Title: Agents that Reason Logically

AIMA: Chapter 6 (Sections 6.1 and 6.2)

Introduction to Artificial Intelligence

CSCE 476-876, Spring 2002

URL: `www.cse.unl.edu/~choueiry/S02-476-876`

Berthe Y. Choueiry (Shu-we-ri)

`choueiry@cse.unl.edu`, `(402)472-5444`

Glossary 7:

Assigned: Friday, Mar 28, 2000

Due: Monday, April 1, 2002

- Background knowledge
- Completeness (of inference)
- Compositional (languages)
- Epistemological level
- Entailment
- Inference
- Knowledge base
- Knowledge level
- Proof
- Semantics (of a logic)
- Sentence
- Soundness (of inference)
- Syntax (of a logic)
- Validity (of a logical sentence)

Outline

- Knowledge bases
- Wumpus world
- Logic in general

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 (of representations) of facts about the world

Access to KB :

Two primitives: (1) ASK and (2) TELL

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

On Knowledge

Declarative knowledge :
The designer tells the agent what it should know about the world, fact by fact

Learning mechanisms :
The agent can generate general knowledge about the world given a series of percepts

Learning in a KB-Agent :
→ allow building fully autonomous agents

Knowledge-Based Agent (I)

Two components:

1. **Knowledge base**:
   Stores the representations of facts of the world
   May contain some background knowledge

2. **Inference mechanism**:
   Generates new facts from the ones stored in KB

Knowledge-Based Agent (II)

Perceives
→ Tells KB about new percepts (new sentences)

Reasons
→ Asks KB about actions to take (inference)

Acts
→ Tells KB about actions (new sentences)

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
```

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

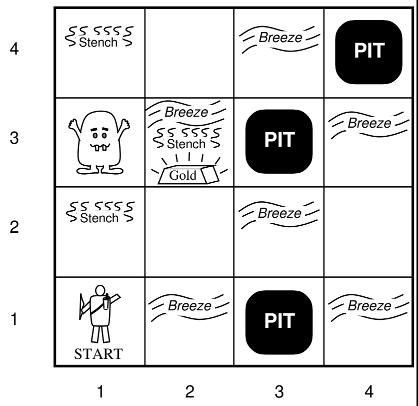
deduce appropriate actions

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

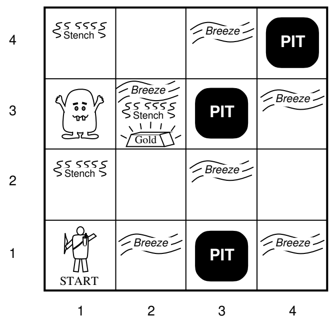


Wumpus World: PAGE description

Percepts : Breeze, Glitter, Smell, Bump, Wumpus' scream

Actions Left turn, Right turn, Forward, Grab, Release, Shoot

Goals Get gold back to start, without entering pit or wumpus square



Environment

Squares adjacent to wumpus are smelly
Squares adjacent to pit are breezy
Glitter if and only if gold is in the same square
Shooting kills the wumpus if you are facing it
Shooting uses up the only arrow
Grabbing picks up the gold if in the same square
Releasing drops the gold in the same square

Wumpus World: Characterization

Is the world:

- Deterministic? Yes, outcome exactly specified
- Fully accessible? No, only local perception
- Static? Yes, Wumpus/Pits don't move
- Discrete? Yes

Wumpus World: Notes

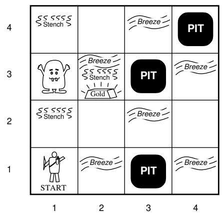
An agent can do well in a single environment: learns the environment, executes rules.

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

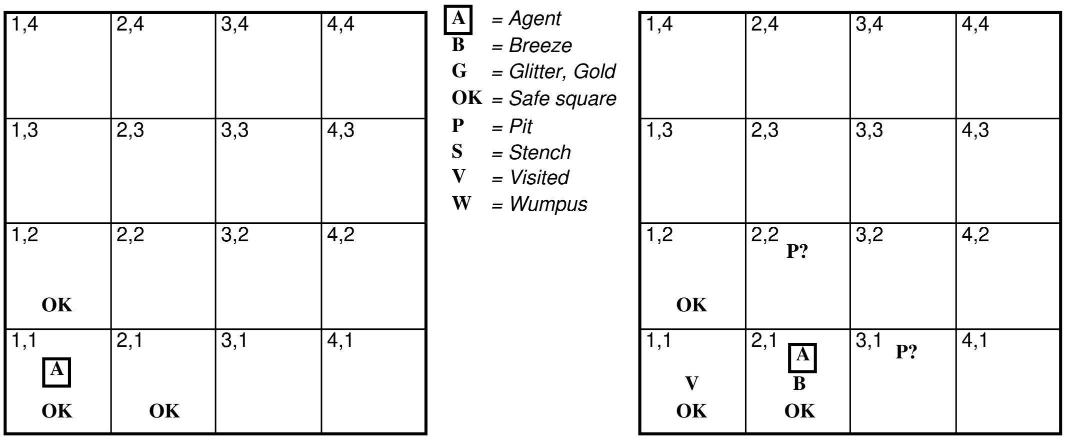
- Constraints: (A) start from position [1, 1], (B) limited to 4×4 grid
- Location of Wumpus and Gold chosen randomly with a uniform distribution (all squares are possible except [1, 1])
- Each square, except [1, 1], can be a pit with probability 0.2
- Terribly bad cases: gold in a pit or surrounded by 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] ⇒ [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 ⇒ ∃ a pit in neighboring squares [1, 1], [2, 2] and [3, 1]. Agent knows no pit in [1, 1] → 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]



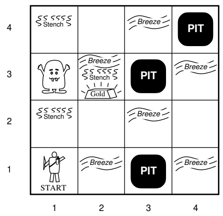
Wumpus World: Acting & Reasoning



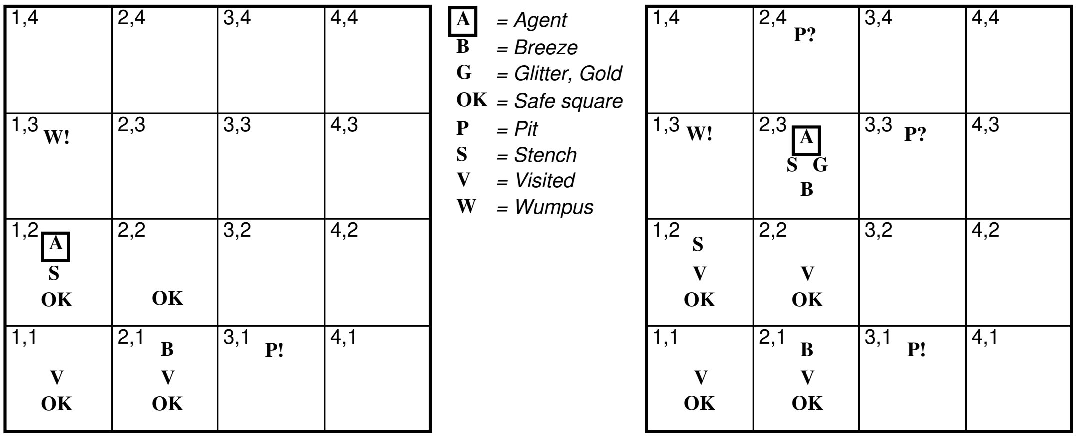
(a) (b)

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



(a) (b)

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 <u>truth</u> of a sentence in a world

Examples, the language of arithmetic:

- Syntax: $x + 2 \geq y$ is a sentence; $x2 + y >$ is not a sentence
- Semantics:
  – True: $x + 2 \geq y$ is true iff the number $x + 2$ is no less than the number y
  – True: $x + 2 \geq y$ is true in a world where $x = 7,\ y = 1$
  – False: $x + 2 \geq y$ 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

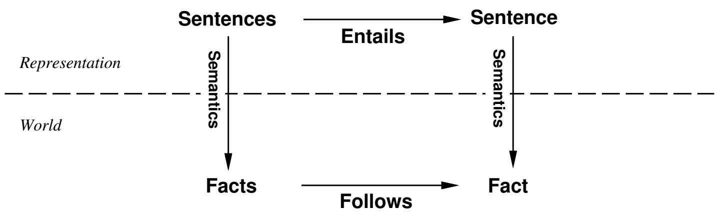
Entailment

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



Knowledge base KB <u>entails</u> sentence α (KB $\models \alpha$)
iff α is true in all worlds where KB is true

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

Inference

A inference proc. can be described by the sentences it can derive

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

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 $\vdash_i \alpha$
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**

(to be cont')