HIDDEN MARKOV MODELS

CSE 511A: Introduction to Artificial Intelligence

Some content and images are from slides created by Dan Klein and Pieter Abbeel for CS188 Intro to AI at UC Berkeley.
All CS188 materials are available at http://ai.berkeley.edu.

BAYESIAN NETWORKS

- Disease: *COVID-19*
- Symptoms: *Fever or chills, cough, shortness of breath or difficulty breathing, fatigue, muscle or body aches, headache, new loss of taste or smell, sore throat, congestion or runny nose, nausea or vomiting, diarrhea*

- Disease: *Flu*
- Symptoms: *Fever or feeling feverish/chills, cough, sore throat, runny or stuffy nose, muscle or body aches, headaches, fatigue, some people may have vomiting and diarrhea, though this is more common in children than adults.*

- Assume that you have a very bad fever and that you are coughing and tired all the time. Which is the more likely prognosis?

BAYESIAN NETWORKS

- Disease: *COVID-19*
- Symptoms: *Fever or chills, cough, shortness of breath or difficulty breathing, fatigue, muscle or body aches, headache, new loss of taste or smell, sore throat, congestion or runny nose, nausea or vomiting, diarrhea*

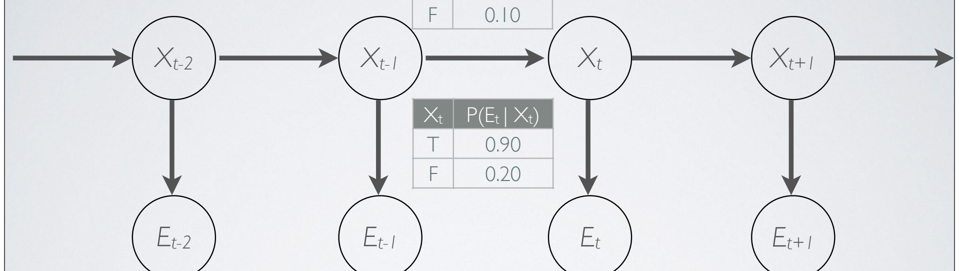
Bayesian networks so far are used for single-shot inference.
What if you want to perform continuous inference?
i.e., perform inference on day 1, day 2, day 3, …
Hint: *Can you still use a Bayesian network for this problem?*

diarrhea, though this is more common in children than adults.

- Assume that you have a very bad fever and that you are coughing and tired all the time. Which is the more likely prognosis?

BAYESIAN NETWORKS

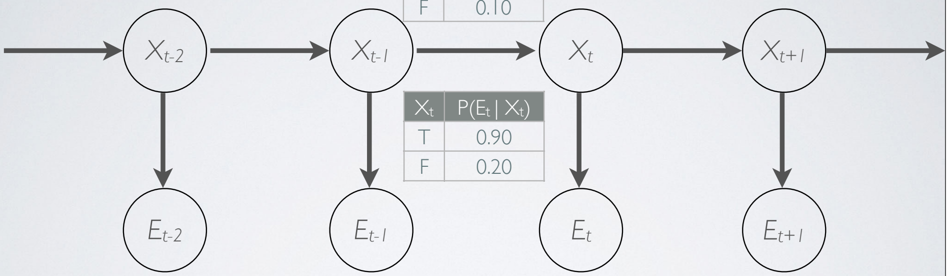
X_{t-1}	$P(X_t \mid X_{t-1})$
T	0.95
F	0.10



X_t	$P(E_t \mid X_t)$
T	0.90
F	0.20

- X_t's: State variables (variables that we want to infer) at each time step *t*. e.g., "have the coronavirus"
- E_t's: Evidence variables (variables that we can observe) at each time step *t*.. e.g., "have fever"
- *Stationary process:* CPTs are identical for all time steps.

HIDDEN MARKOV MODELS

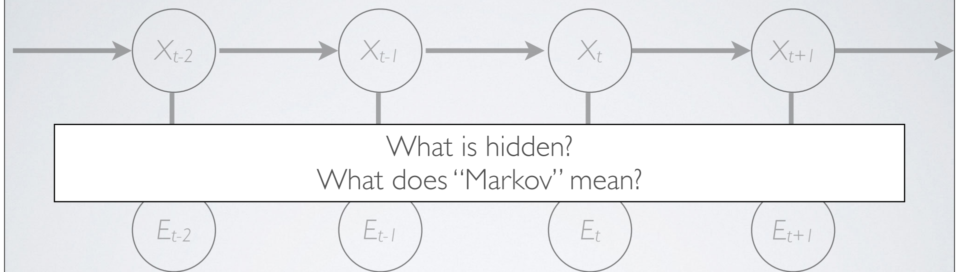


X_{t-1}	$P(X_t \mid X_{t-1})$
T	0.95
F	0.10

X_t	$P(E_t \mid X_t)$
T	0.90
F	0.20

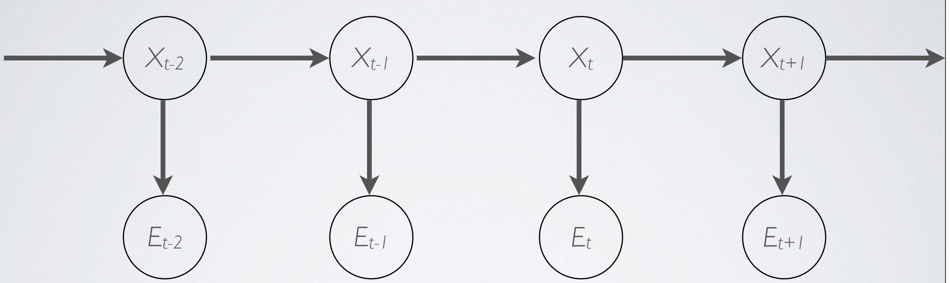
- These kinds of networks are called Hidden Markov Models (HMMs)
- Only *one* state variable.
- Possibly more evidence variables, but usually one.

HIDDEN MARKOV MODELS



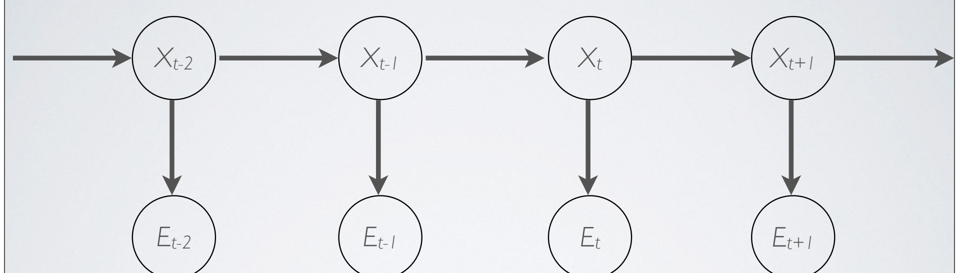
- These kinds of networks are called Hidden Markov Models (HMMs)
- Only *one* state variable.
- Possibly more evidence variables, but usually one.

HIDDEN MARKOV MODELS



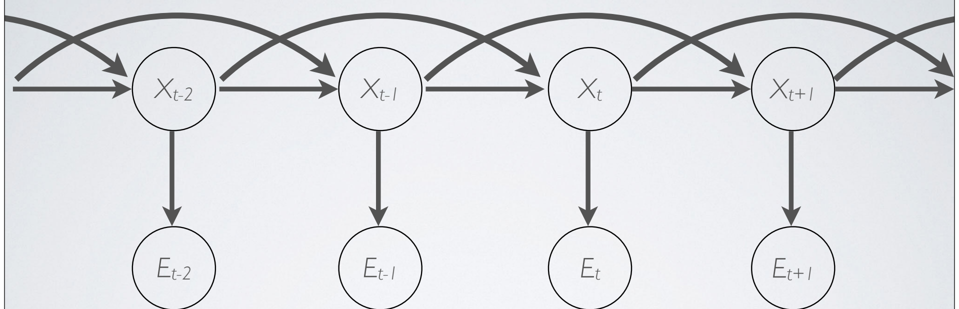
- Hidden: Because the state variables cannot be directly observed.
- Markov: Because the state variables depend only on a *constant* number of previous state variables and NOT the entire history.
  - i.e., $P(X_t \mid X_{t-1}, \ldots, X_1) = P(X_t \mid X_{t-1}, \ldots, X_{t-k})$

HIDDEN MARKOV MODELS



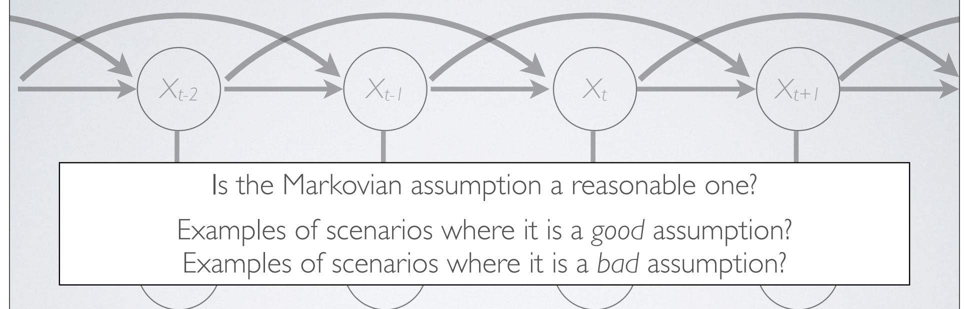
- First-order HMM: State variable depends only on the previous state variable.
  - i.e., $P(X_t \mid X_{t-1}, \ldots, X_1) = P(X_t \mid X_{t-1})$

HIDDEN MARKOV MODELS



- Second-order HMM: State variable depends only on the previous two state variables.
  - i.e., $P(X_t \mid X_{t-1}, \ldots, X_1) = P(X_t \mid X_{t-1}, X_{t-2})$

HIDDEN MARKOV MODELS



Is the Markovian assumption a reasonable one?

Examples of scenarios where it is a *good* assumption?
Examples of scenarios where it is a *bad* assumption?

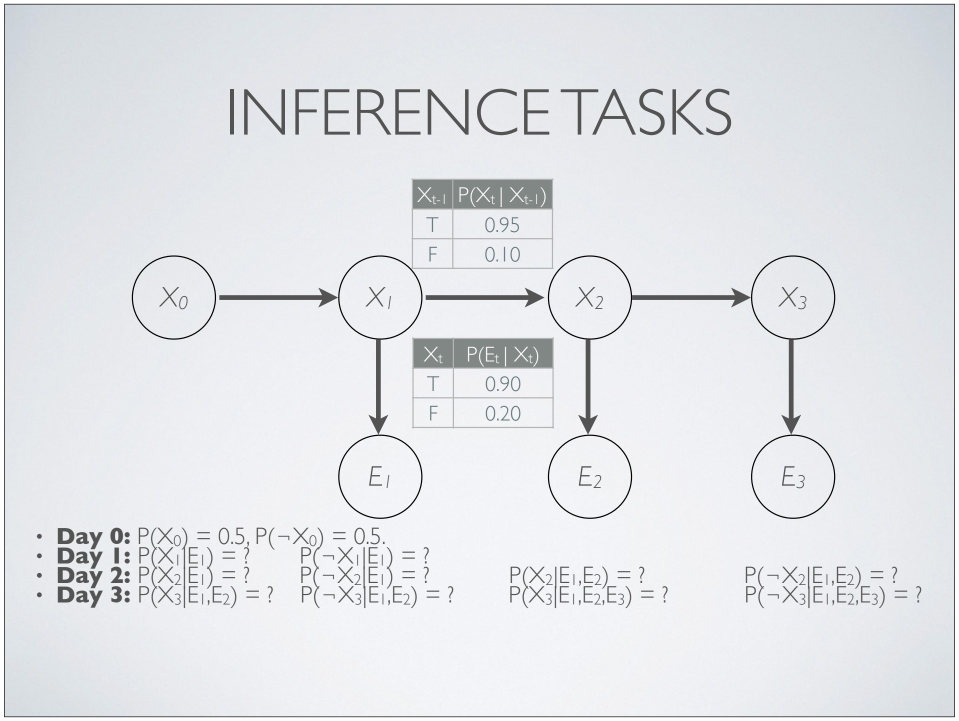
- Second-order HMM: State variable depends only on the previous two state variables.
  - i.e., $P(X_t \mid X_{t-1}, \ldots, X_1) = P(X_t \mid X_{t-1}, X_{t-2})$

INFERENCE TASKS

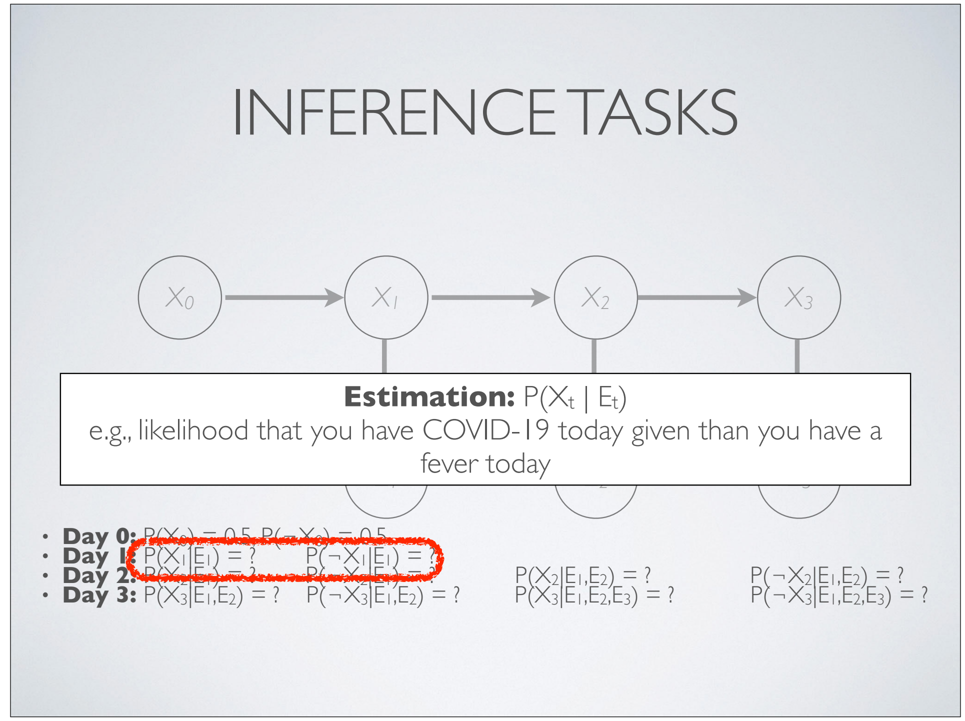
- Estimation: $P(X_t \mid E_t)$
  - e.g., likelihood that you have COVID-19 today given than you have a fever today
- Prediction: $P(X_{t+1} \mid E_t, \ldots, E_1)$
  - e.g., likelihood that you have COVID-19 tomorrow in the future given your fever history
- Filtering: $P(X_t \mid E_t, \ldots, E_1)$
  - e.g., likelihood that you have COVID-19 today given your fever history
- Smoothing: $P(X_k \mid E_t, \ldots, E_1)$
  - e.g., likelihood that you have COVID-19 in the past given your fever history
- Most likely sequence: $\arg\max_{X_1, \ldots, X_t} P(X_1, \ldots, X_t \mid E_t, \ldots, E_1)$
  - e.g., most likely COVID-19 history given your fever history

INFERENCE TASKS

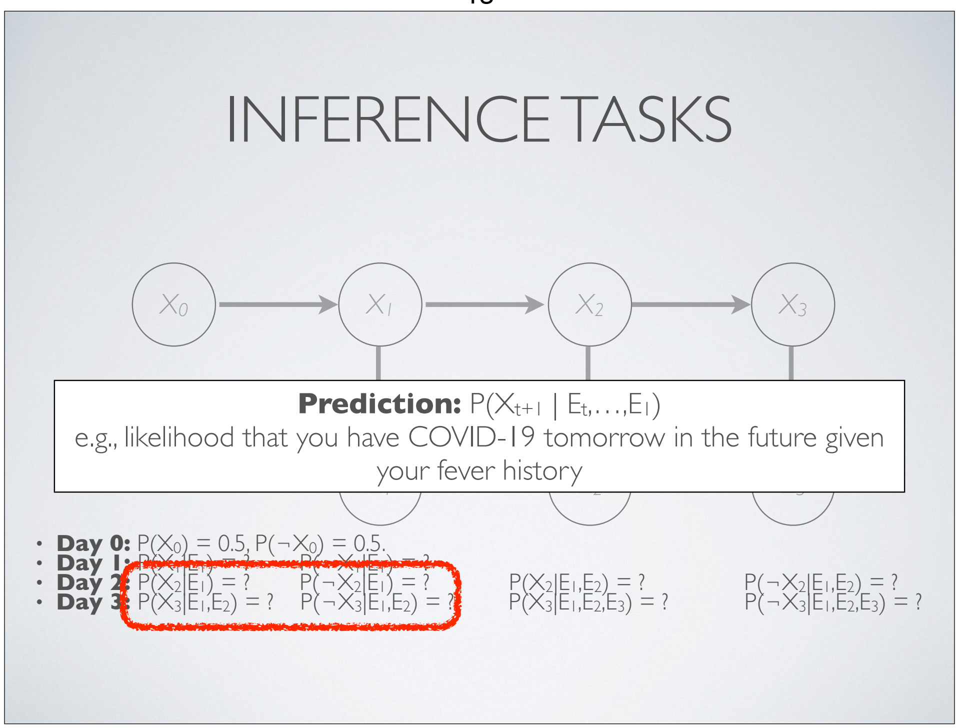
- Estimation: $P(X_t \mid E_t)$
  - e.g., likelihood that you have COVID-19 today given than you have a fever today
- Prediction: $P(X_{t+1} \mid E_t, \ldots, E_1)$
  - e.g., likelihood that you have COVID-19 tomorrow in the future given your fever history
- Filtering: $P(X_t \mid E_t, \ldots, E_1)$
  - e.g., likelihood that you have COVID-19 today given your fever history
- Smoothing: $P(X_k \mid E_t, \ldots, E_1)$
  - e.g., likelihood that you have COVID-19 in the past given your fever history
- Most likely sequence: $\arg\max_{X_1, \ldots, X_t} P(X_1, \ldots, X_t \mid E_t, \ldots, E_1)$
  - e.g., most likely COVID-19 history given your fever history


INFERENCE TASKS
X_{t-1} | $P(X_t \mid X_{t-1})$
T | 0.95
F | 0.10
X_t | $P(E_t \mid X_t)$
T | 0.90
F | 0.20
X_0
X_1
X_2
X_3
E_1
E_2
E_3
Day 0: $P(X_0) = 0.5, P(\neg X_0) = 0.5$
Day 1: $P(X_1|E_1) = ?$ $P(\neg X_1|E_1) = ?$
Day 2: $P(X_2|E_1) = ?$ $P(\neg X_2|E_1) = ?$ $P(X_2|E_1,E_2) = ?$ $P(\neg X_2|E_1,E_2) = ?$
Day 3: $P(X_3|E_1,E_2) = ?$ $P(\neg X_3|E_1,E_2) = ?$ $P(X_3|E_1,E_2,E_3) = ?$ $P(\neg X_3|E_1,E_2,E_3) = ?$

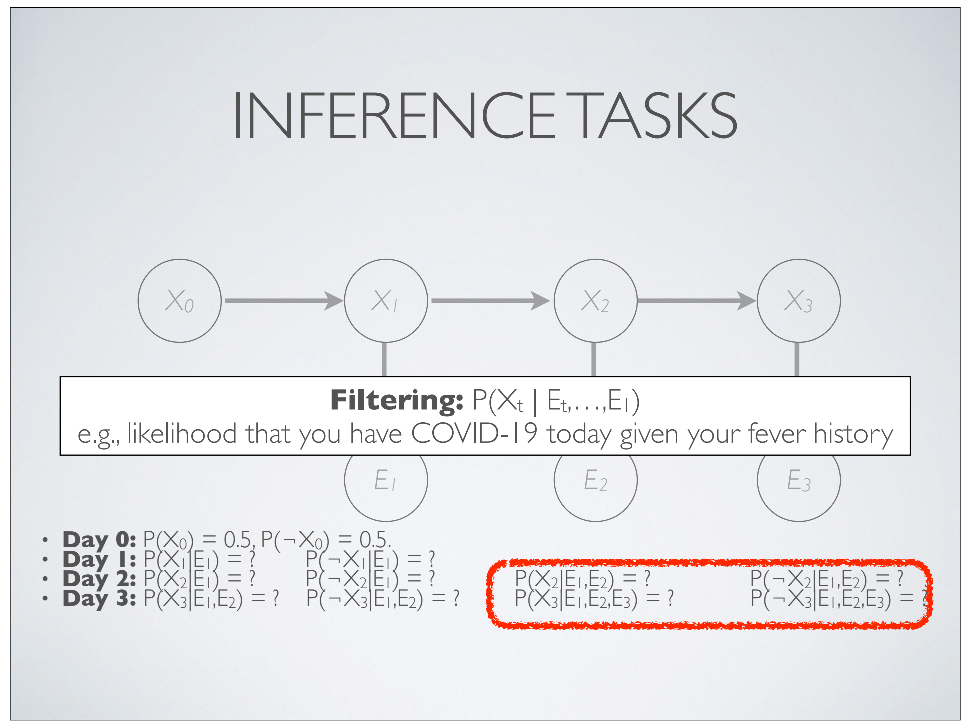
13


INFERENCE TASKS
X_0
X_1
X_2
X_3
Estimation: $P(X_t \mid E_t)$
e.g., likelihood that you have COVID-19 today given than you have a fever today
Day 1: $P(X_1|E_1) = ?$ $P(\neg X_1|E_1) = ?$
$P(X_2|E_1,E_2) = ?$ $P(\neg X_2|E_1,E_2) = ?$
Day 3: $P(X_3|E_1,E_2) = ?$ $P(\neg X_3|E_1,E_2) = ?$ $P(X_3|E_1,E_2,E_3) = ?$ $P(\neg X_3|E_1,E_2,E_3) = ?$

14


INFERENCE TASKS
X_0
X_1
X_2
X_3
Prediction: $P(X_{t+1} \mid E_t,\ldots,E_1)$
e.g., likelihood that you have COVID-19 tomorrow in the future given your fever history
Day 0: $P(X_0) = 0.5, P(\neg X_0) = 0.5$
Day 2: $P(X_2|E_1) = ?$ $P(\neg X_2|E_1) = ?$ $P(X_2|E_1,E_2) = ?$ $P(\neg X_2|E_1,E_2) = ?$
Day 3: $P(X_3|E_1,E_2) = ?$ $P(\neg X_3|E_1,E_2) = ?$ $P(X_3|E_1,E_2,E_3) = ?$ $P(\neg X_3|E_1,E_2,E_3) = ?$

15


INFERENCE TASKS
X_0
X_1
X_2
X_3
Filtering: $P(X_t \mid E_t,\ldots,E_1)$
e.g., likelihood that you have COVID-19 today given your fever history
E_1
E_2
E_3
Day 0: $P(X_0) = 0.5, P(\neg X_0) = 0.5$
Day 1: $P(X_1|E_1) = ?$ $P(\neg X_1|E_1) = ?$
Day 2: $P(X_2|E_1) = ?$ $P(\neg X_2|E_1) = ?$ $P(X_2|E_1,E_2) = ?$ $P(\neg X_2|E_1,E_2) = ?$
Day 3: $P(X_3|E_1,E_2) = ?$ $P(\neg X_3|E_1,E_2) = ?$ $P(X_3|E_1,E_2,E_3) = ?$ $P(\neg X_3|E_1,E_2,E_3) = ?$

16

INFERENCE TASKS

- Estimation: $P(X_t \mid E_t)$
  - e.g., likelihood that you have COVID-19 today given than you have a fever today
- Prediction: $P(X_{t+1} \mid E_t,\ldots,E_1)$
  - e.g., likelihood that you have COVID-19 tomorrow in the future given your fever history
- Filtering: $P(X_t \mid E_t,\ldots,E_1)$
  - e.g., likelihood that you have COVID-19 today given your fever history
- Smoothing: $P(X_k \mid E_t,\ldots,E_1)$
  - e.g., likelihood that you have COVID-19 in the past given your fever history
- Most likely sequence: $\text{argmax}_{X_1,\ldots,X_t} P(X_1,\ldots,X_t \mid E_t,\ldots,E_1)$
  - e.g., most likely COVID-19 history given your fever history

Solved more systematically using the Forward-Backward algorithm
See textbook for description