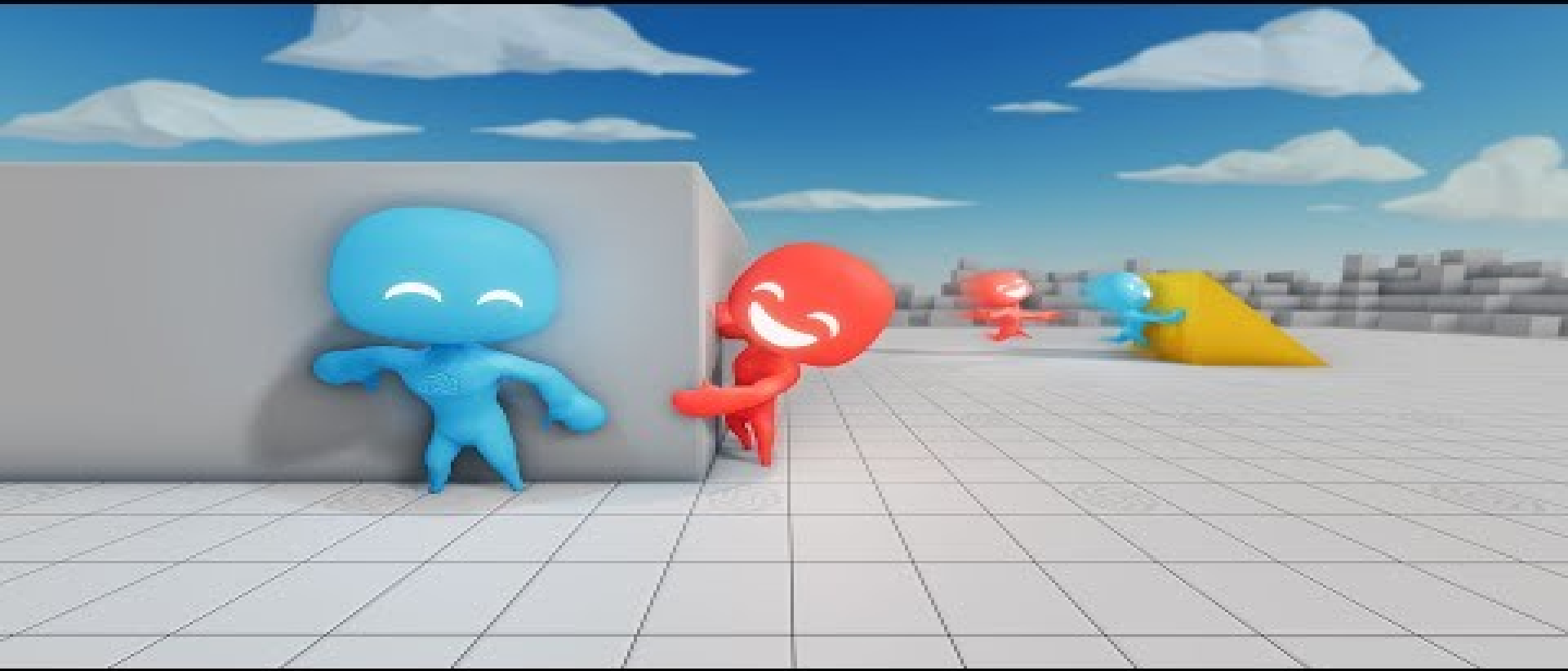
Reinforcement Learning

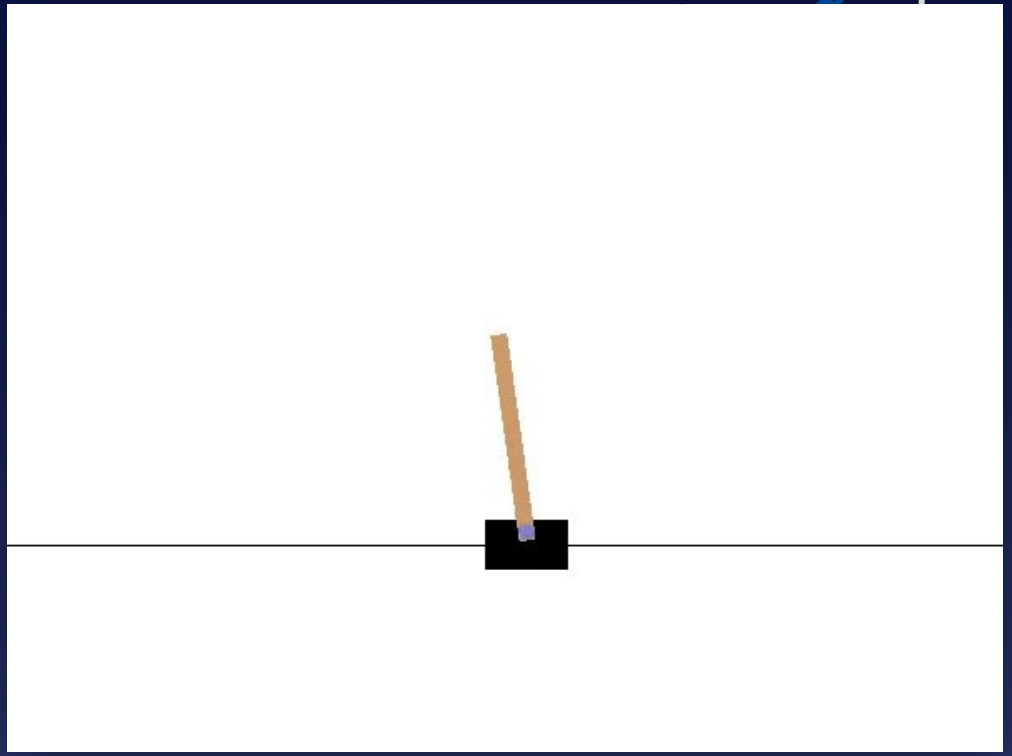
The first steps... with Q-learning!



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 →

	LEANING →	LEANING ←
DRIVE RIGHT →		
← DRIVE LEFT		

Q-Table

Actions ↓ | States →

	LEANING → TURNING ↻	LEANING → TURNING ↺	LEANING ← TURNING ↻	LEANING ← TURNING ↺
DRIVE RIGHT →				
← DRIVE LEFT				

How do we create this table?

Q-Table

Actions ↓ | States →

	LEANING → TURNING ↻	LEANING → TURNING ↺	LEANING ← TURNING ↻	LEANING ← TURNING ↺
DRIVE RIGHT →				
← DRIVE LEFT				

Q-Table

Actions

States

↓	LEANING → TURNING ↻	LEANING → TURNING ↺	LEANING ← TURNING ↻	LEANING ← TURNING ↺
DRIVE RIGHT →	□□□□□	□□□□□	□□□□□	□□□□□
← DRIVE LEFT	□□□□□	□□□□□	□□□□□	□□□□□

Q-Table

Actions

States

	LEANING → TURNING ↻	LEANING → TURNING ↺	LEANING ← TURNING ↻	LEANING ← TURNING ↺
DRIVE RIGHT →				
← DRIVE LEFT				

Choose a random action

Q-Table

Actions

States

	LEANING → TURNING ↻	LEANING → TURNING ↺	LEANING ← TURNING ↻	LEANING ← TURNING ↺
DRIVE RIGHT →				
← DRIVE LEFT				

See where it sends you

Q-Table

Actions ↓	LEANING → TURNING ↻	LEANING → TURNING ↺	LEANING ← TURNING ↻	LEANING ← TURNING ↺
DRIVE RIGHT →	■ ■ ■ ■ ■	□ □ □ □ □	□ □ □ □ □	□ □ □ □ □
← DRIVE LEFT	□ □ □ □ □	□ □ □ □ □	□ □ □ □ □	□ □ □ □ □

States

Learn from your reward

Q-Table

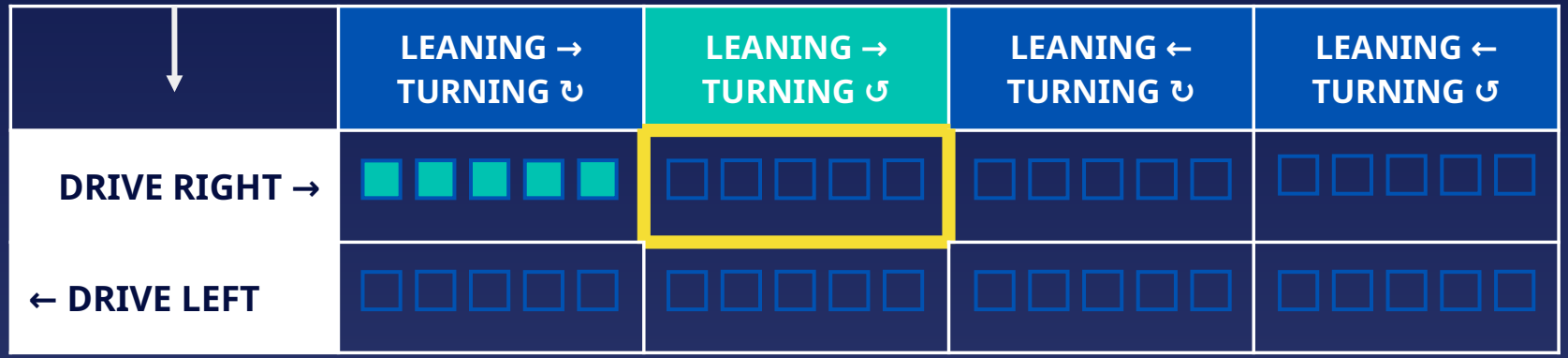
Actions ↓ | States →

	LEANING → TURNING ↻	LEANING → TURNING ↺	LEANING ← TURNING ↻	LEANING ← TURNING ↺
DRIVE RIGHT →	■ ■ ■ ■ ■	□ □ □ □ □	□ □ □ □ □	□ □ □ □ □
← DRIVE LEFT	□ □ □ □ □	□ □ □ □ □	□ □ □ □ □	□ □ □ □ □

Q-Table

Actions

States

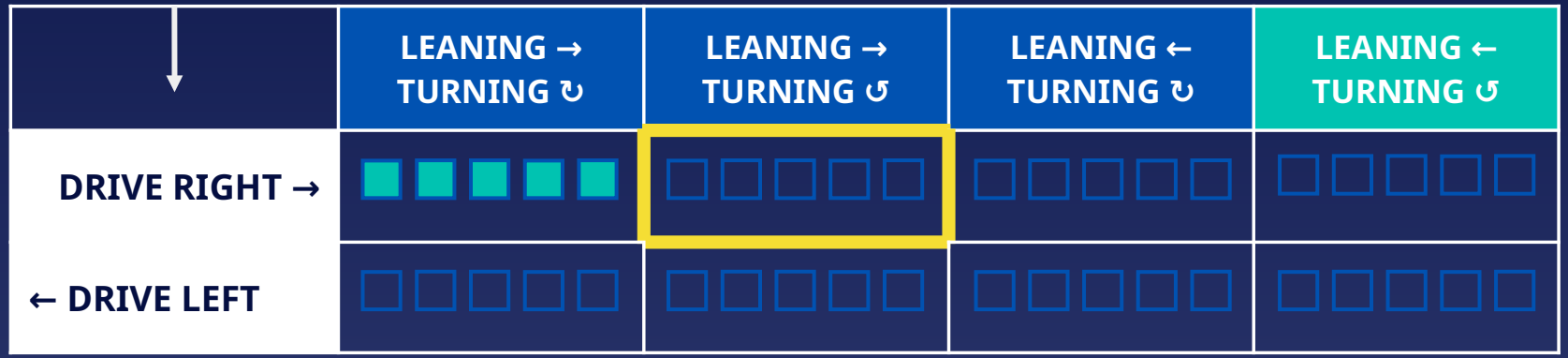


Choose a random action

Q-Table

Actions

States



See where it sends you

Q-Table

Actions

States

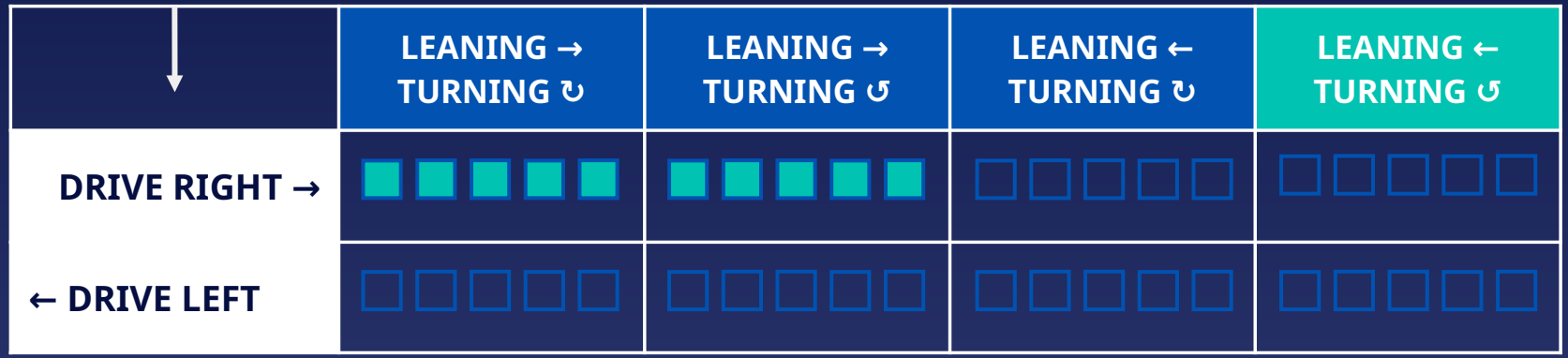
	LEANING → TURNING ↻	LEANING → TURNING ↺	LEANING ← TURNING ↻	LEANING ← TURNING ↺
DRIVE RIGHT →	■■■■■	■■■■■	□□□□□	□□□□□
← DRIVE LEFT	□□□□□	□□□□□	□□□□□	□□□□□

Learn from your reward

Q-Table

Actions

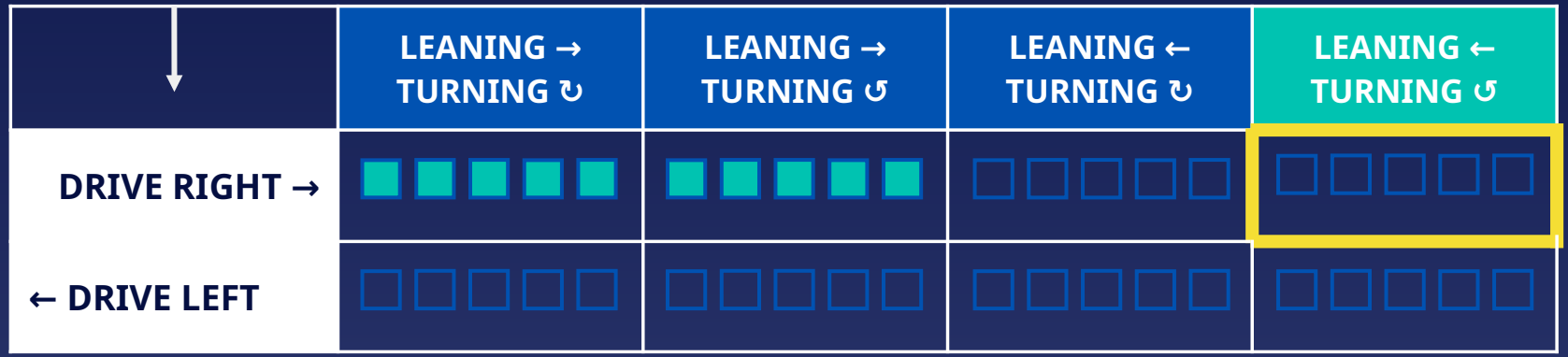
States



Q-Table

Actions

States

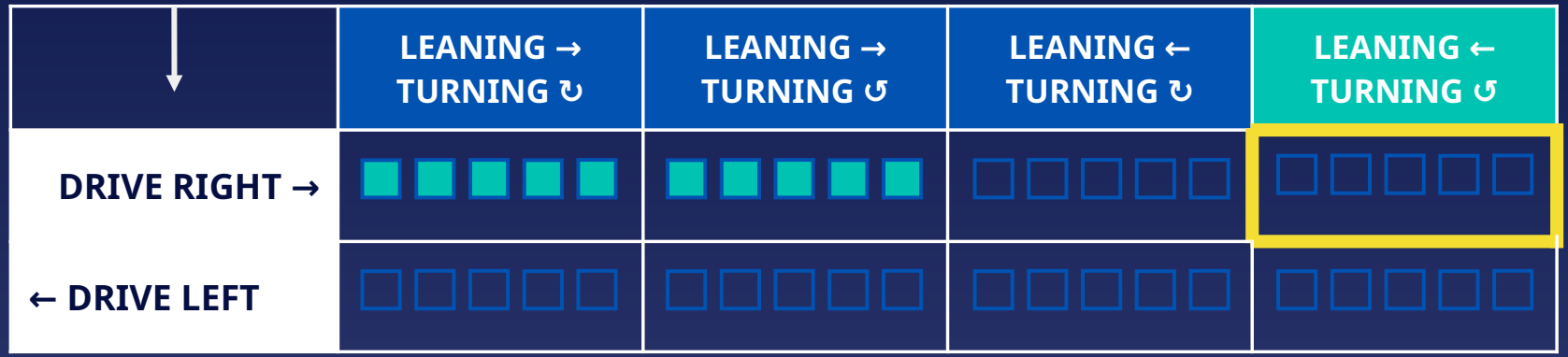


Choose a random action

Q-Table

Actions

States



See where it sends you

Q-Table

Actions

States

	LEANING → TURNING ↻	LEANING → TURNING ↺	LEANING ← TURNING ↻	LEANING ← TURNING ↺
DRIVE RIGHT →	■■■■■	■■■■■	□□□□□	■□□□□
← DRIVE LEFT	□□□□□	□□□□□	□□□□□	□□□□□

Learn from your reward

Q-Table

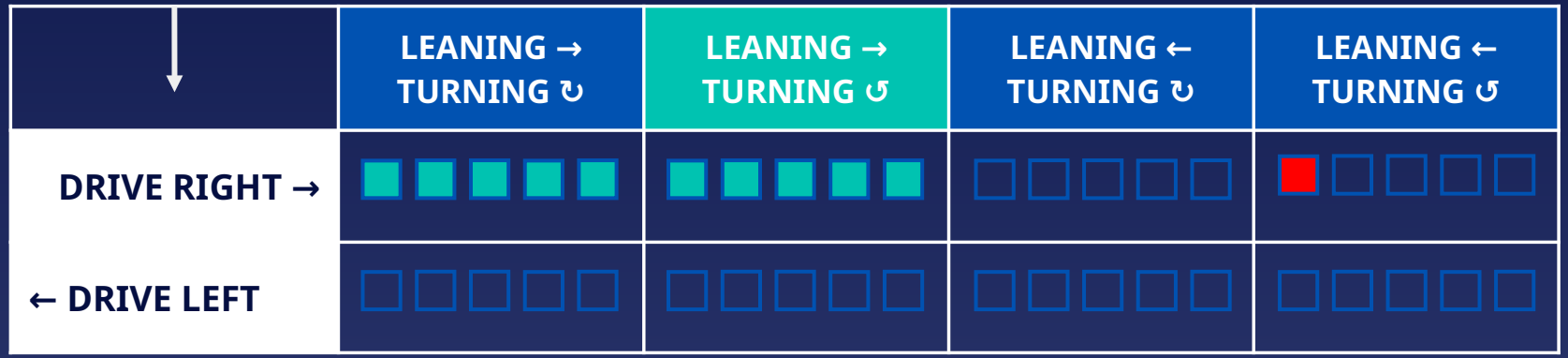
Actions ↓ | States →

	LEANING → TURNING ↻	LEANING → TURNING ↺	LEANING ← TURNING ↻	LEANING ← TURNING ↺
DRIVE RIGHT →				
← DRIVE LEFT				

Q-Table

Actions

States

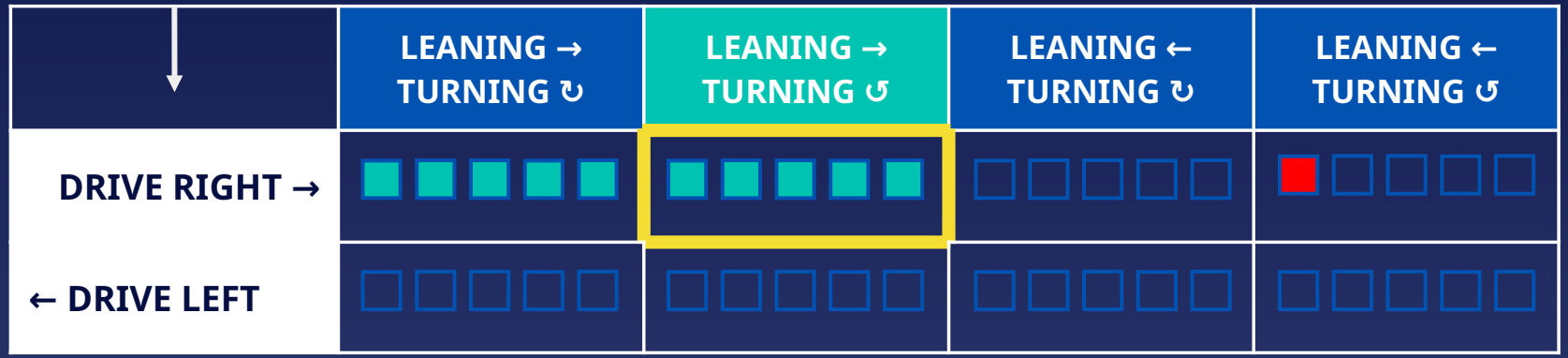


Start over from a random state

Q-Table

Actions

States

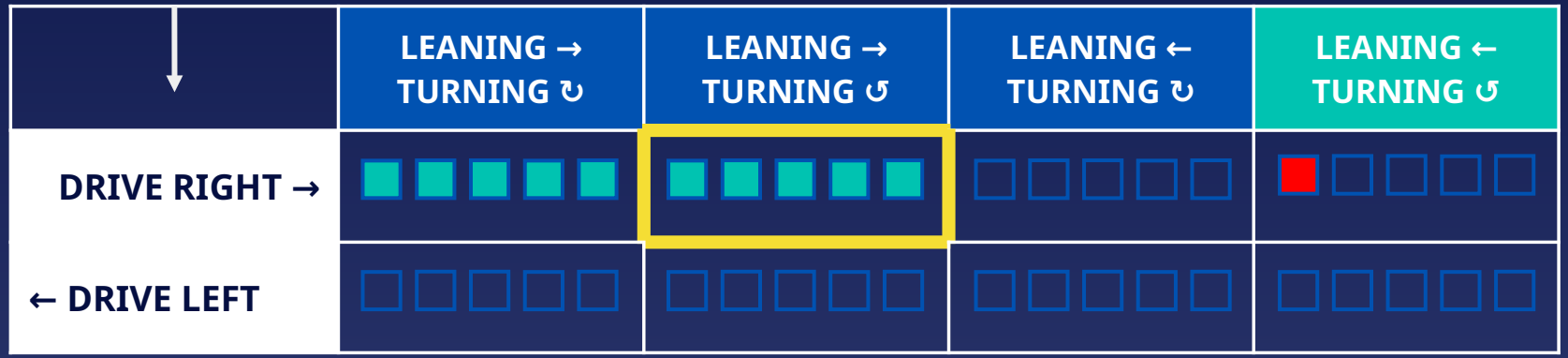


Choose a random action

Q-Table

Actions

States



See where it sends you

Q-Table

Actions ↓	LEANING → TURNING ↻	LEANING → TURNING ↺	LEANING ← TURNING ↻	LEANING ← TURNING ↺
DRIVE RIGHT →	■■■■■	■■■■■	□□□□□	■□□□□
← DRIVE LEFT	□□□□□	□□□□□	□□□□□	□□□□□

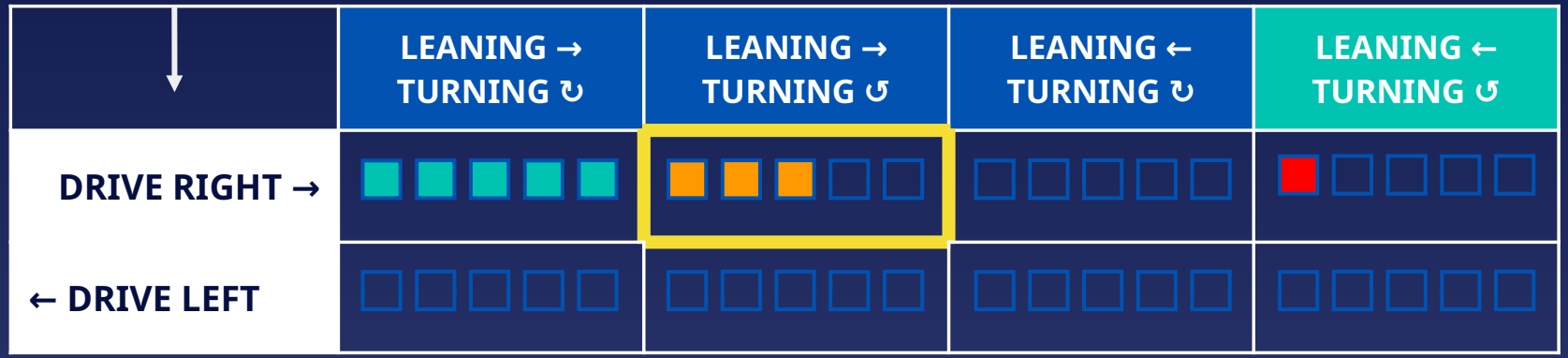
States

Learn from your reward...

Q-Table

Actions

States



...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...)

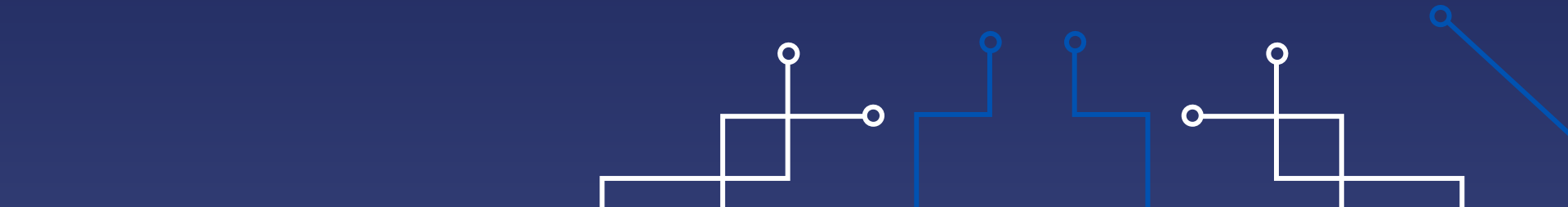
Q-Table

Actions ↓ | States →

	LEANING → TURNING ↻	LEANING → TURNING ↺	LEANING ← TURNING ↻	LEANING ← TURNING ↺
DRIVE RIGHT →				
← DRIVE LEFT				

Leaning Turning

Actio
n

If this made sense, great!

But if not, don’t worry.