CSCI 5582
Artificial Intelligence

Lecture 12
Jim Martin

Today 10/12

• Review
• Basic probability
• Break
• Belief Networks
Review

• Where we are...
  - Agents can use search to find useful actions based on looking into the future
  - Agents can use logic to complement search to represent and reason about
    • Unseen parts of the current environment
    • Past environments
    • Future environments
  - And they can play a mean game of chess

Where we aren’t

• Agents can’t
  - Deal well with uncertain situations (not clear people are all that great at this)
  - Learn
  - See, speak, hear, move, or feel
Exercise

• You go to the doctor and for insurance reasons they perform a test for a horrible disease
• You test positive
• The doctor says the test is 99% accurate
• Do you worry?

An Exercise

• It depends; let’s say...
  - The disease occurs 1 in 10000 folks
  - And that the 99% means that 99 times out a 100 when you give the test to someone without the disease it will return negative
  - And that when you have the disease it always says you are positive
  - Do you worry?
An Exercise

• The test’s false positive rate is 1/100
• Only 1/10000 people have the disease
• If you gave the test to 10000 random people you would have
  - 100 false positives
  - 1 true positive
• Do you worry?

CSCI 5582 Fall 2006

An Exercise

• Do you worry?
  - Yes, I always worry
  - Yes, my chances of having the disease are 100x they were before I went to the doctor
    • Went from 1/10000 to 1/100 (approx)
  - No, I live with a lot of other 1/100 bad things without worrying

CSCI 5582 Fall 2006
Another Example

• You hear on the news...
  - People who attend grad school to get a masters degree have a 10x increased chance of contracting schistosomiasis
• Do you worry?
  - Depends on where you go to grad school

Back to Basics

• Prior (or unconditional) probability
  - Written as $P(A)$
  - For now think of $A$ as a proposition that can turn out to be True or False
  - $P(A)$ is your belief that $A$ is true given that you know nothing else relevant to $A$
Also

- Just as with logic we can create complex sentences with a partially compositional semantics (sort of)...

\[ P(A \land B), P(A \lor B), P(\neg A \lor B) \ldots \]

Basics

- Conditional (or posterior) probabilities
- Written as \( P(A|B) \)
- Pronounced as the probability of \( A \) given \( B \)
- Think of it as your belief in \( A \) given that you know absolutely that \( B \) is true.
And

- \( P(A|B) \)… your belief in \( A \) given that you know \( B \) is true
- **AND** \( B \) is all you know that is relevant to \( A \)

Conditionals Defined

- **Conditionals**
  \[
P(A | B) = \frac{P(A \land B)}{P(B)}
\]

- **Rearranging**
  \[
P(A \land B) = P(A | B)P(B)
\]

- **And also**
  \[
P(A \land B) = P(B | A)P(A)
\]
Conditionals Defined

- Inference means updating your beliefs as evidence comes in
  - $P(A)$... belief in $A$ given that you know nothing else of relevance
  - $P(A|B)$... belief in $A$ once you know $B$ and nothing else relevant
  - $P(A|B^\complement C)$ belief in $A$ once you know $B$ and $C$ and nothing else relevant
Also

- What you’d expect... we can have
  \[ P(A|B^C) \text{ or } P(A^D|E) \text{ or } P(A^B|C^D) \]
  etc...

Joint Semantics

- Joint probability distribution... the equivalent of truth tables in logic
- Given a complete truth table you can answer any question you want
- Given the joint probability distribution over N variables you can answer any question you might want to that involve those variables
Joint Semantics

- With logic you don’t always need the whole truth table; you can use inference methods and compositional semantics
  - I.e. if I know the truth values for A and B, I can retrieve the value of A^B
- With probability, you need the joint to do inference unless you’re willing to make some assumptions

<table>
<thead>
<tr>
<th></th>
<th>Toothache=True</th>
<th>Toothache=False</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cavity = True</td>
<td>0.04</td>
<td>0.06</td>
</tr>
<tr>
<td>Cavity = False</td>
<td>0.01</td>
<td>0.89</td>
</tr>
</tbody>
</table>

- What’s the probability of having a cavity and a toothache?
- What’s the probability of having a toothache?
- What’s the probability of not having a cavity?
- What’s the probability of having a toothache or a cavity?
Note

• Adding up across a row is really a form of reasoning by cases…
• Consider calculating $P(\text{Cavity})$…
  - We know that in this world you either have a toothache or you don’t. I.e toothaches partition the world.
  - So…

Partitioning

$$P(\text{Cavity}) = P(\text{Cavity} \land \text{Toothache}) + P(\text{Cavity} \land \neg \text{Toothache})$$
Combining Evidence

• Suppose you know the values for
  - $P(A|B)=0.2$
  - $P(A|C)=0.05$
  - Then you learn $B$ is true
    • What's your belief in $A$?
  - Then you learn $C$ is true
    • What's your belief in $A$?
Details...

• Where do all the numbers come from?
  - Mostly counting
  - Sometimes theory
  - Sometimes guessing
  - Sometimes all of the above

Numbers

• \( P(A) \) \[ \frac{\text{Count(All As)}}{\text{Count(All Events)}} \]

• \( P(A^\cap B) \) \[ \frac{\text{Count(All A and B together)}}{\text{Count(All Events)}} \]

• \( P(A|B) \) \[ \frac{\text{Count(All A and B Together)}}{\text{Count(All Bs)}} \]
Break

• HW Questions?

Bayes

• We know...
  \[ P(A \cap B) = P(A | B)P(B) \]
  and
  \[ P(A \cap B) = P(B | A)P(A) \]

• So rearranging things
  \[ P(A | B)P(B) = P(B | A)P(A) \]
  \[ P(A | B) = \frac{P(B | A)P(A)}{P(B)} \]
Bayes

• Memorize this

\[ P(A | B) = \frac{P(B | A)P(A)}{P(B)} \]

Bayesian Diagnosis

• Given a set of symptoms choose the best disease (the disease most likely to give rise to those symptoms)
  - I.e. Choose the disease the gives the highest \( P(\text{Disease} | \text{Symptoms}) \) for all possible diseases
• But you probably can’t assess that…
• So maximize this...

\[ P(\text{Disease} | \text{Symptoms}) = \frac{P(\text{Symptoms} | \text{Disease})P(\text{Disease})}{P(\text{Symptoms})} \]
Meningitis

\[ P(S \mid M) = 0.5 \]
\[ P(M) = 0.00002 \]
\[ P(S) = 0.05 \]

so...

\[
P(M \mid S) = \frac{P(S \mid M) P(M)}{P(S)}
\]
\[
= \frac{0.5 \times 0.00002}{0.05}
\]
\[
= 0.0002
\]

Differential Diagnosis

• Why on earth would anyone know \( P(S) \)?
• And do you need to know it?
  – Asking for the most probable disease given some symptoms doesn’t entail knowing the probability of the diseases.
  
  \[
  \text{Argmax}_D P(D \mid S) = P(S \mid D) P(D)/P(S) \]
  
  is the same as
  
  \[
  \text{Argmax}_D P(D \mid S) = P(S \mid D) P(D)
  \]
Well

• What if you needed the exact probability

\[ P(S) = P(S \land M) + P(S \land \neg M) \]

\[ = P(S \mid M)P(M) + P(S \mid \neg M)P(\neg M) \]

Next Time

• Graphical models or Belief Nets
  - Chapter 14

• Quiz is postponed 1 Week
  - Now on 10/26
  - Covers 7, 8, 9, 13 and 14