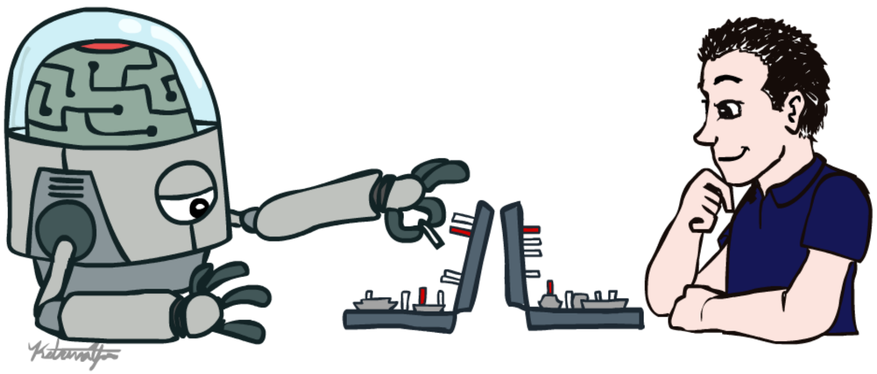
# CSE 3521:
# Introduction to Artificial Intelligence



[Many slides are adapted from the UC Berkeley. CS188 Intro to AI at UC Berkeley and previous CSE 3521 course at OSU.]

THE OHIO STATE UNIVERSITY

# Logical Inference

- **Knowledge-based logical agents**
  - Knowledge base and representation
  - Entailment and inference

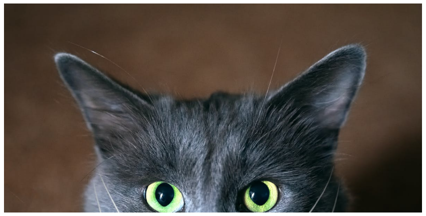
- **Propositional logic**

In which we design agents that can form representations of the world, use a process of inference to derive new representations about the world, and use these new representations to deduce what to do.

In contrast, planning agents find the best action sequence by strategically trial and error in the simulation!

# Knowledge-Based Logical Agents

- Two central concepts
  - Representation of knowledge (state sequences?)
  - Reasoning processes acting on knowledge (e.g., choose actions)
- Play crucial role in “Partially Observable” environments
  - Combine general knowledge with current percepts to infer hidden aspects before acting
- Aids in agent flexibility
  - Learn new knowledge for new tasks
  - Adapt to changes in environment by updating relevant knowledge

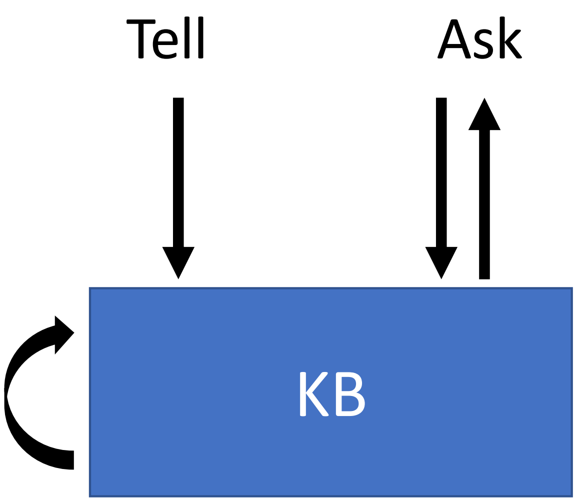


# Knowledge Base

- Central component is its knowledge base (KB)
  - Contains set of "sentences" (in knowledge representation language)
  - KB initially contains some background knowledge

- How to add new information to KB?
  - **TELL** function
  - Inference: deriving new sentences from old ones

- How to query what is known?
  - **ASK** function
  - Answers should follow what has been told to the KB previously

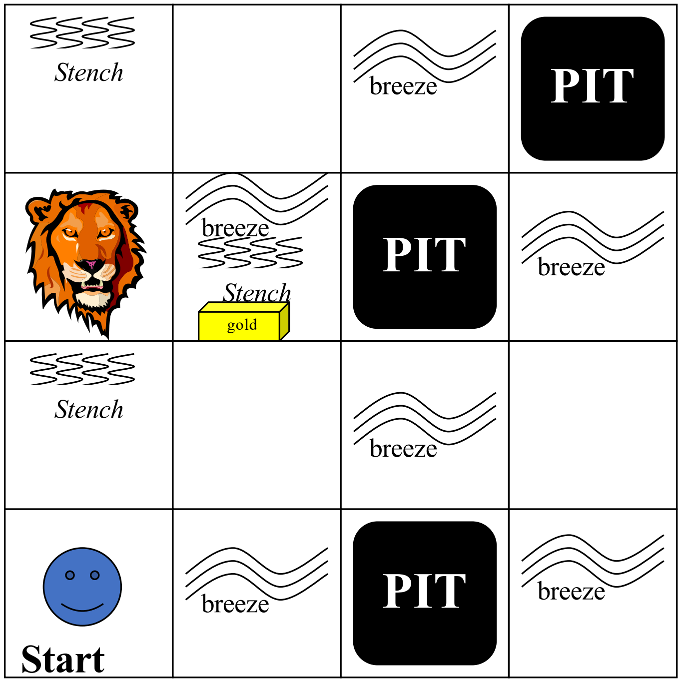


# Logic

- For <u>logical</u> agents, knowledge is <u>definite</u>
  - Each proposition is either “True” or “False”
- Logic has advantage of being simple representation for knowledge-based agents
  - But limited in its ability to handle uncertainty
- We will examine propositional logic and first-order logic

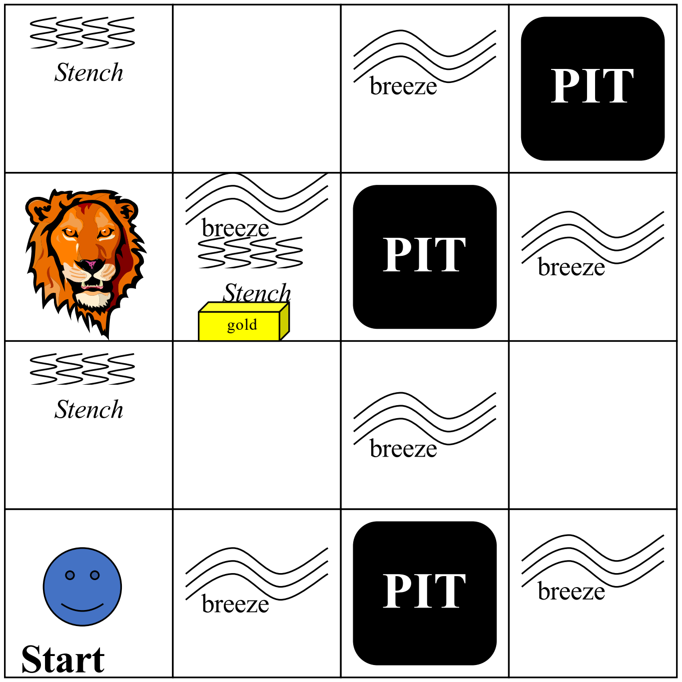
# Wumpus World Environment

- Simple environment to motivate logical reasoning
- Agent explores cave with rooms connected by passageways
- "Wumpus" beast lurking somewhere in cave
  - Eats anyone who enters its room
  - Agent has one arrow (can kill Wumpus)
- Some rooms contain bottomless pits
- Occasional heap of gold present
- Agent task
  - Enter cave, find the gold, return to entrance, and exit



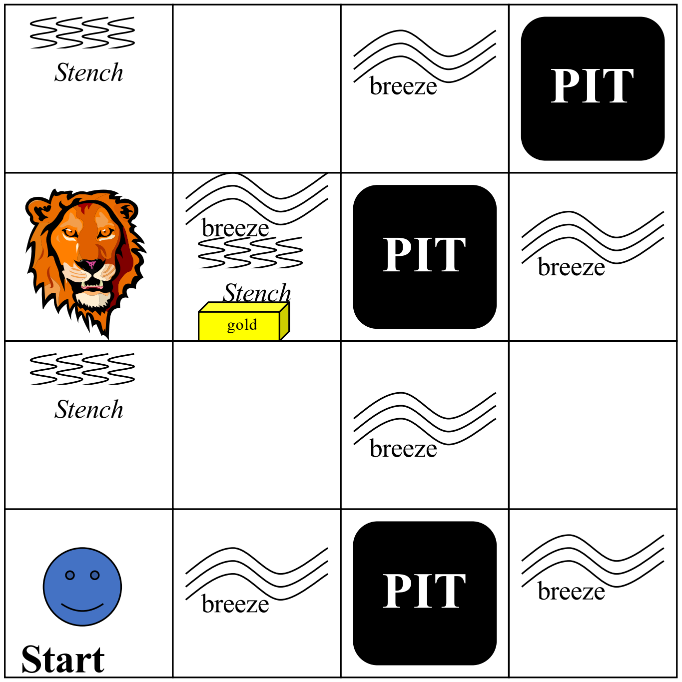
# Wumpus World PEAS Description

- (P)erformance measure
  - Cost of –1000 for falling into pit or being eaten by Wumpus (GAME OVER!)
  - Receive +1000 for picking up gold
  - Cost of –1 for each action taken
  - Cost of –10 for using up the only arrow
- (E)nvironment
  - 4x4 grid of rooms
  - Agent starts in square [1,1]
  - Wumpus and gold locations chosen randomly
  - Probability of square being a pit is .2
    - [0=*no*, …, 0.5=*maybe*, …, 1=*yes*]



# Wumpus World PEAS Description

- (A)ctuators
  - Move forward, turn left, turn right
    - Note: die if enter pit or live wumpus square
  - Grab (gold)
  - Shoot (arrow)
    - Kills wumpus if facing its square
- (S)ensors
  - Nose: squares adjacent to wumpus are “smelly”
  - Skin/hair: Squares adjacent to pit are “breezy”
  - Eye: “Glittery” if and only if gold is in the same square
  - Percepts: [Stench, Breeze, Glitter]



# Exploring a Wumpus World

**A** = agent

**B** = breeze

**G** = glitter, gold

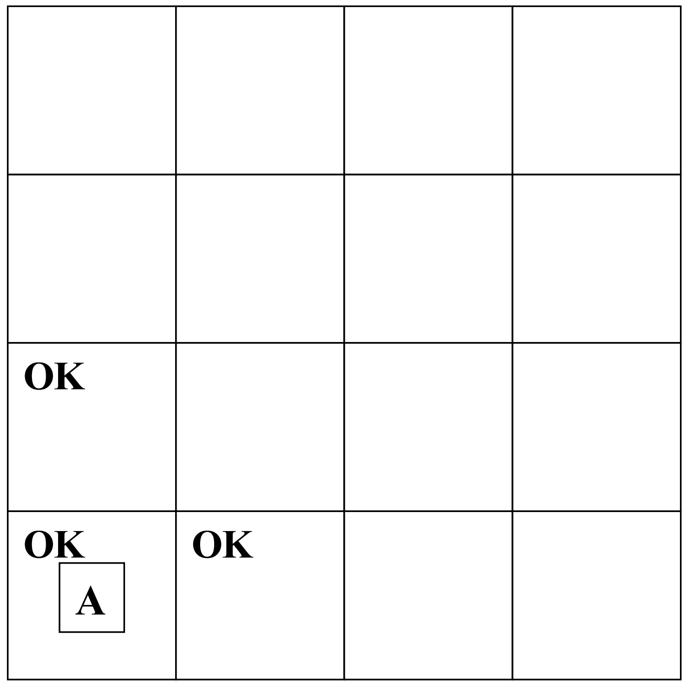
**OK** = safe square

**P** = pit

**S** = stench

**V** = visited

**W** = Wumpus




From local percepts, determines that {(1,1), (1,2), (2,1)} are free from danger.

# Exploring a Wumpus World

**A** = agent

**B** = breeze

**G** = glitter, gold

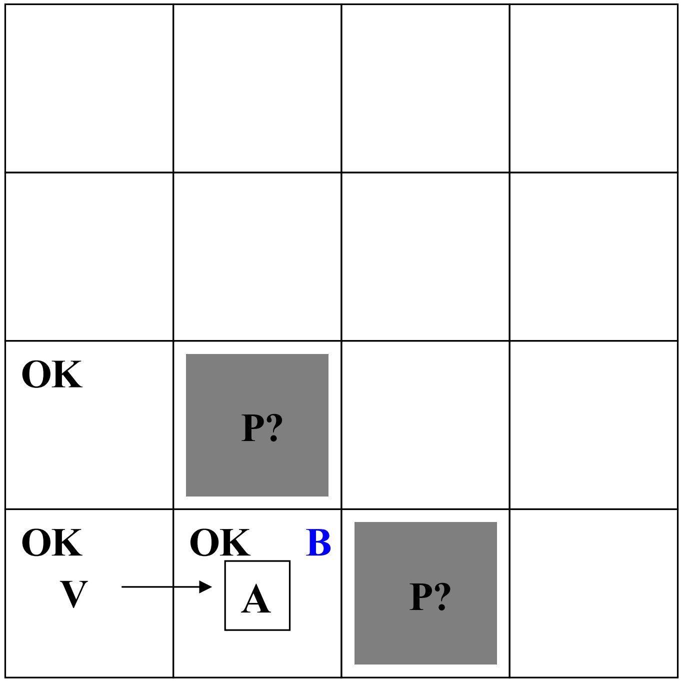
**OK** = safe square

**P** = pit

**S** = stench

**V** = visited

**W** = Wumpus




From breeze percept, determines that (2,2) or (3,1) is a pit. Go back to (1,1) and move up to (1,2).

# Exploring a Wumpus World

**A** = agent

**B** = breeze

**G** = glitter, gold

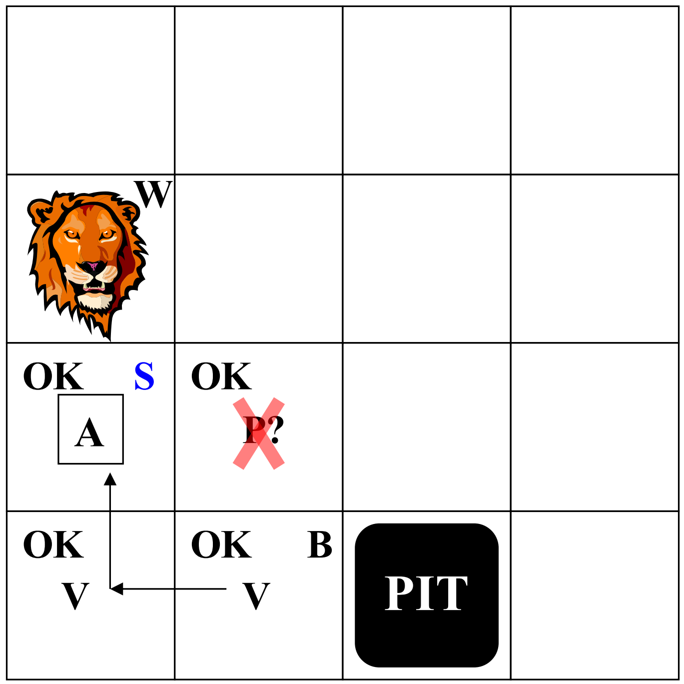
**OK** = safe square

**P** = pit

**S** = stench

**V** = visited

**W** = Wumpus




From stench and no-breeze percept in (1,2), determines that Wumpus in (1,3), pit in (3,1), and (2,2) clear.

# Exploring a Wumpus World

**A** = agent

**B** = breeze

**G** = glitter, gold

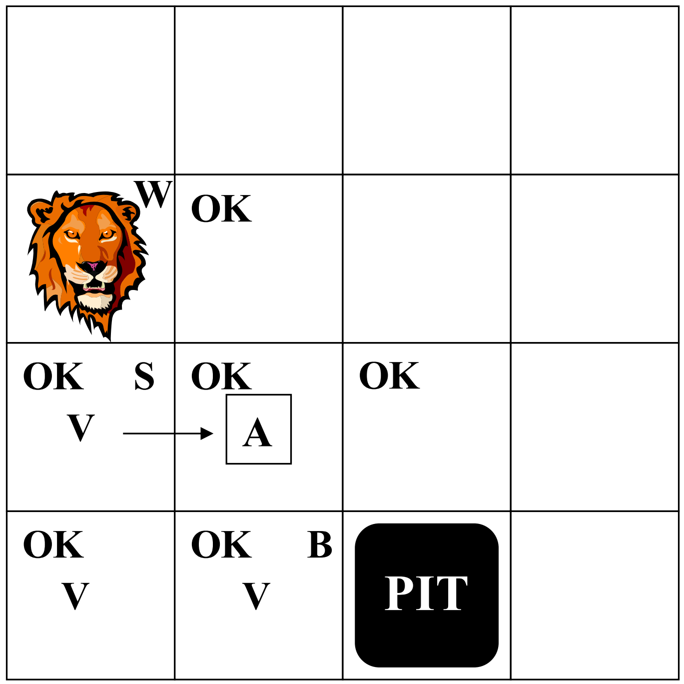
**OK** = safe square

**P** = pit

**S** = stench

**V** = visited

**W** = Wumpus




From local percepts, it is OK to move up or right.

# Exploring a Wumpus World

**A** = agent

**B** = breeze

**G** = glitter, gold

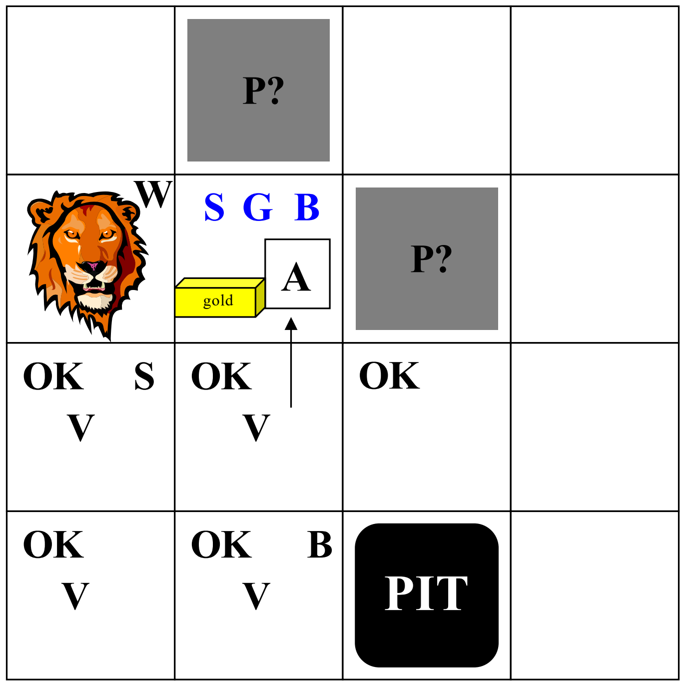
**OK** = safe square

**P** = pit

**S** = stench

**V** = visited

**W** = Wumpus




Found gold! No need to explore further. Time to head back.

# Exploring a Wumpus World

**A** = agent

**B** = breeze

**G** = glitter, gold

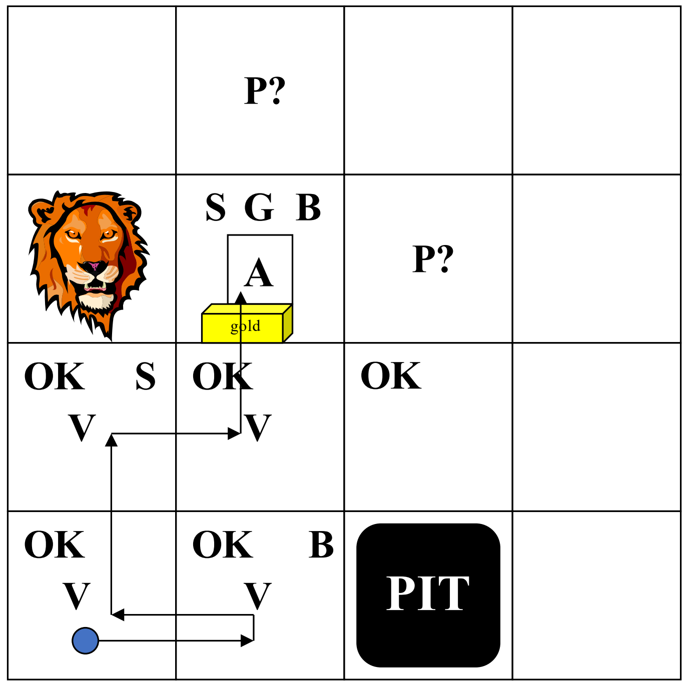
**OK** = safe square

**P** = pit

**S** = stench

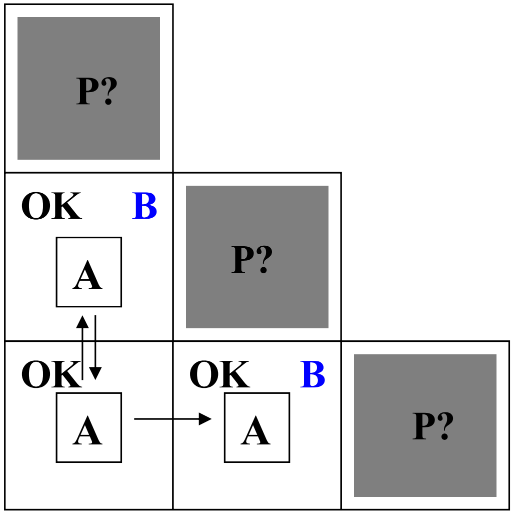
**V** = visited

**W** = Wumpus




Then go home using **OK** squares (retrace route).

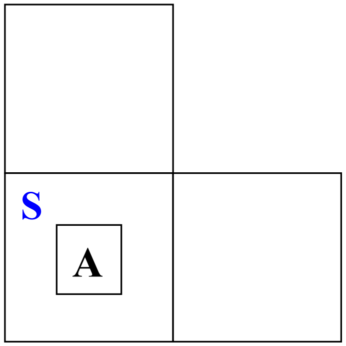
# Tight Spot




<u>Breeze</u> in (1,2) and (2,1) → no safe actions!

Pit may actually only be in (2,2), but can't tell.

# More Tight Spot




<u>Smell</u> in (1,1) $\rightarrow$ Cannot move!

Possible action: shoot arrow straight ahead

# Logical Agent

- Need agent to represent beliefs
  - "There is a pit in (2, 2) or (3, 1)"
  - "There is no Wumpus in (2, 2)"
- Need to make inferences
  - If available information is correct, draw a conclusion that is guaranteed to be correct
- Need <u>representation</u> and <u>reasoning</u>
  - Support the operation of knowledge-based agent

# Knowledge Representation

- For expressing knowledge in computer-tractable form
- Knowledge representation language defined by
  - **Syntax**
    - Defines the possible well-formed configurations of sentences in the language
  - **Semantics**
    - Defines the "meaning" of sentences (need interpreter)
    - Defines the truth of a sentence in a world (or model)

# The Language of Arithmetic

- Syntax: "$x + 2 \geq y$" is a sentence

  "$x2 + y >$" is not a sentence

- Semantics: $x + 2 \geq y$ is true iff the number $x + 2$ is no less than the number $y$

  $x + 2 \geq y$ is True in a world where $x$=7, $y$=1

  $x + 2 \geq y$ is False in a world where $x$=0, $y$=6

# Inference

- Sentence is valid iff it is true under all possible interpretations in all possible worlds
  - Also called tautologies
  - "There is a stench at (1,1) or there is not a stench at (1,1)"
  - "There is an open area in front of me" is not valid in all worlds

- Sentence is satisfiable iff there is some interpretation in some world for which it is true
  - "There is a wumpus at (1,2)" could be true in some situation
  - "There is a wall in front of me and there is no wall in front of me" is unsatisfiable

# Propositional Logic: Syntax

- Syntax of propositional logic defines allowable sentences
- Atomic sentences consists of a single proposition symbol
  - Each symbol stands for proposition that can be True or False
- Symbols of propositional logic
  - Propositional symbols: *P*, *Q*, … (e.g., "Today is Tuesday")
  - Logical constants: *True*, *False*
- Making complex sentences
  - Logical connectives of symbols: $\wedge$, $\vee$, $\Leftrightarrow$, $\Rightarrow$, $\neg$
  - Also have parentheses to enclose each sentence: (…)
- Sentences will be used for inference/problem-solving

# Propositional Logic: Syntax

- *True*, *False*, $S_1$, $S_2$, … are sentences
- If $S$ is a sentence, $\neg S$ is a sentence
  - Not (negation)
- $S_1 \wedge S_2$ is a sentence, also $(S_1 \wedge S_2)$
  - And (conjunction)
- $S_1 \vee S_2$ is a sentence
  - Or (disjunction)
- $S_1 \Rightarrow S_2$ is a sentence (e.g., “Today is Tuesday” implies “Tomorrow is Wednesday”)
  - Implies (conditional)
- $S_1 \Leftrightarrow S_2$ is a sentence
  - Equivalence (biconditional)

# Propositional Logic: Semantics

- Semantics defines the rules for determining the truth of a sentence
  - With respect to a particular model)
    - $\neg S$ is true iff $S$ is false
    - $S_1 \wedge S_2$ is true iff $S_1$ is true <u>and</u> $S_2$ is true
    - $S_1 \vee S_2$ is true iff $S_1$ is true <u>or</u> $S_2$ is true
    - $S_1 \Rightarrow S_2$ is true iff $S_1$ is false <u>or</u> $S_2$ is true
      (is false iff $S_1$ is true <u>and</u> $S_2$ is false)
      (if $S_1$ is true, then claiming that $S_2$ is true, otherwise make no claim)
    - $S_1 \Leftrightarrow S_2$ is true iff $S_1 \Rightarrow S_2$ is true <u>and</u> $S_2 \Rightarrow S_1$ is true ($S_1$ same as $S_2$)

# Semantics in Truth Table Form

| $P$ | $Q$ | $\neg P$ | $P \wedge Q$ | $P \vee Q$ | $P \Rightarrow Q$ | $P \Leftrightarrow Q$ |
|---|---|---|---|---|---|---|
| False | False | True | False | False | True | True |
| False | True | True | False | True | True | False |
| True | False | False | False | True | False | False |
| True | True | False | True | True | True | True |

# Propositional Inference: Enumeration Method

- Truth tables can test for valid sentences
  - True under all possible interpretations in all possible worlds
- For a given sentence, make a truth table
  - Columns as the combinations of propositions in the sentence
  - Rows with all possible truth values for proposition symbols
- If sentence true in every row, then valid

# Propositional Inference: Enumeration Method

- Test $((P \lor H) \land \neg H) \Rightarrow P$

| $P$ | $H$ | $P \lor H$ | $\neg H$ | $(P \lor H) \land \neg H$ | $((P \lor H) \land \neg H) \Rightarrow P$ |
|---|---|---|---|---|---|
| False | False | False | True | False | True |
| False | True | True | False | False | True |
| True | False | True | True | True | True |
| True | True | True | False | False | True |

# Practice

- Test $(P \wedge H) \Rightarrow (P \vee \neg H)$

# Practice

- Test $(P \wedge H) \Rightarrow (P \vee \neg H)$

| $P$ | $H$ | $P \wedge H$ | $\neg H$ | $(P \vee \neg H)$ | $(P \wedge H) \Rightarrow (P \vee \neg H)$ |
|---|---|---|---|---|---|
| False | False | False | True | True | True |
| False | True | False | False | False | True |
| True | False | False | True | True | True |
| True | True | True | False | True | True |

# Simple Wumpus Knowledge Base

- For simplicity, only deal with the pits
- Choose vocabulary
  - 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 [$i$,$j$]
- KB sentences
  - **FACT:** "There is no pit in [1,1]"

    $R_1$: $\neg P_{1,1}$
  - **RULE:** "There is breeze in adjacent neighbor of pit"

    $R_2$: $B_{1,1} \Leftrightarrow (P_{1,2} \vee P_{2,1})$

    $R_3$: $B_{2,1} \Leftrightarrow (P_{1,1} \vee P_{2,2} \vee P_{3,1})$

***Need rule for each square!***

# Wumpus Environment

- Given knowledge base
- Include percepts as move through environment (online)
- Need to "deduce what to do"
- Derive chains of conclusions that lead to the desired goal
  - Use inference rules

$$\frac{\alpha}{\beta}$$

Inference rule: "$\alpha$ derives $\beta$"

***Knowing $\alpha$ is true, then $\beta$ must also be true***

# Inference Rules for Prop. Logic

- <u>Modus Ponens</u>
  - From implication and premise of implication, can infer conclusion

$$\frac{\alpha \Rightarrow \beta, \quad \alpha}{\beta}$$

# Inference Rules for Prop. Logic

- And-Elimination
  - From conjunction, can infer any of the conjuncts

$$\frac{\alpha_1 \land \alpha_2 \land \cdots \land \alpha_n}{\alpha_i}$$

# Inference Rules for Prop. Logic

- <u>And-Introduction</u>
  - From list of sentences, can infer their conjunction

$$\frac{\alpha_1, \alpha_2, \dots, \alpha_n}{\alpha_1 \wedge \alpha_2 \wedge \cdots \wedge \alpha_n}$$

# Inference Rules for Prop. Logic

- Or-Introduction
  - From sentence, can infer its disjunction with anything else

$$\frac{\alpha_i}{\alpha_1 \vee \alpha_2 \vee \cdots \vee \alpha_n}$$

# Inference Rules for Prop. Logic

- Double-Negation Elimination
  - From doubly negated sentence, can infer a positive sentence

$$\frac{\neg\neg\alpha}{\alpha}$$

# Inference Rules for Prop. Logic

- Unit Resolution
  - From disjunction, if one of the disjuncts is false, can infer the other is true

$$\frac{\alpha \vee \beta, \quad \neg \beta}{\alpha}$$

# Inference Rules for Prop. Logic

- <u>Resolution</u>
  - Most difficult because β cannot be both true and false
  - One of the other disjuncts must be true in one of the premises
    - (implication is transitive)

$$\frac{\alpha \vee \beta, \quad \neg\beta \vee \gamma}{\alpha \vee \gamma}$$

| $\neg\beta$ | OR | $\beta$ |
|---|---|---|
| $\dfrac{\alpha \vee \beta, \neg\beta}{\alpha}$ | OR | $\dfrac{\neg\beta \vee \gamma, \beta}{\gamma}$ |
| $\alpha$ | OR | $\gamma$ |

| $\alpha$ | $\beta$ | $\gamma$ | $\alpha \vee \beta$ | $\neg\beta \vee \gamma$ | $\alpha \vee \gamma$ |
|---|---|---|---|---|---|
| T | T | T | T | T | T |
| T | T | F | T | F | T |
| T | F | T | T | T | T |
| T | F | F | T | T | T |
| F | T | T | T | T | T |
| F | T | F | T | F | T |
| F | F | T | F | T | F |
| F | F | F | F | T | F |

# TASK:  Find the Wumpus

Can we infer that the Wumpus is in cell (1,3), given our percepts and environment rules?

**A** = agent

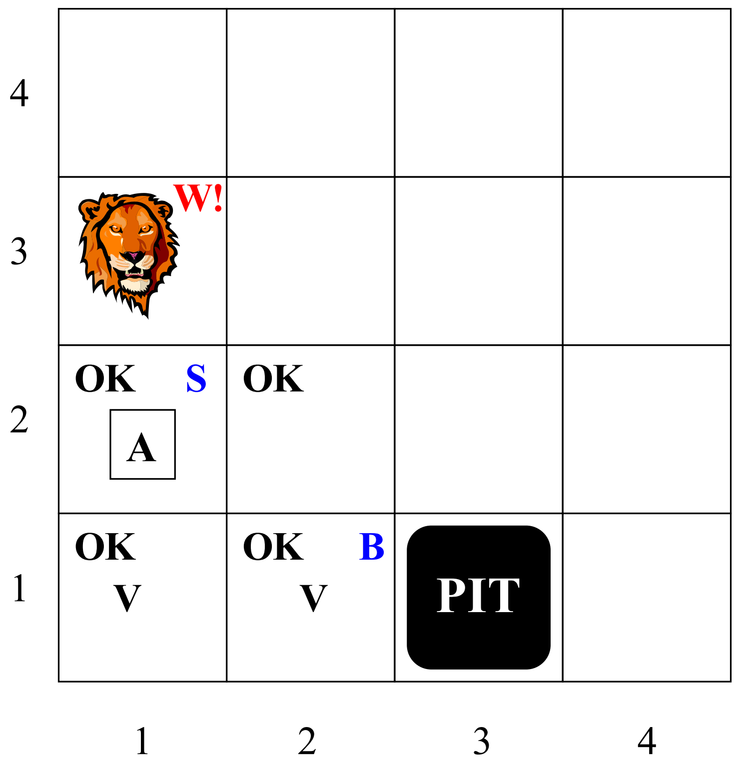
**B** = breeze

**OK** = safe square

**S** = stench

**V** = visited

**W** = wumpus



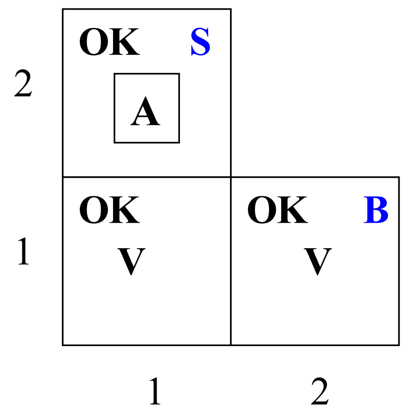
# Wumpus Knowledge Base

- Percept sentences (facts) “at this point”

$\neg S_{1,1}$ $\quad \neg B_{1,1}$

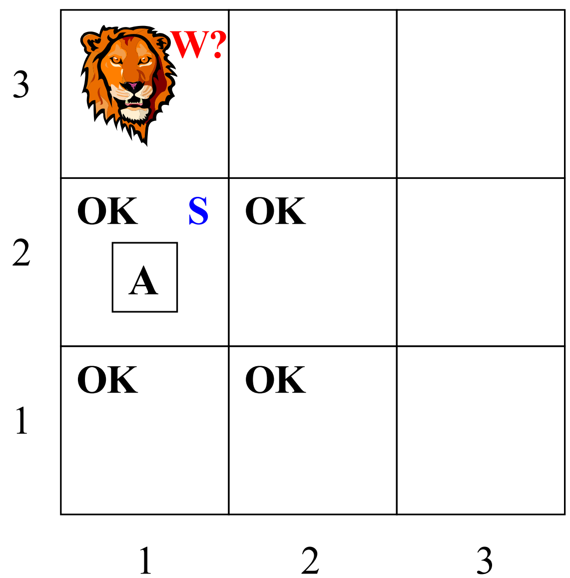
$\neg S_{2,1}$ $\quad B_{2,1}$

$S_{1,2}$ $\quad \neg B_{1,2}$



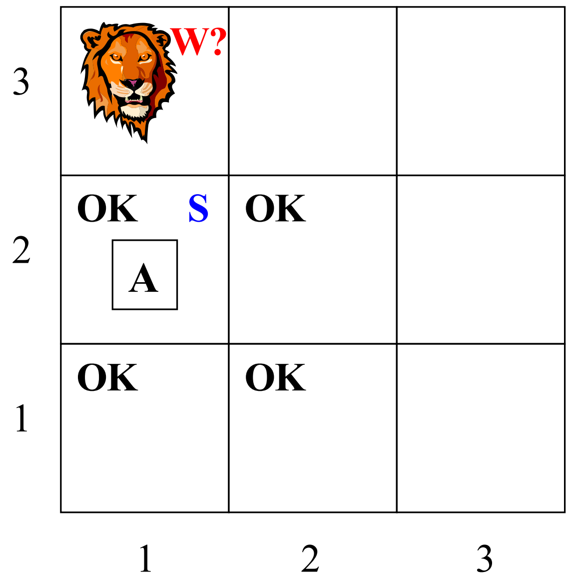
# Environment Rules

$$R_1: \neg S_{1,1} \Rightarrow \neg W_{1,1} \wedge \neg W_{1,2} \wedge \neg W_{2,1}$$



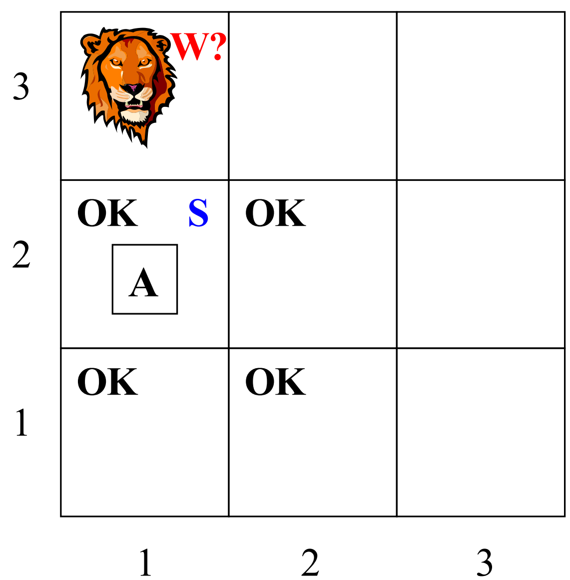
# Environment Rules

$$R_2: \neg S_{2,1} \Rightarrow \neg W_{1,1} \wedge \neg W_{2,1} \wedge \neg W_{2,2} \wedge \neg W_{3,1}$$



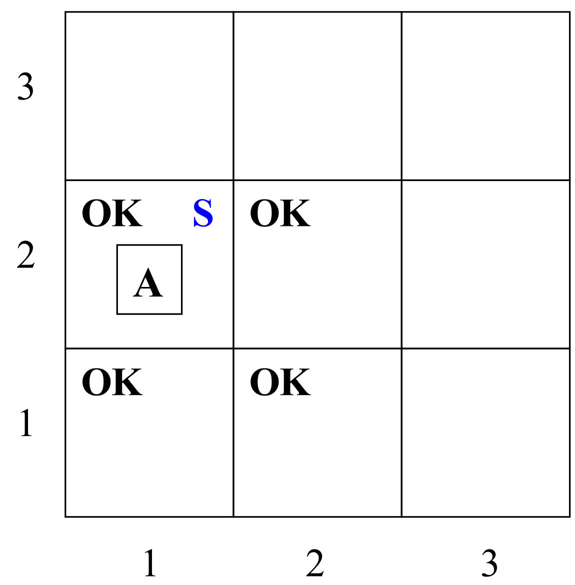
# Environment Rules

$$R_3: \quad \neg S_{1,2} \Rightarrow \neg W_{1,1} \wedge \neg W_{1,2} \wedge \neg W_{2,2} \wedge \neg W_{1,3}$$



# Environment Rules

$$R_4:\ S_{1,2} \Rightarrow W_{1,3} \vee W_{1,2} \vee W_{2,2} \vee W_{1,1}$$



# Conclude $W_{1,3}$ ?

- Does the Wumpus reside in square (1,3) ?
- In other words, can we infer $W_{1,3}$ from our knowledge base?

$$KB \vdash_i W_{1,3}$$

# Conclude $w_{1,3}$ (Step #1)

- Modus Ponens $\frac{\alpha \Rightarrow \beta, \ \alpha}{\beta}$

$R_1$: $\neg S_{1,1} \Rightarrow \neg W_{1,1} \wedge \neg W_{1,2} \wedge \neg W_{2,1}$
Percept: $\neg S_{1,1}$

- Infer

$\neg W_{1,1} \wedge \neg W_{1,2} \wedge \neg W_{2,1}$

# Conclude $w_{1,3}$ (Step #2)

- And-Elimination $\dfrac{\alpha_1 \wedge \alpha_2 \wedge \ldots \wedge \alpha_n}{\alpha_i}$

$$\neg W_{1,1} \wedge \neg W_{1,2} \wedge \neg W_{2,1}$$

- Infer

$$\neg W_{1,1} \quad \neg W_{1,2} \quad \neg W_{2,1}$$

# Conclude $w_{1,3}$ (Step #3)

- Modus Ponens $\frac{\alpha \Rightarrow \beta, \quad \alpha}{\beta}$

  $R_2$: $\neg S_{2,1} \Rightarrow \neg W_{1,1} \wedge \neg W_{2,1} \wedge \neg W_{2,2} \wedge \neg W_{3,1}$
  Percept: $\neg S_{2,1}$

- Infer

  $\neg W_{1,1} \wedge \neg W_{2,1} \wedge \neg W_{2,2} \wedge \neg W_{3,1}$

- And-Elimination $\frac{\alpha_1 \wedge \alpha_2 \wedge \ldots \wedge \alpha_n}{\alpha_i}$

  $\neg W_{1,1}$ $\neg W_{2,1}$ $\neg W_{2,2}$ $\neg W_{3,1}$

# Conclude $w_{1,3}$ (Step #4)

- Modus Ponens $\frac{\alpha \Rightarrow \beta, \;\; \alpha}{\beta}$

$R_4$: $S_{1,2} \Rightarrow W_{1,3} \vee W_{1,2} \vee W_{2,2} \vee W_{1,1}$
Percept: $S_{1,2}$

- Infer

$W_{1,3} \vee W_{1,2} \vee W_{2,2} \vee W_{1,1}$

# Conclude $w_{1,3}$ (Step #5)

- Unit Resolution $\dfrac{\alpha \vee \beta, \ \neg\beta}{\alpha}$

  $W_{1,3} \vee W_{1,2} \vee W_{2,2} \vee W_{1,1}$ from Step #4

  $\neg W_{1,1}$ from Step #2

- Infer

  $W_{1,3} \vee W_{1,2} \vee W_{2,2}$

# Conclude $w_{1,3}$ (Step #6)

- Unit Resolution $\frac{\alpha \vee \beta, \;\; \neg\beta}{\alpha}$

  $W_{1,3} \vee W_{1,2} \vee W_{2,2}$ from Step #5

  $\neg\, W_{2,2}$ from Step #3

- Infer

  $W_{1,3} \vee W_{1,2}$

# Conclude $W_{1,3}$ (Step #7)

- Unit Resolution

$$\frac{\alpha \vee \beta, \ \neg\beta}{\alpha}$$

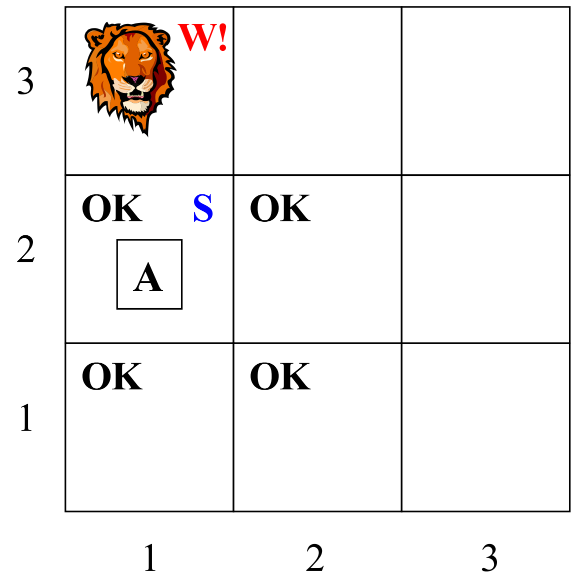
$W_{1,3} \vee W_{1,2}$ from Step #6

$\neg W_{1,2}$ from Step #2

- Infer

$W_{1,3} \rightarrow$ The wumpus is in cell 1,3!!!

# Wumpus in $W_{1,3}$



# Propositional Logic

- Propositional logic commits to existence of facts about the world being represented
  - Simple syntax and semantics
- Proof methods
  - Truth table
  - Inference rules
    - Modus Ponens
    - And-Elimination
    - And/Or-Introduction
    - Double-Negation Elimination
    - Unit Resolution
    - Resolution
- Propositional logic quickly becomes impractical