

Title: Logical Agents

AIMA: Chapter 7 (Sections 7.1, 7.2, and 7.3)

Introduction to Artificial Intelligence

CSCE 476-876, Fall 2020

**URL:** [cse.unl.edu/~choueiry/F20-476-876](http://cse.unl.edu/~choueiry/F20-476-876)

Berthe Y. Choueiry (Shu-we-ri)

(402)472-5444

## Outline

- Wumpus world: motivating example
- Knowledge bases
- Logic for Knowledge Representation & Reasoning
  - Syntax
  - Semantics
  - Inference mechanisms: complexity, completeness

Propositional logic/sentential logic

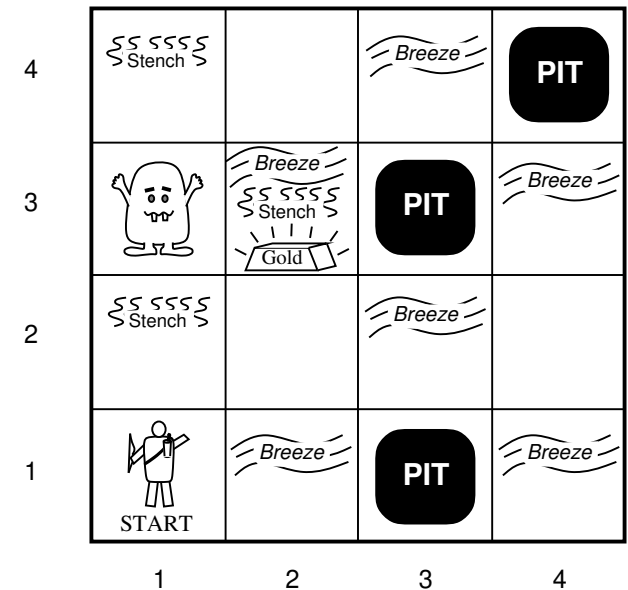
Predicate logic/first-order logic

## Motivating example: The Wumpus world

Early computer game

Agent explores a cave with:

- bottomless pits
- a beast that eats anyone who enters the room, and
- heap of gold to trap



## **PEAS description** of the Wumpus world

**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:** Breeze, Glitter, Smell

**Actuators:** Left turn, Right turn, Forward, Grab, Release, Shoot

## Wumpus World: Characterization

Is the world:

- Observable? No, only local perception
- Deterministic? Yes, outcome exactly specified
- Episodic? No, sequential at the level of actions
- Static? Yes, Wumpus/Pits don't move
- Discrete? Yes
- Single-agent? Yes, Wumpus considered a natural feature

## Empirical evaluations: single/multiple configuration

An agent can do well in a single environment: learns the environment, executes rules.

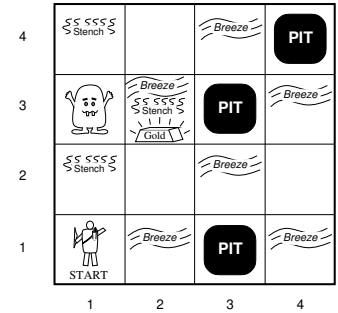
Agent must be tested in a complete class of environments and its average performance must be determined → empirical experiments

- Constraints: start from position [1,1], limited to 4×4 grid
- Location of Wumpus and Gold chosen randomly with a uniform distribution (all squares are possible except [1,1])
- Each square, except [1,1], can be a pit with probability 0.2
- Terribly bad cases: gold in a pit or surrounded by pits

## Wumpus World: Acting & Reasoning

- After receiving initial percepts, agent knows it is in [1,1] and it is OK
- No stench or breeze in [1,1]  $\Rightarrow$  [1,2] and [2,1] are danger-free
- Cautious agent moves only to square it knows it is OK
- Agent moves only to square [2,1], detects breeze  $y \Rightarrow \exists$  a pit in neighboring squares [1,1], [2,2] and [3,1]. Agent knows no pit in [1,1]  $\rightarrow$  Pit indicated in [2,2] and [3,1] with P?
- Not visited OK squares? Only [1,2]. Agent goes to [1,1], proceeds to [1,2]

# Wumpus World: Acting & Reasoning



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

(a)

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

1,4	2,4	3,4	4,4
1,3	2,3	3,3	4,3
1,2	2,2	3,2	4,2
OK	P?		
1,1	2,1	3,1	4,1
V	A	P?	
OK	B		
	OK		

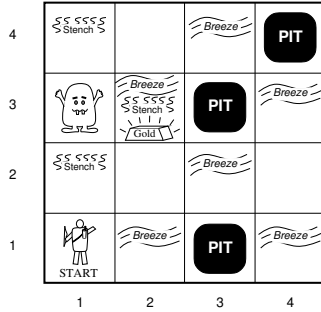
(b)



## Wumpus World: Acting & Reasoning

- Agents detects stench in [1,2]  $\Rightarrow$  Wumpus nearby!  
Possibilities: [1,1], [1,3] or [2,2].  
Agent knows [1,1] is Wumpus-free (Agent was there!)  
Agent can infer [2,2] is Wumpus-free ( $\nexists$  stench in [2,1])  
Agent infers Wumpus is in [1,3] (W!)
- Lack of breeze in [1,2]  $\Rightarrow$  [2,2] is pit-free  
But,  $\exists$  a pit in either [2,2] or [3,1]  $\Rightarrow$   $\exists$  pit in [3,1] (P!)  
Inference combines knowledge gained at different times and places, beyond the abilities of most animals, but Logical Inference can handle this
- Since [2,2] is OK and not visited, Agent moves there
- etc.

# Wumpus World: Acting & Reasoning



1,4	2,4	3,4	4,4
1,3 W!	2,3	3,3	4,3
1,2 <b>A</b> S OK	2,2 OK	3,2	4,2
1,1 V OK	2,1 B V OK	3,1 P!	4,1

(a)

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

1,4	2,4 P?	3,4	4,4
1,3 W!	2,3 <b>A</b> S G B	3,3 P?	4,3
1,2 S V OK	2,2 V OK	3,2	4,2
1,1 V OK	2,1 B V OK	3,1 P!	4,1

(b)

## The point of the Wumpus world

In each case where the agent draws a conclusion from the available information, that conclusion is guaranteed to be correct if the available information is correct.

→ Fundamental property of logical reasoning.

## Knowledge Base

**A fact in the world:** A representation of a fact in the world  
A sentence = a representation of a fact in the world in a  
formal language

**A Knowledge Based (KB):** A set sentences  
A set (of representations) of facts about the world

**Issues:** Access to KB, Representation (language), Reasoning  
(inference)

## Level of Knowledge

Agents can be viewed at various levels:

1. **Epistemological:**

Abstract description of what the agent knows about the world

2. **Logical:**

Encoding of knowledge into sentences

3. **Implementation:**

Actual implementation (lists, arrays, hash tables, etc.)

- Very important for performance of agent
- Irrelevant for higher levels of knowledge

## A simple 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*

14

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

# 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*

15

**Perceives:** Tells KB about new percepts (new sentences)

Representation: MAKE-PERCEPT-SENTENCE

**Access to KB:** Asks KB about actions to take (inference)

Two primitives: ASK and TELL hide reasoning details

**Acts:** Tells KB about actions (new sentences)

Representation: MAKE-ACTION-SENTENCE,  
MAKE-ACTION-QUERY

## Logic in general

**Logics** are formal languages for representing information such that conclusions can be drawn

**Syntax** defines the sentences in the language (grammar)

**Semantics** define the “meaning” of sentences; *i.e.*, define truth of a sentence in a world

Example: the language of arithmetic

- Syntax:  $x + 2 \geq y$  is a sentence;  $x^2 + y >$  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  $x = 0$ ,  $y = 6$



# Types of logic

Logics are characterized by what they commit to as “primitives”

## Ontological commitment :

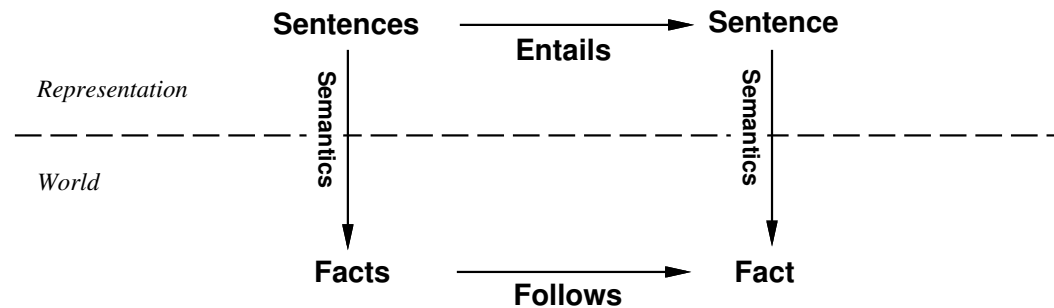
what exists—facts? objects? time? beliefs?

## Epistemological commitment :

what states of knowledge?

Language	Ontological Commitment (What exists in the world)	Epistemological Commitment (What an agent believes about facts)
Propositional logic	facts	true/false/unknown
First-order logic	facts, objects, relations	true/false/unknown
Temporal logic	facts, objects, relations, times	true/false/unknown
Probability theory	facts	degree of belief 0...1
Fuzzy logic	degree of truth	degree of belief 0...1

# Knowledge representation & reasoning



**Facts:** in the world

**Representations:** in the computer

**Reasoning:** process of constructing new representations from old ones

**Proper Reasoning:** ensures new representations correspond to facts that actually follow from facts in the world

## Entailment

Entailment means that one thing follows from another:

$$(\text{KB} \models \alpha)$$

Knowledge base KB entails sentence  $\alpha$

iff  $\alpha$  is true in all worlds where KB is true

Example: KB:  $\{a \wedge b\}$ , then

$$\text{KB} \models a; \text{KB} \models b; \text{KB} \models a \vee b$$

Entailment is a relationship between sentences (i.e., syntax)  
that is based on semantics

$(\alpha \models \beta)$ : the truth of  $\beta$  contains the truth of  $\alpha$

For example:  $(x + y = 4) \models (4 = x + y)$ ,  $(x + y = 4) \models (4 \geq x + y)$ ,  
 $(x + y \geq 4) \not\models (4 = x + y)$ ,

## Models

Logicians typically think in terms of **models**, which are formally structured worlds with respect to which truth can be evaluated

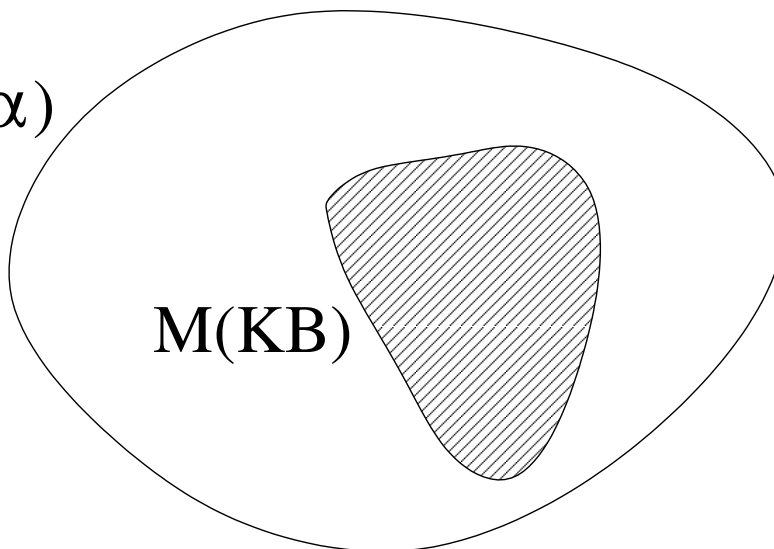
We say  $m$  is a **model** of a sentence  $\alpha$  if  $\alpha$  is true in  $m$

$M(\alpha)$  is the set of all models of  $\alpha$

Then  $KB \models \alpha$  if and only if  $M(KB) \subseteq M(\alpha)$

$M(\alpha)$

$M(KB)$



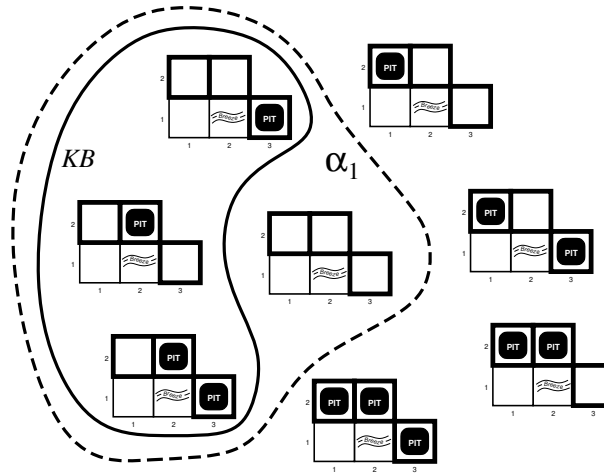
## Entailment in the Wumpus world

Situation: Agent detected nothing in [1,1], breeze in [2,1]  
 $2^3=8$  possible models

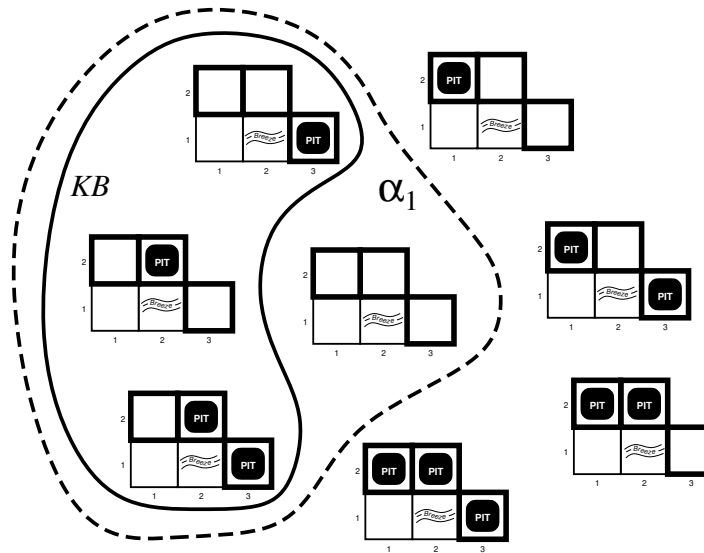
Percepts + the PEAS description = KB

Agent wonders whether pit is in [1,2], [2,2], and [3,1]:  
 Only 3 models where the KB is true

$\alpha_1$  = no pit in [1,2]:  
 $\alpha_1$  is true in 4 models.



## Entailment in the Wumpus world



Consider:  $\alpha_1 = \text{no pit in } [1,2]$ ,  $\alpha_2 = \text{no pit in } [2,2]$

Model checking:  $\text{KB} \models \alpha_1$ ,  $\text{KB} \not\models \alpha_2$

Given KB, agent cannot conclude whether  $\alpha_2$  holds or not

Entailment can be used to derive conclusions: Inference

Inference here is done by model checking

# Inference

$\text{KB} \vdash_i \alpha \equiv \alpha$  is derived from KB by procedure  $i$

Consequences of KB are a haystack;  $\alpha$  is a needle.

Entailment = needle in haystack; inference = finding it

**Soundness:**  $i$  is sound if

whenever  $\text{KB} \vdash_i \alpha$ , it is also true that  $\text{KB} \models \alpha$

**Completeness:**  $i$  is complete if

whenever  $\text{KB} \models \alpha$ , it is also true that  $\text{KB} \vdash_i \alpha$

That is, the procedure will answer any question whose answer follows from what is known by the KB

The record of operation of a sound inference procedure is a **proof**

Next, propositional logic: syntax, semantics, and inference