TDDC17

Fö 5

Knowledge Representation I

Intuitions

Propositional Logic

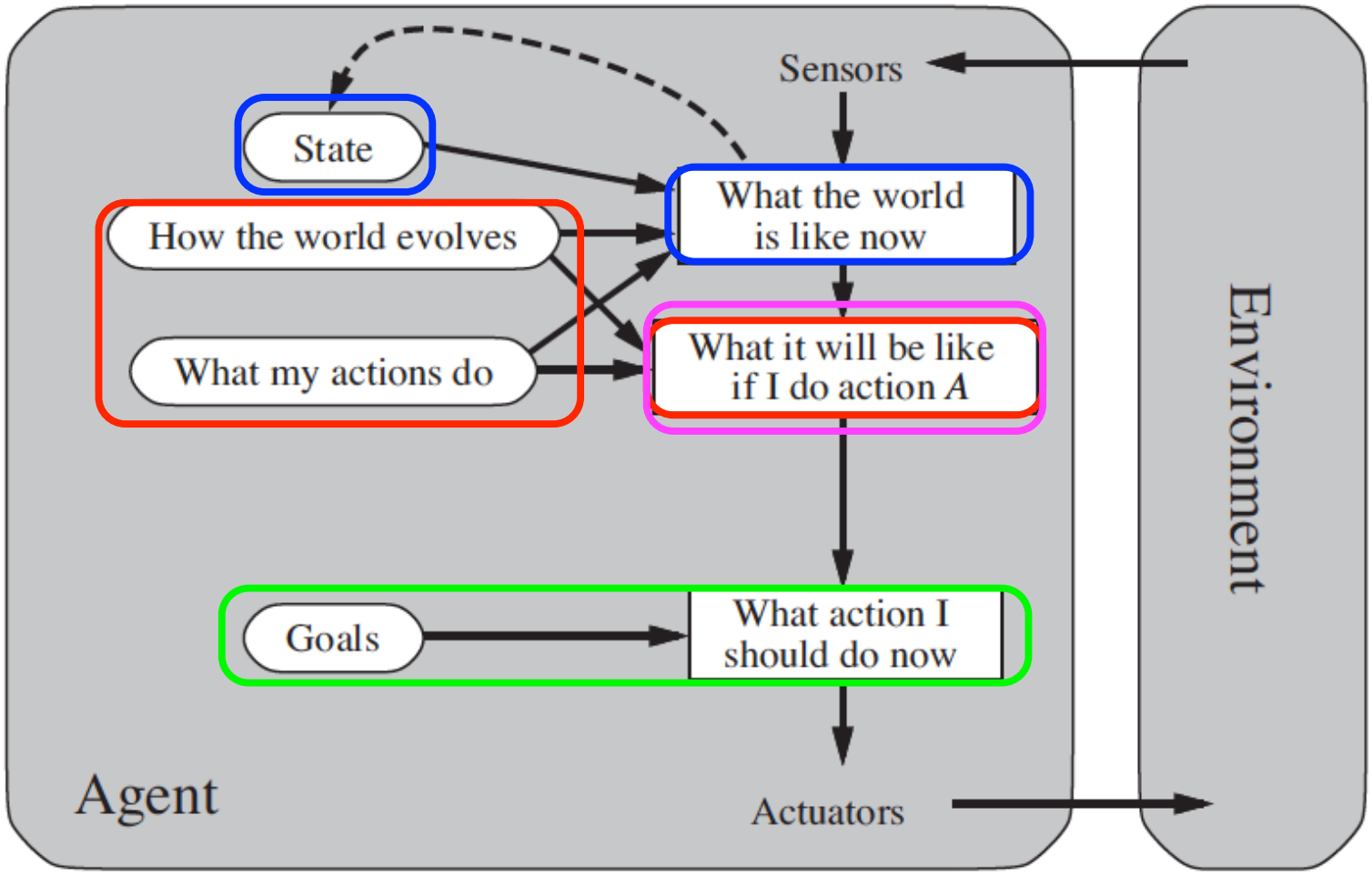
Propositional Theorem Proving

Model-based, Goal-Directed Agents

Keeps track of world-states

Has knowledge about the ways of the world

It is goal directed

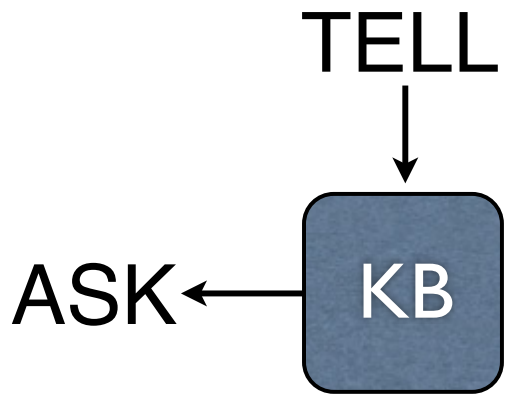


Anticipates by internal simulation/inference

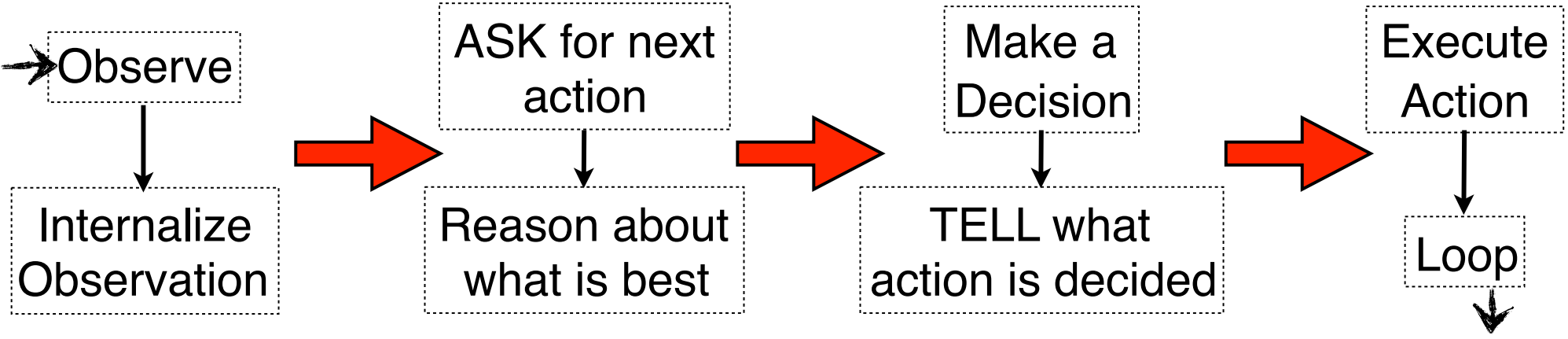
Generic Model-based Agent

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

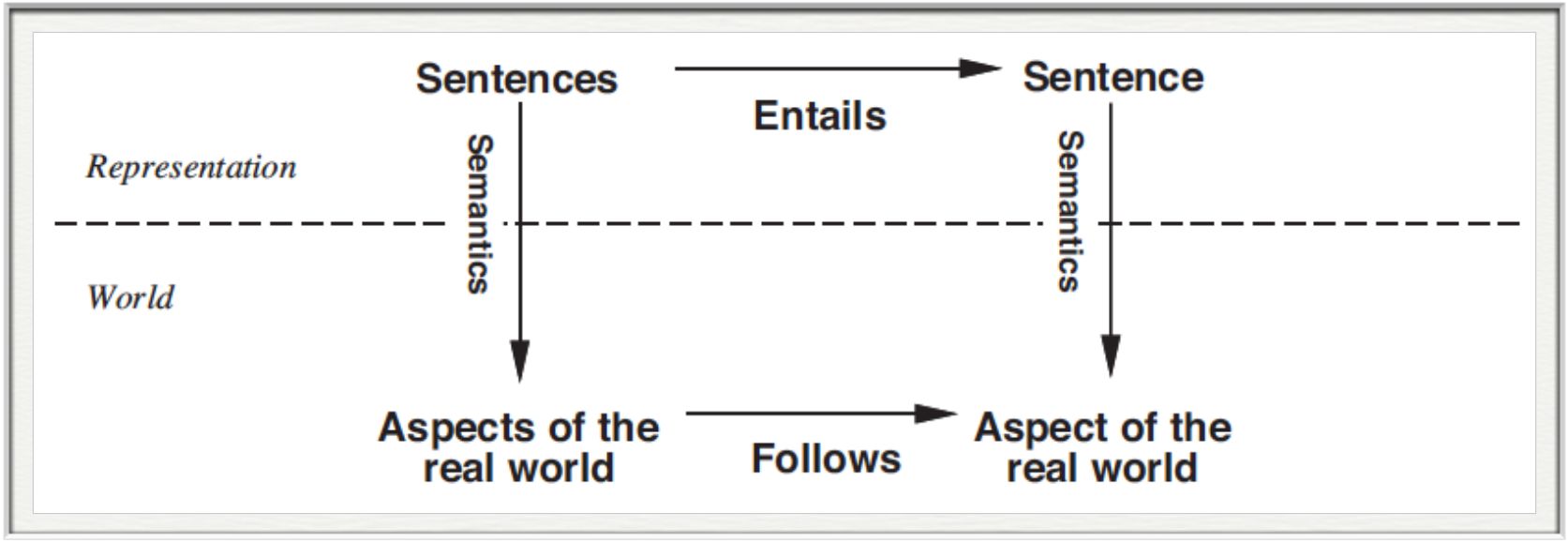
  TELL(KB, MAKE-PERCEPT-SENTENCE(*percept*, t))
  action $\leftarrow$ ASK(KB, MAKE-ACTION-QUERY(t))
  TELL(KB, MAKE-ACTION-SENTENCE(*action*, t))
  $t \leftarrow t + 1$
  return *action*



Advice Taker [McCarthy]



Knowledge Representation and Logic



What is our representation language?
How is it grounded causally in the world?

Truth preservation (soundness) guarantees fidelity of entailments to the world under the assumption that observation sentences (sensing) are correct, in Addition to background knowledge in the KB.

Knowledge Representation Hypothesis

Any mechanically embodied intelligent process will be comprised of structural ingredients that

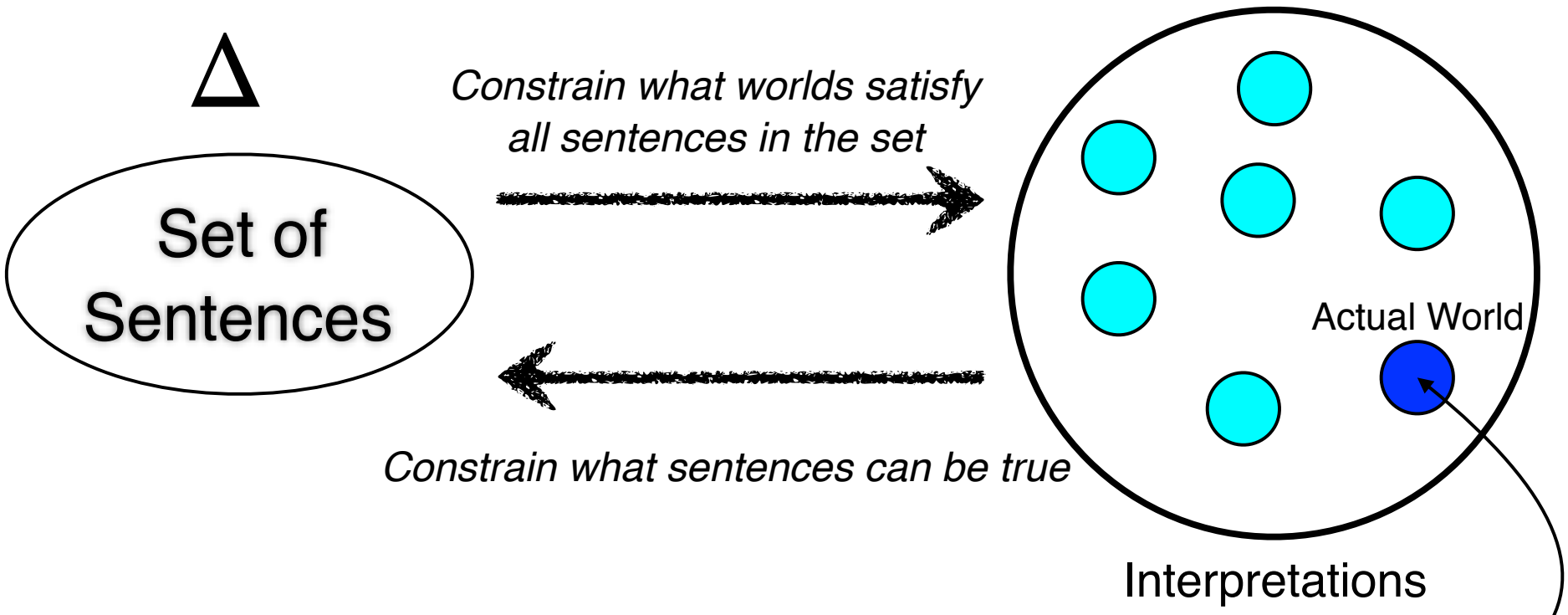
a) we as external observers naturally take to represent a propositional account of the knowledge that the overall process exhibits, and

b) independent of such external semantical attribution, play a formal but causal and essential role in engendering the behavior that manifests that knowledge. [Smith, 1982]

Knowledge as Constraints!

Knowledge Base

Possible Worlds



Syntax

$\Delta \vdash \alpha$

Inference

Semantics

$\Delta \models \alpha$

Entailment

Logic as a Representation Language

What is Logic?

Given a set of facts Δ taken to hold as true about the ”world” and given an assertion α about the ”world”, is there a good argument for believing that α holds based on the initial set of facts Δ?

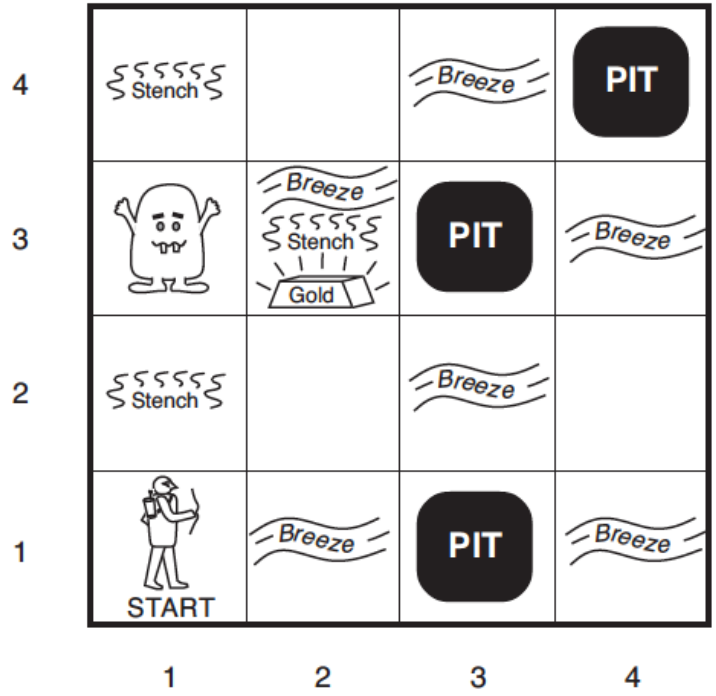
Logic in the general sense is about making distinctions between good arguments and bad arguments and the different criteria that may be used in making this distinction.
Deduction is one such criteria. (There are others!)

Logic in the more restricted sense is about the study of mathematical theories for formalizing the distinction between good/bad arguments and mechanizing ways to make these distinctions

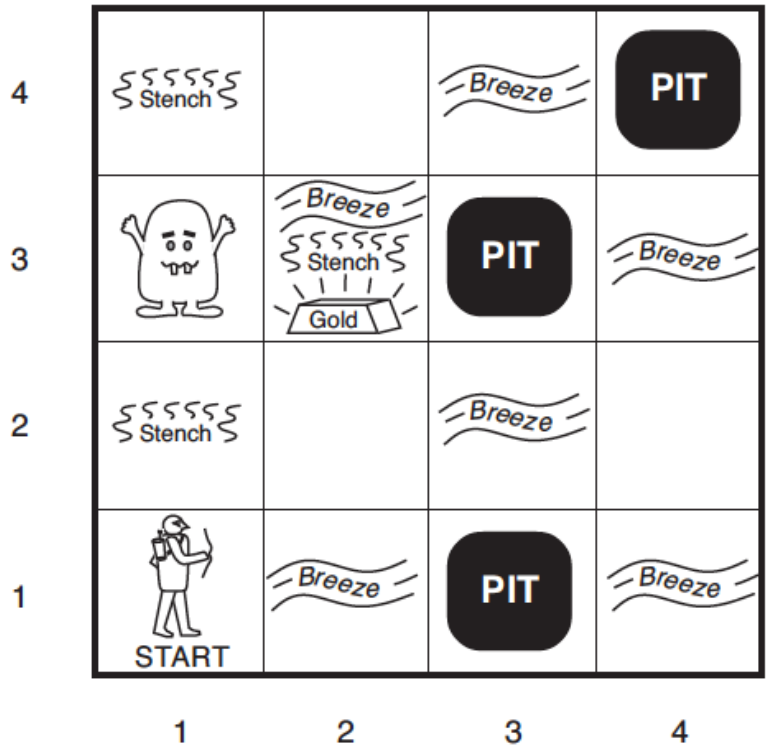
Wumpus World

The Wumpus World is a cave consisting of rooms connected by passageways. Lurking somewhere in the cave is a Wumpus, a beast that eats anyone who enters its room. The wumpus can be shot by an agent, but the agent only has one arrow. Some rooms contain bottomless pits that will trap anyone who wanders into such a room. There is also the possibility of finding a heap of gold.

This is the goal of anyone who enters the Wumpus World.



The Task Environment



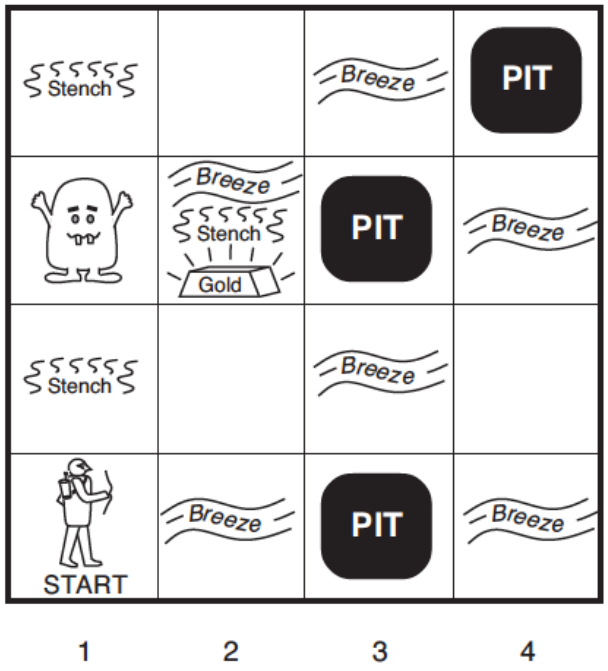
Performance Measure

- **+1000** for picking up gold,
- **-1000** for falling into a pit or being eaten by a Wumpus,
- **-1** for each action taken, and
- **–10** for using an arrow.

Environment

4x4 grid of rooms. Square **[1,1]** is initial state with agent facing to the right. Locations of gold, and wumpus are chosen randomly, with a uniform distribution, from all squares but **[1,1]**.
Each square other than **[1,1]** can contain a pit with probability 0.2.

The Task Environment



Actuators

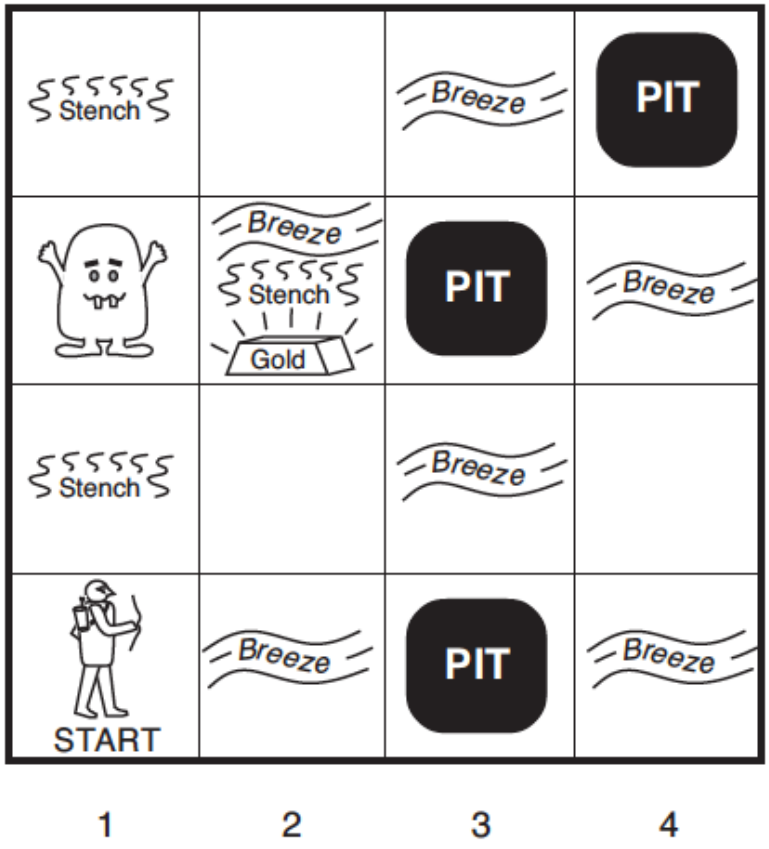
- The agent can **Move forward**, **Turn right** or **left** by 90 degrees
- **Grab** can be used to pick up an object in the same square as the agent.
- **Shoot** can be used to shoot the single arrow in a straight line until it hits something (Wumpus or a boundary wall)

Sensors

- A **stench** is perceived in the square containing a Wumpus or in those directly adjacent (not diagonal) to the Wumpus
- A **breeze** is perceived in a square directly adjacent to a pit
- A **glitter** is perceived in a square with gold in it.
- A **bump** is perceived if an agent walks into a wall.
- When the wumpus dies it emits a horrible **scream**.

An Example: Wumpus World

Reality



Agent A's View

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

Let's Reason!

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	3,2	4,2
1,1 A OK	2,1 OK	3,1	4,1

In $Rm_{1,1}$, there is no breeze or stench:

$$\neg B_{1,1} \wedge \neg S_{1,1}$$

Consequently, $Rm_{2,1}$ and $Rm_{1,2}$ are safe:

$$OK_{2,1} \wedge OK_{1,2}$$

KB:

$$\neg B_{1,1}, \neg S_{1,1}, OK_{1,1}, OK_{2,1}, OK_{1,2}$$

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

A moves to $Rm_{2,1}$ and feels a breeze: $B_{2,1}$

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

$$\neg B_{1,1}, \neg S_{1,1}, OK_{1,1}, OK_{2,1}, OK_{1,2}$$
$$B_{2,1}$$

KB

What can A conclude about pits in its vicinity?

Given , $B_{2,1}$ there may be a Pit in either $Rm_{2,2}$ or $Rm_{3,1}$: $P_{2,2} \vee P_{3,1}$

$$\neg B_{1,1}, \neg S_{1,1}, OK_{1,1}, OK_{2,1}, OK_{1,2}$$
$$B_{2,1} \quad P_{2,2} \vee P_{3,1}$$

KB

Partial Observability as disjunctive information

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 W?	2,3	3,3	4,3
1,2 **A** **S** **OK**	2,2 P? W?	3,2	4,2
1,1 **V** **OK**	2,1 **B** **V** **OK**	3,1 P?	4,1

Since there may be a Pit in either $Rm_{2,2}$ or $Rm_{3,1}$:

$$P_{2,2} \vee P_{3,1}$$

A decides to move back to $Rm_{1,1}$ and then to $Rm_{1,2}$.

A then senses a stench in $Rm_{1,2}$: $S_{1,2}$

$$\neg B_{1,1}, \neg S_{1,1}, OK_{1,1}, OK_{2,1}, OK_{1,2}$$
$$B_{2,1}, P_{2,2} \vee P_{3,1}, S_{1,2}$$

KB

What can A infer about the Wumpus and Pits in the vicinity?

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

$$\neg B_{1,1}, \neg S_{1,1}, OK_{1,1}, OK_{2,1}, OK_{1,2}$$
$$B_{2,1}, P_{2,2} \vee P_{3,1}, S_{1,2}$$

KB

Given $S_{1,2}$, there may be a Wumpus in either $Rm_{1,3}$ or $Rm_{2,2}$: $W_{1,3} \vee W_{2,2}$

$$\neg B_{1,1}, \neg S_{1,1}, OK_{1,1}, OK_{2,1}, OK_{1,2}$$
$$B_{2,1}, P_{2,2} \vee P_{3,1}, S_{1,2}, W_{1,3} \vee W_{2,2}$$

If there was a Wumpus in $Rm_{2,2}$, then A would have sensed a stench in $Rm_{2,1}$, but it didn't. So there is no Wumpus in $Rm_{2,2}$: $\neg W_{2,2}$

$$\neg B_{1,1}, \neg S_{1,1}, OK_{1,1}, OK_{2,1}, OK_{1,2}$$
$$B_{2,1}, P_{2,2} \vee P_{3,1}, S_{1,2}, W_{1,3} \vee W_{2,2}, \neg W_{2,2}$$

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

$$\neg B_{1,1}, \neg S_{1,1}, OK_{1,1}, OK_{2,1}, OK_{1,2}$$
$$B_{2,1}, P_{2,2} \vee P_{3,1}, S_{1,2}, W_{1,3} \vee W_{2,2}, \neg W_{2,2}$$

KB

1,4	2,4	3,4	4,4
1,3 W!	2,3	3,3	4,3
1,2 A S OK	2,2 P?	3,2	4,2
1,1 V OK	2,1 B V OK	3,1 P?	4,1

But $W_{1,3} \vee W_{2,2}$ and $\neg W_{2,2}$
imply $W_{1,3}$, so there is a Wumpus in $R_{1,3}$

$$\frac{W_{1,3} \vee W_{2,2} \quad \neg W_{2,2}}{W_{1,3}}$$

Resolution

$$\frac{\neg W_{2,2} \quad \neg W_{2,2} \rightarrow W_{1,3}}{W_{1,3}}$$

Modus Ponens

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 W!	2,3	3,3	4,3
1,2 A S OK	2,2	3,2	4,2
1,1 V OK	2,1 B V OK	3,1 P	4,1

$$\neg B_{1,1}, \neg S_{1,1}, OK_{1,1}, OK_{2,1}, OK_{1,2}$$
$$B_{2,1}, P_{2,2} \vee P_{3,1}, S_{1,2}, W_{1,3} \vee W_{2,2}, \neg W_{2,2}$$

KB

If there was a Pit in $Rm_{2,2}$, then A would have sensed a breeze in $Rm_{1,2}$, but it didn't. So there is no Pit in $Rm_{2,2}$: $\neg P_{2,2}$

$$\neg B_{1,1}, \neg S_{1,1}, OK_{1,1}, OK_{2,1}, OK_{1,2}$$
$$B_{2,1}, P_{2,2} \vee P_{3,1}, S_{1,2}, W_{1,3} \vee W_{2,2}, \neg W_{2,2}$$
$$\neg P_{2,2}$$

But $P_{2,2} \vee P_{3,1}$ and $\neg P_{2,2}$ imply $P_{3,1}$, so there is a Pit in $R_{3,1}$

$$\neg B_{1,1}, \neg S_{1,1}, OK_{1,1}, OK_{2,1}, OK_{1,2}$$
$$B_{2,1}, P_{2,2} \vee P_{3,1}, S_{1,2}, W_{1,3} \vee W_{2,2}, \neg W_{2,2}$$
$$\neg P_{2,2}, P_{3,1}$$

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

$$\neg B_{1,1}, \neg S_{1,1}, OK_{1,1}, OK_{2,1}, OK_{1,2}$$
$$B_{2,1}, P_{2,2} \vee P_{3,1}, S_{1,2}, W_{1,3} \vee W_{2,2}, \neg W_{2,2}$$
$$\neg P_{2,2}, P_{3,1}, OK_{2,2}$$

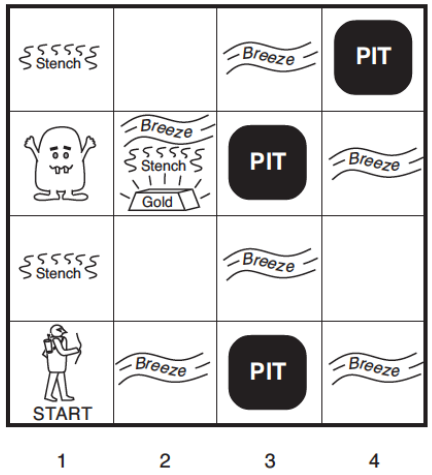
KB

1,4	2,4	3,4	4,4
1,3 W	2,3 OK	3,3	4,3
1,2 S V OK	2,2 A	3,2 OK	4,2
1,1 V OK	2,1 B V OK	3,1 P	4,1

Since there is no stench or breeze in $Rm_{2,2}$, Both $Rm_{2,3}$ and $Rm_{3,2}$ are ok to move to: $OK_{2,3} \wedge OK_{3,2}$

$$\neg B_{1,1}, \neg S_{1,1}, OK_{1,1}, OK_{2,1}, OK_{1,2}$$
$$B_{2,1}, P_{2,2} \vee P_{3,1}, S_{1,2}, W_{1,3} \vee W_{2,2}, \neg W_{2,2}$$
$$\neg P_{2,2}, P_{3,1}, OK_{2,2}, OK_{2,3}, OK_{3,2}$$

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



$$\neg B_{1,1}, \neg S_{1,1}, OK_{1,1}, OK_{2,1}, OK_{1,2}$$
$$B_{2,1}, P_{2,2} \vee P_{3,1}, S_{1,2}, W_{1,3} \vee W_{2,2}, \neg W_{2,2}$$
$$\neg P_{2,2}, P_{3,1}, OK_{2,2}, OK_{2,3}, OK_{3,2}$$

KB

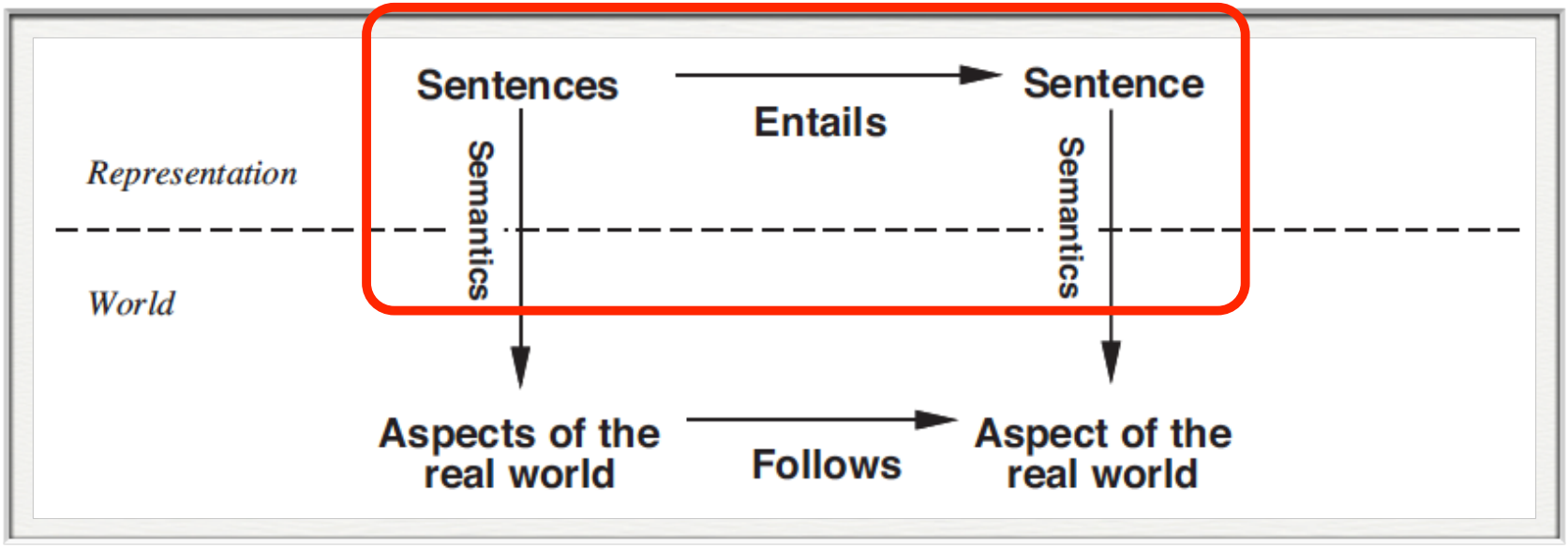
A chooses to move to $Rm_{2,3}$ and then senses a breeze, stench, and gold:

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 OK	4,2
1,1 V OK	2,1 B V OK	3,1 P!	4,1

$$\neg B_{1,1}, \neg S_{1,1}, OK_{1,1}, OK_{2,1}, OK_{1,2}$$
$$B_{2,1}, P_{2,2} \vee P_{3,1}, S_{1,2}, W_{1,3} \vee W_{2,2}, \neg W_{2,2}$$
$$\neg P_{2,2}, P_{3,1}, OK_{2,2}, OK_{2,3}, OK_{3,2}$$
$$B_{2,3}, S_{2,3}, G_{2,3}$$

A picks up the gold, generates a motion plan to get back to [1,1] and wins the game!

Logic as a Representation Language



Propositional Logic

First-Order Logic

Default/ Nonmonotonic Reasoning

Resolution Theorem

Resolution Theorem Proving