AI 2002
Artificial Intelligence

Dr. Hashim Yasin

Knowledge, Reasoning and Logic

Knowledge

Humans know things …

⇒ **the knowledge helps them to do various tasks.**

⇒ **The knowledge has been achieved**

- not by purely reflex mechanisms
- but by the processes of **reasoning**

- In AI, the example is **knowladge-based agent** which contains **set of sentences** referred as **knowledge-base.**

Knowledge-based Agent

For a generic knowledge-based agent:

- A **percept is given** to the agent.
- The agent **adds the percept** to its knowledge base.
- **Perform best action** according to the knowledge base.
- Tells the knowledge base that it has in fact **taken that action**.

Knowledge-based Agent

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

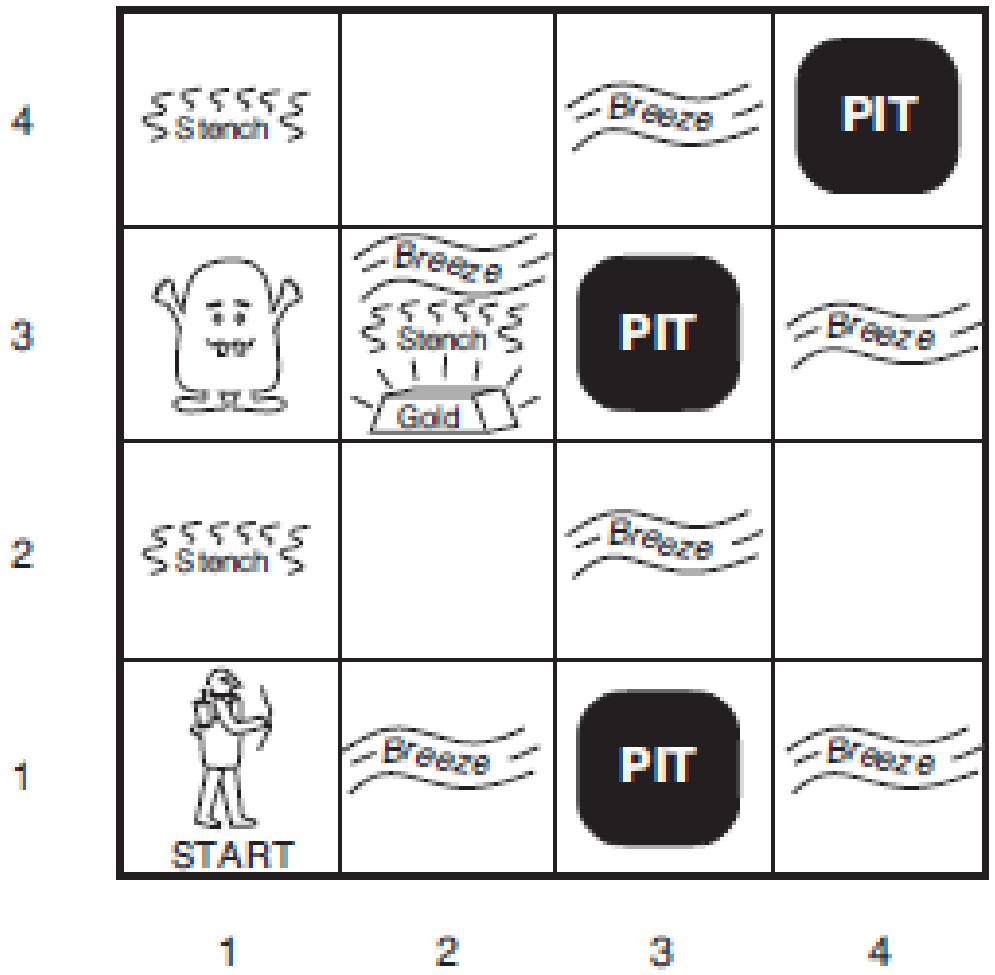
  TELL(KB, MAKE-PERCEPT-SENTENCE(*percept*, t))
  action $\leftarrow$ ASK(KB, MAKE-ACTION-QUERY(t))
  TELL(KB, MAKE-ACTION-SENTENCE(*action*, t))
  $t \leftarrow t + 1$
  return *action*

constructs a **sentence** asserting that the agent ***perceived the given percept*** at time ***t***

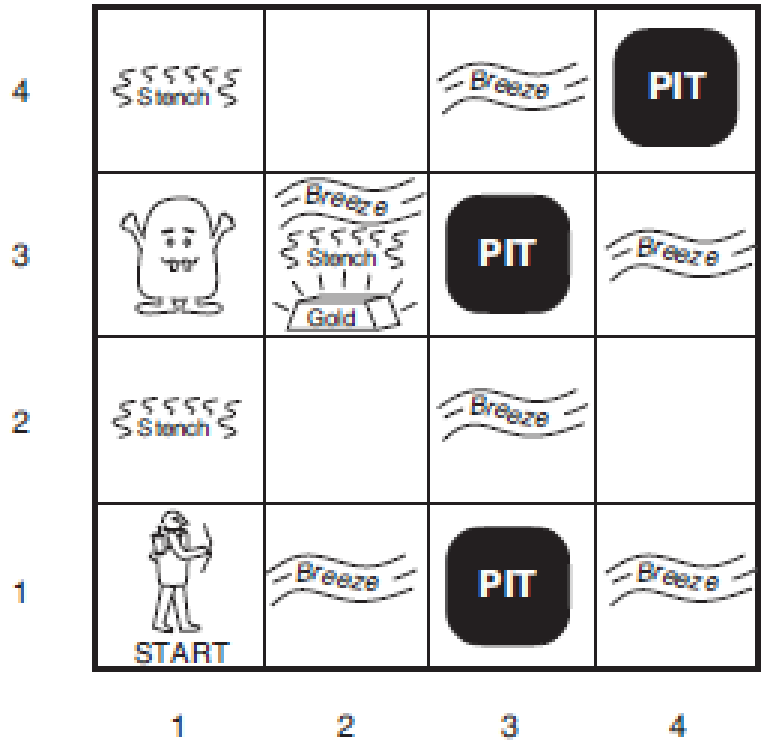
constructs a sentence that asks ***what action should be done*** at time ***t***

constructs a sentence that ***the chosen action was executed*** at time ***t***

The Wumpus World Example

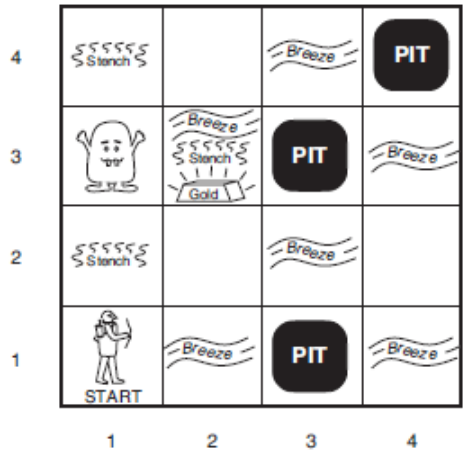


The Wumpus World Example



1,4	2,4	3,4	4,4
1,3	2,3	3,3	4,3
1,2 OK	2,2	3,2	4,2
1,1 A OK	2,1 OK	3,1	4,1

The Wumpus World



The PEAS description for Wumpus World:

Performance measure:

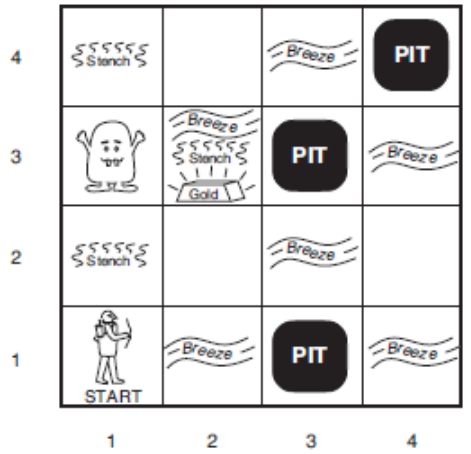
- +1000 for climbing out of the cave with the gold,
- –1000 for falling into a pit or being eaten by the Wumpus,
- –1 for each action taken
- –10 for using up the arrow

Environment:

- A 4×4 grid of rooms. The agent starts in the square labelled [1,1], facing to the right.

The game ends either when the agent dies or when the agent climbs out of the cave.

The Wumpus World



The PEAS description for Wumpus World:

Actuators:

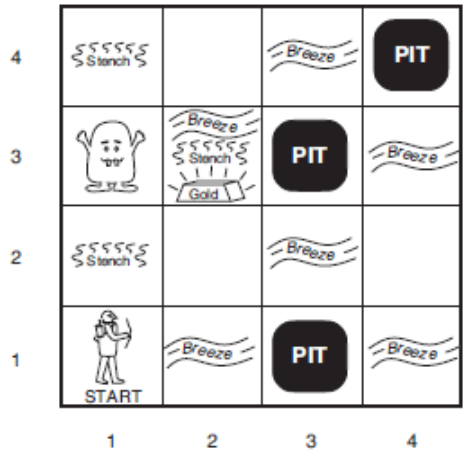
- The agent can move *Forward, TurnLeft by 90◦, TurnRight by 90◦*, grab, shoot

Sensors:

- The square adjacent directly (not diagonally) to the square containing **Wumpus**, the agent will perceive a ***Stench***.
- The squares adjacent to a **pit**, the agent will perceive a ***Breeze***.
- The square with **gold**, the agent will perceive a ***Glitter***.
- An agent **walks into a wall**, it will perceive a ***Bump***.
- When the **Wumpus is killed**, it emits a woeful ***Scream*.**

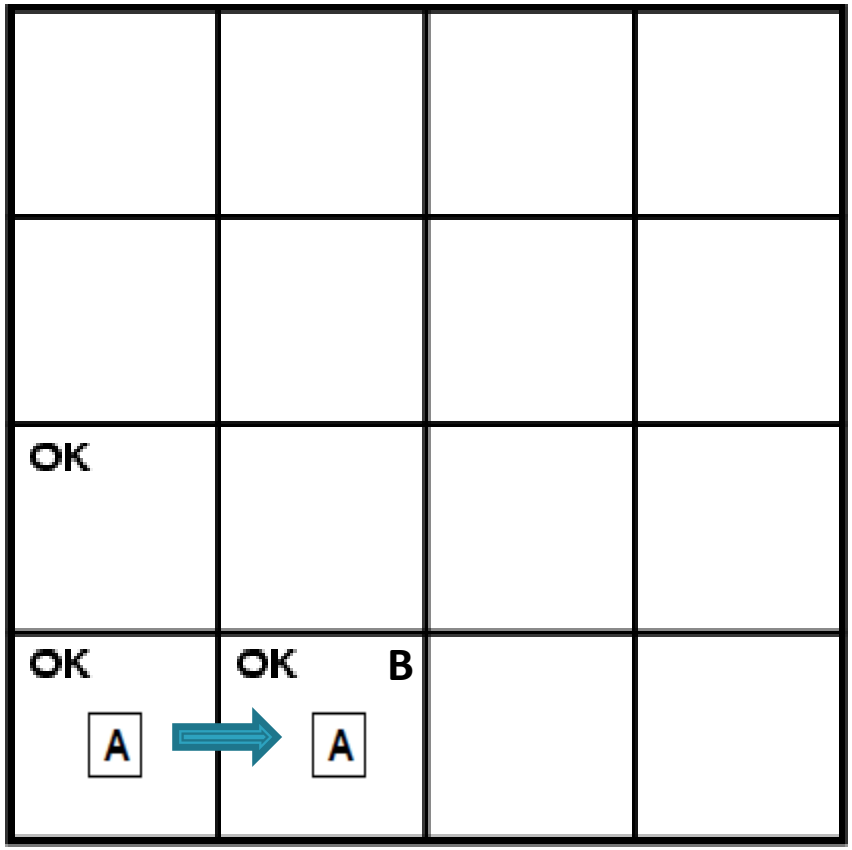
The Wumpus World

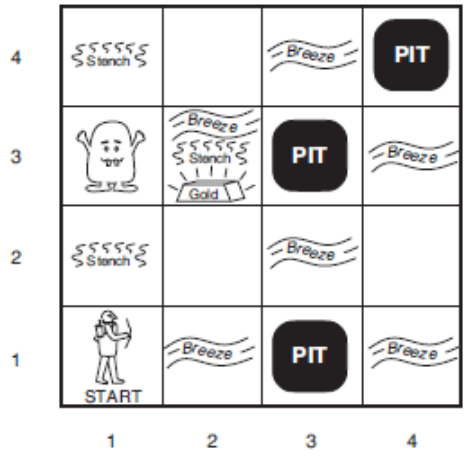
OK			
OK A	OK		



A = Agent
B = Breeze
G = Glitter, Gold
OK = Safe square
P = Pit
S = Stench
V = Visited
W = Wumpus

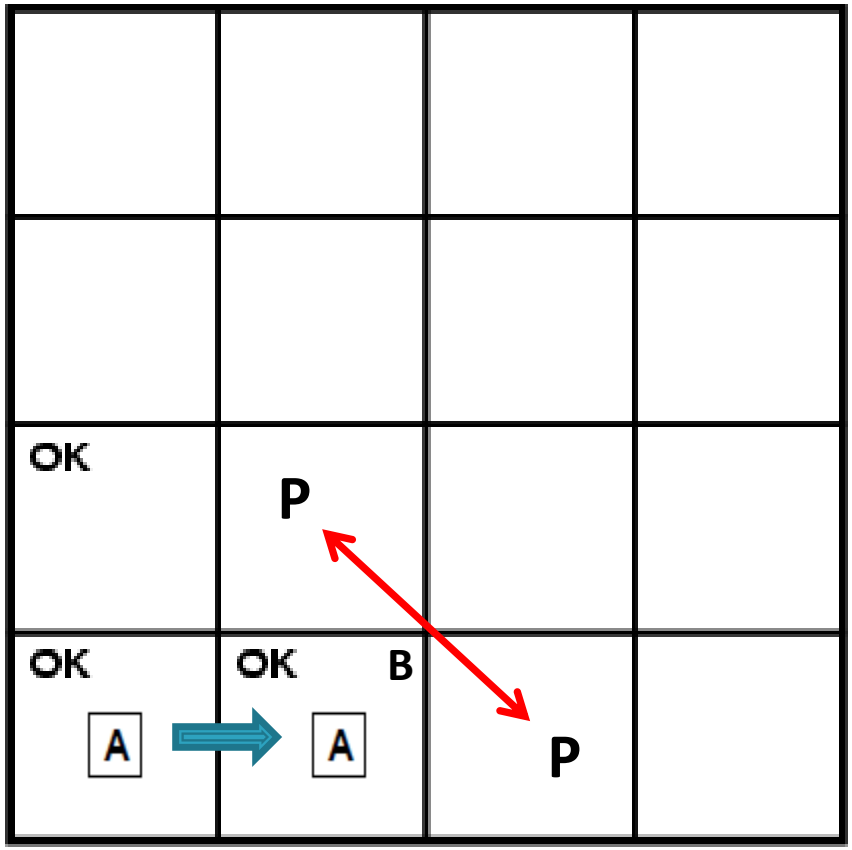
The Wumpus World

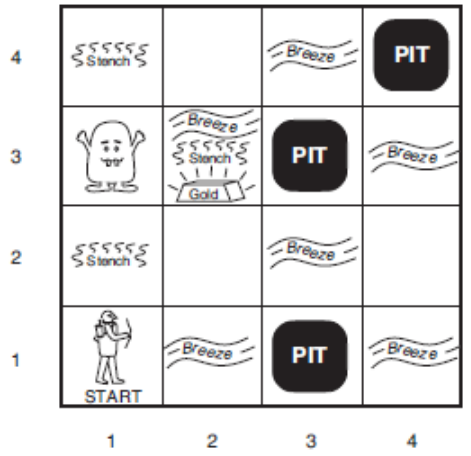




A = Agent
B = Breeze
G = Glitter, Gold
OK = Safe square
P = Pit
S = Stench
V = Visited
W = Wumpus

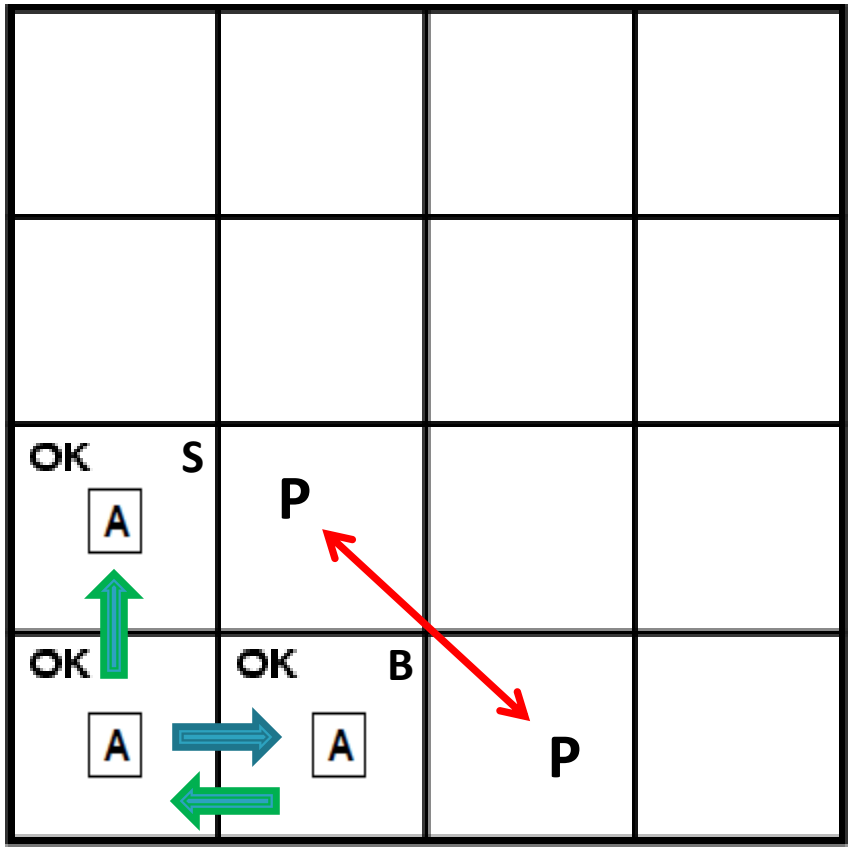
The Wumpus World

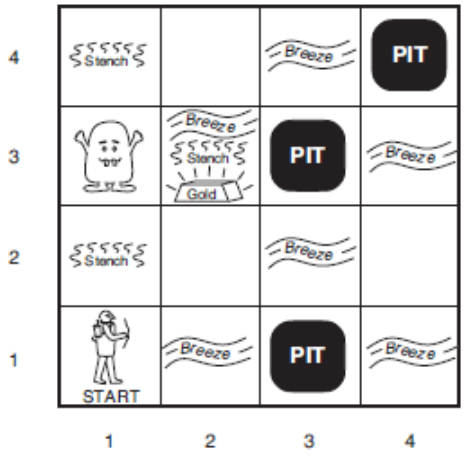




A = *Agent*
B = *Breeze*
G = *Glitter, Gold*
OK = *Safe square*
P = *Pit*
S = *Stench*
V = *Visited*
W = *Wumpus*

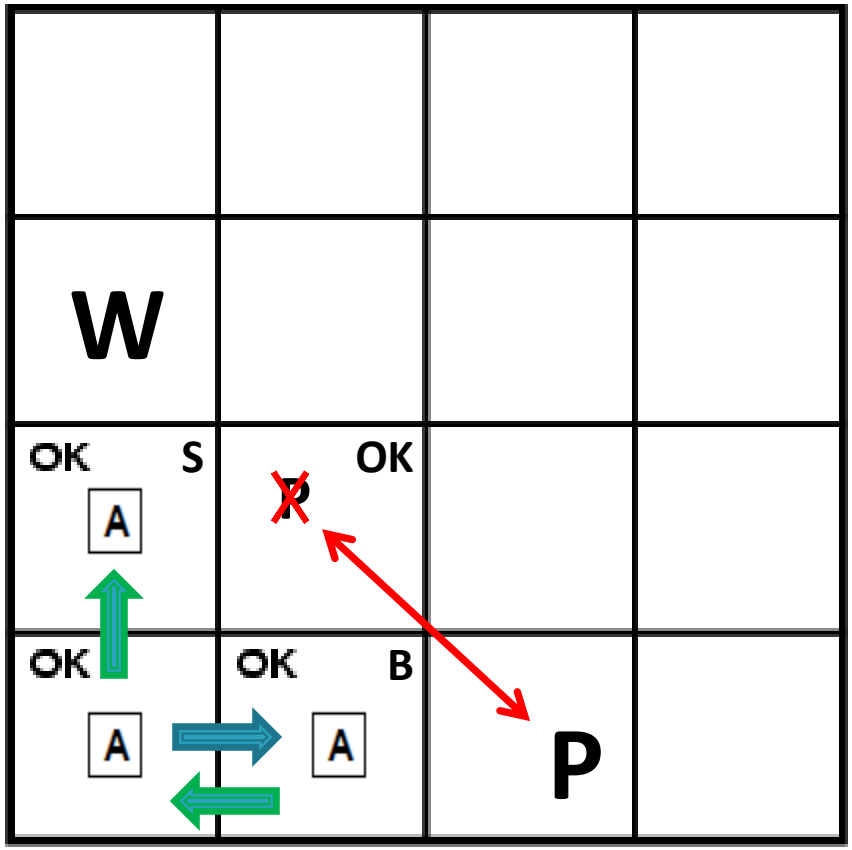
The Wumpus World

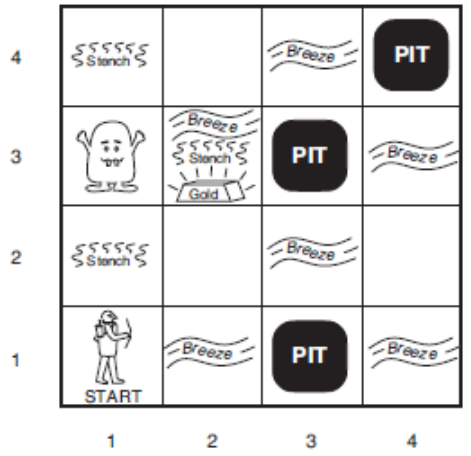




A = *Agent*
B = *Breeze*
G = *Glitter, Gold*
OK = *Safe square*
P = *Pit*
S = *Stench*
V = *Visited*
W = *Wumpus*

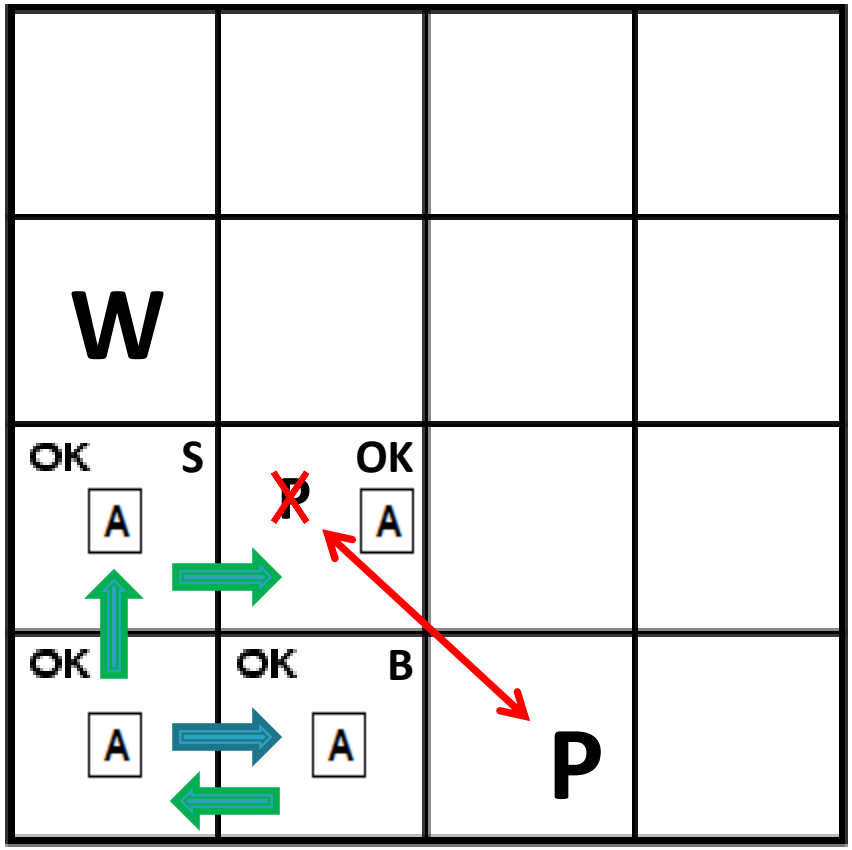
The Wumpus World

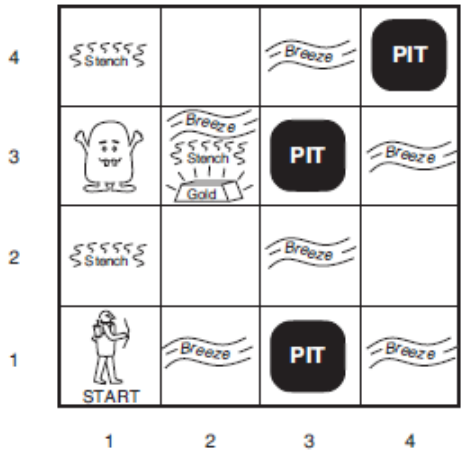




A = Agent
B = Breeze
G = Glitter, Gold
OK = Safe square
P = Pit
S = Stench
V = Visited
W = Wumpus

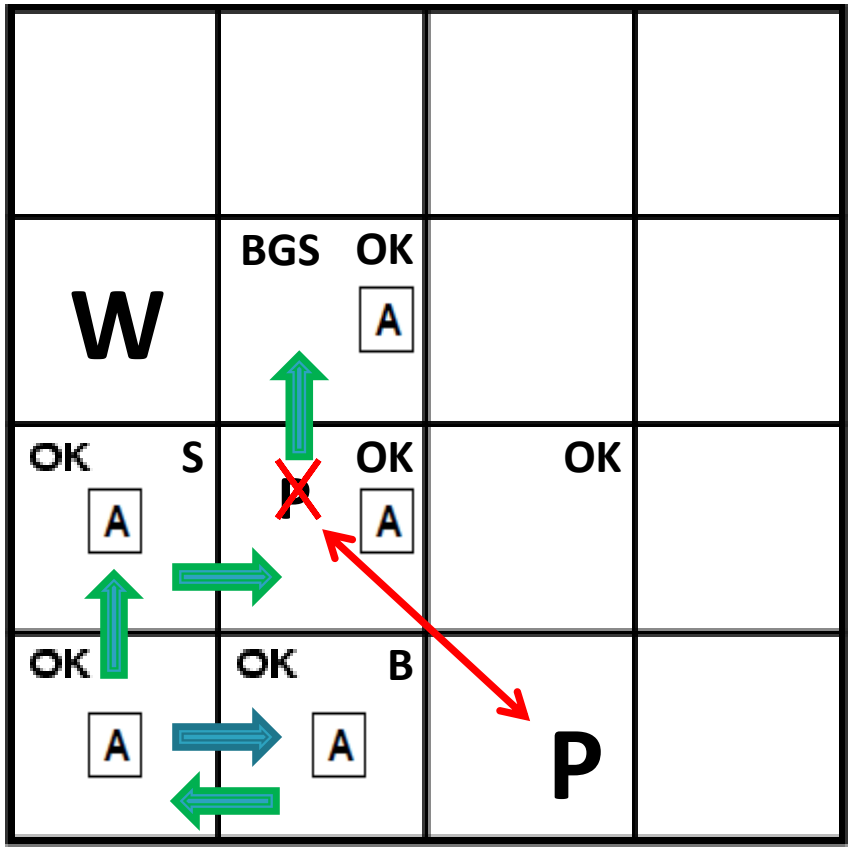
The Wumpus World

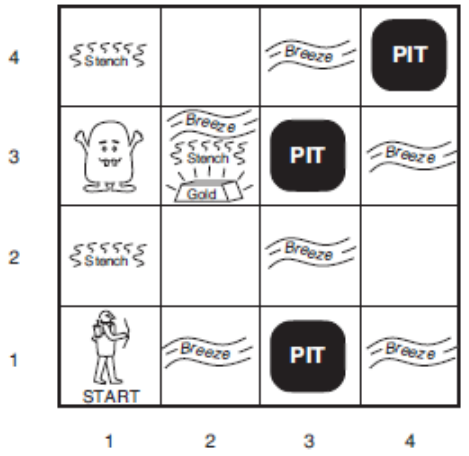




A = Agent
B = Breeze
G = Glitter, Gold
OK = Safe square
P = Pit
S = Stench
V = Visited
W = Wumpus

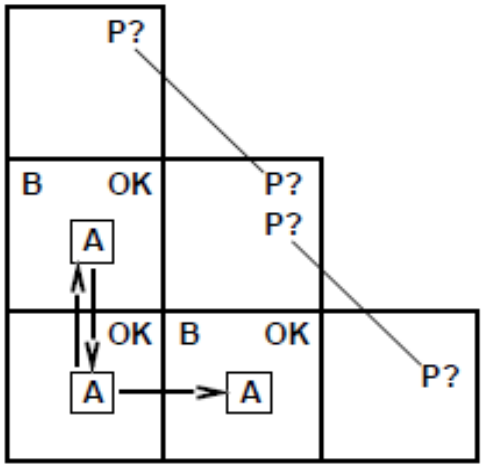
The Wumpus World





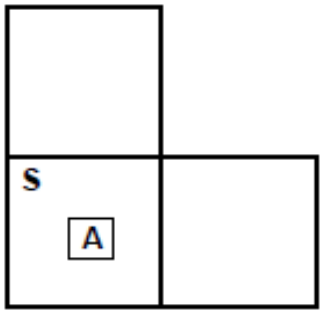
A = Agent
B = Breeze
G = Glitter, Gold
OK = Safe square
P = Pit
S = Stench
V = Visited
W = Wumpus

The Wumpus World



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

Logic

- The knowledge bases consist of sentences.

How to represent these sentances?

- **Logic**, a formal language, is the solution --- a way of manipulating expressions in the language.

- Logic has
  - Syntax
  - Semantics

Logic

Syntax:

What expressions are legal --- what are allowed to write down.

- The notion of syntax is clear enough with the example: "$x + y = 4$" is a well-formed sentence, whereas "$x4y+ =$" is not.

Semantics:

What legal expression means --- meaning of sentences

- the sentence "$x + y = 4$" is **true** in a **world** where x is 2 and y is 2, but **false** in a **world** where x is 1 and y is 1.
- *Syntax is a form and semantics is the content.*

Logic

Semantics:

- The semantics defines the **truth** of each sentence with respect to each **possible world**.
- The term **model** can be used in place of **"possible world."**
- If a sentence α is true in model m, we say that m **satisfies** α or sometimes m is a **model** of α.
- The notation $M(\alpha)$ --- the set of all **model**s of α.

Logic --- Entailment

Entailment:

- means that **one thing follows from another**:

$$\alpha \models \beta$$

- *if and only if, in every model in which* α *is true,* β *is also true*. We can write

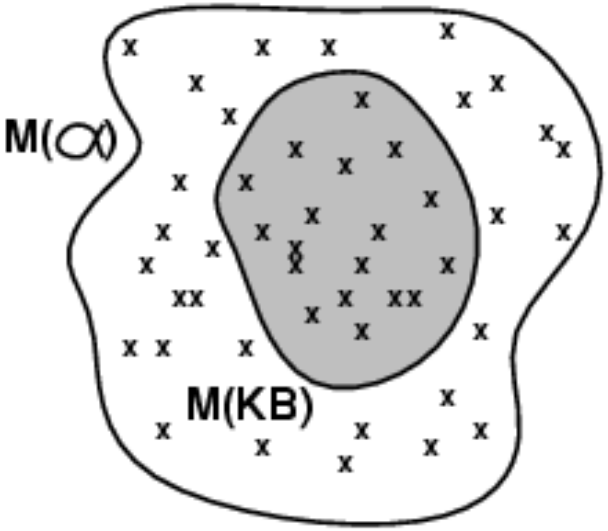
$$\alpha \models \beta \text{ if and only if } M(\alpha) \subseteq M(\beta)$$

- The notation $\subseteq$ means that: if $\alpha \models \beta$, then α is a *stronger assertion than* β

Logic --- Entailment

- We say m is a model of sentence α if α is true in m
- $M(\alpha)$ is the set of all models of α

Then $KB \models \alpha$ iff $M(KB) \subseteq M(\alpha)$

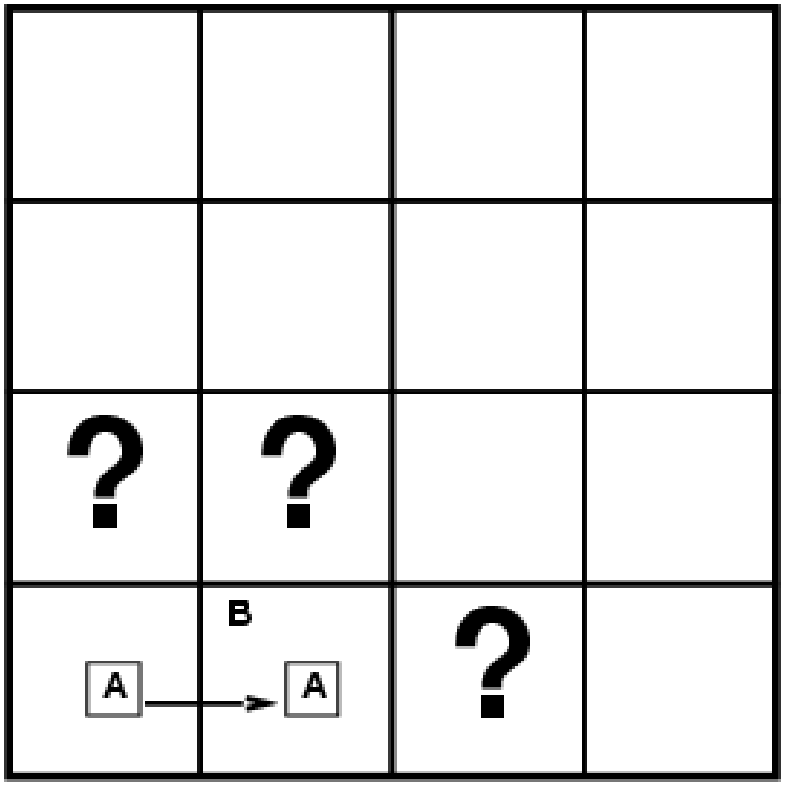


Example:

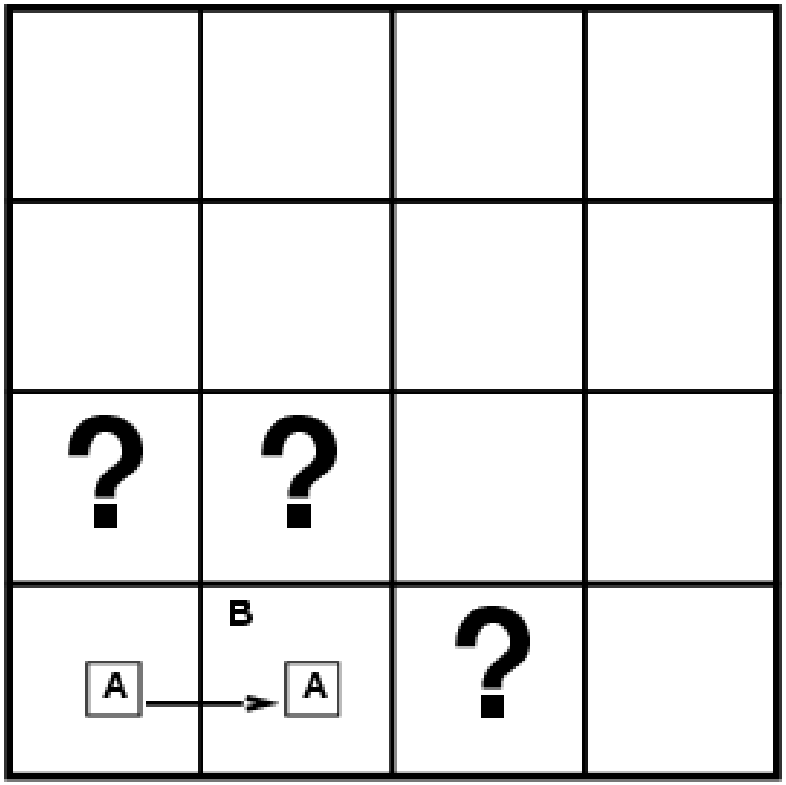
- The sentence $x = 0$ entails the sentence $xy = 0$
  - In any model where x is zero, it is the case that xy is zero (regardless of the value of y)

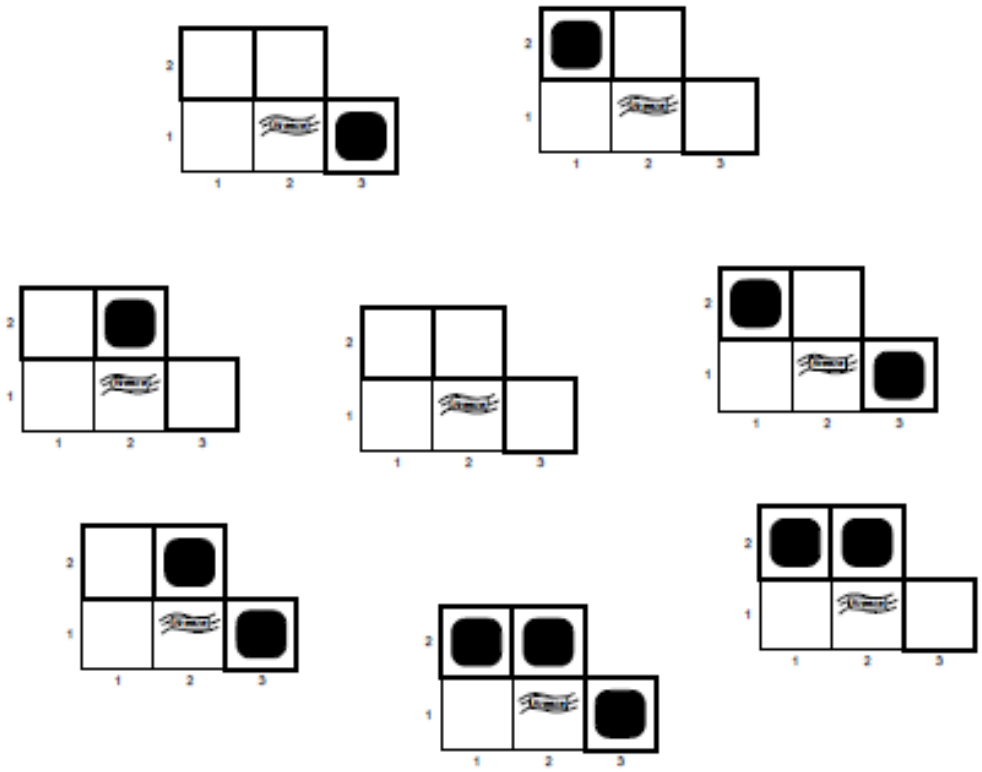
Entailment--- Wumpus World

- Situation after detecting nothing in [1,1], moving right, **breeze** in [1,2]
- Consider possible models for **KB** assuming only pits
- **3 Boolean choices** $\Rightarrow$ 8 possible models



Entailment--- Wumpus World

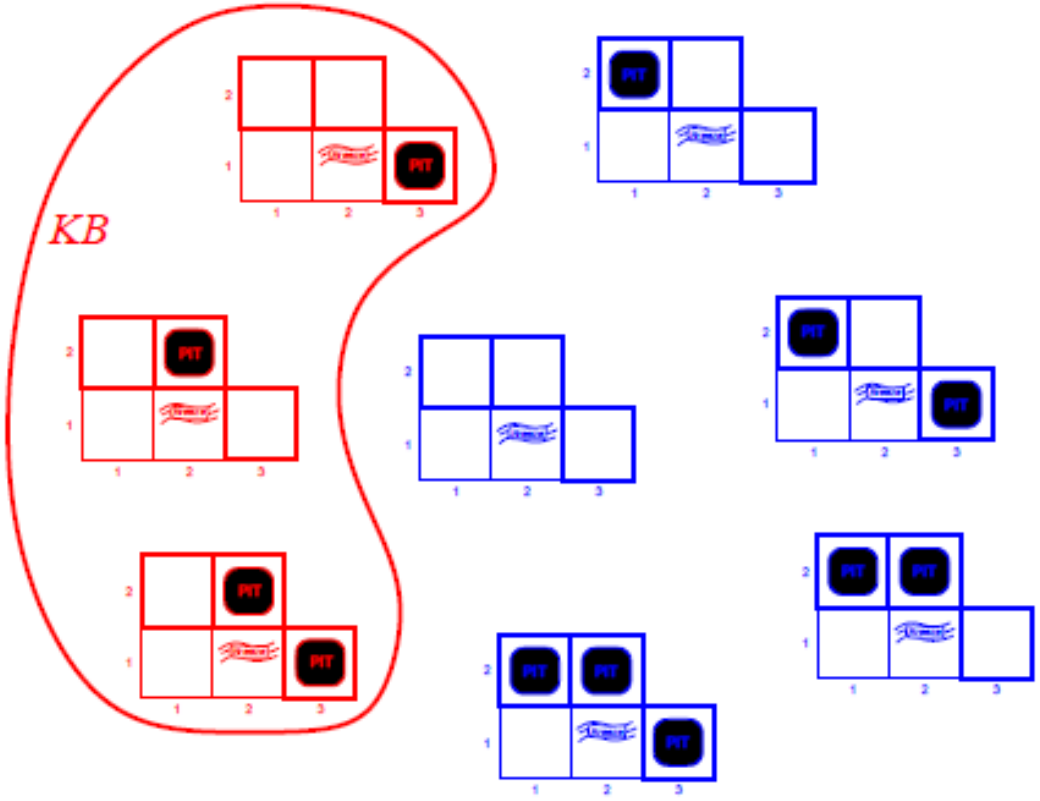




3 Boolean choices ⇒ 8 possible models

regardless of wumpus-world rules

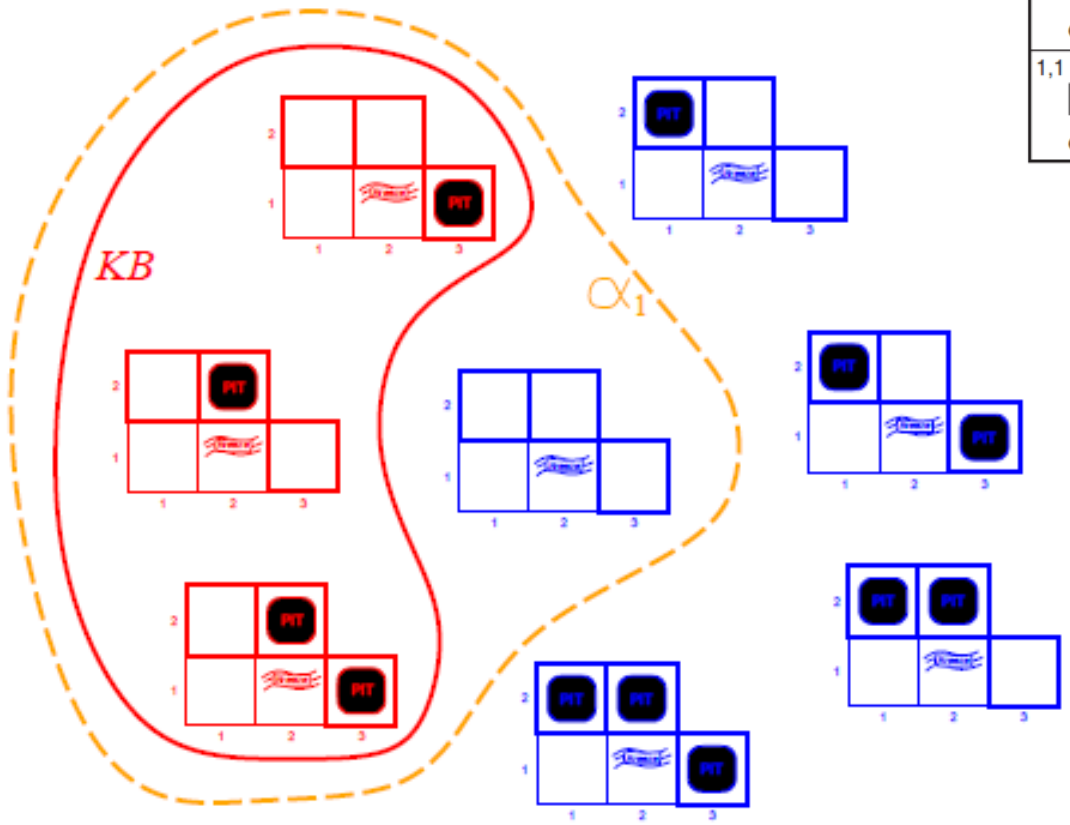
Entailment--- Wumpus World



KB = wumpus-world rules + observations

Entailment--- Wumpus World

1,4	2,4	3,4	4,4
1,3	2,3	3,3	4,3
1,2 OK	2,2	3,2	4,2
1,1 A OK	2,1 OK	3,1	4,1

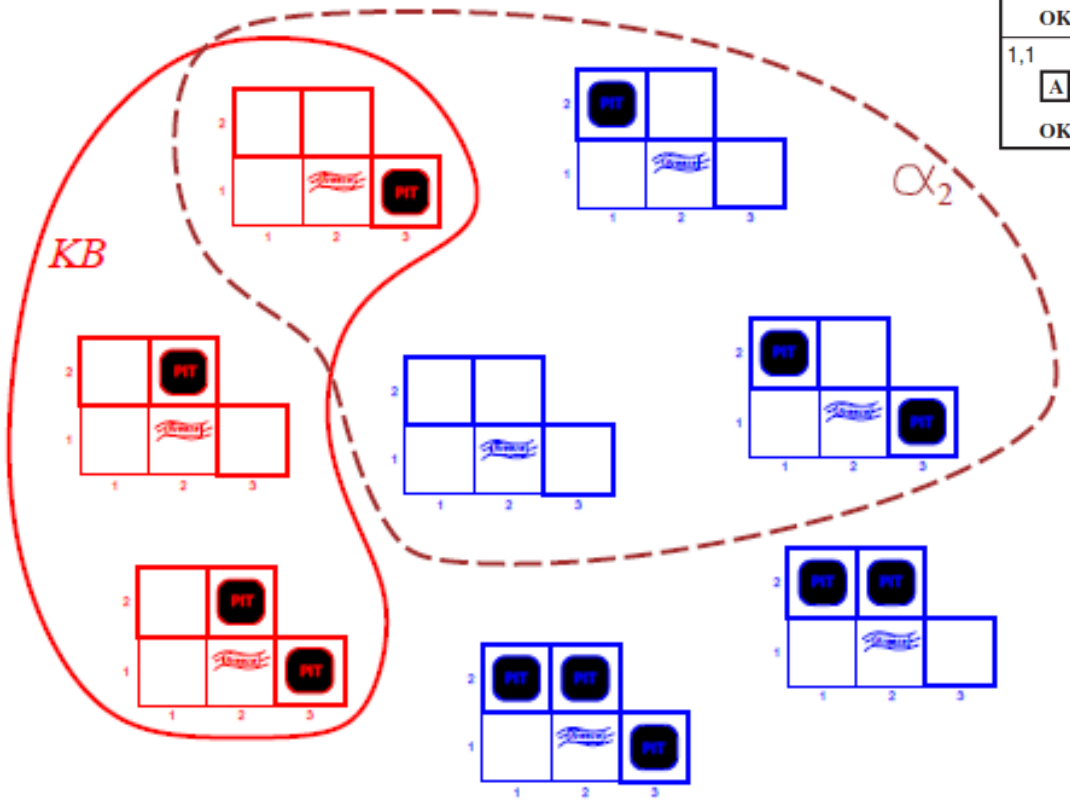


KB = wumpus-world rules + observations

α_1 = "[1,2] is safe", $KB \models \alpha_1$, proved by model checking

Entailment--- Wumpus World

1,4	2,4	3,4	4,4
1,3	2,3	3,3	4,3
1,2 OK	2,2	3,2	4,2
1,1 A OK	2,1 OK	3,1	4,1



KB = wumpus-world rules + observations

α_2 = "[2,2] is safe", $KB \not\models \alpha_2$

Inference

- *If an inference algorithm* $\boldsymbol{i}$ *can derive* $\boldsymbol{\alpha}$ *from* $\boldsymbol{KB}$, we write

$$KB \vdash_i \alpha$$

- which is pronounced "$\boldsymbol{\alpha}$ **is derived from** $\boldsymbol{KB}$ **by** $\boldsymbol{i}$" or "$\boldsymbol{i}$ **derives** $\boldsymbol{\alpha}$ **from** $\boldsymbol{KB}$."

Soundness:

- An inference algorithm that **derives only entailed sentences** is called **sound or truth preserving.**
- Soundness is a highly desirable property.

Completeness:

- An inference algorithm is complete if it can derive any sentence that is **entailed**.

Logic

- **We'll look at two kinds of logic:**

Propositional Logic

which is relatively simple.

First-order Logic

which is more complicated.

Reading Material

- **Artificial Intelligence,** A Modern Approach

  Stuart J. Russell and Peter Norvig

  - **Chapter 7.**