AI

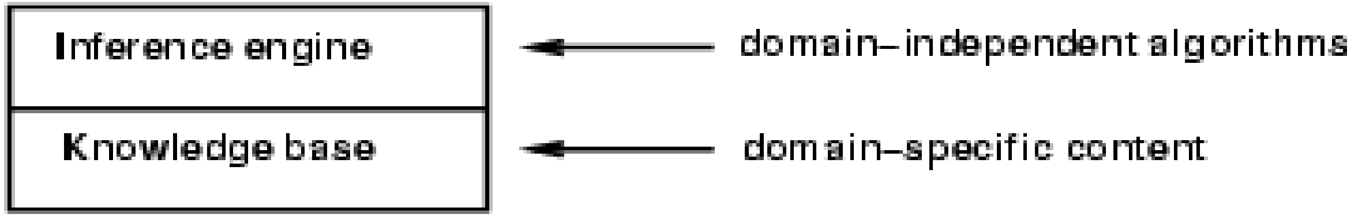
Artificial Intelligence

8.3.1
Logical Agents (Ch. 7)

Outline

- Knowledge-based agents
- Wumpus world
- Logic in general – models and entailment
- Propositional (Boolean) logic
- Equivalence, validity, satisfiability
- Inference rules and theorem proving
  - forward chaining
  - backward chaining
  - resolution

Knowledge Bases



- Knowledge base = set of sentences in a formal language
- Declarative approach to building an agent (or other system):
  - `Tell` it what it needs to know
- Then it can `Ask` itself what to do - answers should follow from the KB
- Agents can be viewed at the knowledge level

  i.e., what they know, regardless of how implemented
- Or at the implementation level
  - i.e., data structures in KB and algorithms that manipulate them

A simple knowledge-based agent

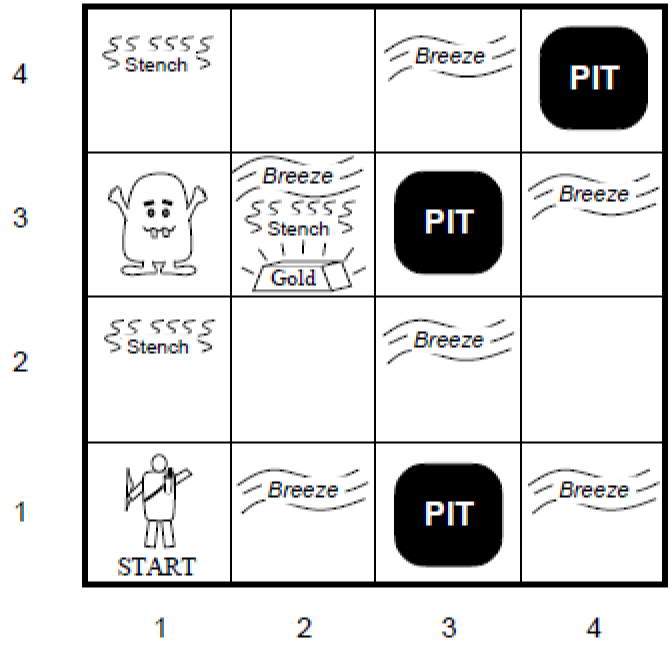
```
function KB-AGENT(percept) returns an action
    static: KB, a knowledge base
            t, a counter, initially 0, indicating time

    TELL(KB, MAKE-PERCEPT-SENTENCE(percept, t))
    action ← ASK(KB, MAKE-ACTION-QUERY(t))
    TELL(KB, MAKE-ACTION-SENTENCE(action, t))
    t ← t + 1
    return action
```

- The agent must be able to:
  - Represent states, actions, etc.
  - Incorporate new percepts
  - Update internal representations of the world
  - Deduce hidden properties of the world
  - Deduce appropriate actions

Wumpus World PEAS description

- Performance measure
  - gold +1000, death -1000
  - -1 per step, -10 for using the arrow
- Environment
  - 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 are facing it
  - Shooting uses up the only arrow
  - Grabbing picks up gold if in same square
  - Releasing drops the gold in same square



- Sensors: Stench, Breeze, Glitter, Bump, Scream
- Actuators: Left turn, Right turn, Forward, Grab, Release, Shoot

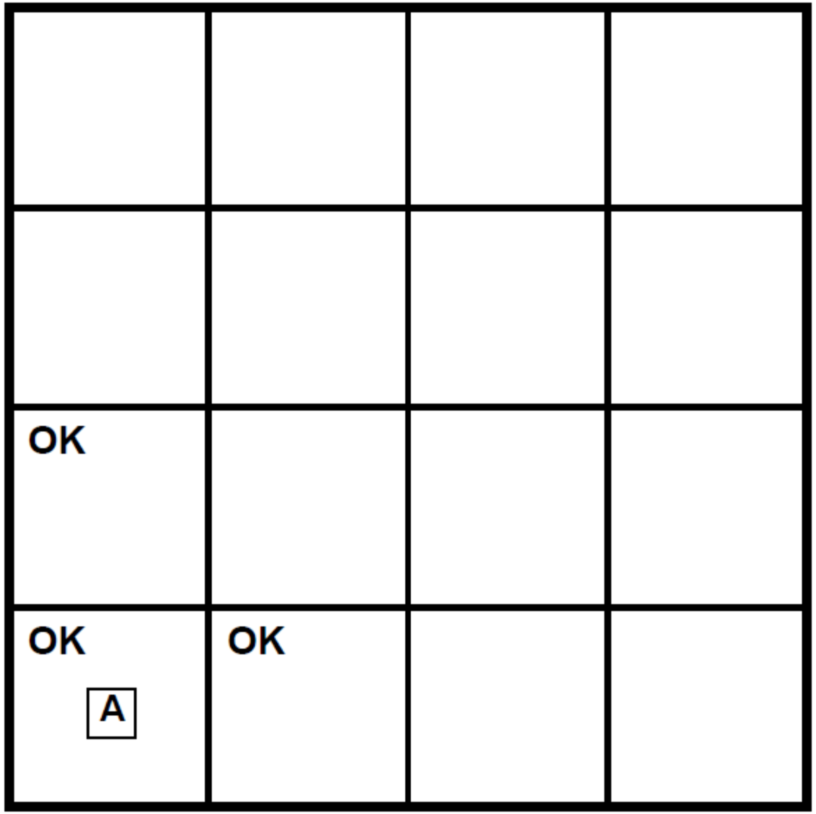
Links

- Minesweeper
  - http://minesweeperonline.com/
- Battleships
  - http://www.battleshiponline.org

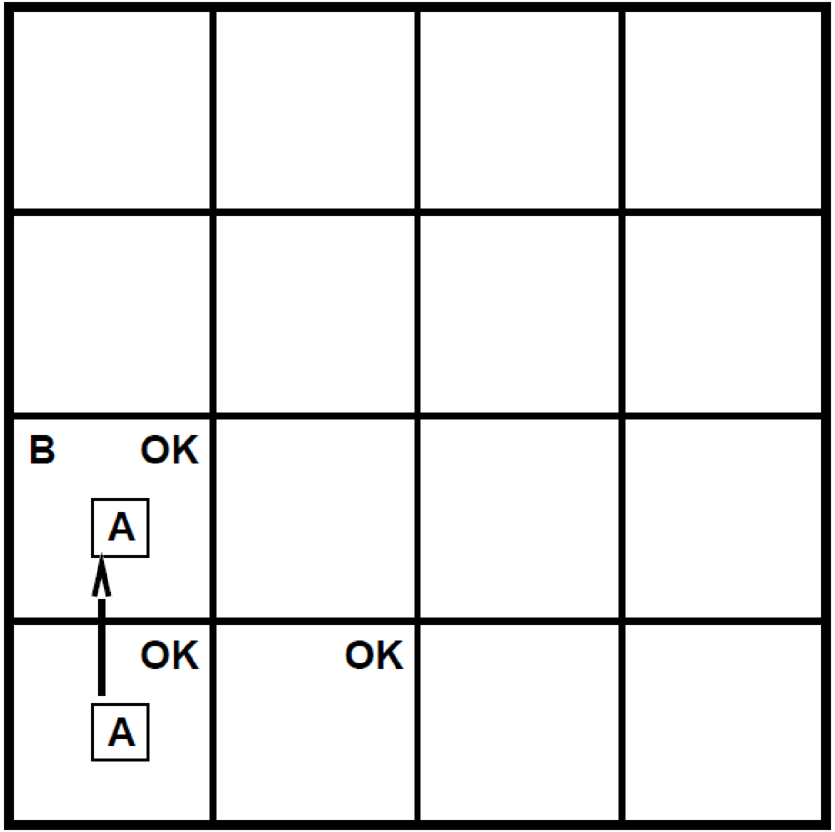
Wumpus world characterization

- Fully Observable
  - No – only local perception
- Deterministic
  - Yes – outcomes exactly specified
- Episodic
  - No – sequential at the level of actions
- Static
  - Yes – Wumpus and Pits do not move
- Discrete
  - Yes
- Single-agent?
  - Yes – Wumpus is essentially a natural feature

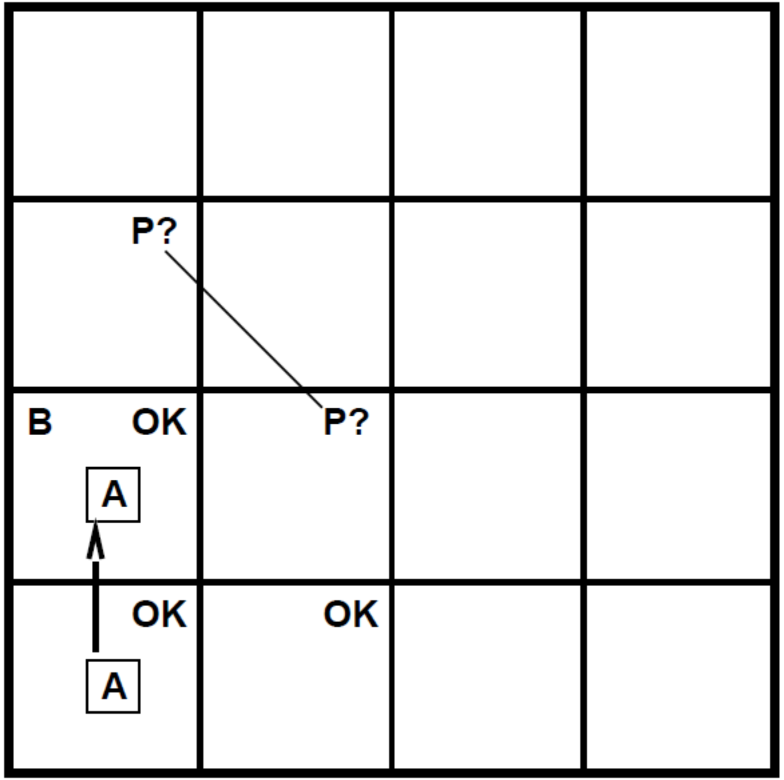
Exploring a wumpus world



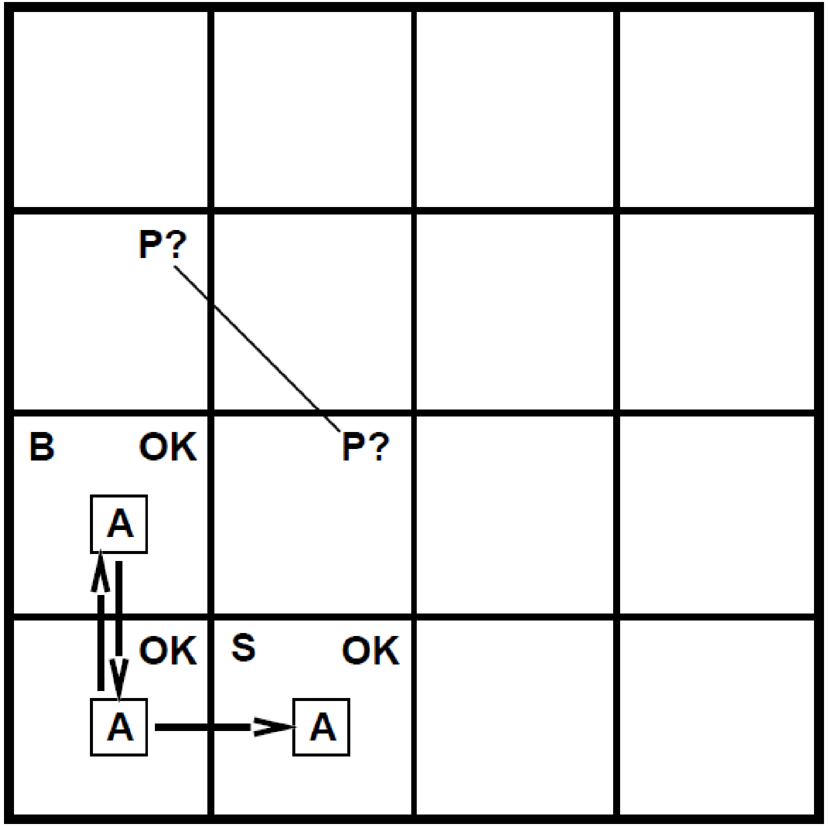
Exploring a wumpus world



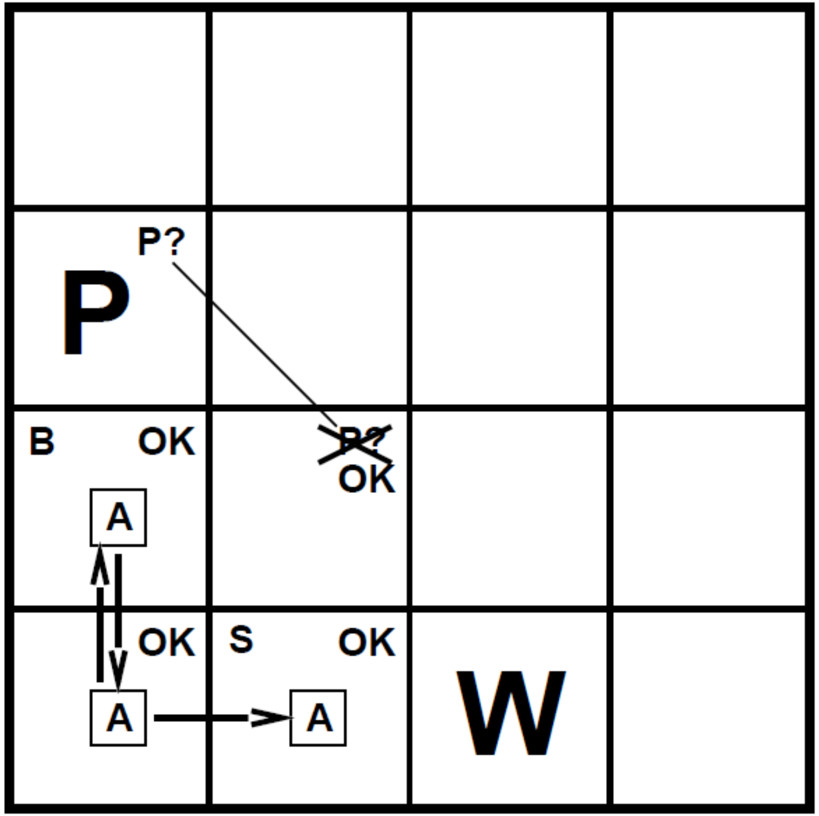
Exploring a wumpus world



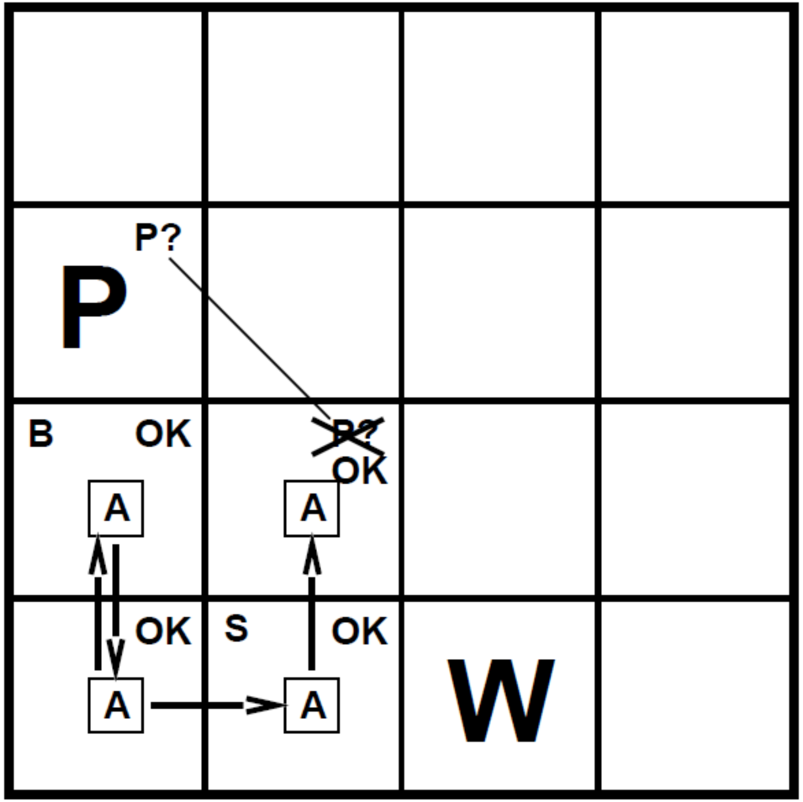
Exploring a wumpus world



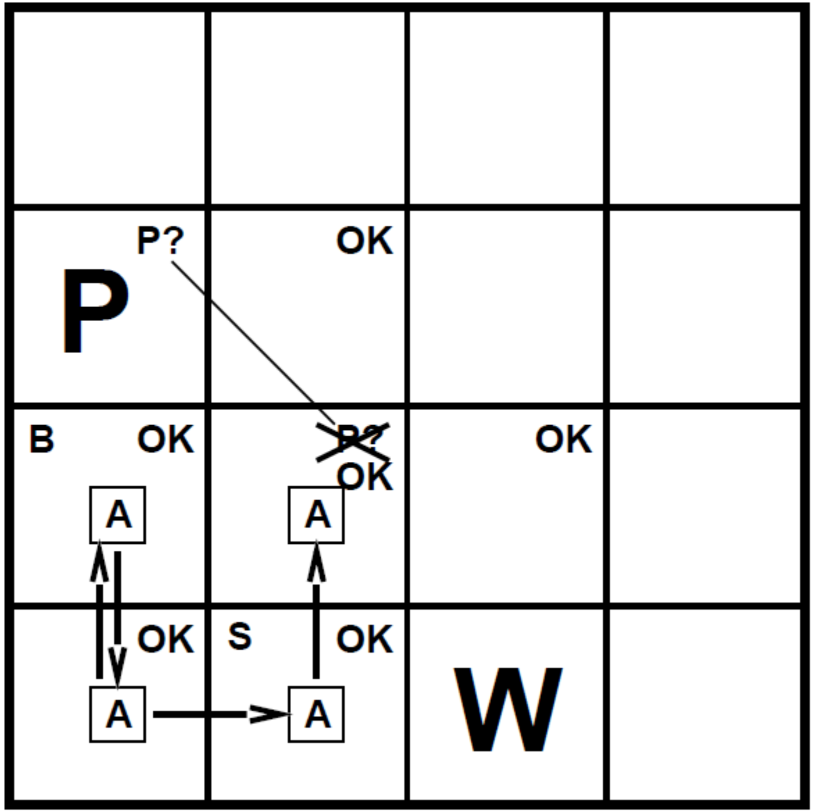
Exploring a wumpus world



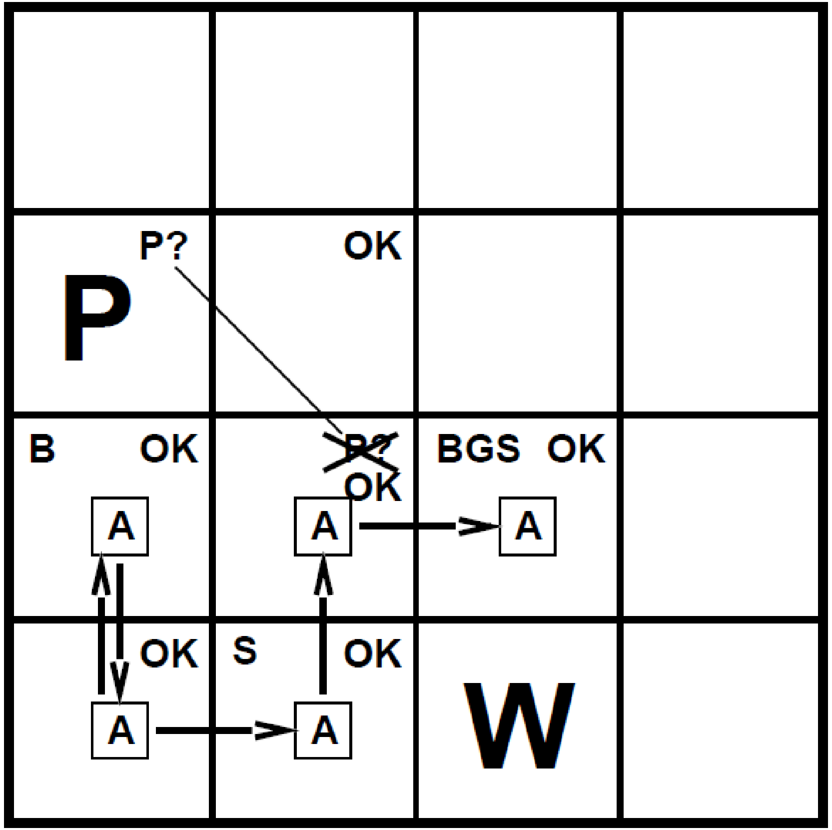
Exploring a wumpus world



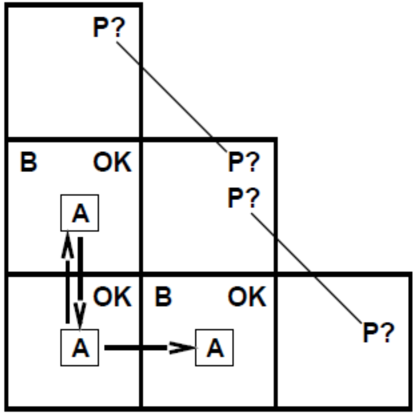
Exploring a wumpus world



Exploring a wumpus world

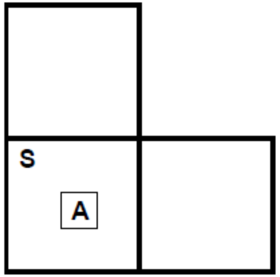


Other tight spots



Breeze in (1,2) and (2,1)
$\Rightarrow$ no safe actions

Assuming pits uniformly distributed,
(2,2) has pit w/ prob 0.86, vs. 0.31



Smell in (1,1)
$\Rightarrow$ cannot move

Can use a strategy of coercion:
shoot straight ahead
wumpus was there $\Rightarrow$ dead $\Rightarrow$ safe
wumpus wasn't there $\Rightarrow$ safe

Logic in general

- Logics are formal languages for representing information such that conclusions can be drawn
- Syntax defines the sentences in the language
- Semantics define the "meaning" of sentences;
  - i.e., define truth of a sentence in a world

- E.g., the language of arithmetic
  - $x+2 \geq y$ is a sentence; $x2+y >$ is not a sentence
  - $x+2 \geq y$ 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 $x = 0, y = 6$

Entailment

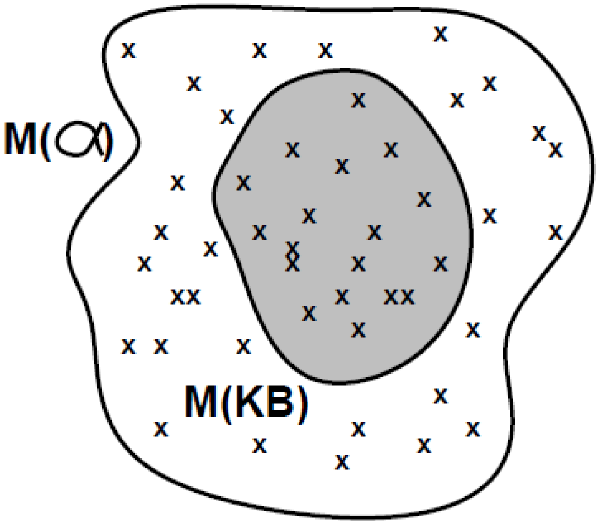
- Entailment means that one thing follows from another
- KB $\models \alpha$
- Knowledge base *KB* entails sentence α if and only if α is true in all worlds where *KB* is true
  - E.g., the KB containing “the Giants won” and “the Reds won” entails “Either the Giants won or the Reds won”
  - E.g., $x+y = 4$ entails $4 = x+y$
  - Entailment is a relationship between sentences (i.e., syntax) that is based on semantics

Models

- Logicians typically think in terms of models, which are formally structured worlds with respect to which truth can be evaluated
- We say *m* is a model of a sentence α if α is true in *m*

- *M(α)* is the set of all models of α

- Then KB $\models$ α iff *M(KB)* $\subseteq$ *M*(α)
  - E.g.
    - *KB* = Giants won and Reds won
    - α = Giants won

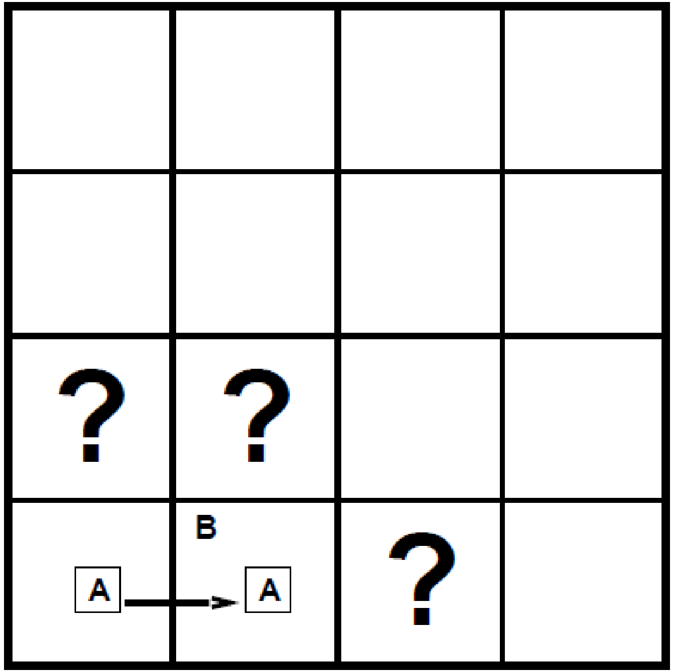


Entailment in the wumpus world

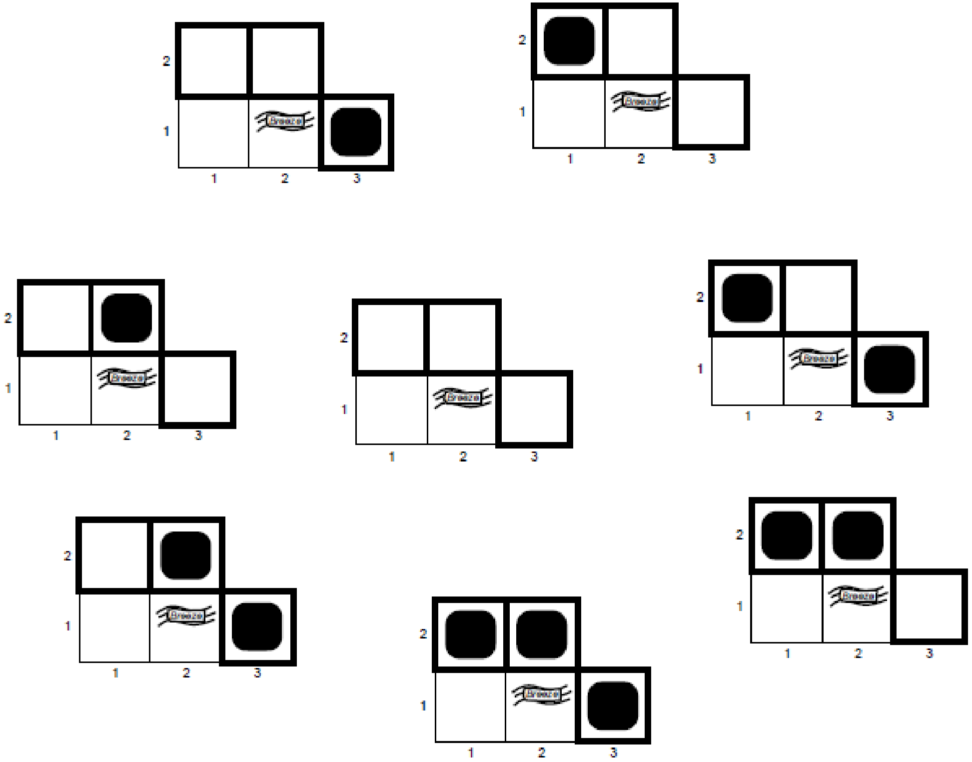
Situation after detecting nothing in [1,1], moving right, breeze in [2,1]

Consider possible models for *KB* assuming only pits

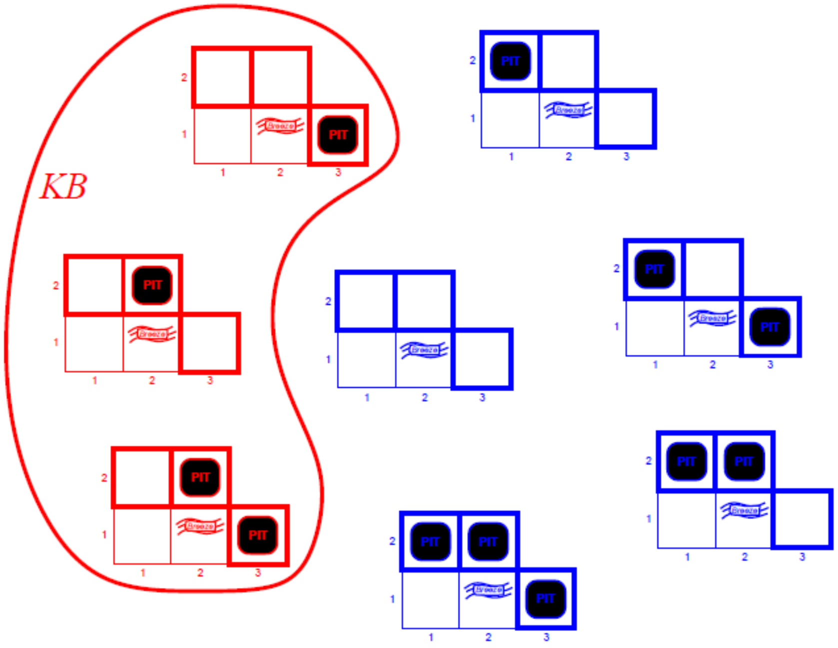
3 Boolean choices $\Rightarrow$ 8 possible models



Wumpus models

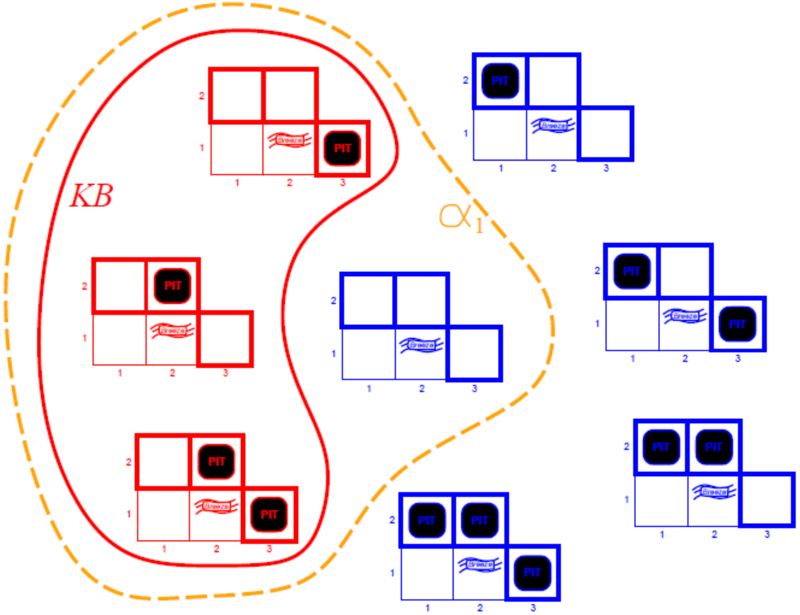


Wumpus models



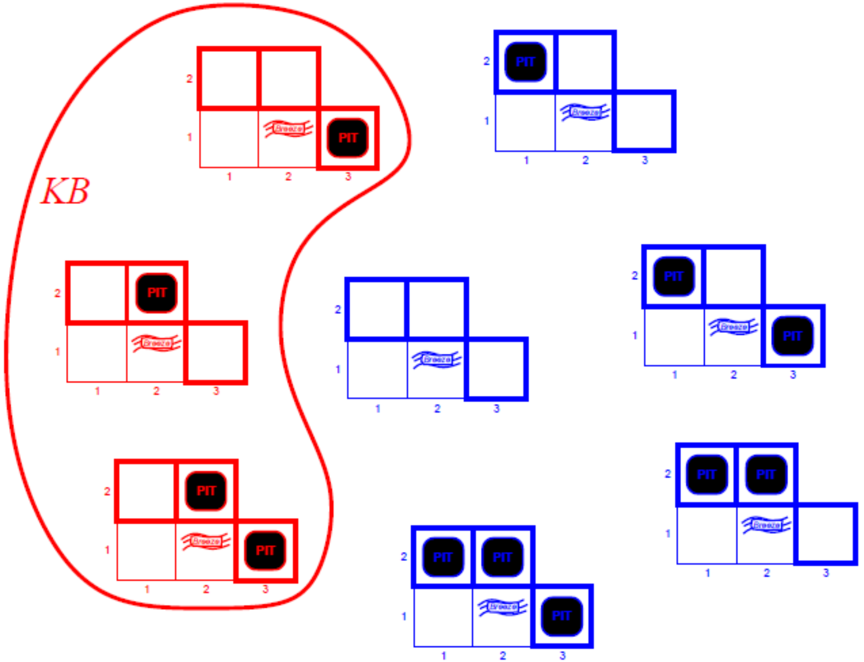
- *KB* = wumpus-world rules + observations

Wumpus models



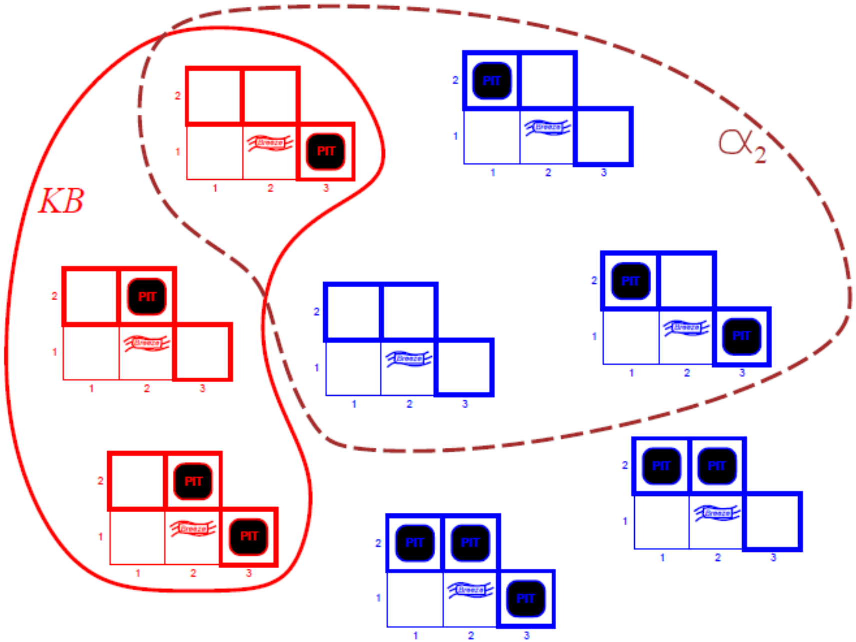
- *KB* = wumpus-world rules + observations
- α_1 = "[1,2] is safe", *KB* $\models \alpha_1$, proved by model checking

Wumpus models



- *KB* = wumpus-world rules + observations

Wumpus models



- *KB* = wumpus-world rules + observations
- α_2 = "[2,2] is safe", $KB \models \alpha_2$

Logic problem

- There are two types of people on an island:

- Knight: Always tells truth.
- Knave: Always lies

- A says: "B is a knight."
- B says: "The two of us are opposite types."

- Determine the types of A and B

Logical representation

- We can describe the puzzle by the following propositions:

  p: A is a knight, tells the truth.

  ^p: A is a knave, lies.

  q: B is a knight, tells the truth.

  ^q: B is a knave, lies.

Reasoning for p

- Suppose p=T:
  - A tells the truth: "B is a knight."
  - So B tells the truth.
  - B said: "The two of us are opposite types.".
  - So A and B are different types.
  - This is false, because both A and B are knights.

Reasoning for ^p

- Suppose p=F:
  - A lies. So B is a knave.
  - So B lies.
  - B said: "The two of us are opposite types.".
  - So A and B are the same type.
  - This holds and we get the conclusion:
  - Both A and B are knaves.

Inference

- $KB \vdash_i \alpha$ = sentence α can be derived from *KB* by procedure *i*
- Soundness: *i* is sound if whenever $KB \vdash_i \alpha$, it is also true that $KB \models \alpha$
- Completeness: *i* is complete if whenever $KB \models \alpha$, it is also true that $KB \vdash_i \alpha$
- Preview: we will define a logic (first-order logic) which is expressive enough to say almost anything of interest, and for which there exists a sound and complete inference procedure.
- That is, the procedure will answer any question whose answer follows from what is known by the *KB*.

Models

- State-based
  - states, actions, costs
  - Used for route finding, game playing
- Variable-based
  - variables, values, domains
  - used for scheduling, tracking, medical diagnosis
- Logic-based
  - logical formulas and inference rules
  - used for theorem proving, verification, reasoning

[Example from Percy Liang]

Historical Notes

- Logic was the dominant paradigm in AI until the 1990s
- Problem 1
  - didn't handle uncertainty (however, one can use probabilities)
- Problem 2
  - didn't address fine tuning from data (however, this can be solved using machine learning)
- Strength
  - provides expressiveness in a compact way

[Example from Percy Liang]

AI