Logical agents

Propositional Logic

Outline

- Knowledge-based Agents
- Wumpus World
- Logic, models, entailment
- Propositional Logic

Logical Agents

- Human intelligence is achieved not by purely reflex mechanisms but by processes of reasoning that operate on internal representations of knowledge
- In AI, this intelligence is embodied in knowledge-based agents
- Agents seen previously are limited and inflexible – cannot do inference or learning of new knowledge
- Develop logic as a general class for representation for KB agents

Knowledge-based Agents (KBAs)

- Knowledge base (KB) = a set of sentences in a formal language, i.e. a computer interpretable language
- A sentence that is known to be true without being derived from other sentences is an axiom
- Add sentences and query a KB by TELL-ing (facts) and ASK-ing (for inference) operators
- Inference – deriving new sentences from old
- Declarative approach is one where a KB agent (or any system) is built by TELL-ing it what it needs to know and it can then operate and perform inferences – what
- Procedural approach encodes desired behaviours directly as a program code – how

KBA - Pseudocode

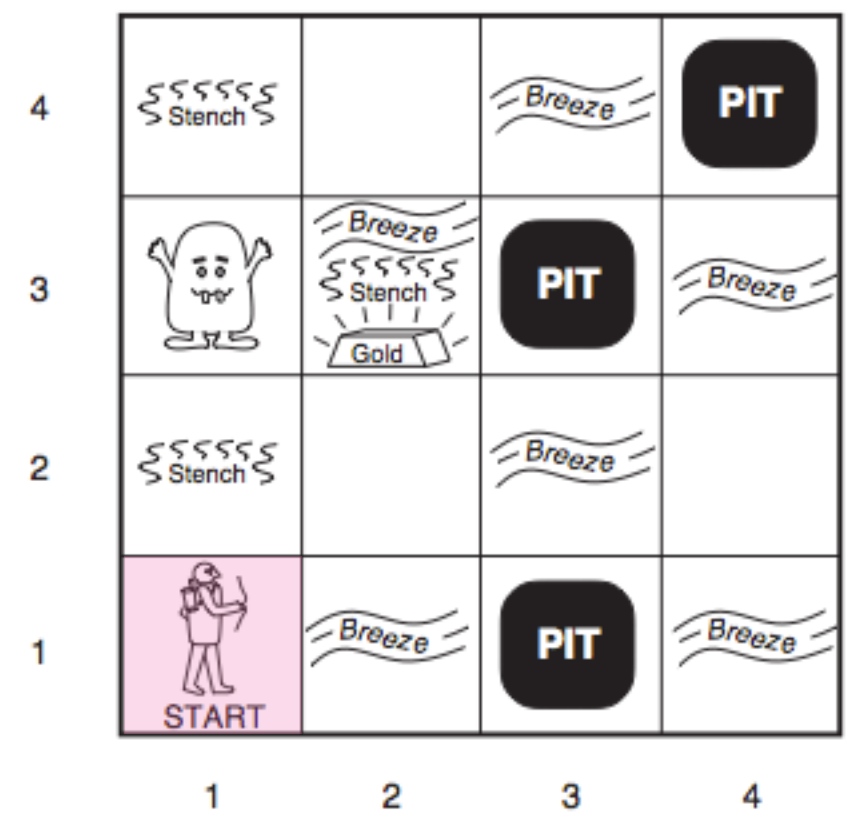
```
function KB-AGENT(percept) returns an action
   persistent: KB, a knowledge base
               t, a counter, initially 0, indicating time

   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
```

- The KB 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

Wumpus World PEAS

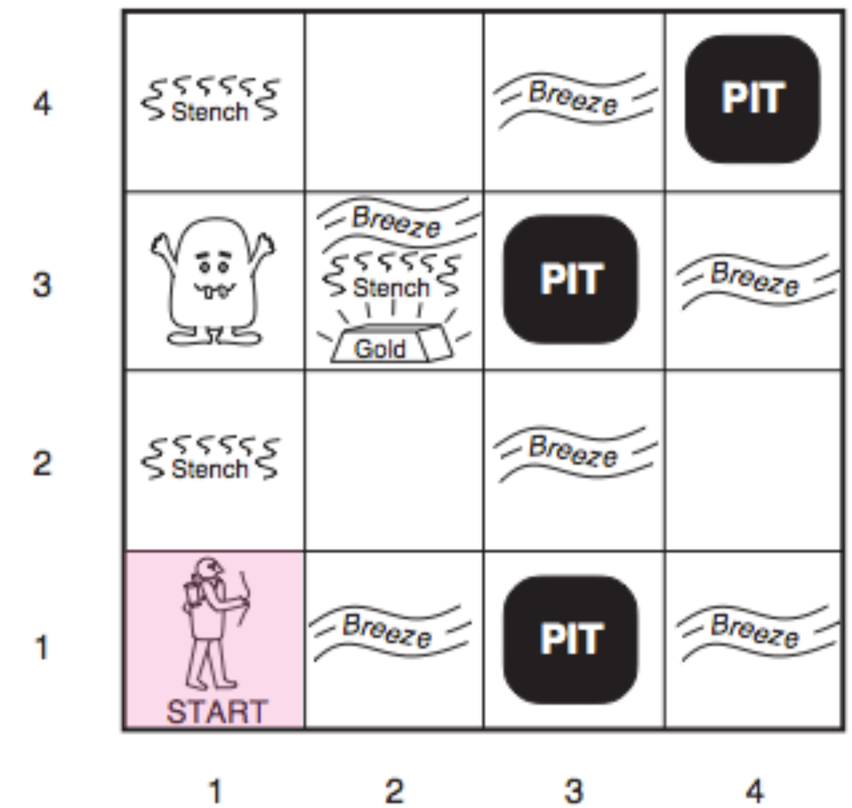
- **Performance measure**
  - gold +1000, death by wumpus or pit -1000
  - -1 for each step, -10 for using arrow
- **Environment**
  - squares adjacent to wumpus are smelly
  - squares adjacent to pit are breezy
  - glitter iff gold is in the same square
  - shooting kills wumpus if you are facing it
  - shooting uses up the only arrow
  - grabbing picks up gold if in same square
  - releasing drops the gold in the same square
- **Actuators (Actions)**
  - TurnLeft 90°, Turn Right 90°, Forward, Grab, Release, Shoot
- **Sensors**
  - Stench, Breeze, Glitter, Bump, Scream (Wumpus dies) **[0, 0, 0, 0, 0]** if all not sensed



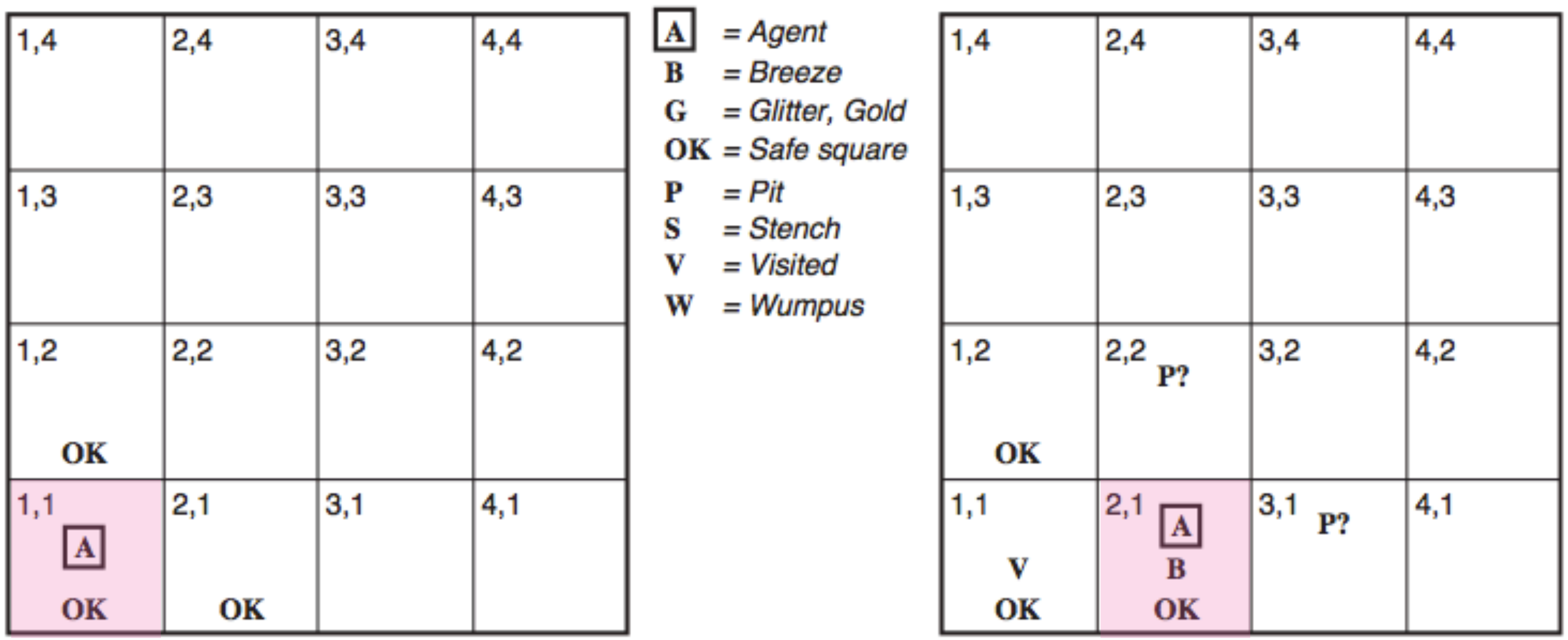
* Q: When does the game end?

Wumpus World Characterisation

- Observable??
  - No – only local perception
- Deterministic??
  - Yes – outcomes exactly specified and determined by current state
- Static??
  - Yes – Wumpus and Pits do not move
- Discrete??
  - Yes – there is a finite number of states
- Single-agent?
  - Yes – Wumpus is a natural feature, not an agent



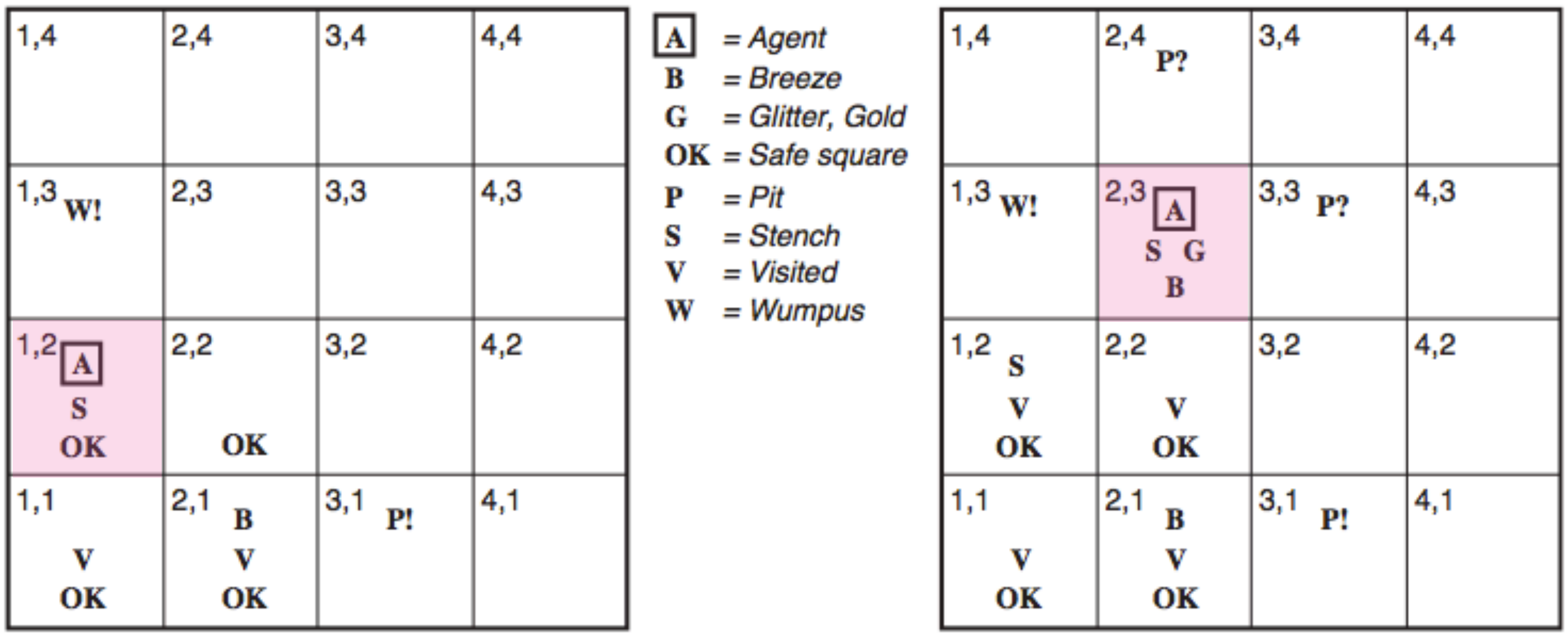
Exploring a Wumpus World



- The first step at [1,1]
- Percept *[0,0,0,0,0] i.e. [None, None, None, None, None]*

- After one move at [2, 1]
- Percept *[0,1,0,0,0] i.e. [None, Breeze, None, None, None]*

Exploring a Wumpus World

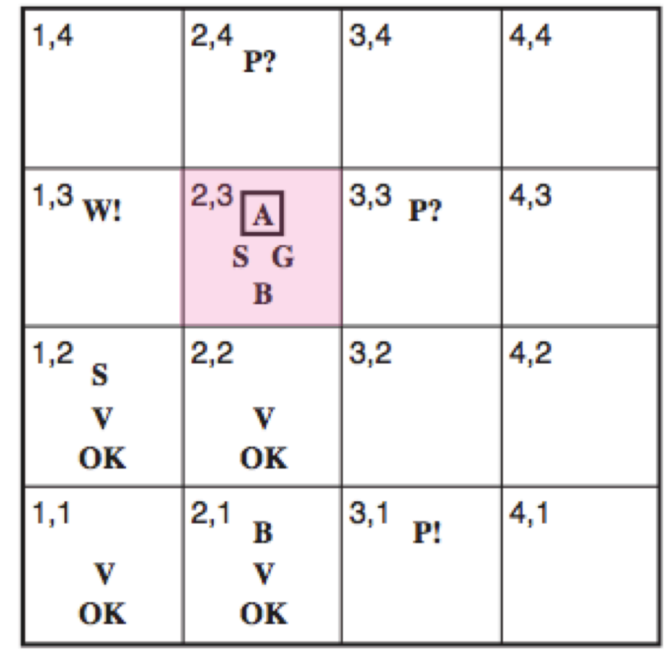


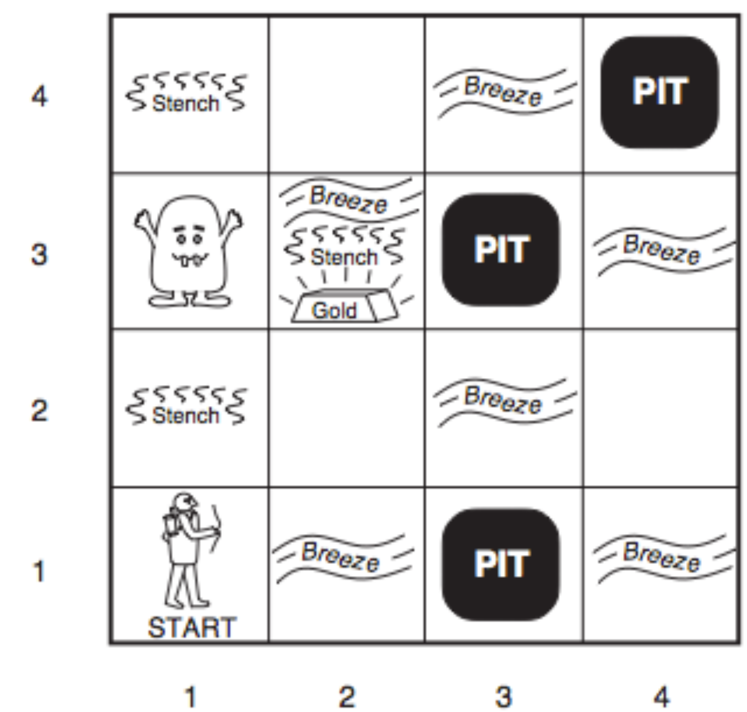
- After third move at [1,2]
- Percept *[1,0,0,0,0] i.e. [Stench, None, None, None, None]*

- After fifth move at [2, 3]
- Percept *[1,1,1,0,0] i.e. [Stench, Breeze, Glitter, None, None]*

WW – Inferencing

- Retreated to [1,1] after realising *Pit* could be in [3,1] or [2,2] and goes to only safe square which is [1,2]
- Stench at [1,2] means that *Wumpus* is either in [1,3] or [2,2]
- But it can't be in [2,2] because there was no stench at [2,1] which it is adjacent to
- Agent infers that *Wumpus* is in [1,3] and moves to [2,2] safely
- Without much inferencing, except that square [2,2] is safe, agent moves to [2,3] where it detects a *Glitter*
- That's enough for it to *Grab* it!





Logic

- Logics are formal languages for representing information such that conclusions can be drawn
- Syntax defines the **sentences (statements)** in the language
- Semantics define the "meaning" of sentences; i.e. define **TRUTH** of a sentence with respect to each possible world
- E.g., in the language of arithmetic

  $x + 2y \geq 2$ is a sentence; $x + 2y$ is not a sentence

  $x + 2y \geq 2$ is true iff the number $x + 2$ is greater than or equal to the number y

  $x + 2y \geq 2$ is true in a world where $x = 7, y = 1$

  $x + 2y \geq 2$ is false in a world where $x = 0, y = 0$

  $x + 2 \geq x + 1$ is true in every world
- Are these well formed logical sentences/propositions?
  - The window is open
  - Moon Jae In is the president of South Korea
  - Are you going to the party?
  - 2+3 = 6
  - Stay at home, please!

Models

- To be precise, the word model should be used in the place of "possible worlds"; models are mathematical abstractions, each of which fixes truth and falsehood of every relevant sentence
- Formally, the possible models are just all the possible assignments of real numbers to x and y in the algebra example
- m satisfies a sentence α (or m is a model of α) if α is true in m
- $M(\alpha)$ is a set of all models of α
- $\alpha : xy = 0$
- $M : x = 0, y = 1;\ x = 0, y = 2;\ \ldots\ x = 0, y = 10{,}000;$
  $\ldots\ x = 1, y = 1;\ x = 3, y = 0;\ \ldots$ // some are True and some are False

Entailment $\models$

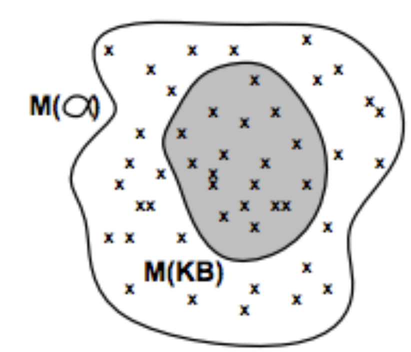
- A sentence or set of sentences, KB entails the sentence α

  $KB \models \alpha$

- Knowledge base KB entails (or logically implies) a sentence α if and only if **in every model in which $\mathbf{KB}$ is true, $\boldsymbol{\alpha}$ is also true**

  $KB \models \alpha \quad \text{iff} \quad M(KB) \subseteq M(\alpha)$

- Examples
  - the KB containing “Spurs won” and “Liverpool won”entails “either Spurs won or Liverpool won”
  - $x + y = 4$ entails $4 = x + y$
  - $x = 0$ entails $xy = 0$



Inference $\vdash$

- $KB \vdash_i \alpha$ sentence α can be derived, deduced or inferred from KB by inference algorithm i
- Entailment = needle α in haystack KB; inference = finding it
- Inference algorithm that derives only entailed sentences is called sound or truth-preserving
  - whenever $KB \vdash_i \alpha$, it is also true that $KB \models \alpha$
- An inference algorithm is complete if it can derive any sentence that is entailed
  - whenever $KB \models \alpha$, it is also true that $KB \vdash_i \alpha$
- Both soundness and completeness are desirable properties, e.g. for model checking

Soundness & Completeness

- Soundness: Everything that is provable in the system is actually true. In a sound system one cannot derive a proposition and its negative from the axioms (e.g. $p \wedge \neg p$ which is a contradiction)
- Completeness: Everything that is true in the system is provable. In a complete system all the true propositions can be derived from the axioms

Soundness and Completeness of a Logical System

- **Soundness**: Everything that is provable in the system is actually true
  - If $KB \vdash \alpha$ then $KB \models \alpha$
  - If something is derivable, then it is valid
  - The inference procedure/proofs don't include anything that's wrong
- **Completeness**: Everything that is true in the system is provable
  - If $KB \models \alpha$ then $KB \vdash \alpha$
  - If something is valid, then it is derivable
  - The inference procedure/proofs don't exclude anything that's correct
- Propositional Logic is sound and complete

Propositional Logic

- Propositional logic is the simplest logic – illustrates basic ideas
- Syntax of well-formed propositional formulas:

The proposition symbols P_1, P_2, Q, $W_{1,3}$, etc. are sentences

If S is a sentence, $\neg S$ is a sentence (negation "not S")

If S_1 and S_2 are sentences, $S_1 \wedge S_2$ is a sentence (conjunction "S_1 and S_2")

If S_1 and S_2 are sentences, $S_1 \vee S_2$ is a sentence (disjunction "S_1 or S_2")

If S_1 and S_2 are sentences, $S_1 \Rightarrow S_2$ is a sentence (implication "S_1 implies S_2" /conditional "If S_1 then S_2")

If S_1 and S_2 are sentences, $S_1 \Leftrightarrow S_2$ is a sentence (bi-conditional / "S_1 if and only if S_2")

Propositional Logic - Semantics

- Each model specifies true/false for each propositional symbol
- Rules for evaluating truth with a respect to a model *m*:

$\neg S$	is true iff	S	is false			
$S_1 \wedge S_2$	is true iff	S_1	is true	AND	S_1	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
i.e.	is false iff	S_1	is true	and	S_2	is false
$S_1 \Leftrightarrow S_2$	is true iff	$S_1 \Rightarrow S_2$	is true	and	$S_2 \Rightarrow S_1$	is true

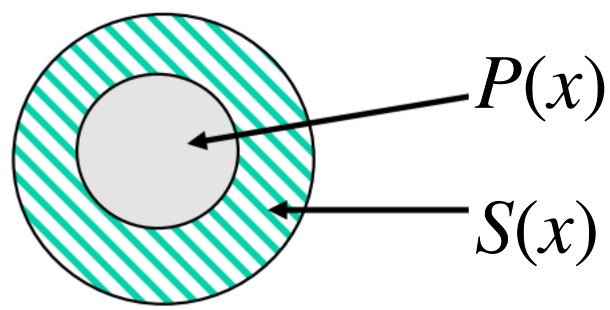
Truth Tables for Connectives

P	Q	$\neg P$	$P \wedge Q$	$P \vee Q$	$P \Rightarrow Q$	$P \Leftrightarrow Q$
F	F	T	F	F	T	T
F	T	T	F	T	T	F
T	F	F	F	T	F	F
T	T	F	T	T	T	T
		negation	and	or	implies	iff

T = True
F = False

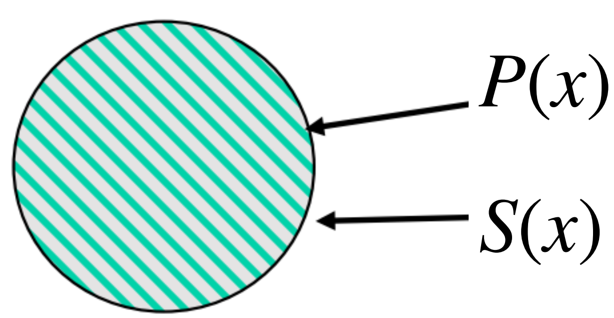
Implication and Biconditional

$S(x)$ is a necessary condition of $P(x)$



$$P(x) \Rightarrow S(x)$$

$S(x)$ is a necessary and sufficient condition of $P(x)$



$$P(x) \Leftrightarrow S(x)$$

Exercise

- Sentences/Statements

  S The store is open today

  M Mary is going to the store today

  J John is going to the store today

- Define propositional sentences using the variables above and logical connectives ($\neg, \wedge, \vee,$) for the following:
- "Either John or Mary (or both) are going to the store today"
- "John is going to the store today, but Mary isn't"
- "The store is open today, **however**, **neither** John nor Mary is going"

Entailment using TT

- $KB \models \alpha$ only if in every model where KB is $True$, α is also $True$
- Let's use truth tables to show if
  - $P \models (P \vee Q)$

P	Q	$P \vee Q$
T	T	T
T	F	T
F	T	T
F	F	F

- Use truth tables to show if
  - $P \models (P \wedge Q)$

Simple KB for Wumpus World

- Symbols (representation)
- $P_{x,y}$ is true if there is a pit in $[x,y]$.
- $W_{x,y}$ is true if there is a wumpus in $[x,y]$.
- $B_{x,y}$ is true if agent perceives breeze in $[x,y]$.
- $S_{x,y}$ is true if agent perceives stench in $[x,y]$.
- Insert some known facts
- There is no pit in [1,1]

R_1: $\neg P_{1,1}$

- A square is breezy if and only if there is a pit in a neighbouring square

R_2: $B_{1,1} \Leftrightarrow P_{1,2} \vee P_{2,1}$

R_3: $B_{2,1} \Leftrightarrow P_{1,1} \vee P_{2,2} \vee P_{3,1}$

- Breeze percepts

R_4: $\neg B_{1,1}$

R_5: $B_{2,1}$

Truth Table for Inference

$$KB \models \neg P_{1,2} \text{ ?}$$

$B_{1,1}$	$B_{2,1}$	$P_{1,1}$	$P_{1,2}$	$P_{2,1}$	$P_{2,2}$	$P_{3,1}$	R_1	R_2	R_3	R_4	R_5	KB
false	*false*	*false*	*false*	*false*	*false*	*false*	*true*	*true*	*true*	*true*	*false*	*false*
false	*false*	*false*	*false*	*false*	*false*	*true*	*true*	*true*	*false*	*true*	*false*	*false*
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
false	*true*	*false*	*false*	*false*	*false*	*false*	*true*	*true*	*false*	*true*	*true*	*false*
false	*true*	*false*	*false*	*false*	*false*	*true*	*true*	*true*	*true*	*true*	*true*	*true*
false	*true*	*false*	*false*	*false*	*true*	*false*	*true*	*true*	*true*	*true*	*true*	*true*
false	*true*	*false*	*false*	*false*	*true*	*true*	*true*	*true*	*true*	*true*	*true*	*true*
false	*true*	*false*	*false*	*true*	*false*	*false*	*true*	*false*	*false*	*true*	*true*	*false*
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
true	*true*	*true*	*true*	*true*	*true*	*true*	*false*	*true*	*true*	*false*	*true*	*false*

KB is true if R_1 through R_5 are true, occurs in just 3 rows out of 128. In those rows, $P_{1,2}$ is false, so there is no pit in [1,2]. But there may or may not be a pit in $P_{2,2}$ or $P_{3,1}$

Summary

- Knowledge-based agents apply inference to KB to derive new information and make decisions
- PEAS for Wumpus World
- Logic, models, entailment and inference
- Soundness and completeness
- Propositional logic

References

- Russel and Norvig, Chapter 7
- Recommended programming language for logic is Prolog