PROPOSITIONAL LOGIC 2

Lara J. Martin (she/they)
TA: Aydin Ayanzadeh (he)

9/28/2023

CMSC 671

By the end of class today, you will be able to:

1. Identify if a knowledge base entails certain statements given the possible worlds
2. Write a proof using rules of inference in propositional logic

Modified from slides by Dr. Cynthia Matuszek & Dr. Chris Callison-Burch

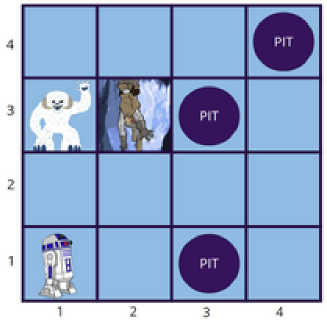
MODULE 2 GROUP PRESENTATIONS

1-page paper summaries due tomorrow 10/4 at **11:59pm**

Group presentations are Thursday 10/5

HW 2 RELEASED

https://laramartin.net/Principles-of-AI/homeworks/logic/logical-agent.html



Wampa World

Homework 2: Hunt the Wampa (10%)

Due October 10, 2023 at 11:59:00 PM on Blackboard.

Materials:

HW2-LogicalAgents.ipynb

Learning Objectives

In this assignment, you will:

- Combine propositional logic rules to create an inference algorithm & knowledge base that can successfully guide the agent (the robot R2-D2) toward its goal
- Analyze the consequences of propositional logic rules on the agent's decision-making process
- Evaluate the effectiveness of your inference algorithm in guiding the agent's behavior in different Wampa World scenarios
- Recognize logical agents in the wild
- Compare logical agents to search algorithms

Part 1: Implement the agent

RECAP

PROPOSITIONAL LOGIC DEFINITIONS

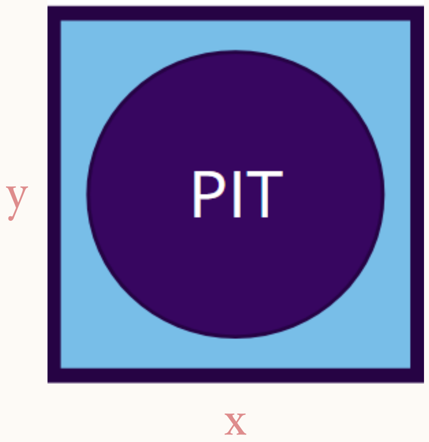
- Symbol: a variable that stands for a statement that must be either True or False
- Sentence: an assertion about the world; in a knowledge representation language
  - Two kinds: axioms and derived sentences
- Inference: deriving new sentences from KB
- M(**α**): all possible worlds where **α** is true
- If **α**⊨β (entailment), **α** is a stronger/more specific statement than β
- If **α**⊢β (inference), β is provable from **α**
- $\mathbf{W_{1,3} \Rightarrow S_{1,2}}$ What is $\mathbf{W_{1,3}}$? What is $\mathbf{S_{1,2}}$?
  implies premise/antecede conclusion/consequence

A model of a KB is an interpretation in which all sentences in KB are true (i.e., like the conjunction of all sentences in the KB)

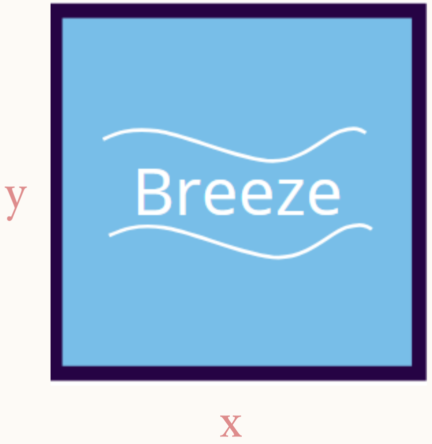
POSSIBLE WORLDS

WAMPA WORLD KB

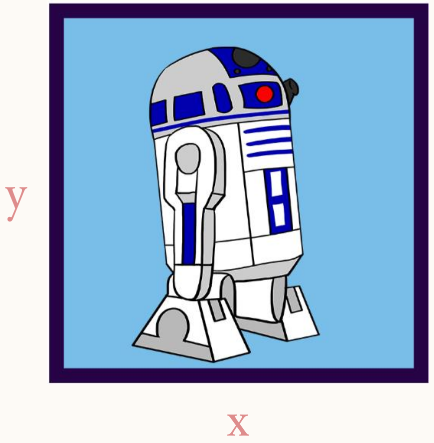
Pxy is true if there is a pit in [x,y]



Bxy is true if there is a breeze in [x,y]

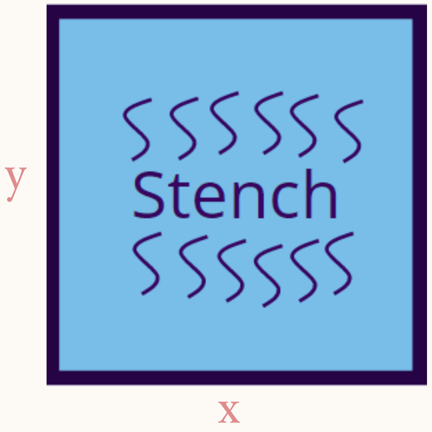


Axy is true if there is an agent in [x,y]



Wxy is true if there is a Wampa in [x,y]

Sxy is true if there is a stench in [x,y]



Symbols for each location [x,y]

WAMPA WORLD KB

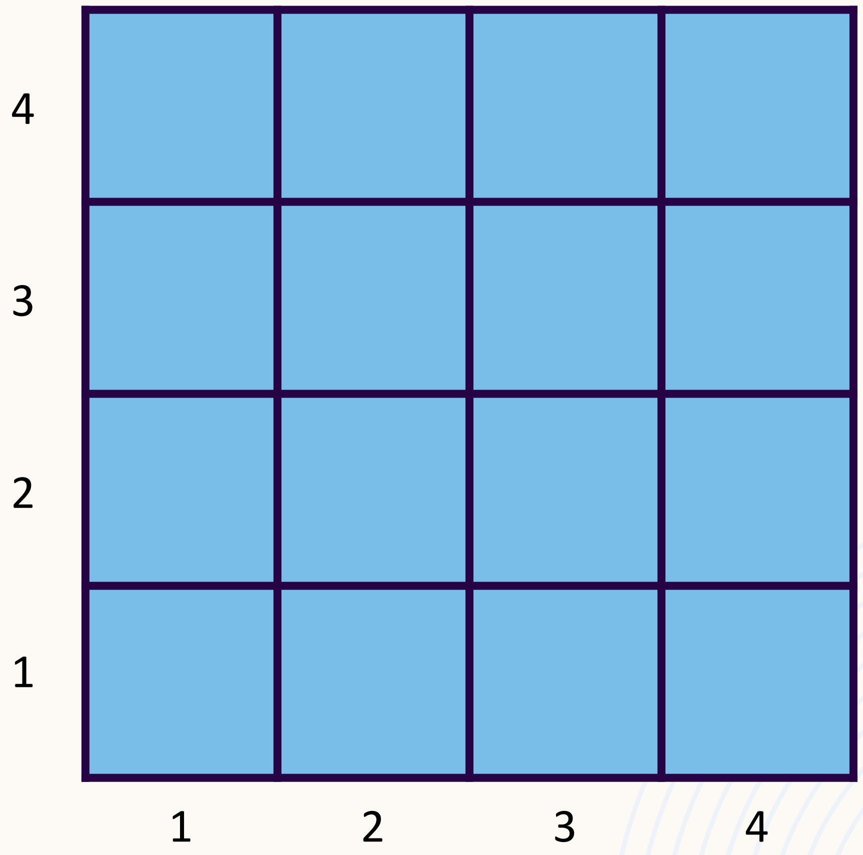
We can construct sentences out of these using logical connectors. We'll label each sentence.

R1: $\neg$ P1,1

R2: B1,1 $\Leftrightarrow$ (P1,2 $\vee$ P2,1)

R3: B2,1 $\Leftrightarrow$ (P1,1 $\vee$ P2,2 $\vee$ P3,1)

These are true of all Wampa Worlds.



WAMPA WORLD KB

We can construct sentences out of these using logical connectors. We'll label each sentence.

R1: $\neg P_{1,1}$

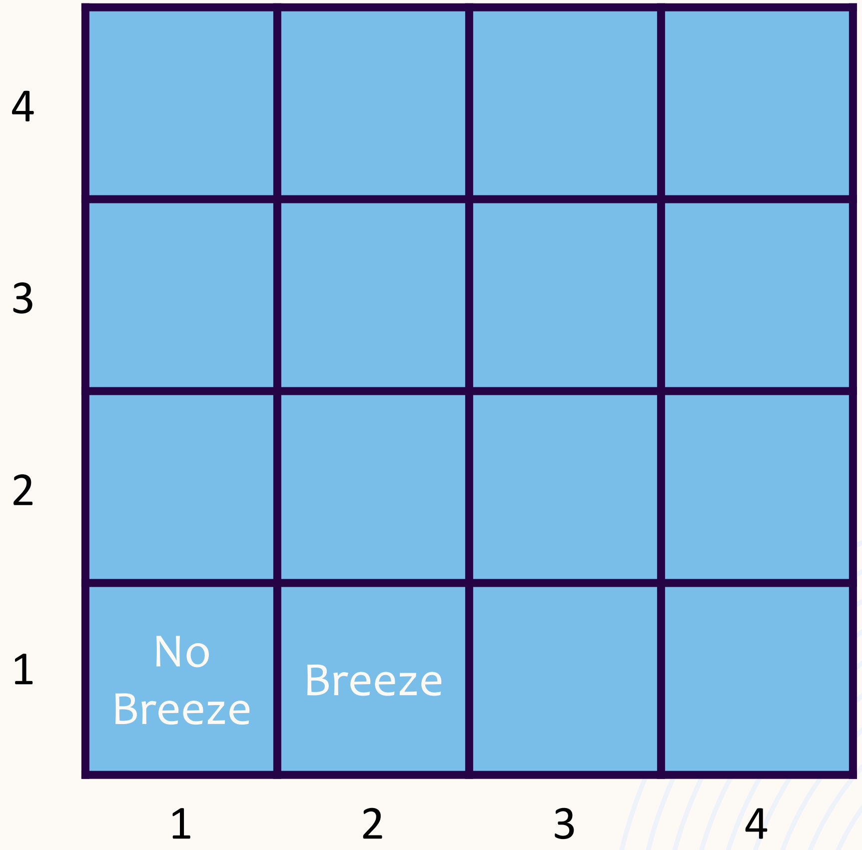
R2: $B_{1,1} \Leftrightarrow (P_{1,2} \lor P_{2,1})$

R3: $B_{2,1} \Leftrightarrow (P_{1,1} \lor P_{2,2} \lor P_{3,1})$

What if we perceive the presence or absence of breeze in [1,1], [2,1]?

R4: $\neg B_{1,1}$

R5: $B_{2,1}$



WAMPA WORLD KB

We can construct sentences out of these using logical connectors. We'll label each sentence.

R1: ¬ P1,1

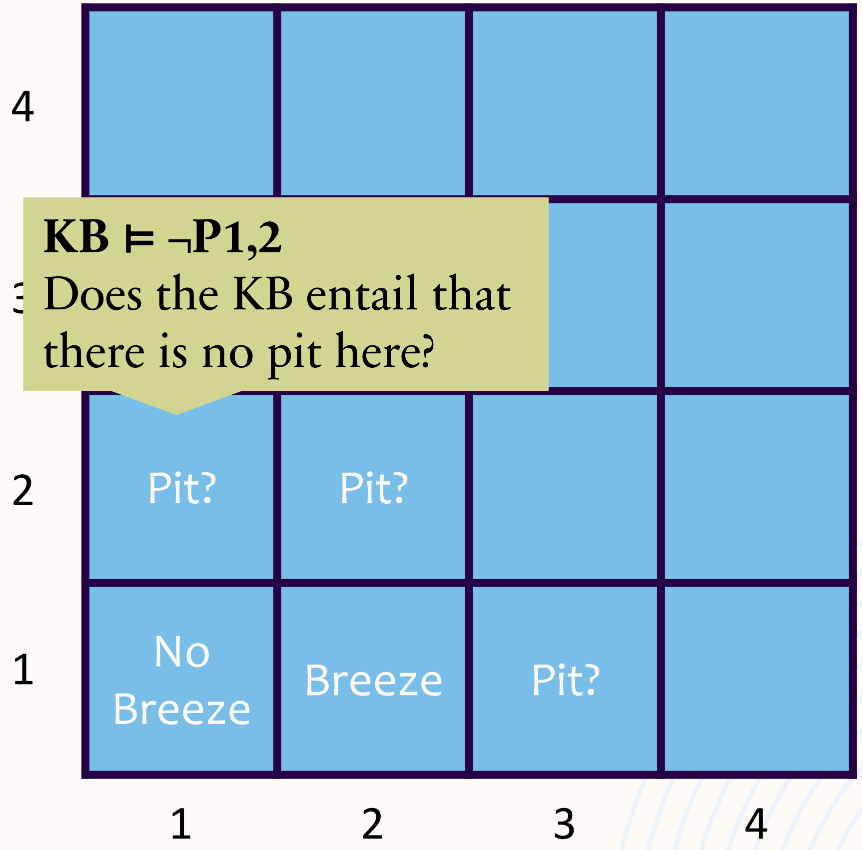
R2: B1,1 ⇔ (P1,2 V P2,1)

R3: B2,1 ⇔ (P1,1 V P2,2 V P3,1)

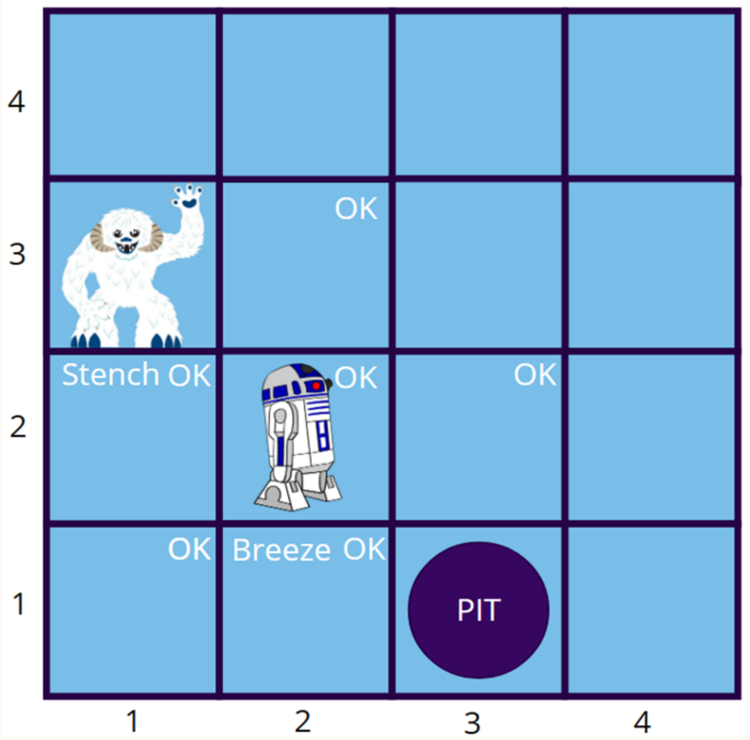
R4: ¬ B1,1

R5: B2,1

Can we mechanically combine the sentences in our KB to prove that a pit exists at (or is absent from) any location?

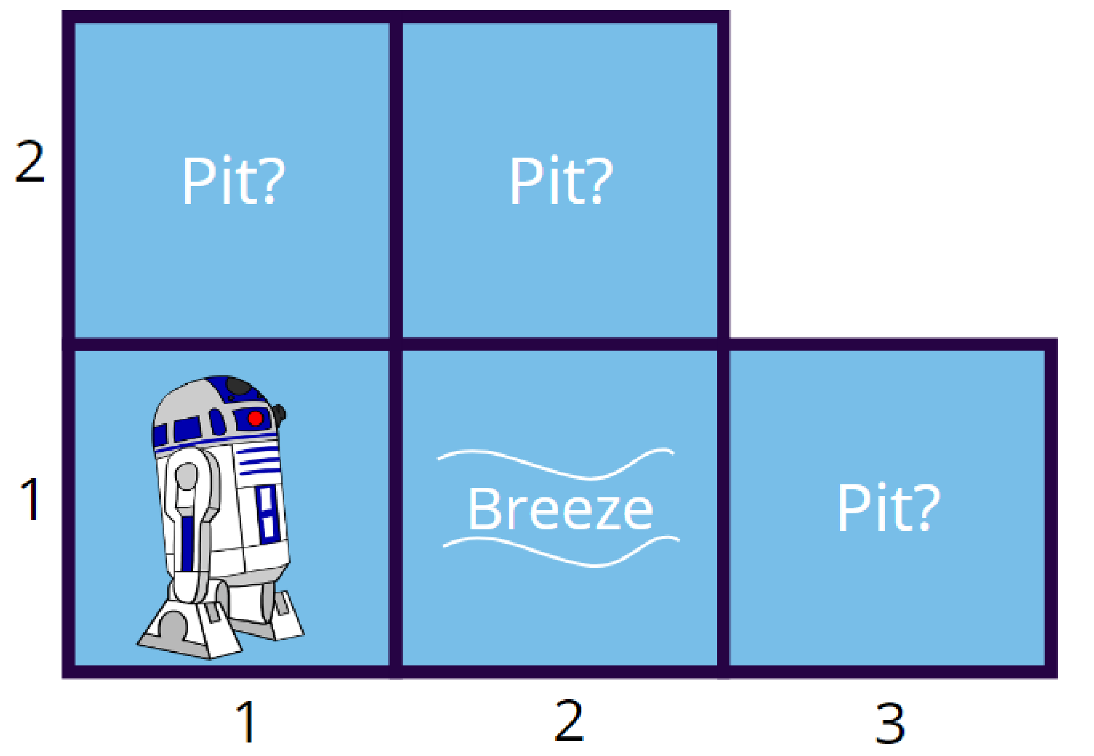


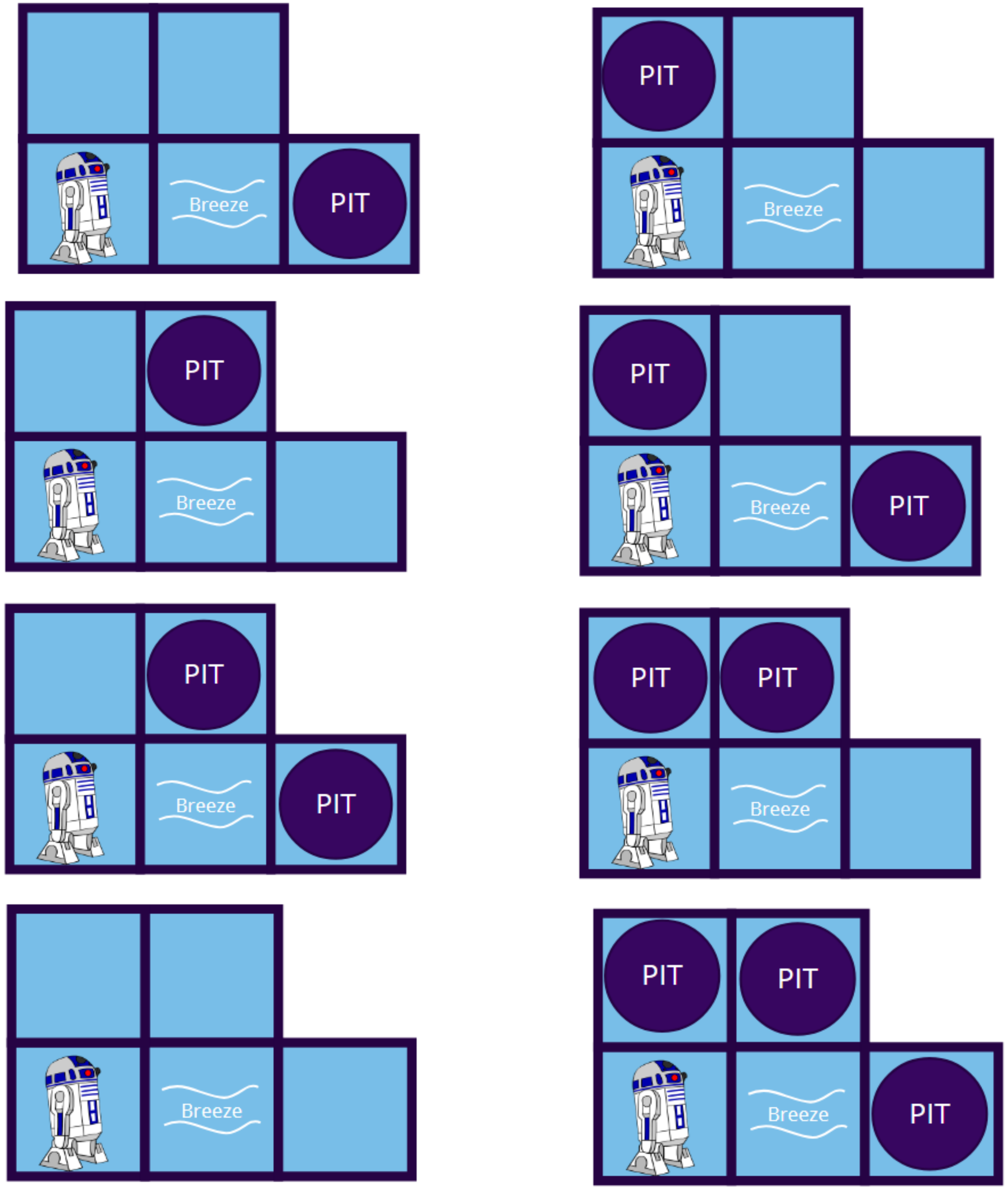
THE "HUNT THE WAMPA" AGENT



- Some atomic propositions:
  - S12 = There is a stench in cell (1,2)
  - B34 = There is a breeze in cell (3,4)
  - W13 = The Wampa is in cell (1,3)
  - V11 = We have visited cell (1,1)
  - OK11 = Cell (1,1) is safe.
  - etc
- Some rules:
  - (Rule 1) $\neg S11 \rightarrow \neg W11 \land \neg W12 \land \neg W21$
  - (Rule 2) $\neg S21 \rightarrow \neg W11 \land \neg W21 \land \neg W22 \land \neg W31$
  - (Rule 3) $\neg S12 \rightarrow \neg W11 \land \neg W12 \land \neg W22 \land \neg W13$
  - (Rule 4) $S12 \rightarrow W13 \lor W12 \lor W22 \lor W11$
- Note that the lack of variables requires us to give similar rules for each cell

POSSIBLE WORLDS

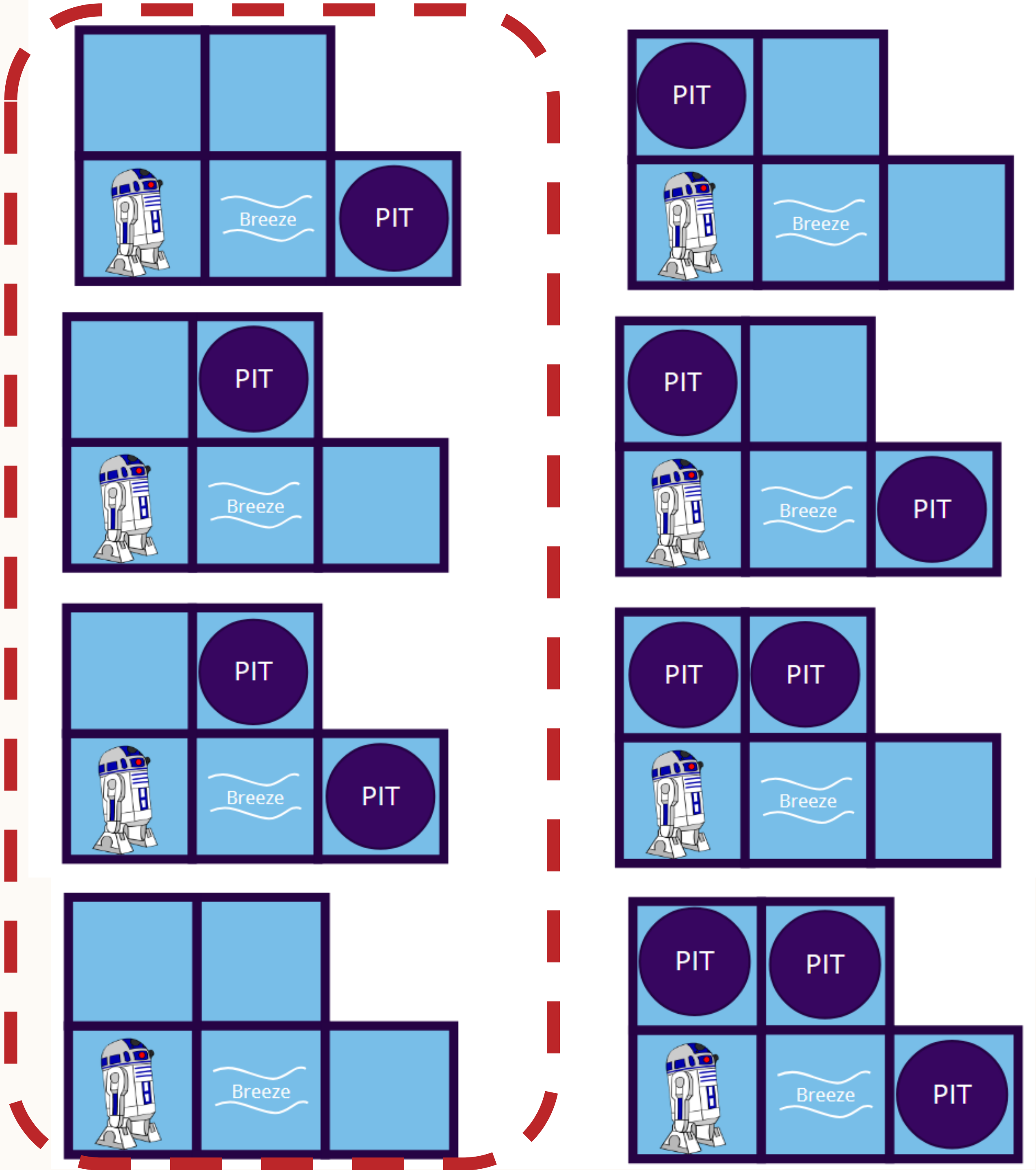




M(**α**) – set of all models **m** where **α** is satisfied

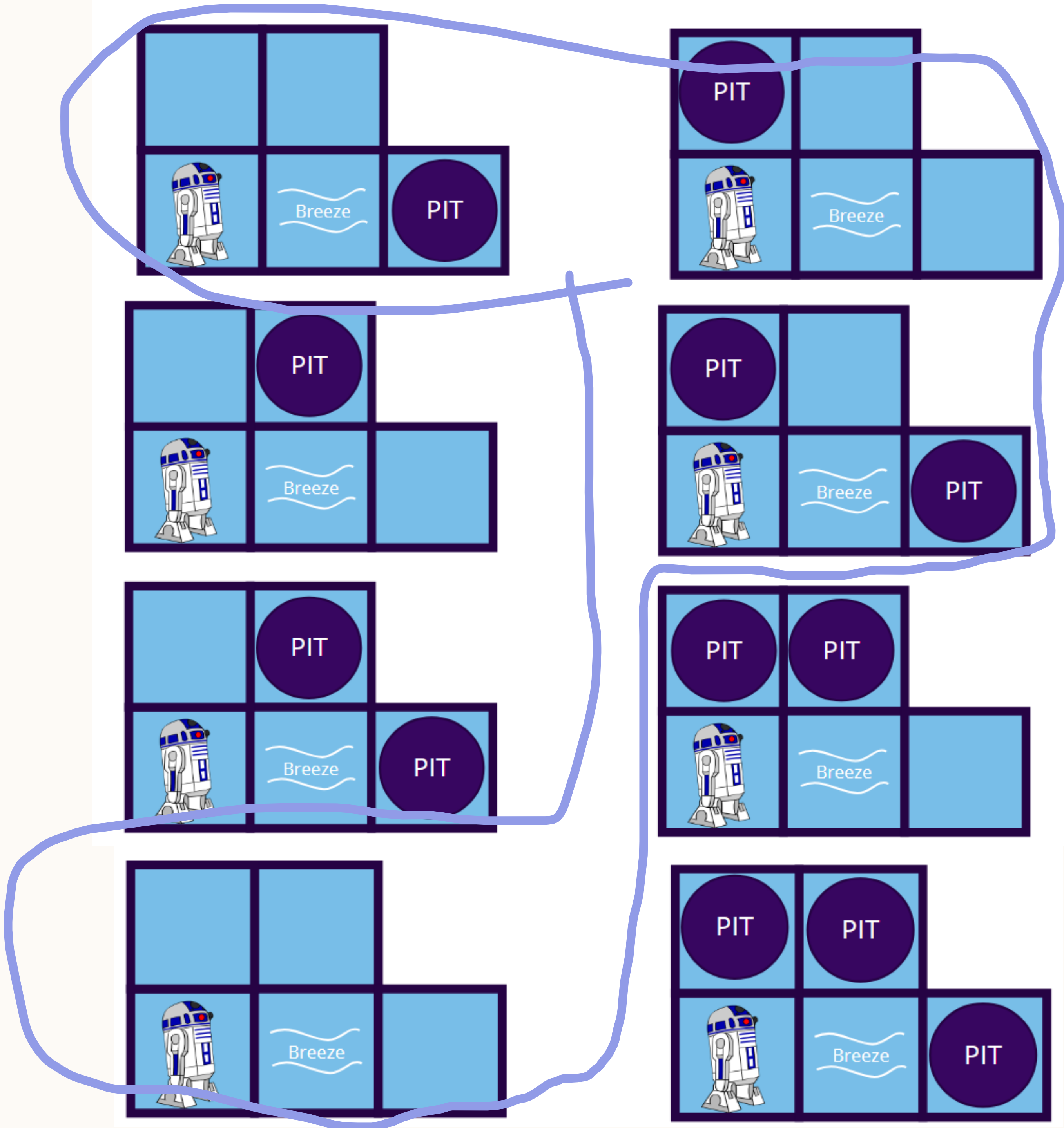
POSSIBLE WORLDS

α_1 = “There is no pit in [1,2]”



POSSIBLE WORLDS

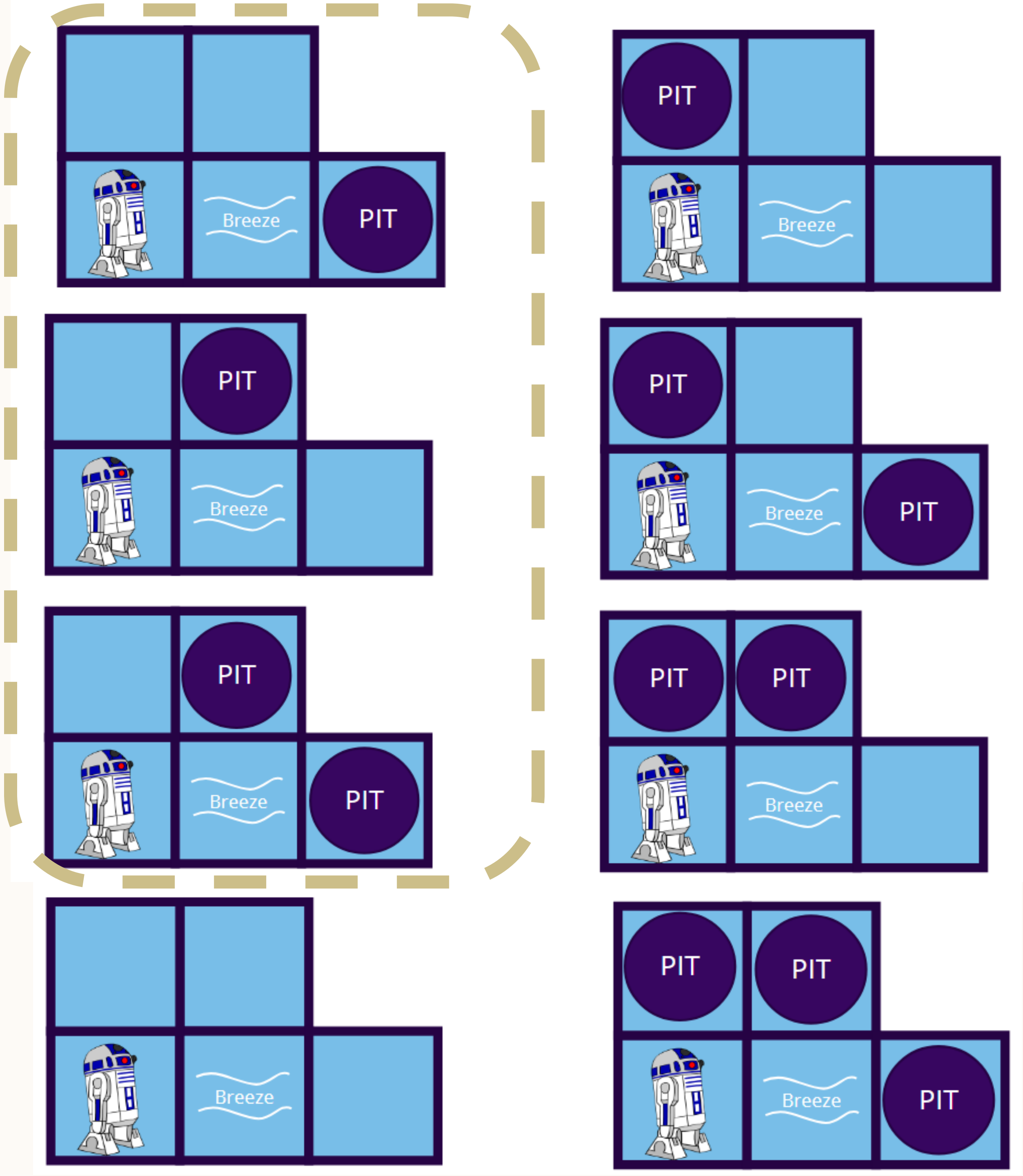
α_1 = "There is no pit in [2,2]"



POSSIBLE WORLDS

KB =

R1: $\neg P_{1,1}$
R2: $B_{1,1} \Leftrightarrow (P_{1,2} \lor P_{2,1})$
R3: $B_{2,1} \Leftrightarrow (P_{1,1} \lor P_{2,2} \lor P_{3,1})$
R4: $\neg B_{1,1}$
R5: $B_{2,1}$



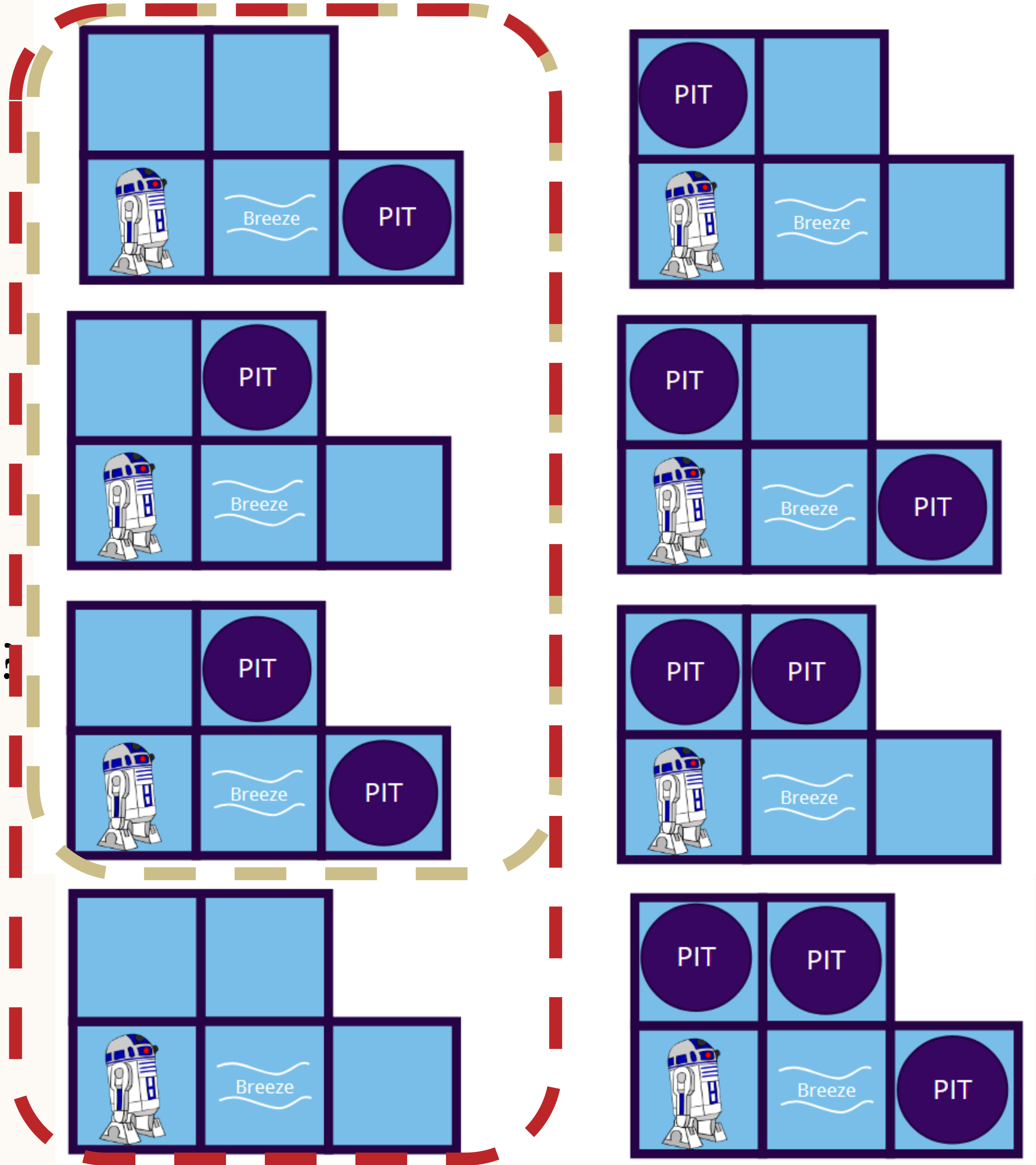
POSSIBLE WORLDS

β ⊨ α if and only if **M(β) ⊆ M(α)**
"**β entails α** if and only if every model in which β is true, **α** is also true"

Does our KB entail that there is no pit in [1,2]?
KB⊨α_1 if and only if **M(KB) ⊆ M(α_1)**

KB =
R1: ¬ P1,1
R2: B1,1 ⇔ (P1,2 V P2,1)
R3: B2,1 ⇔ (P1,1 V P2,2 V P3,1)
R4: ¬ B1,1
R5: B2,1

α_1 = "There is no pit in [1,2]"
KB⊨α_1



POSSIBLE WORLDS

$\beta \vDash \boldsymbol{\alpha}$ if and only if $\mathbf{M}(\beta) \subseteq \mathbf{M}(\boldsymbol{\alpha})$

"**β entails α** if and only if every model in which β is true, **α** is also true"

Does our KB entail that there is **no pit in** [2,2]?

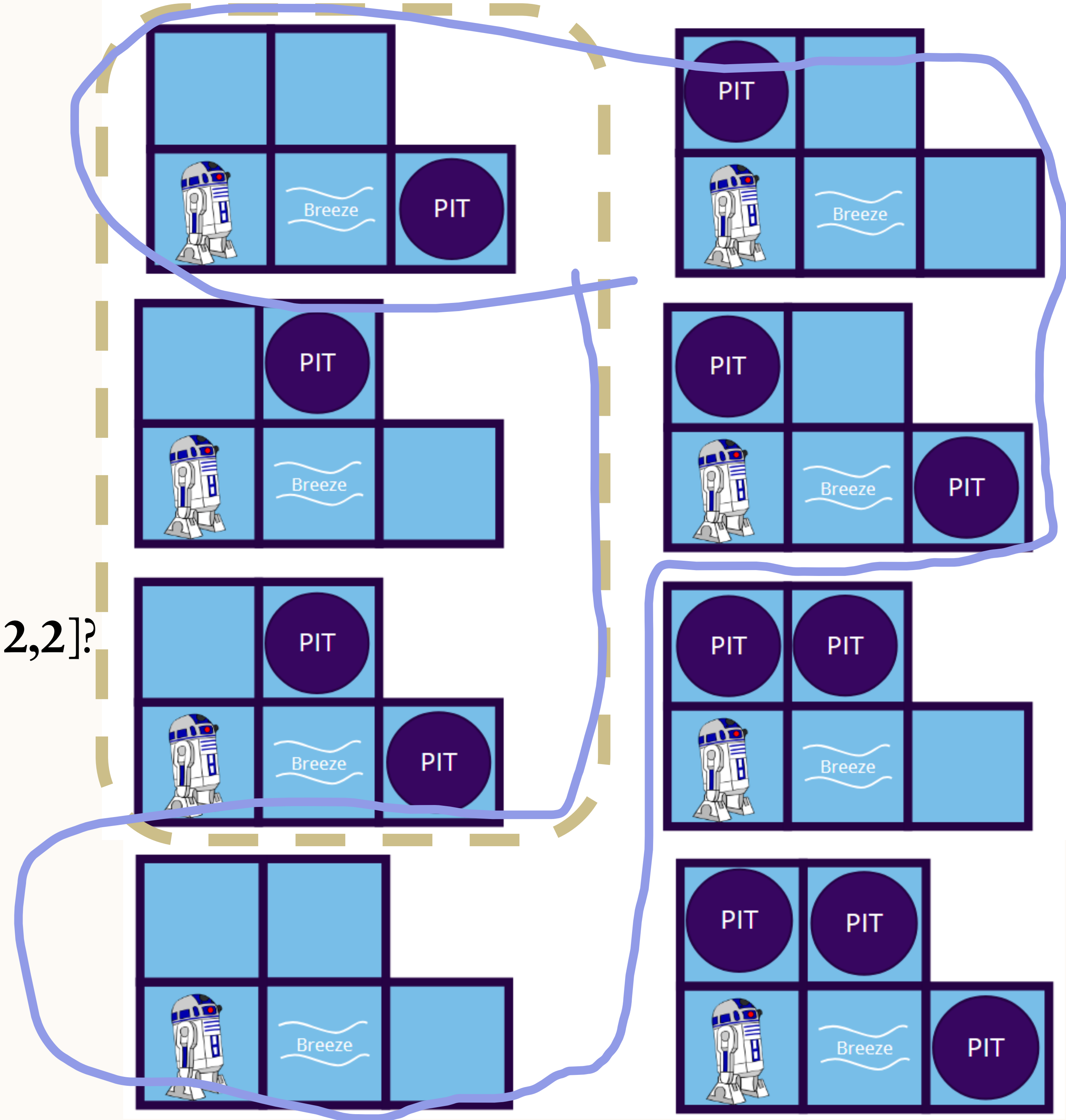
$\mathbf{KB} \vDash \boldsymbol{\alpha}_2$ if and only if $\mathbf{M}(\mathbf{KB}) \subseteq \mathbf{M}(\boldsymbol{\alpha}_2)$

KB =

R1: ¬ P1,1
R2: B1,1 ⇔ (P1,2 V P2,1)
R3: B2,1 ⇔ (P1,1 V P2,2 V P3,1)
R4: ¬ B1,1
R5: B2,1

$\boldsymbol{\alpha}_2$ = "There is no pit in [2,2]"

KB does not entail $\boldsymbol{\alpha}_2$ in some models where KB is true & $\boldsymbol{\alpha}_2$ is false



THEOREM PROVING

SOUND RULES OF INFERENCE

RULE	PREMISES	CONCLUSION
Modus Ponens	$\alpha, \alpha \Rightarrow \beta$	β
And Introduction	α, β	$\alpha \land \beta$
And Elimination	$\alpha \land \beta$	α
Double Negation	$\neg \neg\alpha$	α
Unit Resolution	$\alpha \lor \beta, \neg\beta$	α
Resolution	$\alpha \lor \beta, \neg\beta \lor \gamma$	$\alpha \lor \gamma$
de Morgans	$\neg(\alpha \lor \beta)$	$\neg\alpha \land \neg\beta$
$\lor/\Rightarrow$ Equivalence	$\alpha \Rightarrow \beta$	$\neg\alpha \lor \beta$

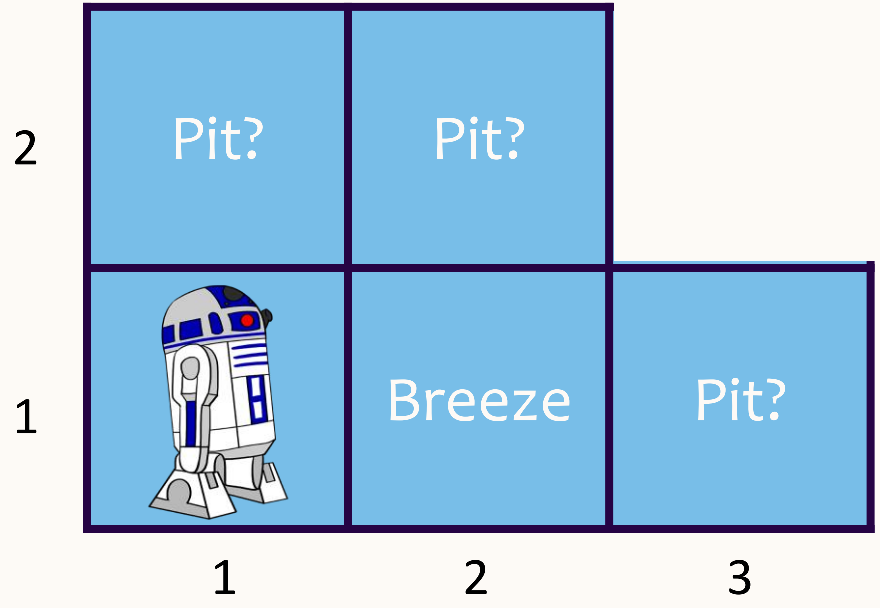
All of the logical equivalence rules can be re-written as inference rules.

INFERENCE EXAMPLE

KB =

R1: $\neg P_{1,1}$
R2: $B_{1,1} \Leftrightarrow (P_{1,2} \lor P_{2,1})$
R3: $B_{2,1} \Leftrightarrow (P_{1,1} \lor P_{2,2} \lor P_{3,1})$
R4: $\neg B_{1,1}$
R5: $B_{2,1}$

Prove whether or not there is a pit in [1,2].



INFERENCE EXAMPLE

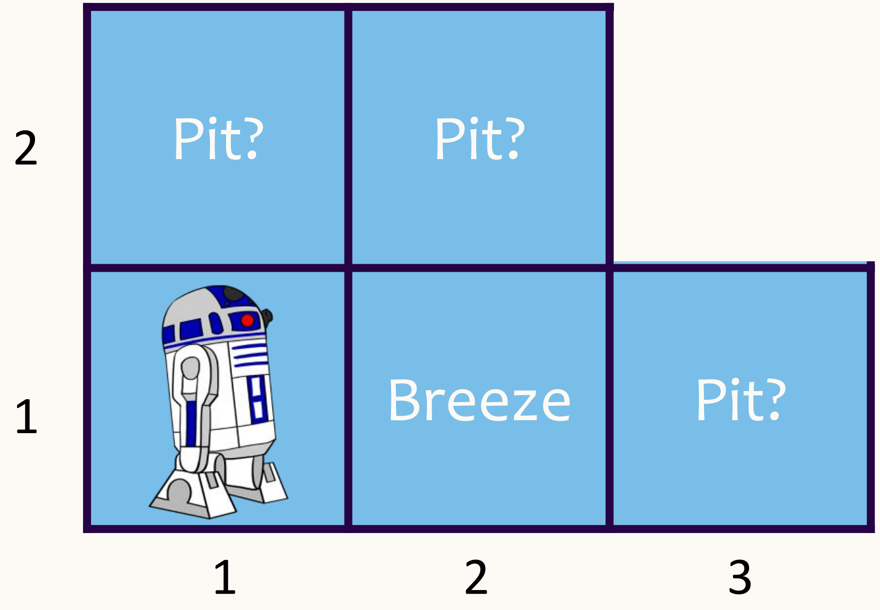
KB =
- R1: $\neg P_{1,1}$
- R2: $B_{1,1} \Leftrightarrow (P_{1,2} \lor P_{2,1})$
- R3: $B_{2,1} \Leftrightarrow (P_{1,1} \lor P_{2,2} \lor P_{3,1})$
- R4: $\neg B_{1,1}$
- R5: $B_{2,1}$

Biconditional Elimination:

$$\frac{\alpha \Leftrightarrow \beta}{(\alpha \Rightarrow \beta) \land (\beta \Rightarrow \alpha)}$$

Apply biconditional Elimination to R2 to get R6.

R6: $(B_{1,1} \Rightarrow (P_{1,2} \lor P_{2,1})) \land ((P_{1,2} \lor P_{2,1}) \Rightarrow B_{1,1})$



Monotonicity: if **KB⊨α** then **KB ∧ β ⊨α**
We can safely add to the KB, without invalidating anything else that we inferred.

INFERENCE EXAMPLE

KB =

$R1: \neg P1,1$
$R2: B1,1 \Leftrightarrow (P1,2 \vee P2,1)$
$R3: B2,1 \Leftrightarrow (P1,1 \vee P2,2 \vee P3,1)$
$R4: \neg B1,1$
$R5: B2,1$

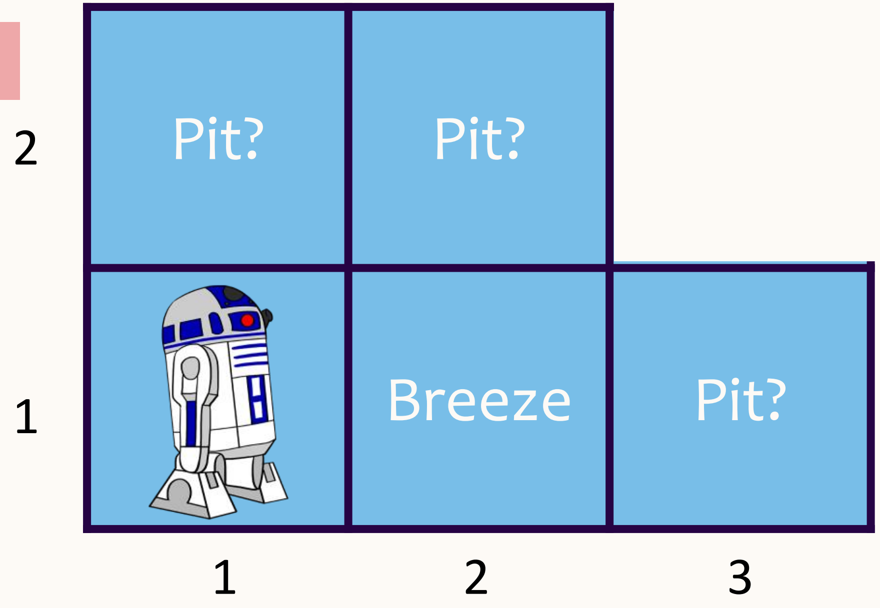
$R6: (B1,1 \Rightarrow (P1,2 \vee P2,1)) \wedge ((P1,2 \vee P2,1) \Rightarrow B1,1)$

And Elimination:

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

Apply And-Elimination to R6 to get R7.

$R7: ((P1,2 \vee P2,1) \Rightarrow B1,1)$



INFERENCE EXAMPLE

KB =

R1: $\neg P1,1$
R2: $B1,1 \Leftrightarrow (P1,2 \lor P2,1)$
R3: $B2,1 \Leftrightarrow (P1,1 \lor P2,2 \lor P3,1)$
R4: $\neg B1,1$
R5: $B2,1$

Logical equivalence for contrapositives:

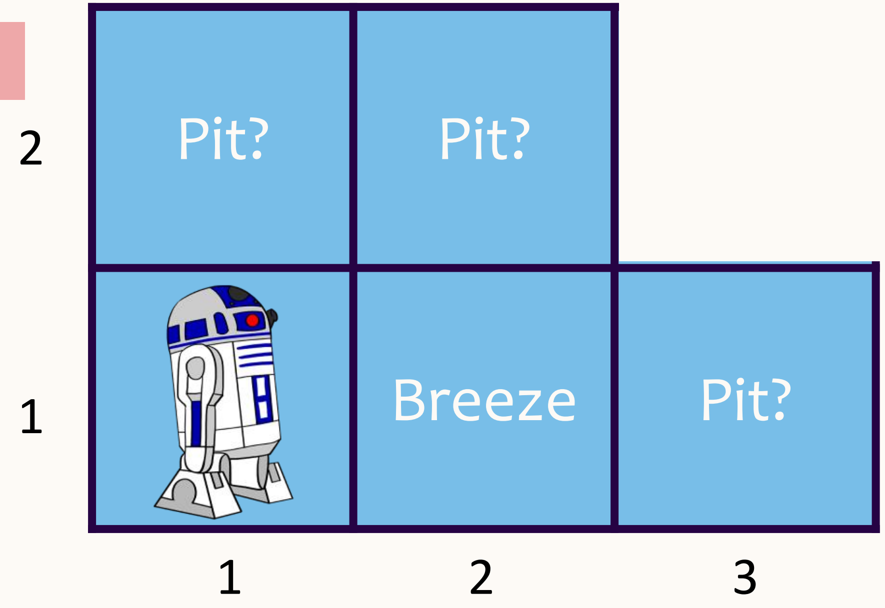
$$\frac{(\alpha \Rightarrow \beta)}{(\neg \beta \Rightarrow \neg \alpha)}$$

R6: $(B1,1 \Rightarrow (P1,2 \lor P2,1)) \land ((P1,2 \lor P2,1) \Rightarrow B1,1)$

R7: $((P1,2 \lor P2,1) \Rightarrow B1,1)$

Logical equivalence for contrapositives applied to R7 gives R8.

R8: $(\neg B1,1 \Rightarrow \neg (P1,2 \lor P2,1))$



INFERENCE EXAMPLE

KB =
- R1: $\neg P_{1,1}$
- R2: $B_{1,1} \Leftrightarrow (P_{1,2} \lor P_{2,1})$
- R3: $B_{2,1} \Leftrightarrow (P_{1,1} \lor P_{2,2} \lor P_{3,1})$
- R4: $\neg B_{1,1}$
- R5: $B_{2,1}$

R6: $(B_{1,1} \Rightarrow (P_{1,2} \lor P_{2,1})) \land ((P_{1,2} \lor P_{2,1}) \Rightarrow B_{1,1})$

R7: $((P_{1,2} \lor P_{2,1}) \Rightarrow B_{1,1})$

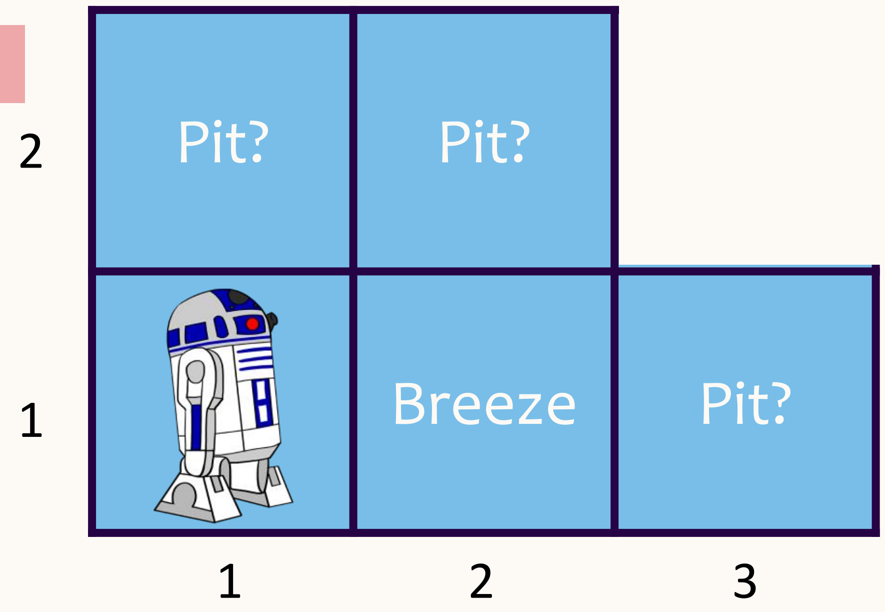
R8: $(\neg B_{1,1} \Rightarrow \neg (P_{1,2} \lor P_{2,1}))$

Apply Modus Ponens to R4 and R8 to get:

R9: $\neg (P_{1,2} \lor P_{2,1}))$

Modus Ponens:

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



INFERENCE EXAMPLE

KB =

R1: $\neg P_{1,1}$
R2: $B_{1,1} \Leftrightarrow (P_{1,2} \lor P_{2,1})$
R3: $B_{2,1} \Leftrightarrow (P_{1,1} \lor P_{2,2} \lor P_{3,1})$
R4: $\neg B_{1,1}$
R5: $B_{2,1}$

R6: $(B_{1,1} \Rightarrow (P_{1,2} \lor P_{2,1})) \land ((P_{1,2} \lor P_{2,1}) \Rightarrow B_{1,1})$

R7: $((P_{1,2} \lor P_{2,1}) \Rightarrow B_{1,1})$

R8: $(\neg B_{1,1} \Rightarrow \neg (P_{1,2} \lor P_{2,1}))$

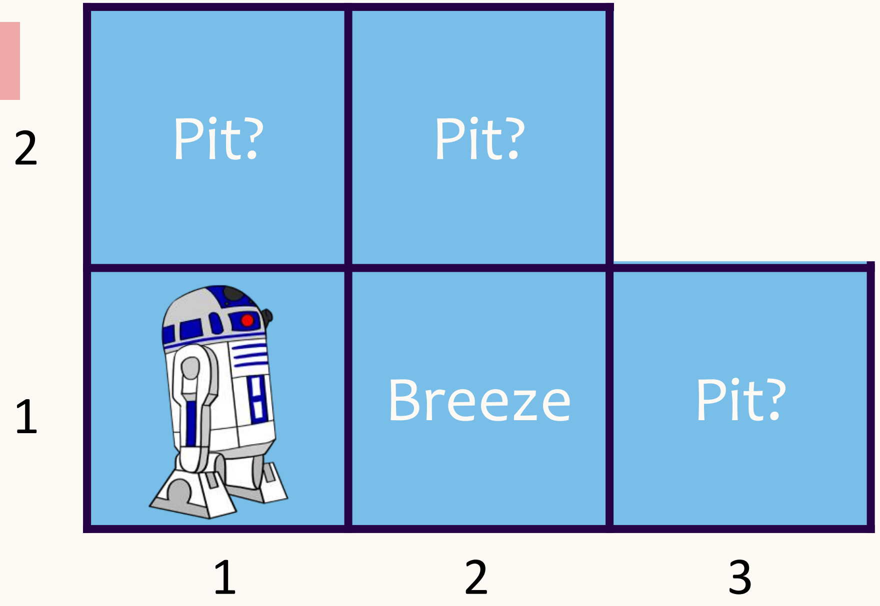
R9: $\neg (P_{1,2} \lor P_{2,1}))$

Apply De Morgan's Rule to R9:

R10: $\neg P_{1,2} \land \neg P_{2,1}$

De Morgan's Rule:

$$\frac{\neg(\boldsymbol{\alpha} \lor \beta)}{(\neg\alpha \land \neg\beta)}$$



INFERENCE EXAMPLE

KB =

R1: $\neg P1,1$
R2: $B1,1 \Leftrightarrow (P1,2 \lor P2,1)$
R3: $B2,1 \Leftrightarrow (P1,1 \lor P2,2 \lor P3,1)$
R4: $\neg B1,1$
R5: $B2,1$

R6: $(B1,1 \Rightarrow (P1,2 \lor P2,1)) \land ((P1,2 \lor P2,1) \Rightarrow B1,1)$

R7: $((P1,2 \lor P2,1) \Rightarrow B1,1)$

R8: $(\neg B1,1 \Rightarrow \neg (P1,2 \lor P2,1))$

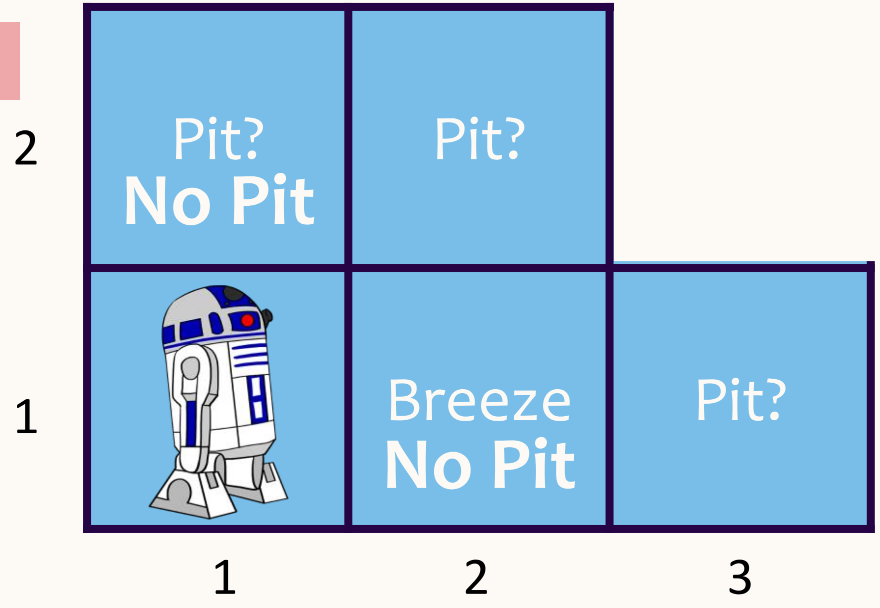
R9: $\neg (P1,2 \lor P2,1))$

R10: $\neg P1,2 \land \neg P2,1$

R11: $\neg P1,2$ R12: $\neg P2,1$

And Elimination:

$$\frac{\alpha \land \beta}{\alpha}$$



Your Mission

Prove that the **Wampa is in (1,3)**, given the observations shown and these rules:

Reminder of Wampa Rules

- If there is no stench in a cell, then there is no Wampa in any adjacent cell
- If there is a stench in a cell, then there is a Wampa in some adjacent cell
- If there is no breeze in a cell, then there is no pit in any adjacent cell
- If there is a breeze in a cell, then there is a pit in some adjacent cell
- If a cell has been visited, it has neither a Wampa nor a pit

FIRST write the propositional rules for the relevant cells (your initial KB)

THEN write the proof steps and indicate what inference rules you used in each step

PROVE IT!

A = Agent
B = Breeze
G = Gasp
OK = Safe square
P = Pit
S = Stench
V = Visited
W = Wampa

V12 S12 ¬B12	V22 ¬S22 ¬B22		
V11 ¬S11 ¬B11	V21 B21 ¬S21		

Inference Rules

Modus Ponens

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

And Introduction

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

And Elimination

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

Double Negation

$$\frac{\neg\neg\alpha}{\alpha}$$

Unit Resolution

$$\frac{\alpha \vee \beta, \neg\beta}{\alpha}$$

Resolution

$$\frac{\alpha \vee \beta, \neg\beta \vee \gamma}{\alpha \vee \gamma}$$

FOR NEXT CLASS

- Read Chapters 8.1.2, 8.2, 8.3, 9.3
- Get ready to present Module 2 (for those with Module 2)
- Start looking at Homework 2