Markov Models and Hidden Markov Models

Robert Platt
Northeastern University

Some images and slides are used from:
1. CS188 UC Berkeley
2. RN, AIMA

Markov Models

We have already seen that an MDP provides a useful framework for modeling stochastic control problems.

Markov Models: model any kind of temporally dynamic system.

Probability recap

- Conditional probability $$P(x|y) = \frac{P(x,y)}{P(y)}$$
- Product rule $$P(x,y) = P(x|y)P(y)$$
- Chain rule $$\begin{aligned} P(X_1, X_2, \dots X_n) &= P(X_1)P(X_2|X_1)P(X_3|X_1, X_2)\dots \\ &= \prod_{i=1}^{n} P(X_i|X_1, \dots, X_{i-1}) \end{aligned}$$
- X, Y independent if and only if: $$\forall x, y : P(x,y) = P(x)P(y)$$
- X and Y are conditionally independent given Z if and only if:

$$X \perp\!\!\!\perp Y | Z \qquad \forall x, y, z : P(x, y|z) = P(x|z)P(y|z)$$

Probability again: Independence

Two random variables, *x* and *y*, are independent when:

$$\forall(x,y), P(x,y) = P(x)P(y) \quad \Longleftrightarrow \quad x \perp\!\!\!\perp y$$

$$x \not\!\perp\!\!\!\perp y$$

The outcomes of two different coin flips are usually independent of each other

Probability again: Independence

If: $$P(x,y) = P(x)P(y)$$

Then: $$P(x) = P(x|y)$$

$$P(y) = P(y|x)$$

Why?

Probability again: Independence

Two random variables, *x* and *y*, are independent when:

$$\forall (x,y), P(x,y) = P(x)P(y) \iff x \perp\!\!\!\perp y$$

$$x \not\perp\!\!\!\perp y$$

The outcomes of two different coin flips are usually independent of each other

Example: Independence

	winter	!winter
snow	0.1	0.1
!snow	0.3	0.5

Example: Independence

	winter	!winter
snow	0.1	0.1
!snow	0.3	0.5

Are snow and winter independent variables?

Example: Independence

	winter	!winter
snow	0.1	0.1
!snow	0.3	0.5

Are snow and winter independent variables?

P(snow) = 0.2

P(winter) = 0.4

Example: Independence

	winter	!winter
snow	0.1	0.1
!snow	0.3	0.5

Are snow and winter independent variables?

P(snow) = 0.2

P(winter) = 0.4

What would the distribution look like if snow, winter were independent?

Conditional independence

Independence: $$\forall (x,y), P(x,y) = P(x)P(y)$$

$$x \perp\!\!\!\perp y$$

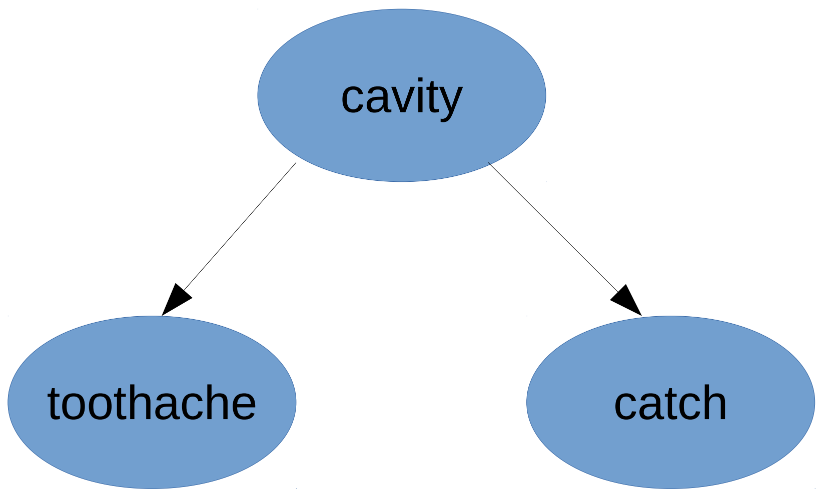
Conditional independence: $$\forall (x,y,z), P(x,y|z) = P(x|z)P(y|z)$$

$$x \perp\!\!\!\perp y | z$$

Equivalent statements of conditional independence: $$P(x|z) = P(x|z,y)$$

$$P(y|z) = P(y|z,x)$$

Conditional independence: example



P(toothache, catch | cavity) = P(toothache | cavity) P(catch | cavity)

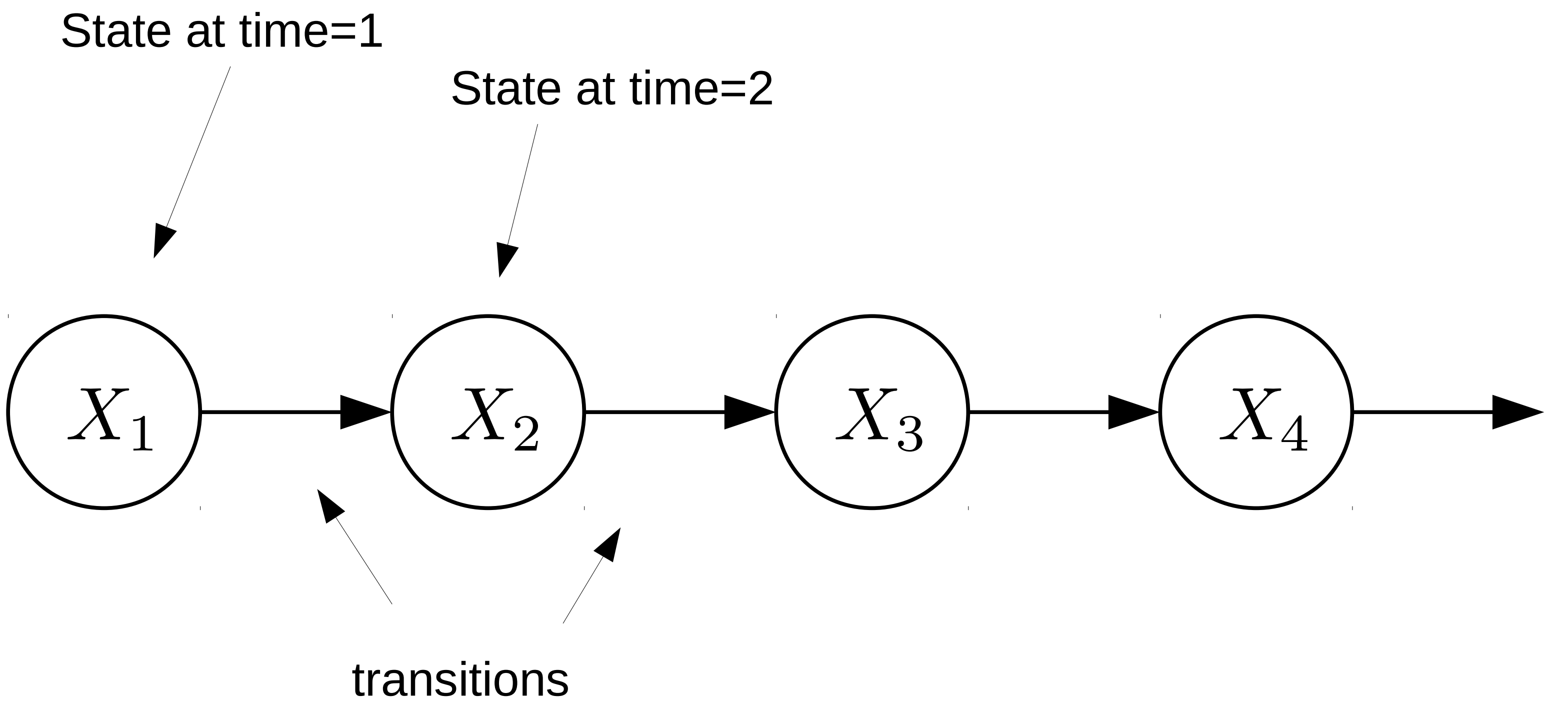
Toothache and catch are conditionally independent given cavity – this is the “common cause” scenario covered in Bayes Nets...

Examples of conditional independence

What are the conditional independence relationships in the following?

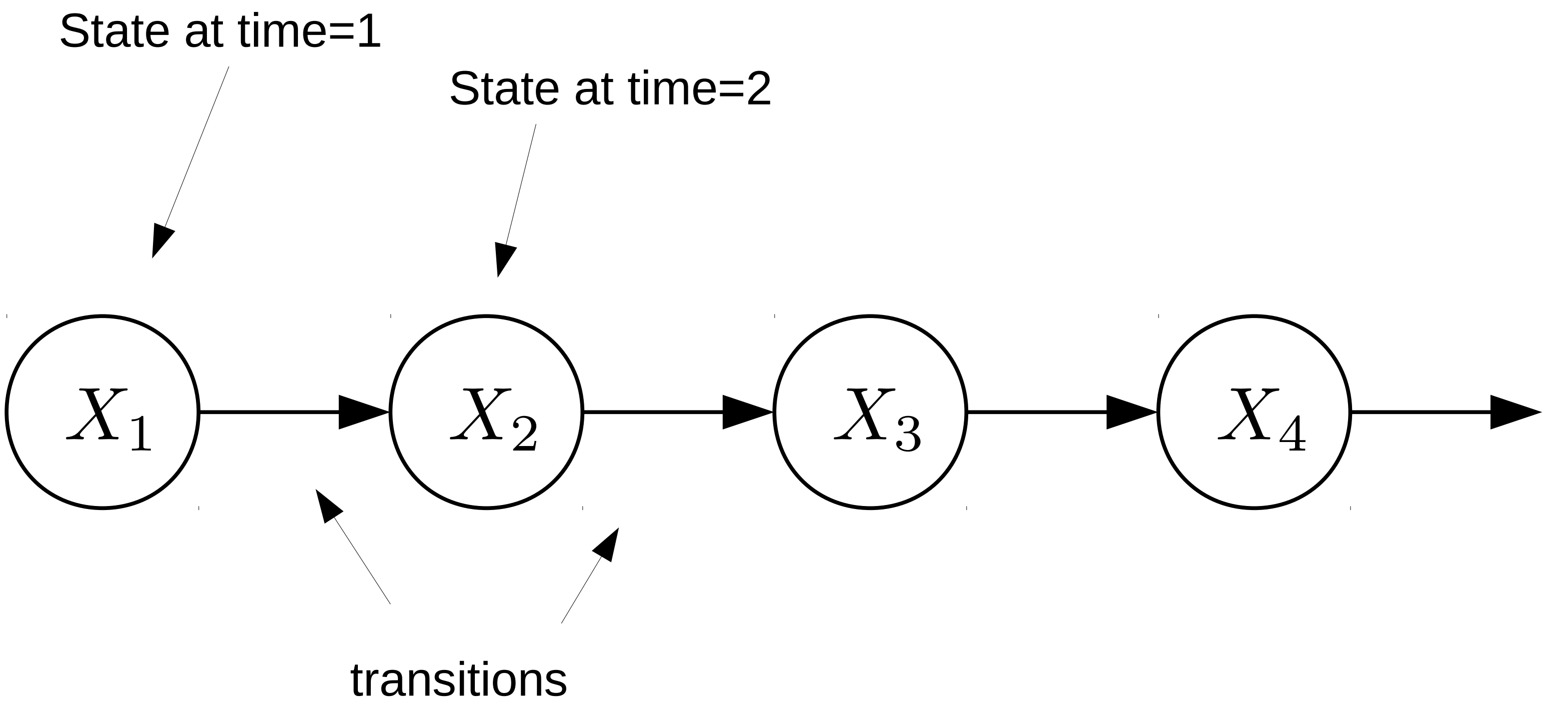
– traffic, raining, late for work
– snow, cloudy, crash
– fire, smoke, alarm

Markov Processes



Markov model can be used to model any sequential time process
– the weather
– traffic
– stock market
– news cycle
...

Markov Processes

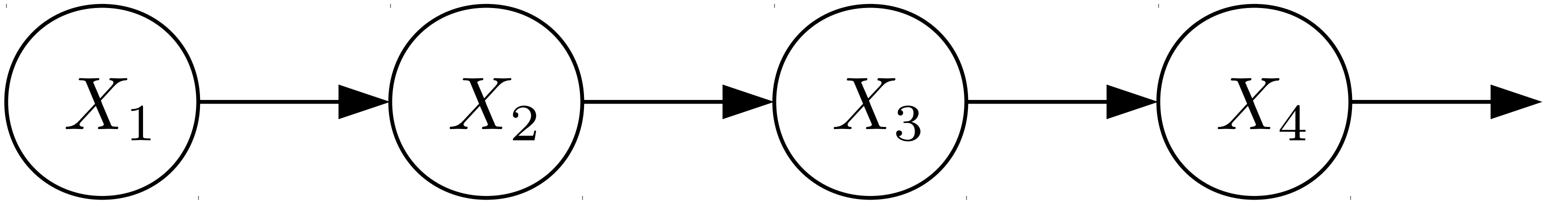


Since this is a Markov process, we assume transitions are Markov:

Process model: $$P(X_t|X_{t-1}) = P(X_t|X_{t-1}, \ldots, X_1)$$

Markov assumption: $$X_t \perp\!\!\!\perp X_{t-2}|X_{t-1}$$

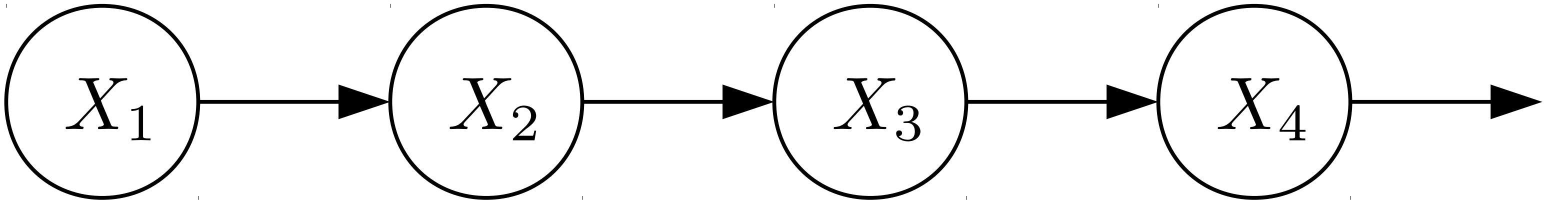
Markov Processes



How do we calculate: $P(X_1, X_2, X_3, X_4) = ?$

$$P(X_1, X_2, X_3, X_4) = P(X_1)P(X_2|X_1)P(X_3|X_2, X_1)P(X_4|X_3, X_2, X_1)$$

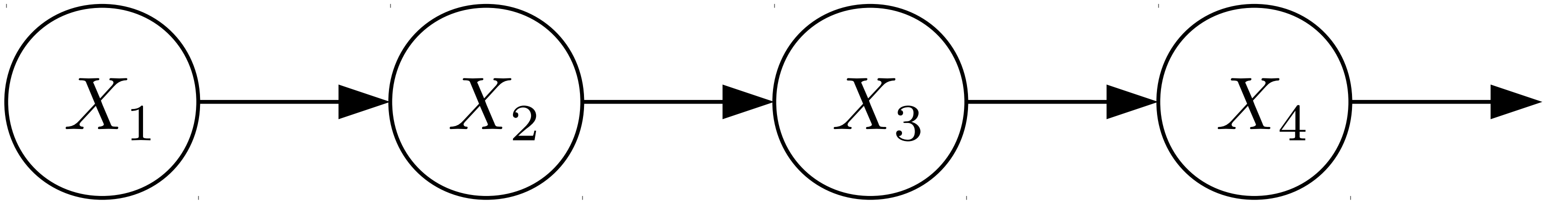
Markov Processes



How do we calculate: $P(X_1, X_2, X_3, X_4) = ?$

$$P(X_1, X_2, X_3, X_4) = \underbrace{P(X_1)P(X_2|X_1)}_{P(X_2, X_1)}P(X_3|X_2, X_1)P(X_4|X_3, X_2, X_1)$$

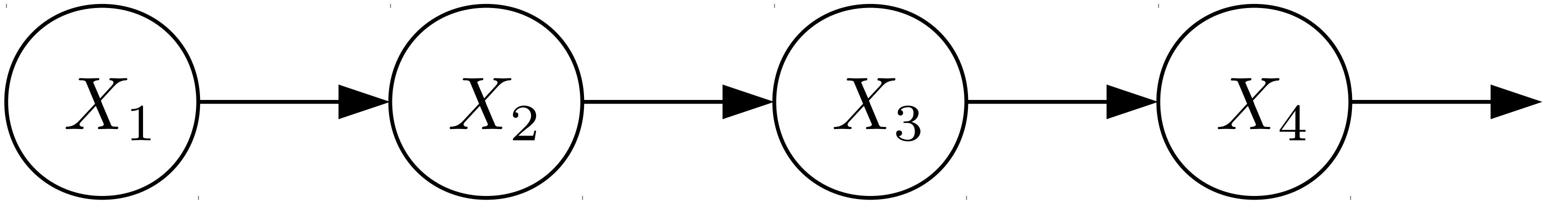
Markov Processes



How do we calculate: $P(X_1, X_2, X_3, X_4) =?$

$$P(X_1, X_2, X_3, X_4) = \underbrace{P(X_1)P(X_2|X_1)P(X_3|X_2, X_1)}_{P(X_3, X_2, X_1)}P(X_4|X_3, X_2, X_1)$$

Markov Processes

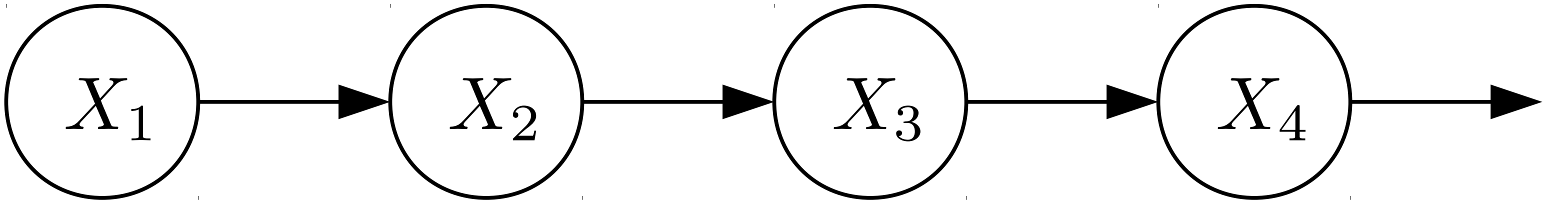


How do we calculate: $P(X_1, X_2, X_3, X_4) = ?$

$$P(X_1, X_2, X_3, X_4) = P(X_1)P(X_2|X_1)P(X_3|X_2, X_1)P(X_4|X_3, X_2, X_1)$$

Can we simplify this expression?

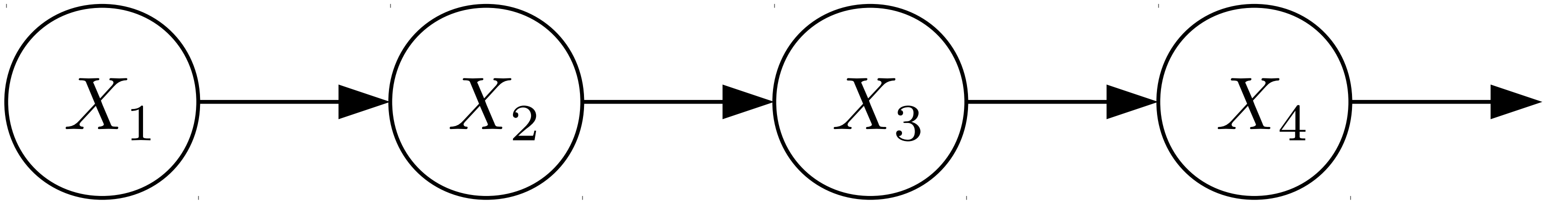
Markov Processes



How do we calculate: $P(X_1, X_2, X_3, X_4) = ?$

$$P(X_1, X_2, X_3, X_4) = P(X_1)P(X_2|X_1)\underbrace{P(X_3|X_2, X_1)}_{P(X_3|X_2)}\underbrace{P(X_4|X_3, X_2, X_1)}_{P(X_4|X_3)}$$

Markov Processes

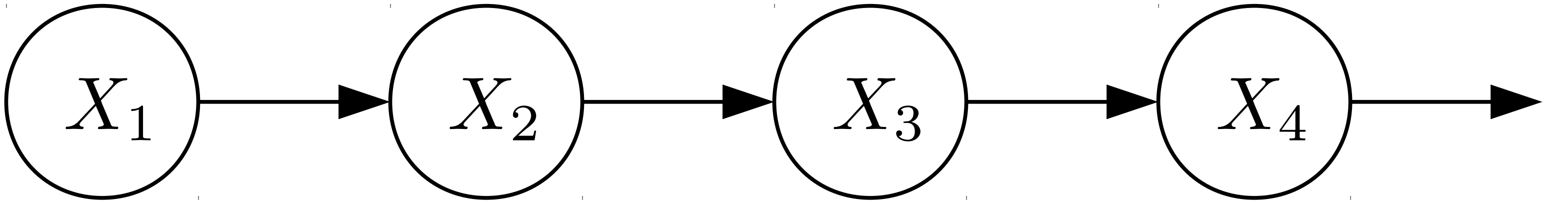


How do we calculate: $P(X_1, X_2, X_3, X_4) = ?$

$$P(X_1, X_2, X_3, X_4) = P(X_1)P(X_2|X_1)P(X_3|X_2, X_1)P(X_4|X_3, X_2, X_1)$$

$$P(X_1, X_2, X_3, X_4) = P(X_1)P(X_2|X_1)P(X_3|X_2)P(X_4|X_3)$$

Markov Processes



How do we calculate: $P(X_1, X_2, X_3, X_4) = ?$

$$P(X_1, X_2, X_3, X_4) = P(X_1)P(X_2|X_1)P(X_3|X_2, X_1)P(X_4|X_3, X_2, X_1)$$

$$P(X_1, X_2, X_3, X_4) = P(X_1)P(X_2|X_1)P(X_3|X_2)P(X_4|X_3)$$

In general: $P(X_1, X_2, \ldots, X_T) = P(X_1) \prod_{t=1}^{T-1} P(X_{t+1}|X_t)$

Markov Processes

$$X_1 \to X_2 \to X_3 \to X_4 \to$$

How do we calculate: $P(X_1, X_2, X_3, X_4) = ?$

$$P(X_1, X_2, X_3, X_4) = P(X_1)P(X_2|X_1)P(X \;\ldots\; X_2, X_1)$$

Process model

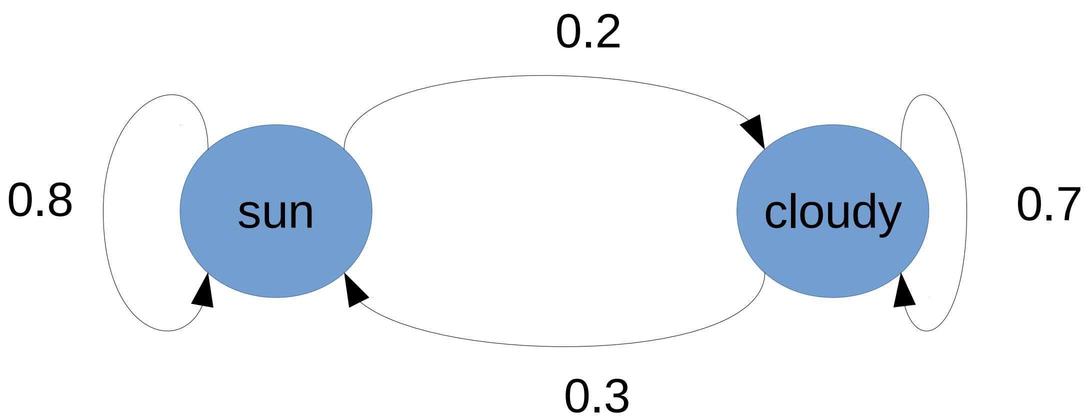
$$P(X_1, X_2, X_3, X_4) = P(X_1)P(X_2|X_1)P(X_3|X_2)P(X_4|X_3)$$

In general: $P(X_1, X_2, \ldots, X_T) = P(X_1) \prod_{t=1}^{T-1} \boxed{P(X_{t+1}|X_t)}$

Markov Processes: example

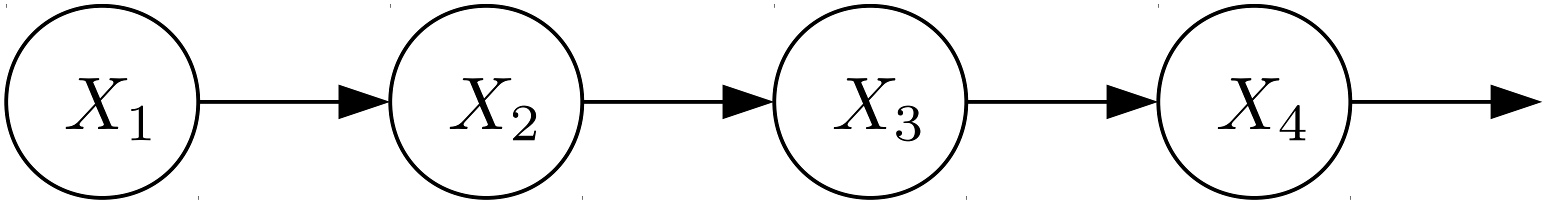
Two states: cloudy, sunny

X_{t-1}	X_t	X_t
sun	sun	0.8
sun	cloudy	0.2
cloudy	sun	0.3
cloudy	cloudy	0.7



$$t \in \{\text{mon}, \text{tues}, \text{weds}, \text{thurs}, \text{fri}\}$$

Simulating dynamics forward



Joint distribution: $P(X_1, X_2, \ldots, X_T) = P(X_1) \prod_{t=1}^{T-1} P(X_{t+1}|X_t)$

But, suppose we want to predict the state at time T, given a prior distribution at time 1?

$$P(X_2) = \sum_{X_1} P(X_1) P(X_2|X_1)$$

$$P(X_3) = \sum_{X_2} P(X_2) P(X_3|X_2)$$

$$\vdots$$

$$P(X_T) = \sum_{X_{T-1}} P(X_{T-1}) P(X_T|X_{T-1})$$

Simulating dynamics forward

Suppose is it sunny on mon... $P(x_1) = 1$

Prob sunny tues
$$\begin{aligned} P(x_2) &= P(x_2|x_1)P(x_1) \\ &= 0.8 \end{aligned}$$

Prob sunny weds
$$\begin{aligned} P(x_3) &= P(x_3|x_2)P(x_2) + P(x_3|\bar{x}_2)P(\bar{x}_2) \\ &= 0.64 + 0.06 = 0.7 \end{aligned}$$

Prob sunny thurs
$$\begin{aligned} P(x_4) &= P(x_4|x_3)P(x_3) + P(x_4|\bar{x}_3)P(\bar{x}_3) \\ &= 0.56 + 0.09 = 0.65 \end{aligned}$$

Prob sunny fri
$$\begin{aligned} P(x_5) &= P(x_5|x_4)P(x_4) + P(x_5|\bar{x}_4)P(\bar{x}_4) \\ &= 0.52 + 0.105 = 0.625 \end{aligned}$$

$$P(x_\infty) = 0.6$$

Simulating dynamics forward

Suppose is it cloudy on mon... $P(x_1) = 0$

Prob sunny tues
$$\begin{aligned} P(x_2) &= P(x_2|x_1)P(x_1) + P(x_2|\bar{x}_1)P(\bar{x}_1) \\ &= 0 + 0.3 = 0.3 \end{aligned}$$

Prob sunny weds
$$\begin{aligned} P(x_3) &= P(x_3|x_2)P(x_2) + P(x_3|\bar{x}_2)P(\bar{x}_2) \\ &= 0.24 + 0.21 = 0.45 \end{aligned}$$

Prob sunny thurs
$$\begin{aligned} P(x_4) &= P(x_4|x_3)P(x_3) + P(x_4|\bar{x}_3)P(\bar{x}_3) \\ &= 0.36 + 0.165 = 0.53 \end{aligned}$$

Prob sunny fri
$$\begin{aligned} P(x_5) &= P(x_5|x_4)P(x_4) + P(x_5|\bar{x}_4)P(\bar{x}_4) \\ &= 0.424 + 0.141 = 0.565 \end{aligned}$$

$$P(x_\infty) = 0.6$$

Simulating dynamics forward

Suppose is it cloudy on mon... $P(x_1) = 0$

Prob sunny tues

$$\begin{aligned} P(x_2) &= P(x_2|x_1)P(x_1) + P(x_2|\bar{x}_1)P(\bar{x}_1) \\ &= 0 + 0.3 = 0.3 \end{aligned}$$

Prob sunny weds

$$\begin{aligned} P(x_3) &= P(x_3|x_2)P(x_2) + P(x_3|\bar{x}_2)P(\bar{x}_2) \\ &= 0.24 + 0.21 = 0.45 \end{aligned}$$

Prob sunny thurs

$$\begin{aligned} P(x_4) &= P(x_4|x_3)P(x_3) + P(x_4|\bar{x}_3)P(\bar{x}_3) \\ &= 0.36 + 0.165 = 0.53 \end{aligned}$$

Prob sunny

$x_5|\bar{x}_4)P(\bar{x}_4)$

5

Converge to same distribution regardless of starting point
– called the "stationary distribution"

$$\boxed{P(x_\infty) = 0.6}$$

An aside: the stationary distribution

How might you calculate the stationary distribution?

Let: $p_t = \begin{pmatrix} p(sun) \\ p(cloudy) \end{pmatrix}$ $\quad T = \begin{pmatrix} 0.8 & 0.3 \\ 0.2 & 0.7 \end{pmatrix}$

Then: $p_{t+1} = Tp_t$

$$p_{t+n} = T^n p_t$$

Stationary distribution is the value for p such that: $p = Tp$

An aside: the stationary distribution

How might you calculate the stationary distribution?

Let: $p_t = \begin{pmatrix} p(sun) \\ p(cloudy) \end{pmatrix}$ $\quad T = \begin{pmatrix} 0.8 & 0.3 \\ 0.2 & 0.7 \end{pmatrix}$

Then: $p_{t+1} = Tp_t$

$$p_{t+n} = T^n p_t$$

Stationary distribution is the value for p such that: $p = Tp$

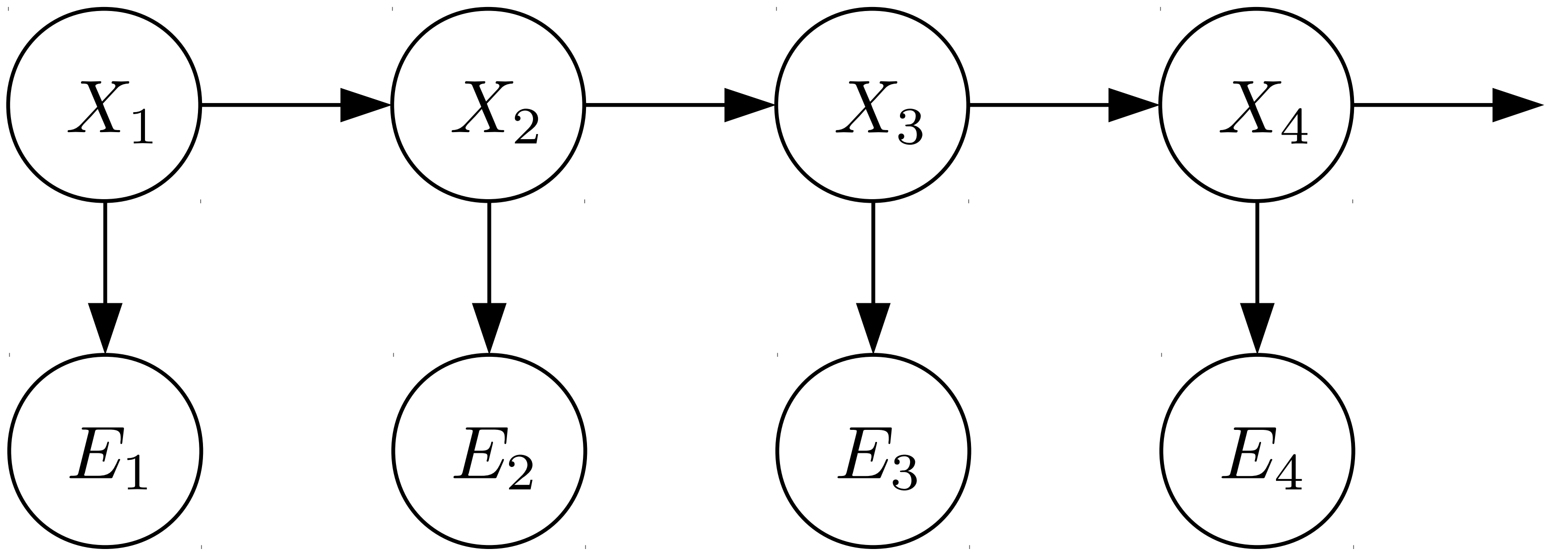
How calculate p that satisfies this eqn?

Hidden Markov Models (HMMs)

Hidden Markov Models:

– extension of the Markov model
– state is assumed to be “hidden”

Hidden Markov Models (HMMs)



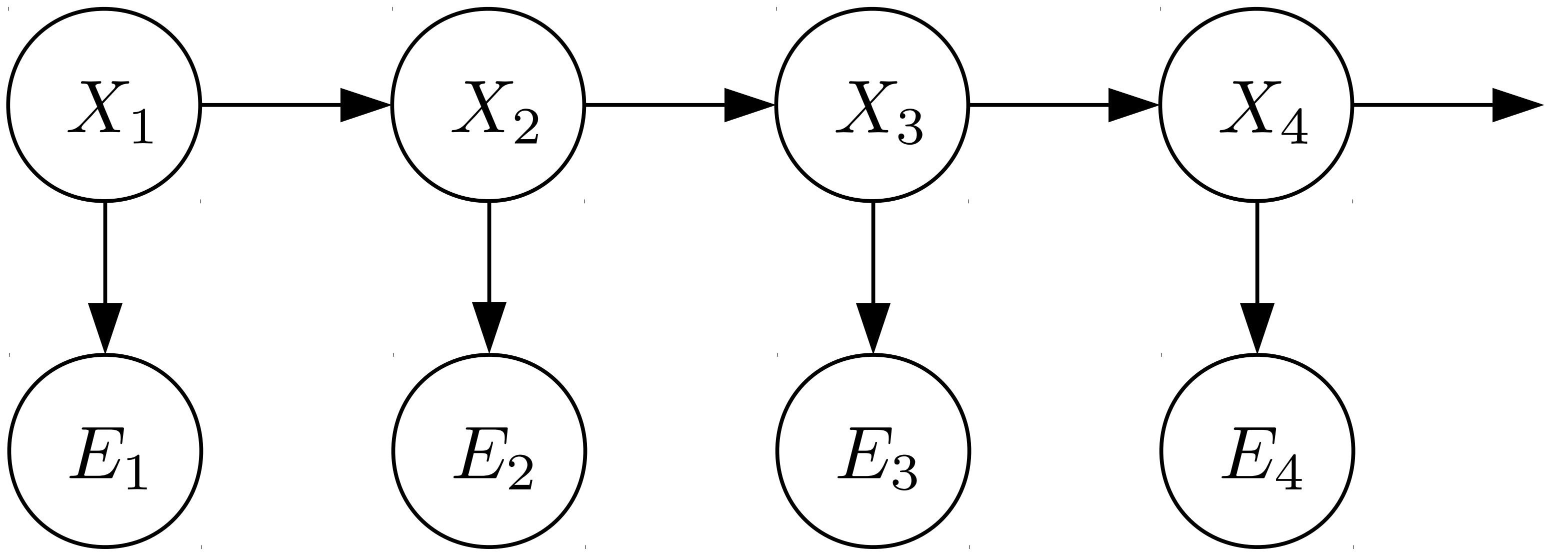
Called an “emission”

State, X_t , is assumed to be unobserved

However, you get to make one observation, E_t , on each timestep.

<u>Examples:</u>
– speech to text; tracking in computer vision' robot localization

Hidden Markov Models (HMMs)



Sensor Markov Assumption: the current observation depends only on current state:

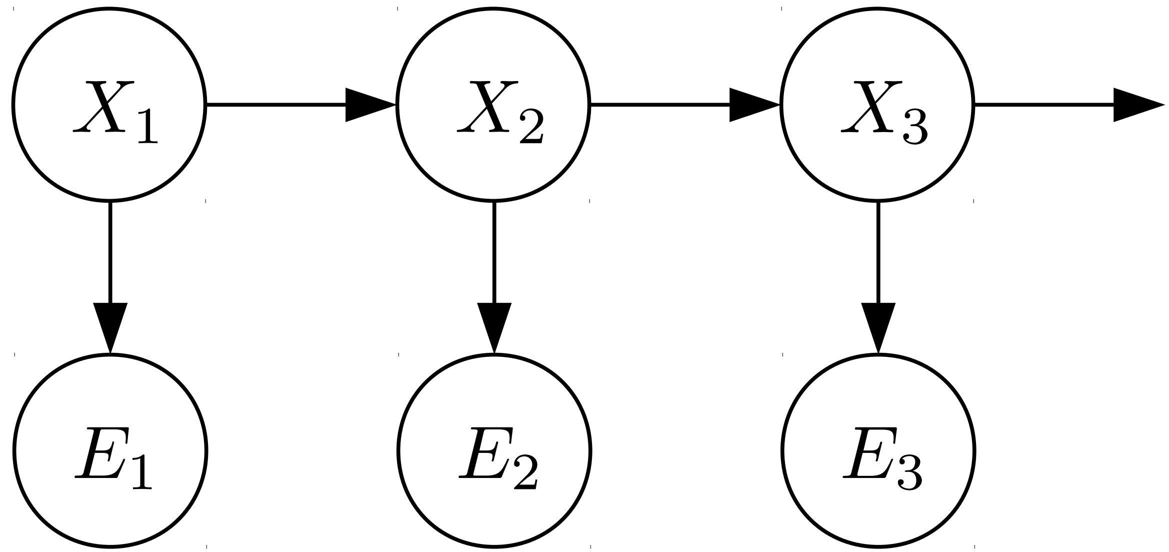
$$P(E_t|X_t, X_{t-1}, \ldots, X_1) = P(E_t|X_t)$$

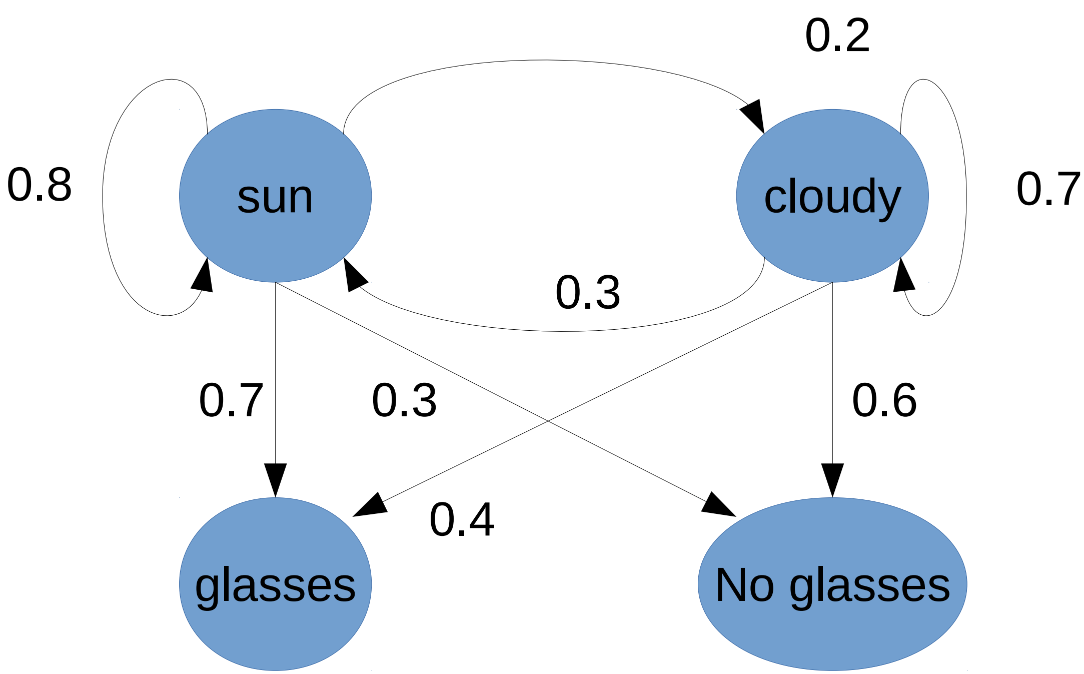
$$E_t \perp\!\!\!\perp X_{t-1}|X_t$$

HMM example

$X \in \{\text{sun}, \text{cloudy}\}$
(state is unobserved)

$E \in \{\text{glasses}, \text{no glasses}\}$
(only observations are observed)





You live underground...

Every day, you're boss comes in either wearing sunglasses or not

Can you infer whether it's sunny out based on whether you see the glasses over a sequence of days?

– e.g. what's the prob it's sunny out today if you've seen your boss wear glasses three days in a row?

HMM Filtering

Given a prior distribution, $P(X_1)$, and a series of observations, $E_1, \dots, E_T$, calculate the posterior distribution: $P(X_t | E_1, \dots, E_T)$

Two steps:

$$\boxed{P(X_t | E_{1:t})} \longrightarrow \boxed{P(X_{t+1} | E_{1:t})} \longrightarrow \boxed{P(X_{t+1} | E_{1:t+1})}$$

Process update   Observation update

The Kalman filter is perhaps the most famous instance of this idea

HMM Filtering

Given a prior distribution, $P(X_1)$, and a series of observations, $E_1, \ldots, E_T$, calculate the posterior distribution: $P(X_t | E_1, \ldots, E_T)$

Two steps:

$$\begin{array}{ccccc} B(X_t) & & B'(X_t) & & B(X_{t+1}) \\ \boxed{P(X_t | E_{1:t})} & \xrightarrow[\text{Process update}]{} & \boxed{P(X_{t+1} | E_{1:t})} & \xrightarrow[\text{Observation update}]{} & \boxed{P(X_{t+1} | E_{1:t+1})} \end{array}$$

HMM Filtering

Given a prior distribution, $P(X_1)$, and a series of observations, $E_1, \ldots, E_T$, calculate the posterior distribution: P

“Beliefs”

Two steps:

$B(X_t)$ $\quad$ $B'(X_t)$ $\quad$ $B(X_{t+1})$

$$P(X_t|E_{1:t}) \longrightarrow P(X_{t+1}|E_{1:t}) \longrightarrow P(X_{t+1}|E_{1:t+1})$$

Process update $\quad$ Observation update

Process update

$$\overset{B(X_t)}{\boxed{P(X_t|e_{1:t})}} \longrightarrow \overset{B'(X_t)}{\boxed{P(X_{t+1}|e_{1:t})}}$$

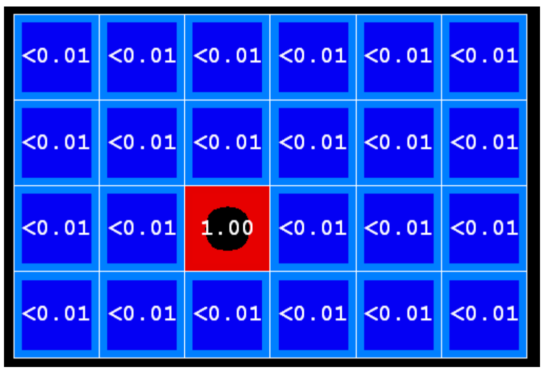
$$P(X_{t+1}|e_{1:t}) = \sum_{X_t} P(X_{t+1}|X_t, e_{1:t}) P(X_t|e_{1:t})$$

$$B'(X_{t+1}) = \sum_{X_t} P(X_{t+1}|X_t, e_{1:t}) B(X_t)$$

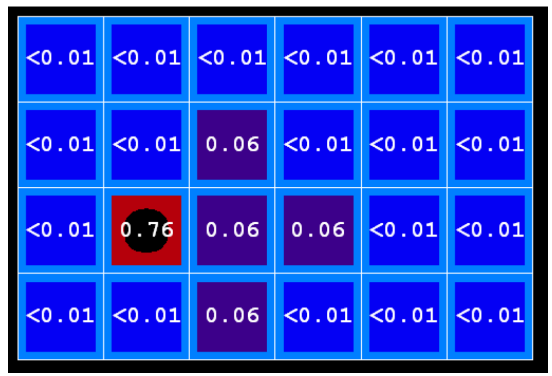
This is just forward simulation of the Markov Model

Process update: example

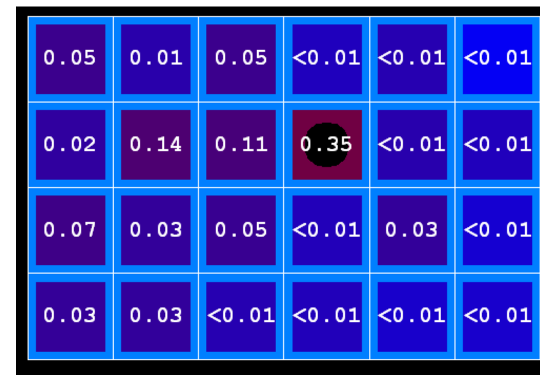
$$B'(X_{t+1}) = \sum_{X_t} P(X_{t+1}|X_t, e_{1:t})B(X_t)$$



T = 1



T = 2



T = 5

Completely certain about ghost position at T=1

A little less certain on the next time step...

By now, we've almost completely lost track of the ghost...

If we only do the process update, then we typically lose information over time – when might this not be true?

Observation update

$$B'(X_t) \qquad\qquad B(X_{t+1})$$

$$\boxed{P(X_{t+1}|e_{1:t})} \longrightarrow \boxed{P(X_{t+1}|e_{1:t+1})}$$

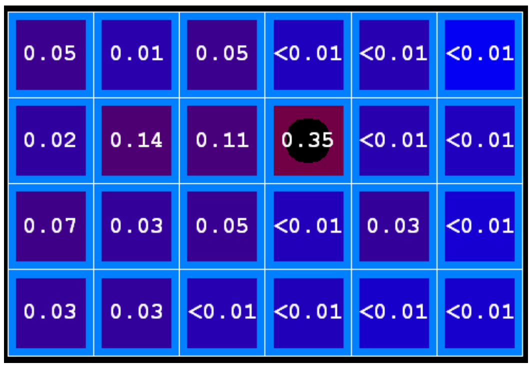
$$P(X_{t+1}|e_{1:t+1}) = \eta P(e_{t+1}|X_{t+1})P(X_{t+1}|e_{1:t})$$

$$B(X_{t+1}) = \eta P(e_{t+1}|X_{t+1})B'(X_{t+1})$$

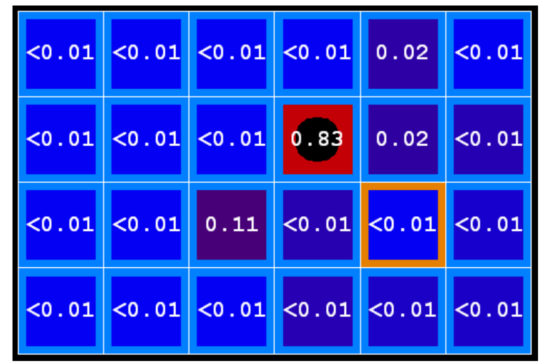
Where $\eta = \dfrac{1}{P(e_{t+1})}$ is a normalization factor

Observation update

$$B(X_{t+1}) = \eta P(e_{t+1}|X_{t+1})B'(X_{t+1})$$



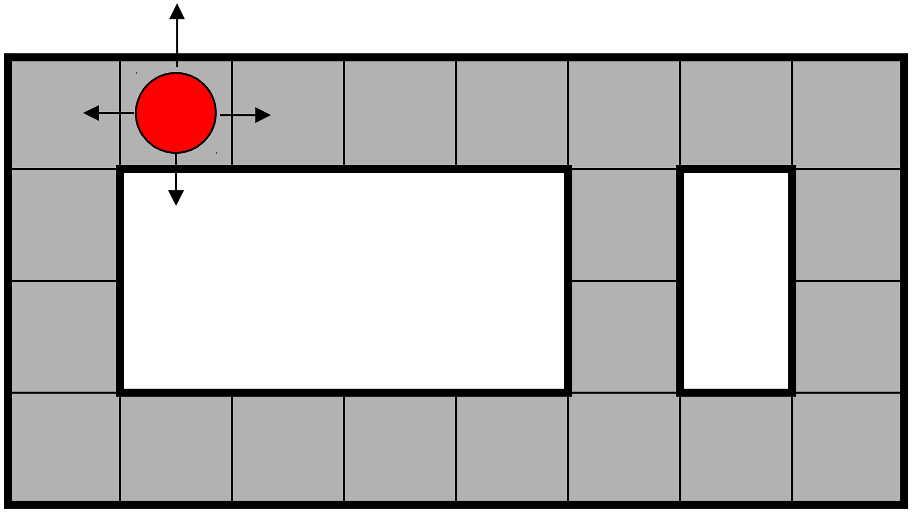
Before observation



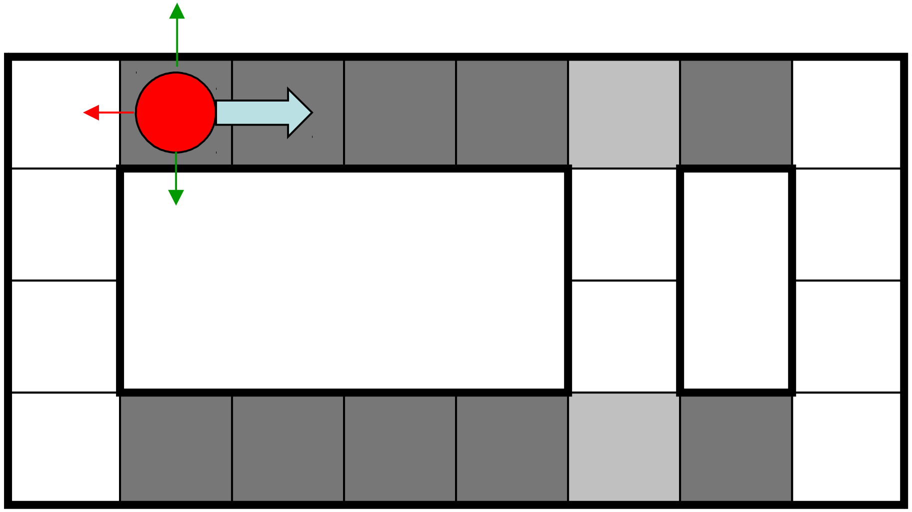
After observation

Observations enable the system to gain information
– a single observation may not determine system state exactly
– but, the more observations, the better

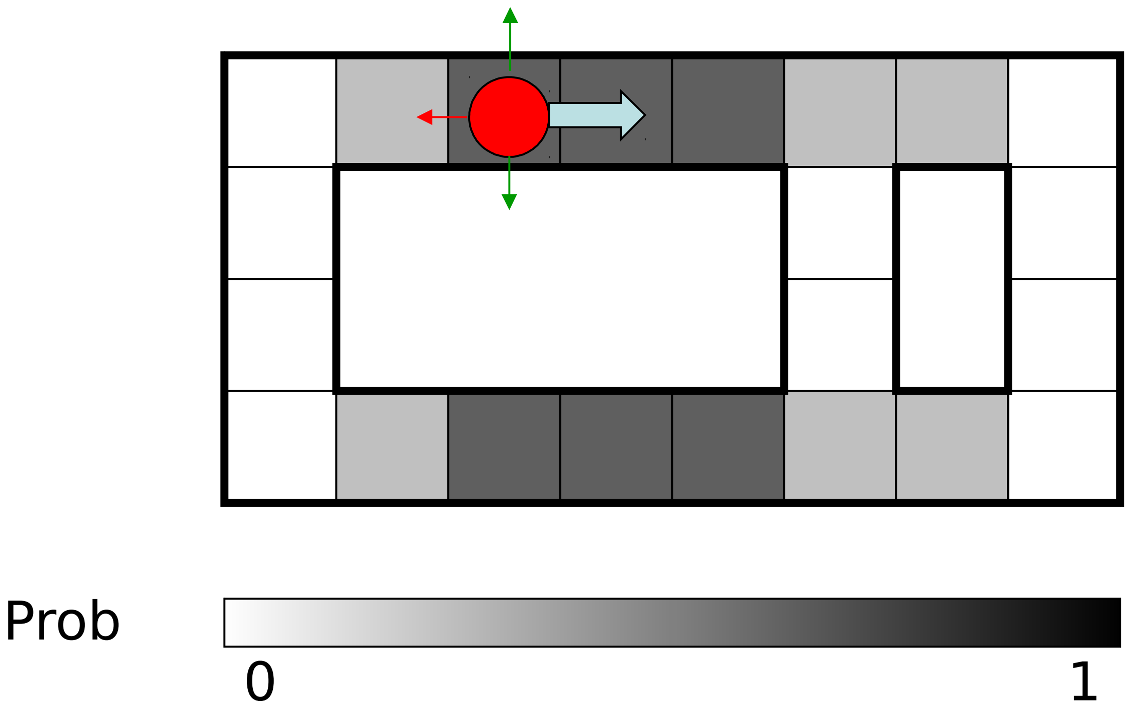
Robot localization example



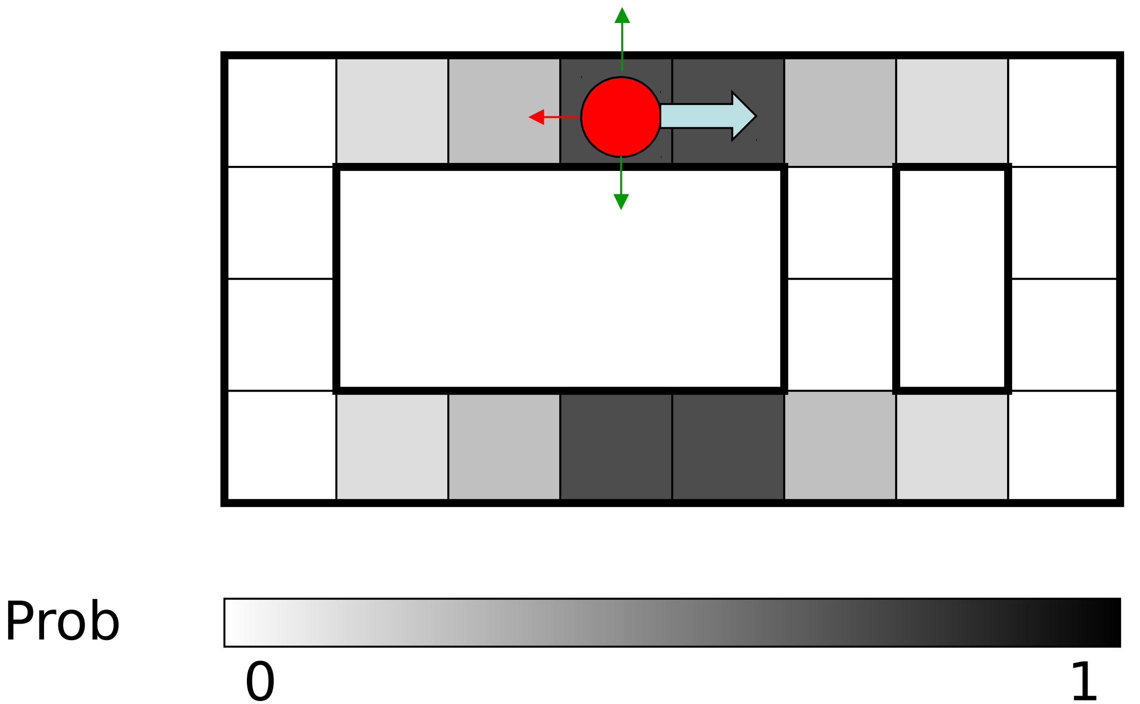
Robot localization example



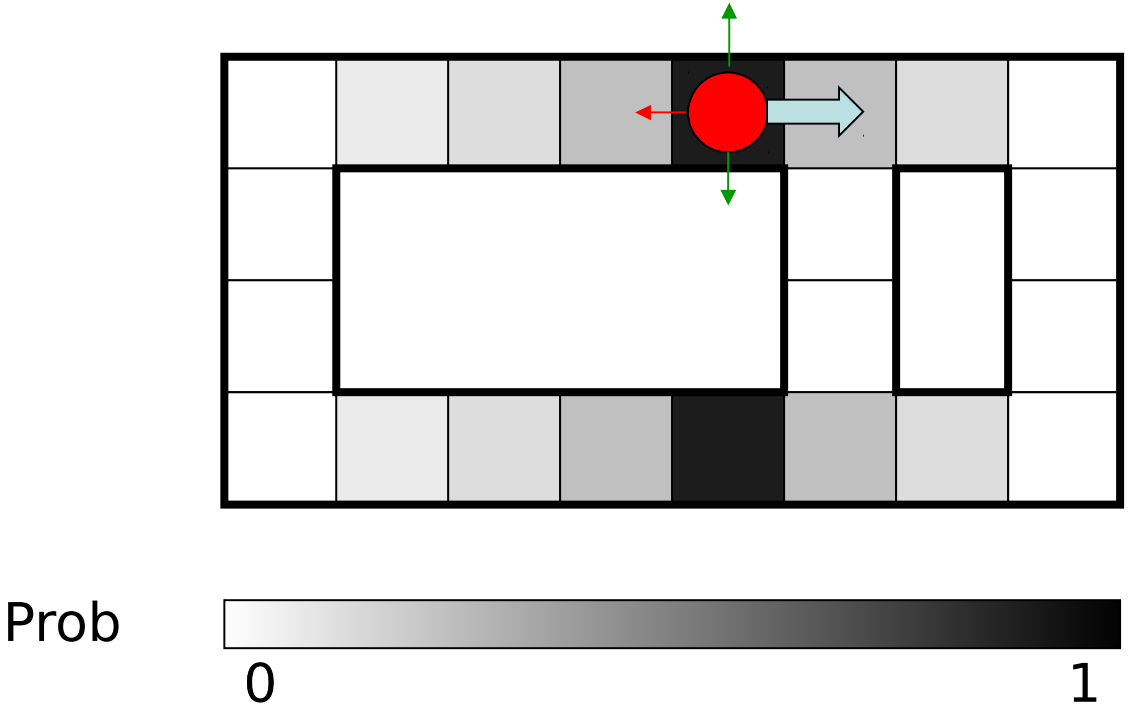
Robot localization example



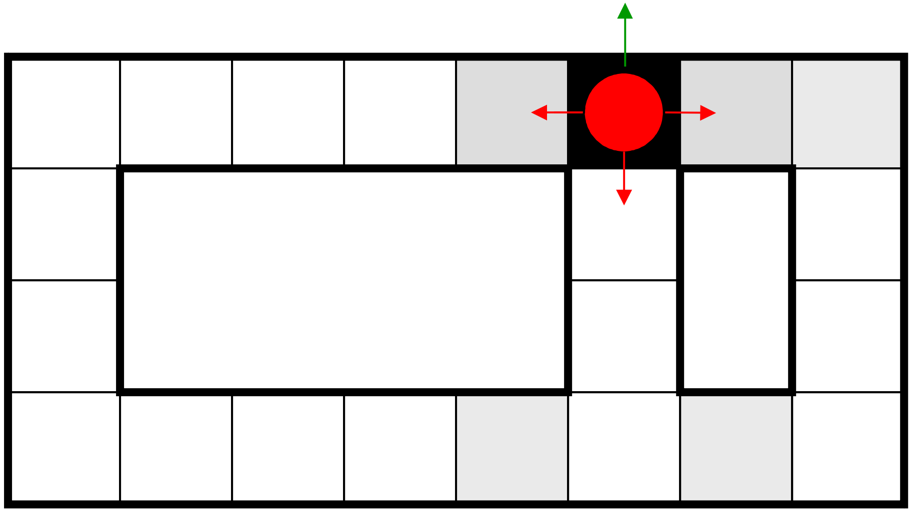
Robot localization example



Robot localization example



Robot localization example

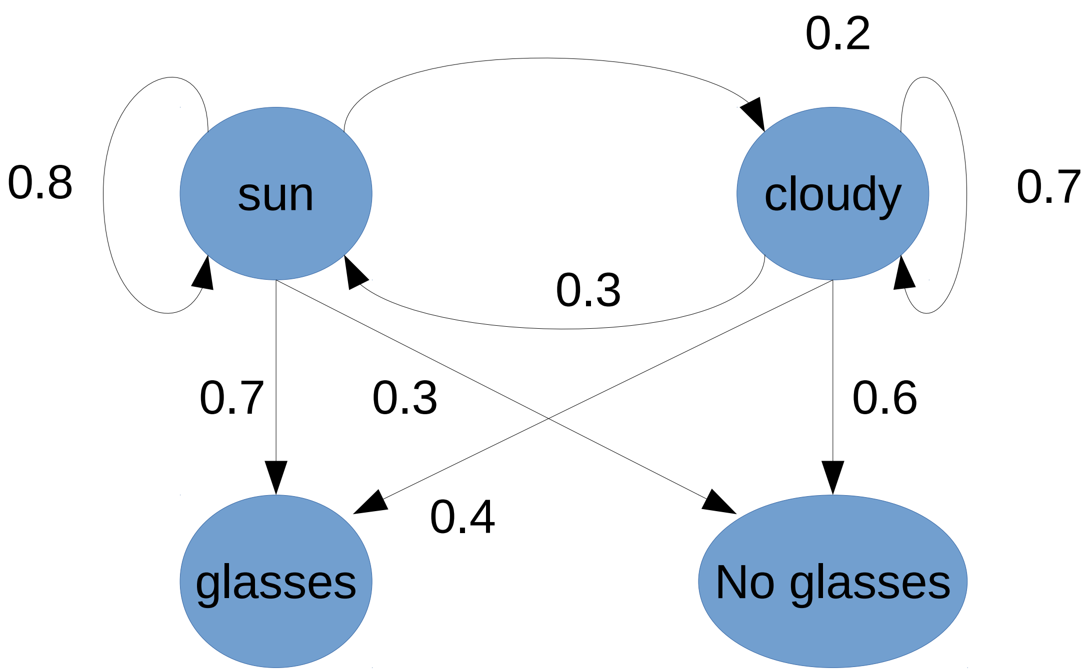


Prob

0

1

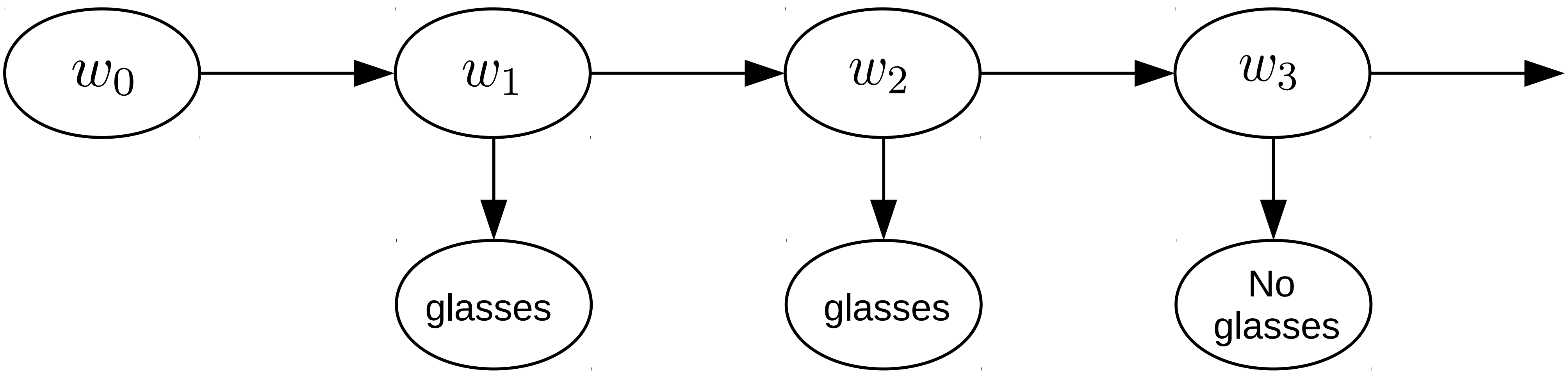
Weather HMM example



Weather HMM example

X_{t-1}	X_t	X_t
sun	sun	0.8
sun	cloudy	0.2
cloudy	sun	0.3
cloudy	cloudy	0.7

X_t	P(g_t|X_t)
sun	0.7
cloudy	0.4

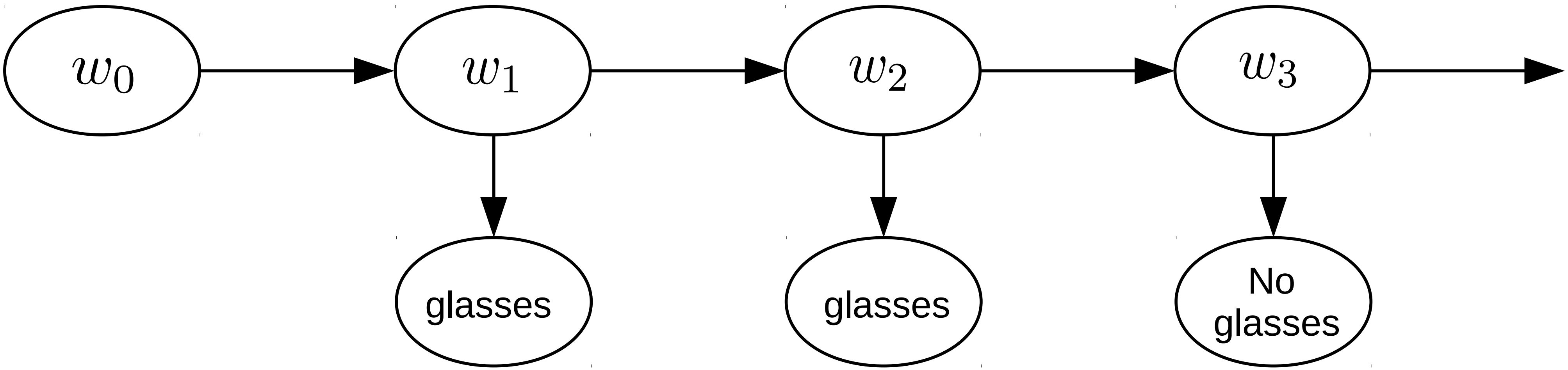


w_t	P(w_t)
sun	0.5
cloudy	0.5

Weather HMM example

X_{t-1}	X_t	X_t
sun	sun	0.8
sun	cloudy	0.2
cloudy	sun	0.3
cloudy	cloudy	0.7

X_t	P(g_t|X_t)
sun	0.7
cloudy	0.4



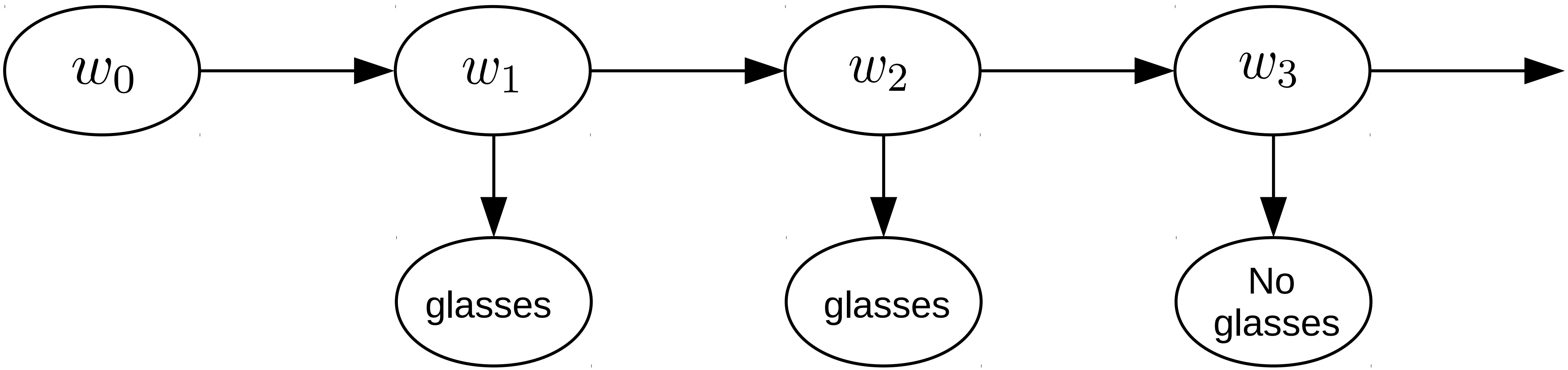
w_t	P(w_t)
sun	0.5
cloudy	0.5

w_t	P(w_t)
sun	?
cloudy	?

Weather HMM example

X_{t-1}	X_t	X_t
sun	sun	0.8
sun	cloudy	0.2
cloudy	sun	0.3
cloudy	cloudy	0.7

X_t	P(g_t|X_t)
sun	0.7
cloudy	0.4



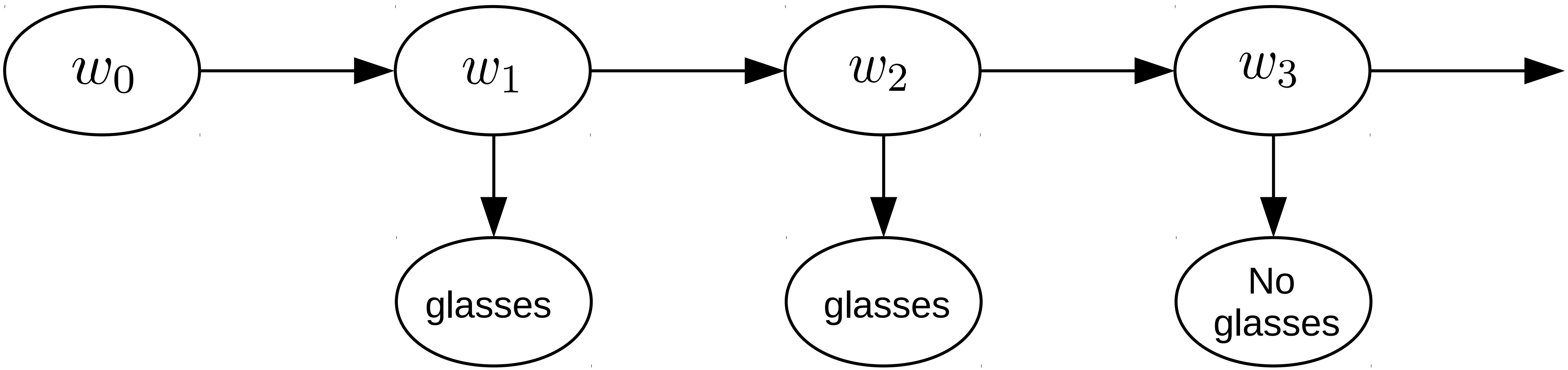
w_t	P(w_t)
sun	0.5
cloudy	0.5

w_t	P(w_t)
sun	0.55
cloudy	0.45

Weather HMM example

X_{t-1}	X_t	X_t
sun	sun	0.8
sun	cloudy	0.2
cloudy	sun	0.3
cloudy	cloudy	0.7

X_t	P(g_t|X_t)
sun	0.7
cloudy	0.4



w_t	P(w_t)
sun	0.5
cloudy	0.5

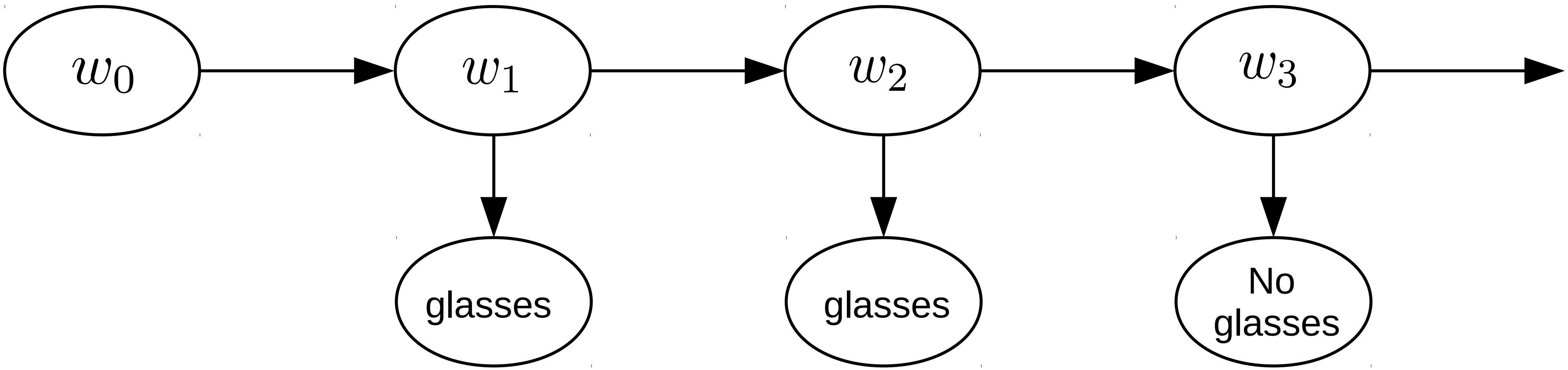
w_t	P(w_t)
sun	0.55
cloudy	0.45

w_t	P(w_t)
sun	?
cloudy	?

Weather HMM example

X_{t-1}	X_t	X_t
sun	sun	0.8
sun	cloudy	0.2
cloudy	sun	0.3
cloudy	cloudy	0.7

X_t	P(g_t|X_t)
sun	0.7
cloudy	0.4



w_t	P(w_t)
sun	0.5
cloudy	0.5

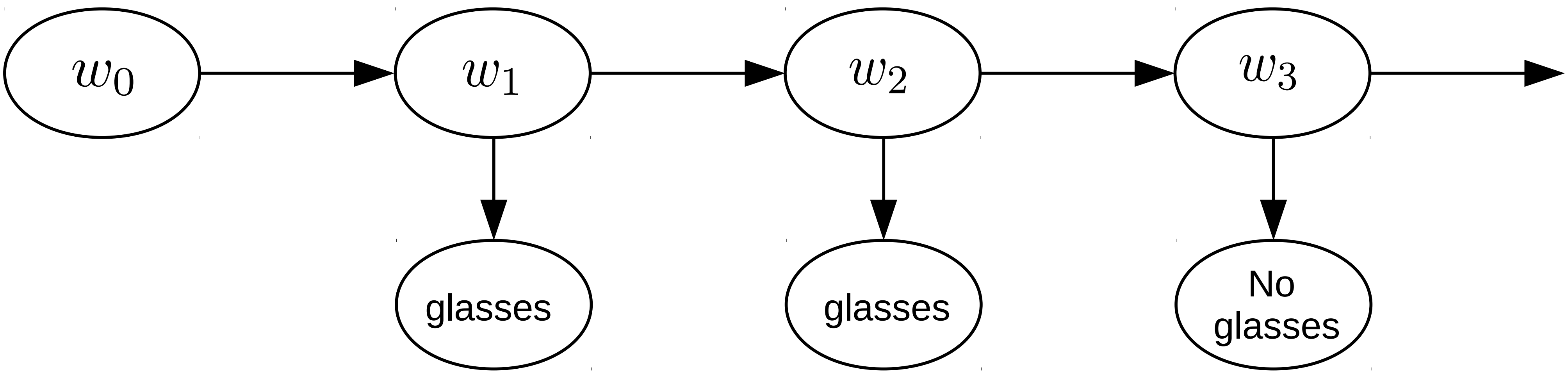
w_t	P(w_t)
sun	0.55
cloudy	0.45

w_t	P(w_t)
sun	0.68
cloudy	0.31

Weather HMM example

X_{t-1}	X_t	X_t
sun	sun	0.8
sun	cloudy	0.2
cloudy	sun	0.3
cloudy	cloudy	0.7

X_t	P(g_t|X_t)
sun	0.7
cloudy	0.4



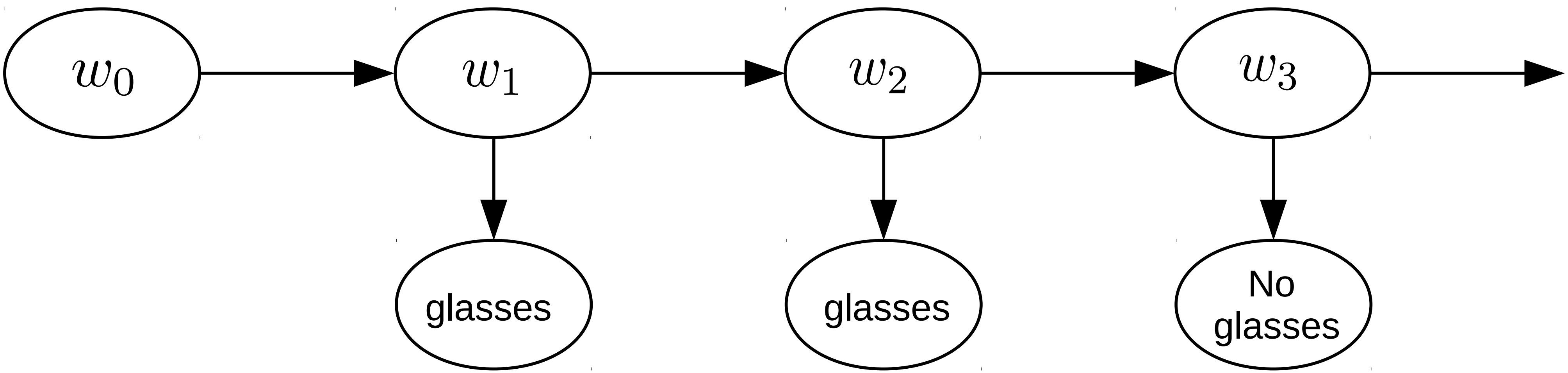
w_t	P(w_t)
sun	?
cloudy	?

w_t	P(w_t)
sun	0.68
cloudy	0.31

Weather HMM example

X_{t-1}	X_t	X_t
sun	sun	0.8
sun	cloudy	0.2
cloudy	sun	0.3
cloudy	cloudy	0.7

X_t	P(g_t|X_t)
sun	0.7
cloudy	0.4



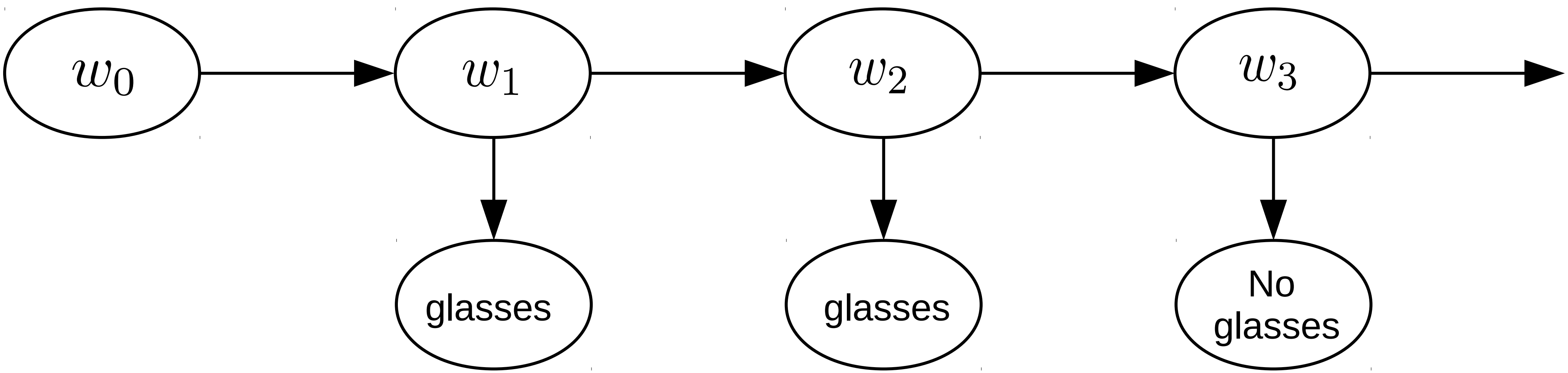
w_t	P(w_t)
sun	0.64
cloudy	0.36

w_t	P(w_t)
sun	0.68
cloudy	0.31

Weather HMM example

X_{t-1}	X_t	X_t
sun	sun	0.8
sun	cloudy	0.2
cloudy	sun	0.3
cloudy	cloudy	0.7

X_t	P(g_t|X_t)
sun	0.7
cloudy	0.4



w_t	P(w_t)
sun	0.64
cloudy	0.36

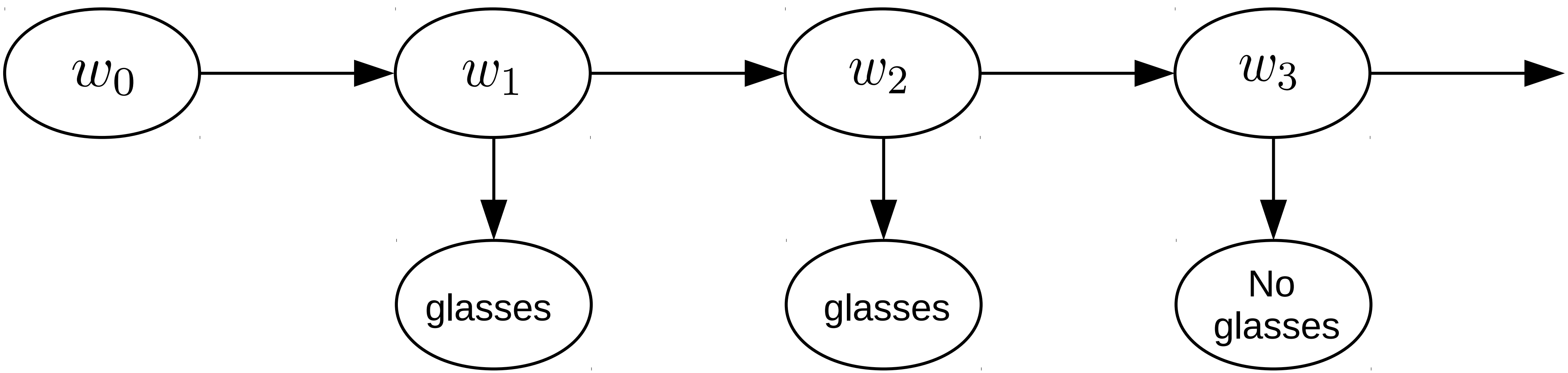
w_t	P(w_t)
sun	0.68
cloudy	0.31

w_t	P(w_t)
sun	?
cloudy	?

Weather HMM example

X_{t-1}	X_t	X_t
sun	sun	0.8
sun	cloudy	0.2
cloudy	sun	0.3
cloudy	cloudy	0.7

X_t	P(g_t|X_t)
sun	0.7
cloudy	0.4

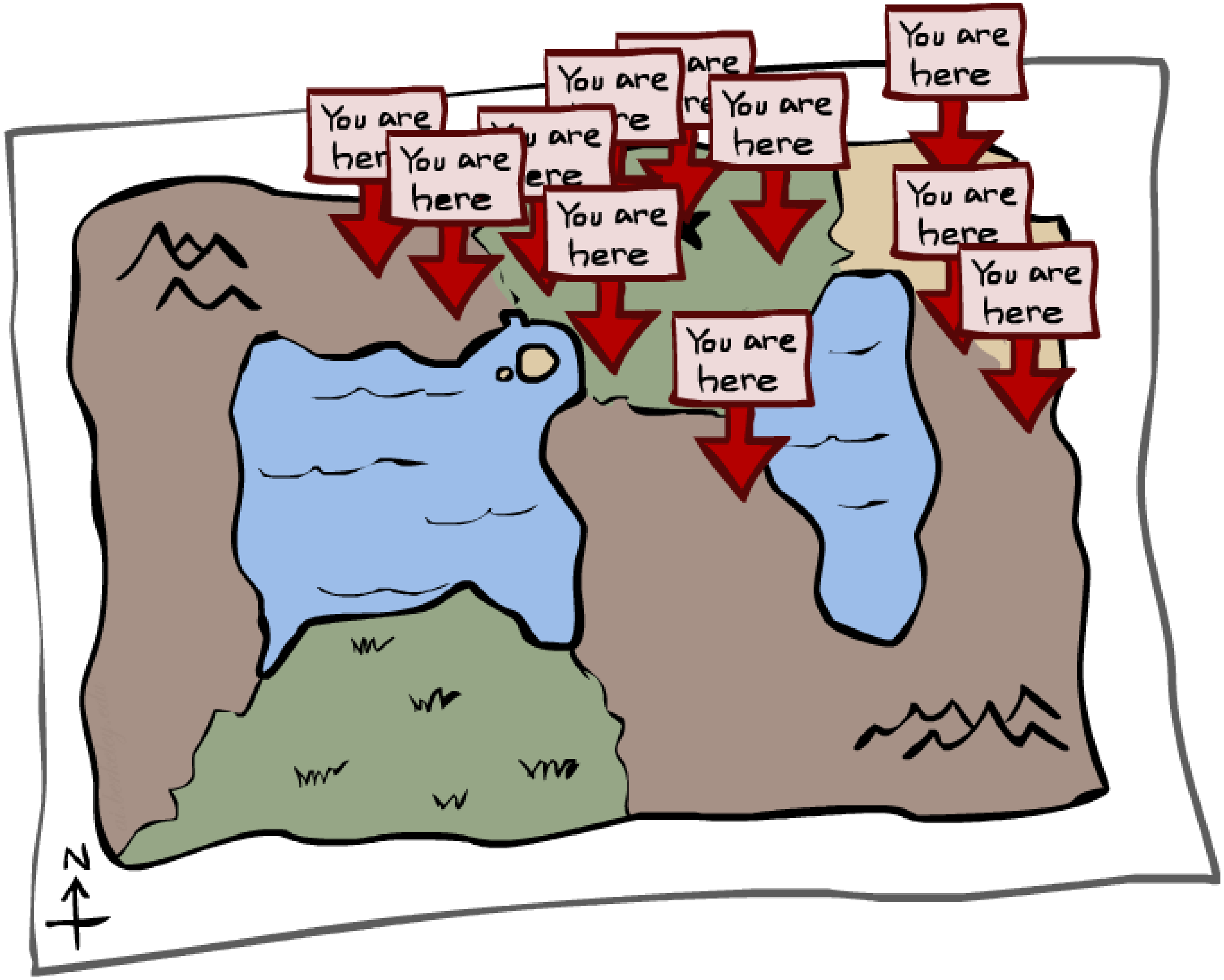


w_t	P(w_t)
sun	0.64
cloudy	0.36

w_t	P(w_t)
sun	0.68
cloudy	0.31

w_t	P(w_t)
sun	0.76
cloudy	0.24

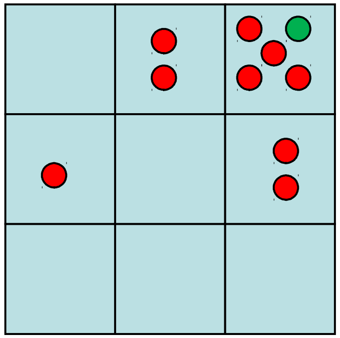
Particle Filtering



Slide: Berkeley CS188 course notes (downloaded Summer 2015)

Representation: Particles

- Our representation of P(X) is now a list of N particles (samples)
  - Generally, N << |X|
  - Storing map from X to counts would defeat the point



- P(x) approximated by number of particles with value x
  - So, many x may have P(x) = 0!
  - More particles, more accuracy

Particles:
(3,3)
(2,3)
(3,3)
(3,2)
(3,3)
(3,2)
(1,2)
(3,3)
(3,3)
(2,3)

- For now, all particles have a weight of 1

Particle Filtering: Elapse Time

- Each particle is moved by sampling its next position from the transition model

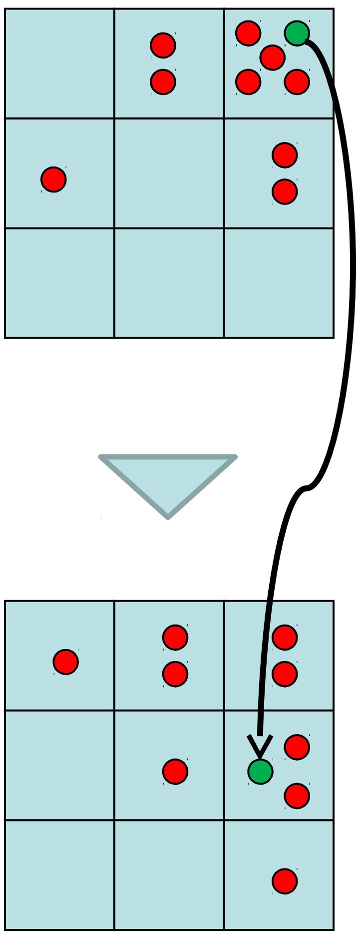
$$x' = \text{sample}(P(X'|x))$$

  - This is like prior sampling – samples' frequencies reflect the transition probabilities
  - Here, most samples move clockwise, but some move in another direction or stay in place

- This captures the passage of time
  - If enough samples, close to exact values before and after (consistent)

Particles:
(3,3)
(2,3)
(3,3)
(3,2)
(3,3)
(3,2)
(1,2)
(3,3)
(3,3)
(2,3)

Particles:
(3,2)
(2,3)
(3,2)
(3,1)
(3,3)
(3,2)
(1,3)
(2,3)
(3,2)
(2,2)



Slide: Berkeley CS188 course notes (downloaded Summer 2015)

Particle Filtering: Observe

- Slightly trickier:
  - Don’t sample observation, fix it
  - Similar to likelihood weighting, downweight samples based on the evidence

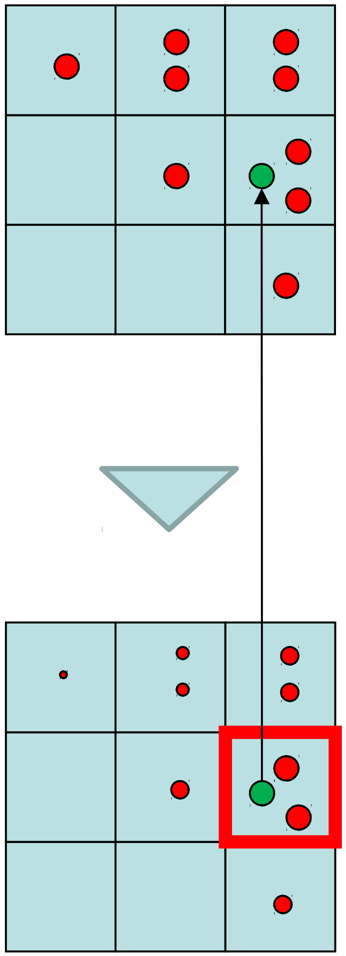
$$w(x) = P(e|x)$$

$$B(X) \propto P(e|X)B'(X)$$

  - As before, the probabilities don’t sum to one, since all have been downweighted (in fact they now sum to (N times) an approximation of P(e))

Particles:
(3,2)
(2,3)
(3,2)
(3,1)
(3,3)
(3,2)
(1,3)
(2,3)
(3,2)
(2,2)

Particles:
(3,2) w=.9
(2,3) w=.2
(3,2) w=.9
(3,1) w=.4
(3,3) w=.4
(3,2) w=.9
(1,3) w=.1
(2,3) w=.2
(3,2) w=.9
(2,2) w=.4



Slide: Berkeley CS188 course notes (downloaded Summer 2015)

Particle Filtering: Resample

- Rather than tracking weighted samples, we resample

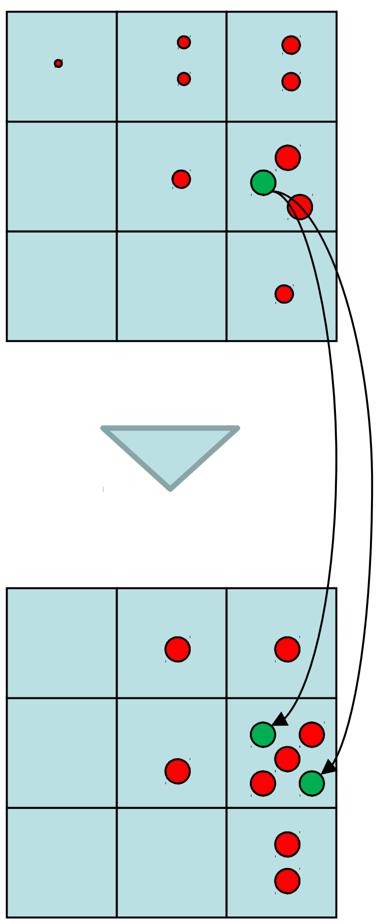
- N times, we choose from our weighted sample distribution (i.e. draw with replacement)

- This is equivalent to renormalizing the distribution

- Now the update is complete for this time step, continue with the next one

Particles:
(3,2) w=.9
(2,3) w=.2
(3,2) w=.9
(3,1) w=.4
(3,3) w=.4
(3,2) w=.9
(1,3) w=.1
(2,3) w=.2
(3,2) w=.9
(2,2) w=.4

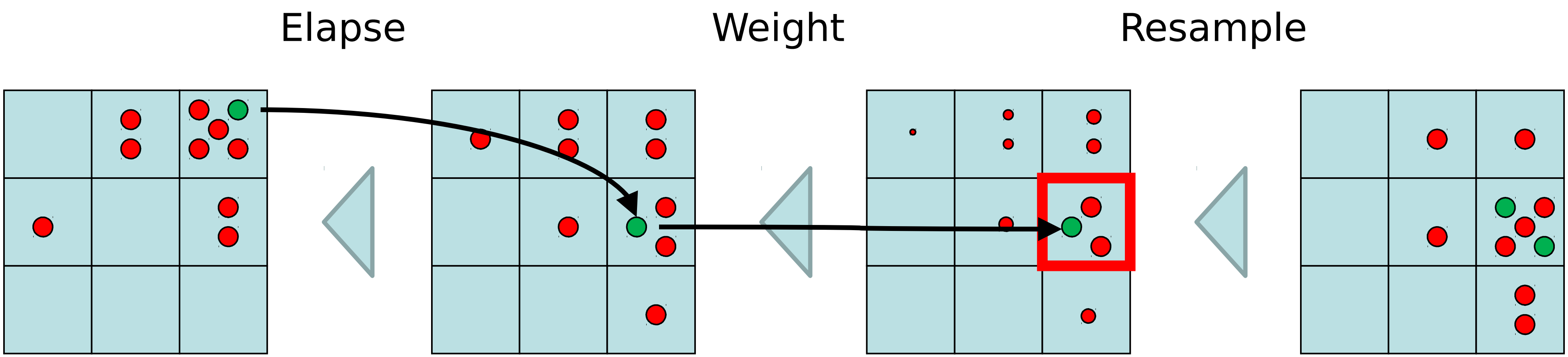
(New) Particles:
(3,2)
(2,2)
(3,2)
(2,3)
(3,3)
(3,2)
(1,3)
(2,3)
(3,2)
(3,2)



Slide: Berkeley CS188 course notes (downloaded Summer 2015)

Recap: Particle Filtering

- Particles: track samples of states rather than an explicit distribution



Particles:
(3,3)
(2,3)
(3,3)
(3,2)
(3,3)
(3,2)
(1,2)
(3,3)
(3,3)
(2,3)

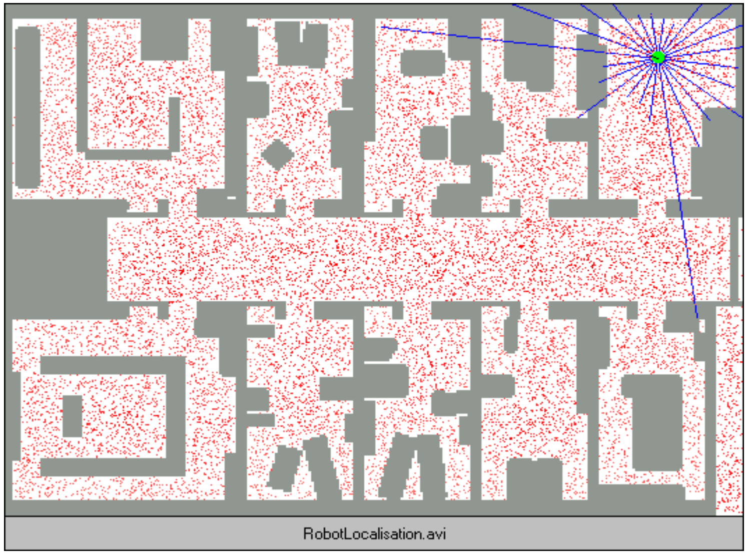
Particles:
(3,2)
(2,3)
(3,2)
(3,1)
(3,3)
(3,2)
(1,3)
(2,3)
(3,2)
(2,2)

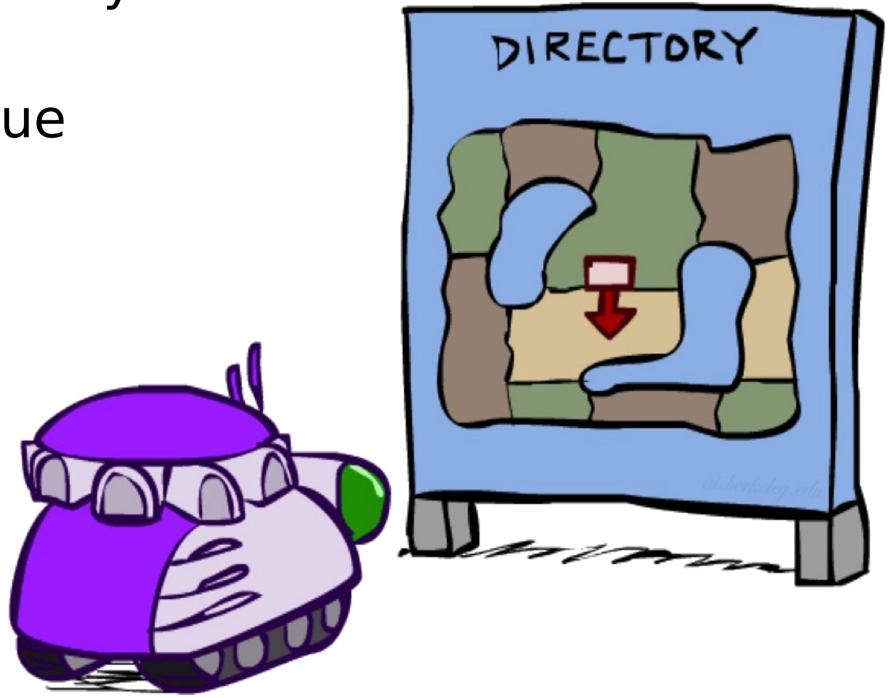
Particles:
(3,2) w=.9
(2,3) w=.2
(3,2) w=.9
(3,1) w=.4
(3,3) w=.4
(3,2) w=.9
(1,3) w=.1
(2,3) w=.2
(3,2) w=.9
(2,2) w=.4

(New) Particles:
(3,2)
(2,2)
(3,2)
(2,3)
(3,3)
(3,2)
(1,3)
(2,3)
(3,2)
(3,2)

Robot Localization

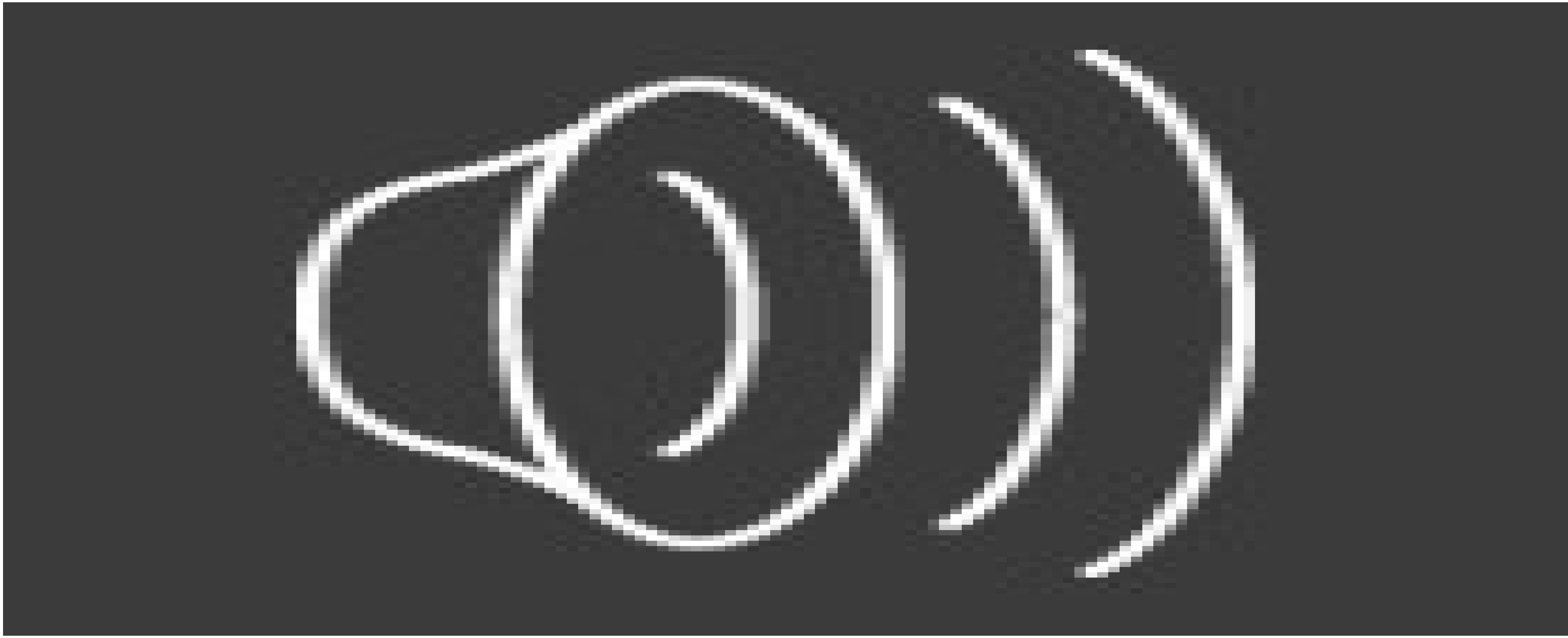
- In robot localization:
  - We know the map, but not the robot’s position
  - Observations may be vectors of range finder readings
  - State space and readings are typically continuous (works basically like a very fine grid) and so we cannot store B(X)
  - Particle filtering is a main technique



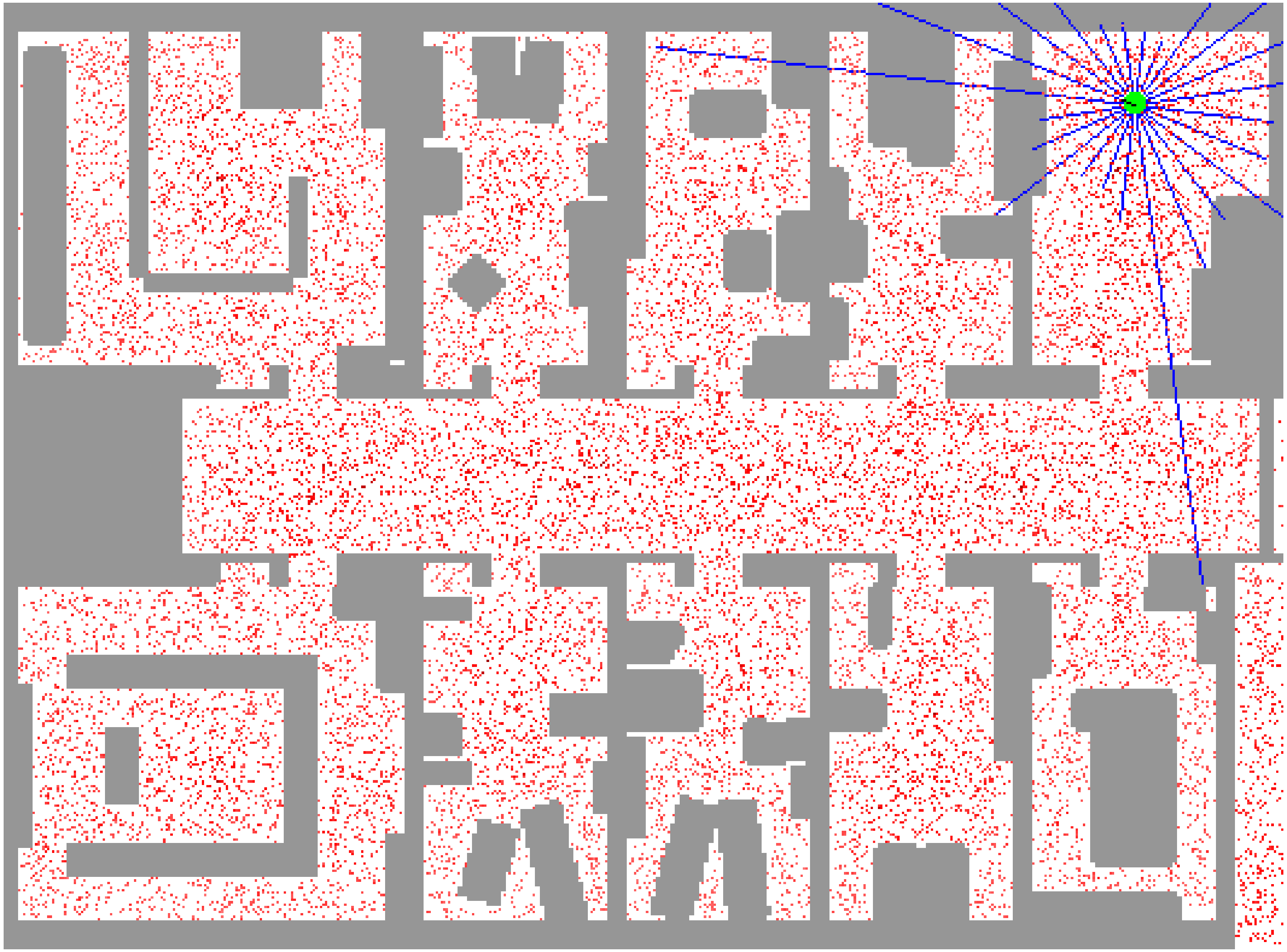


Slide: Berkeley CS188 course notes (downloaded Summer 2015)

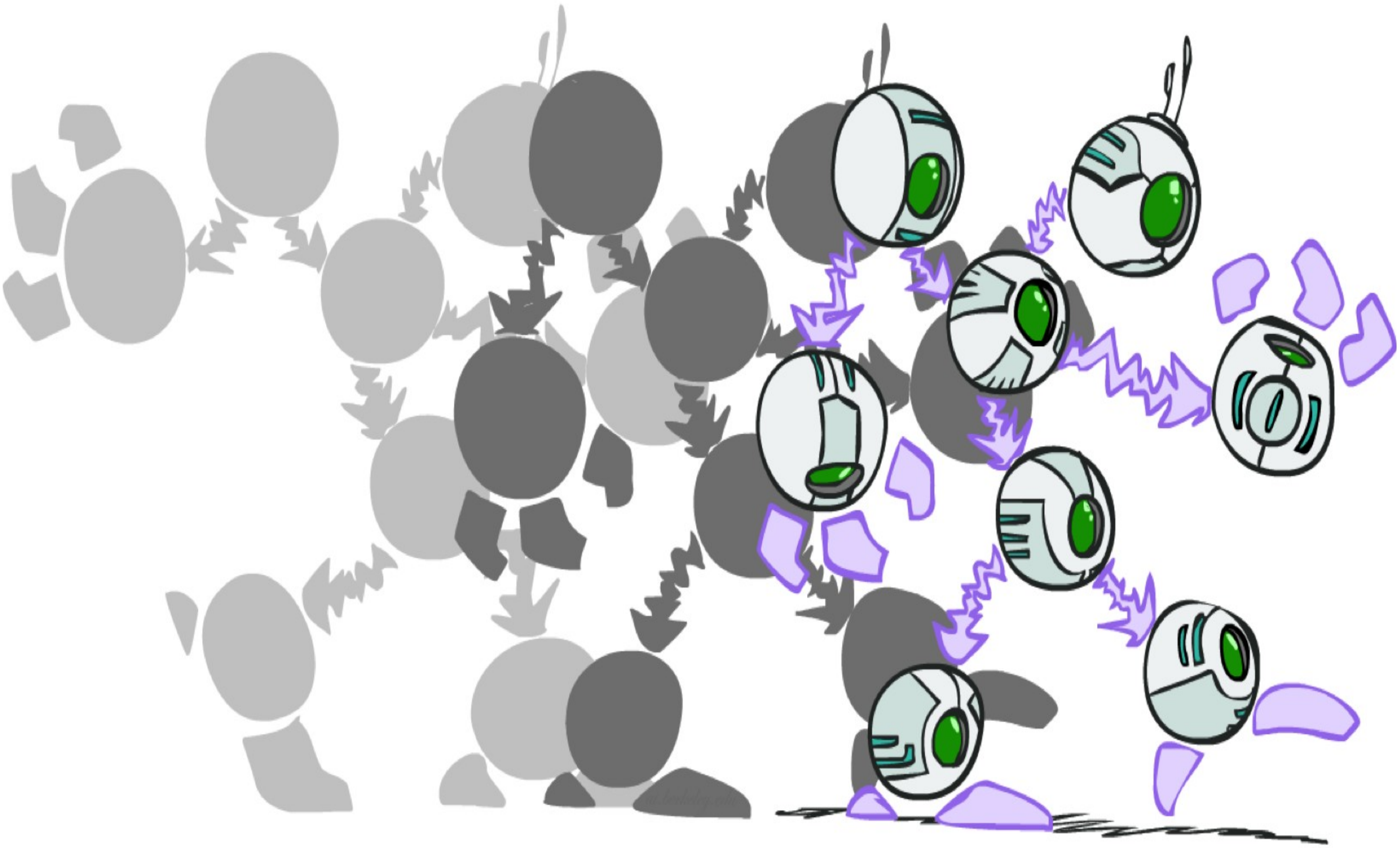
Particle Filter Localization (Sonar)



Particle Filter Localization (Laser)

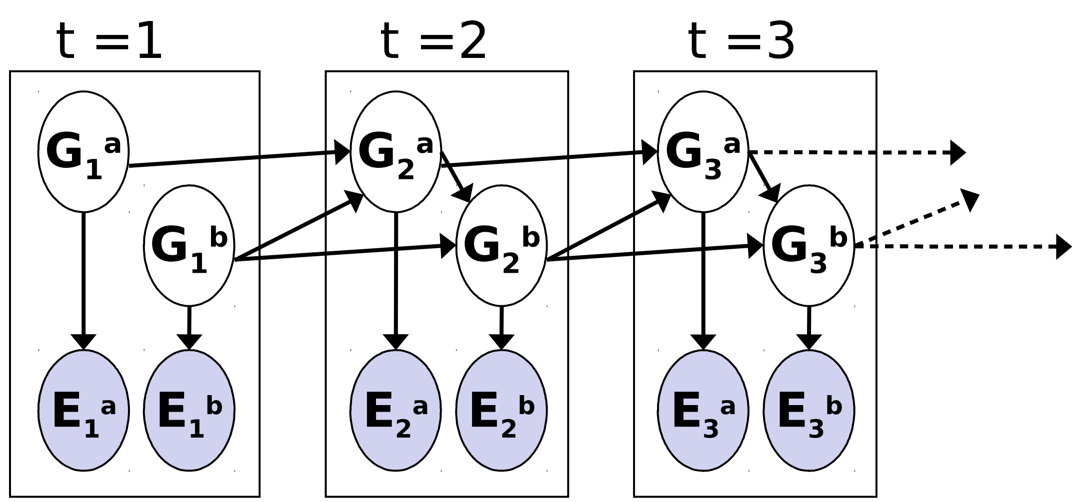


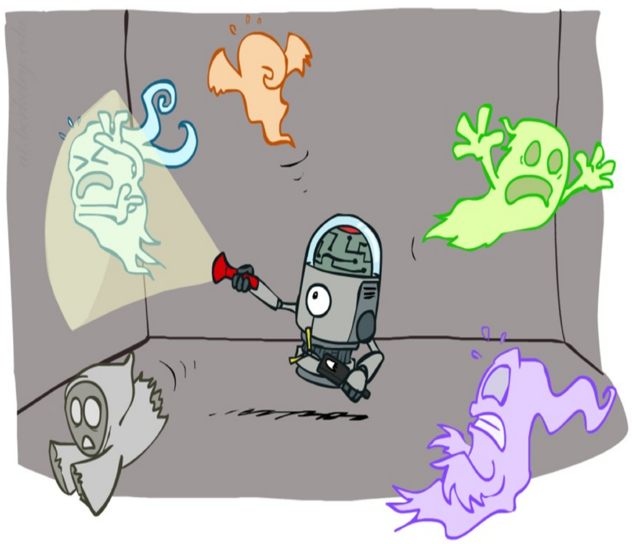
Dynamic Bayes Nets



Dynamic Bayes Nets (DBNs)

- We want to track multiple variables over time, using multiple sources of evidence
- Idea: Repeat a fixed Bayes net structure at each time
- Variables from time *t* can condition on those from *t-1*





- Dynamic Bayes nets are a generalization of HMMs

DBN Particle Filters

- A particle is a complete sample for a time step
- **Initialize**: Generate prior samples for the t=1 Bayes net
  - Example particle: $\mathbf{G_1^a} = (3,3)$ $\mathbf{G_1^b} = (5,3)$
- **Elapse time**: Sample a successor for each particle
  - Example successor: $\mathbf{G_2^a} = (2,3)$ $\mathbf{G_2^b} = (6,3)$
- **Observe**: Weight each *entire* sample by the likelihood of the evidence conditioned on the sample
  - Likelihood: $P(\mathbf{E_1^a} \mid \mathbf{G_1^a}) * P(\mathbf{E_1^b} \mid \mathbf{G_1^b})$
- **Resample:** Select prior samples (tuples of values) in proportion to their likelihood