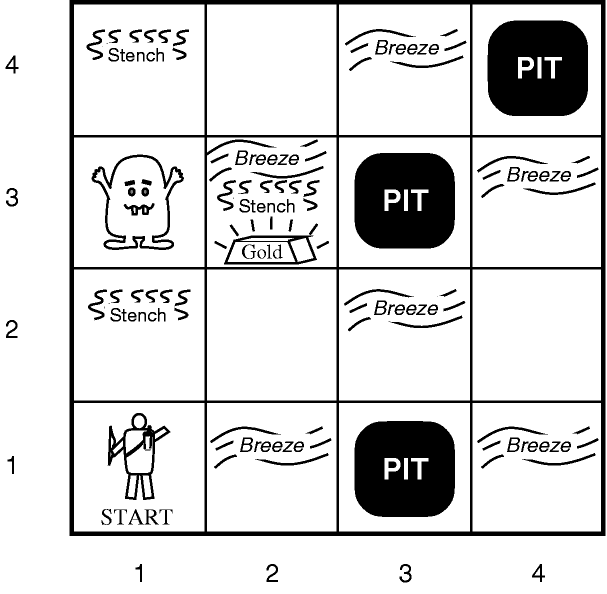
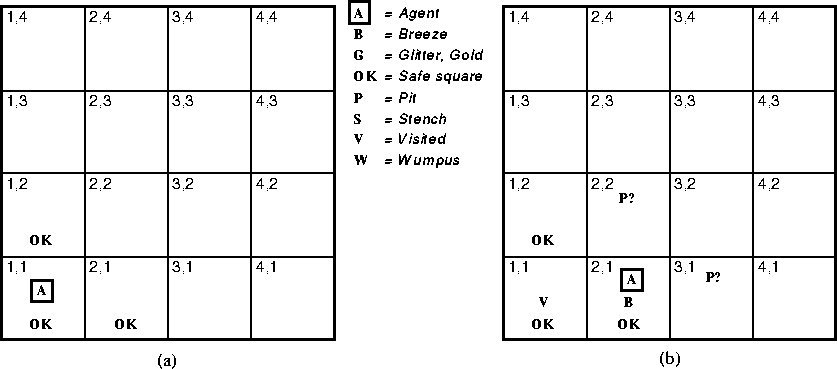
**Week 7 Topic Outline – Logical Agents**

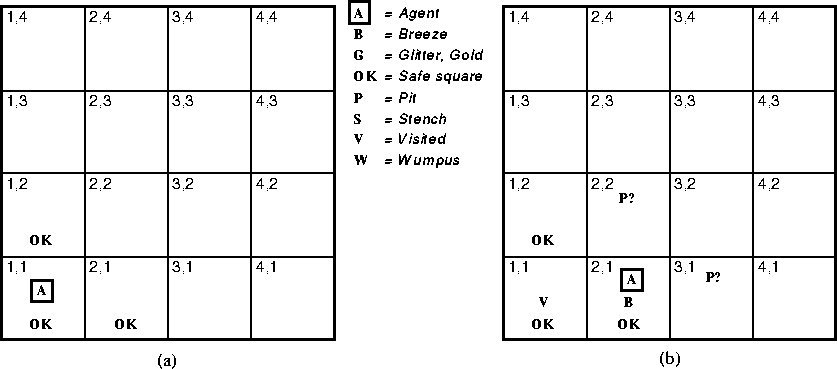
1. **Knowledge-Based Agents**
   1. Agents need knowledge about the world in order to reach good decisions.
   2. “Knowledge” is contained in the form of sentences that are stored in a **knowledge base (KB)**
   3. A Knowledge-Based Agent is a combination of the knowledge base and an inference mechanism.
      1. Stores sentences about the world in its knowledge base
      2. Uses the inference mechanism to infer new sentences
      3. Uses these sentences to decide what action to take
   4. These sentences are sometimes called **axioms**
2. **The Wumpus World**
   1. Wumpus World is a cave consisting of rooms connected by passageways.
   2. Lurking somewhere in the cave is the terrible Wumpus, a beast that eats anyone who enters its room.
   3. The Wumpus can be shot by an agent, but the agent has only one arrow.
   4. Some rooms contain bottomless pits that will trap anyone who wanders into these rooms (except the Wumpus, which is too big to fall in).
   5. There’s a huge pile of gold somewhere in the cave.
   6. **Performance measure:**
      1. +1000 for climbing out of the cave with the gold
      2. -1000 for falling into a pit or being eaten by the Wumpus
      3. -1 for each action taken
      4. -10 for using up the arrow
      5. The game ends when the agent climbs out of the cave or dies
   7. **Environment:**
      1. 4x4 grid of rooms
      2. Agent always starts in the square labeled [1, 1], facing to the right
      3. Wumpus and gold locations are set randomly, with uniform distribution, from the squares other than the start.
      4. Each square other than the start can be a pit, with probability 0.2.
   8. **Actuators:**
      1. Agent can move *Forward, TurnLeft* 90 degrees*, TurnRight* 90 degrees*.*
      2. Agent dies if it moves into a square with a pit or a live Wumpus.
      3. Agent is safe (but smelly) if it enters a square with a dead Wumpus.
      4. Agent remains still if it tries to move forward into a wall.
      5. *Grab* can be used to pick up the gold if it is in the same square as the agent
      6. *Shoot* can be used to fire an arrow in a straight line in the direction the agent is facing. The arrow continues until it either hits (kills) the Wumpus or hits a wall. *Shooting* when the agent has no arrow has no effect.
      7. *Climb* can be used to exit the cave, but only in square [1, 1].
   9. **Sensors:**
      1. In the square containing the Wumpus and in the directly (no diagonally) adjacent squares, the agent will perceive a *Stench*.
      2. In the squares directly adjacent to a pit, the agent will perceive a *Breeze.*
      3. In the square where the gold is, the agent will perceive a *Glitter.*
      4. When an agent walks into a wall, it will perceive a *Bump.*
      5. When the Wumpus is killed, it emits a woeful *Scream* that can be perceived anywhere in the cave.
      6. The percepts will be given to the agent program in the form of a list of five symbols; for examples, if there is a stench and a breeze, but no glitter, bump, or scream, the agent program will get [*Stench, Breeze, None, None, None*].
   10. **Characteristics**
       1. Discrete or Continuous?
          1. Discrete data only take on particular values and no values in between (number of cars or siblings a person has)
          2. Continuous data can take on any value on a range
       2. Static or Dynamic?
       3. Single- or Multi-agent?
       4. Partially or Fully Observable

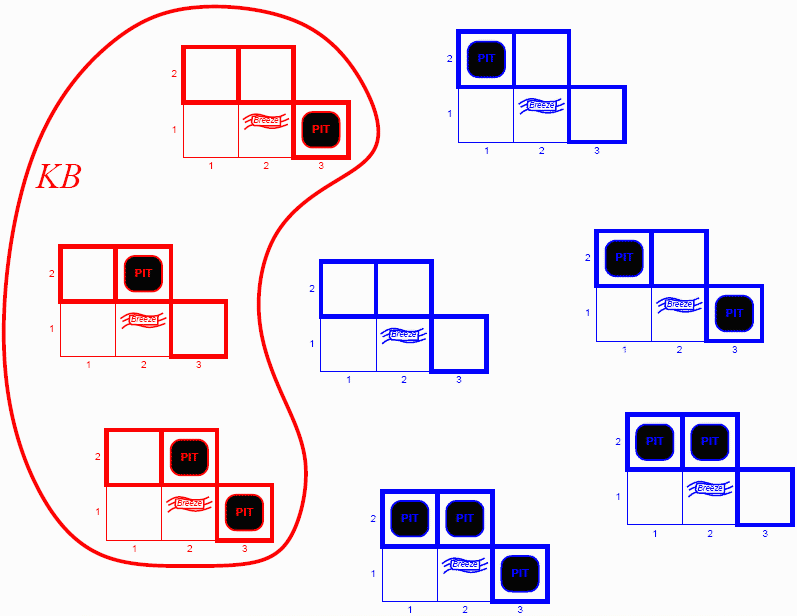


1. **Chances of maximizing our performance measure**
   1. The main challenge is its initial ignorance of the configuration of the environment
   2. Overcoming this ignorance seems to require logical reasoning.
   3. In most instances of the Wumpus world, it is possible for the agent to retrieve the gold safely.
   4. Occasionally, the agent must choose between going home empty-handed and risking death to find the gold.
   5. About 21% of the environments are utterly unfair, because the gold is in a pit or surrounded by pits.
2. **Our Knowledge-Based Agent in the Wumpus World**
   1. The agent’s initial knowledge base contains the rules of the environment (as described above)
   2. It knows that it is in [1, 1] and that [1, 1,] is a safe square (denoted with an “A” and “OK”
   3. The first percept is [*None, None, None, None, None*], from which the agent can conclude that its neighboring squares, [1, 2] and [2, 1] are free of dangers (marked “OK”).
   4. This is represented by figure (a) below.



* 1. Let’s suppose the agent decides to move forward to [2, 1].
  2. The agent perceives a breeze (“B”) in the game, so there must be a pit in [2, 2] or [3, 1], or both (“P?”).
  3. This is represented by figure (b) above.

1. **Play Wumpus World**
2. **Break**
3. **Representing Logic as Sentences**
   1. Every sentence must be either true or false – there is no gray area
   2. “x + y = 6” is a well-formed sentence
      1. True when x = 3 and y = 3
      2. False when x = 1 and y = 2
      3. Etc.
   3. **Entailment** – when one sentence follows logically from another sentence.
      1. A Model is a term that refers to a possible world or environment
      2. The sentence “x = 0” entails the sentence “xy = 0”. Any model (world/environment) where x = 0, any product of x and another value will be 0, regardless of y’s value.
      3. A |= B mean sentence A entails sentence B
      4. Look at figure (b) above. The agent has detected nothing in [1, 1] and a breeze in [2, 1]. These percepts, combined with the agent’s knowledge of the rules of the Wumpus world, constitute the Knowledge Base (KB).
      5. The agent is interested in whether the adjacent squares [1, 2], [2, 2] and [3, 1] contain pits.
      6. Each of the three squares may or may not contain a pit, so there are 23 = 8 possible models.
      7. These models are represented below:



2

1

5

4

3

8

6

7

* + 1. The KB can be thought of as a set of sentences (or one sentence that contains everything the agent knows). This is shown in red above.
    2. The KB is false in models that contradict what the agent knows and true when they do not.
  1. **Which of the models in the above image are false, and which are true?**
  2. Let’s consider two possibilities
     1. A1 = “There is no pit in [1, 2]”
     2. A2 = “There is no pit in [2, 2]”
     3. Compare the sentence A1 with the **knowledge base (red)** and **models (blue)**. What can you conclude?
     4. *In every model in which the knowledge base in true, A1 is true.*
        1. **Which model number satisfies that?**
        2. Hence, KB |= A1 : there is no pit in [1, 2]
     5. What about A2? *In some models in which KB is true, A2 is false* 
        1. **Which model number satisfies that?**
        2. Hence, KB V= A2 : the agent cannot conclude that there is no pit in [2, 2] (nor can it conclude that there is).
  3. The idea of using logical inferences to create additions to the KB is **entailment.**

1. **Propositional Logic**
   1. Syntax and connectors on page 244
   2. The semantics defines the rules for determining the truth of a sentence with respect to a particular model. In propositional logic, a model assigns true or false for every proposition symbol.
   3. Example: If the sentences in the KB make use of the proposition symbols P1,2, P2,2, and P3,1, then one possible model is:
      1. M1 = { P1,2 = false, P2,2 = false, and P3,1 = true}
   4. Again, with three proposition symbols, there are 23 = 8 possible models
2. **Implementation**
   1. How would we go about coding something like this?
   2. Take 5-10 minutes and plan out how this would be done?
   3. Be ready to share your ideas.