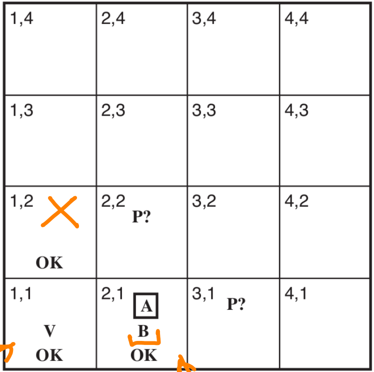
Wupus-world inference example

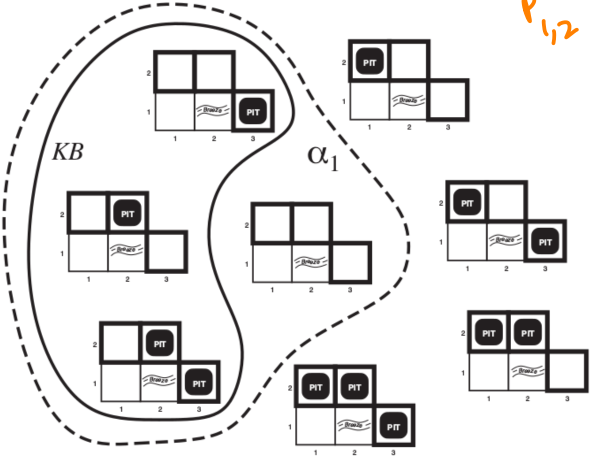


{x, y}

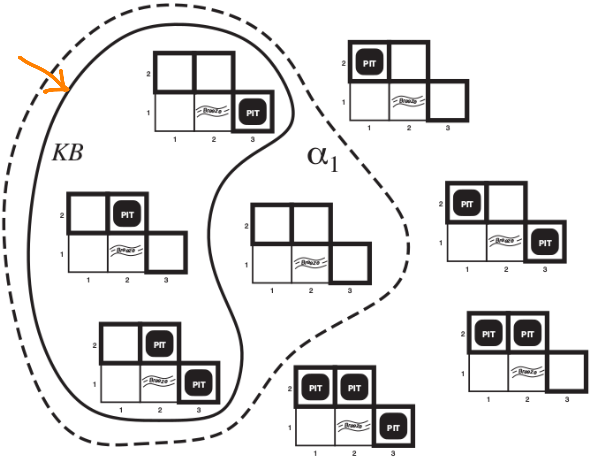
- KB contains agent's percepts (in the first 2 steps) and rules of the Wumpus world
- Agent wants to know whether pit is present in [1,2] and [2,2].
- $\alpha_1 \equiv$ "No pit in [1,2]"
- $\alpha_2 \equiv$ "No pit in [2,2]"
- $KB \models \alpha_1$?
- $KB \models \alpha_2$?

$KB \models \alpha_1$?

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

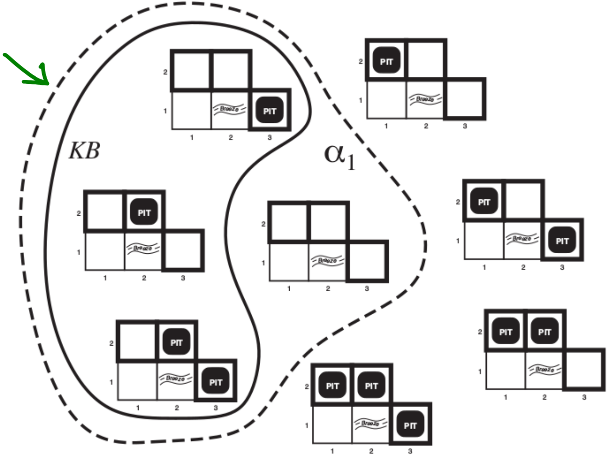


$KB \models \alpha_1$?



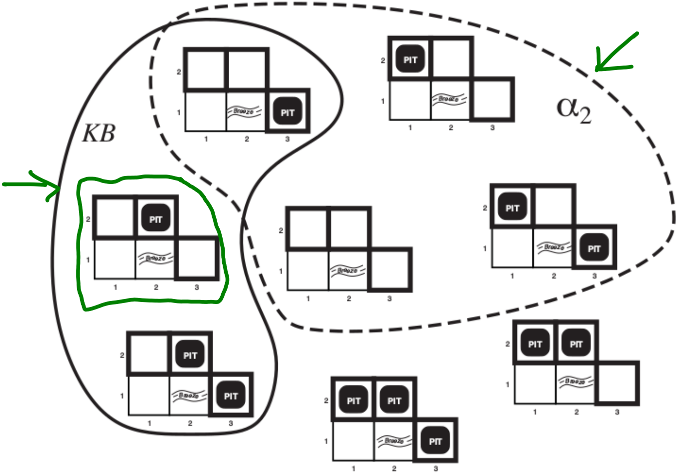
Models in which KB is True: $\{P_{1,2} = F, P_{2,2} = T, P_{3,1} = T\}$, etc.

$KB \models \alpha_1$?

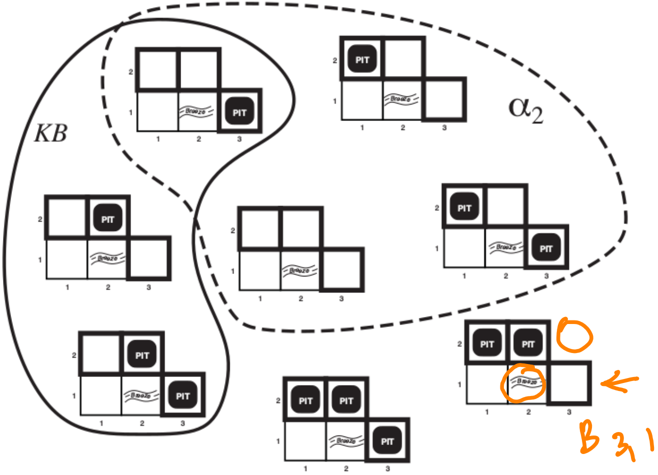


Models in which KB is True: $\{P_{1,2} = F, P_{2,2} = T, P_{3,1} = T\}$, etc.
Models in which α_1 is True: $\{P_{1,2} = F, P_{2,2} = F, P_{3,1} = F\}$, etc.

$KB \models \alpha_2$?

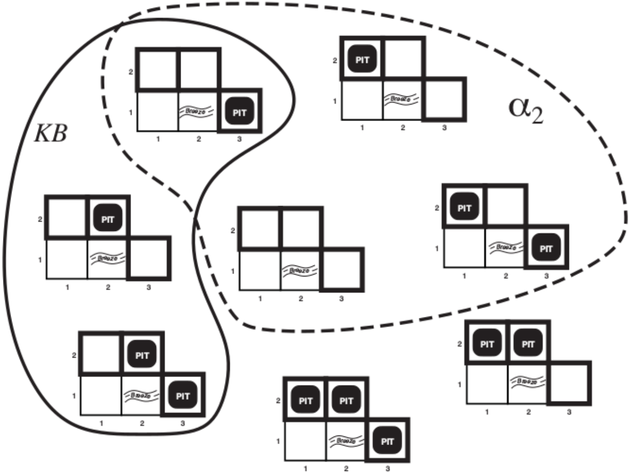


$KB \models \alpha_2$?



Models in which KB is True: $\{P_{1,2} = F, P_{2,2} = T, P_{3,1} = T\}$, etc.

$KB \models \alpha_2$?



Models in which KB is True: $\{P_{1,2} = F, P_{2,2} = T, P_{3,1} = T\}$, etc.
Models in which α_2 is True: $\{P_{1,2} = T, P_{2,2} = F, P_{3,1} = T\}$, etc.

Logic

1. Model
2. $\underbrace{M(\alpha_1)}$

$m(\alpha_1)$

$\alpha \models \beta$

3. Entailment ($\alpha \models \beta$)
4. $\alpha \models \beta$ iff $\underbrace{M(\alpha)} \subseteq \underbrace{M(\beta)}$

Propositional Logic

$$
\begin{array}{rcl}
Sentence & \rightarrow & AtomicSentence \mid \; ComplexSentence \\
AtomicSentence & \rightarrow & True \mid \; False \mid \; P \mid \; Q \mid \; R \mid \; \ldots \\
ComplexSentence & \rightarrow & (\; Sentence \;) \mid \; [\; Sentence \;] \\
& \mid & \neg\; Sentence \\
& \mid & Sentence \land Sentence \\
& \mid & Sentence \lor Sentence \\
& \mid & Sentence \;\; \Rightarrow \;\; Sentence \\
& \mid & Sentence \;\; \Leftrightarrow \;\; Sentence \\
\\
\textsc{Operator Precedence} & : & \neg, \land, \lor, \Rightarrow, \Leftrightarrow
\end{array}
$$

Propositional Logic Connectives

NEGATION ¬ (not). A sentence such as $\neg W_{1,3}$ is called the **negation** of $W_{1,3}$. A **literal** is either an
LITERAL atomic sentence (a **positive literal**) or a negated atomic sentence (a **negative literal**).

CONJUNCTION ∧ (and). A sentence whose main connective is $\wedge$, such as $W_{1,3} \wedge P_{3,1}$, is called a **conjunction**; its parts are the **conjuncts**.

DISJUNCTION ∨ (or). A sentence using $\vee$, such as $(W_{1,3} \wedge P_{3,1}) \vee W_{2,2}$, is a **disjunction** of the **disjuncts** $(W_{1,3} \wedge P_{3,1})$ and $W_{2,2}$.

IMPLICATION ⇒ (implies). A sentence such as $(W_{1,3} \wedge P_{3,1}) \Rightarrow \neg W_{2,2}$ is called an **implication** (or conditional).
PREMISE Its **premise** or **antecedent** is $(W_{1,3} \wedge P_{3,1})$, and its **conclusion** or **consequent**
CONCLUSION is $\neg W_{2,2}$.
RULES

BICONDITIONAL ⇔ (if and only if). The sentence $W_{1,3} \Leftrightarrow \neg W_{2,2}$ is a **biconditional**.

Propositional Logic Connectives

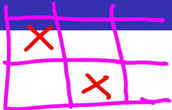
NEGATION LITERAL	$\neg$ (not). A sentence such as $\neg W_{1,3}$ is called the **negation** of $W_{1,3}$. A **literal** is either an atomic sentence (a **positive literal**) or a negated atomic sentence (a **negative literal**).
CONJUNCTION	$\wedge$ (and). A sentence whose main connective is $\wedge$, such as $W_{1,3} \wedge P_{3,1}$, is called a **conjunction**; its parts are the **conjuncts**.
DISJUNCTION	$\vee$ (or). A sentence using $\vee$, such as $(W_{1,3} \wedge P_{3,1}) \vee W_{2,2}$, is a **disjunction** of the **disjuncts** $(W_{1,3} \wedge P_{3,1})$ and $W_{2,2}$.
IMPLICATION PREMISE CONCLUSION RULES	$\Rightarrow$ (implies). A sentence such as $(W_{1,3} \wedge P_{3,1}) \Rightarrow \neg W_{2,2}$ is called an **implication** (or conditional). Its **premise** or **antecedent** is $(W_{1,3} \wedge P_{3,1})$, and its **conclusion** or **consequent** is $\neg W_{2,2}$.
BICONDITIONAL	$\Leftrightarrow$ (if and only if). The sentence $W_{1,3} \Leftrightarrow \neg W_{2,2}$ is a **biconditional**.

Semantics of PL

Simple Knowledge Base

$P_{x,y}$ is true if there is a pit in $[x, y]$.
$W_{x,y}$ is true if there is a wumpus in $[x, y]$, dead or alive.
$B_{x,y}$ is true if the agent perceives a breeze in $[x, y]$.
$S_{x,y}$ is true if the agent perceives a stench in $[x, y]$.

Simple Knowledge Base



$P_{x,y}$ is true if there is a pit in $[x, y]$.
$W_{x,y}$ is true if there is a wumpus in $[x, y]$, dead or alive.
$B_{x,y}$ is true if the agent perceives a breeze in $[x, y]$.
$S_{x,y}$ is true if the agent perceives a stench in $[x, y]$.

KB:

$$
\begin{aligned}
R_1 &: \quad \neg P_{1,1} \\
R_2 &: \quad B_{1,1} \Leftrightarrow (P_{1,2} \vee P_{2,1}) \\
R_3 &: \quad B_{2,1} \Leftrightarrow (P_{1,1} \vee P_{2,2} \vee P_{3,1}) \\
R_4 &: \quad \neg B_{1,1} \\
R_5 &: \quad B_{2,1}
\end{aligned}
$$

$R_6 : \quad B_{1,3}$

Simple Knowledge Base

7

2

Does KB $\models P_{1,2}$?

Simple Knowledge Base

$KB \models \alpha$

Does KB $\models P_{1,2}$?
Does KB $\models P_{2,2}$?

Model Checking

$KB \nvDash P_{1,2}$

$B_{1,1}$	$B_{2,1}$	$P_{1,1}$	$P_{1,2}$	$P_{2,1}$	$P_{2,2}$	$P_{3,1}$	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*	<u>*true*</u>
false	*true*	*false*	*false*	*false*	*true*	*false*	*true*	*true*	*true*	*true*	*true*	<u>*true*</u>
false	*true*	*false*	*false*	*false*	*true*	*true*	*true*	*true*	*true*	*true*	*true*	<u>*true*</u>
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*

Figure 7.9 A truth table constructed for the knowledge base given in the text. *KB* is true if R_1 through R_5 are true, which occurs in just 3 of the 128 rows (the ones underlined in the right-hand column). In all 3 rows, $P_{1,2}$ is false, so there is no pit in [1,2]. On the other hand, there might (or might not) be a pit in [2,2].

$KB \nvDash P_{2,2}$

$KB \nvDash \neg P_{2,2}$

$KB \vDash \neg P_{1,2}$

Logical inference algorithms

- Model checking

Logical inference algorithms

- Model checking
- Inference algorithm $(KB \vdash_i \alpha)$ (algorithm i derives α from KB)

Logical inference algorithms

- Model checking
- Inference algorithm ($KB \vdash_i \alpha$) (algorithm i derives α from KB)
- Soundness

$m(KB) \subseteq m(\alpha)$

$KB \vdash_i \alpha \quad \Rightarrow \quad KB \models \alpha$

$\Leftarrow$

Logical inference algorithms

- Model checking
- Inference algorithm ($KB \vdash_i \alpha$) (algorithm i derives α from KB)
- Soundness
- Completeness

Logical equivalences

$$
\begin{aligned}
(\alpha \wedge \beta) &\equiv (\beta \wedge \alpha) && \text{commutativity of } \wedge \\
(\alpha \vee \beta) &\equiv (\beta \vee \alpha) && \text{commutativity of } \vee \\
((\alpha \wedge \beta) \wedge \gamma) &\equiv (\alpha \wedge (\beta \wedge \gamma)) && \text{associativity of } \wedge \\
((\alpha \vee \beta) \vee \gamma) &\equiv (\alpha \vee (\beta \vee \gamma)) && \text{associativity of } \vee \\
\neg(\neg\alpha) &\equiv \alpha && \text{double-negation elimination} \\
(\alpha \Rightarrow \beta) &\equiv (\neg\beta \Rightarrow \neg\alpha) && \text{contraposition} \\
(\alpha \Rightarrow \beta) &\equiv (\neg\alpha \vee \beta) && \text{implication elimination} \\
(\alpha \Leftrightarrow \beta) &\equiv ((\alpha \Rightarrow \beta) \wedge (\beta \Rightarrow \alpha)) && \text{biconditional elimination} \\
\neg(\alpha \wedge \beta) &\equiv (\neg\alpha \vee \neg\beta) && \text{De Morgan} \\
\neg(\alpha \vee \beta) &\equiv (\neg\alpha \wedge \neg\beta) && \text{De Morgan} \\
(\alpha \wedge (\beta \vee \gamma)) &\equiv ((\alpha \wedge \beta) \vee (\alpha \wedge \gamma)) && \text{distributivity of } \wedge \text{ over } \vee \\
(\alpha \vee (\beta \wedge \gamma)) &\equiv ((\alpha \vee \beta) \wedge (\alpha \vee \gamma)) && \text{distributivity of } \vee \text{ over } \wedge
\end{aligned}
$$

Inference rules

Modus Ponens :

Inference rules

Modus Ponens :

$$\frac{\alpha \Rightarrow \beta, \qquad \alpha}{\beta}$$

Inference rules

Modus Ponens :

$$\frac{\alpha \Rightarrow \beta, \qquad \alpha}{\beta}$$

And-Elimination :

Inference rules

Modus Ponens :

$$\frac{\alpha \Rightarrow \beta, \qquad \alpha}{\beta}$$

And-Elimination :

$$\frac{\alpha \wedge \beta}{\alpha}$$

Inference rules

Modus Ponens :

$$\frac{\alpha \Rightarrow \beta, \qquad \alpha}{\beta}$$

And-Elimination :

$$\frac{\alpha \wedge \beta}{\alpha}$$

Resolution :

Inference rules

$\neg a$

$a \vee b$

$a \vee \neg b \vee \neg c$

Modus Ponens :

$$\frac{\alpha \Rightarrow \beta, \qquad \alpha}{\beta}$$

And-Elimination :

$$\frac{\alpha \wedge \beta}{\alpha}$$

Resolution :

C_1 C_2

$$\frac{\underbrace{a \vee b \vee \neg c}_{C_1}, \qquad \underbrace{c \vee d}_{C_2}}{\underbrace{a \vee b \vee d}}$$

Concepts for inference

Concepts for inference

- Logical equivalences ($\equiv$)

$$M(\alpha_1) = M(\alpha_2) \quad \text{iff} \quad \alpha_1 \equiv \alpha_2$$

Concepts for inference

- Logical equivalences ($\equiv$)
- Validity or Tautology

α

Concepts for inference

- Logical equivalences ($\equiv$)
- Validity or Tautology
- Deduction theorem
  $\alpha \models \beta$ if and only if $\overbrace{(\alpha \Rightarrow \beta)}$ is valid. or tautology

Concepts for inference

- Logical equivalences ($\equiv$)
- Validity or Tautology
- Deduction theorem
  $\alpha \models \beta$ if and only if ______________ is valid.
- Monotonicity

Concepts for inference

- Logical equivalences ($\equiv$)
- Validity or Tautology
- Deduction theorem
  $\alpha \models \beta$ if and only if ________________ is valid.
- Monotonicity
  - Suppose $KB \models \alpha$. Is it possible to add a sentence to KB such that $KB' \not\models \alpha$?

Concepts for inference

- Logical equivalences ($\equiv$)
- Validity or Tautology
- Deduction theorem

  $\alpha \models \beta$ if and only if ________________ is valid.
- Monotonicity
  - Suppose $KB \models \alpha$. Is it possible to add a sentence to KB such that $KB \not\models \alpha$?

  Suppose KB' is obtained by adding more sentences to KB.

Concepts for inference

- Logical equivalences ($\equiv$)
- Validity or Tautology
- Deduction theorem

  $\alpha \models \beta$ if and only if ________________ is valid.
- Monotonicity
  - Suppose $KB \models \alpha$. Is it possible to add a sentence to KB such that $KB \not\models \alpha$?

  Suppose KB' is obtained by adding more sentences to KB.

$$\underbrace{M(KB)} \subseteq \underbrace{M(\alpha)}$$

Concepts for inference

- Logical equivalences ($\equiv$)
- Validity or Tautology
- Deduction theorem

  $\alpha \models \beta$ if and only if ________________ is valid.
- Monotonicity
  - Suppose $KB \models \alpha$. Is it possible to add a sentence to KB such that $KB \not\models \alpha$?

  Suppose KB' is obtained by adding more sentences to KB.

$$M(KB) \subseteq M(\alpha) \quad \text{— 1}$$
$$M(KB') \subseteq M(KB) \quad \text{— 2}$$

Concepts for inference

KB:

$\neg\alpha$

- Logical equivalences ($\equiv$)
- Validity or Tautology
- Deduction theorem
  $\alpha \models \beta$ if and only if ________________ is valid.
- Monotonicity
  - Suppose $KB \models \alpha$. Is it possible to add a sentence to KB such that $KB \not\models \alpha$?

  Suppose KB' is obtained by adding more sentences to KB.

$$M(KB) \subseteq M(\alpha)$$
$$M(KB') \subseteq M(KB)$$
$$\therefore M(KB') \subseteq M(\alpha)$$

$m(KB')$