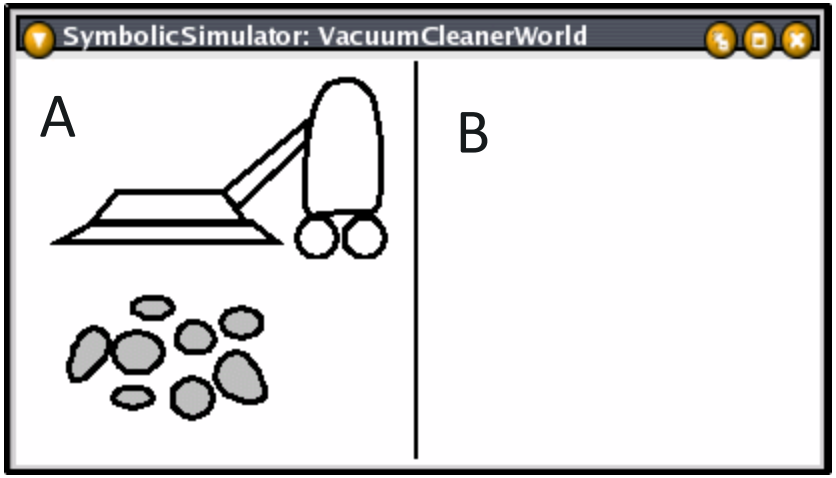
Propositional Logic

knowledge-based agents, logic, theorem proving

CS 4100/5100
Foundations of AI

A Return to Vacuum World

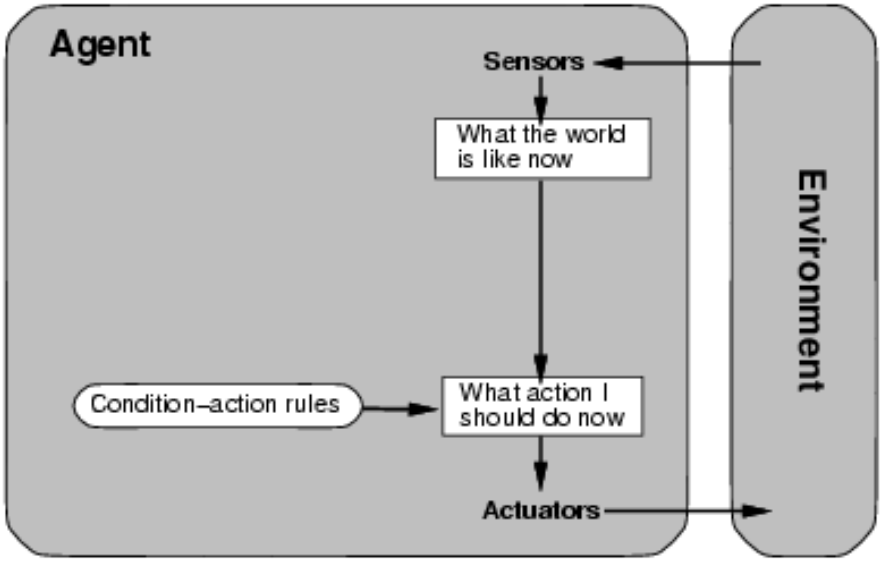


Percepts: location and contents (e.g., [A, Dirty])

Actions: Left, Right, Suck, NoOp

Production Rules: [A, Dirty] -> Suck; [B, Dirty] -> Suck;
[A, Clean] -> Right; [B, Clean] -> Left

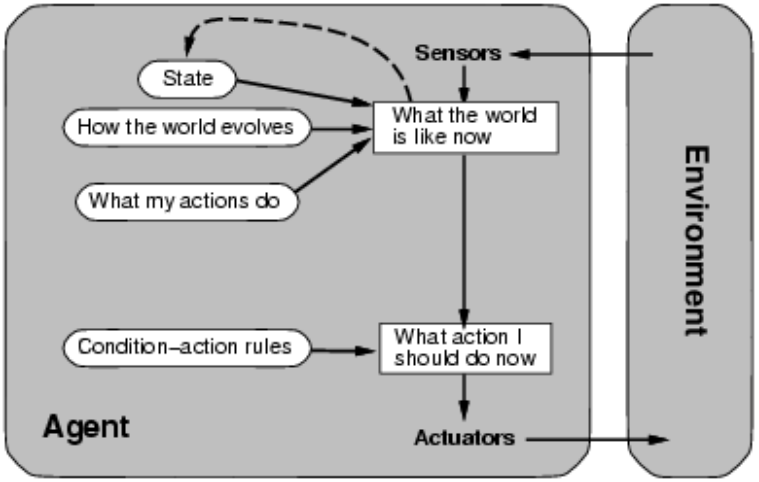
The Vacuum as a Reflex Agent



- Only knows current state
- No explicit knowledge of whole environment
- How does this scale?

The Vacuum as a Knowledge-Based Agent

- Formal language for expressing knowledge
  - **Knowledge representation**
- Knowledge base design: what is known?
  - **Ontology design**
- Algorithms to use and update the knowledge base
  - **Automated inference**



Knowledge Representation: Formal Logic

- **Declarative** approach
  - Beliefs are a set of sentences
- Sentences are believed true by the agent
- More sentences can be added
- Sentences can be retracted

Limitations of Logic

- Belief must be definite and consistent
- Dynamic worlds are more awkward
  - Modeling intuitive knowledge
- Human behavior and emotion

Formal Logic

- Still very useful!
- Simple to compute
- Well-understood rules for reasoning
- Let's start with **propositional logic**

Logics

- **Syntax**: how sentences are formed
  - Algebra: x + y = 2 is ok, x2 + = y is not
  - English: I love Star Trek is ok, Star I Trek love is not

- **Semantics**: when sentences are true, and what sentences mean
  - Algebra: x + y = 4 is true for x = 2, y = 2
    x = 1, y = 3, ….
  - I love Star Trek is true for I = Gillian, …

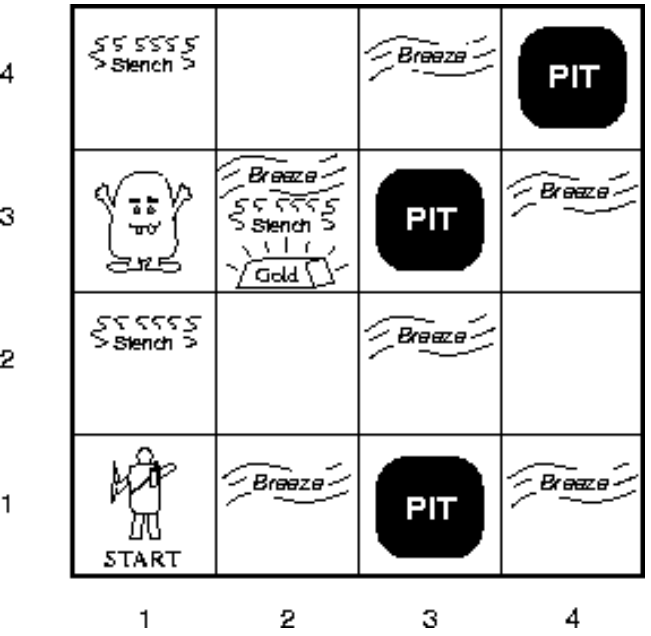
Propositional Logic: Syntax

- **Proposition symbols**: X, Y, PhoneRinging, …
  - Special symbols with fixed meaning: **True**, **False**
- **Atomic sentences**: single proposition symbol
- Complex sentences: combine propositional symbols with **connectives**

Propositional Logic: Syntax

- ¬ not
- ∧ and
- ∨ or
- ⇒ implies
- ⇔ if and only if
- ⊕ xor

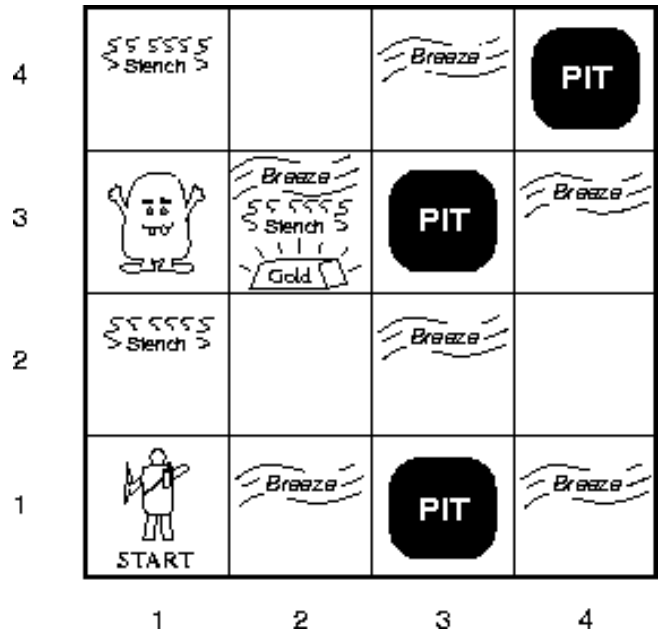
Wumpus World



- **Environment**
  - 4x4 grid – agent starts at [1, 1]
  - 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 face it
  - Shooting uses the only arrow
  - Grabbing picks up gold in the same square
  - Climbing exits the cave if at [1,1]
- **Actions**: Forward, TurnLeft, TurnRight, Grab, Shoot, Climb
- **Percepts**: Stench, Breeze, Glitter, Bump, Scream

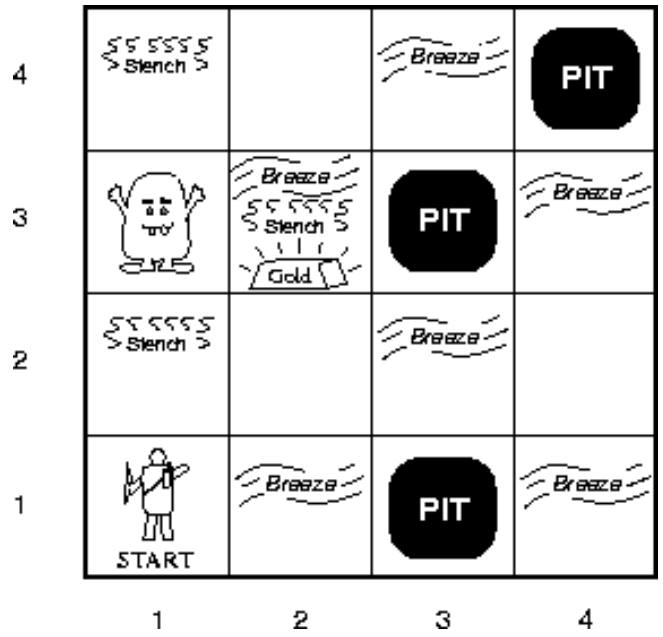
Wumpus World: Proposition Symbols

- World Representation:
  - $P_{x,y}$
  - $W_{x,y}$
- Agent Perception:
  - $S_{x,y}$
  - $B_{x,y}$



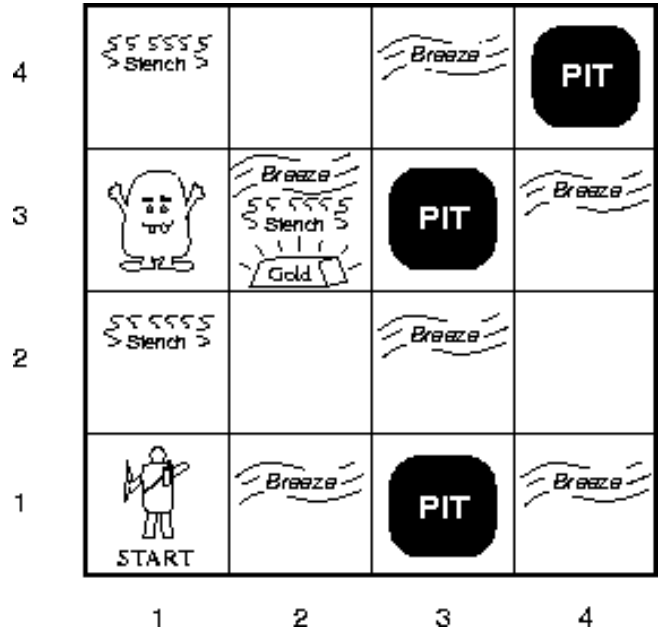
Wumpus World: Axioms

- $\sim P_{1,1}$.
- $B_{1,1} \leftrightarrow (P_{1,2} \vee P_{2,1})$.
- $B_{2,1} \leftrightarrow (P_{1,1} \vee P_{2,2} \vee P_{3,1})$.
- $B_{3,2} \leftrightarrow (P_{2,3} \vee P_{3,3} \vee P_{4,2} \vee P_{3,1})$.



Wumpus World: Percepts

- $\sim B_{1,1}$.
- $B_{2,1}$.



Propositional Model: Semantics

- **Models:** possible worlds
  - Assigned values to all proposition symbols
  - Truth tables

X	Y	¬X	X ∧ Y	X ⇒ Y
False	False	True	False	True
False	True	True	False	True
True	False	False	False	False
True	True	False	True	True

More on Semantics

- If α is true in model M, then M is "a model of" α
  - also: M *satisfies* α
  - notation: M(α)
- Entailment (⊨): one sentences follows logically from another
  - α ⊨ β if and only if M(α) and M(β)
  - in every model in which α is true, β is also true
  - M(α) ⊆ M(β)
- Necessary principle for **inference**

Entailment and Knowledge Bases

- The knowledge base (KB) used by our agent is a set of sentences
  - A database of things the agent knows
- KB entails sentence α iff α is true in all worlds where the KB is true

Logical Inference

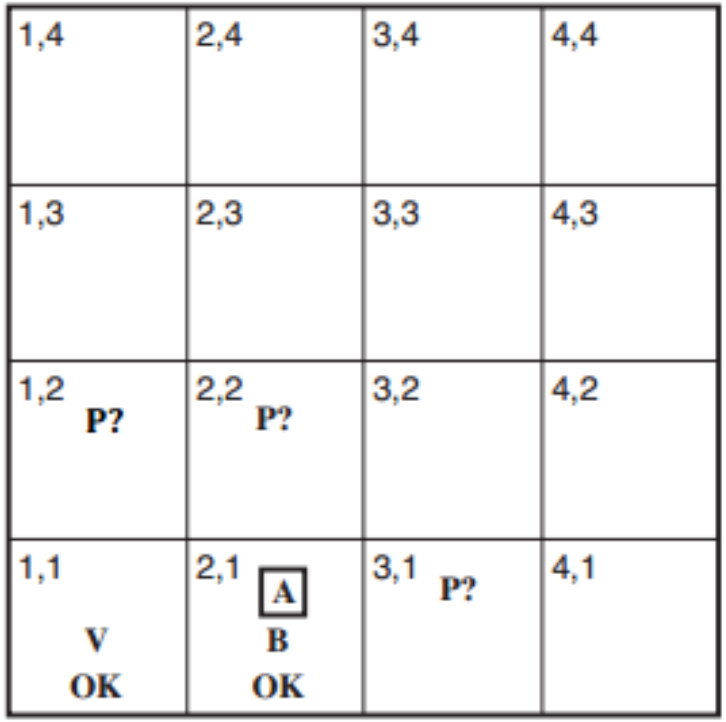
- KB $\vdash_i \alpha$ means "α is derived from KB by algorithm i"

- Inference algorithm that derives only entailed sentences is called **sound**

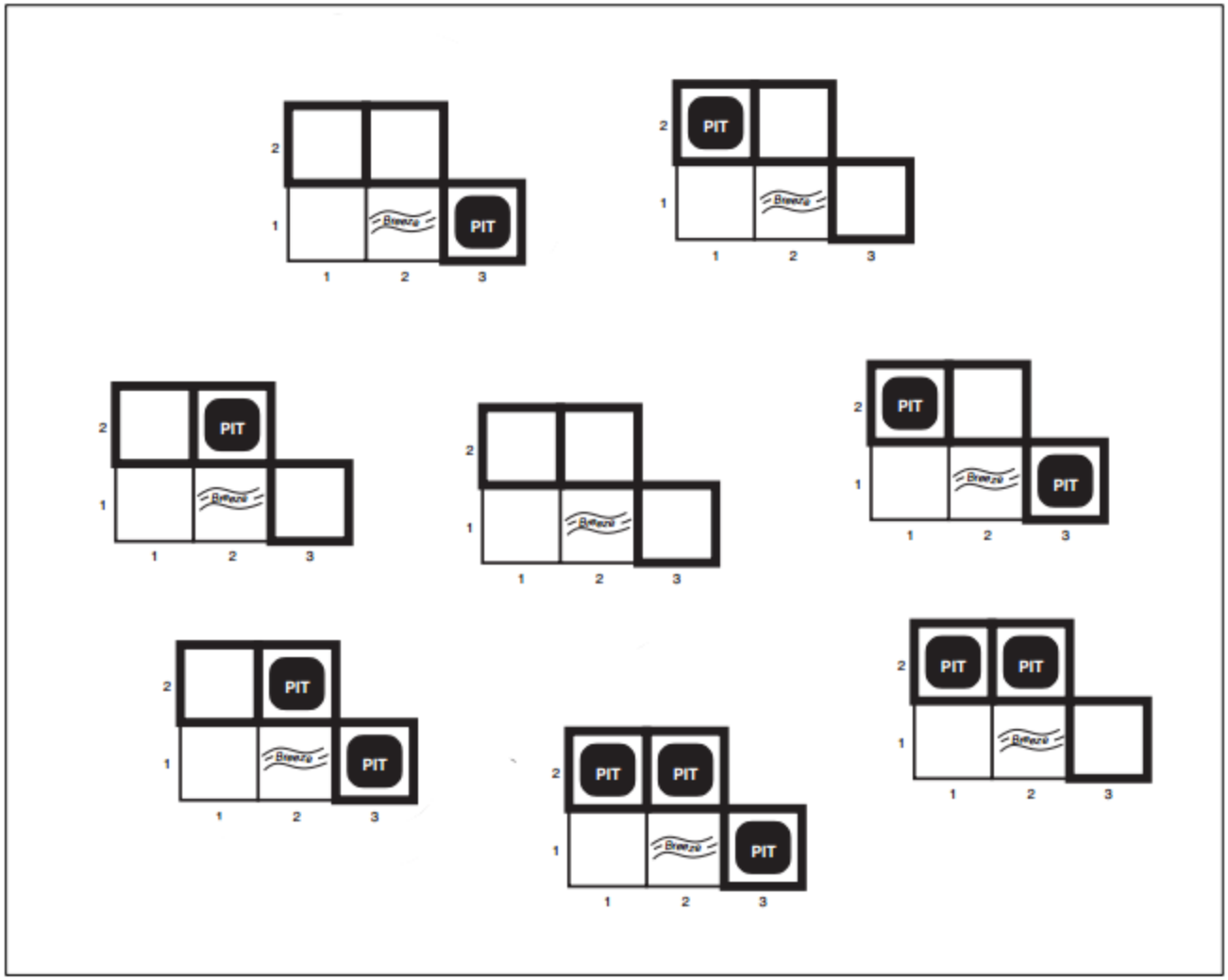
- Inference algorithm that can derive all entailed sentences is called **complete**

Model Checking

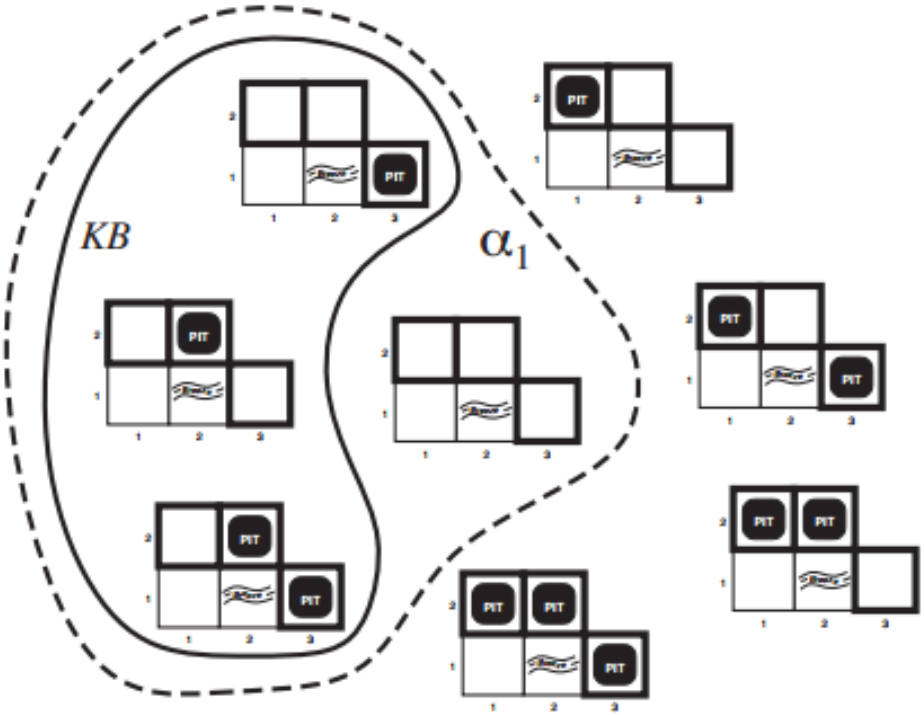
- Check every model to see if α is true!
  - Yes, this is really slow.
  - $O(2^n)$
- Action sequence: detect nothing at [1,1], move right, detect breeze at [2,1]
- What can we say about location of pits in adjacent squares?
  - 3 boolean choices (pit or no pit) -> 2^3 = 8 possible worlds



Wumpus Models (Possible Worlds)

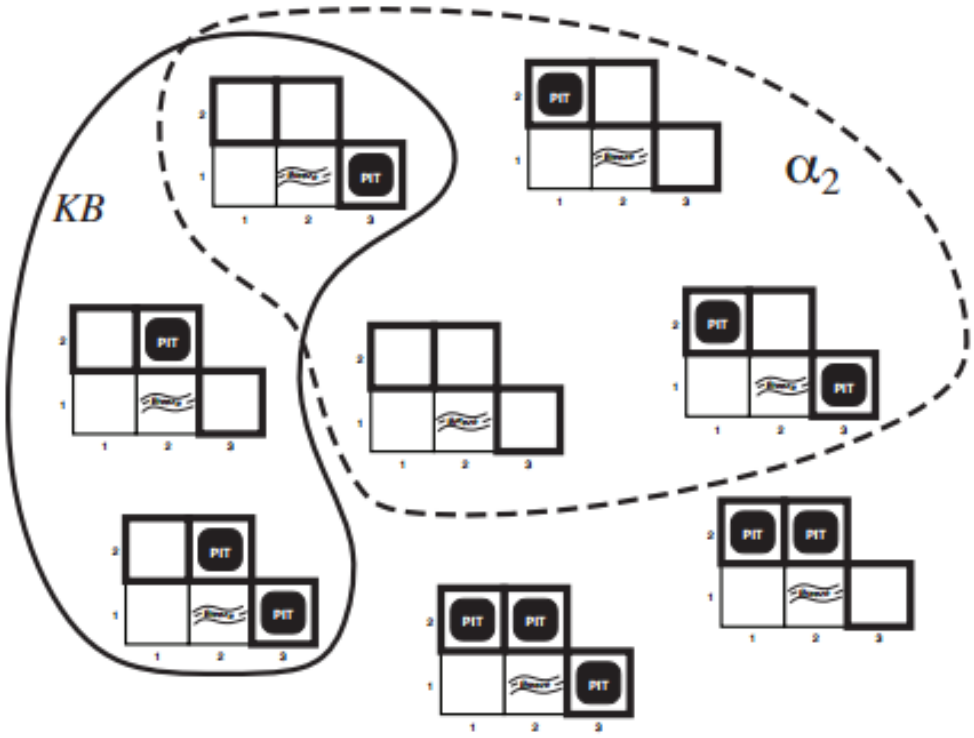


Wumpus Models



- KB = wumpus world rules + observations
- $\alpha_1 = \sim P_{1,2}$.
- KB entails α_1 by model checking

Wumpus Models



- KB = wumpus world rules + observations
- $\alpha_2 = \sim P_{2,2}$.
- KB does **not** entail α_2 by model checking