Knowledge-Based Agents

- A knowledge-based agent has a base of knowledge about the world
- It uses that base to make new inferences about the world
- From those inferences it decides what to do

Knowledge Base

- The knowledge base is a set of sentences in a particular formal language
- On receiving a percept, the knowledge-based agent operates as follows:
  - It TELLS its knowledge base about the percept
  - It ASKS the knowledge base what to do
  - It TELLS the knowledge base about its action
  - And then it performs the action

Knowledge-Based Agents

- A knowledge-based agent must be able to:
  - Represent states, actions, etc
  - Incorporate new percepts into its representation
  - Update its internal representation of the world
  - Deduce hidden properties of the world
  - Deduce appropriate actions
- Many of these abilities depend on being able to infer new sentences from old
**KB Agent Architecture**

- A knowledge-based agent has three implementation levels:
  - Knowledge level
  - Inference level
  - Implementation level

**Wumpus World: PEAS Description**

- Performance measure:
  - gold +1000
  - death -1000
  - -1 per step
  - -10 for shooting an arrow

- Environmental percepts:
  - Square adjacent to wumpus smell
  - Square adjacent to pit is breezy
  - Square glitters if gold is present
  - Shooting kills wumpus if facing it
  - Shooting uses the only arrow
  - Grabbing picks up gold if present
  - Releasing drops gold

- Sensors can detect:
  - Stench
  - Breeze
  - Glitter
  - Bump
  - Scream

- Actuators allow:
  - Left turn
  - Right turn
  - Forward
  - Grab
  - Release
  - Shoot

**Wumpus World Properties**

- Fully observable?
  - No: agent can only perceive current square

- Deterministic?
  - Yes: all possible outcomes exactly specified

- Episodic?
  - No: sequential actions

- Static?
  - Yes: wumpus and pits do not move

- Discrete?
  - Yes: split into individual squares

- Single agent?
  - Yes: wumpus is not an agent
Logical Agents

- A *logical agent* is a particular type of knowledge-based agent that uses *logic* to deduce new facts about the world and which actions to take.
- If a logical agent draws a conclusion from available information, that conclusion is guaranteed to be correct if the information it was drawn from is correct.

Logic

- A *logic* is a formal language for representing information as sequences of symbols
  - “*Sentences*”
- The *syntax* of a logic defines the valid sentences.
- The *semantics* of a logic define their meaning
  - In many cases, whether they are true or not

Example: Arithmetic

- Symbols: + - × = < ≤ ≥ [0-9] [a-z]
- Syntax:
  - Valid sentences:
    - 1 + 1 = 2
    - 2 × 2 = 5
    - x + 2 ≥ y
  - Invalid sentences:
    - 1 × 1 =
    - x + 2 ≥ 3
- Semantics:
  - True sentences:
    - 1 + 1 = 2
    - x + 2 ≥ y, if the number x+2 is no less than the number y
  - False sentences:
    - 2 × 2 = 5

Worlds and Models

- Semantics are defined with respect to a particular state of the world
  - E.g. x + 2 ≥ y is true in a world where x = 3 and y = 2
  - x + 2 ≥ y is false in a world where x = 2 and y = 5
- A *model* is a formal representation of the state of the world
  - X=3, y=2 is a model; x=2, y=5 is a different model
- “M is a model of α” means sentence α is true in model M
  - M(α) is the set of all possible models of α
  - E.g. all values of x and y that make a = “x+2 ≥ y” true
Entailment

- If a sentence $\beta$ logically follows from a sentence $\alpha$, we say that $\alpha$ entails $\beta$
  
  Notation: $\alpha \vdash \beta$

- Formal definition:
  
  $\alpha$ entails $\beta$ if and only if $\beta$ is true in all models in which $\alpha$ is true
  
  $\alpha \vdash \beta \equiv M(\alpha) \subseteq M(\beta)$

- Example:
  
  $x + y = 4$ entails $y = 4 - x$

Entailment in Wumpus World

- Situation: agent detects nothing in [1,1], moves right, detects a breeze in [2,1]
  
  Knowledge base: rules + [1,1] empty; [2,1] breeze

- Consider all possible models of KB
  
  - No stench, so no wumpus nearby
  
    - For simplicity, consider pits only
  
  - Three squares we can infer something about; 8 possible models

Wumpus World Models

Models of Knowledge Base

- Knowledge base = rules + observations
- $M(KB)$ = set of models in which the knowledge base is true
**Entailment**

- $\alpha_1 = \text{“[1,2] is safe”}
- $M(KB) \subseteq M(\alpha_1)$, therefore $KB \models \alpha_1$

**Non-entailment**

- $\alpha_2 = \text{“[2,2] is safe”}
- $M(KB) \not\subseteq M(\alpha_2)$, therefore $KB \not\models \alpha_2$

**Inference**

- Entailment just says that a sentence $\beta$ is consistent with a sentence $\alpha$
  - It says nothing about how to derive $\beta$ from $\alpha$
- **Inference** is the process of deducing a true sentence $\beta$ from a given sentence $\alpha$
  - Notation: $\alpha \vdash \beta$
  - “$\beta$ is derived from $\alpha$ by inference procedure $i$”
  - Usually omit the $i$ for simplicity

**Properties of Inference Procedures**

- **Soundness**: an inference procedure is *sound* if every sentence it can derive is entailed
  - If $\alpha \vdash \beta$ implies $\alpha \models \beta$, then $i$ is sound
- **Completeness**: an inference procedure is *complete* if every entailed sentence can be derived
  - If $\alpha \models \beta$ implies $\alpha \vdash \beta$, then $i$ is complete
• The knowledge base is syntactic
• Sensors create the connection between it and the real world—the semantics