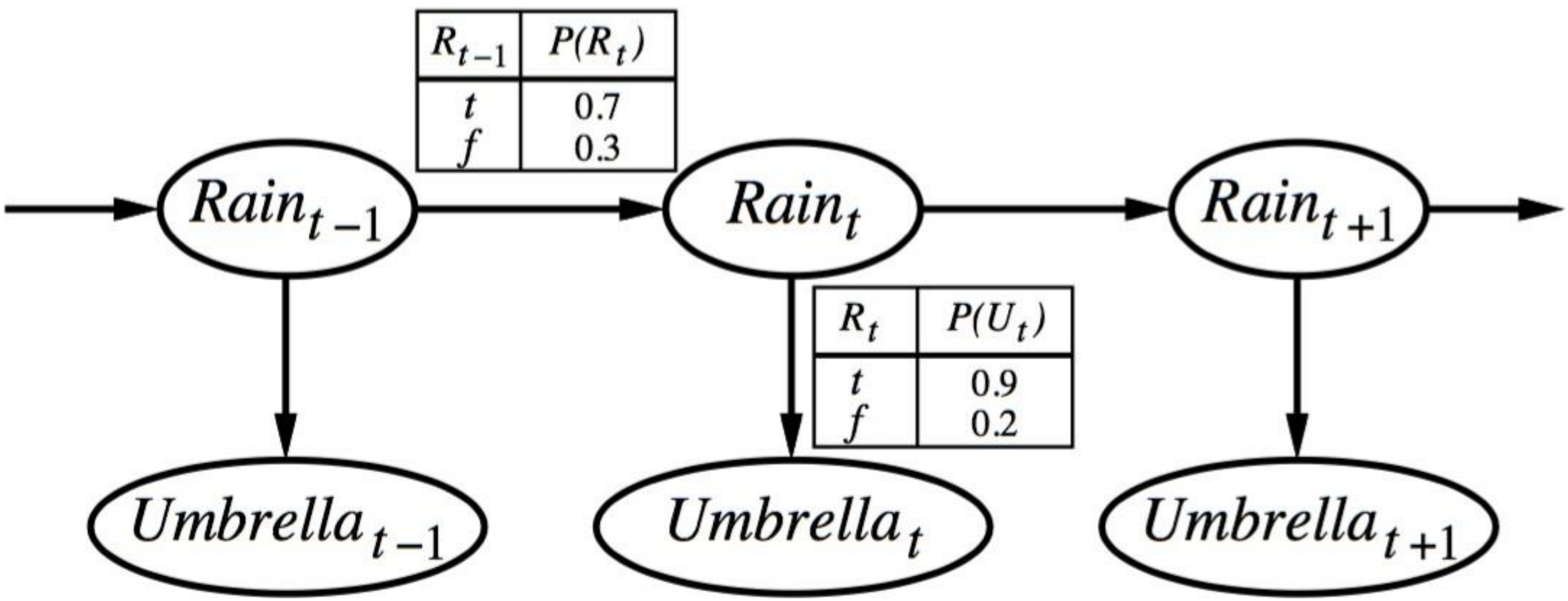
Example: is it raining outside?

- You are the security guard permanently located at a secret underground installation.
- You cannot see the weather outside.
- Everyday, you see the director arriving with or without an umbrella.
- At day t+1, the director arrived with an umbrella. Is it raining outside?

Day	1	2	3	…	t	t+1
Observed umbrella?	✓	✓	✗	…	✗	✓

Example: is it raining outside?

- A commonly used temporal model for this kind of problem:



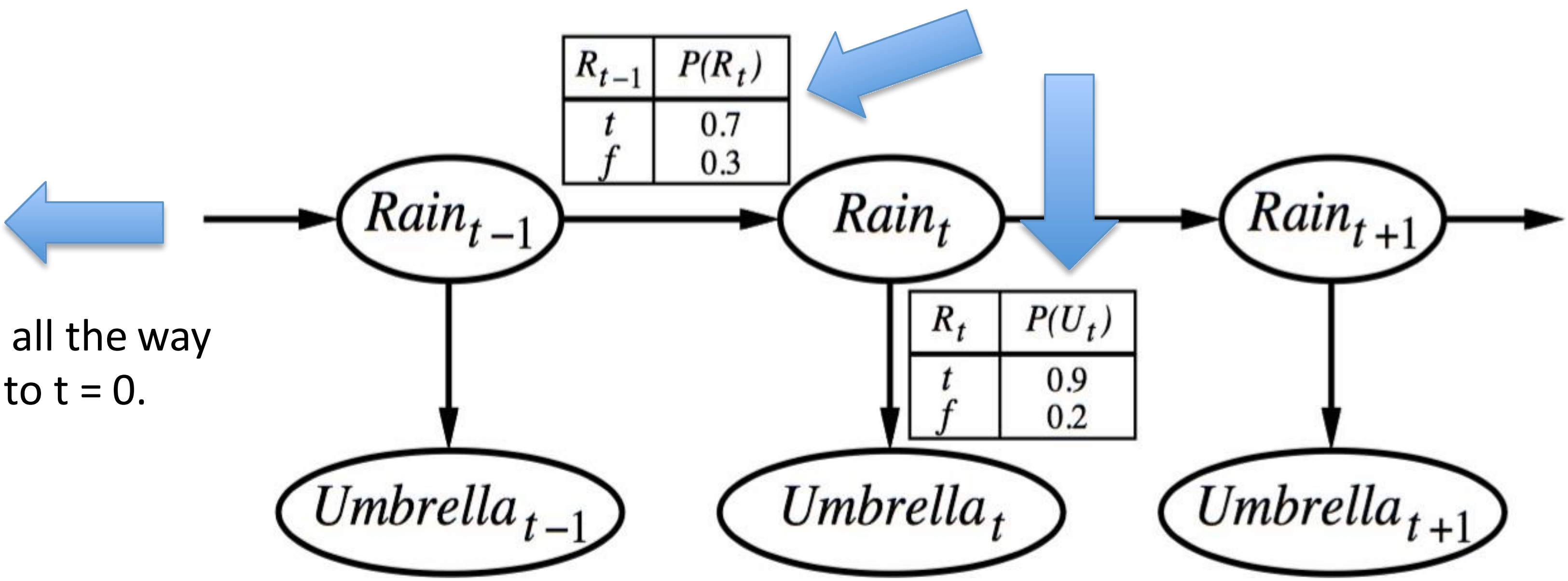
R_{t-1}	$P(R_t)$
t	0.7
f	0.3

R_t	$P(U_t)$
t	0.9
f	0.2

Example: is it raining outside?

- This is just a Bayesian Network with the concept of time.

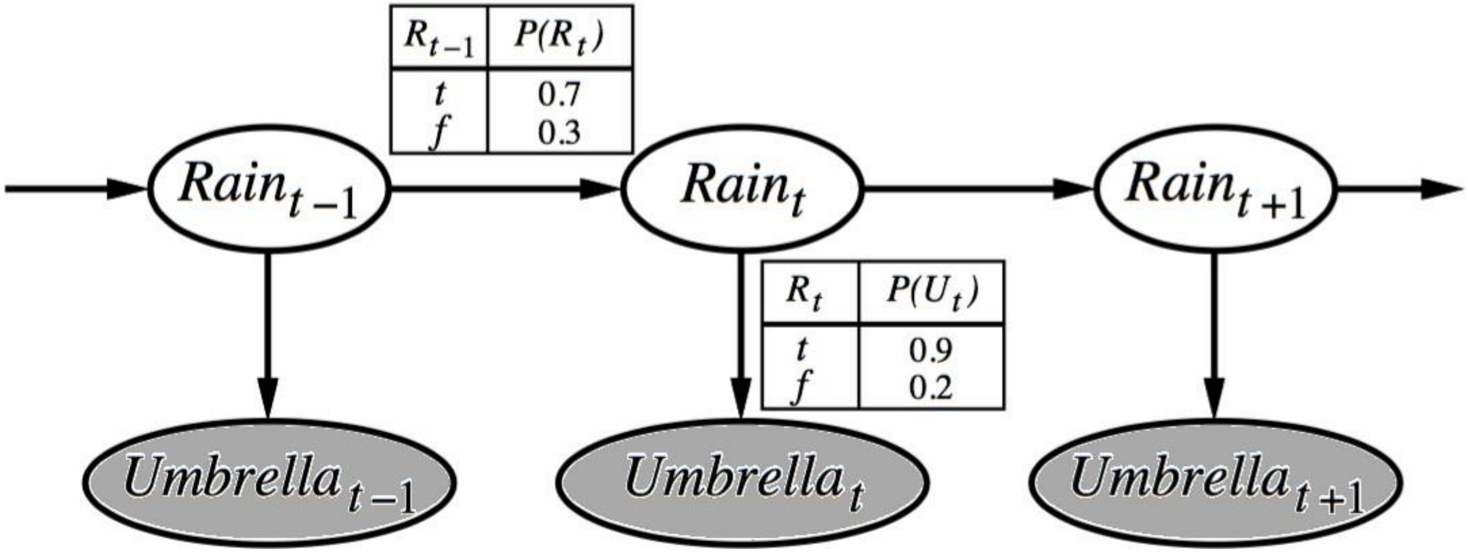
Conditional probability tables



Goes all the way back to t = 0.

- Variables = { R_0, R_1, …, R_{t+1}, U_1, …, U_{t+1} }.

Example: is it raining outside?

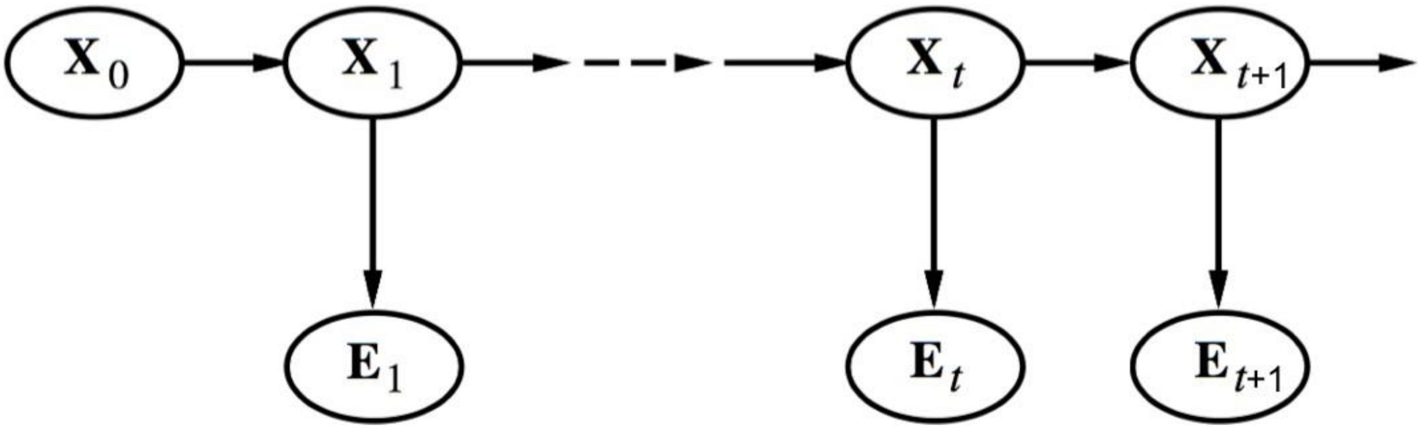


- You have observed evidences(Umbrella) $\{u_1, \ldots, u_{t+1}\}$ = {true,true,false,…,false,true}.
- You want to calculate the probability

$$P(R_{t+1} | u_1, \ldots, u_{t+1})$$

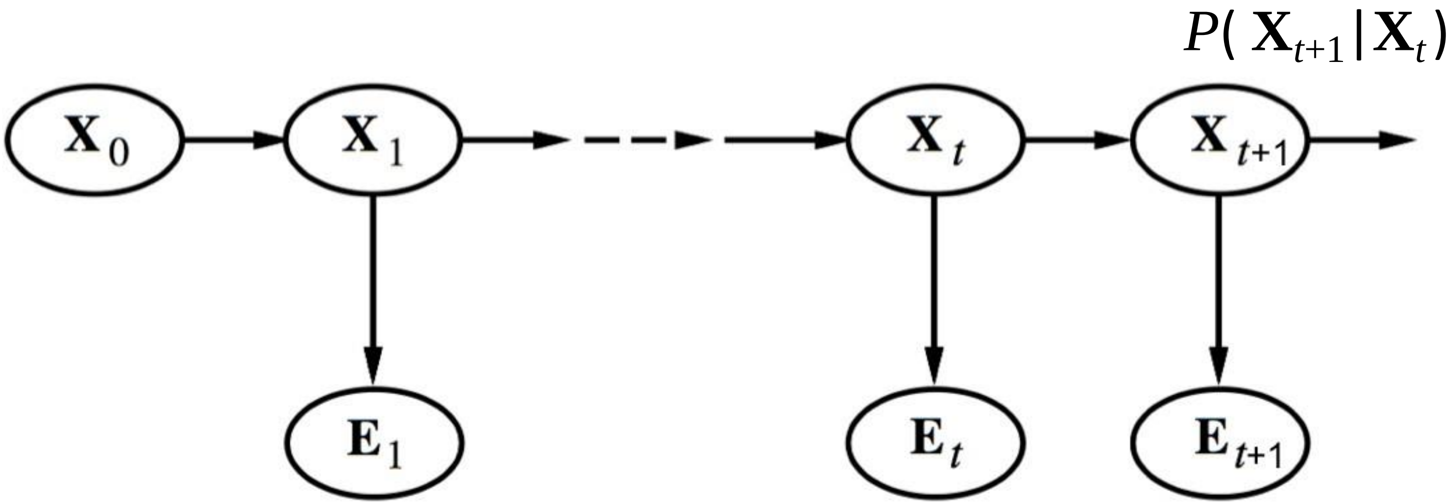
for R_{t+1} = true and R_{t+1} = false.
- This is a special kind of probabilistic inference called filtering.

The general case



- State variables $\{ \mathbf{X}_0, \mathbf{X}_1, \ldots, \mathbf{X}_{t+1} \}$.
- Evidence variables $\{ \mathbf{E}_1, \ldots, \mathbf{E}_{t+1} \}$.
- By convention, we assume $\mathbf{X}_t$ starts at t=0 while $\mathbf{E}_t$ starts at t=1.

The general case



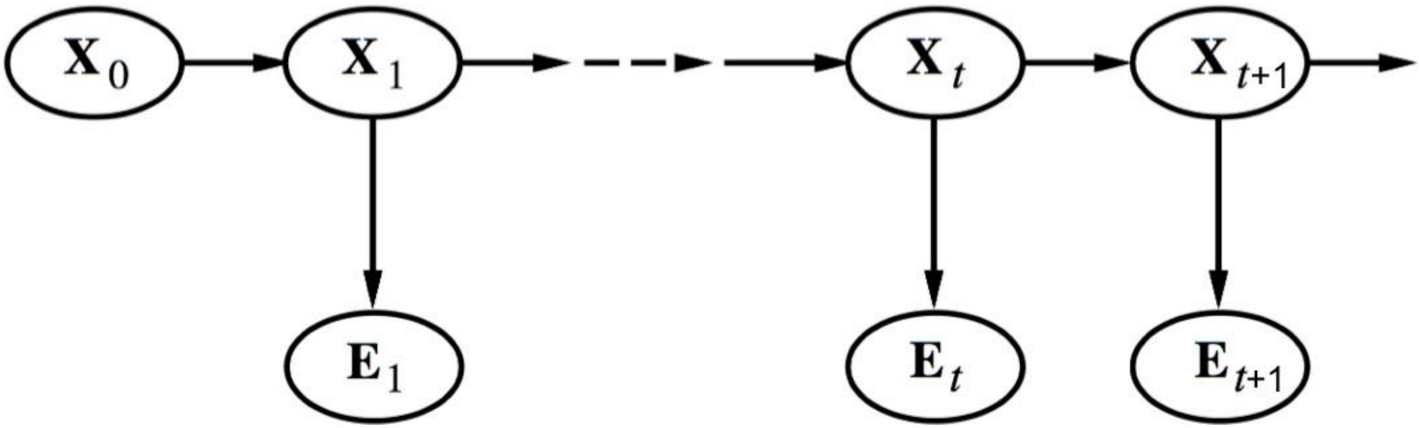
- State transition model

$$\mathrm{P}(\mathbf{X}_{t+1} | \mathbf{X}_0, \ldots, \mathbf{X}_t) = \mathrm{P}(\mathbf{X}_{t+1} | \mathbf{X}_t)$$

- **First order Markov assumption**: the present state depends only on the immediate previous state.

The general case

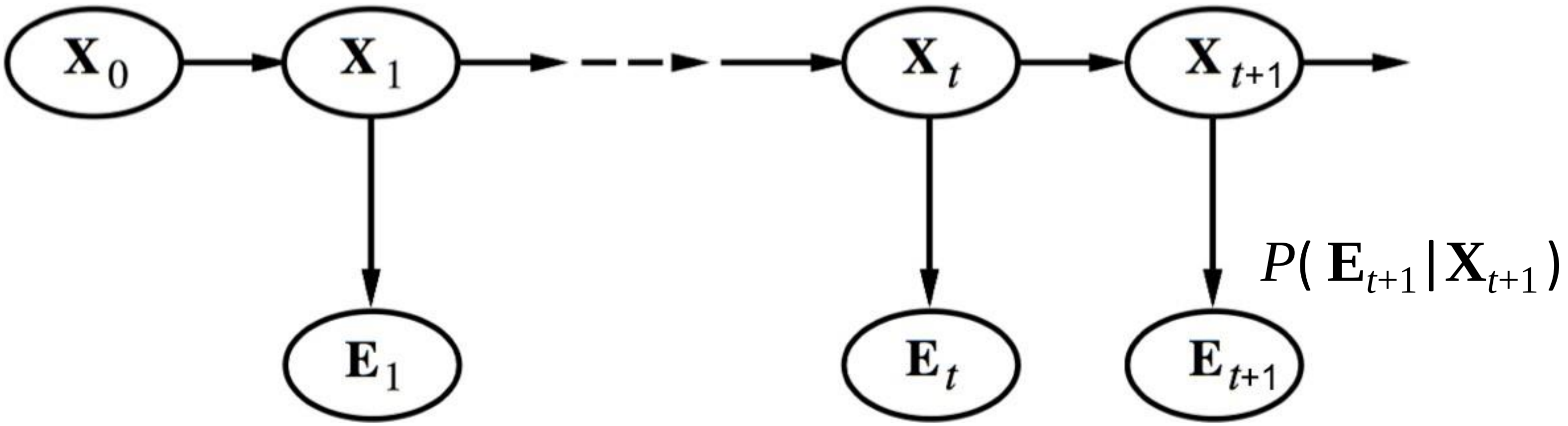
$P(\mathbf{X}_{t+1}|\mathbf{X}_t)$



Assume the state changes are caused by a **stationary process**—that is, a process of change that is governed by laws that do not themselves change over time.

$P(\mathbf{X}_t|\mathbf{X}_{t-1})$ is the same for all t.

The general case



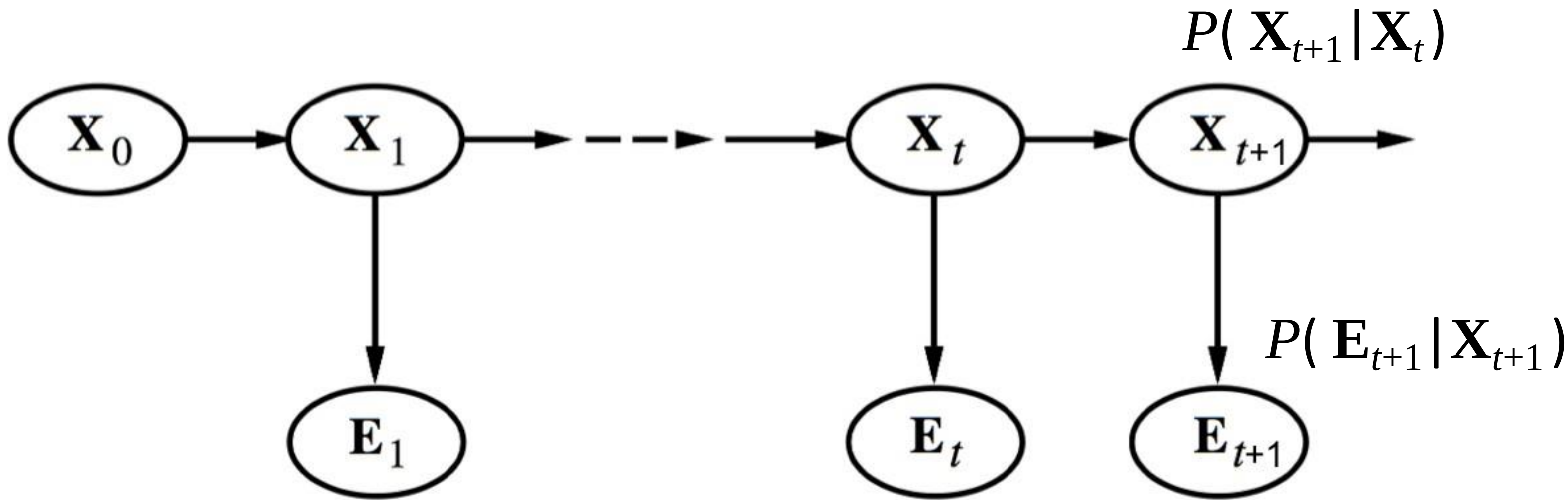
- Observation model (Sensor model)

$$P(\mathbf{E}_{t+1} | \mathbf{X}_{0:t+1}, \mathbf{E}_{0:t}) = P(\mathbf{E}_{t+1} | \mathbf{X}_{t+1})$$

- Sensor Markov assumption: the probability of observing $\mathbf{E}_t$ depends only on the state $\mathbf{X}_t$.

*Note: $\mathbf{X}_{0:t} = \mathbf{X}_0, \mathbf{X}_1, \ldots, \mathbf{X}_t$

The general case



$$\mathbf{P}(\mathbf{X}_{0:t}, \mathbf{E}_{1:t}) = \mathbf{P}(\mathbf{X}_0) \prod_{i=1}^{t} \mathbf{P}(\mathbf{X}_i \mid \mathbf{X}_{i-1}) \, \mathbf{P}(\mathbf{E}_i \mid \mathbf{X}_i)$$

Inference: Filtering

- We have observed $\mathbf{e}_1, \ldots, \mathbf{e}_{t+1} = \mathbf{e}_{1:t+1}$. We wish to calculate

$$\boxed{\mathbf{P}(\mathbf{X}_{t+1} \mid \mathbf{e}_{1:t+1})}$$

$$\mathbf{P}(\mathbf{X}_{t+1} \mid \mathbf{e}_{1:t+1}) = \mathbf{P}(\mathbf{X}_{t+1} \mid \mathbf{e}_{1:t}, \mathbf{e}_{t+1}) \quad \text{(dividing up the evidence)}$$

$$= \alpha\, \mathbf{P}(\mathbf{e}_{t+1} \mid \mathbf{X}_{t+1}, \mathbf{e}_{1:t})\, \mathbf{P}(\mathbf{X}_{t+1} \mid \mathbf{e}_{1:t}) \quad \text{(using Bayes' rule)}$$

Set $A=\mathbf{X}_{t+1},$ $B = \mathbf{e}_{1:t},$ $C=\mathbf{e}_{t+1},$

$P(\mathbf{X}_{t+1|}\ \mathbf{e}_{1:t+1})$

$$=P(A|B,C) = \frac{P(A,B,C)}{P(B,C)} = \frac{P(C|A,B)P(A,B)}{P(B,C)} = \frac{P(C|A,B)P(A|B)P(B)}{P(B,C)}$$

$$= \alpha P(C|A,B)P(A|B) = \alpha\, P(\mathbf{e}_{t+1}|\mathbf{X}_{t+1}, \mathbf{e}_{1:t})P(\mathbf{X}_{t+1}|\mathbf{e}_{1:t})$$

Inference: Filtering

- We have observed $\mathbf{e}_1, \ldots, \mathbf{e}_{t+1} = \mathbf{e}_{1:t+1}$. We wish to calculate

$$\boxed{\mathbf{P}(\mathbf{X}_{t+1} \mid \mathbf{e}_{1:t+1})}$$

$$\begin{aligned}
\mathbf{P}(\mathbf{X}_{t+1} \mid \mathbf{e}_{1:t+1}) &= \mathbf{P}(\mathbf{X}_{t+1} \mid \mathbf{e}_{1:t}, \mathbf{e}_{t+1}) \quad \text{(dividing up the evidence)} \\
&= \alpha\, \mathbf{P}(\mathbf{e}_{t+1} \mid \mathbf{X}_{t+1}, \mathbf{e}_{1:t})\, \mathbf{P}(\mathbf{X}_{t+1} \mid \mathbf{e}_{1:t}) \quad \text{(using Bayes' rule)} \\
&= \alpha\, \underline{\mathbf{P}(\mathbf{e}_{t+1} \mid \mathbf{X}_{t+1})}\, \mathbf{P}(\mathbf{X}_{t+1} \mid \mathbf{e}_{1:t}) \quad \text{(by the sensor Markov assumption).}
\end{aligned}$$

Get from CPT

Sensor Markov assumption:
the probability of observing $\mathbf{E}_t$ depends only on the state $\mathbf{X}_t$

Inference: Filtering

- We have observed $\mathbf{e}_1, \ldots, \mathbf{e}_{t+1} = \mathbf{e}_{1:t+1}$. We wish to calculate

$$\boxed{\mathbf{P}(\mathbf{X}_{t+1} \mid \mathbf{e}_{1:t+1})}$$

$$\begin{aligned}
\mathbf{P}(\mathbf{X}_{t+1} \mid \mathbf{e}_{1:t+1}) &= \mathbf{P}(\mathbf{X}_{t+1} \mid \mathbf{e}_{1:t}, \mathbf{e}_{t+1}) \quad \text{(dividing up the evidence)} \\
&= \alpha\, \mathbf{P}(\mathbf{e}_{t+1} \mid \mathbf{X}_{t+1}, \mathbf{e}_{1:t})\, \mathbf{P}(\mathbf{X}_{t+1} \mid \mathbf{e}_{1:t}) \quad \text{(using Bayes' rule)} \\
&= \alpha\, \underline{\mathbf{P}(\mathbf{e}_{t+1} \mid \mathbf{X}_{t+1})}\, \mathbf{P}(\mathbf{X}_{t+1} \mid \mathbf{e}_{1:t}) \quad \text{(by the sensor Markov assumption).}
\end{aligned}$$

?

Get from CPT

Inference: Filtering

$\mathbf{P}(\mathbf{X}_{t+1} \mid \mathbf{e}_{1:t})$ =?

Set A=$\mathbf{X}_{t+1,}$ B = $\mathbf{e}_{1:t,}$ C=$\mathbf{e}_{t+1,}$ D= $\mathbf{X}_t$

$$\mathbf{P}(\mathbf{X}_{t+1} \mid \mathbf{e}_{1:t}) = P(A \mid B)$$

$$= \frac{P(A,B)}{P(B)} = \frac{\sum_D P(A,B,D)}{P(B)} = \frac{\sum_D P(A|B,D)P(D|B)P(B)}{P(B)}$$

$$= \frac{P(B)\sum_D P(A|B,D)P(D|B)}{P(B)}$$

$$= \sum_D P(A|B,D)P(D|B)$$

$$= \sum_{\mathbf{X}_t} P(\mathbf{X}_{t+1}|\mathbf{e}_{1:t}, \mathbf{X}_t)P(\mathbf{X}_t|\mathbf{e}_{1:t})$$

One-step prediction

Inference: Filtering

- We have observed $\mathbf{e}_1, \ldots, \mathbf{e}_{t+1} = \mathbf{e}_{1:t+1}$. We wish to calculate

$$\boxed{\mathbf{P}(\mathbf{X}_{t+1} \mid \mathbf{e}_{1:t+1})}$$

$$\begin{aligned}
\mathbf{P}(\mathbf{X}_{t+1} \mid \mathbf{e}_{1:t+1}) &= \mathbf{P}(\mathbf{X}_{t+1} \mid \mathbf{e}_{1:t}, \mathbf{e}_{t+1}) \quad \text{(dividing up the evidence)} \\
&= \alpha\, \mathbf{P}(\mathbf{e}_{t+1} \mid \mathbf{X}_{t+1}, \mathbf{e}_{1:t})\, \mathbf{P}(\mathbf{X}_{t+1} \mid \mathbf{e}_{1:t}) \quad \text{(using Bayes' rule)} \\
&= \alpha\, \mathbf{P}(\mathbf{e}_{t+1} \mid \mathbf{X}_{t+1})\, \mathbf{P}(\mathbf{X}_{t+1} \mid \mathbf{e}_{1:t}) \quad \text{(by the sensor Markov assumption).} \\
&= \alpha\, \mathbf{P}(\mathbf{e}_{t+1} \mid \mathbf{X}_{t+1}) \sum_{\mathbf{x}_t} \mathbf{P}(\mathbf{X}_{t+1} \mid \mathbf{x}_t, \mathbf{e}_{1:t}) P(\mathbf{x}_t \mid \mathbf{e}_{1:t}) \\
&= \alpha\, \mathbf{P}(\mathbf{e}_{t+1} \mid \mathbf{X}_{t+1}) \sum_{\mathbf{x}_t} \mathbf{P}(\mathbf{X}_{t+1} \mid \mathbf{x}_t) P(\mathbf{x}_t \mid \mathbf{e}_{1:t}) \quad \text{(Markov assumption).}
\end{aligned}$$

First order Markov assumption: the present state depends only on the immediate previous state.

Inference: Filtering

- We have observed $\mathbf{e}_1$, ..., $\mathbf{e}_{t+1}$ = $\mathbf{e}_{1:t+1}$. We wish to calculate

$$\boxed{\mathbf{P}(\mathbf{X}_{t+1} \mid \mathbf{e}_{1:t+1})}$$

$$\mathbf{P}(\mathbf{X}_{t+1} \mid \mathbf{e}_{1:t+1}) = \mathbf{P}(\mathbf{X}_{t+1} \mid \mathbf{e}_{1:t}, \mathbf{e}_{t+1}) \quad \text{(dividing up the evidence)}$$

$$= \alpha\, \mathbf{P}(\mathbf{e}_{t+1} \mid \mathbf{X}_{t+1}, \mathbf{e}_{1:t})\, \mathbf{P}(\mathbf{X}_{t+1} \mid \mathbf{e}_{1:t}) \quad \text{(using Bayes' rule)}$$

$$= \alpha\, \mathbf{P}(\mathbf{e}_{t+1} \mid \mathbf{X}_{t+1})\, \mathbf{P}(\mathbf{X}_{t+1} \mid \mathbf{e}_{1:t}) \quad \text{(by the sensor Markov assumption).}$$

$$= \alpha\, \mathbf{P}(\mathbf{e}_{t+1} \mid \mathbf{X}_{t+1}) \sum_{\mathbf{x}_t} \mathbf{P}(\mathbf{X}_{t+1} \mid \mathbf{x}_t, \mathbf{e}_{1:t}) P(\mathbf{x}_t \mid \mathbf{e}_{1:t})$$

$$= \alpha\, \underbrace{\mathbf{P}(\mathbf{e}_{t+1} \mid \mathbf{X}_{t+1})}_{\text{Observation model}} \sum_{\mathbf{x}_t} \underbrace{\mathbf{P}(\mathbf{X}_{t+1} \mid \mathbf{x}_t)}_{\text{Transition model}} P(\mathbf{x}_t \mid \mathbf{e}_{1:t}) \quad \text{(Markov assumption).}$$

Inference: Filtering

- We have observed $\mathbf{e}_1, \ldots, \mathbf{e}_{t+1} = \mathbf{e}_{1:t+1}$. We wish to calculate

$$\boxed{\mathbf{P}(\mathbf{X}_{t+1} \mid \mathbf{e}_{1:t+1})}$$

$$
\begin{aligned}
\boxed{\mathbf{P}(\mathbf{X}_{t+1} \mid \mathbf{e}_{1:t+1})} &= \mathbf{P}(\mathbf{X}_{t+1} \mid \mathbf{e}_{1:t}, \mathbf{e}_{t+1}) \quad \text{(dividing up the evidence)} \\
&= \alpha\, \mathbf{P}(\mathbf{e}_{t+1} \mid \mathbf{X}_{t+1}, \mathbf{e}_{1:t})\, \mathbf{P}(\mathbf{X}_{t+1} \mid \mathbf{e}_{1:t}) \quad \text{(using Bayes' rule)} \\
&= \alpha\, \mathbf{P}(\mathbf{e}_{t+1} \mid \mathbf{X}_{t+1})\, \mathbf{P}(\mathbf{X}_{t+1} \mid \mathbf{e}_{1:t}) \quad \text{(by the sensor Markov assumption).} \\
&= \alpha\, \mathbf{P}(\mathbf{e}_{t+1} \mid \mathbf{X}_{t+1}) \sum_{\mathbf{x}_t} \mathbf{P}(\mathbf{X}_{t+1} \mid \mathbf{x}_t, \mathbf{e}_{1:t}) P(\mathbf{x}_t \mid \mathbf{e}_{1:t}) \\
&= \alpha\, \mathbf{P}(\mathbf{e}_{t+1} \mid \mathbf{X}_{t+1}) \sum_{\mathbf{x}_t} \mathbf{P}(\mathbf{X}_{t+1} \mid \mathbf{x}_t) \boxed{P(\mathbf{x}_t \mid \mathbf{e}_{1:t})} \quad \text{(Markov assumption).}
\end{aligned}
$$

Has the same form! but at one time step before.

This process is called recursive estimation.

Inference: Filtering

- We have observed $\mathbf{e}_1$, ..., $\mathbf{e}_{t+1} = \mathbf{e}_{1:t+1}$. We wish to calculate

$$\boxed{\mathbf{P}(\mathbf{X}_{t+1} \mid \mathbf{e}_{1:t+1})}$$

$$\mathbf{P}(\mathbf{X}_{t+1} \mid \mathbf{e}_{1:t+1}) = \mathbf{P}(\mathbf{X}_{t+1} \mid \mathbf{e}_{1:t}, \mathbf{e}_{t+1}) \quad \text{(dividing up the evidence)}$$

$$= \alpha\, \mathbf{P}(\mathbf{e}_{t+1} \mid \mathbf{X}_{t+1}, \mathbf{e}_{1:t})\, \mathbf{P}(\mathbf{X}_{t+1} \mid \mathbf{e}_{1:t}) \quad \text{(using Bayes' rule)}$$

$$= \alpha\, \mathbf{P}(\mathbf{e}_{t+1} \mid \mathbf{X}_{t+1})\, \mathbf{P}(\mathbf{X}_{t+1} \mid \mathbf{e}_{1:t}) \quad \text{(by the sensor Markov assumption).}$$

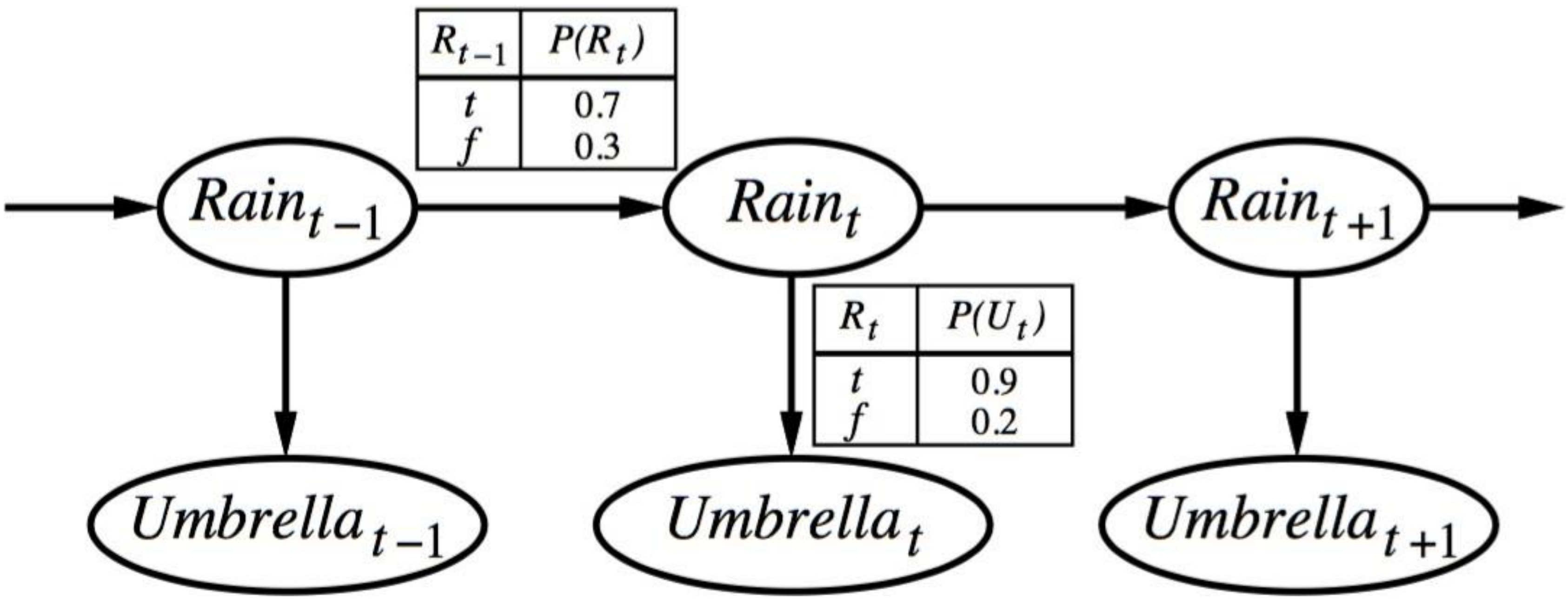
$$= \alpha\, \mathbf{P}(\mathbf{e}_{t+1} \mid \mathbf{X}_{t+1}) \sum_{\mathbf{x}_t} \mathbf{P}(\mathbf{X}_{t+1} \mid \mathbf{x}_t, \mathbf{e}_{1:t}) P(\mathbf{x}_t \mid \mathbf{e}_{1:t})$$

$$= \underbrace{\alpha\, \mathbf{P}(\mathbf{e}_{t+1} \mid \mathbf{X}_{t+1}) \sum_{\mathbf{x}_t} \mathbf{P}(\mathbf{X}_{t+1} \mid \mathbf{x}_t) P(\mathbf{x}_t \mid \mathbf{e}_{1:t})} \quad \text{(Markov assumption).}$$

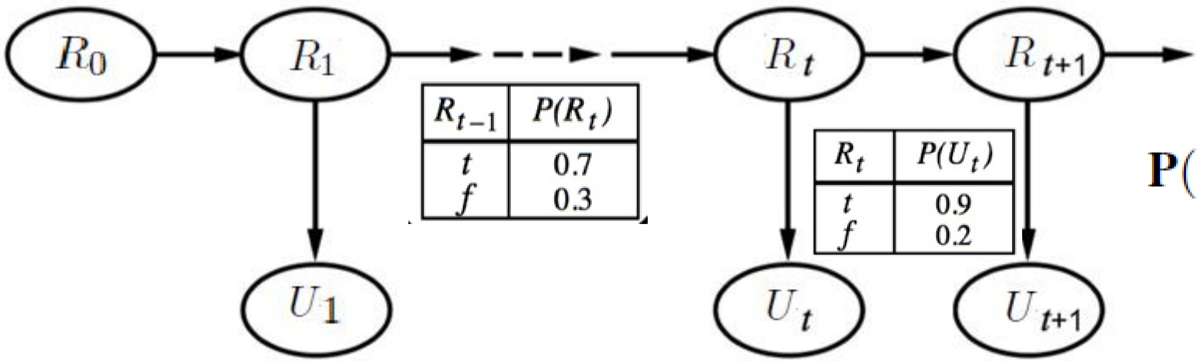
Combing the prediction with the new evidence is called update.

Example: is it raining outside?

- Form it as a first-order Markov process:



Example: is it raining outside?



R_{t-1}	$P(R_t)$
t	0.7
f	0.3

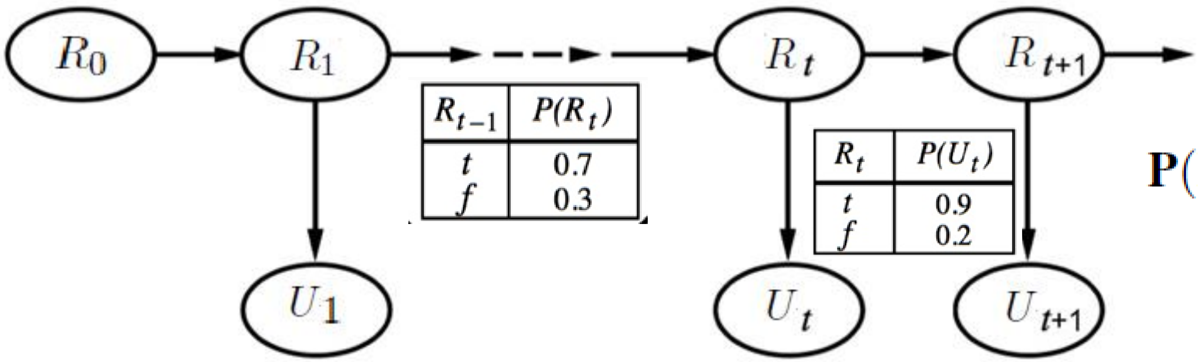
R_t	$P(U_t)$
t	0.9
f	0.2

$$\begin{aligned}\mathbf{P}(\mathbf{X}_{t+1} \mid \mathbf{e}_{1:t+1}) \\ &= \alpha\, \mathbf{P}(\mathbf{e}_{t+1} \mid \mathbf{X}_{t+1}) \\ &\quad \sum_{\mathbf{x}_t} \mathbf{P}(\mathbf{X}_{t+1} \mid \mathbf{x}_t) P(\mathbf{x}_t \mid \mathbf{e}_{1:t})\end{aligned}$$

- On day 0, we have no observations, only the security guard's prior beliefs; let's assume that consists of $\mathbf{P}(R_0) = \langle 0.5, 0.5 \rangle$.
- On day 1, the umbrella appears, so $U_1 = true$.

$$\begin{aligned}\mathbf{P}(R_1 \mid u_1) &= \alpha\, \mathbf{P}(u_1 \mid R_1) \sum_{r_0} \mathbf{P}(R_1 \mid r_0) P(r_0) \\ &= \alpha \langle 0.9, 0.2 \rangle \Big(\langle 0.7, 0.3 \rangle \times 0.5 + \langle 0.3, 0.7 \rangle \times 0.5 \Big) \\ &= \alpha \langle 0.9, 0.2 \rangle \langle 0.5, 0.5 \rangle \\ &= \alpha \langle 0.45, 0.1 \rangle \approx \langle 0.818, 0.182 \rangle\end{aligned}$$

Example: is it raining outside?



$$\mathbf{P}(\mathbf{X}_{t+1} \mid \mathbf{e}_{1:t+1}) = \alpha\, \mathbf{P}(\mathbf{e}_{t+1} \mid \mathbf{X}_{t+1}) \sum_{\mathbf{x}_t} \mathbf{P}(\mathbf{X}_{t+1} \mid \mathbf{x}_t) P(\mathbf{x}_t \mid \mathbf{e}_{1:t})$$

- On day 2, the umbrella appears, so $U_2 = true$.

$$\begin{aligned}
\mathbf{P}(R_2 \mid u_1, u_2) &= \alpha\, \mathbf{P}(u_2 \mid R_2) \sum_{r_1} \mathbf{P}(R_2 \mid r_1) P(r_1 \mid u_1) \\
&= \alpha\, \langle 0.9, 0.2 \rangle \big(\langle 0.7, 0.3 \rangle \times 0.818 + \langle 0.3, 0.7 \rangle \times 0.182 \big) \\
&= \alpha\, \langle 0.9, 0.2 \rangle \langle 0.627, 0.373 \rangle \\
&= \alpha\, \langle 0.565, 0.075 \rangle \approx \langle 0.883, 0.117 \rangle
\end{aligned}$$

Can keep on going as new observations are made.

Hidden Markov Model (HMM)

- A HMM is obtained if X_t and E_t for all t are single discrete random variables.

  e.g., "is it raining outside?" is a HMM.
- In a HMM, the **transition model** can be encoded in an SxS matrix

$$\mathbf{T} = \begin{bmatrix} \mathbf{T}[1,1] & \mathbf{T}[1,2] & \cdots & \mathbf{T}[1,S] \\ \mathbf{T}[2,1] & \cdots & \cdots & \vdots \\ \vdots & \cdots & \cdots & \vdots \\ \mathbf{T}[S,1] & \cdots & \cdots & \mathbf{T}[S,S] \end{bmatrix}$$

where S is the number of possible values of X_t, and

$$\mathbf{T}[i, j] = P(X_t = j \mid X_{t-1} = i)$$

Hidden Markov Model (HMM)

- Given evidence e_t at time step t, the observation model can be encoded in an SxS diagonal matrix

$$\mathbf{O}_t = \begin{bmatrix} \mathbf{O}_t[1,1] & 0 & \cdots & 0 \\ 0 & \mathbf{O}_t[2,2] & \cdots & \vdots \\ \vdots & \cdots & \mathbf{O}_t[3,3] & 0 \\ 0 & \cdots & 0 & \mathbf{O}_t[S,S] \end{bmatrix}$$

where

$$\mathbf{O}_t[\mathrm{i}, \mathrm{i}] = P(e_t \mid X_t = i)$$

Hidden Markov Model (HMM)

- Recursive estimation in HMM can be computed as

$$\mathbf{f}_{t+1} = \alpha \mathbf{O}_{t+1} \mathbf{T}^{T} \mathbf{f}_{t}$$

where

$\mathbf{f}_{t+1}$ is column vector form of $P(\mathrm{X}_{t+1} \mid \mathrm{e}_{1:t+1})$

$\mathbf{f}_{t}$ is column vector form of $P(\mathrm{X}_{t} \mid \mathrm{e}_{1:t})$

Hidden Markov Model (HMM)

- Recursive estimation in HMM can be computed as

$$\mathbf{f}_{t+1} = \alpha \mathbf{O}_{t+1} \mathbf{T}^{T} \mathbf{f}_t$$

$$\begin{bmatrix} \mathbf{f}_{t+1}[1] \\ \\ \\ \mathbf{f}_{t+1}[S] \end{bmatrix} = \alpha \begin{bmatrix} \mathbf{O}_{t+1}[1,1] & 0 & & 0 \\ 0 & \mathbf{O}_{t+1}[2,2] & & \vdots \\ \vdots & \cdots & \cdots & 0 \\ 0 & \cdots & & \mathbf{O}_{t+1}[S,S] \end{bmatrix} \begin{bmatrix} \mathbf{T}[1,1] & \mathbf{T}[2,1] \cdots & \mathbf{T}[S,1] \\ \mathbf{T}[1,2] & \cdots & \vdots \\ \vdots & & \\ \mathbf{T}[1,S] & & \mathbf{T}[S,S] \end{bmatrix} \begin{bmatrix} \mathbf{f}_t[1] \\ \\ \\ \mathbf{f}_t[S] \end{bmatrix}$$

$$\mathbf{f}_{t+1}[i] = \alpha \mathbf{O}_t[\mathrm{i, i}] \sum_j \mathbf{T}[j,i] \mathbf{f}_{t+}[i]$$

$$\mathbf{P}(\mathbf{X}_{t+1} \mid \mathbf{e}_{1:t+1}) = \alpha \mathbf{P}(\mathbf{e}_{t+1} \mid \mathbf{X}_{t+1}) \sum_{\mathbf{x}_t} \mathbf{P}(\mathbf{X}_{t+1} \mid \mathbf{x}_t) P(\mathbf{x}_t \mid \mathbf{e}_{1:t})$$

Example: is it raining outside?

- From transition model:

$$\mathbf{T} = \begin{bmatrix} 0.7 & 0.3 \\ 0.3 & 0.7 \end{bmatrix}$$

R_{t-1}	$P(R_t)$
t	0.7
f	0.3

$\mathbf{T}[1,1]=P(R_t=true \mid R_{t-1}=true)$ $\mathbf{T}[1,2]=P(R_t=true \mid R_{t-1}=false)$

$\mathbf{T}[2,1]=P(R_t=false \mid R_{t-1}=true)$ $\mathbf{T}[2,2]=P(R_t=false \mid R_{t-1}=false)$

- At $t=0$, rain or no rain is equally likely, so

$$\mathbf{f}_0 = \begin{bmatrix} 0.5 \\ 0.5 \end{bmatrix}$$

Example: is it raining outside?

- From transition model:

$$\mathbf{T} = \begin{bmatrix} 0.7 & 0.3 \\ 0.3 & 0.7 \end{bmatrix}$$

R_{t-1}	$P(R_t)$
t	0.7
f	0.3

- At $t=0$, rain or no rain is equally likely, so

$$\mathbf{f}_0 = \begin{bmatrix} 0.5 \\ 0.5 \end{bmatrix}$$

- At $t=1$, umbrella is observed, so

$$\mathbf{O}_1 = \begin{bmatrix} 0.9 & 0 \\ 0 & 0.2 \end{bmatrix}$$

R_t	$P(U_t)$
t	0.9
f	0.2

$\mathbf{O}[1,1]=P(E_t=true|\ R_{t-1}=true)$ $\quad$ $\mathbf{O}[2,2]=P(E_t=true|\ R_{t-1}=false)$

Example: is it raining outside?

- From transition model:

$$\mathbf{T} = \begin{bmatrix} 0.7 & 0.3 \\ 0.3 & 0.7 \end{bmatrix}$$

R_{t-1}	$P(R_t)$
t	0.7
f	0.3

- At $t=0$, rain or no rain is equally likely, so

$$\mathbf{f}_0 = \begin{bmatrix} 0.5 \\ 0.5 \end{bmatrix}$$

- At $t=1$, umbrella is observed, so

$$\mathbf{O}_1 = \begin{bmatrix} 0.9 & 0 \\ 0 & 0.2 \end{bmatrix}$$

R_t	$P(U_t)$
t	0.9
f	0.2

- Filtering result at $t=1$ is

$$\mathbf{f}_1 = \alpha \begin{bmatrix} 0.9 & 0 \\ 0 & 0.2 \end{bmatrix} \begin{bmatrix} 0.7 & 0.3 \\ 0.3 & 0.7 \end{bmatrix}^T \begin{bmatrix} 0.5 \\ 0.5 \end{bmatrix} = \alpha \begin{bmatrix} 0.45 \\ 0.1 \end{bmatrix} \approx \begin{bmatrix} 0.818 \\ 0.182 \end{bmatrix}$$

Example: is it raining outside?

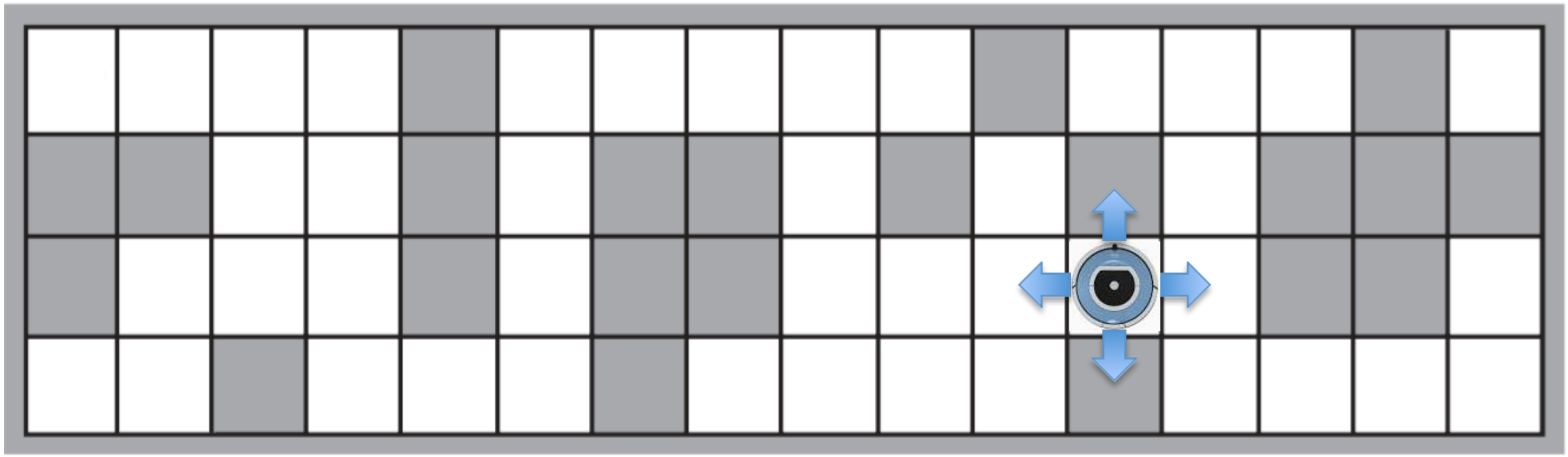
- At t=2, umbrella is observed, so

$$\mathbf{O}_2 = \begin{bmatrix} 0.9 & 0 \\ 0 & 0.2 \end{bmatrix}$$

- Filtering result at t=2 is

$$\mathbf{f}_2 = \alpha \begin{bmatrix} 0.9 & 0 \\ 0 & 0.2 \end{bmatrix} \begin{bmatrix} 0.7 & 0.3 \\ 0.3 & 0.7 \end{bmatrix}^T \begin{bmatrix} 0.45 \\ 0.1 \end{bmatrix} = \alpha \begin{bmatrix} 0.3105 \\ 0.041 \end{bmatrix} \approx \begin{bmatrix} 0.883 \\ 0.117 \end{bmatrix}$$

Example: robot localisation

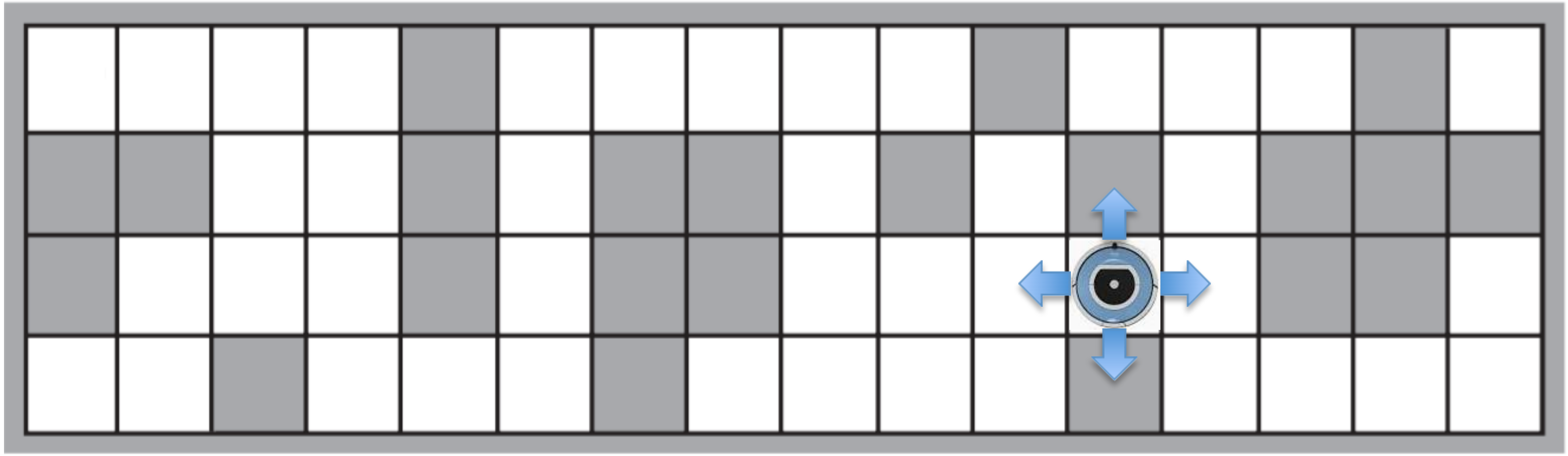


- A robot is moving in a 2D map with obstacles. Its sensor detects the directions are blocked by obstacles at the robot's location.
- A sequence of sensor readings is observed:

Time step	1	2	3	…	t	t+1
Blocked directions	N S W	N S	N	…	S E W	S E

- At time step t+1, where is the robot?

Example: robot localisation



- State variable represents the robot location:

$$X_t \in \{1, 2, \ldots, S\}$$

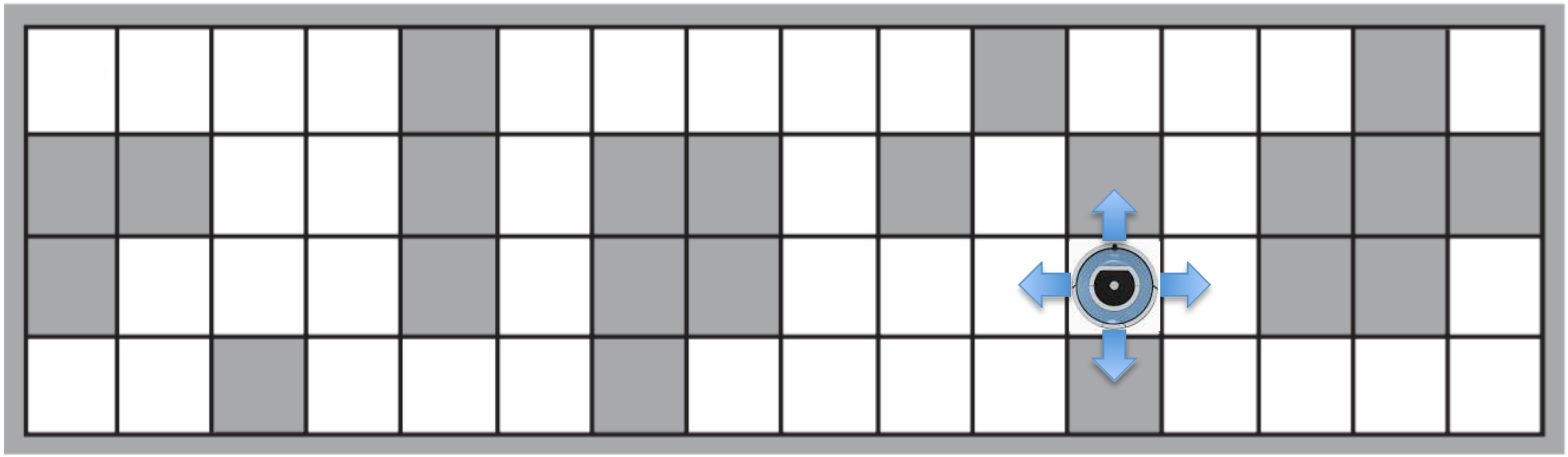
S=42=64 squares-22 blocked

- Sensor reading $E_t = e_t$: observed obstacles.

Time step	1	2	3	…	*t*	*t*+1
Blocked directions	N S W	N S	N	…	S E W	S E

E_t has 16 possible values

Example: robot localisation



- Assuming random walk, the **transition model** is

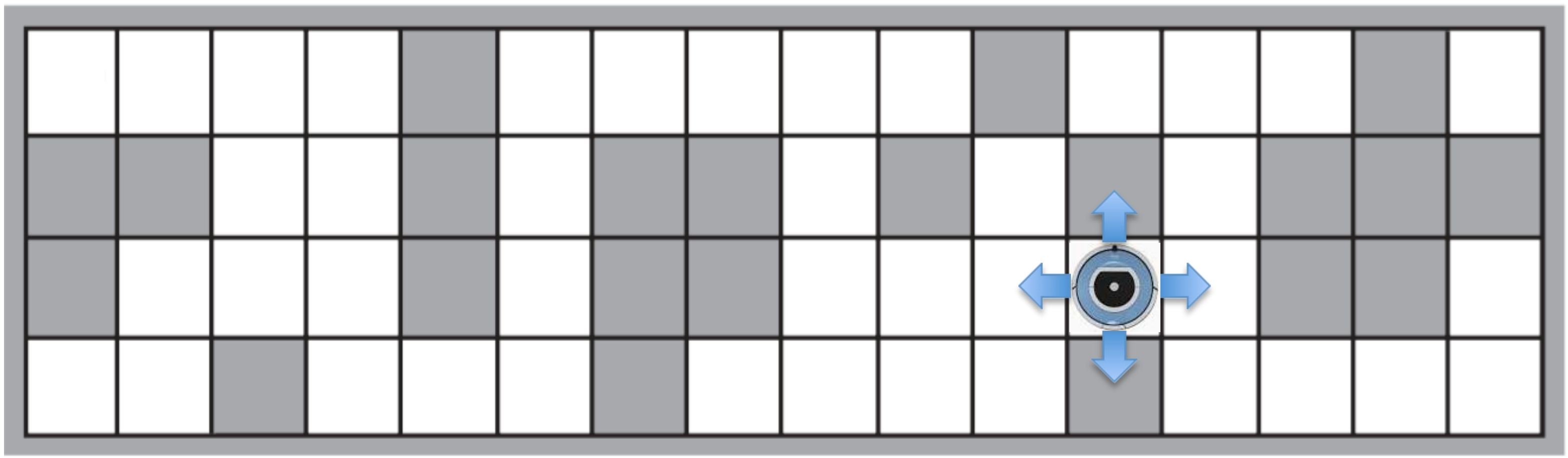
$$P(X_{t+1}=j \mid X_t=i) = \mathbf{T}_{ij} = \begin{cases} 1/N(i) & \text{if } j \in \text{NEIGHBOURS(i)} \\ 0 & \text{otherwise} \end{cases}$$

where $\text{NEIGHBOURS}(i)$ = set of empty neighbours of cell i.

N(i) = number of neighbours of cell i.

$\mathbf{T}$ has 42×42=1764 entries

Example: robot localisation



- The sensor's error rate is ϵ and error occurs independently in the four directions. This gives the **observation model**

$$P(E_t = e_t \mid X_t = i) = (1-\epsilon)^{4-d_{it}} \epsilon^{d_{it}}$$

where d_{it} is the number of directions that are wrong given location $X_t = i$ and sensor reading $E_t = e_t$

- Example: at the robot's position in the map above, the probability of observing

$e_t =$ *NSW* is $(1-\epsilon)^3 \epsilon^1$

Example: robot localisation

- Assume the robot is equally likely to be at any square at $t = 0$, i.e., $\mathbf{f}_0$ is uniform. $P(X_0 = i) = 1/n$

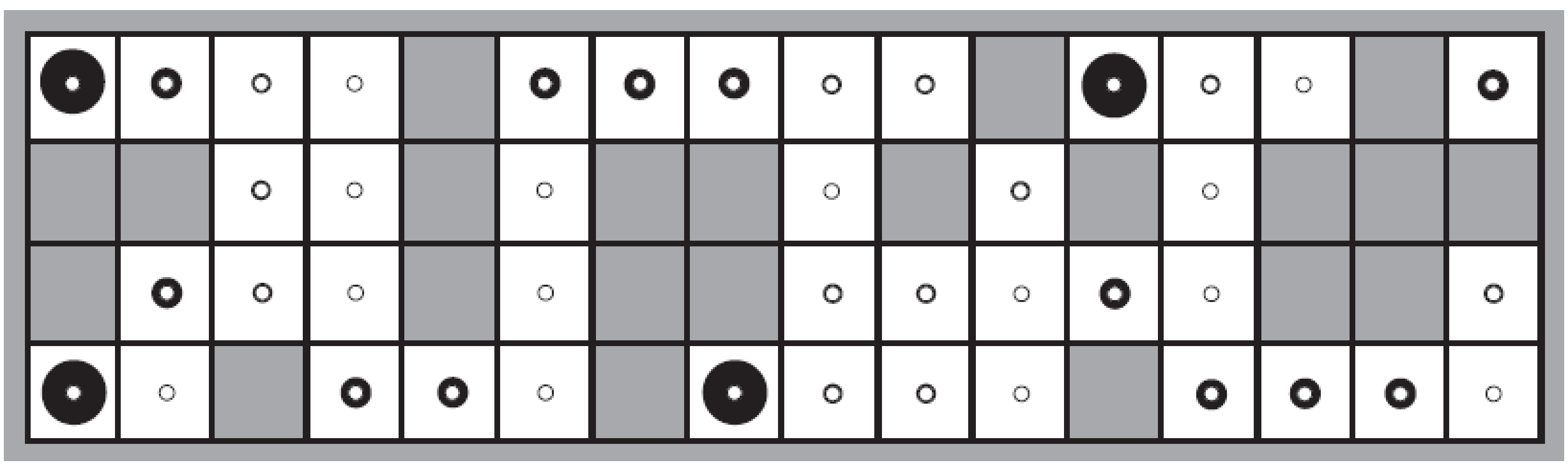
$$\mathbf{f}_0 = \underset{42 \times 1}{\begin{bmatrix} 1/42 \\ 1/42 \\ \cdots \\ \\ \\ 1/42 \end{bmatrix}}$$

Example: robot localisation

- Assume the robot is equally likely to be at any square at $t = 0$, i.e., $\mathbf{f}_0$ is uniform. $P(X_0 = i) = 1/n$

- After observing $E_1 = NSW$,

$$\mathbf{f}_1 = P(X_1 \mid E_1 = NSW) = \alpha \mathbf{O}_1 \mathbf{T}^T \mathbf{f}_0$$

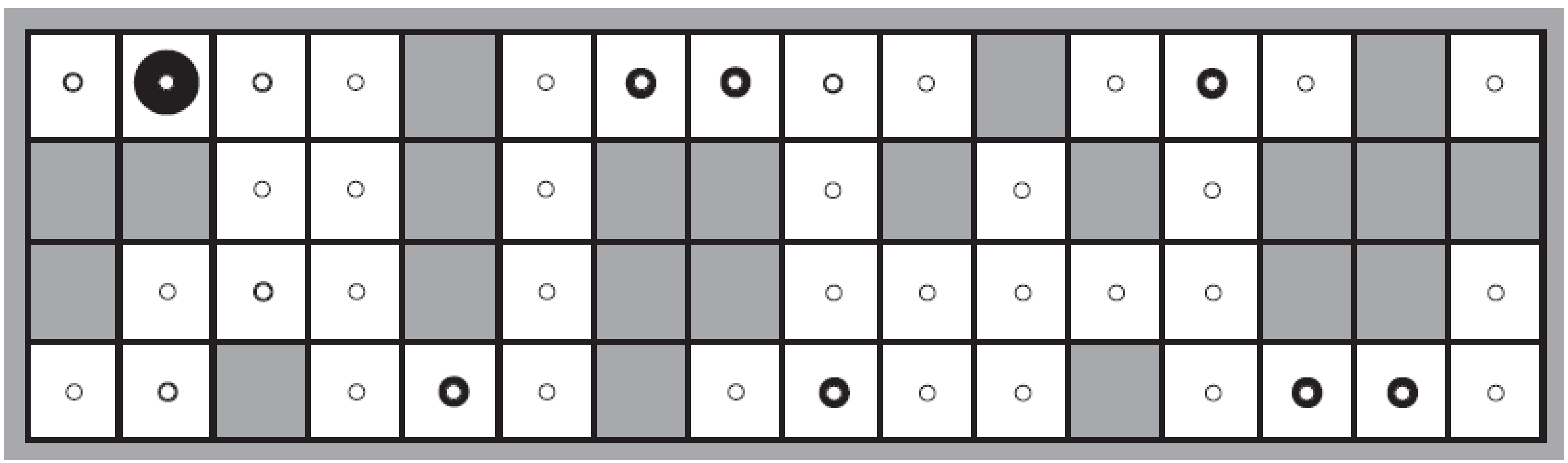


Posterior distribution over robot location after E_1= NSW $\epsilon = 0.2$

Example: robot localisation

- After observing $E_2 = NS$,

$$\mathbf{f}_2 = P(X_2 \mid E_1 = NSW, E_2 = NS) = \alpha \mathbf{O}_2 \mathbf{T}^T \mathbf{f}_1$$



Posterior distribution over robot location after E_1= NSW, E_2= NS