Tufts

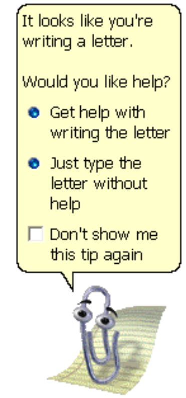
Lecture 14:
Dynamic Bayes Nets

Artificial Intelligence (CS 131)

1

Uses of Bayes Nets

- Older versions of MS Office used Bayes Nets to run the "Intelligent Assistant" program, including animated paper-clip, "Clippy"
- Tracked user behavior to see if it should suggest help, and to determine what sort of help the user might need
- Probably the least popular Bayes Net in the history of mankind!



2

Microsoft Assistants

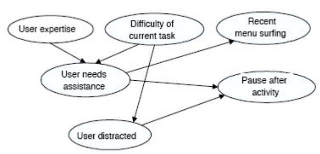
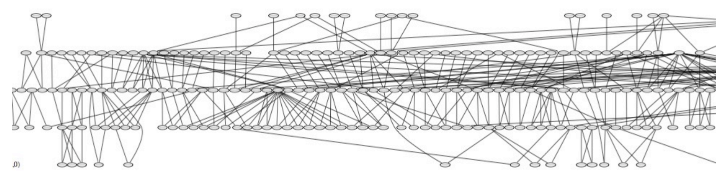


Figure 2: A portion of a Bayesian user model for inferring the likelihood that a user needs assistance, considering profile information as well as observations of recent activity.

- The BN was used to predict the "Needs Assistance" variable
- Reasoning based on prior distributions over how hard certain things were to do in Office, and on how expert users were likely to be
- Also used evidence, taken from things like menu use, clicking, waiting, re-doing or un-doing things…

3

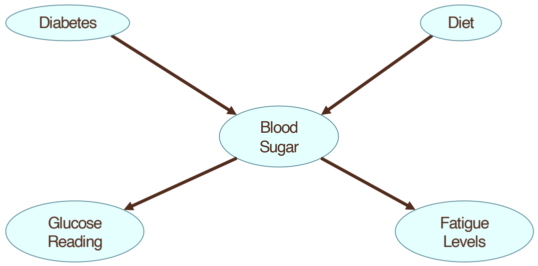
A More Successful Example



- BN for tracking genetic tendencies behind pig diseases (something very important to the farm industry)

4

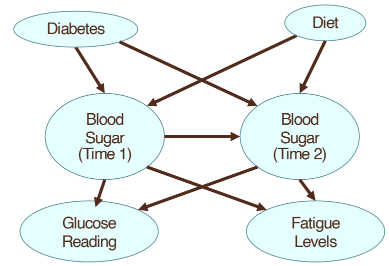
A Bayesian Model of Diabetes



- A simple Bayesian model of diabetes and blood sugar
- Blood sugar *changes over time*, based on its prior level

5

Modeling Diabetes Over Time



- To build in temporal relations, we could duplicate the blood-sugar node, adding in effects over time
- This provides a more complicated model, and becomes cumbersome when we are dealing with long time-spans, and many different time-steps

6

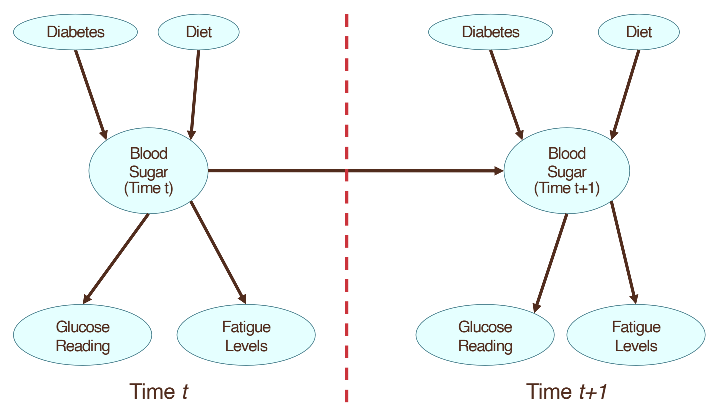
Temporal Probability Models

- We would like to keep our model simple, and still represent long stretches of time
  - This is made possible where we have processes that are essentially the *same* at every time-step
- Our models represent change over time:

1. States of model depend upon states at *previous* time(s)
2. Observations and other local variables at any point in time depend only upon the *current* state

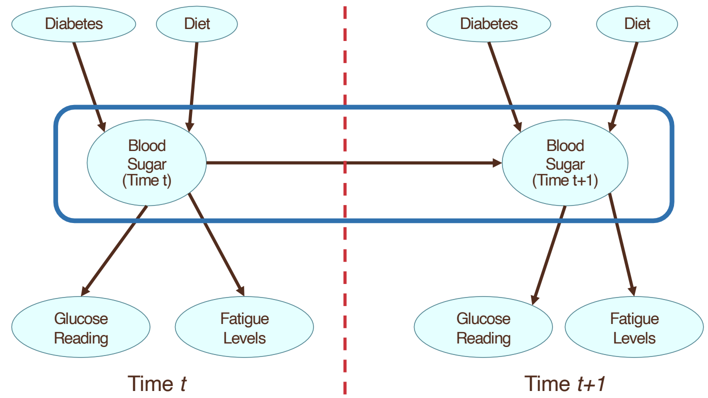
7

Dynamic BN: a Temporal Model



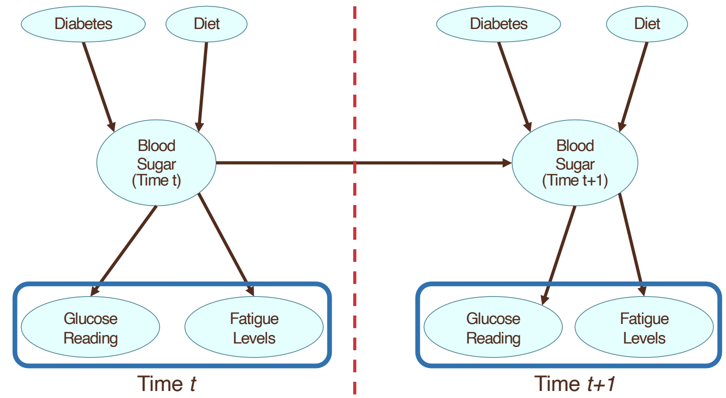
8

States Depend upon Prior States



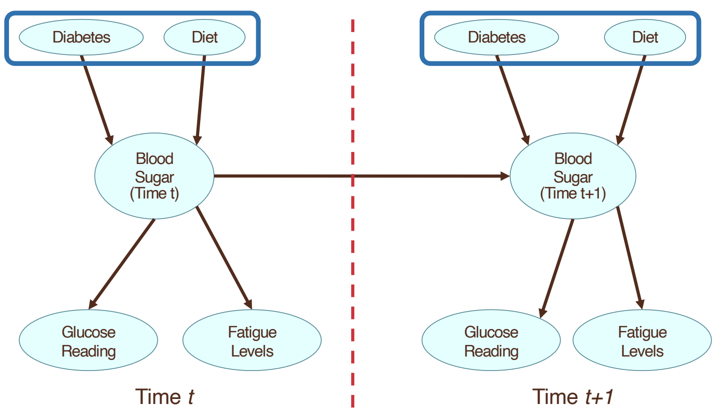
9

Observations Depend upon the Current State of the System Alone



10

Local Variables Are Not Observed, and Do Not Depend upon Prior Time-steps



11

Specifying Dynamic Bayes Nets (DBNs)

- We must specify all of the following:
  1. Transition model: probability of going from one state, X_t, to another, X_{t+1}, at the next time-step
  2. Observation model: probability of an observation, E_t, based on the current state
  3. Prior distribution on initial state: $P(X_0)$

- This then defines a complete joint distribution:

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

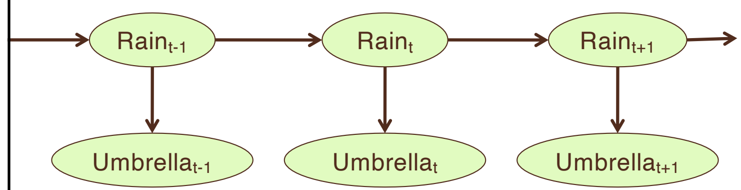
12

Some Basic Assumptions

- Stationary process: the system works the same way at every point in time
  - Even if variables change over time, they change in the same way
- Markov assumption: current system state depends only upon *some finite number* of previous states
  - Basic (*first-order*) Markov Process: state of system only depends on the *one* state immediately before
  - *Second-order*: state depends upon the prior *two* states
  - n-order: depends upon prior n prior states

13

Transition Model



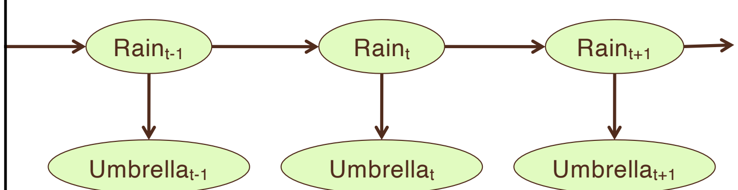
- Probability that system state will change over time
  - Stationary assumption: probabilities are the same for all times t
  - Markov assumption: state X_t is conditionally independent of all other states, given evidence about those that influence it
  - For a 1st-order model, this means:

$$P(X_t \mid X_{0:t-1}) = P(X_t \mid X_{t-1})$$

$X_{0:t-1}$ abbreviates: $X_0, X_1, X_2, \ldots, X_{t-1}$

14

Observation Model



- Probability of observation, E_t, given current state, X_t
  - E_t is conditionally independent of everything else, given evidence about the local state:

$$P(E_t \mid X_{0:t},\, E_{0:t-1}) = P(E_t \mid X_t)$$

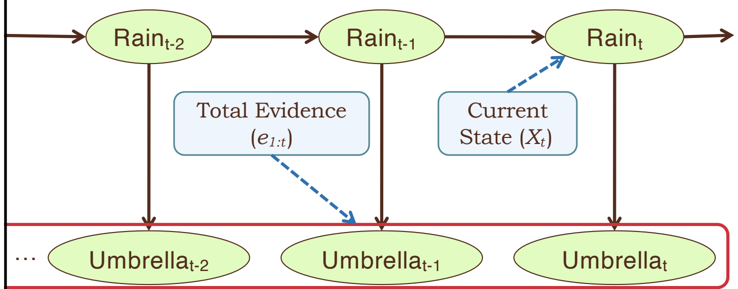
15

Forms of Inference in DBNs

- Filtering: $\mathbf{P}(X_t \mid e_{1:t})$
  - Expresses a belief state about the current state, given sequence of past evidence
- Prediction: $\mathbf{P}(X_{t+k} \mid e_{1:t}),\ k > 0$
  - Used to evaluate *future* possibilities based upon past evidence
  - Used in *planning* for the future
- Smoothing: $\mathbf{P}(X_k \mid e_{1:t}),\ 0 \leq k < t$
  - Gives a better estimate of *past* states based on *new* evidence
- Most likely explanation: $\arg\max_{x_{1:t}} \mathbf{P}(x_{1:t} \mid e_{1:t})$
  - Gives state-sequence that is most likely based upon history
  - Useful in many tasks like speech recognition, robot localization

16

Inference: Filtering



- What is the probability of current state, given all evidence up to that point? $P(X_t \mid e_{1:t})$

Recursive Belief Updates

- The sequence of things we have seen over time ($e_{1:t}$) is our observation history
  - Nobody has infinite memory, however
  - When we want to figure out how likely some state is, given our observation history, we don't want to have to remember *every* observation we ever had
- We look for a way to update *incrementally*, so that we keep only a current probability at every time-step

Recursive Belief Updates

- Best if we can make our *current* belief about the state of the system a function of our *latest observation* and our *old* belief (i.e., about what the state was *before* the latest observation):

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

Recursive Belief Updates

- We can get the *old* belief as a function of the one *before it*:

$$P(X_{t+1} \mid e_{1:t+1}) = f(e_{t+1}, \boxed{P(X_t \mid e_{1:t})})$$

$$\boxed{f(e_t,\, P(X_{t-1} \mid e_{1:t-1}))}$$

Recursive Belief Updates

- We then repeat this for the time step before *that*:

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

$$f(e_t,\, \boxed{P(X_{t-1} \mid e_{1:t-1})})$$

$$\boxed{f(e_{t-1},\, P(X_{t-2} \mid e_{1:t-2}))}$$

21

Recursive Belief Updates

- And so on, until we get down to the base level:

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

$$f(e_t,\, P(X_{t-1} \mid e_{1:t-1}))$$

$$f(e_{t-1},\, \boxed{P(X_{t-2} \mid e_{1:t-2})})$$

(*t* – 3 more steps)

$$\boxed{f(e_1,\, P(X_0))}$$

22

Updating Beliefs Over Time

- All of this work is necessary only if we *wait* until lots of evidence comes in first, however
- If we keep track of our belief-state *from the beginning*, we can update our current beliefs based only upon the *immediate prior* belief

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

3. …and compute a belief state about what the world is like at time *t + 1*

2. …we can take our latest observational evidence…

1. If we have *already* saved a belief state about what the world was like at time *t*…

23