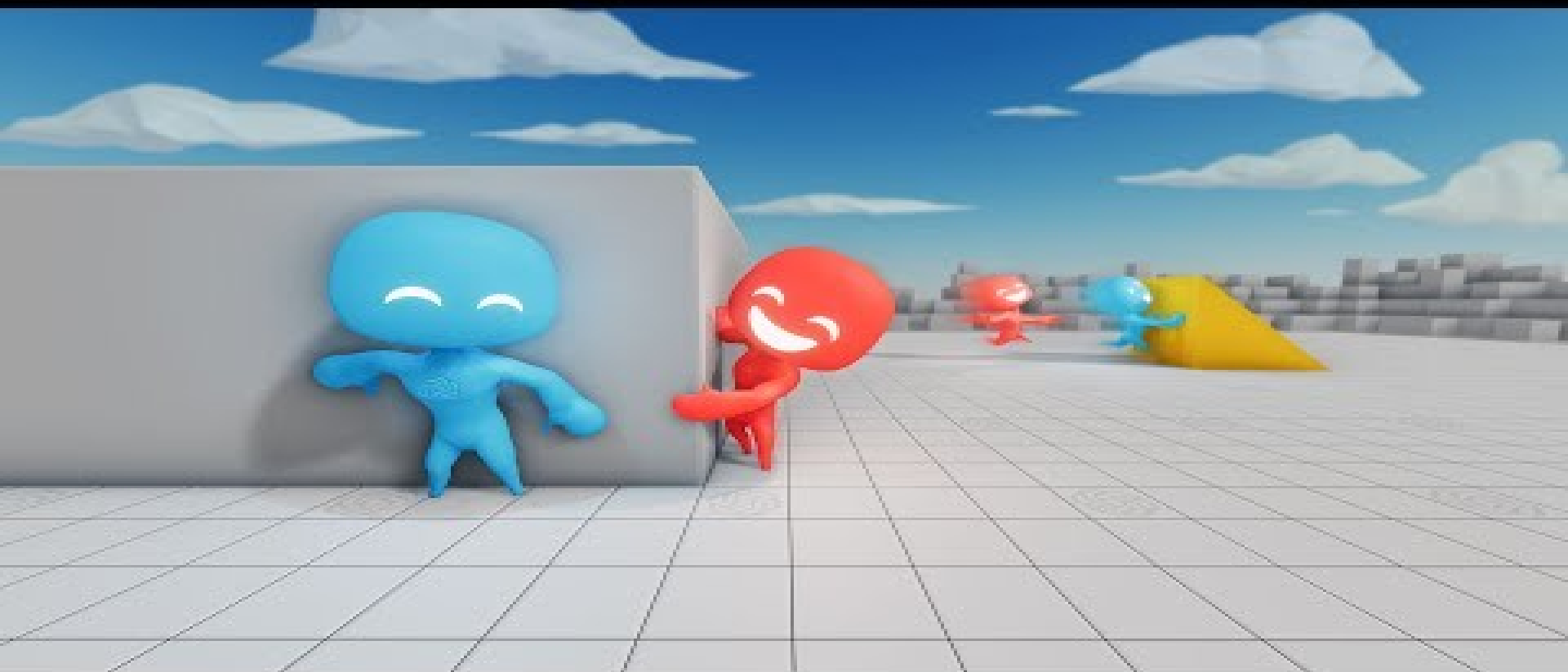


The background is a dark blue gradient. It is decorated with white and yellow circuit-like lines. Some lines are straight, while others are stepped or curved. Small circles, representing solder points or components, are placed at the ends of these lines. A prominent yellow line starts from the left edge, goes up, then right, ending in a yellow circle. Another yellow line starts from the top left, goes right, then down, ending in a yellow circle. A blue line starts from the top center, goes down, then right, ending in a blue circle. On the right side, there is a large circular graphic with concentric white and blue rings. Inside these rings, there is a small white circle with a yellow border, and a black dot with a white border. A white line with a black dot at its end points towards this central target. 

# Reinforcement Learning

The first steps... with Q-learning!





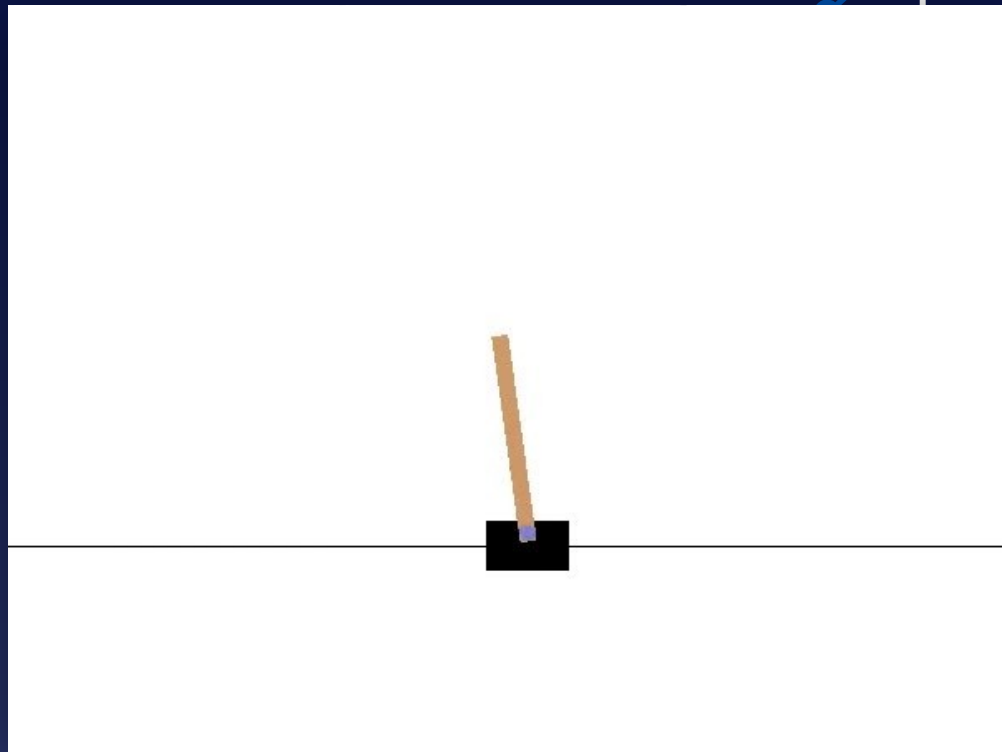
Reinforcement learning allows an  
**agent** to learn how to behave in  
an **environment**.





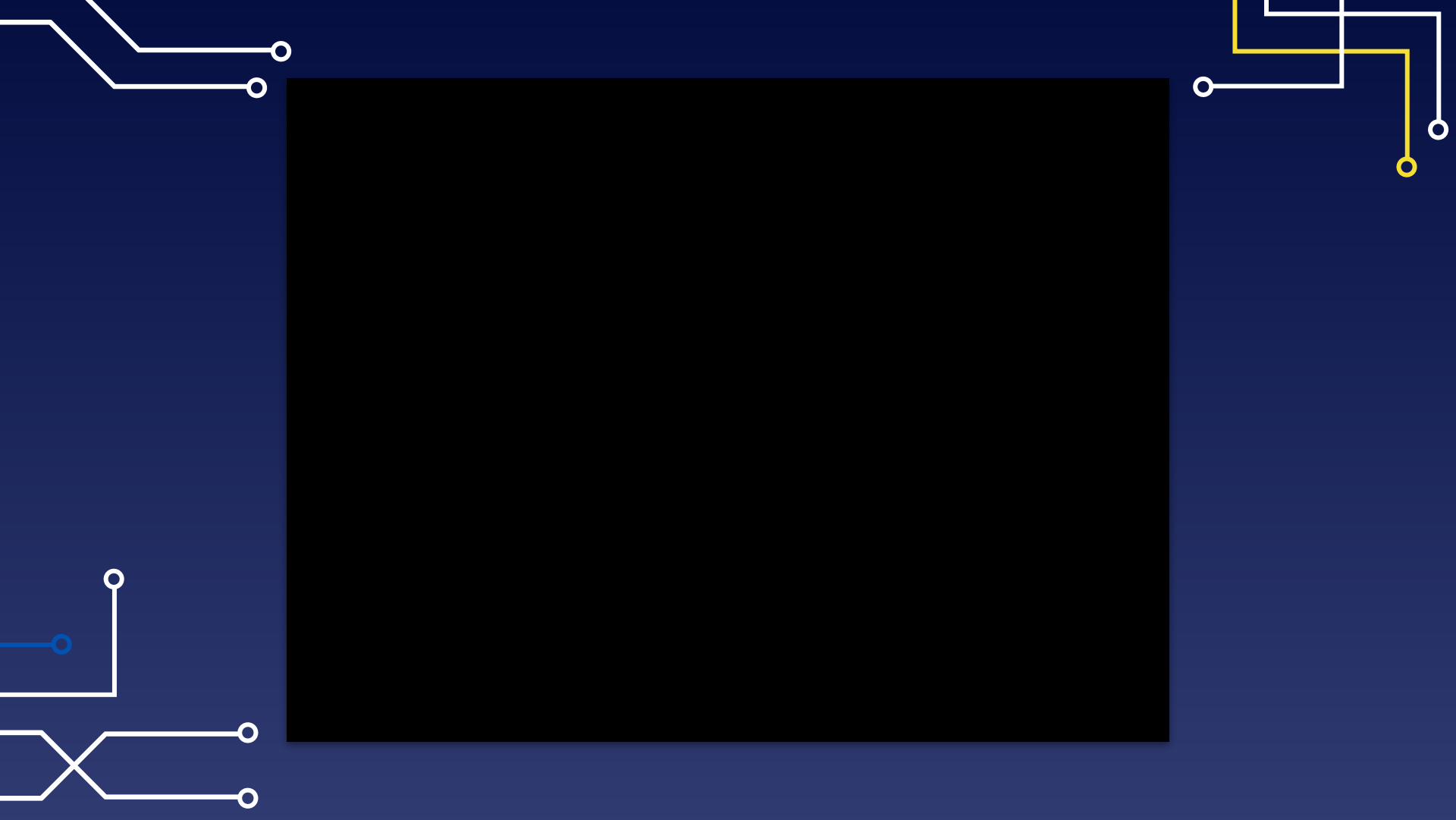
# OpenAI Gym

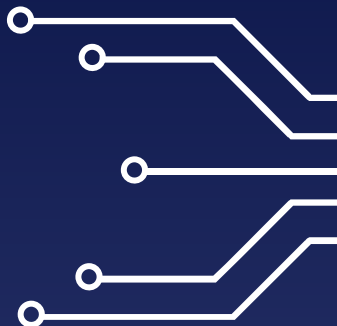

An open-source set of  
**environments** for reinforcement  
learning



**CartPole-v1**







How does the agent know  
how well it is doing?

We give it **rewards!**





# Two Competing Goals



## Exploration

Experiment and try new things! This is how you discover new strategies.

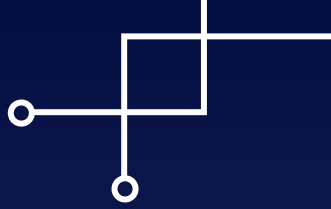


## Exploitation

Apply what you've learned to do as well as possible.



# Q-Learning










# Q-Table

Actions










States

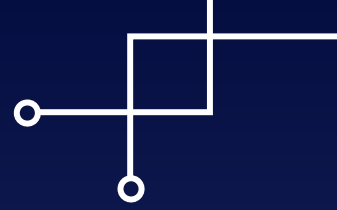
|  | States   |   |
|---|--|---|
|   | LEANING →  | LEANING ←   |
| DRIVE RIGHT →   |  |  |
| ← DRIVE LEFT  |  |  |

# Q-Table

Actions

States

|  | LEANING →<br>TURNING ↻  | LEANING →<br>TURNING ↺   | LEANING ←<br>TURNING ↻  | LEANING ←<br>TURNING ↺  |
|---|---|--|---|---|
| DRIVE RIGHT →   |  |  |  |  |
| ← DRIVE LEFT  |  |  |  |  |












How do we  
create this  
table?



# Q-Table

Actions










States

|  | LEANING →<br>TURNING ∩  | LEANING →<br>TURNING ∪   | LEANING ←<br>TURNING ∩  | LEANING ←<br>TURNING ∪  |
|---|---|--|---|---|
| DRIVE RIGHT →   |  |  |  |  |
| ← DRIVE LEFT  |  |  |  |  |

# Q-Table

Actions

States

|  | LEANING →<br>TURNING ↻  | LEANING →<br>TURNING ↺   | LEANING ←<br>TURNING ↻  | LEANING ←<br>TURNING ↺  |
|---|---|--|---|---|
| DRIVE RIGHT →   |  |  |  |  |
| ← DRIVE LEFT  |  |  |  |  |

# Q-Table

Actions

States

|               | LEANING →<br>TURNING ↷   | LEANING →<br>TURNING ↶   | LEANING ←<br>TURNING ↷   | LEANING ←<br>TURNING ↶   |
|---------------|--|--|--|--|
| DRIVE RIGHT → | <div><div></div><div></div><div></div><div></div><div></div></div> | <div><div></div><div></div><div></div><div></div><div></div></div> | <div><div></div><div></div><div></div><div></div><div></div></div> | <div><div></div><div></div><div></div><div></div><div></div></div> |
| ← DRIVE LEFT  | <div><div></div><div></div><div></div><div></div><div></div></div> | <div><div></div><div></div><div></div><div></div><div></div></div> | <div><div></div><div></div><div></div><div></div><div></div></div> | <div><div></div><div></div><div></div><div></div><div></div></div> |

Choose a random action

# Q-Table

Actions

States









|               | LEANING →<br>TURNING ↷   | LEANING →<br>TURNING ↶   | LEANING ←<br>TURNING ↷   | LEANING ←<br>TURNING ↶   |
|---------------|--|--|--|--|
| DRIVE RIGHT → | <div><div></div><div></div><div></div><div></div><div></div></div> | <div><div></div><div></div><div></div><div></div><div></div></div> | <div><div></div><div></div><div></div><div></div><div></div></div> | <div><div></div><div></div><div></div><div></div><div></div></div> |
| ← DRIVE LEFT  | <div><div></div><div></div><div></div><div></div><div></div></div> | <div><div></div><div></div><div></div><div></div><div></div></div> | <div><div></div><div></div><div></div><div></div><div></div></div> | <div><div></div><div></div><div></div><div></div><div></div></div> |

See where it sends you

# Q-Table

Actions

States

|               | LEANING →<br>TURNING ↷  | LEANING →<br>TURNING ↶   | LEANING ←<br>TURNING ↷  | LEANING ←<br>TURNING ↶  |
|---------------|---|--|---|---|
| DRIVE RIGHT → |  |  |  |  |
| ← DRIVE LEFT  |  |  |  |  |

Learn from your reward














# Q-Table

Actions

States

|  | LEANING →<br>TURNING ↻  | LEANING →<br>TURNING ↻   | LEANING ←<br>TURNING ↻  | LEANING ←<br>TURNING ↻  |
|---|---|--|---|---|
| DRIVE RIGHT →   |  |  |  |  |
| ← DRIVE LEFT  |  |  |  |  |

# Q-Table

Actions

States

|               | LEANING →<br>TURNING ↷ | LEANING →<br>TURNING ↶ | LEANING ←<br>TURNING ↷ | LEANING ←<br>TURNING ↶ |
|---------------|------------------------|------------------------|------------------------|------------------------|
| DRIVE RIGHT → | ■ ■ ■ ■ ■              | □ □ □ □ □              | □ □ □ □ □              | □ □ □ □ □              |
| ← DRIVE LEFT  | □ □ □ □ □              | □ □ □ □ □              | □ □ □ □ □              | □ □ □ □ □              |

Choose a random action

# Q-Table

Actions

States

|               | LEANING →<br>TURNING ↷ | LEANING →<br>TURNING ↶ | LEANING ←<br>TURNING ↷ | LEANING ←<br>TURNING ↶ |
|---------------|------------------------|------------------------|------------------------|------------------------|
| DRIVE RIGHT → | ■ ■ ■ ■ ■              | □ □ □ □ □              | □ □ □ □ □              | □ □ □ □ □              |
| ← DRIVE LEFT  | □ □ □ □ □              | □ □ □ □ □              | □ □ □ □ □              | □ □ □ □ □              |

See where it sends you

# Q-Table

Actions

States

|               | LEANING →<br>TURNING ↷ | LEANING →<br>TURNING ↶ | LEANING ←<br>TURNING ↷ | LEANING ←<br>TURNING ↶ |
|---------------|------------------------|------------------------|------------------------|------------------------|
| DRIVE RIGHT → | ■ ■ ■ ■ ■              | ■ ■ ■ ■ ■              | □ □ □ □ □              | □ □ □ □ □              |
| ← DRIVE LEFT  | □ □ □ □ □              | □ □ □ □ □              | □ □ □ □ □              | □ □ □ □ □              |










Learn from your reward



# Q-Table

Actions

States

|  | LEANING →<br>TURNING ↷  | LEANING →<br>TURNING ↶   | LEANING ←<br>TURNING ↷  | LEANING ←<br>TURNING ↶  |
|---|---|--|---|---|
| DRIVE RIGHT →   |  |  |  |  |
| ← DRIVE LEFT  |  |  |  |  |

# Q-Table

Actions

States

|               | LEANING →<br>TURNING ↷ | LEANING →<br>TURNING ↶ | LEANING ←<br>TURNING ↷ | LEANING ←<br>TURNING ↶ |
|---------------|------------------------|------------------------|------------------------|------------------------|
| DRIVE RIGHT → | ■ ■ ■ ■ ■              | ■ ■ ■ ■ ■              | □ □ □ □ □              | □ □ □ □ □              |
| ← DRIVE LEFT  | □ □ □ □ □              | □ □ □ □ □              | □ □ □ □ □              | □ □ □ □ □              |

Choose a random action

# Q-Table

Actions

States










|               | LEANING →<br>TURNING ↷ | LEANING →<br>TURNING ↶ | LEANING ←<br>TURNING ↷ | LEANING ←<br>TURNING ↶ |
|---------------|------------------------|------------------------|------------------------|------------------------|
| DRIVE RIGHT → | ■ ■ ■ ■ ■              | ■ ■ ■ ■ ■              | □ □ □ □ □              | □ □ □ □ □              |
| ← DRIVE LEFT  | □ □ □ □ □              | □ □ □ □ □              | □ □ □ □ □              | □ □ □ □ □              |

See where it sends you

# Q-Table

Actions

States

|  | States  |  |   |   |
|---|---|--|---|---|
|   | LEANING →<br>TURNING ↷  | LEANING →<br>TURNING ↶   | LEANING ←<br>TURNING ↷  | LEANING ←<br>TURNING ↶  |
| DRIVE RIGHT →   |  |  |  |  |
| ← DRIVE LEFT  |  |  |  |  |












Learn from your reward



# Q-Table

Actions










States

|  | LEANING →<br>TURNING ∪  | LEANING →<br>TURNING ∩   | LEANING ←<br>TURNING ∪  | LEANING ←<br>TURNING ∩  |
|---|---|--|---|---|
| DRIVE RIGHT →   |  |  |  |  |
| ← DRIVE LEFT  |  |  |  |  |

# Q-Table

Actions

States









|  | LEANING →<br>TURNING ↷  | LEANING →<br>TURNING ↻   | LEANING ←<br>TURNING ↷  | LEANING ←<br>TURNING ↻  |
|---|---|--|---|---|
| DRIVE RIGHT →   |  |  |  |  |
| ← DRIVE LEFT  |  |  |  |  |

Start over from a random  
state

# Q-Table

Actions

States

|               | LEANING →<br>TURNING ↷  | LEANING →<br>TURNING ↶   | LEANING ←<br>TURNING ↷  | LEANING ←<br>TURNING ↶  |
|---------------|---|--|---|---|
| DRIVE RIGHT → |  |  |  |  |
| ← DRIVE LEFT  |  |  |  |  |

Choose a random action

# Q-Table

Actions

States









|               | LEANING →<br>TURNING ↷ | LEANING →<br>TURNING ↶ | LEANING ←<br>TURNING ↷ | LEANING ←<br>TURNING ↶ |
|---------------|------------------------|------------------------|------------------------|------------------------|
| DRIVE RIGHT → | ■ ■ ■ ■ ■              | ■ ■ ■ ■ ■              | □ □ □ □ □              | ■ □ □ □ □              |
| ← DRIVE LEFT  | □ □ □ □ □              | □ □ □ □ □              | □ □ □ □ □              | □ □ □ □ □              |

See where it sends you

# Q-Table

Actions

States

|               | LEANING →<br>TURNING ↷  | LEANING →<br>TURNING ↶   | LEANING ←<br>TURNING ↷  | LEANING ←<br>TURNING ↶  |
|---------------|---|--|---|---|
| DRIVE RIGHT → |  |  |  |  |
| ← DRIVE LEFT  |  |  |  |  |









Learn from your reward...



# Q-Table

Actions

States

|               | LEANING →<br>TURNING ↷  | LEANING →<br>TURNING ↶   | LEANING ←<br>TURNING ↷  | LEANING ←<br>TURNING ↶  |
|---------------|---|--|---|---|
| DRIVE RIGHT → |  |  |  |  |
| ← DRIVE LEFT  |  |  |  |  |

...and learn from your new state



# The Process

Start the simulation from a random state. Then...

1. Choose a random action
2. See where it sends you
3. Learn from the reward you get

# The Process

Start the simulation from a random state. Then...

1. Choose a random action
2. See where it sends you
3. Learn from the reward you get  
...and the new state you end up in



# The Process


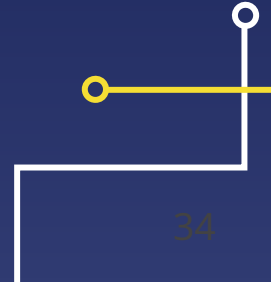
Start the simulation from a random state. Then...

1. Choose a random action  
...or a good action (once you're a little smarter)
2. See where it sends you
3. Learn from the reward you get  
...and the new state you end up in

# The Process



Start the simulation from a random state. Then...

1. Choose a random action  
...or a good action (once you're a little smarter)
  2. See where it sends you
  3. Learn from the reward you get  
...and the new state you end up in  
...learning more slowly over time
- 
- 












(I cheated a tiny  
bit...)



# Q-Table

Actions

States

|  | LEANING →<br>TURNING ↷  | LEANING →<br>TURNING ↶   | LEANING ←<br>TURNING ↷  | LEANING ←<br>TURNING ↶  |
|---|---|--|---|---|
| DRIVE RIGHT →   |  |  |  |  |
| ← DRIVE LEFT  |  |  |  |  |

Learning Turning

Action  
n





**If this made sense, great!**

But if not, don't worry.

