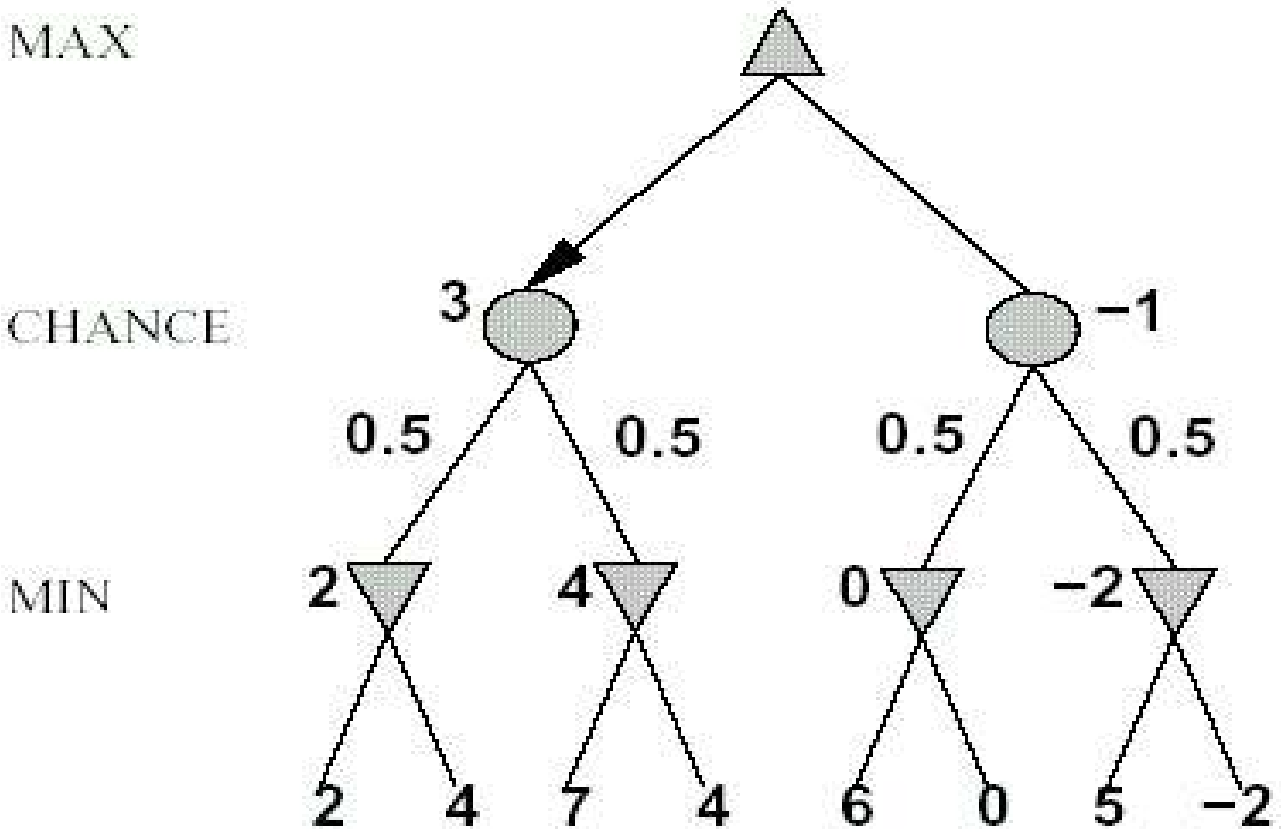
CSC421 Intro to Artificial Intelligence

UNIT 08:
Logical Agents

Deterministic games in practice

- Checkers: 1994 ended 40-year reign of human champion Marion Tinsley
- Chess: 1997 Gary Kasparov – Deep Blue 200 milliion positions/second – some lines of search up to 40 ply
- Othello: human champions refuse to compete with computers because they are too good
- Go: Human champions refuse to play with computers, who are too bad. Branching factor b > 300 pattern databases

Non-deterministic games



EXPECTMINIMAX

```
...
if state is a MAX node then
        return the highest EXPECTIMINIMAX-VALUE of SUCCESSORS(state)
if state is a MIN node then
        return the lowest EXPECTIMINIMAX-VALUE of SUCCESSORS(state)
if state is a chance node then
        return average of EXPECTIMINIMAX-VALUE of SUCCESSORS(state)
...
```

Like MinMax with addition of CHANCE nodes – gives perfect play

Nondeterministic games in practice

- Dice rolls increase b
- As depth increases, probability of reaching a given node shrinks
- Alpha-beta doesn't help much
- TDGammon – depth-2 search + very good EVAL = world champion level

Outline

- Knowledge-based agents
- Wumpus world
- Logic in general
- Humans are among other things information processors
  - One of the strengths of human information processing is our ability to represent and manipulate logical information
- Poll about courses in Logic and CS

Knowledge bases

Inference engine	Domain independent algorithms
Knowledge base	Domain-specific content

Knowledge base = set of “sentences” in a formal language

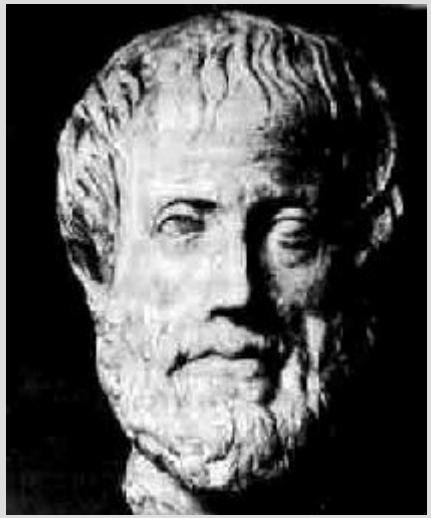
Declarative approach to building an agent:
TELL it what it needs to know

Then ASK itself what to do – answers should follow from the KB

Knowledge level
Implementation level

Aristotle (384-322 BC)

- Concept of proof = series immediately obvious reasoning steps
- One of the many important contributions of Aristotle:
  - Step of proof is obvious based on form rather than content
- Examples
  - All x are y
  - All y are z
  - Therefore all x are z
- X = dogs, Y = mammals, Z = animals
- X = Accords, Y = Hondas, Z = Japanese



(Un)Sound patterns

- What is a good pattern ?
- Example
  - All X are Y
  - Some Y are Z
  - Therefore, some X are Z
- Is this a “sound” pattern ?
- Any examples ?
- Dedection
  - A “correct”(sound) pattern must always lead to correct conclusions i.e. Conclusions that are correct whenever the premises are true

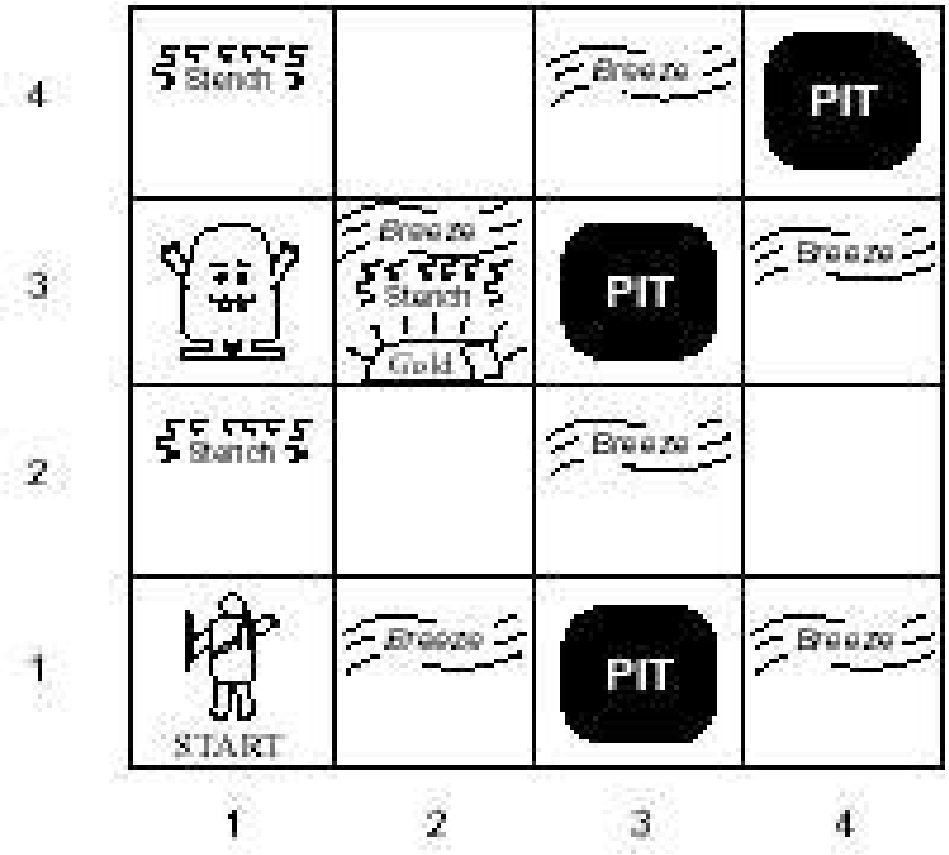
KB agent

```
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
```

Represent states/actions
Incorporate new precepts
Update internal representations of the world
Deduce hidden properties of the world
Deduce appropriate actions

Wumpus World



PEAS description

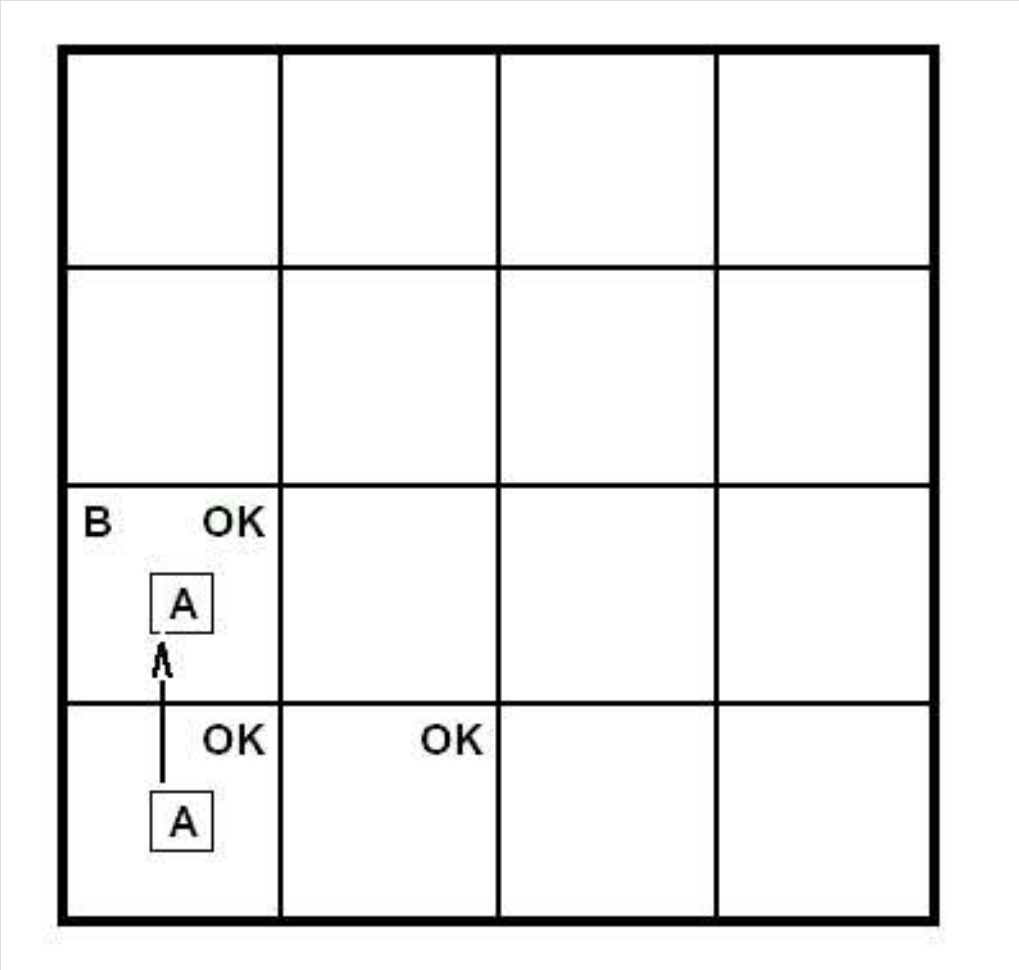
Exploring Wumpus

OK			
OK A	OK		

Percept:
[None, None, None, None, None]

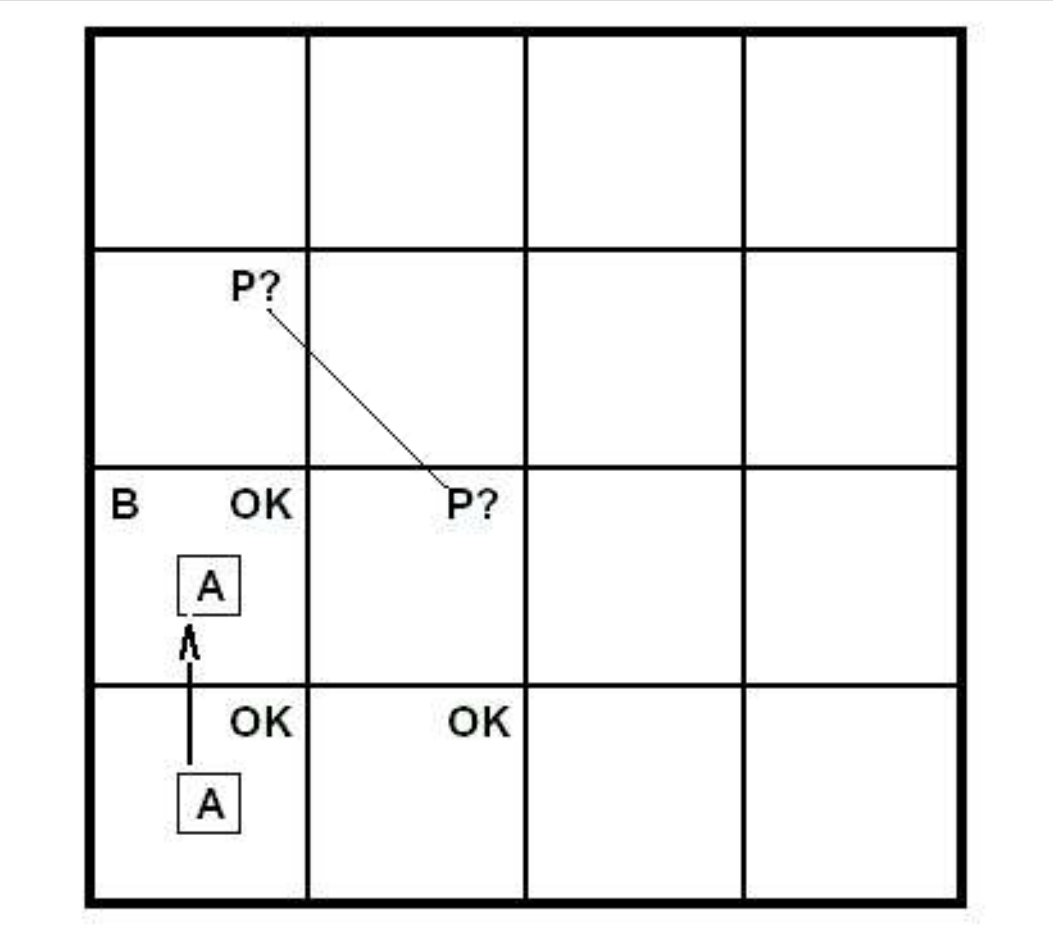
Percept =
- Stench
- Breeze
- Glitter
- Bump
- WumpusDead

Exploring Wumpus

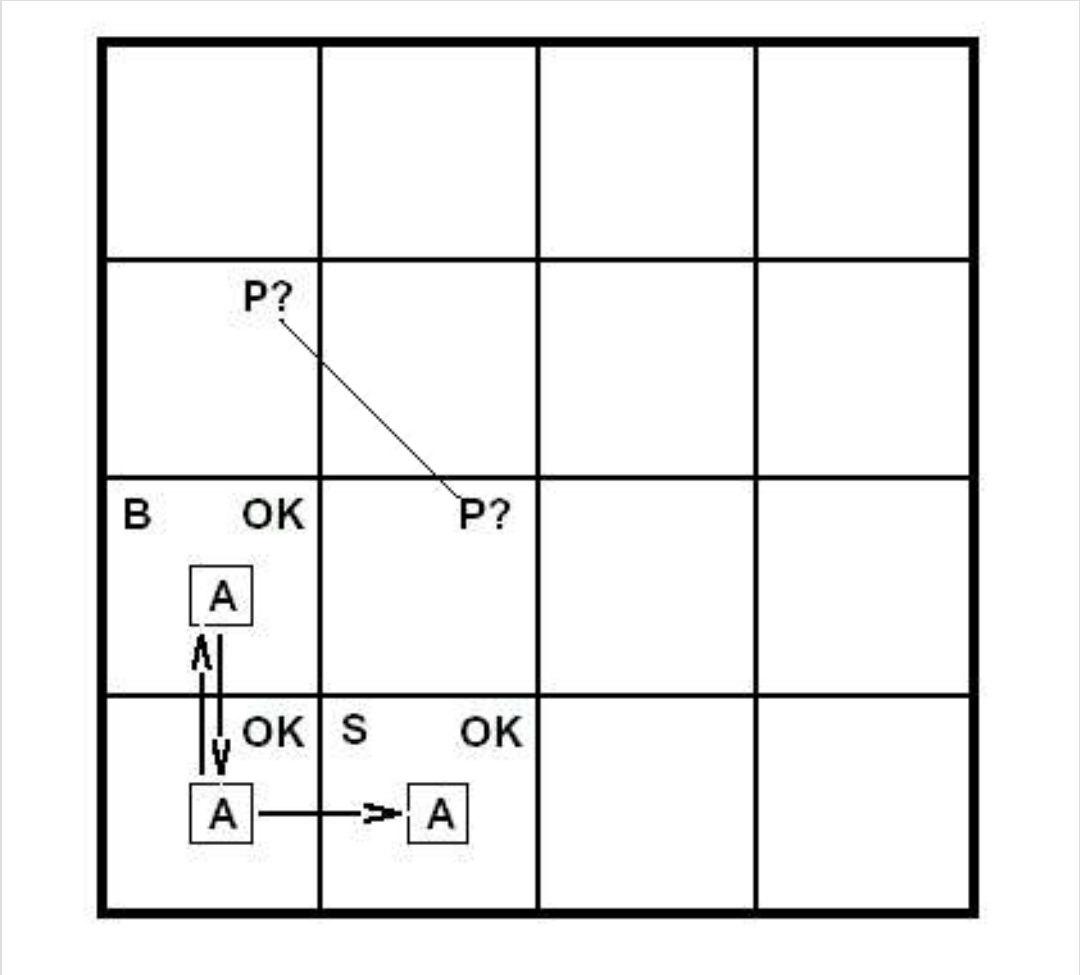


[none, breeze, none, none, none]

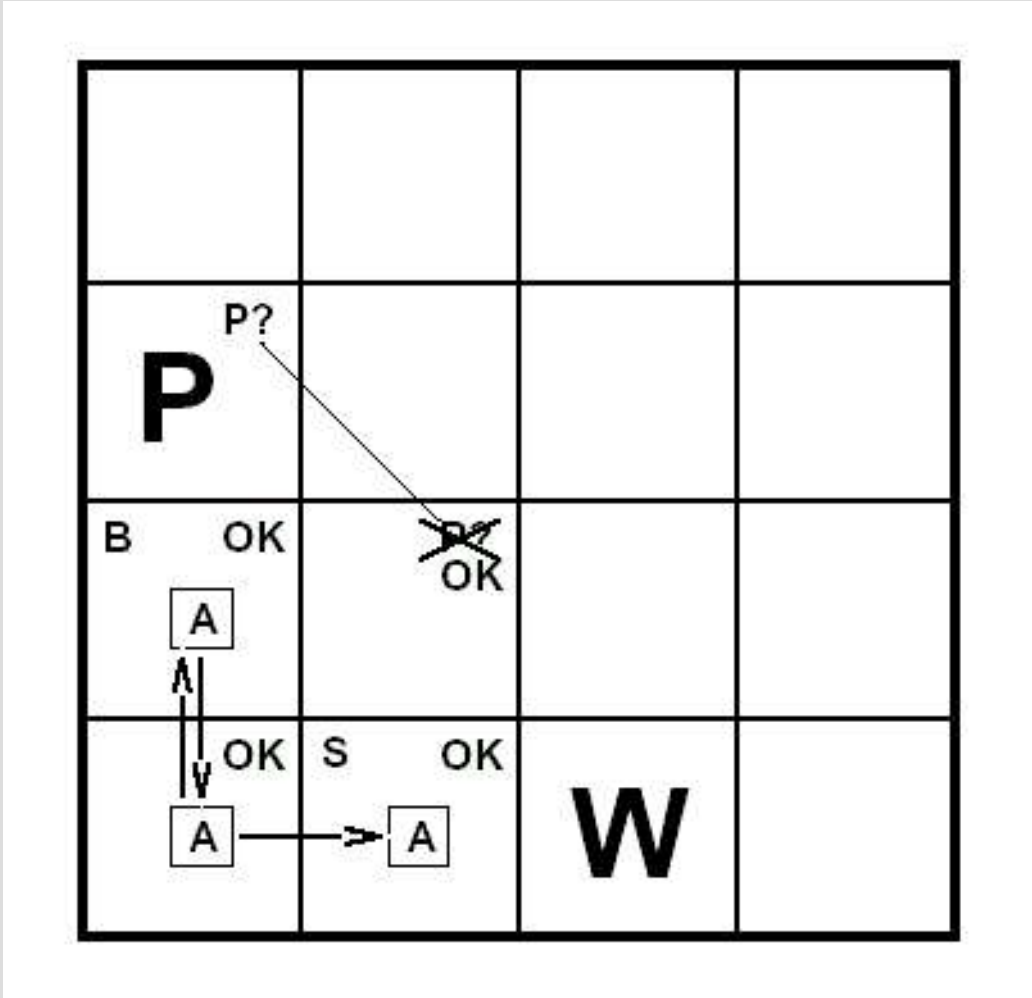
Exploring Wumpus



Exploring Wumpus



Exploring Wumpus



The agent has "deduced" the location of the pit and the wumpus without falling into a horrible death or being eaten alive by the hungry wumpus

Logics

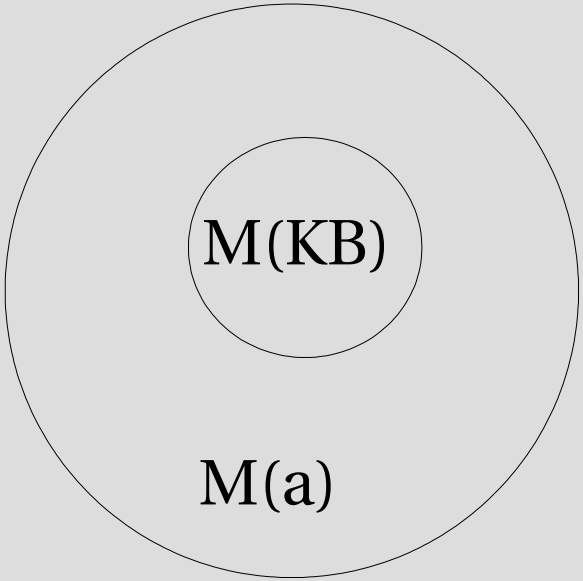
- Formal languages for encoding information
- Legal transformations
- Syntax defines the sentences in the language
- Semantics define the “meaning” of a sentence i.e define the truth of a sentence in a world
- For example
  - x + 2 >= y is true in a world where x = 5 and y = 2
  - x + 2 >= y is false in a world where x = 2 and y = 10

Entailment

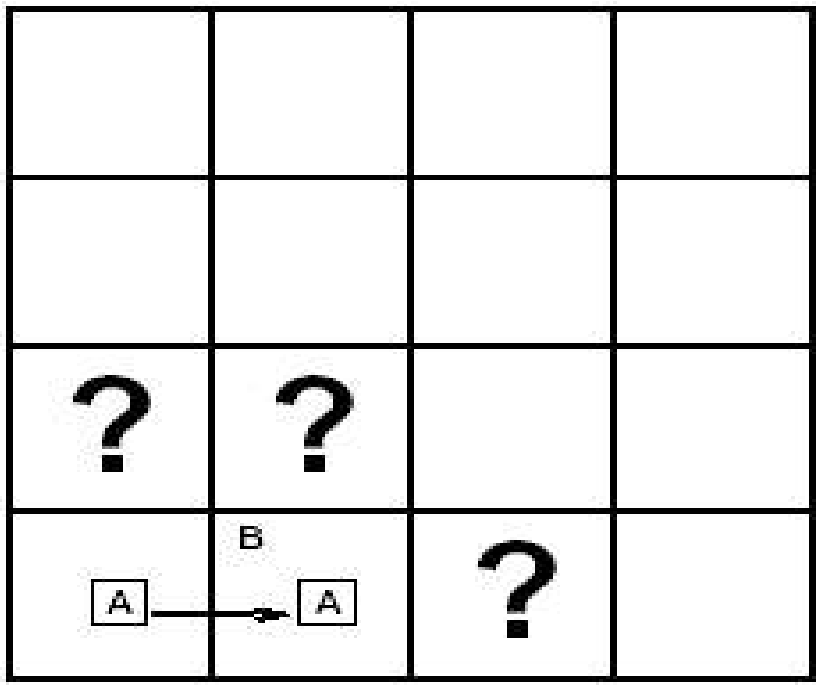
- Entailment means that one thing follows from another:
  - KB |= a
- KB entails sentence a iff a is true in all worlds where the KB is true
- X + Y = 4 entail X – 4 = Y
- Entailment is a relationship between sentences (syntax) that is based on semantics

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 a if a is true in m
- M(a) is the set of all models of a
- KB |= a iff M(KB) $\subseteq$ M(a)
- KB = Giants won and Reds Won
  - a = Giants won

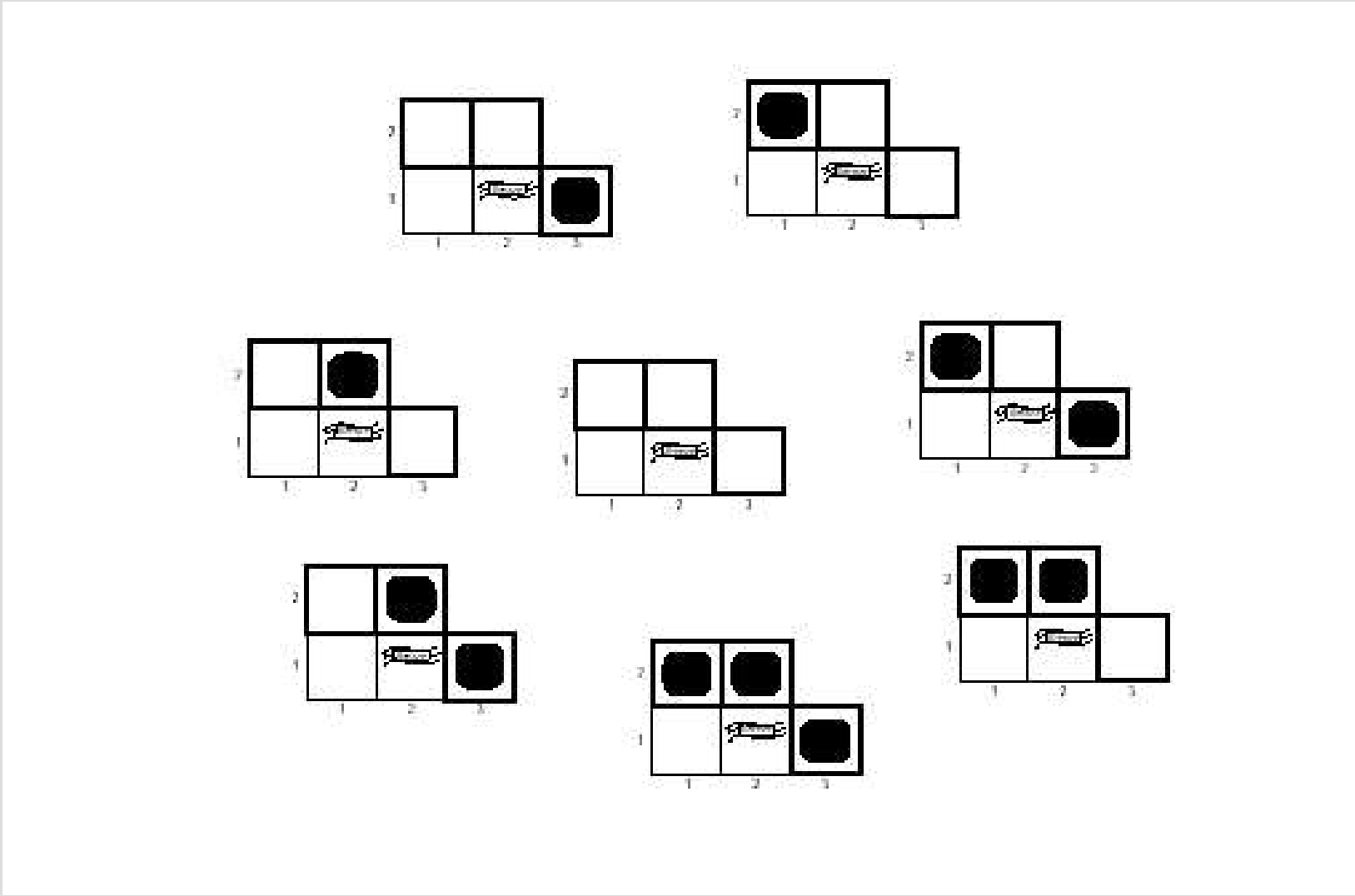


Entailment in the wumpus world

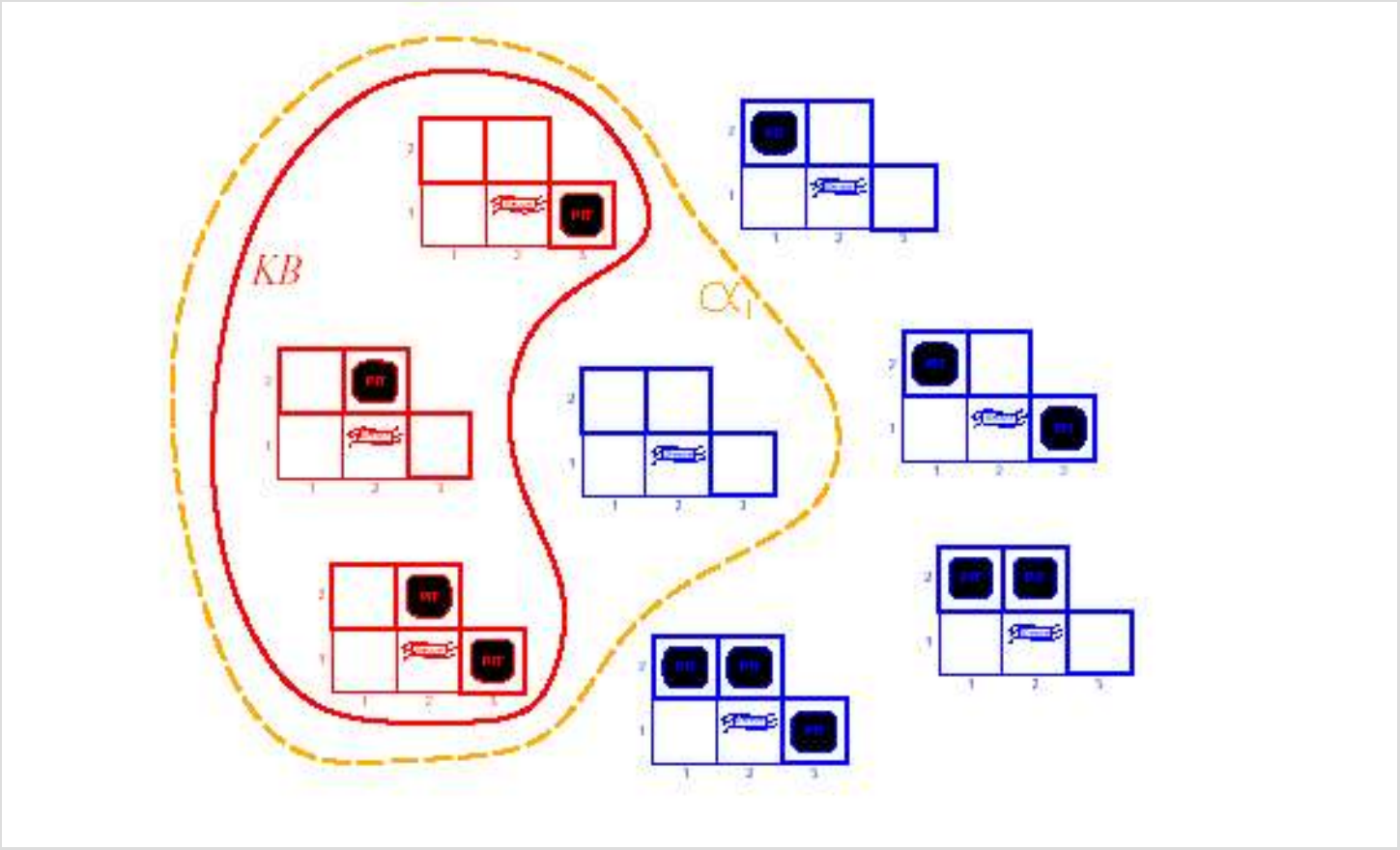


Possible models for ?
assuming only pits =
3 boolean choices
8 possible models

Wumpus Models



Wumpus models



KB = wumpus-world rules + observations
a1 = "[1,2]" is safe, KB |= a1 proved by model checking
What about a2 = "[2,2]" ?

Inference

- KB $|=_i$ a sentence a can be derived by procedure i
- Consequences of KB are haystack, a is needle
  - Entailment: needle in haystack
  - Inference: finding it
- Sound: whenever KB $|=_i$ a it is also true that KB $|=$ a
- Completeness: i is complete if whenever KB $|=$ a it is also true that KB $|=_i$ a