Tufts

Lecture 08:
Propositional Logic

Artificial Intelligence (CS 131)

1

Knowledge-Based Agents

Image source: Russell & Norvig (2021)

- We can build AI agents that act in the world based upon incoming data, supplemented by their knowledge of the world

Actions are chosen based upon *KB*, a knowledge-base

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

Incoming data (percepts) and chosen actions are also recorded to that knowledge-base

2

Using Logic for Knowledge Bases

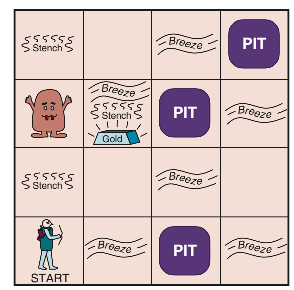
- A logic is a formal language:
  - Represents facts that we know about the world: each such fact is a sentence in our logical language
  - Allows us to generate *new* facts based on what we already know by performing inference
- A logic has two basic components:
  - Syntax: defines what *counts* as a sentence of the language
  - Semantics: defines what makes a sentence *true* or *false*

3

The Wumpus World (PEAS)

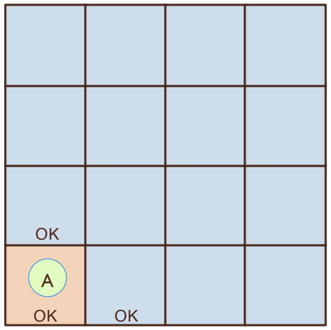
Image source: Russell & Norvig (2021)

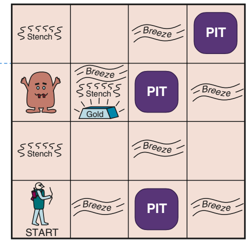
- Performance:
  - gold = +1000; death = –1000
  - step = –1; use arrow = –10
- Environment:
  - Square next to Wumpus: *Smelly*
  - Square next to Pit: *Breezy*
  - Gold squares: *Glitter*
  - *Shooting* Wumpus kills it (if facing it)
  - *Shooting* Wumpus uses up the only arrow
  - Agent can *grab* gold to pick it up
- Actuators:
  - Forward, TurnLeft, TurnRight, Grab, Shoot, Climb
- Sensors/Percepts:
  - [Stench, Breeze, Glitter, Bump, Scream]



4

Exploring the Wumpus World





(A) : Agent

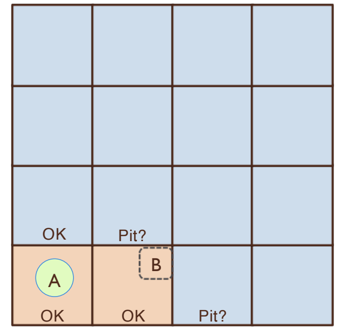
- Initial percepts:

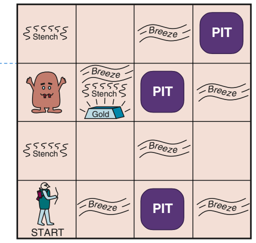
$$[None, None, None, None, None]$$

- Agent can infer that current and adjacent squares are all empty (and OK to enter)

5

Exploring the Wumpus World





(A) : Agent

[B] : Breeze

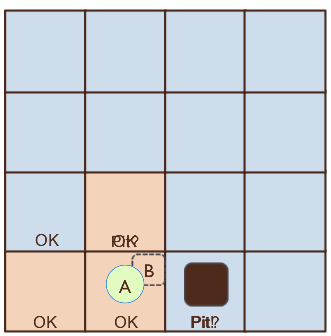
- On moving right, the agent receives percepts:

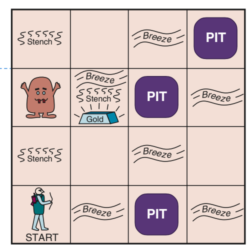
$$[None, Breeze, None, None, None]$$

- Agent can now infer that a Pit is adjacent, although it does not know where

6

Exploring the Wumpus World





(A) : Agent

[B] : Breeze

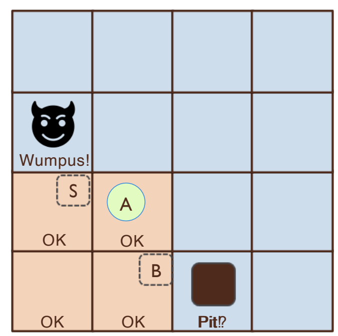
- On moving up, the agent again receives percepts:

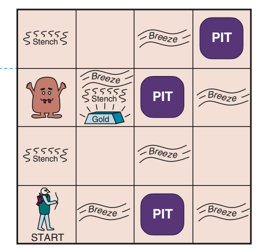
$$[None, None, None, None, None]$$

- Agent can now infer that there must be a Pit at bottom right

7

Exploring the Wumpus World





(A) : Agent

[B] : Breeze [S] : Stench

- Finally, moving left, the agent receives:

$$[Stench, None, None, None, None]$$

- Agent can now infer that there must be a Wumpus above current location

8

Inference & Logical Entailment

- An inference rule tells an agent how to draw conclusions based upon what they already know
- Valid rules encode a semantic relationship between what we infer and and what is already in our knowledge base

- We say that a sentence α follows from our knowledge base, *KB*, or that *KB* entails α:

$$KB \models \alpha \;\Leftrightarrow\; \alpha \text{ is true in every situation in which } KB \text{ is true}$$

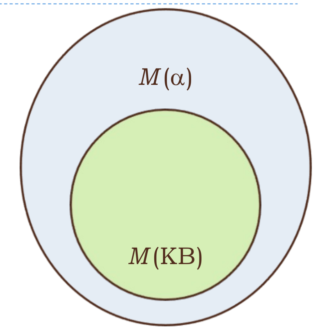
- For example:

$$\{\text{Breeze at } [2,1]\} \models (\text{Pit at } [3,1]) \text{ OR } (\text{Pit at } [2,2])$$

9

Semantic Models

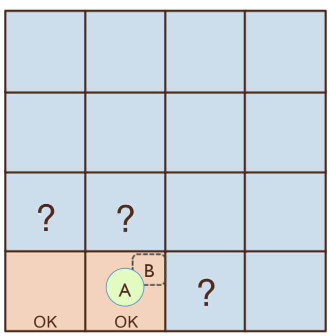
- A formal semantics defines truth and entailment in terms of a logical model
  - Circumstances under which a sentence is true or false
  - Model m is a *model of* sentence α if α is true in m
  - $M(\alpha)$ = *set of all* models of α
- KB entails α whenever *all* models of KB are *also* models of α



$$KB \models \alpha \;\Leftrightarrow\; M(KB) \subseteq M(\alpha)$$

10

Models for the Wumpus World



A : Agent

B : Breeze

- The agent starts at bottom left, sensing *nothing* adjacent to it:

$[None, None, None, None, None]$

- On moving right, the agent senses a breeze:

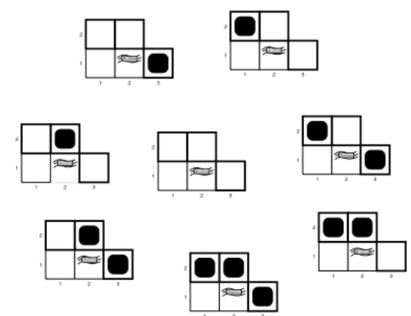
$[None, Breeze, None, None, None]$

- Based on this sequence of percepts, there exist a total of 8 possible models representing the presence or absence of pits in each of the 3 adjacent squares

11

Models for the Wumpus World

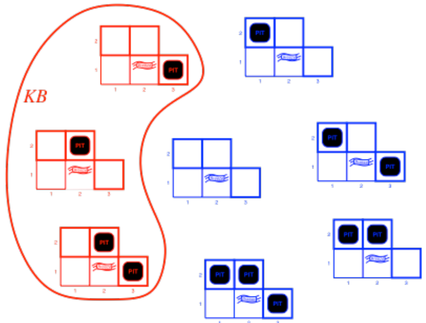
Image source: Russell & Norvig (2021)



- With 8 basic models, our knowledge base will consist of all of those that are consistent with our observation of the breeze, given the rules defining the Wumpus World environment

12

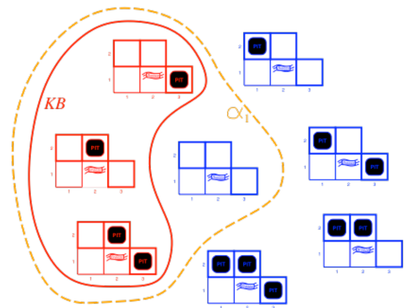
Models for the Wumpus World



- With 8 basic models, our knowledge base will consist of all of those that are consistent with our observation of the breeze, given the rules defining the Wumpus World environment

13

Models for the Wumpus World

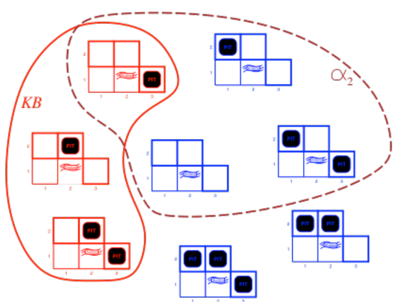


- α_1 = "[1,2] is safe" is entailed by our knowledge base, due to the containment relationship between the models

$$\boxed{M(KB) \subseteq M(\alpha_1)}$$

14

Models for the Wumpus World



- α_2 = "[2,2] is safe" is *not* entailed, however, since there are conditions in which KB is true, but α_2 is still false

$$\boxed{M(KB) \not\subseteq M(\alpha_2)}$$

15

Syntax of Propositional Logic (PL)

- A basic formal language for representing simple statements of fact (propositions)
- A set of logical connectives and a simple recursive syntax (grammar) for their combination into sentences
  - Basic proposition symbols, $P_1, P_2, \ldots$, are sentences
  - If S is a sentence, $\neg S$ is a sentence (negation)
  - If S_1 and S_2 are sentences, $S_1 \wedge S_2$ is a sentence (conjunction)
  - If S_1 and S_2 are sentences, $S_1 \vee S_2$ is a sentence (disjunction)
  - If S_1 and S_2 are sentences, $S_1 \Rightarrow S_2$ is a sentence (implication)
  - If S_1 and S_2 are sentences, $S_1 \Leftrightarrow S_2$ is a sentence (biconditional)

16

Semantics of Propositional Logic (PL)

- A PL model m is an assignment of true/false values to whatever basic propositional symbols exist
  - P_1 = *True*, P_2 = *False*, P_3 = *True*, etc.
  - For n propositional symbols, 2^n such models are possible
- Given an assignment to basic symbols, all more complex sentences are true/false according to semantic rules:

- $\neg S$ is true iff (*if and only if*) S is false
- $S_1 \wedge S_2$ is true iff S_1 is true and S_2 is true
- $S_1 \vee S_2$ is true iff S_1 is true or S_2 is true
- $S_1 \Rightarrow S_2$ is true iff S_1 is false or S_2 is true
- $S_1 \Leftrightarrow S_2$ is true iff $S_1 \Rightarrow S_2$ is true and $S_2 \Rightarrow S_1$ is true

17

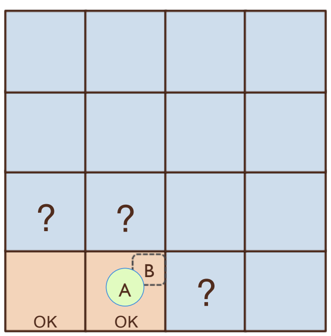
Truth Tables for PL Semantics

P	Q	$\neg P$	$P \wedge Q$	$P \vee Q$	$P \Rightarrow Q$	$P \Leftrightarrow Q$
T	T	F	T	T	T	T
T	F	F	F	T	F	F
F	T	T	F	T	T	F
F	F	T	F	F	T	T

- The basic semantic rules of *PL* correspond to a simple set of truth tables, each of which gives the result of applying one connective to one or more propositional symbols
- Each connective is defined by its truth function: takes one or more truth values as input and outputs another truth value

18

Logic in the Wumpus World



A : Agent

B : Breeze

- We can use PL to express the situation of the agent:
  - Let $P_{i,j}$ be true if there is a pit in location $[i, j]$
  - Let $B_{i,j}$ be true if there is a breeze in location $[i, j]$
  - The relevant KB is: $\{\neg P_{1,1}, \neg P_{2,1}, \neg B_{1,1}, B_{2,1}\}$
- PL can also express Wumpus World rules like "Pits cause breezes in adjacent squares":

$$B_{1,1} \Leftrightarrow (P_{1,2} \vee P_{2,1})$$
$$B_{2,1} \Leftrightarrow (P_{1,1} \vee P_{2,2} \vee P_{3,1})$$

19

Truth-Functional Analysis (01)

- Based on the basic truth tables, we can assign a semantic value to any complex PL sentence, recursively

$$P \quad Q \qquad \neg(P \vee Q) \Leftrightarrow (\neg P \wedge \neg Q)$$

20

Truth-Functional Analysis (02)

- Based on the basic truth tables, we can assign a semantic value to any complex PL sentence, recursively

P	Q	$\neg$	$(P \lor Q)$	$\Leftrightarrow$	$(\neg P$	$\land$	$\neg Q)$
T	T						
T	F						
F	T						
F	F						

Truth-Functional Analysis (03)

- Based on the basic truth tables, we can assign a semantic value to any complex PL sentence, recursively

P	Q	$\neg$	$(P \lor Q)$	$\Leftrightarrow$	$(\neg P$	$\land$	$\neg Q)$
T	T				F		F
T	F				F		T
F	T				T		F
F	F				T		T

Truth-Functional Analysis (04)

- Based on the basic truth tables, we can assign a semantic value to any complex PL sentence, recursively

P	Q	$\neg$	$(P \lor Q)$	$\Leftrightarrow$	$(\neg P$	$\land$	$\neg Q)$
T	T		T		F	F	F
T	F		T		F	F	T
F	T		T		T	F	F
F	F		F		T	T	T

Truth-Functional Analysis (05)

- Based on the basic truth tables, we can assign a semantic value to any complex PL sentence, recursively

P	Q	$\neg$	$(P \lor Q)$	$\Leftrightarrow$	$(\neg P$	$\land$	$\neg Q)$
T	T	F	T		F	F	F
T	F	F	T		F	F	T
F	T	F	T		T	F	F
F	F	T	F		T	T	T

Truth-Functional Analysis (06)

- Based on the basic truth tables, we can assign a semantic value to any complex PL sentence, recursively

P	Q	$\neg$	$(P \lor Q)$	$\Leftrightarrow$	$(\neg P$	$\land$	$\neg Q)$
T	T	F	T	T	F	F	F
T	F	F	T	T	F	F	T
F	T	F	T	T	T	F	F
F	F	T	F	T	T	T	T

25

Valid and Satisfiable Sentences

P	Q	$\neg$	$(P \lor Q)$	$\Leftrightarrow$	$(\neg P$	$\land$	$\neg Q)$
T	T	F	T	T	F	F	F
T	F	F	T	T	F	F	T
F	T	F	T	T	T	F	F
F	F	T	F	T	T	T	T

- An expression like this is valid: true in *any possible model* that assigns truth values to propositional symbols
- In addition, a sentence like this is satisfiable, i.e. there exists *some model* that makes it true

26

Valid and Satisfiable Sentences

- Key relationships to remember:

α is satisfiable $\Leftrightarrow$ $\neg\,\alpha$ is *not* valid

α is *not* satisfiable $\Leftrightarrow$ $\neg\,\alpha$ is valid

27