CMSC 671
Fall 2010

Thu 11/04/10

Probabilistic Reasoning over Time
Chapter 15.1 – 15.2, 15.7

Prof. Laura Zavala, laura.zavala@umbc.edu, ITE 373, 410-455-8775

Time and Uncertainty

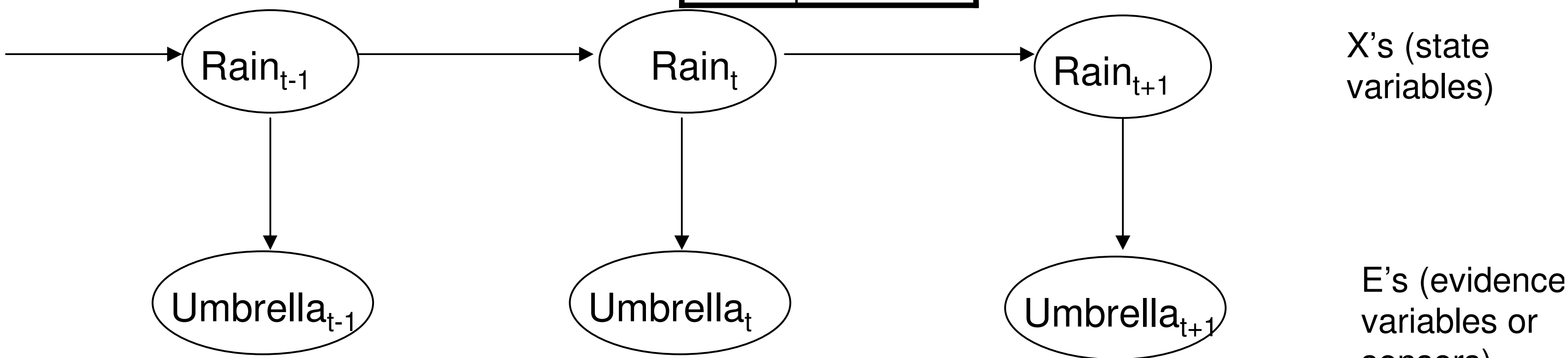
- The world changes, we need to track and predict it
- Examples: diabetes management, traffic monitoring
- Basic idea: copy state and evidence variables for each time step
  - X_t – set of unobservable state variables at time t
    - e.g., $BloodSugar_t$, $StomachContents_t$
  - E_t – set of evidence variables at time t
    - e.g., $MeasuredBloodSugar_t$, $PulseRate_t$, $FoodEaten_t$
- Assumes **discrete** time steps

States and Observations

- Process of change is viewed as series of snapshots, each describing the state of the world at a particular time
- Each time slice involves a set or random variables indexed by t:
  1. the set of unobservable state variables X_t
  2. the set of observable evidence variable E_t
- The observation at time t is $E_t = e_t$ for some set of values e_t
- The notation $X_{a:b}$ denotes the set of variables from X_a to X_b

Example

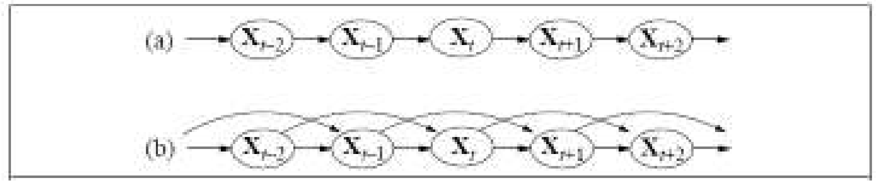
R_{t-1}	$P(R_t \mid R_{t-1})$
T	0.7
F	0.3



R_t	$P(U_t \mid R_t)$
T	0.9
F	0.2

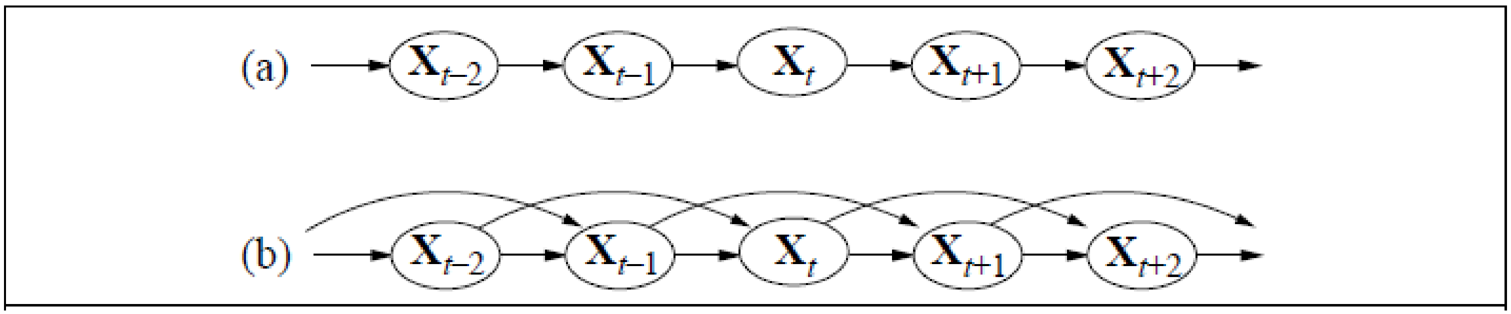
- What is the transition model?
- What is the sensor model?

Stationary Process/Markov Assumption



- Markov Assumption: X_t depends on some previous X_is
  - First-order Markov process:
    $P(X_t|X_{0:t-1}) = P(X_t|X_{t-1})$
  - kth order: depends on previous k time steps
- Sensor Markov assumption:
  $$P(E_t|X_{0:t}, E_{0:t-1}) = P(E_t|X_t)$$
- Assume stationary process: transition model $P(X_t|X_{t-1})$ and sensor model $P(E_t|X_t)$ are the same for all t
  - In a **stationary process**, the changes in the world state are governed by laws that do not themselves change over time
  - The process of change doesn't change

First-order and second-order Markov processes



Complete Joint Distribution

- Given:
  - Transition model: $P(X_t|X_{t-1})$
  - Sensor model: $P(E_t|X_t)$
  - Prior probability: $P(X_0)$
- Then we can specify complete joint distribution:
  - Full joint distribution for BN (slide 10 last class)

$$\boldsymbol{P}(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n) = \Pi_{i=1}^{n} \boldsymbol{P}(\boldsymbol{x}_i \mid \boldsymbol{\pi}_i)$$

  - Using that equation, for any t:

$$P(X_0, X_1, \ldots, X_t, E_1, \ldots, E_t) = P(X_0) \prod_{i=1}^{t} P(X_i \mid X_{i-1}) P(E_i \mid X_i)$$

Inference Tasks

- **Filtering or monitoring**: $P(X_t|e_1,\ldots,e_t)$
  computing current belief state, given all evidence to date
- **Prediction**: $P(X_{t+k}|e_1,\ldots,e_t)$
  computing prob. of some future state
- **Smoothing**: $P(X_k|e_1,\ldots,e_t)$
  computing prob. of past state (hindsight)
- **Most likely explanation**:
  $\arg\max_{x1,..xt} P(x_1,\ldots,x_t|e_1,\ldots,e_t)$
  given sequence of observation, find sequence of states that is most likely to have generated those observations.

Examples

- **Filtering:** What is the probability that it is raining today, given all the umbrella observations up through today?
- **Prediction:** What is the probability that it will rain the day after tomorrow, given all the umbrella observations up through today?
- **Smoothing:** What is the probability that it rained yesterday, given all the umbrella observations through today?
- **Most likely explanation:** if the umbrella appeared the first three days but not on the fourth, what is the most likely weather sequence to produce these umbrella sightings?

Learning

- Besides the inference tasks, we can also learn the transition and sensor models from observations.
- **EM algorithm (chapter 20)**
  - Models are updated with estimates from Inference
    - What transitions occurred and what states generated the sensors readings
  - Updated model provides new estimates
  - The process iterates to convergence

Filtering

- **Filtering**: $P(X_t|e_1,\ldots,e_t)$ computing current belief state, given all evidence to date
- **Example**: What is the probability that it is raining today, given all the umbrella observations up through today?
  - We use recursive estimation to compute $P(X_{t+1} \mid e_{1:t+1})$ as a function of e_{t+1} and $P(X_t \mid e_{1:t})$
  - We can write this as follows:

$$
\begin{aligned}
P(X_{t+1} \mid e_{1:t+1}) &= P(X_{t+1} \mid e_{1:t}, e_{t+1}) \\
&= \alpha P(e_{t+1} \mid X_{t+1}, e_{1:t}) P(X_{t+1} \mid e_{1:t}) \\
&= \alpha P(e_{t+1} \mid X_{t+1}) P(X_{t+1} \mid e_{1:t}) \\
&= \alpha P(e_{t+1} \mid X_{t+1}) \sum_{x_t} P(X_{t+1} \mid x_t) P(x_t \mid e_{1:t})
\end{aligned}
$$

  - This leads to a recursive definition
    - $f_{1:t+1} = \alpha \text{FORWARD}(f_{1:t:t}, e_{t+1})$

Prediction

- **Prediction**: $P(X_{t+k}|e_1,\ldots,e_t)$ computing probability of some future state
- **Example**: What is the probability that it will rain the day after tomorrow, given all the umbrella observations up through today?

- Filtering without the addition of new evidence (e_{t+1})

$P(X_{t+1} | e_{1:t})$ instead of $P(X_{t+1} | e_{1:+1})$

Smoothing

- **Smoothing**: $P(X_k|e_1,\ldots,e_t)$ computing probability of past state (hindsight)
- **Example**: What is the probability that it rained yesterday, given all the umbrella observations through today?
  - Compute $P(X_k|e_{1:t})$ for $0 <= k < t$
  - Using a backward message $b_{k+1:t} = P(E_{k+1:t} | X_k)$, we obtain
    - $P(X_k|e_{1:t}) = \alpha f_{1:k} b_{k+1:t}$
  - The backward message can be computed using

$$b_{k+1:t} = \sum_{x_{k+1}} P(e_{k+1} | x_{k+1}) P(e_{k+2:t} | x_{k+1}) P(x_{k+1} | X_k)$$

  - This leads to a recursive definition
    - $B_{k+1:t} = \alpha \mathrm{BACKWARD}(b_{k+2:t}, e_{k+1:t})$

Probabilistic Temporal Models

- Hidden Markov Models (HMMs)
  - One single state variable (umbrella example is an HMM)
  - For problems with more than one variable, vars are combined into a single “megavariable” with tuples of values. E.g. The state var. for the vacuum world (localization of a robot) is the set of empty squares
- Kalman Filters
  - Handling continuous variables
- Dynamic Bayesian Networks (DBNs)
  - Any number of state variables and evidence variables
  - Includes the previous two