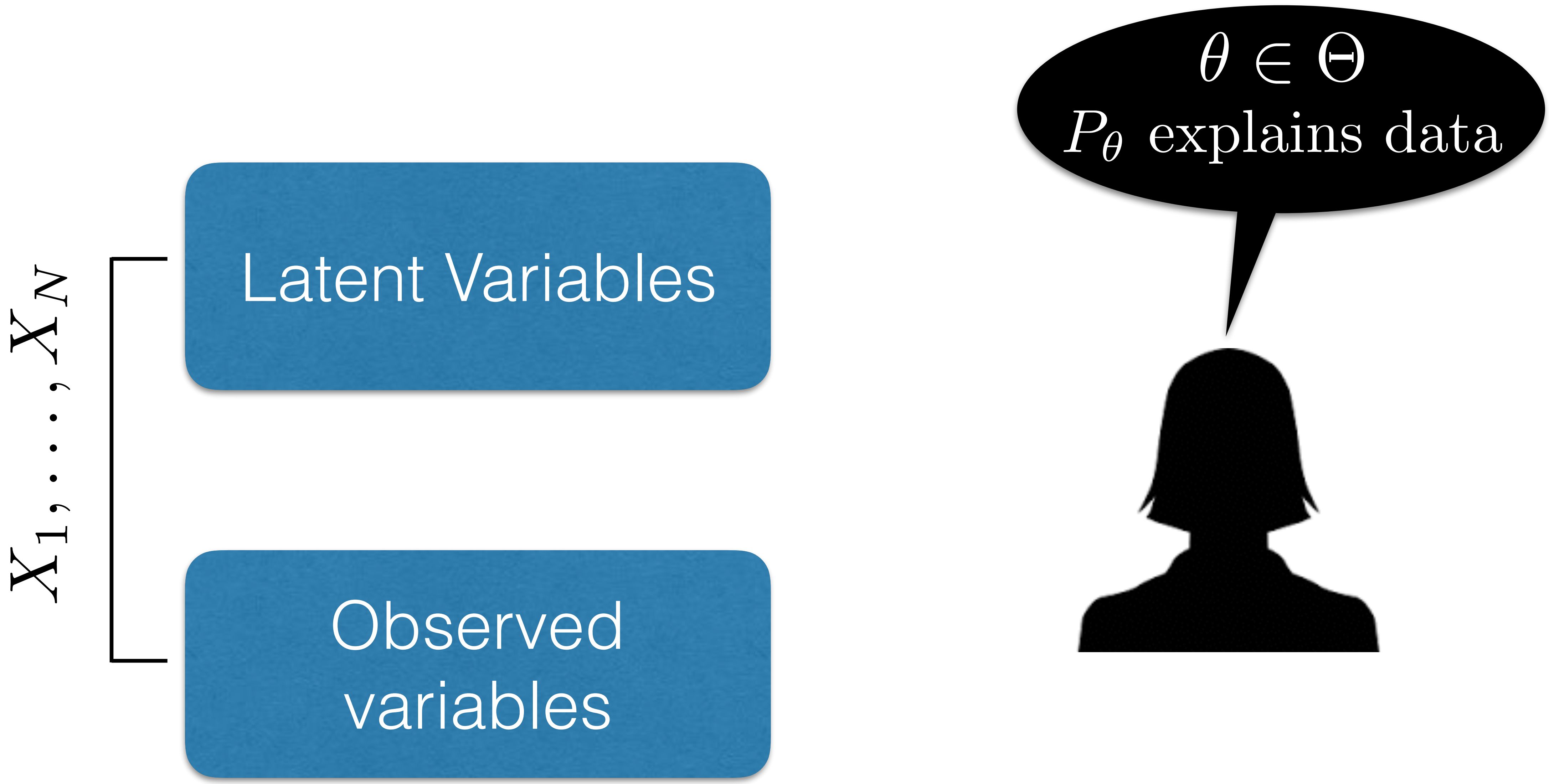
Machine Learning for Data Science (CS4786)
Lecture 19

Hidden Markov Models

Probabilistic Model



Graphical Models

- Abstract away the parameterization specifics
- Focus on relationship between random variables

Relationship Between Variables

Let $X = (X_1, \dots, X_N)$ be the random variables of our model (both latent and observed)

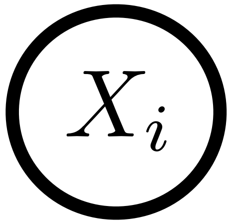
- Joint probability distribution over variable can be complex esp. if we have many complexly related variables
- Can we represent relation between variables in conceptually simpler fashion?
- We often have prior knowledge about the dependencies (or conditional (in)dependencies) between variables

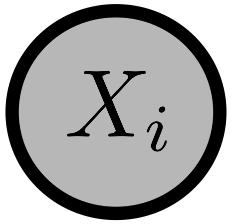
Graphical Models

- A graph whose nodes are variables $X_1, \ldots, X_N$
- Graphs are an intuitive way of representing relationships between large number of variables
- Allows us to abstract out the parametric form that depends on θ and the basic relationship between the random variables.

Draw a picture for the generative story
that explains what generates what.

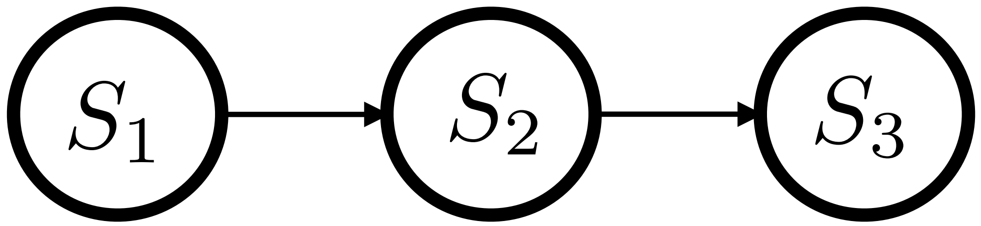
Graphical Models

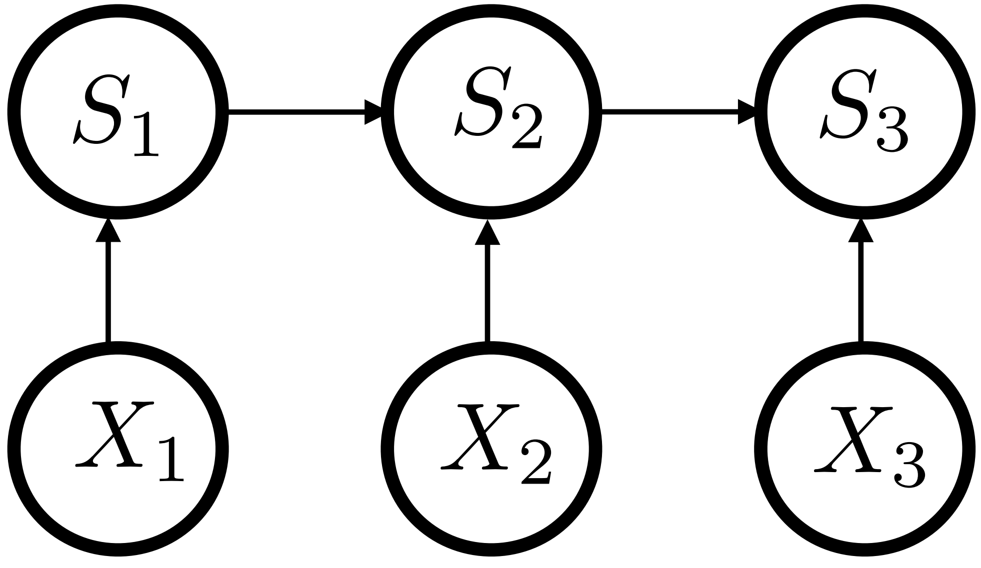
- Variables X_i is written as 

 if X_i is observed
- Variables X_i is written as 

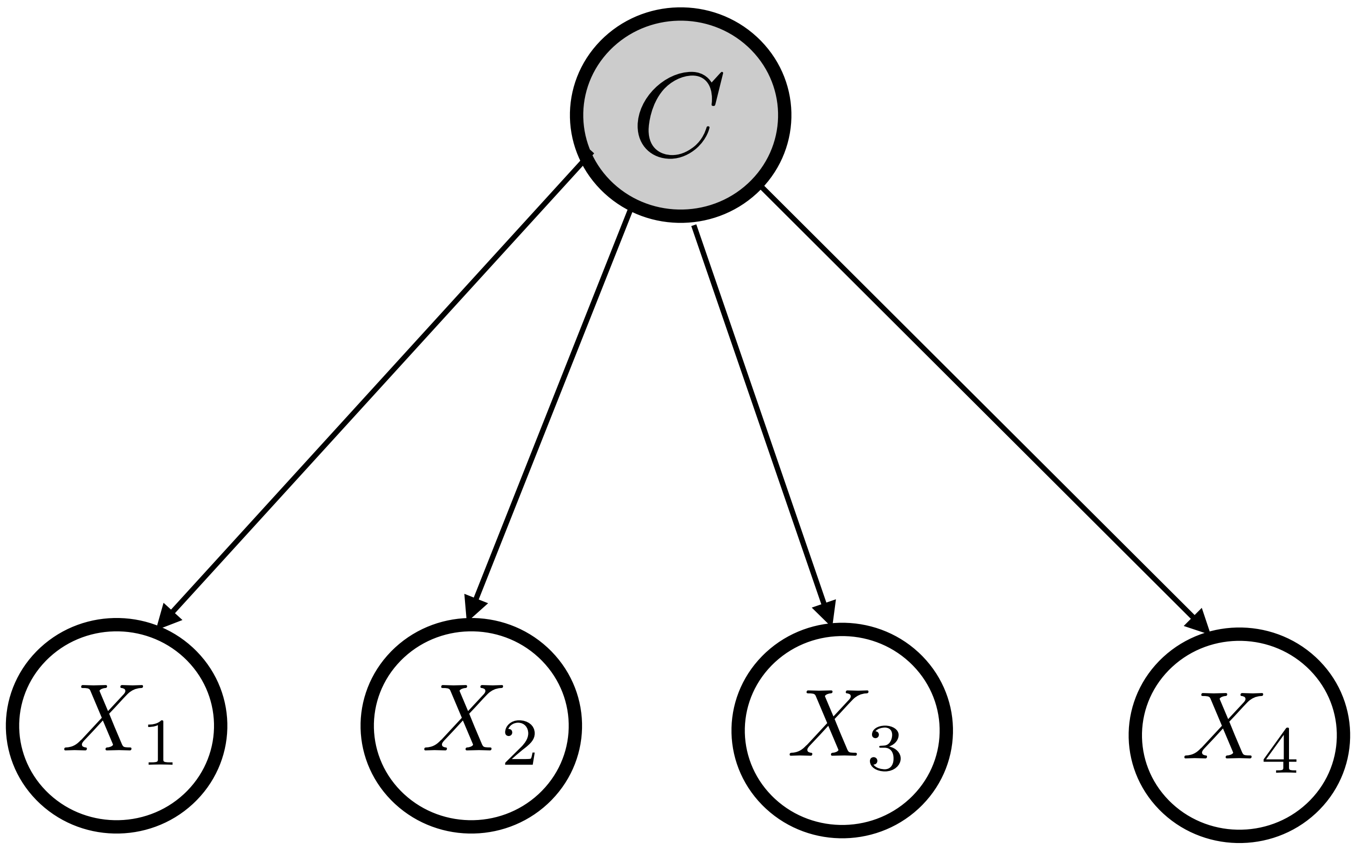
 if X_i is latent
- Parameters are often left out (its understood and not explicitly written out). If present they don't have bounding objects
- An directed edge  is drawn connecting every parent to its child (from parent to child)

Example: Sum of Coin Flips



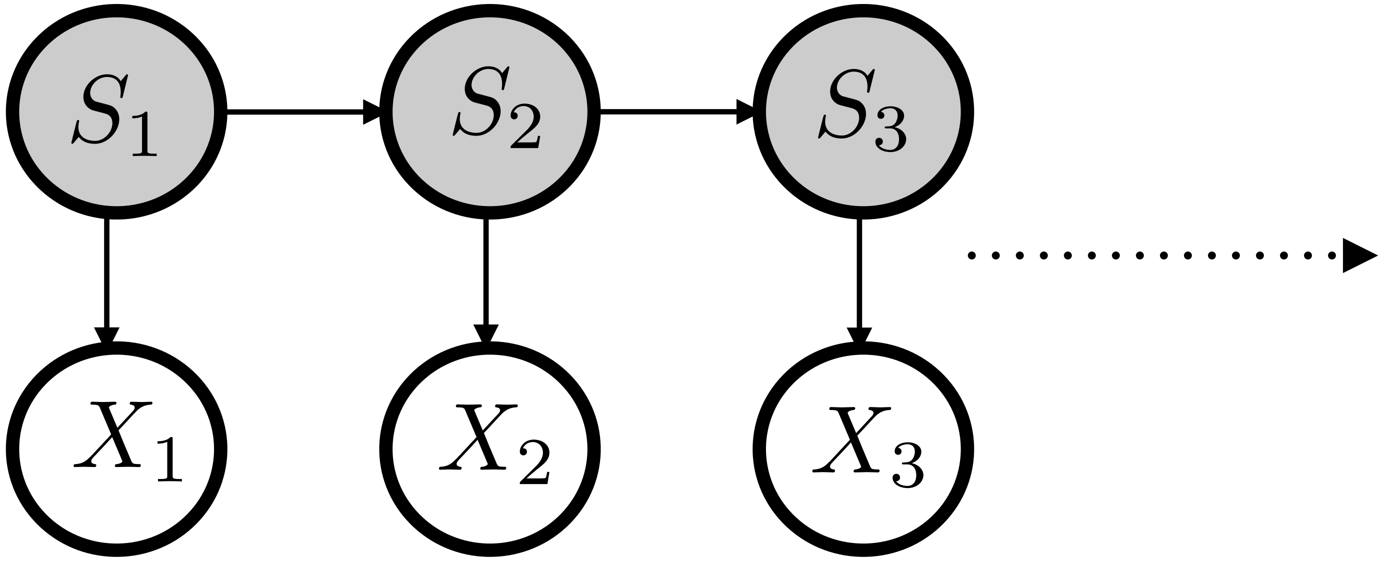


Example: Naive Bayes Classifier



Eg. Spam classification

Example: Hidden Markov Model

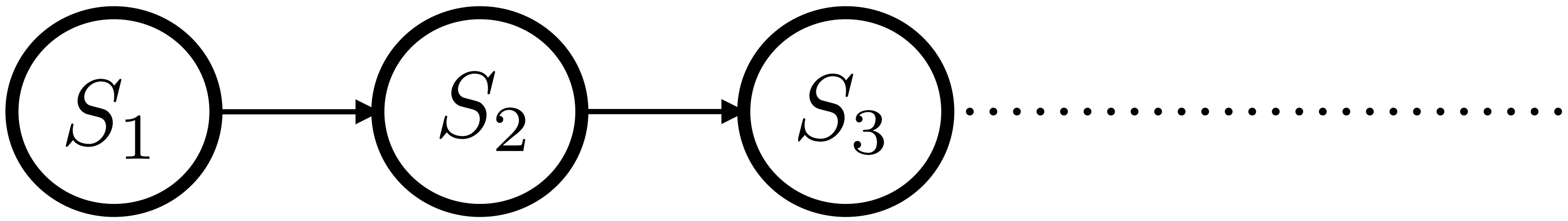


Hidden Markov Model (HMM)

- Speech recognition
- Natural language processing models
- Robot localization
- User attention modeling
- Medical monitoring

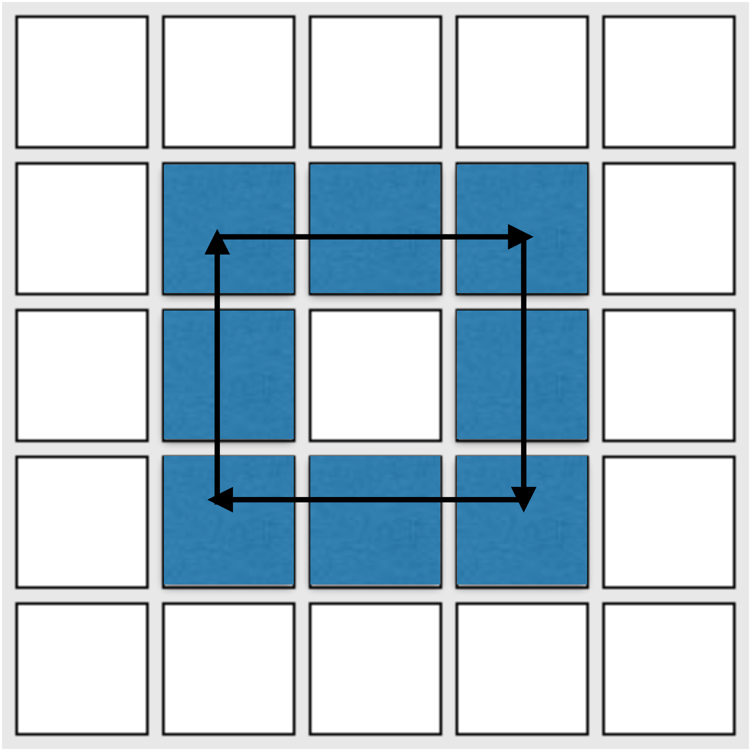
Time! … sequence of observations

Markov Model



- Each node is identically distributed given its predecessor (stationary)
- The values the nodes take are called states
- Parameters?
  - $P(S_1)$ the initial probability table
  - $P(S_t|S_{t-1})$ the transition probabilities

Markov Model



Bot tends to follow outlined path, but with some probability jumps to arbitrary neighbor

- Number of states: 25 (one for each location)
- For white boxes probability of jumping to any of the 4 neighbors is same 1/4
- For Blue boxes, probability of following path is 0.9 and jumping to some other neighbor is 0.0333333

Markov Model

- If we observe the bot long enough, we get an estimate of its behavior (the transition table of jumping from state to state)

- If we observe enough number of times, we can also estimate initial distribution over states

Markov Model

- Inference question: what is probability that we will be in state k at time t? $P(S_t = k)$?

Answer:

$$P(S_t = k) = \sum_{s_1=1}^{K} \cdots \sum_{s_{t-1}=1}^{K} P(S_1 = s_1, \ldots, S_{t-1} = s_{t-1}, S_t = k)$$

$$= \sum_{s_1=1}^{K} \cdots \sum_{s_{t-1}=1}^{K} \prod_{i=1}^{t-1} \left(P(S_i = s_i | S_{i-1} = s_{i-1}) \times P(S_t = k | S_{t-1} = s_{t-1})\right)$$

For every t we can repeat the above or…

$$P(S_t = k) = \sum_{s_{t-1}=1}^{K} P(S_t = k | S_{t-1} = s_{t-1}) P(S_{t-1} = s_{t-1})$$

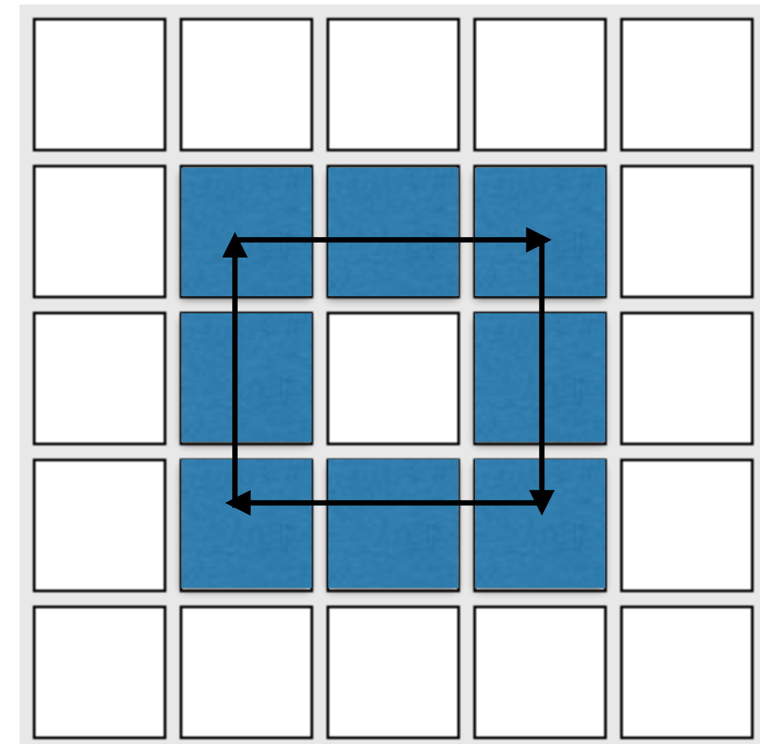
recursively compute probability of previous state

Markov Model

- As time goes by, $P(S_t = k)$ approaches a fixed distribution called stationary distribution
- Without any further observations, you are unlikely to find the bot on a new run (only by luck)

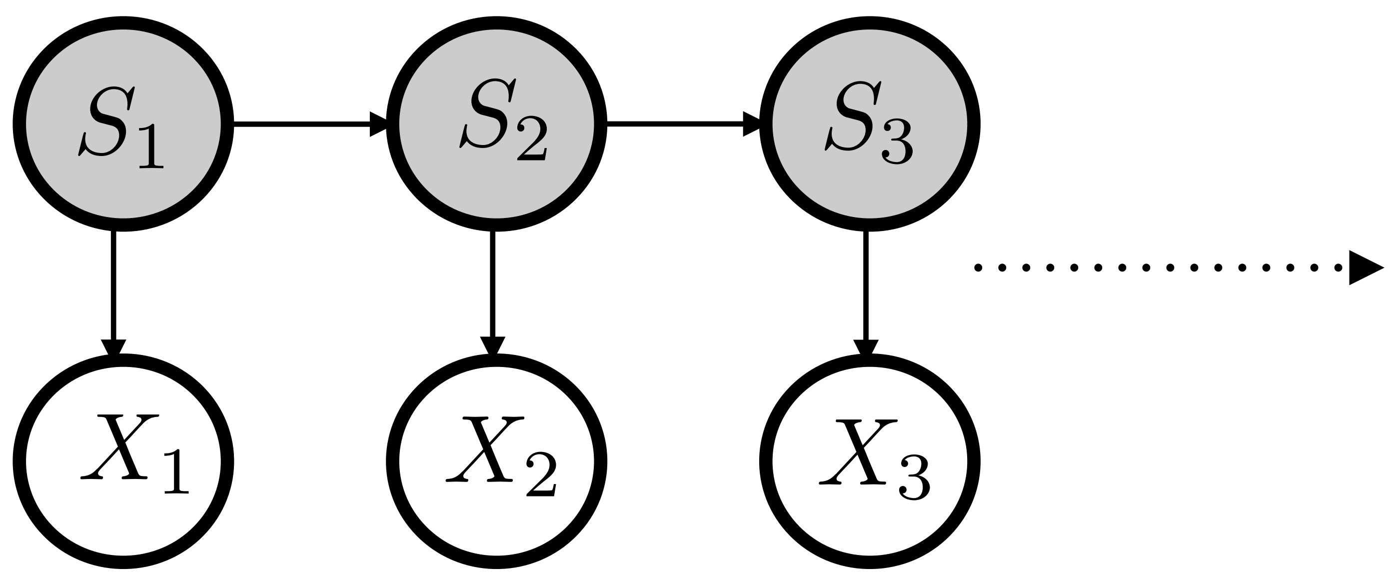
Hidden Markov Model (HMM)

Same example:



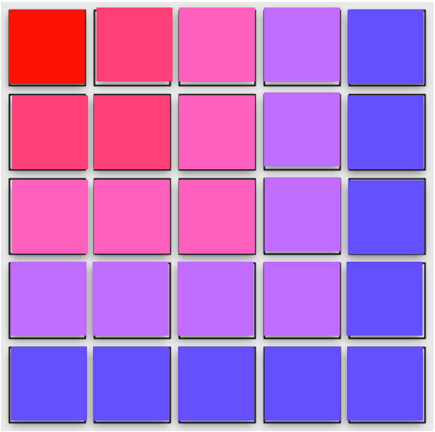
But you don't observe location (dark room)

You hear how close the bot is!



X_t's are loudness of what you hear

HIDDEN MARKOV MODEL (HMM)



- Both during the initial training/estimation phase, you never see the bot you only hear it

- But you hear it at any point in time

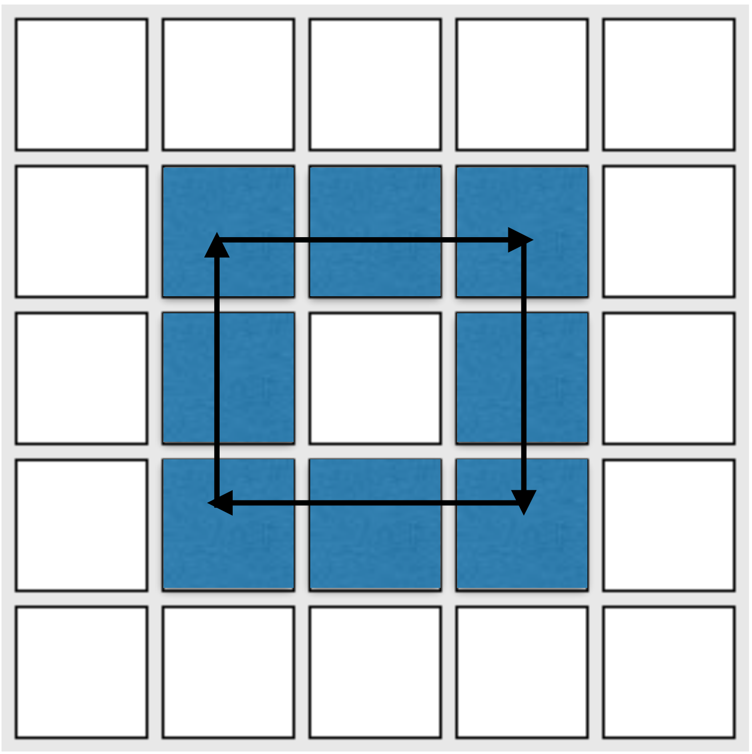
- We will come back to learning next class.

- What is probability that bot will be in state k at time t given the entire sequence of observations?

$$P(S_t = k | X_1, \ldots, X_N)?$$

HIDDEN MARKOV MODEL (HMM)

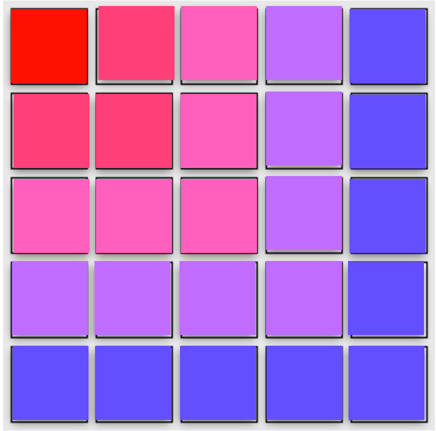
Same example:



But you don't observe location (dark room)

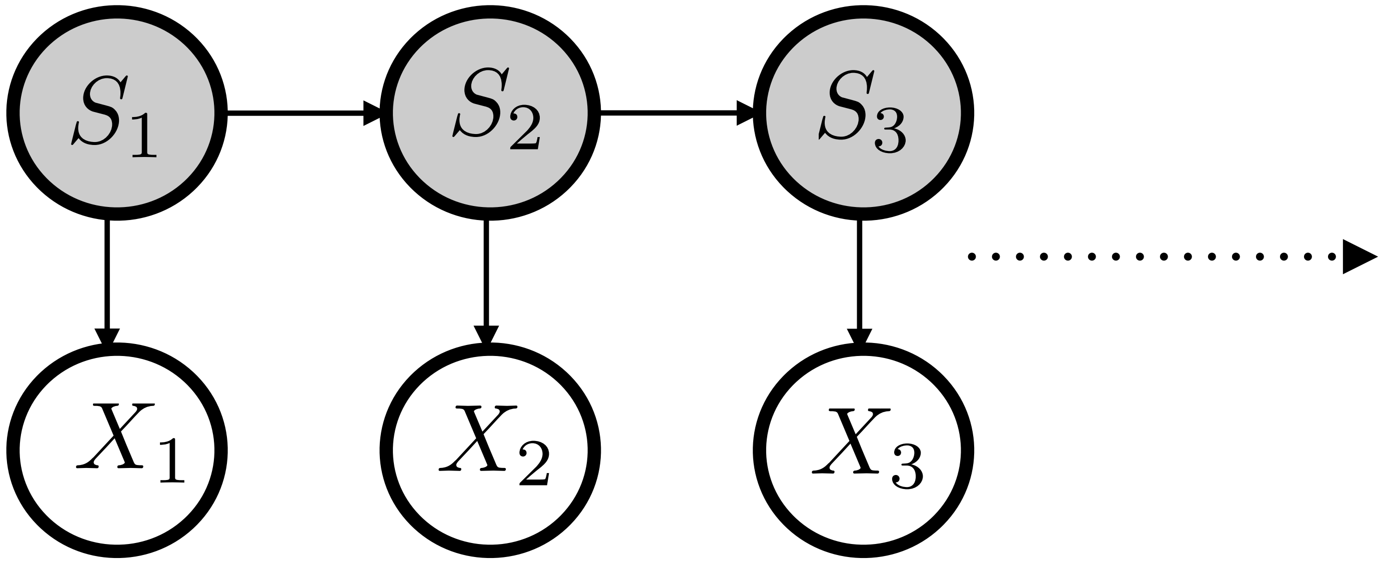
You hear how close the bot is!

What you hear:



Can you catch the Bot?

Hidden Markov Model (HMM)

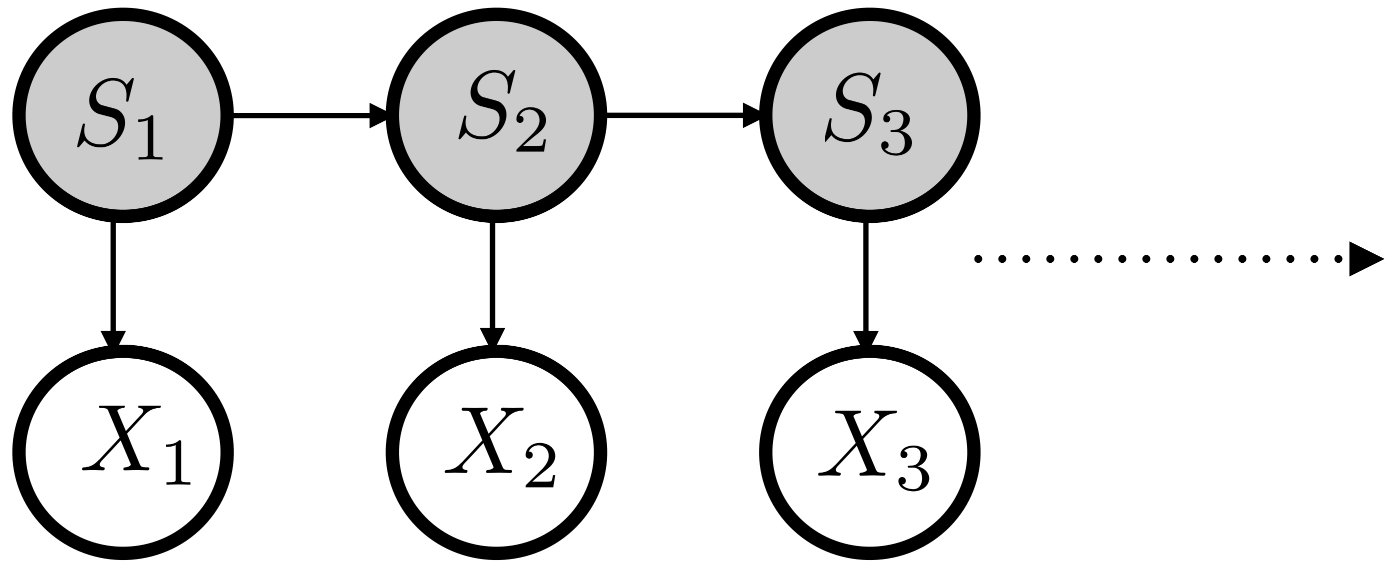


X_t's are what you hear (observation)

S_t's are the unseen locations (states)

Eg: for n x n grid we have, $K = n^2$ states

Number of alphabets = 5
(colors you can observe)



What are the parameters?

Hidden Markov Model (HMM)

- What is probability that bot will be in location k at time t given the entire sequence of observations?

$$P(S_t = k | X_1, \ldots, X_N)?$$

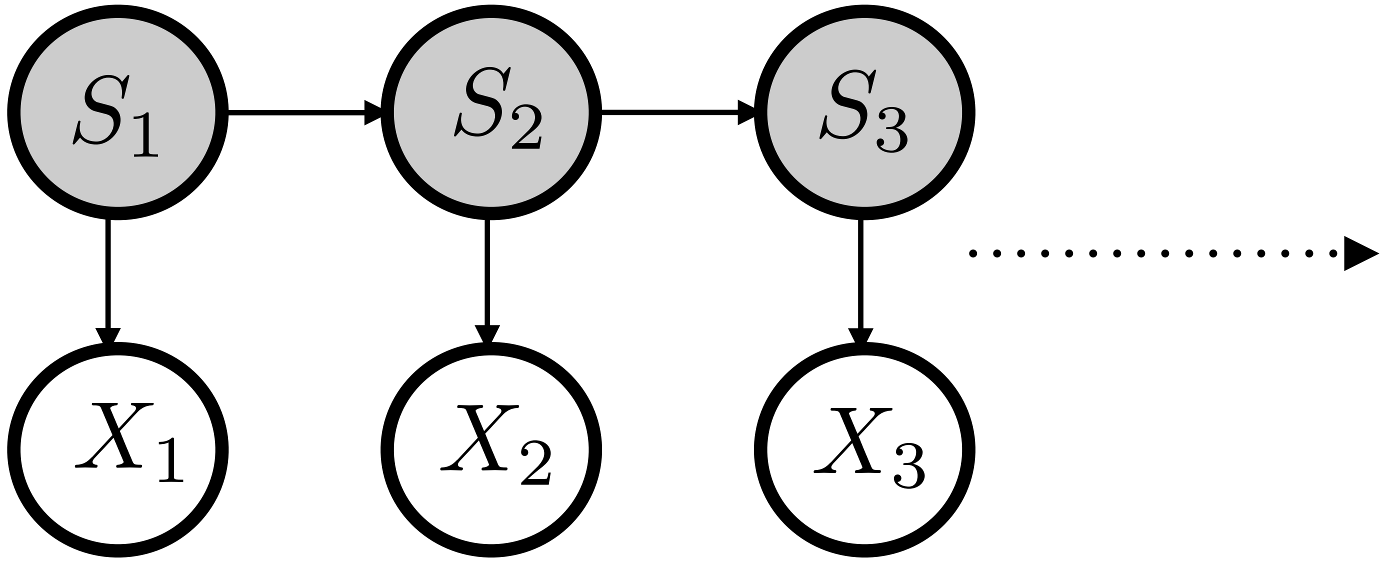
Inference in HMM

$$
\begin{aligned}
&P(S_t = k | X_1, \ldots, X_N) \\
&\quad \propto P(X_{t+1}, \ldots, X_N | S_t = k, X_1, \ldots, X_t) P(S_t = k | X_1, \ldots, X_t) \\
&\quad \propto P(X_{t+1}, \ldots, X_N | S_t = k, X_1, \ldots, X_t) P(S_t = k, X_1, \ldots, X_t) \\
&\quad \propto P(X_{t+1}, \ldots, X_N | S_t = k, X_1, \ldots, X_t) P(X_t | S_t = k, X_1, \ldots, X_{t-1}) P(S_t = k, X_1, \ldots, X_{t-1}) \\
&\quad \propto P(X_{t+1}, \ldots, X_N | S_t = k) P(X_t | S_t = k) P(S_t = k, X_1, \ldots, X_{t-1})
\end{aligned}
$$

We know $P(X_t | S_t = k)$'s and $P(S_t | S_{t-1})$

Compute $P(X_{t+1}, \ldots, X_N)$ and $P(S_t = k, X_1, \ldots, X_{t-1})$ recursively.

Inference in HMM

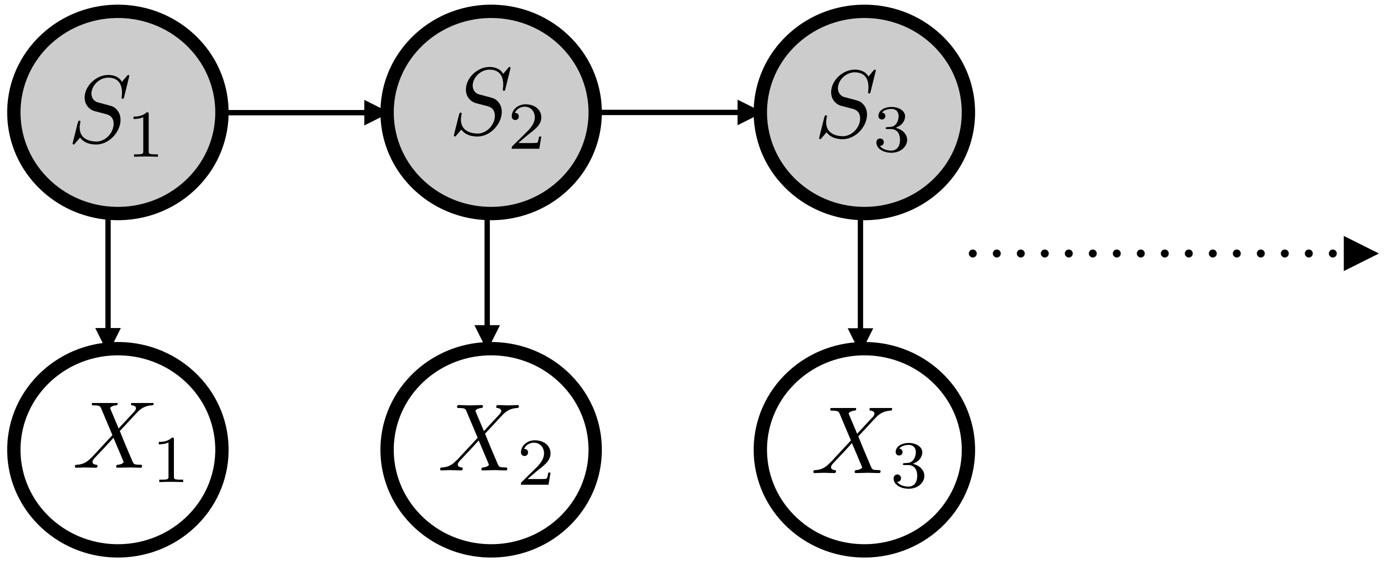


$$\text{message}_{S_{t-1} \mapsto S_t}(k) = P(S_t = k, X_1, \ldots, X_{t-1})$$

$$\text{message}_{S_{t+1} \mapsto S_t}(k) = P(X_n, \ldots, X_{t+1} | S_t = k)$$

$$P(S_t = k | X_1, \ldots, X_n) \propto \text{message}_{S_{t-1} \mapsto S_t}(k) \times \text{message}_{S_{t+1} \mapsto S_t}(k) \times P(X_t | S_t = k)$$

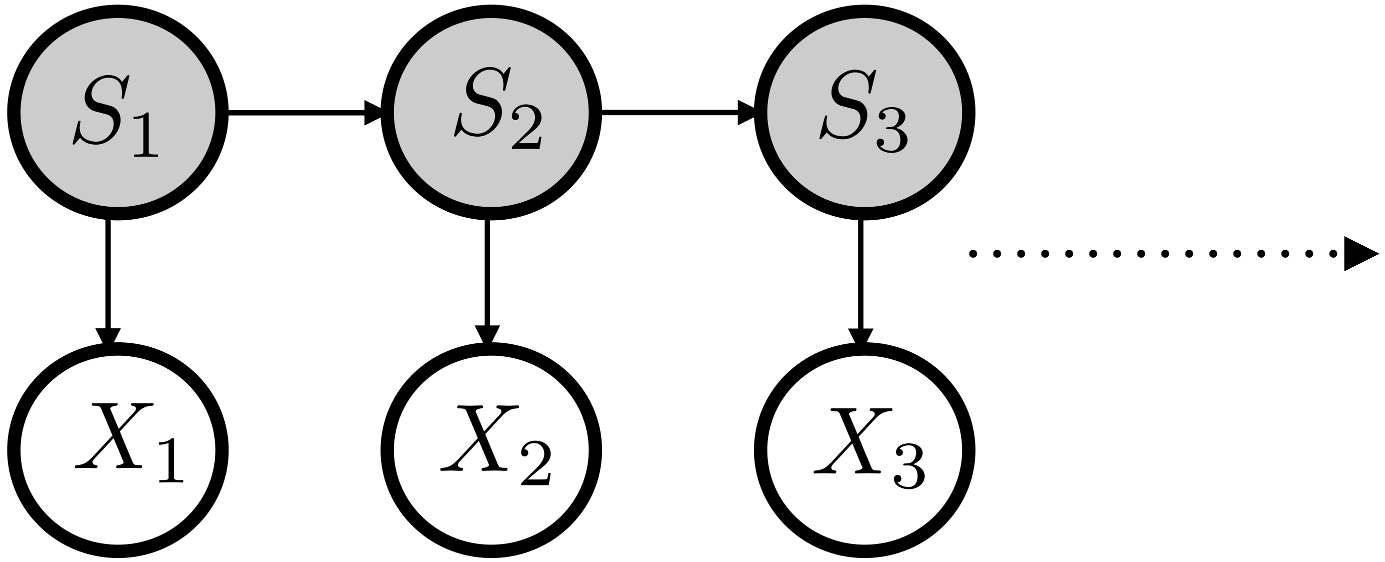
Inference in HMM



$$\text{message}_{S_{t-1} \mapsto S_t}(k) = P(S_t = k, X_1, \ldots, X_{t-1})$$

$$\text{message}_{S_{t+1} \mapsto S_t}(k) = P(X_n, \ldots, X_{t+1} | S_t = k)$$

Inference in HMM



$$\text{message}_{S_{t-1} \mapsto S_t}(k) = P(S_t = k, X_1, \ldots, X_{t-1})$$

$$\text{message}_{S_{t+1} \mapsto S_t}(k) = P(X_n, \ldots, X_{t+1} | S_t = k)$$

Learning Parameters for HMM

- Now that we have algorithm for inference, what about learning
- Given observations, how do we estimate parameters for HMM? Three guesses …