ICS Summer Academy Session II
Topic 7: Reasoning Over Time

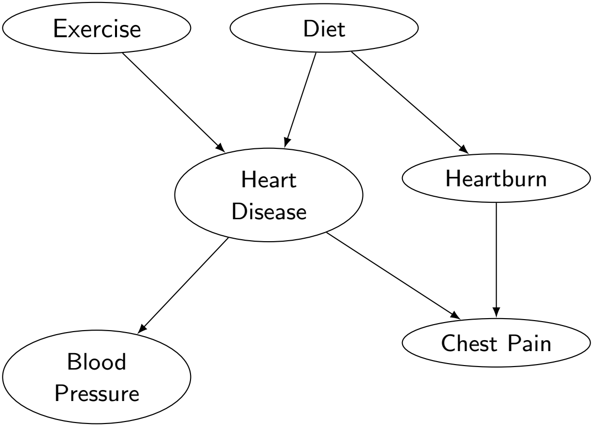
Michael Shindler

Reasoning Over Time

Observations from Thursday/Friday last week:

- Probability distributions can be quite large
- Independence (and conditional) reduce probabilities needed
- Bayesian Networks represent dependencies among variables
  Edges: X has a direct influence on Y
- Today: what about influence over time?

Observed and hidden variables



Example of Reasoning Over Time: "Umbrella World"

- Suppose weather is "rainy" XOR "sunny."
- Suppose tomorrow's weather depends only on today's

Here is a problem you might want to solve:

- Today is rainy
- If today is rainy, tomorrow is rainy 50% of the time
- If today is sunny, tomorrow is sunny 80% of the time
- If today is rainy, what is the probability ten days from now is rainy?

How to Solve in Probability Class?

- Today (day 0) is rainy.
- Probability tomorrow:
  - Rainy: 0.5
  - Sunny: 0.5
- How could day two be rainy? Sunny?
- Probability day two?
  - Rainy:
  - Sunny:

Stationary Markov Chain

Definition

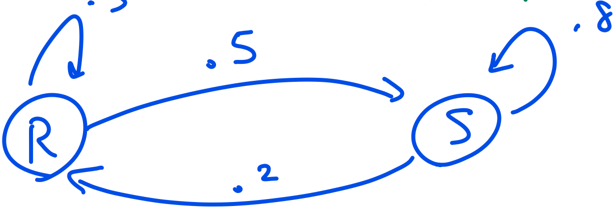
Given ordered random variables $X_1, X_2, \cdots, X_t, \cdots, X_T$, called *states*,

- **Transition probability**

$$P(X_t = \text{value}' | X_{t-1} = \text{value}) \quad a_{ij}$$

- **Initial probability**

$$P(X_1 = \text{value}) \quad \pi_i$$



Text Prediction

This is another prediction of Stationary Markov Chain

- "States" are words
- Probability of "next word" depends on previous word
- How to estimate the probabilities?
  - Suppose all sentences length T
  - We have a file with **MANY** sentences
- We need to compute

$$P(X_1 = x_1, X_2 = x_2, \cdots, X_T = x_T)$$

We use the Markov property to factor

$$P(X_1 = x_1, X_2 = x_2, \cdots, X_T = x_T) =$$

$$P(X_1 = x_1) \prod_{t=2}^{T} P(X_t = x_t | X_{t-1} = x_{t-1})$$

Maximum likelihood estimation

We can use Maximum likelihood estimation

$$\pi_i = \frac{\#\text{of sequences starting with } i}{\#\text{of sequences}}$$

and

$$a_{ij} = \frac{\#\text{of transitions starting with } i \text{ but ending with } j}{\#\text{of transitions starting with } i}$$

Suppose we have two possible states $X_t \in \{0, 1\}$, and we have observed the following 3 sequences

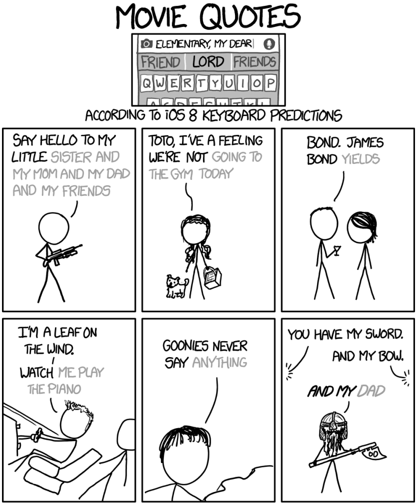
$$\begin{matrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{matrix}$$

Give estimates for each π_i and each a_{ij}.

$\pi_0 = 1/3 \quad \pi_1 = \frac{2}{3}$

$a_{00} = 1/3 \quad a_{01} = 2/3$

$a_{10} = 1/6 \quad a_{11} = 5/6$



XKCD # 1427; More actual results:

- 'Hello. My name is Inigo Montoya. You [are the best. The best thing ever]',
- 'Revenge is a dish best served [by a group of people in my room]', and
- 'They may take our lives, but they'll never take our [money].'

But we need more context...

We have assumed the following Markov property

$$P(X_t|X_1, X_2, \cdots, X_{t-1}) = P(X_t|X_{t-1})$$

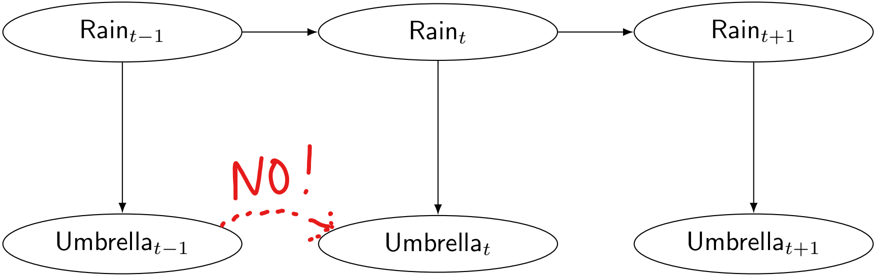
that is why we are only concerning with ourselves the *immediate* history.

We can extend to use more previous states

$$P(X_t|X_1, X_2, \cdots, X_{t-1}) = P(X_t|X_{t-1}, X_{t-2}, \cdots, X_{t-H})$$

Observations vs Causes

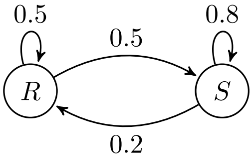
R_{t-1}	$P(R_t \mid R_{t-1})$
t	0.7
f	0.3



R_t	$P(U_t \mid R_t)$
t	0.9
f	0.2

Most likely path in the umbrella example

Today's Weather	$P(\text{Umbrella}\mid\text{today's weather})$
Rainy	0.9
Not Rainy	0.2



Question: This week, my neighbor had umbrella days 1,2,4 and 5. What is the most likely sequence of weather this past week?

Algorithm to solve

Define $\delta_t(j)$: likelihood of most likely path ending at state j at time t.
Let $\psi_t(j)$ be the state I was in at time $t-1$.

Algorithm to solve

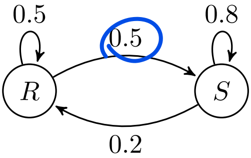
Define $\delta_t(j)$: likelihood of most likely path ending at state j at time t.
Let $\psi_t(j)$ be the state I was in at time $t-1$.

Suppose I tell you $\psi_t(j) = s_{t-1}$ for some s_{t-1}.
What is the likelihood it gets to state s at time t?

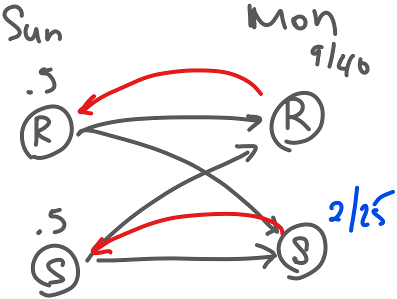
- $\delta_{t-1}(S_{t-1})$ rain/sun the day before
- $P(S_{t-1} \rightarrow S_j)$
- $P(U_t \mid S_j)$

Solving the Umbrella question

Today's Weather	$P(\text{Umbrella}\mid\text{today's weather})$
Rainy	0.9
Not Rainy	0.2



Let's say Sunday has a 50/50 chance of having been rainy or sunny.



$\delta_1(R_1)$:

$\frac{1}{2} \cdot \frac{1}{2} \cdot 9/10 = 9/40$

$\cancel{\frac{1}{2}} \cdot \frac{2}{10} \cdot \frac{9}{10} = 9/100$

$\delta_1(S_1)$: $\frac{1}{2} \cdot \frac{1}{2} \cdot \frac{2}{10} = \frac{1}{20}$

$\frac{1}{2} \cdot \frac{8}{10} \cdot \frac{2}{10} = 2/25$

Erlang/JavaScript example

E	$p(E \mid X = \text{Erlang})$
happy	$4/5$
angry	$1/5$

E	$p(E \mid X = \text{JavaScript})$
happy	$1/4$
angry	$3/4$