Logical Agents

• Knowledge-based Agents
• Wumpus World Redux
• Logic
• Propositional Logic
• Inference

Knowledge Bases

• The next step!
• Knowledge base (KB) – set of statements in some kind of formal language
• Essentially declarative (just tell it what it wants to know!)
• Then query the KB for information/moves.
• KB doesn’t have to be prepped ahead of time!

Simple KB Agent

function KB-Agent(percept) returns an action
    set KB, a knowledge base
    a, a counter, initially 0, indicating time
    TELL(KB, MAKE-PERCEPT-Sentence(percept, a))
    action = ASK(KB, MAKE-ACTION-Sentence(a))
    a = a + 1
    return action

• Agent must be able to:
  o Represent states, actions, etc.
  o Incorporate new percepts
  o Update internal representations of the world
  o Deduce hidden properties of the world
  o Deduce appropriate action

Good Problem? Wumpus World!

• Recall – classic Wumpus World
  o 1 Wumpus
  o 1 arrow
  o 4x4 grid of rooms
  o 4 pits
  o Get the agent to the gold and back out
• Perfect KB problem! Potential solution – keep track of all percepts and infer what rooms have what things?

Wumpus World Execution
Some Other Wumpus World Notes

- Could be stuck without a safe move:

- Solution: shoot in a direction to determine safety!
- What about:

How do we represent the world?

- Logic! Great way of representing information to be used in an inference engine.
- Syntax defines the sentences in the language
- Semantics define the meaning, or truth of a sentence.
- Example: arithmetic relations
  - $x + 2 \leq y$
  - $x + 2 < y$
  - $x + 2 \leq y$ is true if $x = 7$ and $y = 1$, but not if $x = 1$ and $y = 7$

Entailment and Models

- Entailment – one thing follows from another: $KB \models \alpha$
- $KB$ entails $\alpha$ iff $\alpha$ is true in all worlds that the $KB$ is true.
- Consider a baseball game: Atlanta Braves vs. Detroit Tigers – the $KB$ here entails “either the Braves won or the Tigers won.” Knowing the result from Friday, we know that the Tigers won, so the $KB$ further entails “the Tigers won.”

Entailment and Models

- Models model a world, or multiple worlds.
- $m$ is a model of a sentence $\alpha$ if $\alpha$ is true in $m$. Note: $m$ is not a KB, it is just a representation of worlds in which $\alpha$ is true!
- $M(\alpha)$ is the set of all models of $\alpha$
- So – $KB \models \alpha$ iff $M(KB) \subseteq M(\alpha)$

Wumpus World Entailment

- Consider:

Consider possible model for ? spaces – consider only pits. 3 Boolean choices, 8 possible models.
Exercise!

7.1 – Wumpus World modelling + entailment

Inference

- KB |- \alpha, sentence \alpha can be inferred from the KB via procedure i.
- Inference is how you find your entailment statements!
- Soundness – i is sound if: whenever KB |- \alpha, it is also true that KB |= \alpha
- Completeness – is is complete if: whenever KB |= \alpha, it is also true that KB |- \alpha
- What we will do: define a logic (first-order) that can be used to express anything of interest, and define a procedure for inference from that logic.

Propositional Logic: Syntax

- P_1 and P_2 and so on are sentences.
- If S is a sentence, so is \neg S
- If S_1 and S_2 are sentences, so is S_1 \land S_2
- If S_1 and S_2 are sentences, so is S_1 \lor S_2
- If S_1 and S_2 are sentences, so is S_1 \Rightarrow S_2
- If S_1 and S_2 are sentences, so is S_1 \Leftrightarrow S_2

Propositional Logic: Semantics

- Each model specifies true/false for each propositional sentence.
- Basic rules for evaluating truth with respect to a model m:
  - \neg S is true iff S is false
  - S_1 \land S_2 is true iff S_1 is true and S_2 is true
  - S_1 \lor S_2 is true iff S_1 is true or S_2 is true
  - S_1 \Rightarrow S_2 is true iff S_1 is false or S_2 is true
  - i.e., is false iff S_1 is true and S_2 is false
  - S_1 \Leftrightarrow S_2 is true if S_1 \Rightarrow S_2 and S_2 \Rightarrow S_1
- A simple recursive process evaluates arbitrary sentences.
Let $P_{i,j}$ mean that there is a pit at square $i,j$

Let $B_{i,j}$ mean that there is a breeze at square $i,j$

Our KB:

- $\neg P_{1,1}$
- $B_{2,1}$

“Pits cause breezes in adjacent squares”

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

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

“A square is breezy iff there is an adjacent pit”

Some logic fun: 7.2 followed by 7.4