Today 10/10

- Finish FOL
  - FW and BW chaining
- Limitations of truth conditional logic
- Break
- Basic probability
Inference

• Inference in FOL involves showing that some sentence is true, given a current knowledge-base, by exploiting the semantics of FOL to create a new knowledge-base that contains the sentence in which we are interested.

Inference Methods

• Proof as Generic Search
• Proof by Modus Ponens
  – Forward Chaining
  – Backward Chaining
• Resolution
• Model Checking

CSCI 5582 Fall 2006
Generic Search

- States are snapshots of the KB
- Operators are the rules of inference
- Goal test is finding the sentence you’re seeking
  - I.e. Goal states are KBs that contain the sentence (or sentences) you’re seeking

Example

- Harry is a hare \( Hare(Harry) \)
- Tom is a tortoise \( Tortoise(Tom) \)
- Hares outrun tortoises
  \[
  \forall x, y Hare(x) \land Tortoise(y) \rightarrow Outruns(x, y)
  \]
- Harry outruns Tom?
Tom and Harry

- And introduction
  \[ Harry(Hare) \land Tortoise(Tom) \]
- Universal elimination
  \[ Hare(Harry) \land Tortoise(Tom) \rightarrow Outruns(Harry, Tom) \]
- Modus ponens
  \[ Outruns(Harry, Tom) \]

What’s wrong?

- The branching factor caused by the number of operators is huge
- It’s a blind (undirected) search
So...

- So a reasonable method needs to control the branching factor and find a way to guide the search...
- Focus on the first one first

Forward Chaining

- When a new fact \( p \) is added to the KB
  - For each rule such that \( p \) unifies with part of the premise
    - If all the other premises are known
    - Then add consequent to the KB

This is a data-driven method.
Backward Chaining

• When a query \( q \) is asked
  - If a matching \( q' \) is found return substitution list
  - Else For each rule whose consequent matches \( q \), attempt to prove each antecedent by backward chaining

This is a goal-directed method. And it's the basis for Prolog.

CSCI 5582 Fall 2006

---

Backward Chaining

1. \( \text{Tortoise}(x) \land \text{Slug}(y) \rightarrow \text{Faster}(x, y) \)
2. \( \text{Slimy}(z) \land \text{Creeps}(z) \rightarrow \text{Slug}(z) \)
3. \( \text{Tortoise}(\text{Tom}) \)
4. \( \text{Slimy}(\text{Steve}) \)
5. \( \text{Creeps}(\text{Steve}) \)

Is Tom faster than someone?

CSCI 5582 Fall 2006
Notes

• Backward chaining is not abduction; we are not inferring antecedents from consequents.
• The fact that you can’t prove something by these methods doesn’t mean it’s false. It just means you can’t prove it.

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

Problems with Logic

• Monotonicity
• Modularity
• Abduction
Monotonicity

• Some of the problems we noted stemmed from the notion of monotonicity.
  – Once something is true it has to stay true

Monotonicity

• Within a truth-conditional logic there are three ways to deal with this.
  – Make sure you never assert anything that will need to change its truth value
  – Allow things to change but provide a way to roll back the state of the knowledge-base to what it was before
    • This is known as truth-maintenance
  – Allow complex state representations (agent in location x at time y)
Modularity

• Two kinds
  - Locality
  - Detachment
• These make logic work; they’re not really consistent with uncertain reasoning

• Detachment means that you don’t need to care about how you came to know that $A$ is true to use modus ponens to conclude $B$.
• Locality means that you don’t care what else is going on in the KB. As long as you know those two facts you can conclude $B$. 
Abduction

- Abduction means concluding things about antecedents given knowledge of consequents.

\[
B \\
A \rightarrow B
\]
**Illustrative Example**

- You know
  - Meningitis $\rightarrow$ Stiff necks
  - Stiff neck $\rightarrow$ Car accident

- Patient says they’ve been in a car accident
  - What does a backward chainer say?

- Diagnostic test says a patient has meningitis
  - What does a forward chainer say?

**Example**

- Well you can restrict the kb
  - All causal or all diagnostic rules
    - Meningitis $\rightarrow$ Stiff Neck
    - Car accident $\rightarrow$ Stiff Neck
    - Or
    - Stiff Neck $\rightarrow$ Meningitis
    - Stiff Neck $\rightarrow$ Car accident
Example

• But that precludes a useful form of bi-directional reasoning (explaining away)

Bidirectional Inference

• I tell you I sort of have a stiff neck
  - What happens to your belief in...
    • The idea I was in a car accident?
    • The idea I have meningitis?

• Now I tell you I was in a car accident
  - What happens to your belief in...
    • The idea that I really do have a stiff neck?
    • The idea I have meningitis?
So

- Formally, what you just did was
  - You know
    - A→B
    - A→C
  - I told you C
  - Your belief in A went up
  - Your belief in B went down

Basic Probability

- Syntax and Semantics
  - Syntax is easy
  - Semantics can be messy
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?

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
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

Break

• HW Questions?
• HW Questions?
  - How to represent facts you know to be true (so we guarantee they have the right value in satisfying models)?
  - WalkSat as implemented will flip the values of these known facts.
    • Is that a problem?
    • If so how to fix it.
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)...

$$P(A \land B), P(A \lor B), P(\neg A \lor B)$$
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
  \[ P(A \mid B) = \frac{P(A \land B)}{P(B)} \]

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

- And also
  \[ P(A \land B) = P(B \mid A)P(A) \]
Inference

- 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
  - $P(A|B^C)$ belief in $A$ once you know $B$ and $C$ and nothing else relevant

Also

- What you’d expect... we can have $P(A|B^C)$ or $P(A^D|E)$ 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 need the 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
Joint

<table>
<thead>
<tr>
<th>Cavity</th>
<th>Toothache= True</th>
<th>Toothache= False</th>
</tr>
</thead>
<tbody>
<tr>
<td>True</td>
<td>0.04</td>
<td>0.06</td>
</tr>
<tr>
<td>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 \( A \)?
Combining Evidence

Details...

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

- \( P(A) \) \( \frac{\text{Count(All As)}}{\text{Count(All Events)}} \)
- \( P(A \cap B) \) \( \frac{\text{Count(All A and B together)}}{\text{Count(All Events)}} \)
- \( P(A|B) \) \( \frac{\text{Count(All A and B Together)}}{\text{Count(All Bs)}} \)

Bayes

- We know... \( P(A \land B) = P(A \mid B)P(B) \) and \( P(A \land B) = P(B \mid A)P(A) \)
- So rearranging things \( P(A \mid B)P(B) = P(B \mid A)P(A) \)
\[ P(A \mid B) = \frac{P(B \mid A)P(A)}{P(B)} \]
Bayes

• Memorize this

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

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
  \(P(\text{Disease} \mid \text{Symptoms})\) for all possible diseases
• But you probably can’t assess that...
• So maximize this...

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

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

so....

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

Well

• What if you needed the exact probability

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