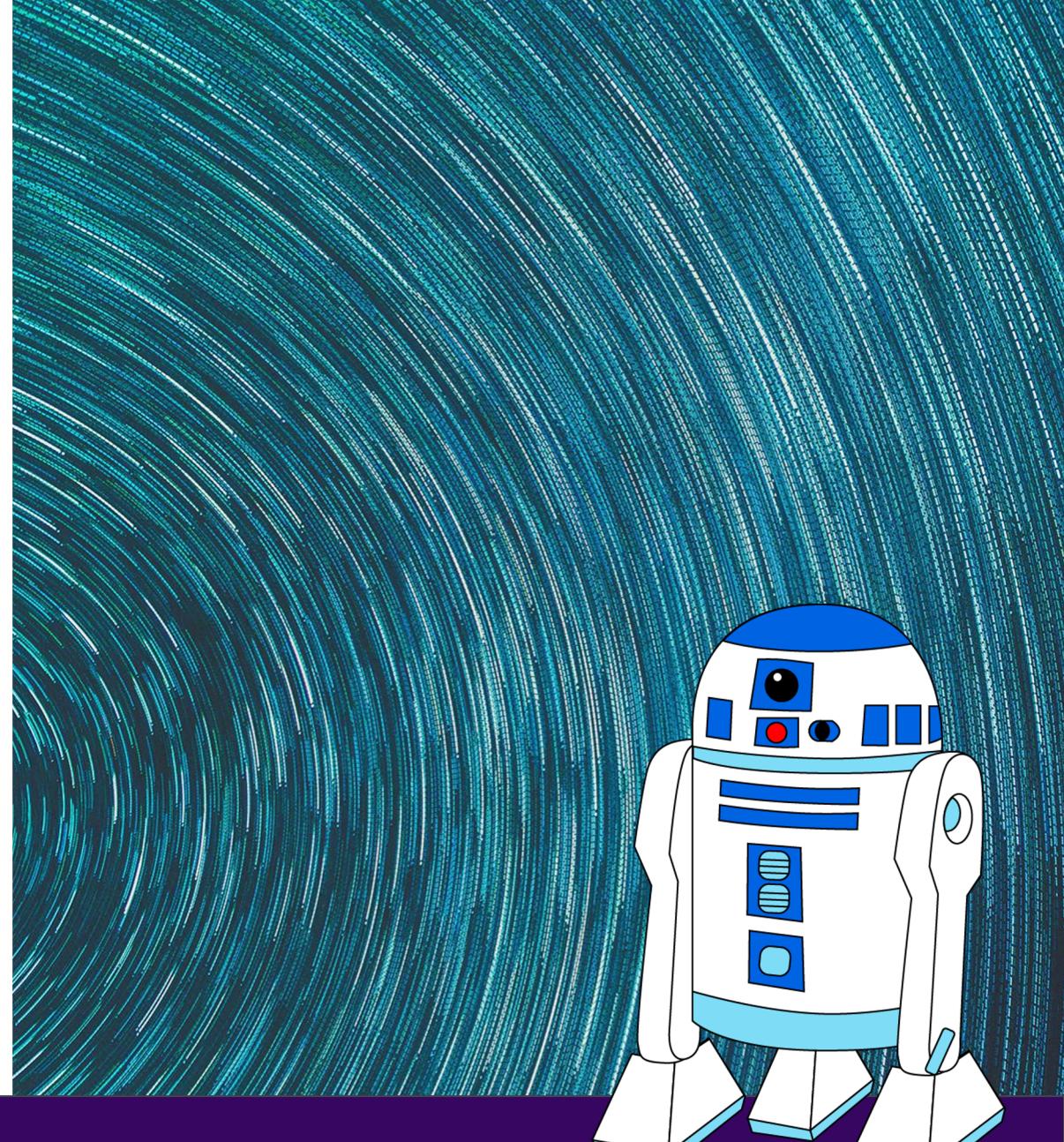


CIS 421/521:
ARTIFICIAL INTELLIGENCE

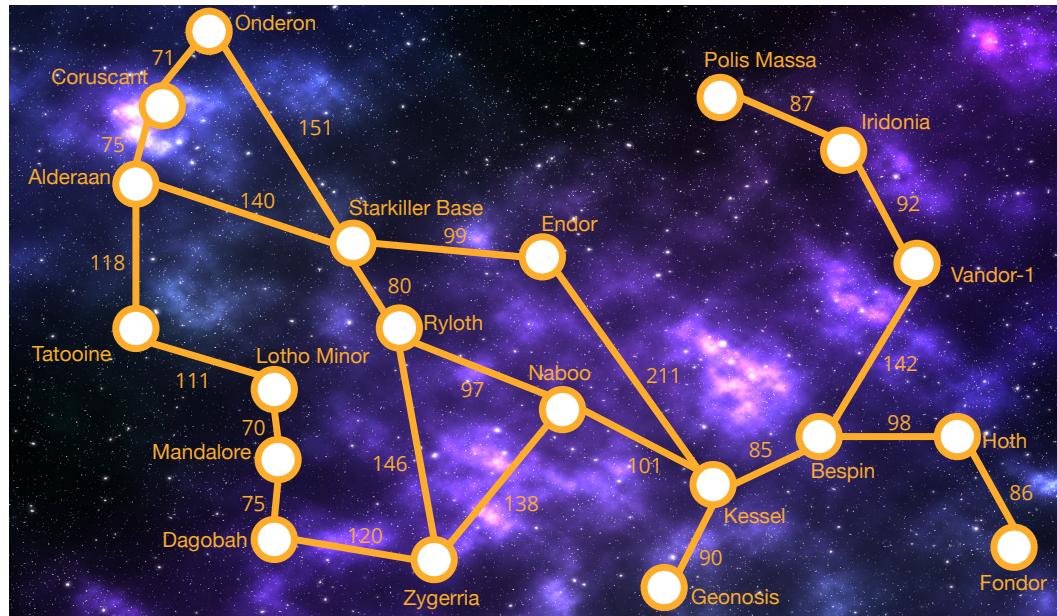
Logical Agents



Knowledge-based Agents

Knowledge-based agents use a process of **reasoning** over an internal **representation** of knowledge to decide what action to take.

So far, our **problem-solving** agents have performed a **search** over **states** in order to find a plan. The representation of states has been **atomic**.



Limited to commands like
“Navigate to Kessel”

“Take me to the nearest
habitable planet where I can
store my perishable cargo”

Knowledge-based Agents

A central component of a knowledge-based agent is a **knowledge base** or KB.

A KB contains a set of **sentence** that are written in a **knowledge representation language**. The sentence contains some assertion about the world.

Natural language sentences	Knowledge representation language sentence
<i>Hoth is a planet</i>	planet(hoth)
<i>Hoth is habitable</i>	habitable(hoth)
<i>Hoth is far from its sun</i>	far_from(hoth, sol)
<i>If a planet is far from its sun then it is cold</i>	planet(x) and sun(y) and far_from(x,y) → cold(x)

Knowledge-based Agents

There are two kinds of sentences:

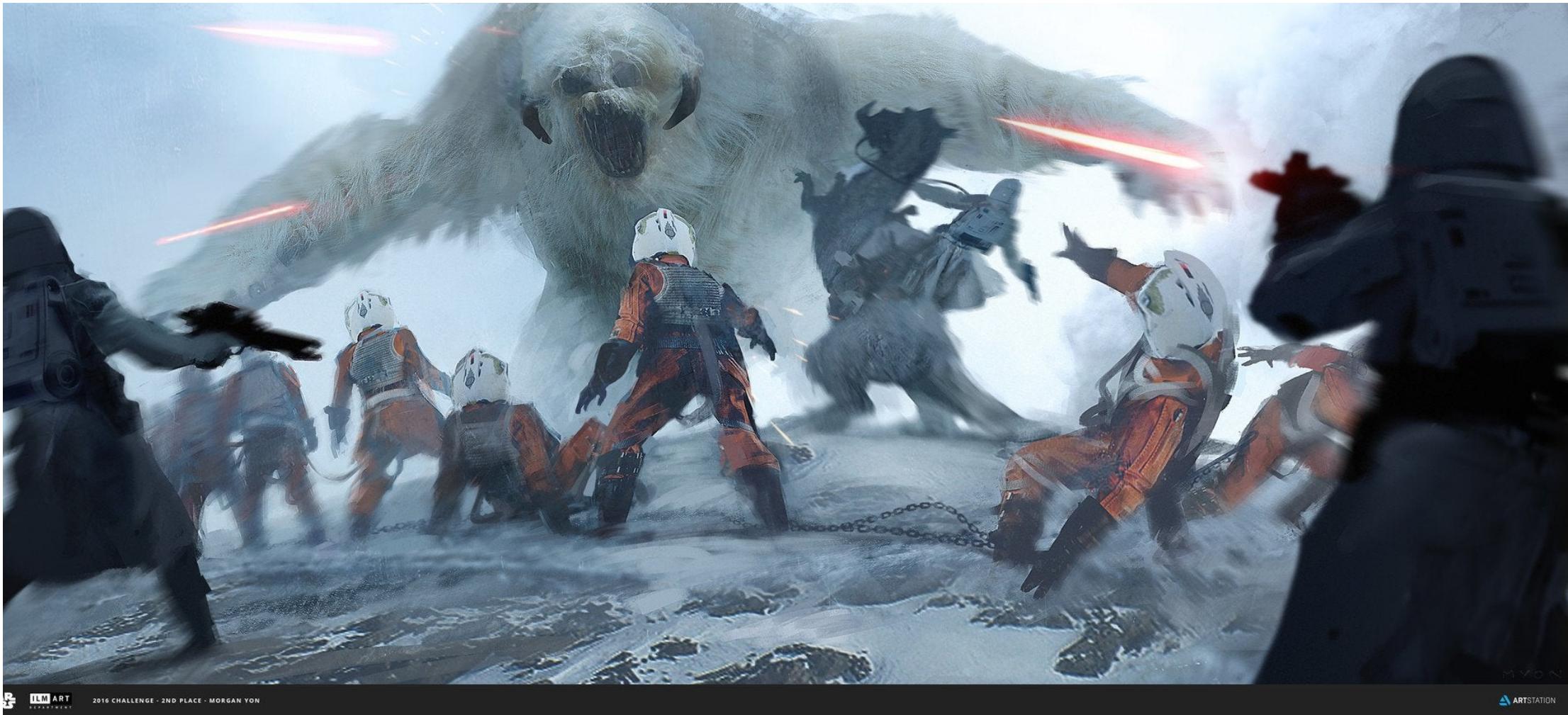
- **Axioms** – a sentence that is given
- **Derived sentences** – a new sentence that is derived from others sentences

The process of deriving new sentences from old sentences is called **inference**.

A KB can initially contain some **background information** about the world, and a knowledge-based agent can add to the information in the KB through its observations of the world.

In addition to asserting new knowledge into its KB, a knowledge-based agent can also query the KB and ask it to derive new knowledge in order to select what action it should take.

Hunt the Wampas



ILM ART
DEPARTMENT

2016 CHALLENGE - 2ND PLACE - MORGAN YON

M.YON

ARTSTATION

Wampa World

Our **knowledge-based agent**, R2D2, explores **a cave** consisting of **rooms** connected by passageways.

Lurking somewhere in the cave is the **Wampa**, a beast that eats any agent that enters its room.

Some rooms contain bottomless **pits** that trap any agent that wanders into the room.

In one room is master **Luke**.

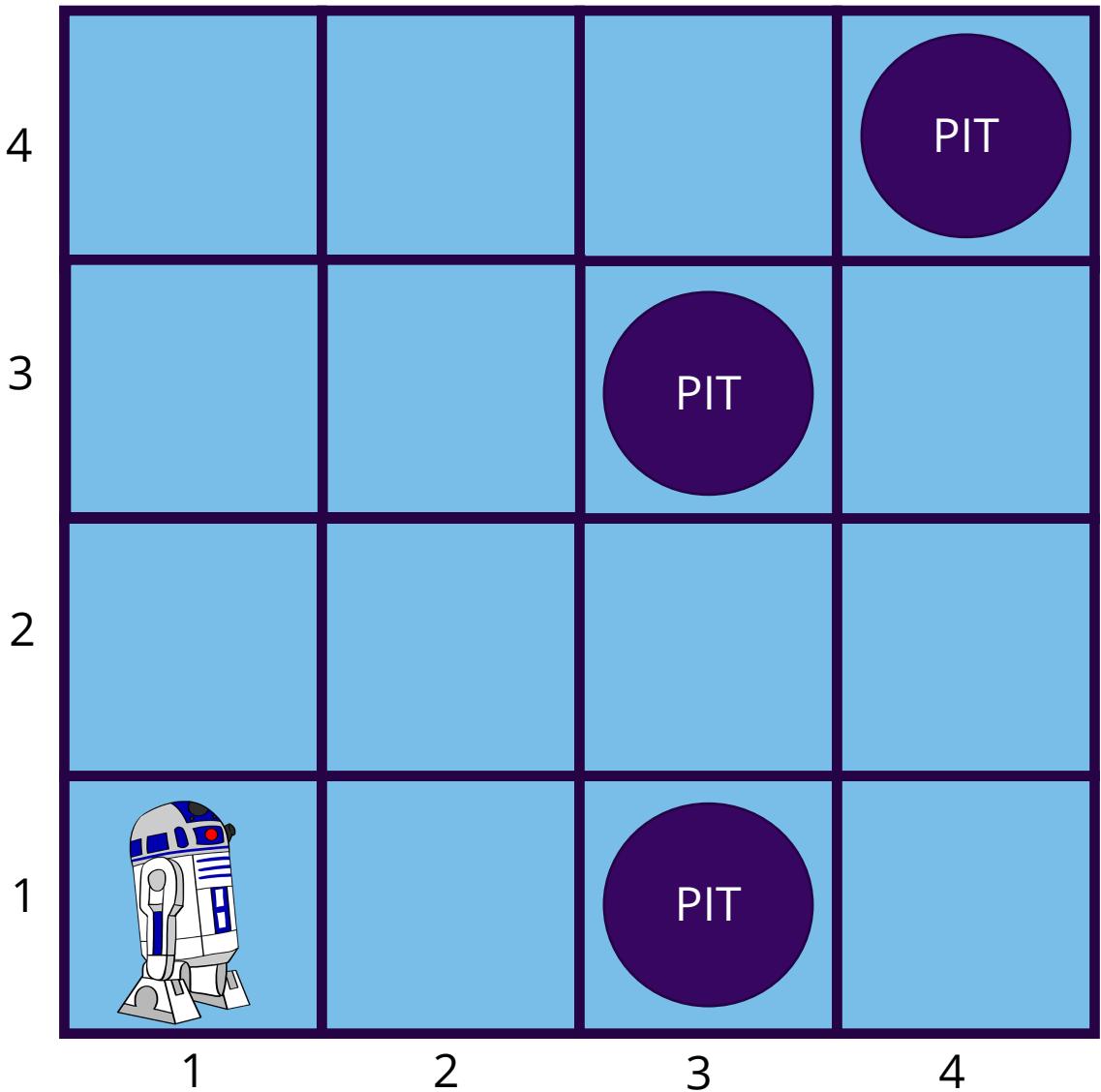
The goal is:

- collect Luke
- exit the world
- without being eaten



Wampa World

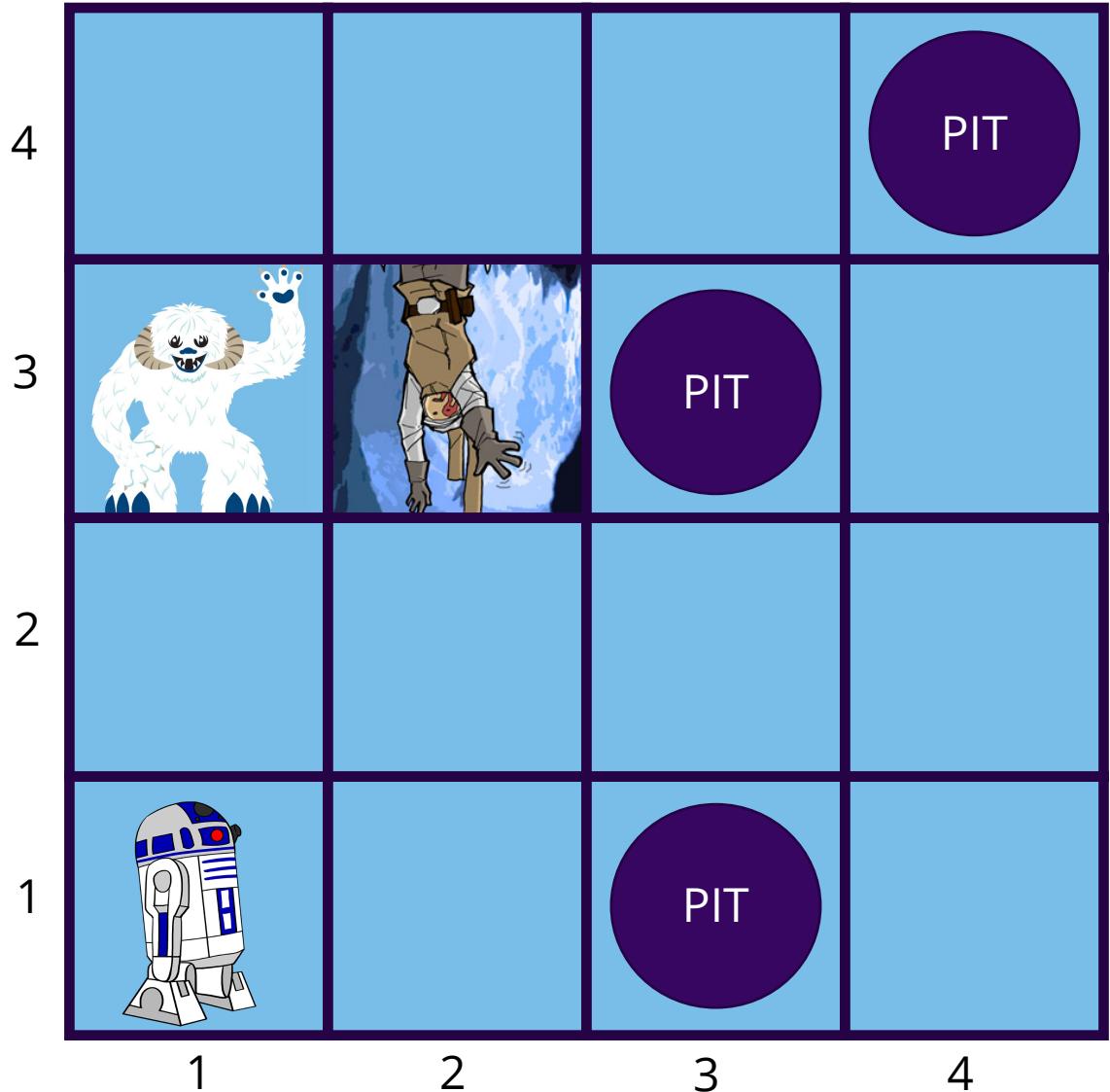
Environment: A 4x4 grid of rooms. The agent starts in the square [1,1]. Wampa and Luke are randomly placed in other squares. Each square can be pit with 20% probability.



Wampa World

Performance measure:

- +1000 points for rescuing Luke and leaving the cave
- 1000 for falling into a pit or being eaten by the Wampa
- 1 for each action taken
- 10 for using up your blaster fire



Wampa World

Actuators:

R2 can move *Forward*, *TurnLeft*, *Turn right*.

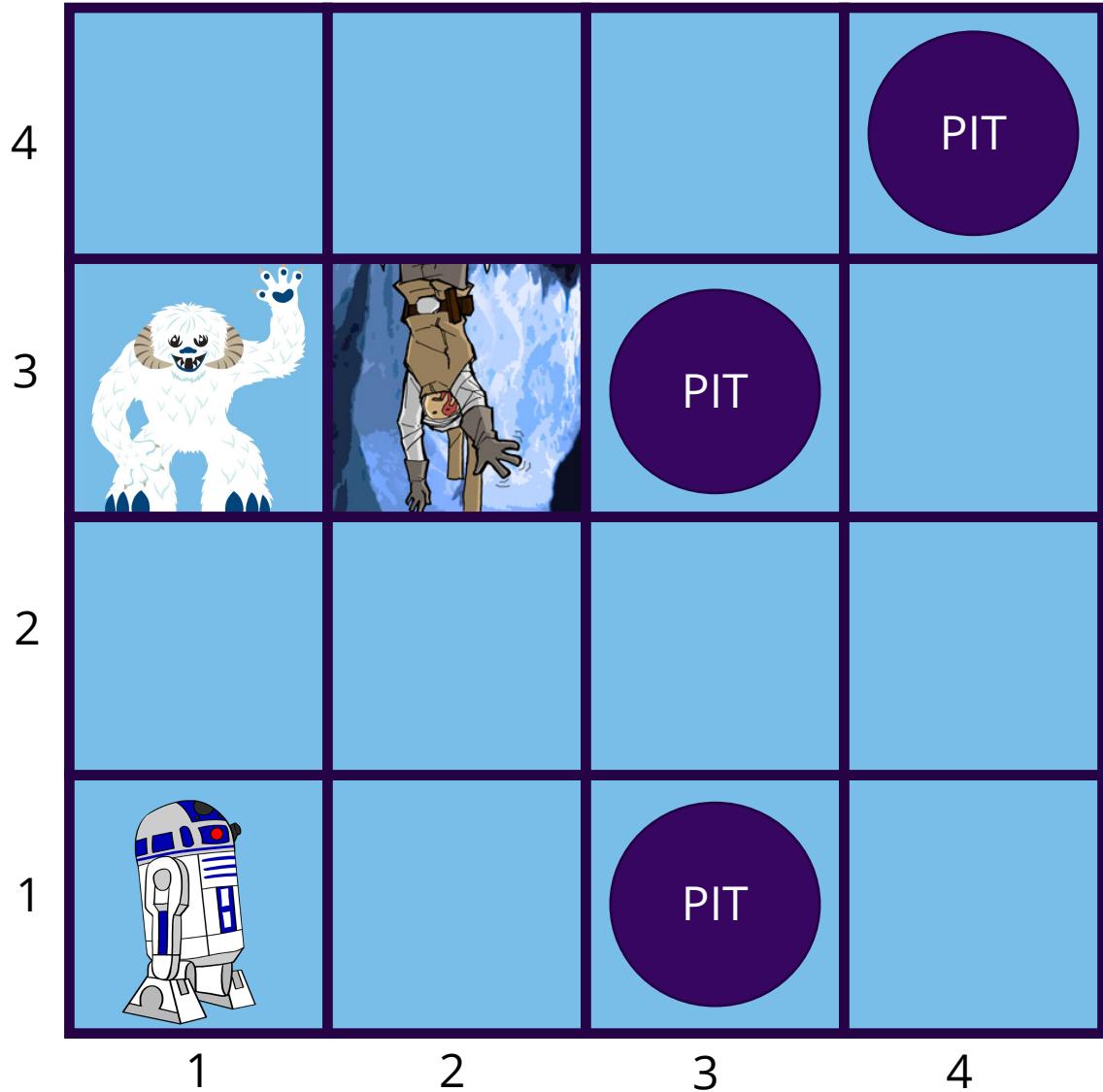
Agent dies if it moves into a pit or a Wumpus square.

Grab can pick up Luke.

Shoot fires blaster bolt in a straight line in the direction that R2D2 is facing.

If the blaster hits the Wampa, it dies. R2 only has enough power for one shot.

Climb gets R2 out of the cave but only works in [1, 1]



Wampa World

Sensors:

In each square adjacent to the Wampa, R2D2's olfactory sensor perceives a *Stench*

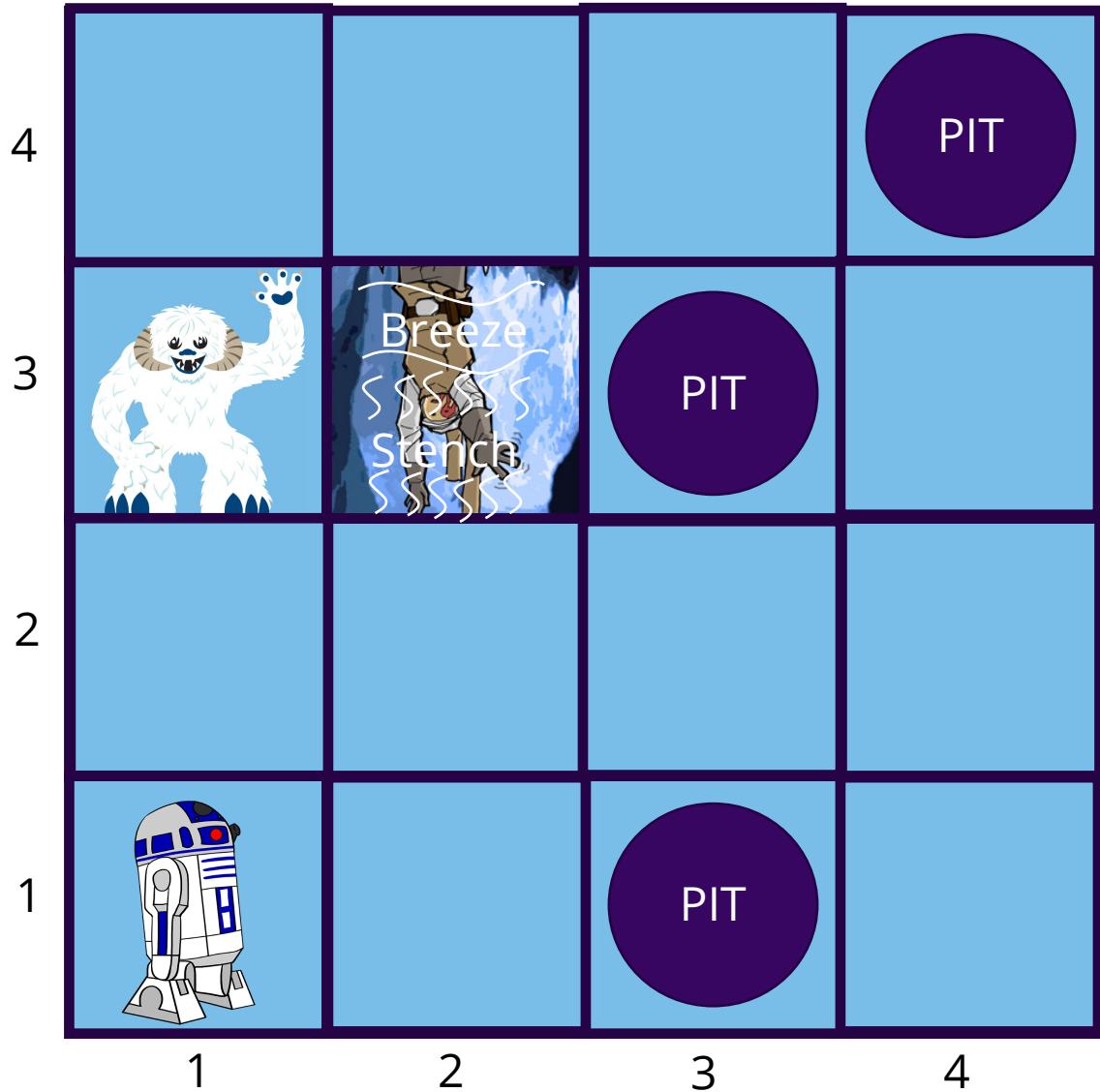
In each square adjacent to a pit, R2D2's wind sensor perceives a *Breeze*

In the square with Luke, R2D2's audio sensor perceives a *Groan*

When R2D2 walks into a wall it perceives a *Bump*

When the Wampa is killed , R2D2's audio sensor perceives a *Scream*

Percept=[*Stench*, *Breeze*, *None*, *None* *None*]



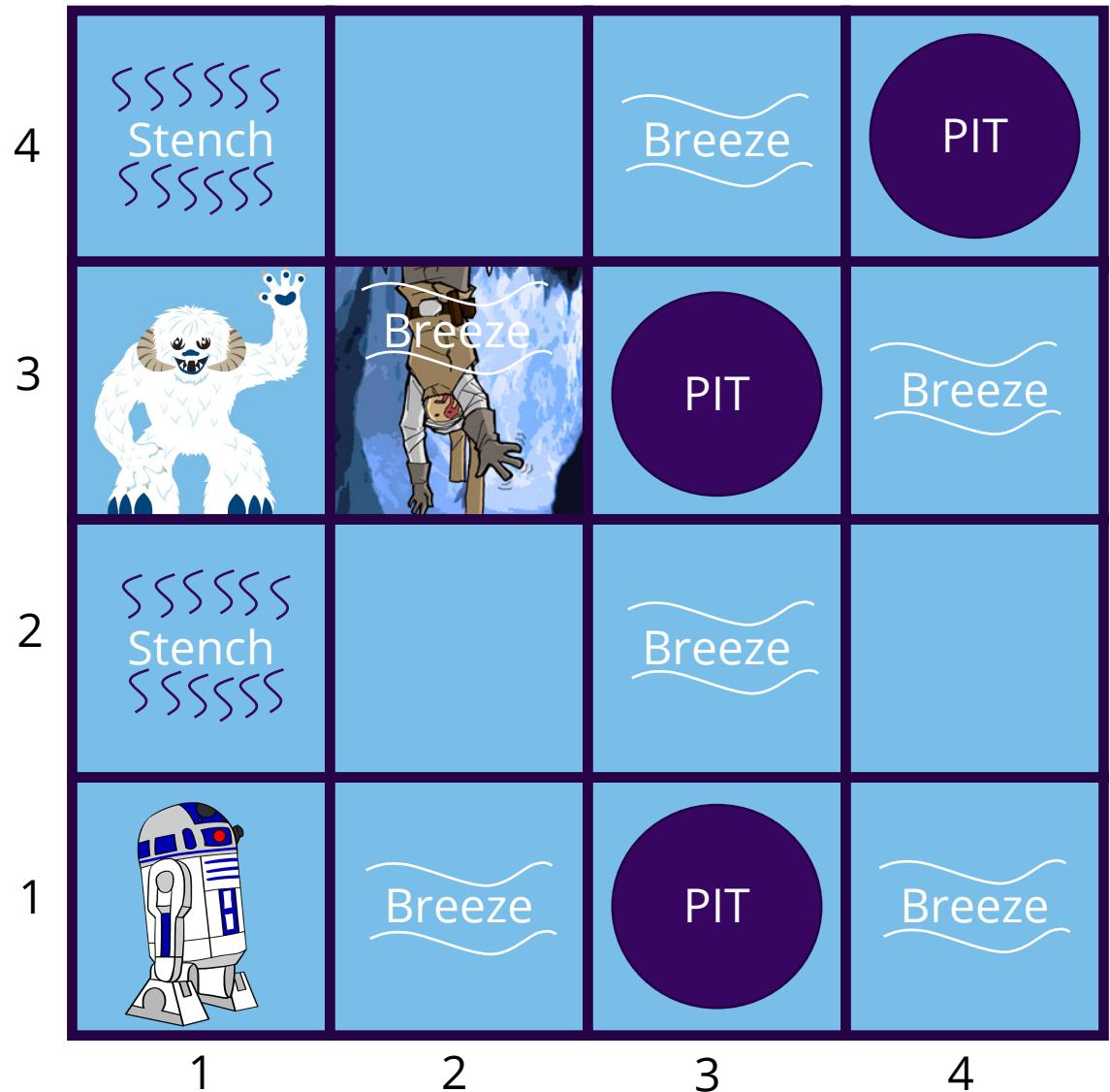
Wampa World

Deterministic, discrete, static, single-agent (Wampa doesn't move)

Sequential because reward doesn't come for many steps

Partially observable because some parts of the state are not directly perceptible:

- Location of Luke, Wampa, and pits aren't directly observable.

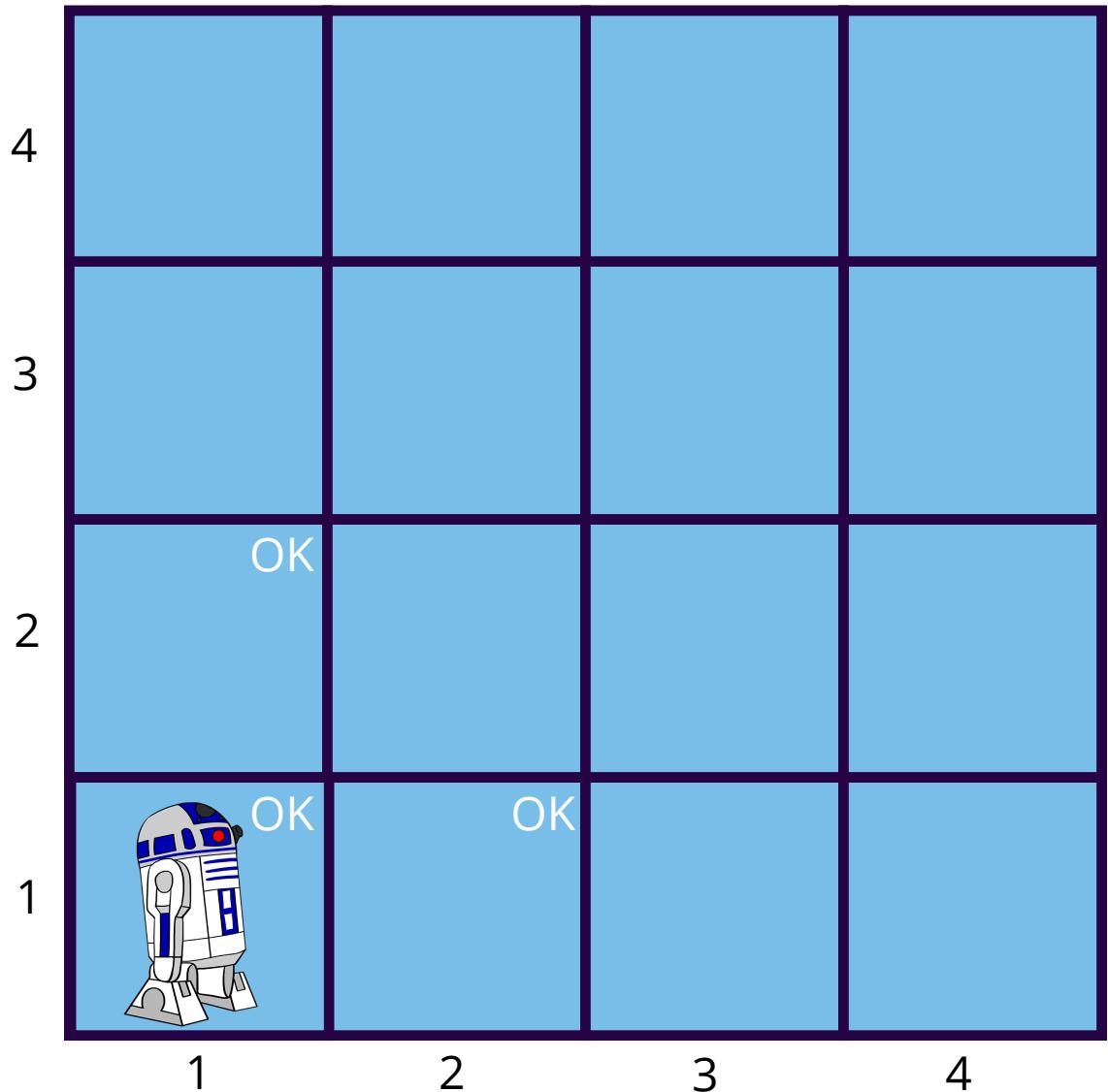


Wampa World Walkthrough

R2D2 starts in [1,1]

Percept=[None, None, None, None None]

What can we conclude about [1,2] and [2,1]?

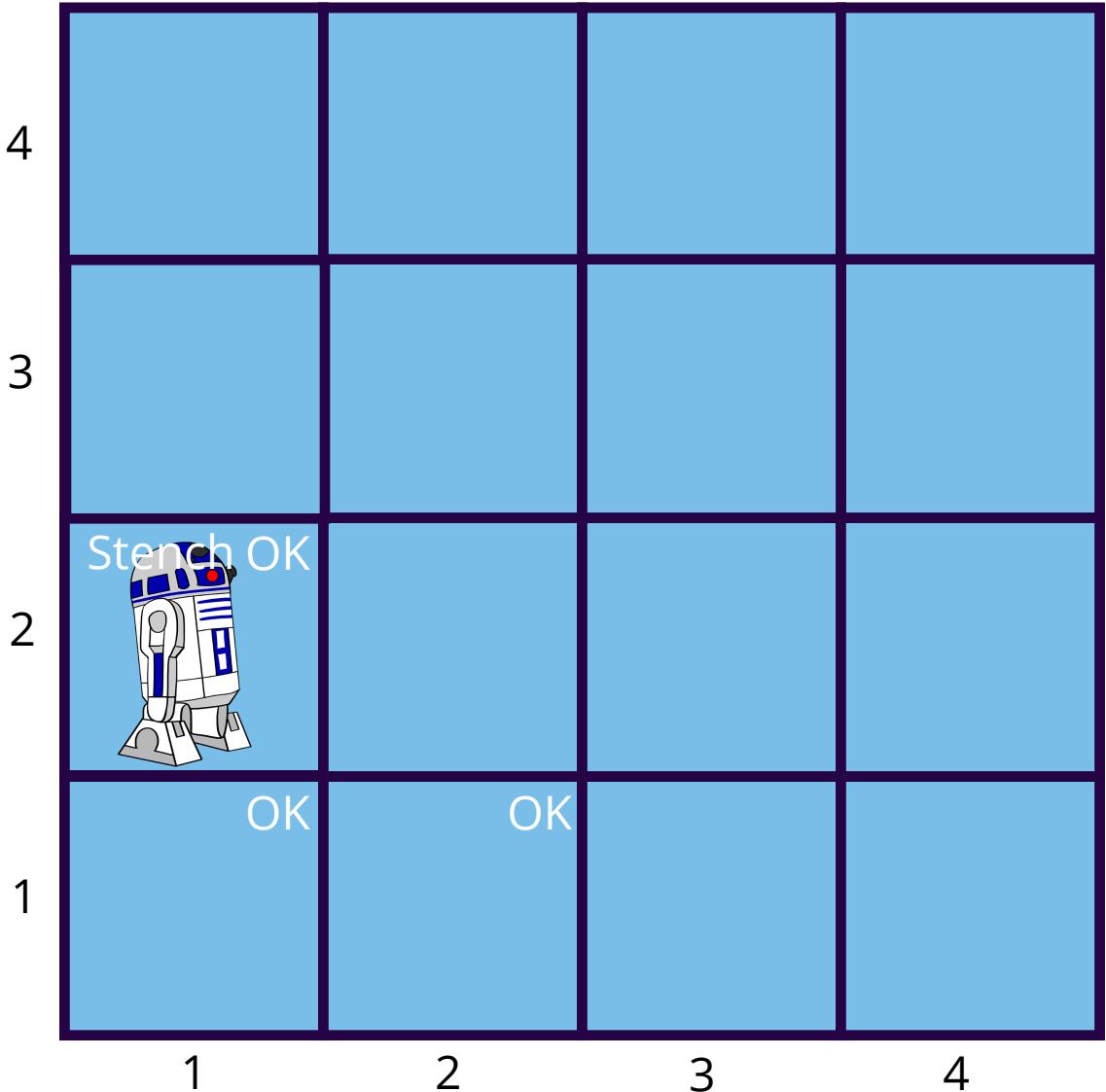


Wampa World Walkthrough

R2D2 moves to [1,2]

Percept=[*Stench*, *None*, *None*, *None* *None*]

What can we conclude about [3,1]?



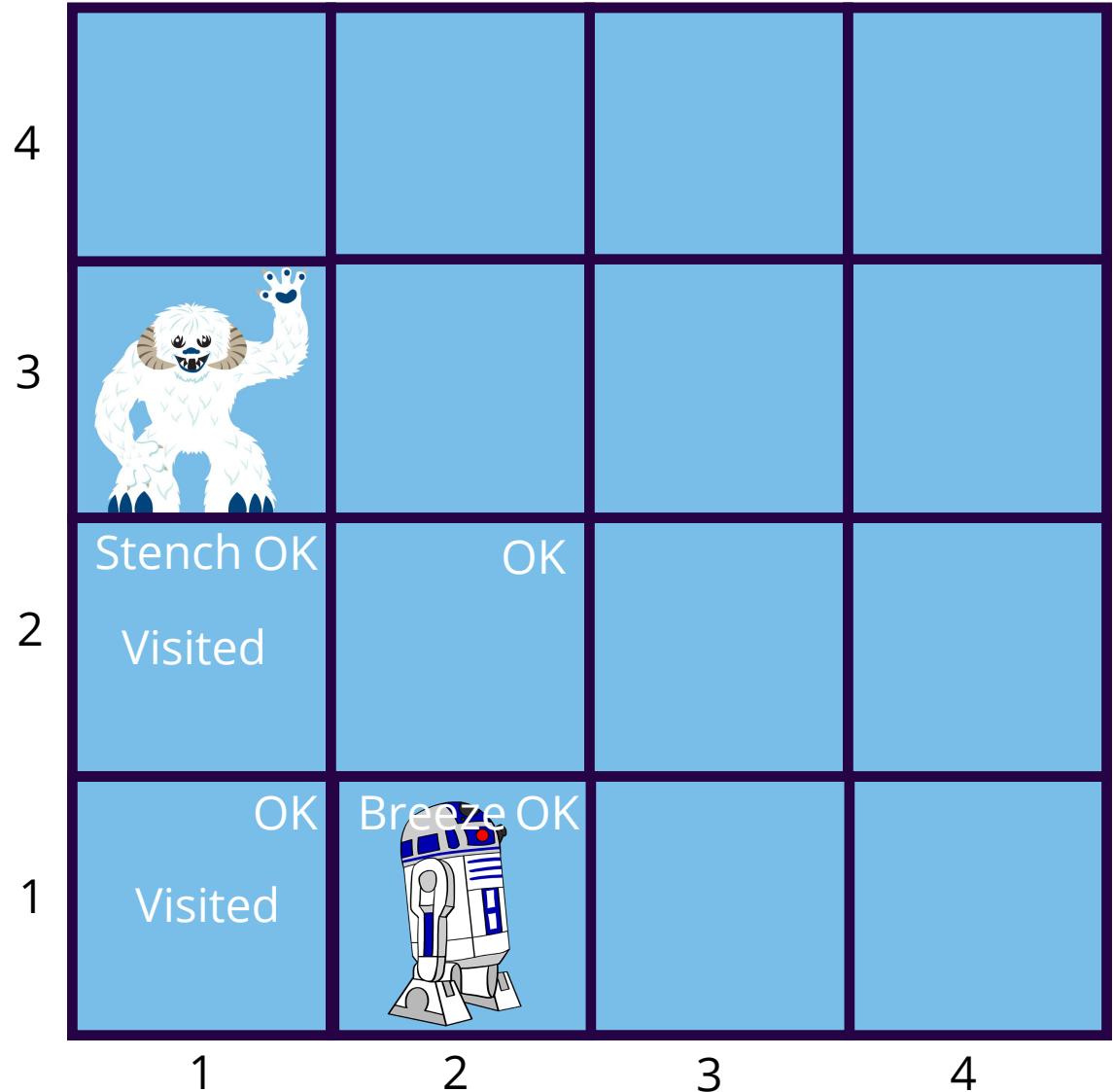
Wampa World Walkthrough

R2D2 moves to [2,1]

Percept=[None, Breeze, None, None None]

What can we conclude about [3,1]?

What can we conclude about [2,2]?

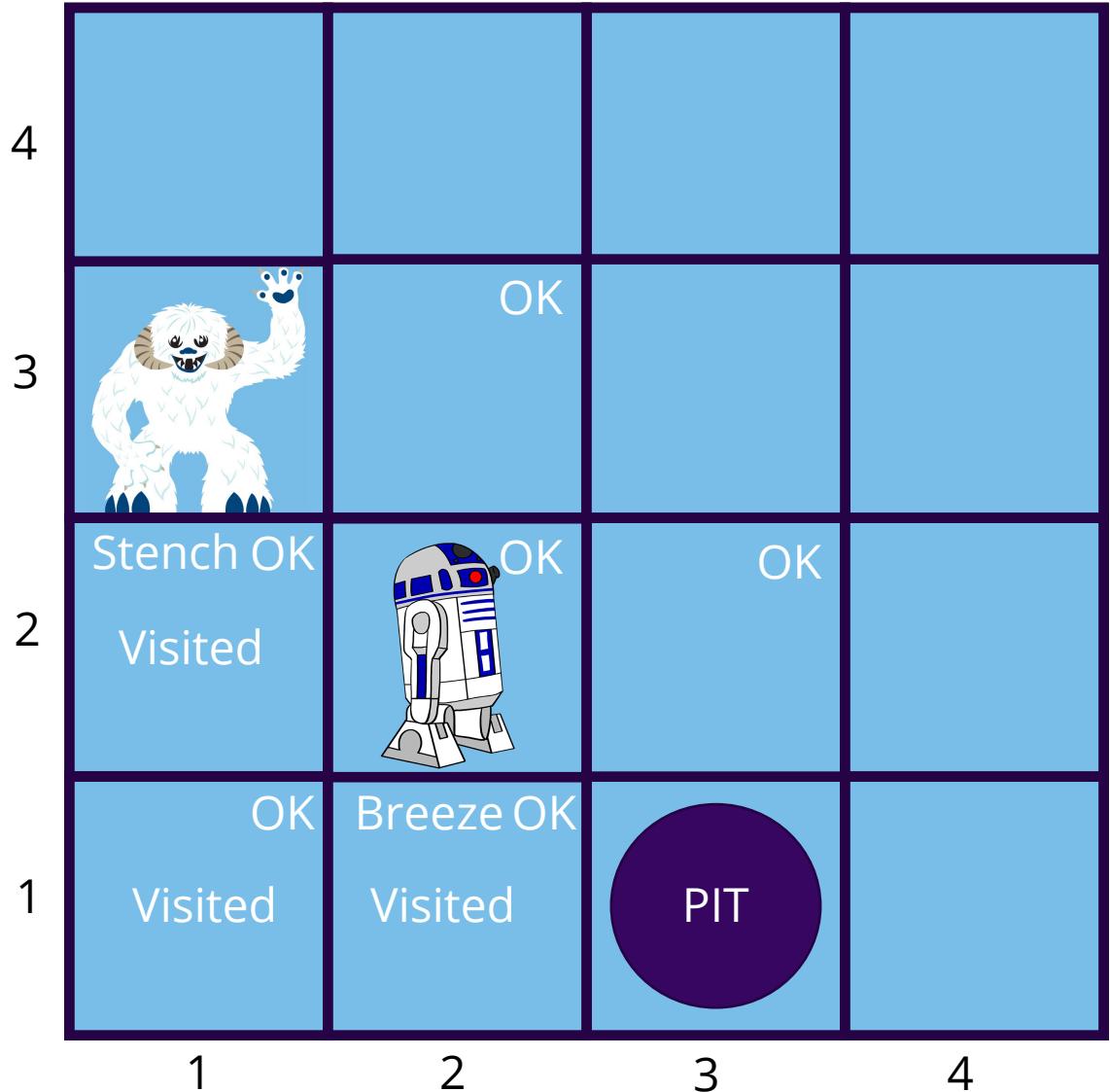


Wampa World Walkthrough

R2D2 moves to [2,2]

Percept=[None, None, None, None None]

What can we conclude about [3,2] and [2,3]?



Wampa World Walkthrough

R2D2 moves to [2,2]

Percept=[*Stench, Breeze, Gasp, None None*]

Who is here?

What is in [2,4] and [3,3]?

		Pit?	
	Stench OK Breeze	Pit?	
	Stench OK Visited	OK Visited	OK
1	OK Visited	OK Visited	PIT