Logical Agents and Propositional Logic

Dr. Demetrios Glinos
University of Central Florida

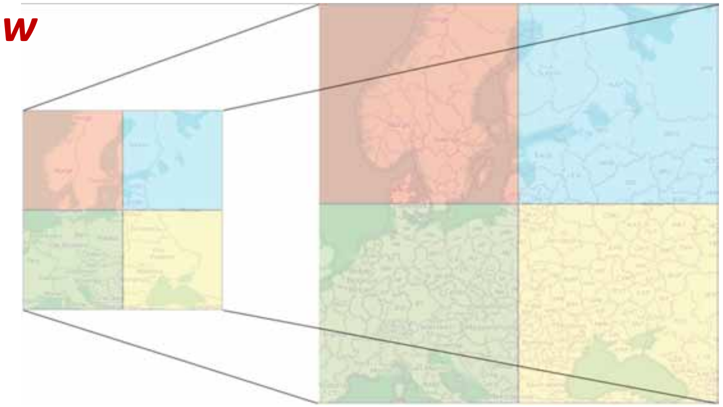
CAP4630 – Artificial Intelligence

Outline

- Knowledge Concepts
- The Wumpus World
- Logic Concepts
- Propositional Logic

Knowledge

- Humans make decisions based on what we ***know***
- So do our computational agents
- But *where* is the knowledge in our agents?



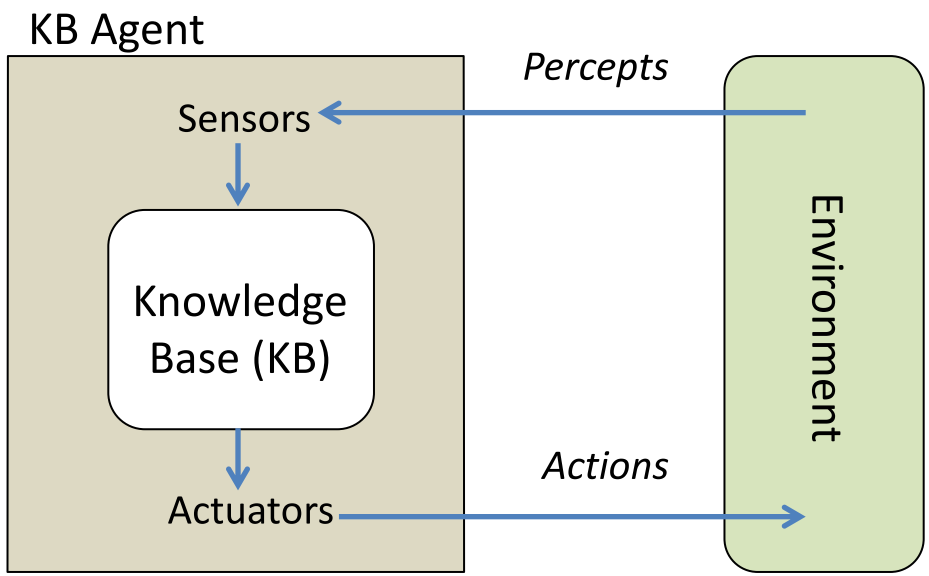
- Basic search agents:
    - knowledge is in the successor function
- Constraint satisfaction agent:
    - knowledge is in the constraints
    - knowledge is in the domains
- Adversarial agents:
    - knowledge is in the transition function
    - knowledge is in the reward function

Reasoning

- We need some kind of repository for all the knowledge we accumulate

- Need a ***knowledge base (KB)***
  - a collection of assertions
  - in some representation language
- May start with "background knowledge"
  - Similar to instinct?
- Must be able to add new sentences to the KB
  - from percepts
  - by inferring new sentences from existing ones

Knowledge-Based Agent



KB Agent Processing:

- *Add percept to KB*
- *Query KB for best action*
- *Add action taken to KB*

Knowledge Level

- KB Agent is fully specified by
  - what it knows
  - what its goals are
  - we assume logic is the reasoning mechanism
- This is independent of its implementation
- Building up the KB
  - by telling it what it needs to know
    - declaratively
    - procedurally
  - by enabling it to learn on its own
    - learning by examples
    - learning from experience

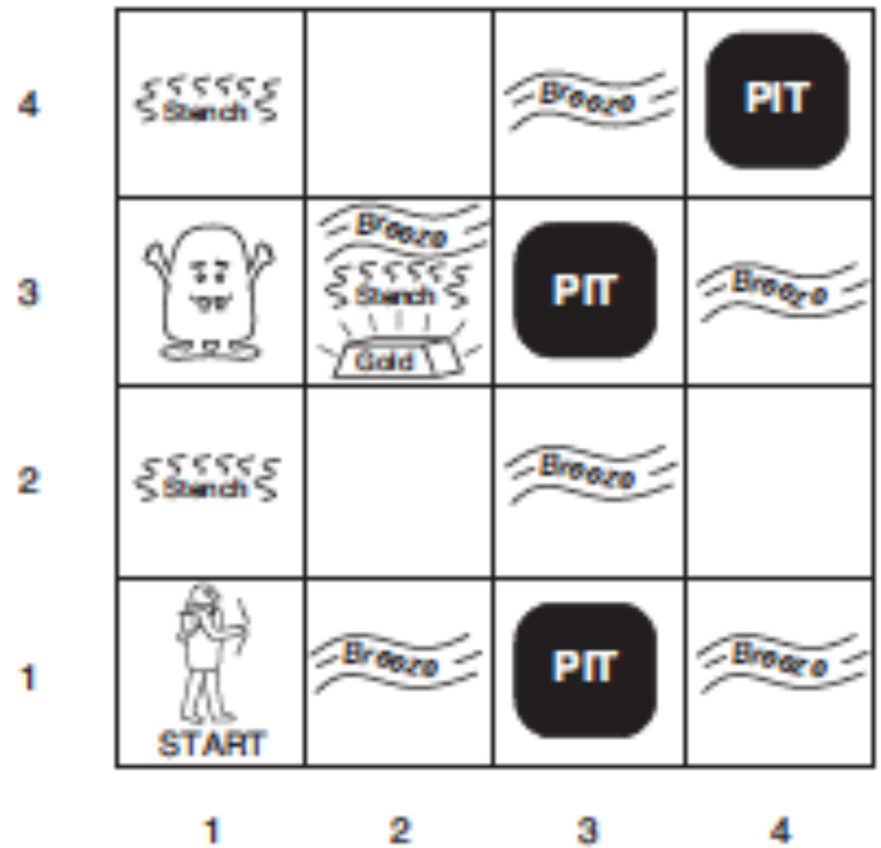
The Wumpus World

Performance:
- +1000 climb out with gold
- -1000 die from Wumpus or pit
- -1 each move, -10 shoot

Environmant
- 4 x 4 grid, random pits (prob .2)
- random Wumpus & gold locations
- agent starts in 1,1 with only 1 arrow
- must climb out from 1,1

Actuators
- Fwd, Back, Left, Right
- Grab, Shoot, Climb

Sensors
- stench around Wumpus
- breeze around pits
- glitter where the gold is
- scream when Wumpus killed
- bump when hit wall (don’t move)

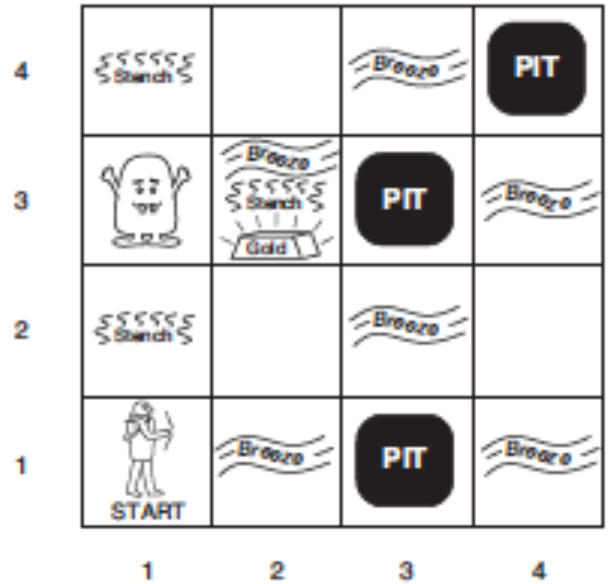


Navigating the Wumpus World

Goal: Get the gold and get out

The problem: Don't know where the Wumpus, the gold, and the pits are

Logical reasoning is needed to solve this based on percepts:



Percepts of form: < Stench, Breeze, Glitter, Bump, Scream >

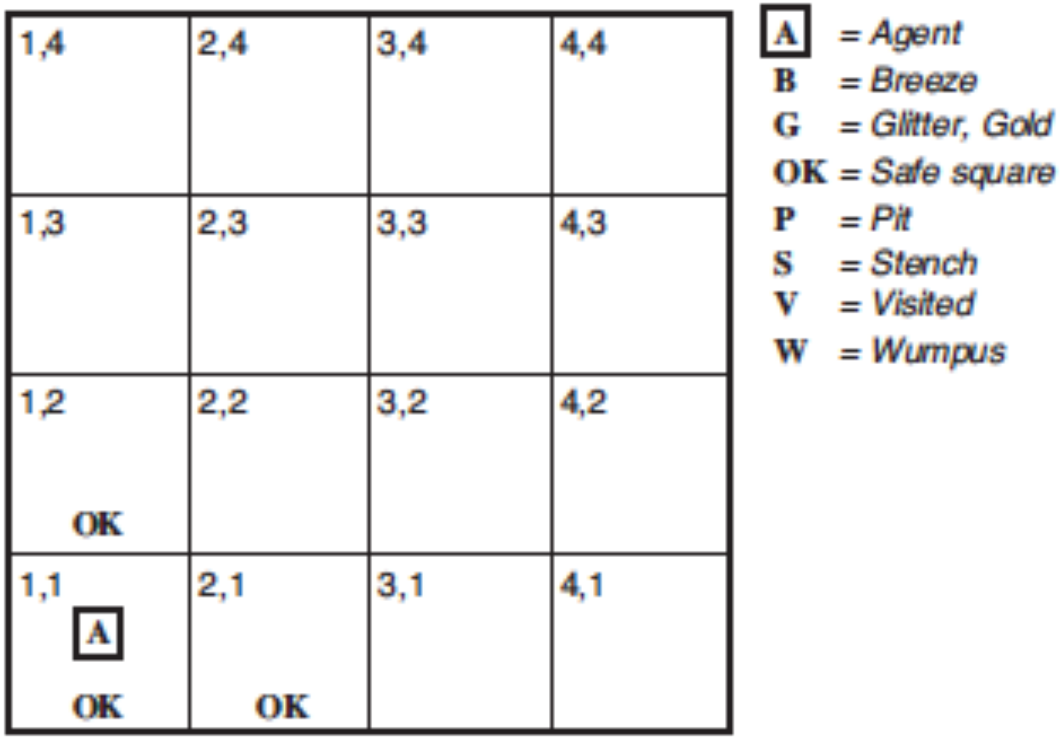
Example: in location (2,1) we have < F, T, F, F, F >

(Note: we follow our text and use column-first notation)

Example: Wumpus game

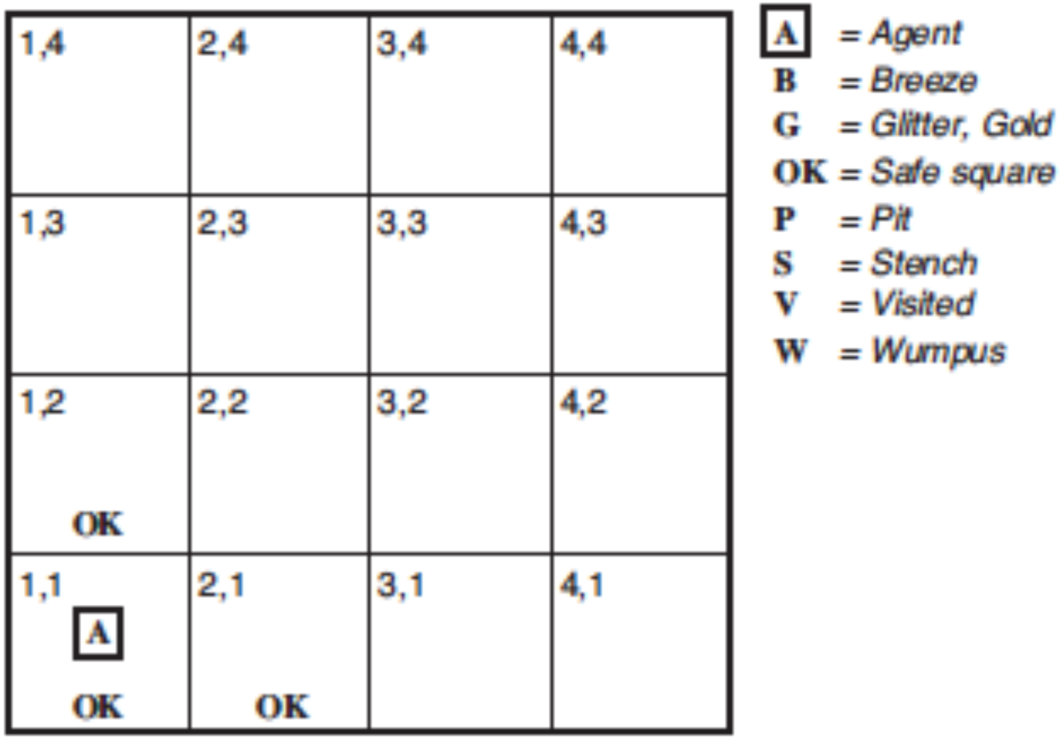
Start in 1,1 with no stench or breeze

What does the agent infer?



Example: Wumpus game

Start in 1,1 with no stench or breeze



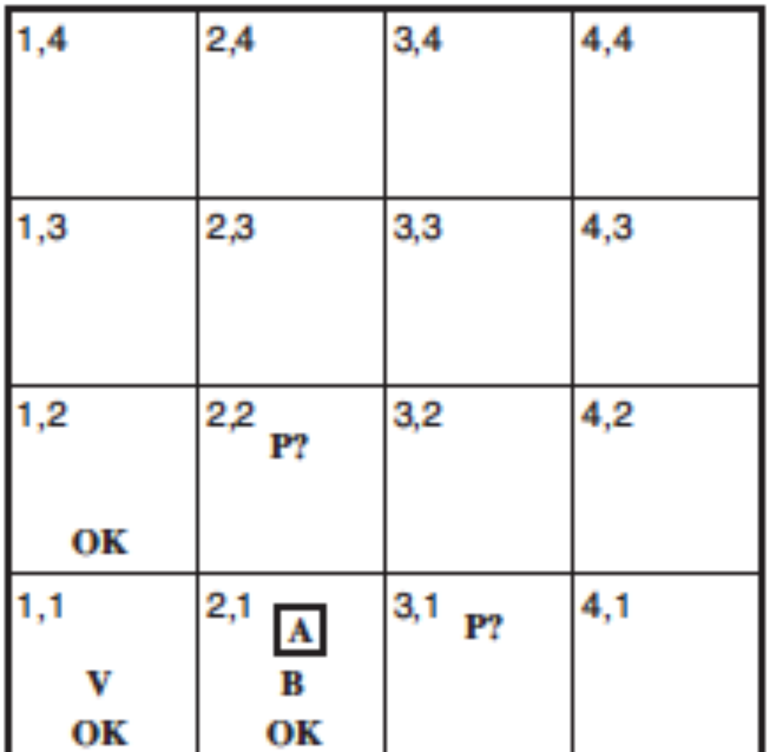
What agent infers:

- no Wumpus or pit in 1,2 and 2,1

Example: Wumpus game

Move to 2,1 and perceive breeze

What does the agent infer?



Example: Wumpus game

Move to 2,1 and perceive breeze

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 $\boxed{A}$ B OK	3,1 P?	4,1

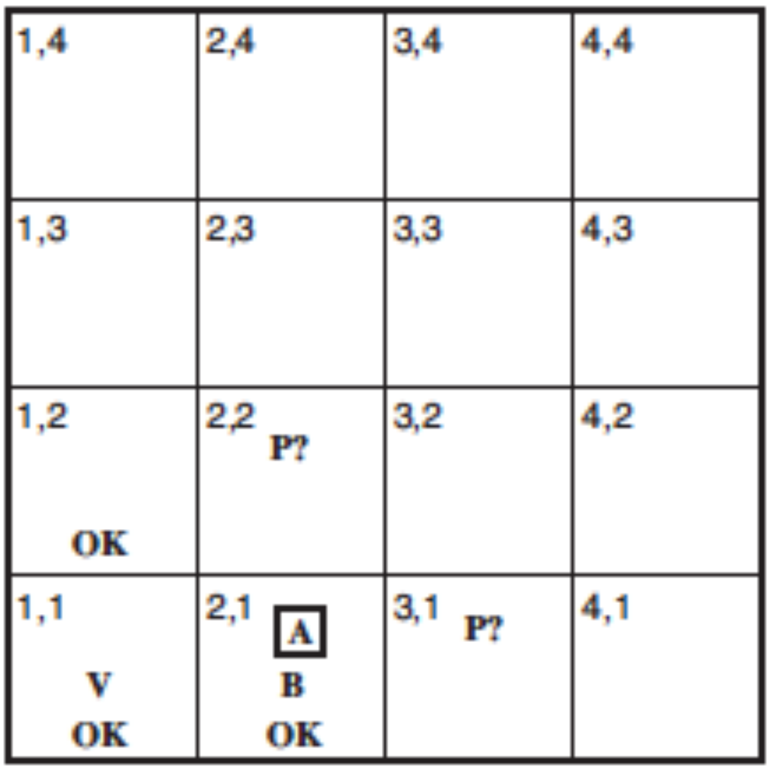
What agent infers:

- possible pit in either or both 2,2 and 3,1

What should the agent do?

Example: Wumpus game

Move to 2,1 and perceive breeze



What agent infers:

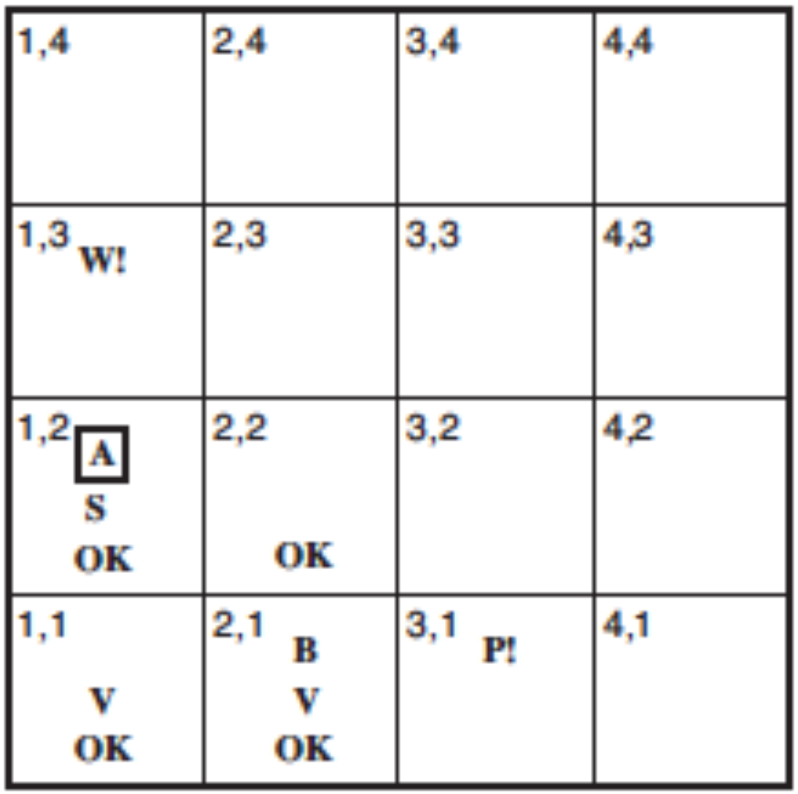
- possible pit in either or both 2,2 and 3,1

→ good idea to backtrack and try 1,2

Example: Wumpus game

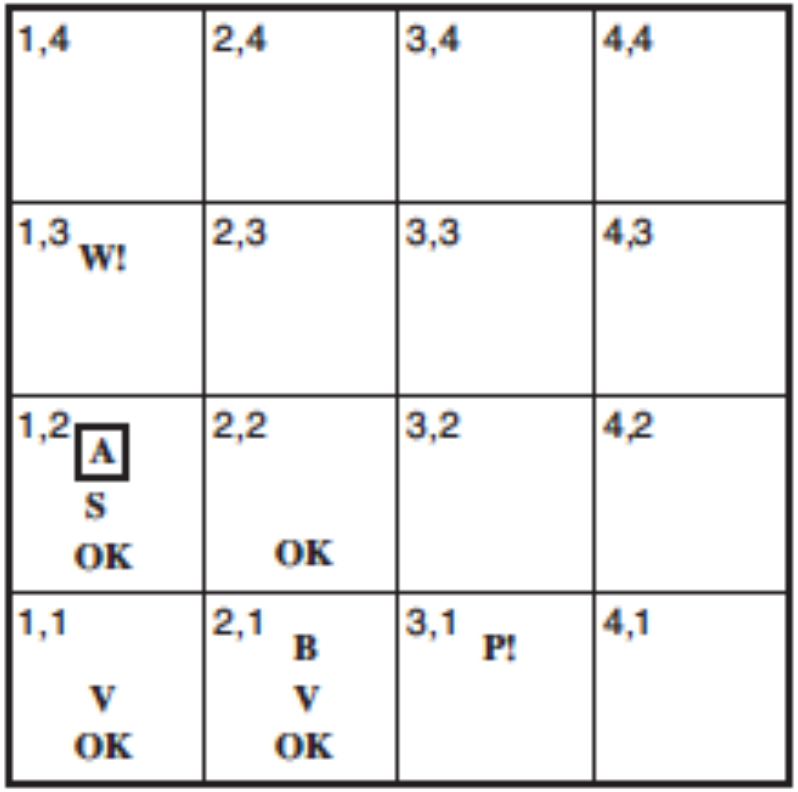
Move to 1,2 and perceive stench

What does the agent infer?



Example: Wumpus game

Move to 1,2 and perceive stench



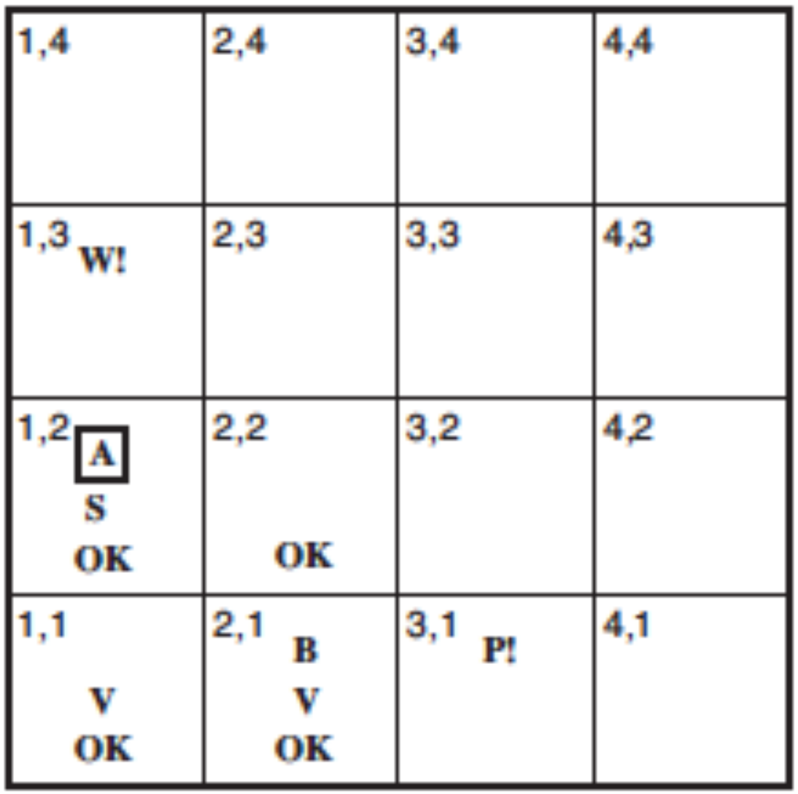
What agent infers from percept:

- stench, so Wumpus in 1,3 or 2,2

So, what else can the agent infer from this?

Example: Wumpus game

Move to 1,2 and perceive stench



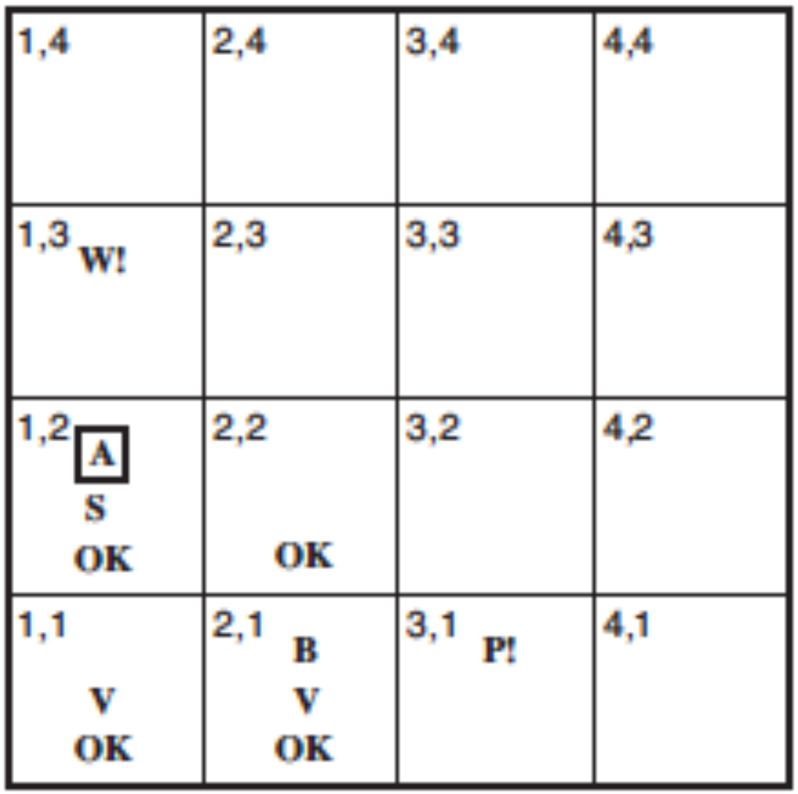
What agent infers:

- stench, so Wumpus in 1,3 or 2,2
    - but no Wumpus in 2,2
    - so Wumpus in 1,3

What else did the agent get from the percept?

Example: Wumpus game

Move to 1,2 and perceive stench



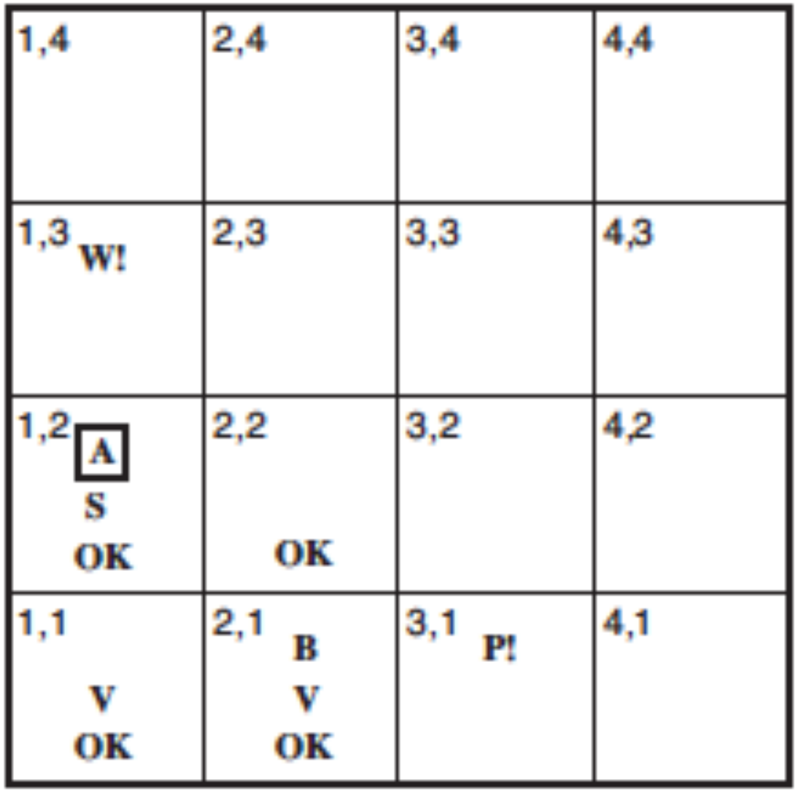
What agent infers:

- stench, so Wumpus in 1,3 or 2,2
  - but no Wumpus in 2,2
  - so Wumpus in 1,3
- no breeze, so no pit in 1,3 or 2,2

So, what can the agent infer from this?

Example: Wumpus game

Move to 1,2 and perceive stench



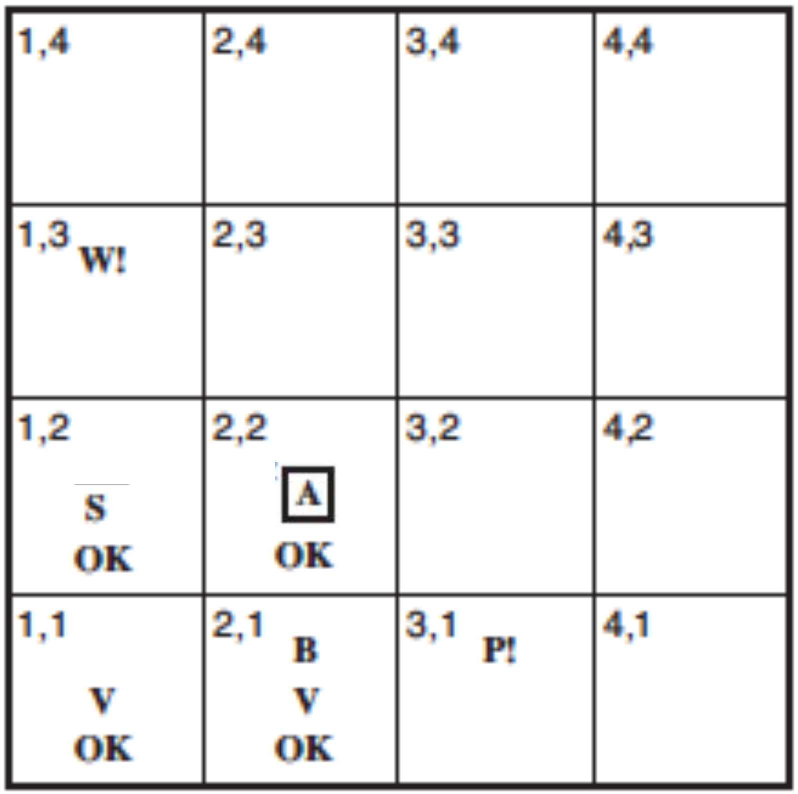
What agent infers:

- stench, so Wumpus in 1,3 or 2,2
  - but no Wumpus in 2,2
  - so Wumpus in 1,3
- no breeze, so no pit in 1,3 or 2,2
  - so 3,1 is a pit
  - so 2,2 is safe

So, what should the agent do?

Example: Wumpus game

Move to 2,2 no stench or breeze



What agent infers:

- no breeze, so no pit in 2,3 or 3,2
- we already know where the Wumpus is

→ Safe to try 2,3 and 3,2

Example: Wumpus game

Move to 2,3 and perceive glitter, stench and breeze

1,4	2,4 P?	3,4	4,4
1,3 W!	2,3 $\boxed{A}$ 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

What agent infers:

- breeze, so possible pit in 2,4 and/or 3,3 (but who cares now?)
- glitter, so gold in 2,3

→ Plan:

- grab gold
- return to 1,1
- climb out

Logic Concepts

- Syntax
  - for each particular representation scheme
  - tells us which sentences are *valid*
    - example, “x + y = 4” is valid in algebra, but “x 4 y + =“ is not

- Semantics
  - defines the *meaning* of a sentence
  - defines the *truth* of each sentence *with respect to each possible world*
    - example: “x + y = 4” is true in a world where x=2 and y=2, but not in a world where x=3 and y=8.

- Standard logics:
  - a sentence must be either T or F in each possible world
  - often different T/F value for different worlds (as in algebra example above)

Models and Truth

P	Q	$P \wedge Q$	$\neg(P \wedge Q)$
T	T	T	F
T	F	F	T
F	T	F	T
F	F	F	T

- A **model** is a possible world
  - assigns T/F value to each sentence
- If a sentence **α** is true in a model **m**, then we say:
  - model m ***satisfies*** sentence α
  - or, equivalently, m is a ***model of*** α
  - and **M(α)** denotes the set of ***all*** models of α
- logical entailment
  - if sentence β ***follows*** from sentence α, we say that α **entails** β
  - $\alpha \models \beta$ means "α entails β"
  - formally, $\alpha \models \beta$ if and only if $M(\alpha) \subseteq M(\beta)$
  - here, α is a ***stronger*** assertion than β (it rules out more possible worlds)

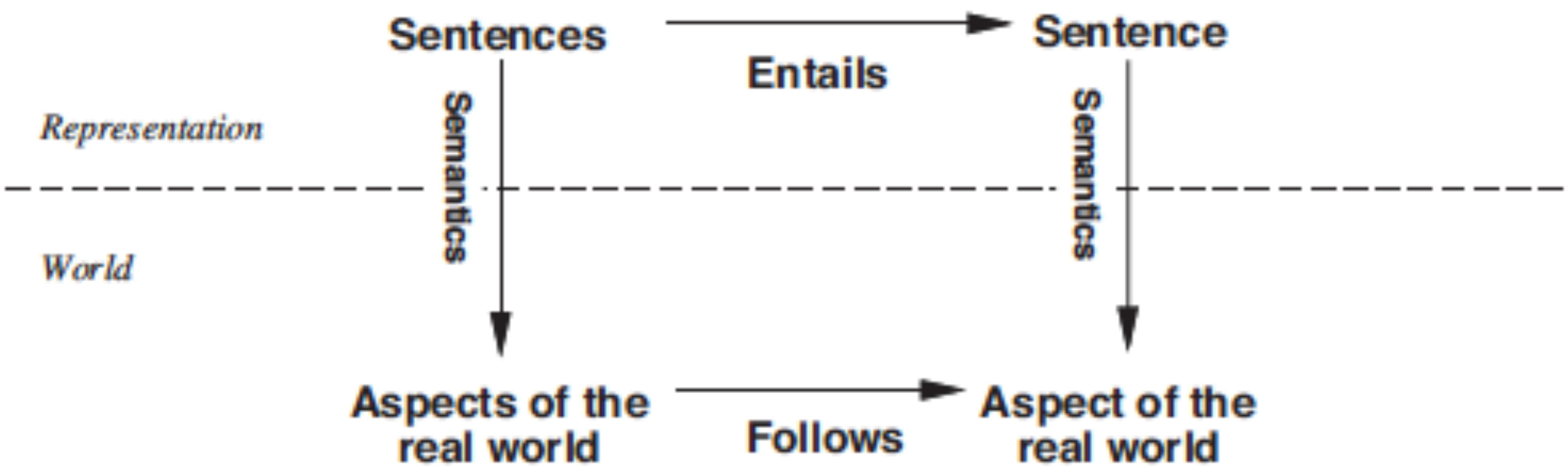
Logical Inference

- Logical inferencing involves determining whether sentence α is entailed by the KB

- We can do this by **model checking** (*"It must be true"*)
  - enumerate all models consistent with KB
  - verifying that α is true for *all* of them, i.e., that $M(KB) \subseteq M(\alpha)$

- We can also do this using **inference algorithms**
  - denote by $\vdash_k$
  - if inference algorithm k can derive α from KB, we write $\mathbf{KB} \vdash_k \boldsymbol{\alpha}$ and we say that *"α is derived from the KB by k"* or *"k derives α from KB"*
  - inference algorithms that derive only entailed sentences are called *sound* or *truth-preserving*
  - an inference algorithm has the *completeness* property if it can derive ***any*** sentence that is entailed

Knowledge-Based Reasoning



Basis of KB reasoning:

> ***If*** KB represents true statements about the real world, ***then*** any sentence derived from KB also says something true about the real world

Grounding

- **Question:** How do we know that the KB represents the real world?

- **Answer:**
  - by construction
    - from sensors: When perceive stench, create appropriate assertion
  - by sound inferences
    - from learning: From understanding, based on experience

- Learning is fallible

  - need good learning procedures
  - our model of the world may not capture all essential details
    - e.g., Wumpuses may cause stench, except on February 29 in leap years when they take their baths.

Propositional Logic

- The logic of *sentences* that can be assigned truth values (propositions)
  - e.g., “The Wumpus is in 1,3”, “I like spanakopita”, “Peter was here yesterday”
- Not all sentences are propositions
  - e.g., “Where is my car?”, “Turn right at the light”, “Hello”
- **Atomic sentence**
  - not composed from simpler sentences
  - usually represented by a single *symbol* (usually, a letter of the alphabet)
- **Complex sentence**
  - composed from simpler sentences using parentheses and *logical connectives*

Grammar for Propositional Logic

Sentence	$\rightarrow$	AtomicSentence | ComplexSentence
AtomicSentence	$\rightarrow$	*True* | *False* | P | Q | R | ...
ComplexSentence	$\rightarrow$	(Sentence) | [Sentence]
		| $\neg$ Sentence
		| Sentence $\land$ Sentence
		| Sentence $\lor$ Sentence
		| Sentence $\Rightarrow$ Sentence
		| Sentence $\Leftrightarrow$ Sentence

OPERATOR PRECEDENCE : $\neg, \land, \lor, \Rightarrow, \Leftrightarrow$

Terminology

Negation (logical "not") : $\neg$

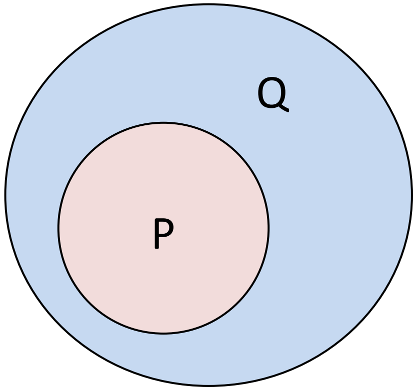
Literal : either an atomic sentence or a negated atomic sentence

Conjunction (logical "and"): $\wedge$

Disjunction (logical "or"): $\vee$

Implication ("implies"): $\Rightarrow$

Biconditional ("if and only if"): $\Leftrightarrow$



Semantics

- In propositional logic, a model *fixes* the truth values for every propositional symbol
- For a finite set of propositions, we can enumerate all models in a truth table

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

Example: KB for the Wumpus World

- Let P = pit, W = Wumpus, B = breeze, and S = stench
- Let us use subscripts to represent the location, e.g., $P_{3,2}$

- Consider this KB:

R_1:	$\neg P_{1,1}$	no pit in 1,1
R_2:	$B_{1,1} \Leftrightarrow (P_{1,2} \vee P_{2,1})$	breeze iff pit in adjacent square
R_3:	$B_{2,1} \Leftrightarrow (P_{1,1} \vee P_{2,2} \vee P_{3,1})$	
R_4:	$\neg B_{1,1}$	no breeze in 1,1
R_5:	$B_{2,1}$	breeze in 2,1

The above is sufficient to derive: $\neg P_{1,2}$

Derivation

Q: How can we derive $\neg P_{1,2}$ from our KB?

The KB:

R_1:	$\neg P_{1,1}$	no pit in 1,1
R_2:	$B_{1,1} \Leftrightarrow (P_{1,2} \vee P_{2,1})$	breeze iff pit in adjacent square
R_3:	$B_{2,1} \Leftrightarrow (P_{1,1} \vee P_{2,2} \vee P_{3,1})$	
R_4:	$\neg B_{1,1}$	no breeze in 1,1
R_5:	$B_{2,1}$	breeze in 2,1

Derivation

Q: How can we derive $\neg P_{1,2}$ from our KB?

The KB:

R_1:	$\neg P_{1,1}$	no pit in 1,1
R_2:	$B_{1,1} \Leftrightarrow (P_{1,2} \lor P_{2,1})$	breeze iff pit in adjacent square
R_3:	$B_{2,1} \Leftrightarrow (P_{1,1} \lor P_{2,2} \lor P_{3,1})$	
R_4:	$\neg B_{1,1}$	no breeze in 1,1
R_5:	$B_{2,1}$	breeze in 2,1

A: Use R_4 and R_2

R_4 says $B_{1,1}$ is False

So, by R_2, $(P_{1,2} \lor P_{2,1})$ is also False

So, $P_{1,2}$ and $P_{2,1}$ are both False

Therefore $P_{1,2}$ is false

Model-Checking Inference Algorithm

- enumerate ***all*** of the models consistent with the KB
- check that the desired proposition is true for ***every*** such model
- KB is true in a model if every proposition in KB is true in that model

R_1:	$\neg P_{1,1}$	no pit in 1,1
R_2:	$B_{1,1} \Leftrightarrow (P_{1,2} \lor P_{2,1})$	breeze iff pit in adjacent square
R_3:	$B_{2,1} \Leftrightarrow (P_{1,1} \lor P_{2,2} \lor P_{3,1})$	
R_4:	$\neg B_{1,1}$	no breeze in 1,1
R_5:	$B_{2,1}$	breeze in 2,1

Q: Why are there are only 3 rows in the world of this KB ?

$B_{1,1}$	$B_{2,1}$	$P_{1,1}$	$P_{1,2}$	$P_{2,1}$	$P_{2,2}$	$P_{3,1}$	R1	R2	R3	R4	R5	KB
												f
												f
												...
												f
F	T	F	F	F	F	T	T	T	T	T	T	true
F	T	F	F	F	T	F	T	T	T	T	T	true
F	T	F	F	F	T	T	T	T	T	T	T	true
												f
												...
												f

Model-Checking Inference Algorithm

- enumerate ***all*** of the models consistent with the KB
- check that the desired proposition is true for ***every*** such model
- KB is true in a model if every proposition in KB is true in that model

R_1:	$\neg P_{1,1}$	no pit in 1,1
R_2:	$B_{1,1} \Leftrightarrow (P_{1,2} \vee P_{2,1})$	breeze iff pit in adjacent square
R_3:	$B_{2,1} \Leftrightarrow (P_{1,1} \vee P_{2,2} \vee P_{3,1})$	
R_4:	$\neg B_{1,1}$	no breeze in 1,1
R_5:	$B_{2,1}$	breeze in 2,1

Q: Why are there are only 3 rows in the world of this KB ?

$B_{1,1}$	$B_{2,1}$	$P_{1,1}$	$P_{1,2}$	$P_{2,1}$	$P_{2,2}$	$P_{3,1}$	R1	R2	R3	R4	R5	KB
												f
												f
												…
												f
F	T	F	F	F	F	T	T	T	T	T	T	true
F	T	F	F	F	T	F	T	T	T	T	T	true
F	T	F	F	F	T	T	T	T	T	T	T	true
												f
												…
												f

Example 1: Model-Checking

Question: In a world where p = True, q = False, and r = True, is the proposition (p $\vee$ q) $\rightarrow$ r entailed?

P	Q	R	P $\vee$ Q	P $\vee$ Q $\rightarrow$ R
T	F	T	T	T

Answer: Yes, by using the truth tables for the logical connectives

Example 2: Model-Checking

Question: In a world where p = True and ($q \lor r$) = True, is the proposition ($p \lor q$) → r entailed?

P	Q	R	Q ⋁ R	P ⋁ Q	P ⋁ Q → R
T	T	T	T	T	T
T	T	F	T	T	F
T	F	T	T	T	T

Answer: No, since the sentence is not true in all models consistent with the KB