Logical Agents

Russell and Norvig, chapter 7

1

Knowledge based agents

Agents need to be able to:

- Store information about their environment
- Update and reason about that information

2

Knowledge based agents

Agents need to be able to:

- Store information about their environment
- Update and reason about that information

To achieve that we will introduce:

- A knowledge base (KB): a list of facts that are known to the agent.
- Rules to infer new facts from old facts using rules of inference.

3

Knowledge based agents

Agents need to be able to:

- Store information about their environment
- Update and reason about that information

To achieve that we will introduce:

- A knowledge base (KB): a list of facts that are known to the agent.
- Rules to infer new facts from old facts using rules of inference.
- Logic provides the natural language for this.

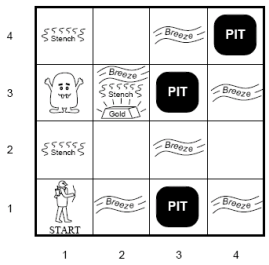
4

Knowledge Bases

- Knowledge base:
  - set of sentences in a formal language.
- Declarative approach to building an agent:
  - `Tell` it what it needs to know.
  - `Ask` it what to do → answers should follow from the KB.

5

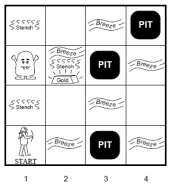
The Wumpus World



6

The Wumpus World

- Performance measure
  - gold: +1000, death: -1000
  - -1 per step, -10 for using the arrow
- Environment
  - Squares adjacent to wumpus: smelly
  - Squares adjacent to pit: 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
- Sensors: Stench, Breeze, Glitter, Bump
- Actuators: Left turn, Right turn, Forward, Grab, Release,Shoot



7

The Wumpus World Environment

- Fully Observable? No, only local perception
- Deterministic? Yes, outcome exactly specified
- Static? Yes, Wumpus and pits do not move
- Discrete? Yes
- Single-agent? Yes

8

Exploring a 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

(a)

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

Initial state

9

Exploring a wumpus world

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

(b)

After 1 move

Breezes are next to pits.
Which is the pit?

10

Exploring a wumpus world

After 3 moves

Stench here
Means Wumpus here

1,4	2,4	3,4	4,4
1,3 W!	2,3	3,3	4,3
1,2 A 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

11

Exploring a wumpus world

After 5 moves

Found gold,
Avoided Wumpus

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 A 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)

12

Logic

- We used logical reasoning to find the gold.
- Formal language for representing information such that conclusions can be drawn
- Syntax defines the sentences in the language
- Semantics define the "meaning" of sentences;
  - i.e., define truth of a sentence in a world

13

Logic

- Syntax defines the sentences in the language
- Semantics define the "meaning" of sentences;
  - i.e., define truth of a sentence in a world
- E.g., the language of arithmetic
  - Syntax: x+2 ≥ y is a sentence; x2+y > {} is not a sentence
  - Semantics: x+2 ≥ y is true in a world where x = 7, y = 1

14

Entailment

- Entailment means that one thing follows from another:

$$KB \models \alpha$$

Knowledge base *KB entails* sentence α
if and only if
α is true in all worlds where *KB* is true

  - Example: in a knowledgebase of arithmetic x+y = 4 entails 4 = x+y

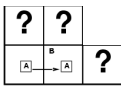
15

Models

- Model: a formally structured world with respect to which truth can be evaluated
  - Example: a model for x+y=4 is an assignment of values to x and y
  - It is true in a world where x is 2 and y is 2

16

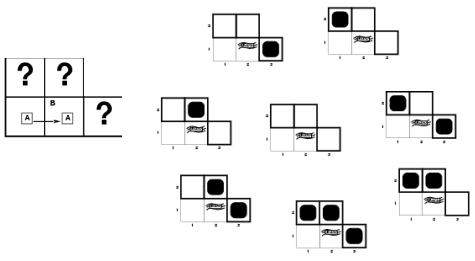
Entailment in the wumpus world



- Let's consider possible models for wumpus *KB,* restricting our attention to pits, and focusing on a region of the wumpus world.
- Situation after:
  nothing in [1,1], moving right, breeze in [2,1]

17

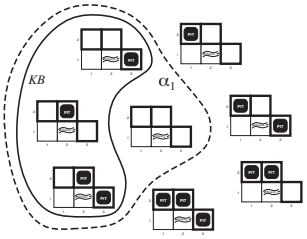
Wumpus models



All possible models

18

Exploring a wumpus world



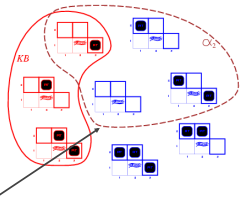
α_1 = "no pit in [1,2]"

KB = all possible wumpus-worlds consistent with the observations and behavior of Wumpus world.

KB $\models \alpha_1$ can be proved by model checking

19

Exploring a wumpus world



α_2 = "no pit in [2,2]", *KB* $\not\models \alpha_2$

20

Logical inference

- The notion of entailment can be used for logic inference.
  - Model checking: check all possible models
  - Is this a good inference method?
- *KB* $\vdash_i \alpha$ - sentence α can be derived from *KB* by procedure *i*
- If an algorithm only derives entailed sentences it is called *sound* or *truth preserving*.
  - Otherwise it just makes things up.

  i is sound if whenever KB $\vdash_i \alpha$ *it is also true that KB* $\models \alpha$
- Completeness: the algorithm can derive any sentence that is entailed.

  i is complete if whenever KB $\models \alpha$ *it is also true that KB* $\vdash_i \alpha$

21

Logic

Needed:

Representation: formalism for storing knowledge.

Reasoning: mechanism for deriving new knowledge from old: *deduction*

22

Propositional logic

- Propositional logic is a simple logic based on *propositions*:
  - Propositions are either *true* or *false*
- Examples of propositions
  - Your textbook is green
  - The sky is falling

23

Propositional logic

- Propositions: statements of fact
  - "It is raining" becomes **RAINING**
- Connectives: operators on propositions
  - "If it is raining, then it is not sunny." becomes **RAINING** → ¬ **SUNNY**

24

Propositional logic: Syntax

- The proposition symbols P_1, P_2 etc are sentences
- If S is a sentence, ¬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)

25

Truth Values

This sentence…	… is true when (iff)
S	S is true
¬S	S is false
(S_1 ^ S_2)	S_1 is true and S_2 is true
(S_1 v S_2)	At least one of S_1 or S_2 is true
($S_1 \Rightarrow S_2$)	S_1 is false or S_2 is true
($S_1 \Leftrightarrow S_2$)	S_1 & S_2 have the same truth value

26

Truth tables for connectives

P	Q	$\neg P$	$P \wedge Q$	$P \vee Q$	$P \Rightarrow Q$	$P \Leftrightarrow Q$
false	*false*	*true*	*false*	*false*	*true*	*true*
false	*true*	*true*	*false*	*true*	*true*	*false*
true	*false*	*false*	*false*	*true*	*false*	*false*
true	*true*	*false*	*true*	*true*	*true*	*true*

OR: P or Q is true or both are true.
XOR: P or Q is true but not both.

Implication is always true when the premises are False!

27

Evaluating truth value

- Simple recursive process evaluates an arbitrary sentence, e.g.,

$\neg P_{1,2} \wedge (P_{2,2} \vee P_{3,1}) = \mathit{true} \wedge (\mathit{true} \vee \mathit{false}) = \mathit{true} \wedge \mathit{true} = \mathit{true}$

28

Question

- Does {A⇒B, B} entail {A}?
  - Written as {A⇒B, B} ├ {A}
  - How do you know if this is true?

A	B	(A⇒B)^B	A
T	T	T	T
T	F	F	T
F	T	T	F
F	F	F	F

No, {A} is not entailed

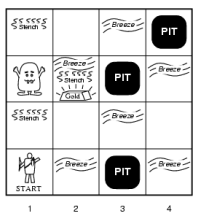
Wumpus world sentences

Let $P_{i,j}$ be true if there is a pit in [i, j].
Let $B_{i,j}$ be true if there is a breeze in [i, j].

start: ¬ $P_{1,1}$
¬ $B_{1,1}$

"Pits cause breezes in adjacent squares"

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



30

Truth tables for inference

$B_{1,1}$	$B_{2,1}$	$P_{1,1}$	$P_{1,2}$	$P_{2,1}$	$P_{2,2}$	$P_{3,1}$	KB	α_1
false	*false*	*false*	*false*	*false*	*false*	*false*	*false*	*true*
false	*false*	*false*	*false*	*false*	*false*	*true*	*false*	*true*
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
false	*true*	*false*	*false*	*false*	*false*	*false*	*false*	*true*
false	*true*	*false*	*false*	*false*	*false*	*true*	*true*	*true*
false	*true*	*false*	*false*	*false*	*true*	*false*	*true*	*true*
false	*true*	*false*	*false*	*false*	*true*	*true*	*true*	*true*
false	*true*	*false*	*false*	*true*	*false*	*false*	*false*	*true*
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
true	*true*	*true*	*true*	*true*	*true*	*true*	*false*	*false*

31

Inference by enumeration

- Enumeration of all models is sound and complete.
- For n symbols, time complexity is $O(2^n)$.
- Need a smarter way to do inference!

32