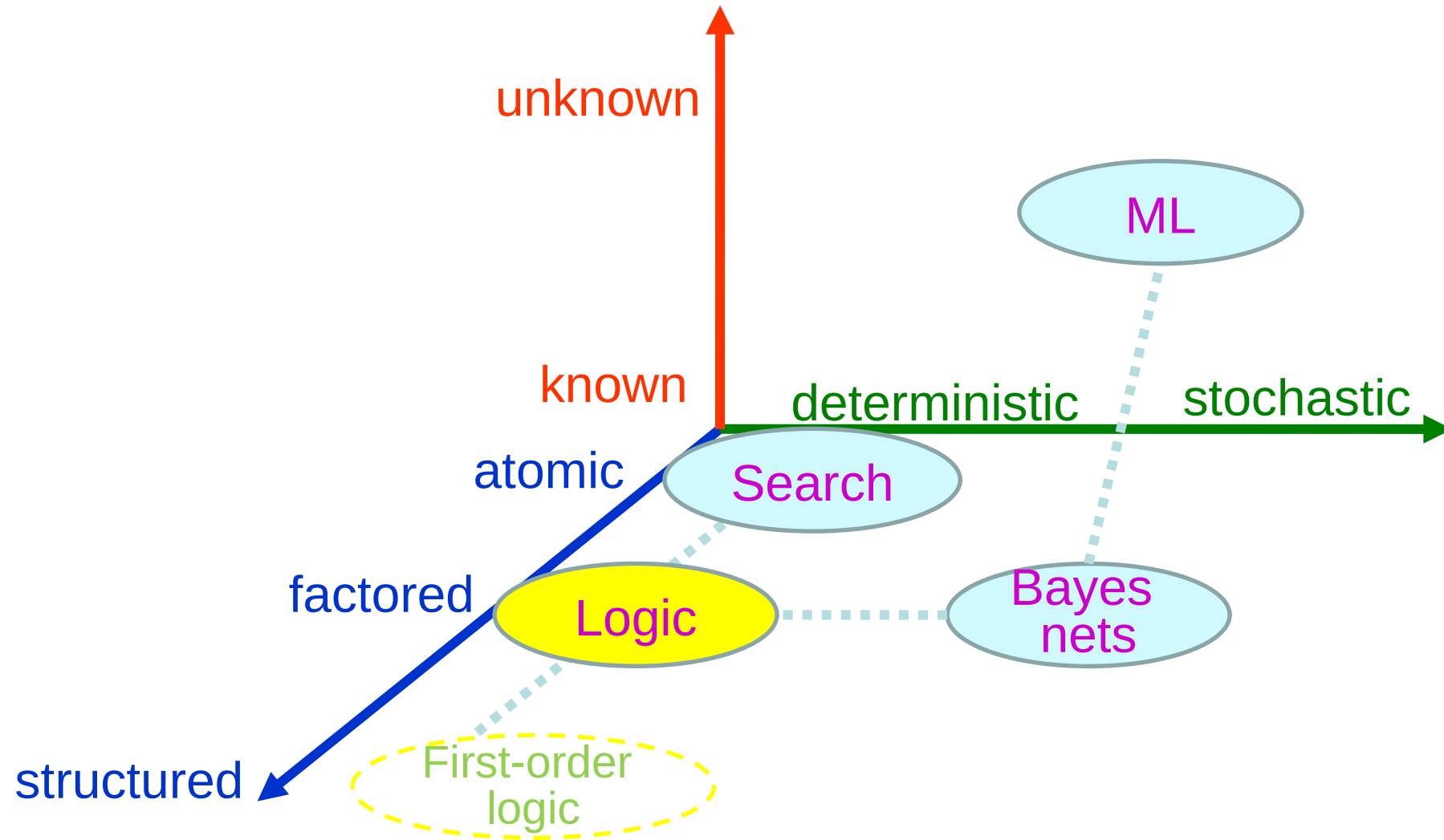


Artificial Intelligence

Section 7 : **Logic Agent**

Outline of the course



Outline

- 7.1 基于知识的智能体
- 7.2 Wumpus world
- 7.3 知识的逻辑表示和推理
- 7.4 命题逻辑：一种简单的逻辑
- 7.5 命题逻辑定理证明
 - 归结原理

基于知识的智能体

- Agents 通过感知、学习和语言来获取以下知识
 - 行动的影响（“转移模型”）
 - 世界如何影响传感器（“传感器模型”）
 - 世界当前状态
- 可以追踪一个部分可见的世界
- 能够制定实现目标的计划
- ...

Knowledge base

Inference engine

Domain-specific facts

Generic code

基于知识的智能体

- 知识库 ***KB*** = set of sentences in a formal language
- 构建智能体的声明性方法：
 - **Tell** : Agent 告诉知识库需要知道的一切
 - **Ask** : Agent 询问知识库要执行什么动作？（回答应该是基于知识库的）
- 这两个过程都可能涉及推理，即从原有语句中推导出新的语句。还可以使得 Agent 具有学习能力

A knowledge-based agent

function KB-AGENT(percept) **returns** an action

persistent: **KB**, a knowledge base

t, an integer, initially 0

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

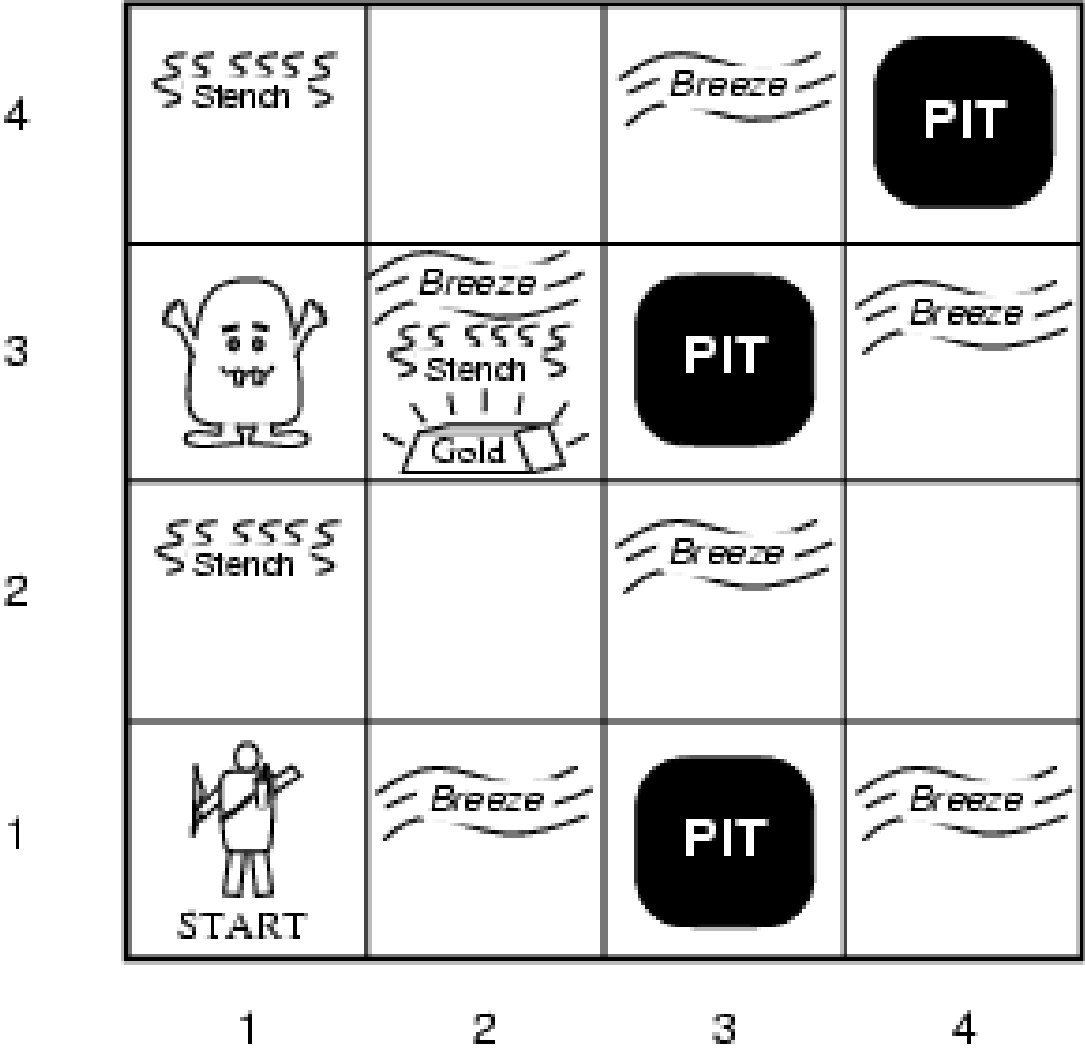
return action

- 智能体应该能够：
 - 表示状态、行为等等
 - 融入新的感知
 - 更新对周围世界的内部表示
 - 推断周围世界的隐藏属性
 - 推断出正确的行为

Outline

- 7.1 基于知识的智能体
- 7.2 Wumpus world
- 7.3 知识的逻辑表示和推理
- 7.4 命题逻辑：一种简单的逻辑
- 7.5 命题逻辑定理证明
 - 归结原理

Wumpus World



Wumpus 世界是由多个房间组成并连接起来的山洞，在洞穴的某处隐藏着一只 **Wumpus**，它会吃掉进入它房间的任何人。

Agent 可以射杀 Wumpus，但是只有一支箭。

某些房间是无底洞，任何人漫游到这些房间都会被无底洞吞噬。

生活在该环境的唯一希望是存在发现一堆金子的可能性。

Wumpus World PEAS description

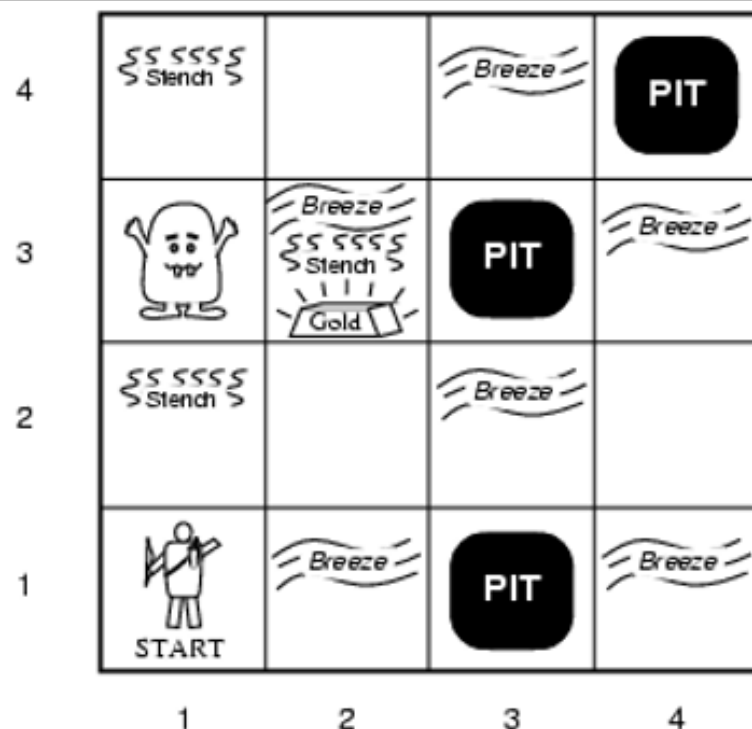
- **Performance measure**

- 黄金 +1000, 死亡 -1000
- 每步 -1, 使用箭 -10

- **Environment**

- 与怪兽相邻的方块是臭的
- 无底黑洞旁边的广场有微风
- 亮光和金子在同一个方块中
- 如果你面对怪兽射击，会杀死它
- 射击会消耗唯一的箭
- 如果和金子在同一个方块中，抓取可以获得金子
- 如果和金子在同一个方块中，放手可以扔掉金子

- **Sensors:** 臭味发射器，风扇，发光器，撞击器，尖叫声
- **Actuators:** 向左转，向右转，向前，抓取，放手，射击

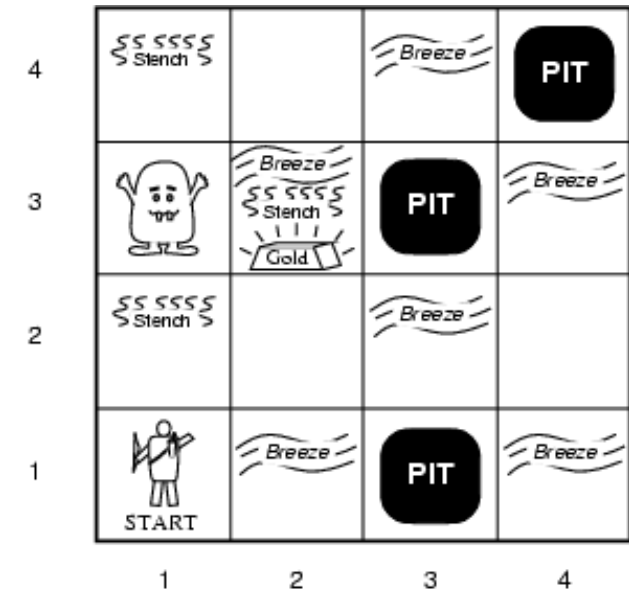


Exploring a wumpus world

首先，定义感知（状态描述）： [Stench, Breeze, Glitter, Crash, Yell]

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



The first percept is [None, None, None, None, None]

conclude [1,2] and [2,1] are OK

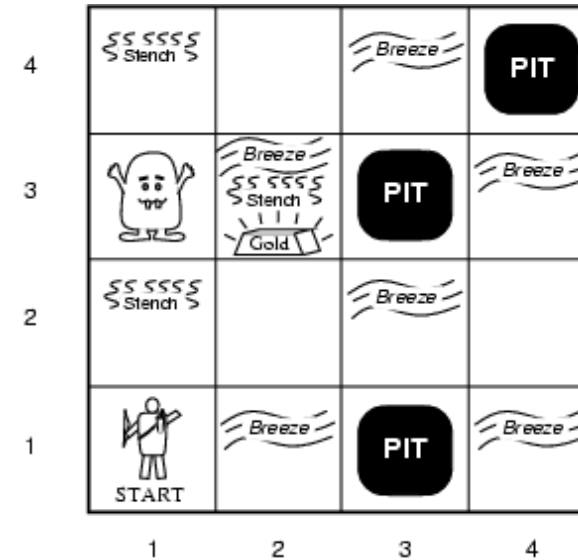
Action : [2,1]

Exploring a wumpus world

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

percept [None,Breeze,None,None,None] in [2,1]

Action : [1,2]



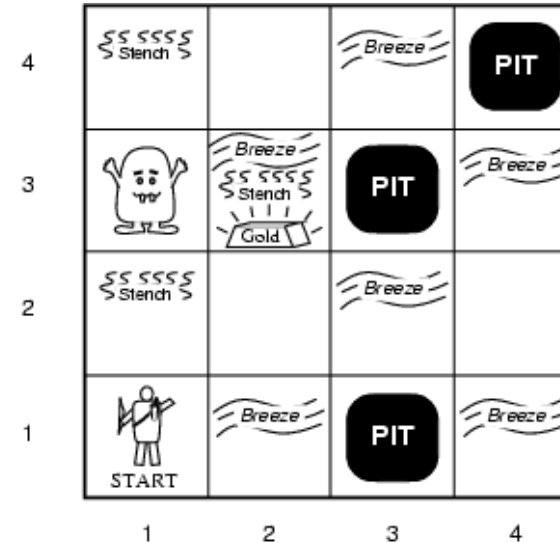
there must be a pit in [2,2] or [3,1] or both

Exploring a wumpus world

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

percept [Stench,None,None,None,None] in [1,2]

Action : [2,2]



wumpus is in [1,3].

the lack of a breeze in [1,2] implies no pit in [2,2]

neither a pit nor a wumpus in [2,2], so it is OK

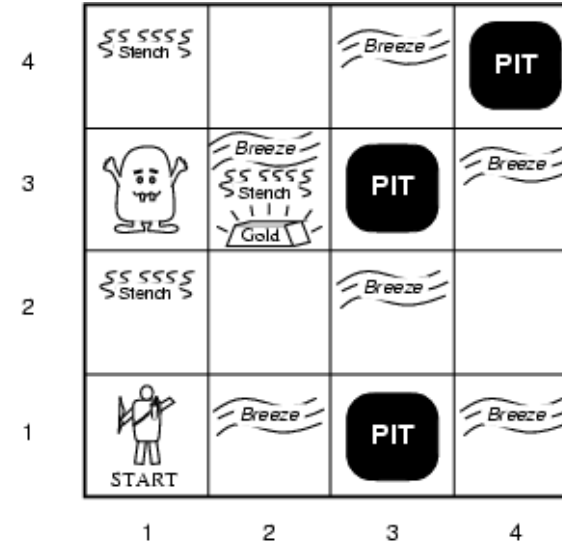
pit in [3,1]

Exploring a wumpus world

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

percept [None, None, None, None, None] in [2,2]

Action : [2,3]

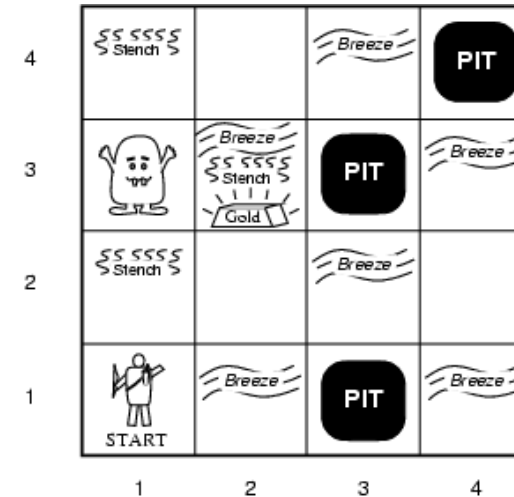


[2,3] is OK

[3,2] is OK

Exploring a wumpus world

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



percept [Stench,Breeze,Glitter,None,None] In [2,3]

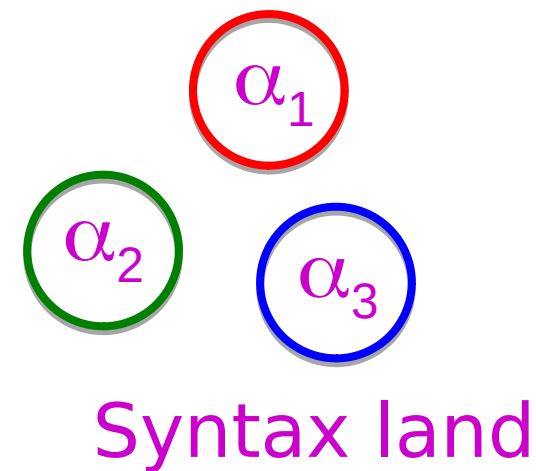
the agent detects a glitter, so it should grab the gold and then return home.

Outline

- 7.1 基于知识的智能体
- 7.2 Wumpus world
- 7.3 知识的逻辑表示和推理
- 7.4 命题逻辑：一种简单的逻辑
- 7.5 命题逻辑定理证明
 - 归结原理

什么是逻辑？

- 逻辑：处理事实的形式化系统，目的得到正确结论
 - 用来区分真和假的工具 – Averroes (12 世纪)
- 语法：构造知识库有效句子的规则
 - 例如， $x + 2 \geq y$ 是有效的算术语句， $\geq x2y +$ 就不是
- 语义：句子的“含义”，逻辑语句和真实世界之间关系
 - 具体而言，语义决定了句子的真假



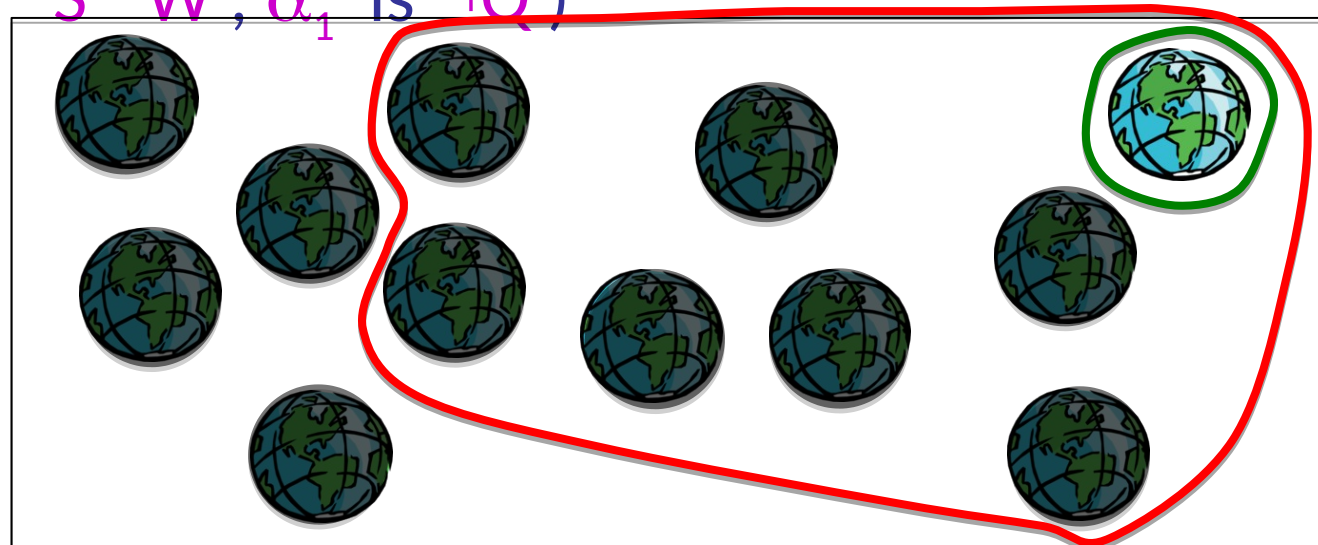
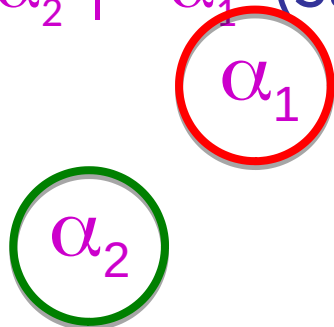
Semantics land

逻辑推理

- 蕴含：意味着一样东西是从另一样东西中派生出来的
- KB 蕴含句子 α 当且仅当 α 在所有 KB 为真的世界里为真

$$KB \models \alpha$$

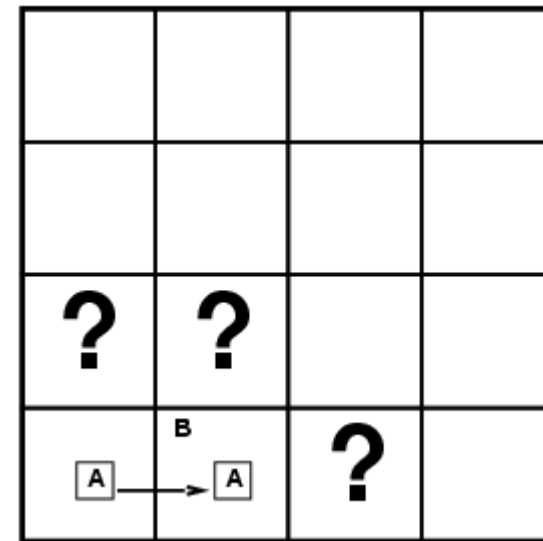
- 例如, $\alpha_2 \models \alpha_1$ (Say α_2 is $\neg Q \wedge R \wedge S \wedge W$; α_1 is $\neg Q$)



蕴含是基于语义的句子之间的关系

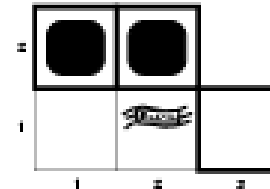
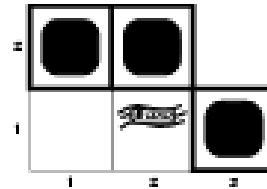
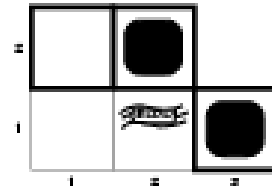
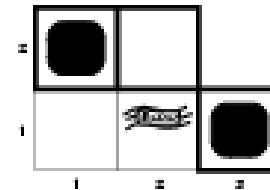
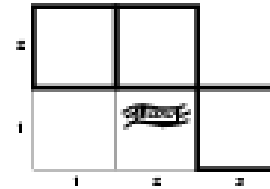
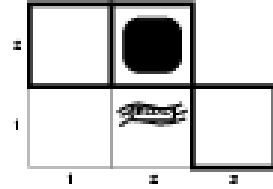
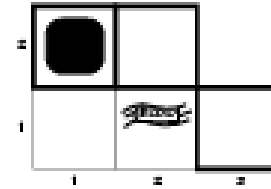
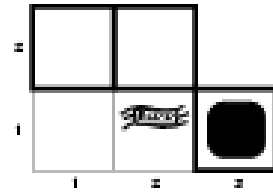
在怪兽世界中的蕴含规则

- 检测到 [1,1] 中没有东西，向右移动，[2,1] 中有**微风**
- 问题：是否有无底洞？
- 知识库 KB ： **8 个可能的情况**（布尔选择）

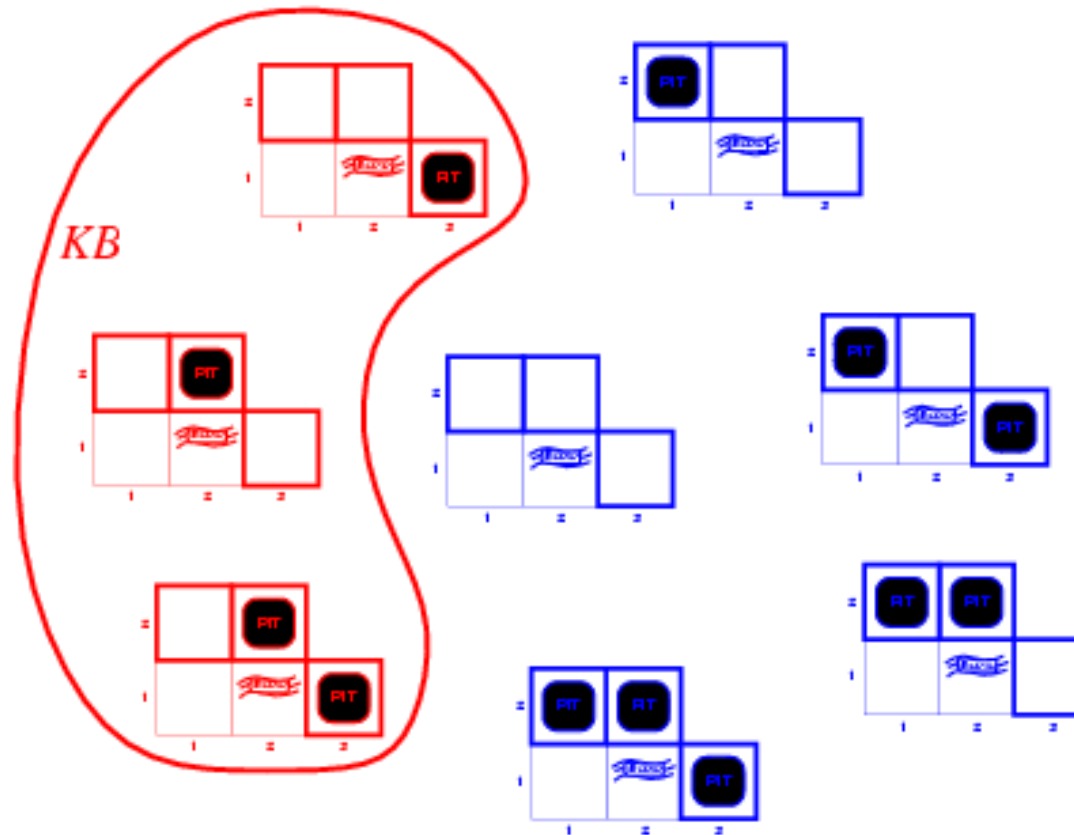


Wumpus models

?	?		
<div>A</div>	<div>B</div>	?	

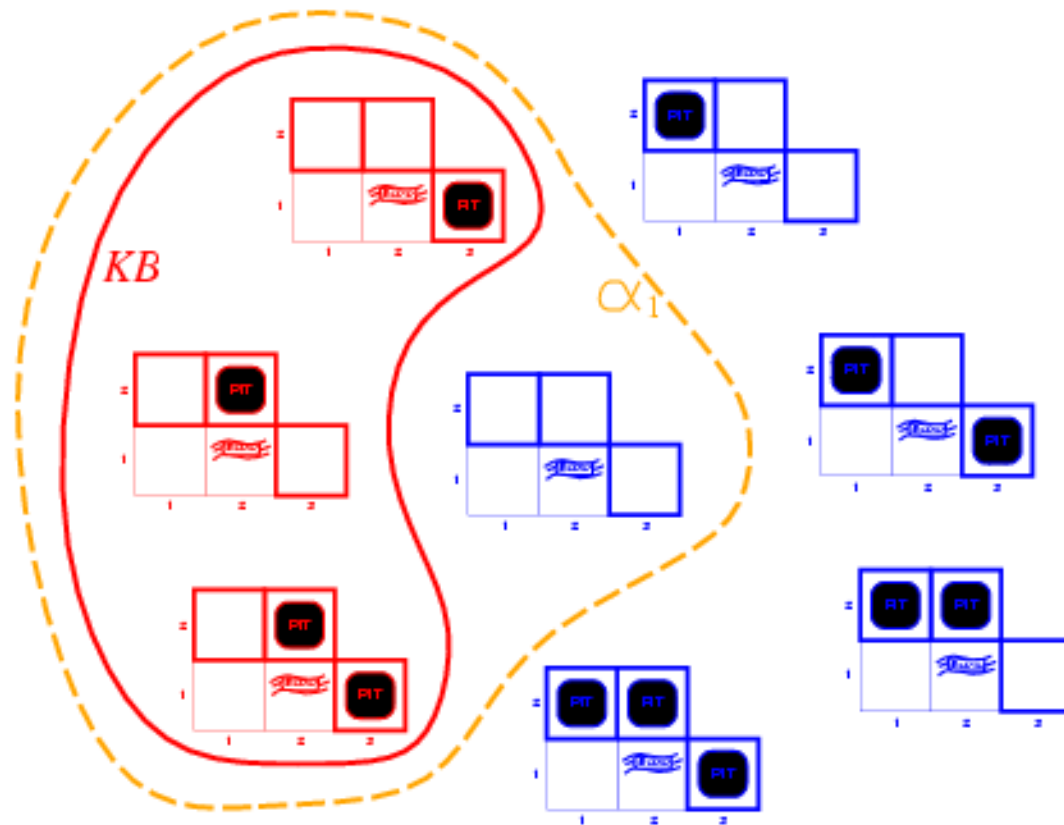


Wumpus models



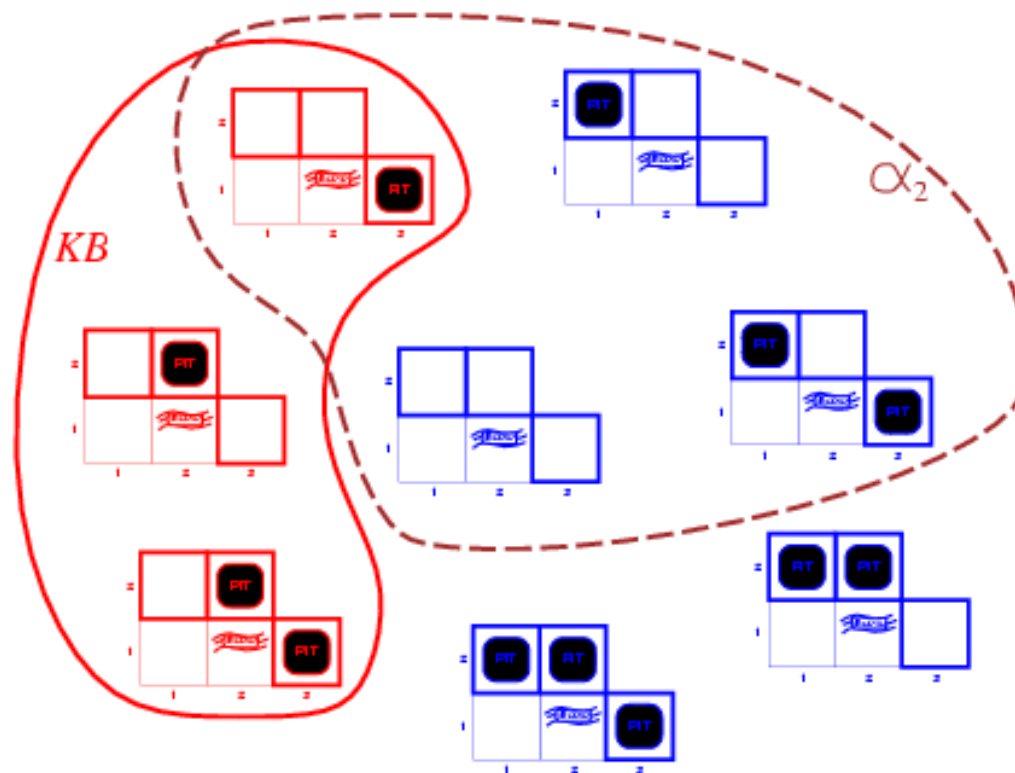
- $KB = \text{wumpus-world rules} + \text{observations}$

Wumpus models



- KB = wumpus-world rules + observations
- α_1 = "[1,2] is safe", 模型验证证明 $KB \models \alpha_1$

Wumpus models



- KB = wumpus-world rules + observations
- α_2 = "[2,2] is safe", $KB \models \alpha_2$ 是否成立？
- $KB \not\models \alpha_2$: Agent 无法得出 [2,2] 中没有无底洞

Outline

- 7.1 基于知识的智能体
- 7.2 Wumpus world
- 7.3 知识的逻辑表示和推理
- 7.4 命题逻辑：一种简单的逻辑
- 7.5 命题逻辑定理证明
 - 归结原理

命题

- **命题**：能判断真假的陈述句。是具有唯一真值的陈述句或判断结果唯一的陈述句
- **命题的真值**：判断的结果，真值的取值真与假二者取一
- **真命题**：真值为真的命题
- **假命题**：真值为假的命题
 - 注意：感叹句、祈使句、疑问句都不是命题
 - 陈述句中的悖论以及判断结果不唯一确定的也不是命题

命题逻辑：语法

	原子命题	复合命题
$Sentence \rightarrow$	$AtomicSentence$	$ComplexSentence$
$AtomicSentence \rightarrow$	$True \mid False \mid P \mid Q \mid R \mid \dots$	
$ComplexSentence \rightarrow$	$(Sentence) \mid [Sentence]$	
	$\neg Sentence$	取非
	$Sentence \wedge Sentence$	合取
	$Sentence \vee Sentence$	析取
	$Sentence \Rightarrow Sentence$	蕴含
	$Sentence \Leftrightarrow Sentence$	双向蕴含

OPERATOR PRECEDENCE : $\neg, \wedge, \vee, \Rightarrow, \Leftrightarrow$

Figure 7.7 A BNF (Backus–Naur Form) grammar of sentences in propositional logic, along with operator precedences, from highest to lowest.

$\neg, \wedge, \vee, \Rightarrow, \Leftrightarrow$

逻辑连接词优先级从高到低

命题逻辑：语义

复合命题 有 5 条规则：(P and Q 是模型 m 中的任意子句)

$\neg P$ is true iff P is false in m

$P \wedge Q$ is true iff both P is **and** Q are true in m

$P \vee Q$ is true iff either P **or** Q is true in m

$P \Rightarrow Q$ is true iff **unless** P is true and Q is false in m

$P \Leftrightarrow Q$ is true iff P and Q are **both** true or **both** false in m

语义定义了用于判定**模型**（可能世界）中语句真值的规则

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

Figure 7.8 Truth tables for the five logical connectives. To use the table to compute, for example, the value of $P \vee Q$ when P is true and Q is false, first look on the left for the row where P is *true* and Q is *false* (the third row). Then look in that row under the $P \vee Q$ column to see the result: *true*.

⇒ 的真值可能不太符合直觉： $P \text{ implies } Q$ 或者 if P then Q

例如，5 是偶数 \square Sam 很聪明 is true or false ?

⇒ 为假的唯一条件： P 为真而 Q 为假