Knowledge Representation I (Propositional Logic)

Some KR Languages

- Propositional Logic
- Predicate Calculus
- Frame Systems
- Rules with Certainty Factors
- Bayesian Belief Networks
- Influence Diagrams
- Semantic Networks
- Concept Description Languages
- Nonmonotonic Logic

In Fact...

- All popular knowledge representation systems are equivalent to (or a subset of) Logic
  - Either Propositional Logic
  - Or Predicate Calculus
- Probability Theory

473 Topics

- Inference
- Logic
- Knowledge Representation
- Supervised Learning
- Reinforcement Learning
- Planning
- Search
- Problem Spaces
- Agency
- Perception
- NLP
- Robotics
- Multi-agent
- Probability
Knowledge bases

- Knowledge base = set of sentences in a formal language
- Declarative approach to building an agent (or other system): Tell it what it needs to know

Then it can ask itself what to do - answers should follow from the KB
- Agents can be viewed at the knowledge level i.e., what they know, regardless of how implemented
- Or at the implementation level i.e., data structures in KB and algorithms that manipulate them

A simple knowledge-based agent

- The agent must be able to:
  - Represent states, actions, etc.
  - Incorporate new percepts
  - Update internal representations of the world
  - Deduce hidden properties of the world
  - Deduce appropriate actions

Wumpus World PEAS description

- Performance measure:
  - gold +1000, death -1000
  - -1 per step, -10 for using the arrow

- Environment:
  - Squares adjacent to wumpus are smelly
  - Squares adjacent to pit are breezy
  - Glitter iff gold is in the same square
  - Shooting kills wumpus if you are facing it
  - Shooting uses up the only arrow
  - Grabbing picks up gold if in same square
  - Releasing drops the gold in same square

Wumpus world characterization

- Fully Observable?
- Deterministic?
- Episodic?
- Static?
- Discrete?
- Single-agent?
Wumpus world characterization

- **Fully Observable** No - only local perception
- **Deterministic** Yes - outcomes exactly specified
- **Episodic** No - sequential at the level of actions
- **Static** Yes - Wumpus and Pits do not move
- **Discrete** Yes
- **Single-agent** Yes - Wumpus is essentially a natural feature
Exploring a wumpus world

Logic in general

• Logics are formal languages for representing information such that conclusions can be drawn
  • Syntax defines the sentences in the language
  • Semantics define the "meaning" of sentences:
    • i.e., define truth of a sentence in a world

Entailment

• Entailment means that one thing follows from another:
  \[ \text{KB} \models \alpha \]

• Knowledge base \( \text{KB} \) entails sentence \( \alpha \) if and only if \( \alpha \) is true in all worlds where \( \text{KB} \) is true
  
  E.g., the KB containing "the Giants won" and "the Reds won" entails "Either the Giants won or the Reds won"
  
  E.g., \( x+y = 4 \) entails \( 4 = x+y \)

  Entailment is a relationship between sentences (i.e., syntax) that is based on semantics

Models

• Logicians typically think in terms of models, which are formally structured worlds with respect to which truth can be evaluated
  
  We say \( m \) is a model of a sentence \( \alpha \) if \( \alpha \) is true in \( m \)
  
  \( M(\alpha) \) is the set of all models of \( \alpha \)
  
  Then \( \text{KB} \models \alpha \) iff \( M(\text{KB}) \subseteq M(\alpha) \)
  
  E.g. \( \text{KB} = \text{Giants won and Reds won} \) entails \( \alpha = \text{Giants won} \)
Situation after detecting nothing in [1,1], moving right, breeze in [2,1]

Consider possible models for \( KB \) (only pits)

3 Boolean choices \( \Rightarrow \) 8 possible models

\( KB = \) wumpus-world rules + observations

\( KB \models \alpha \), proved by model checking
Wumpus models

- $KB$ = wumpus-world rules + observations

$\alpha_2 = \{2,2\}$ is safe, $KB \models \alpha_2$

Missing Elements

- How does an agent reason about the wumpus world?
- How do we map truth/information between the real (wumpus) world and our representation?

Inference
- How do we map truth/information between the real (wumpus) world and our representation?

Semantics
Inference

- \( KB \models a \) = sentence \( a \) can be derived from \( KB \) by procedure \( i \)
- **Soundness**: \( i \) is sound if whenever \( KB \models a \), it is also true that \( KB \models a \)
- **Completeness**: \( i \) is complete if whenever \( KB \models a \), it is also true that \( KB \models a \)
- **Preview**: we will define a logic (first-order logic) which is expressive enough to say almost anything of interest, and for which there exists a sound and complete inference procedure.
  - That is, the procedure will answer any question whose answer follows from what is known by the \( KB \).

Semantics

- **Syntax**: a description of the legal arrangements of symbols (Def "sentences")
- **Semantics**: what the arrangement of symbols means in the world

Propositional Logic

- **Syntax**: Atomic sentences: True, False, \( P, Q, \ldots \)
  - Connectives: \( \land, \lor, \neg, \Rightarrow \)
- **Semantics**: Truth Tables
- **Inference**:
  - Modus Ponens
  - Resolution
  - DPLL
  - GSAT
Propositional Logic: **SEMANTICS**

- "Interpretation" (or "possible world")
  - Assignment to each variable either T or F
  - Assignment of T or F to each connective via defns

<table>
<thead>
<tr>
<th></th>
<th>T</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>P</td>
<td>T</td>
<td>F</td>
</tr>
<tr>
<td>Q</td>
<td>T</td>
<td>F</td>
</tr>
</tbody>
</table>

P ∨ Q

P ∧ Q

¬P


Wumpus world sentences

- Let $P_{i,j}$ be true if there is a pit in [i, j].
- Let $B_{i,j}$ be true if there is a breeze in [i, j].

$\neg P_{1,1}$

$\neg B_{1,1}$

$B_{2,1}$

- "Pits cause breezes in adjacent squares"

$B_{1,1} \Leftrightarrow (P_{1,2} \lor P_{2,1})$

$B_{2,1} \Leftrightarrow (P_{1,1} \lor P_{2,2} \lor P_{3,1})$

Truth tables for inference

<table>
<thead>
<tr>
<th>$B_{1,1}$</th>
<th>$B_{2,1}$</th>
<th>$P_{1,1}$</th>
<th>$P_{1,2}$</th>
<th>$P_{2,1}$</th>
<th>$P_{2,2}$</th>
<th>$P_{3,1}$</th>
<th>$KB$</th>
<th>$\alpha_1$</th>
</tr>
</thead>
<tbody>
<tr>
<td>false</td>
<td>false</td>
<td>false</td>
<td>false</td>
<td>false</td>
<td>false</td>
<td>false</td>
<td>false</td>
<td>true</td>
</tr>
<tr>
<td>false</td>
<td>false</td>
<td>false</td>
<td>false</td>
<td>false</td>
<td>false</td>
<td>false</td>
<td>false</td>
<td>true</td>
</tr>
<tr>
<td>false</td>
<td>true</td>
<td>false</td>
<td>false</td>
<td>false</td>
<td>false</td>
<td>false</td>
<td>false</td>
<td>true</td>
</tr>
<tr>
<td>false</td>
<td>true</td>
<td>false</td>
<td>false</td>
<td>false</td>
<td>false</td>
<td>false</td>
<td>false</td>
<td>true</td>
</tr>
<tr>
<td>false</td>
<td>true</td>
<td>false</td>
<td>false</td>
<td>false</td>
<td>false</td>
<td>false</td>
<td>false</td>
<td>true</td>
</tr>
<tr>
<td>false</td>
<td>true</td>
<td>false</td>
<td>false</td>
<td>false</td>
<td>false</td>
<td>false</td>
<td>false</td>
<td>true</td>
</tr>
<tr>
<td>false</td>
<td>true</td>
<td>false</td>
<td>false</td>
<td>false</td>
<td>false</td>
<td>false</td>
<td>false</td>
<td>true</td>
</tr>
<tr>
<td>false</td>
<td>true</td>
<td>false</td>
<td>false</td>
<td>false</td>
<td>false</td>
<td>false</td>
<td>false</td>
<td>true</td>
</tr>
<tr>
<td>true</td>
<td>true</td>
<td>true</td>
<td>true</td>
<td>true</td>
<td>true</td>
<td>true</td>
<td>true</td>
<td>true</td>
</tr>
</tbody>
</table>

$\alpha_1 = "[1,2] is safe"$

Validity and satisfiability

A sentence is valid if it is true in all models:

- e.g., True, $A \lor \neg A$, $A \Rightarrow A$, $(A \land (A \Rightarrow B)) \Rightarrow B$

Validity is connected to inference via the Deduction Theorem:

$KB \vdash \alpha$ if and only if $(KB \Rightarrow \alpha)$ is valid

A sentence is satisfiable if it is true in some model:

- e.g., $A \lor B$, $A \land B$

A sentence is unsatisfiable if it is true in no models:

- e.g., $A \lor \neg A$

Satisfiability is connected to inference via the following:

$KB \vdash \alpha$ if and only if $(KB \land \neg \alpha)$ is unsatisfiable