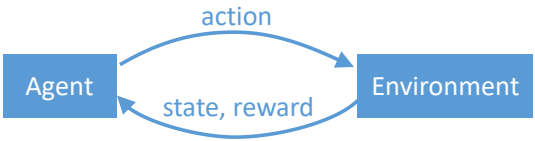
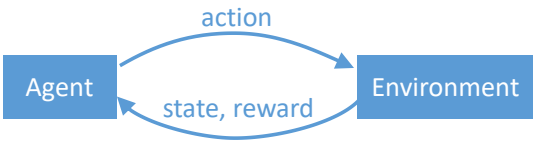
Motivation — Sequential Decision Making



Markov Decision Processes

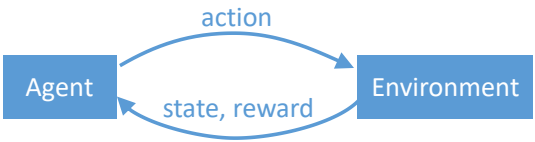
Motivation — Sequential Decision Making



Markov Decision Processes

1. Observe state $s \in S$;

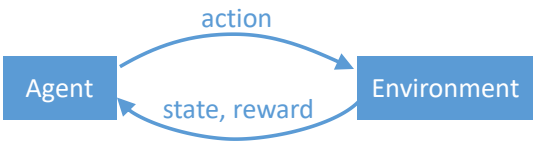
Motivation — Sequential Decision Making



Markov Decision Processes

1. Observe state $s \in S$;
2. Pick a discrete action $a \in A$;

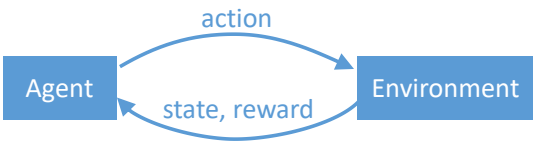
Motivation — Sequential Decision Making



Markov Decision Processes

1. Observe state $s \in S$;
2. Pick a discrete action $a \in A$;
3. Transition to a next state $s' \sim \mathbb{P}\left(s'|s,a\right)$;

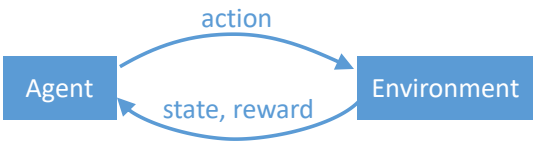
Motivation — Sequential Decision Making



Markov Decision Processes

1. Observe state $s \in S$;
2. Pick a discrete action $a \in A$;
3. Transition to a next state $s' \sim \mathbb{P}(s'|s,a)$;
4. Receive a bounded reward $r \in [0,1]$ drawn from $\mathbb{P}(r|s,a)$.

Motivation — Sequential Decision Making



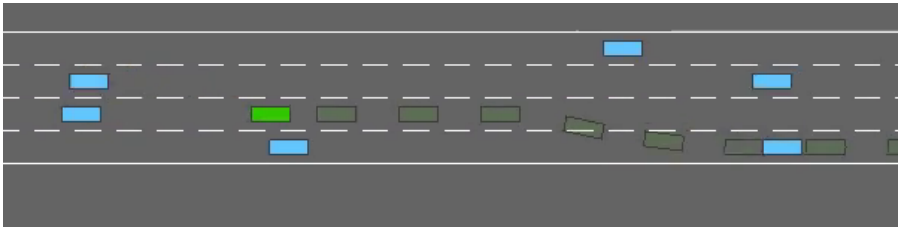
Markov Decision Processes

1. Observe state $s \in S$;
2. Pick a discrete action $a \in A$;
3. Transition to a next state $s' \sim \mathbb{P}(s'|s,a)$;
4. Receive a bounded reward $r \in [0,1]$ drawn from $\mathbb{P}(r|s,a)$.

Objective: maximise $\mathrm{V} = \mathbb{E}\left[\sum_{t=0}^{\infty} \gamma^t r_t\right]$

Motivation — Example

The `highway-env` environment

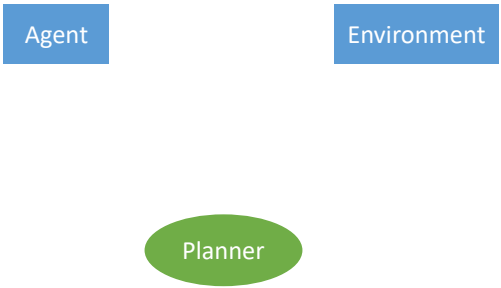


We want to handle stochasticity.

Motivation — How to solve MDPs?

Online *Planning*

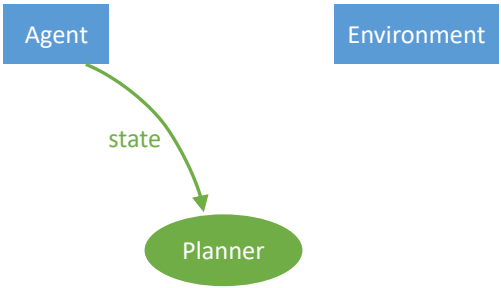
- we have access to a generative model:
  - ↳ yields samples of $s', r \sim \mathbb{P}(s', r|s, a)$ when queried



Motivation — How to solve MDPs?

Online *Planning*

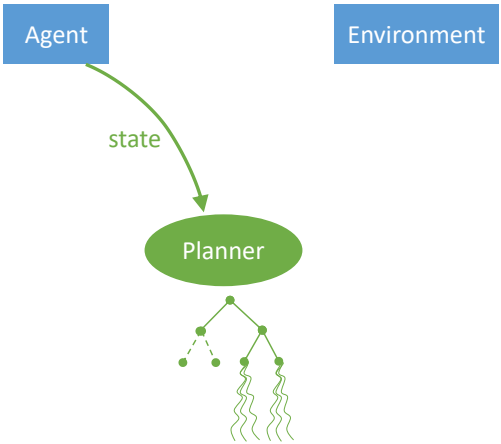
- we have access to a generative model:
  - ↳ yields samples of $s', r \sim \mathbb{P}(s', r|s, a)$ when queried



Motivation — How to solve MDPs?

Online *Planning*

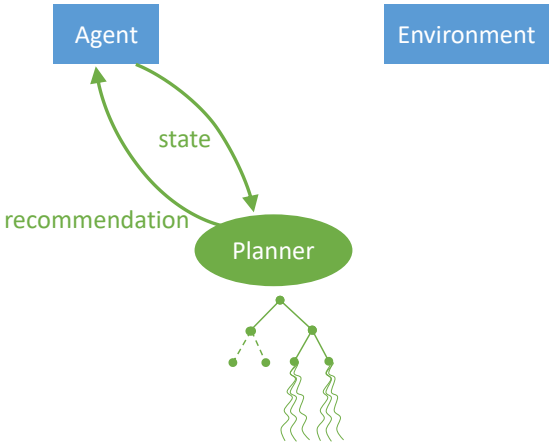
- we have access to a generative model:
  - ↳ yields samples of $s', r \sim \mathbb{P}(s', r|s, a)$ when queried



Motivation — How to solve MDPs?

Online *Planning*

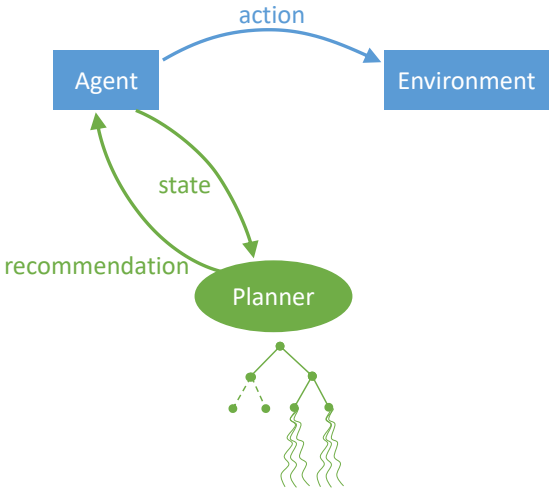
- we have access to a generative model:
  - ↳ yields samples of $s', r \sim \mathbb{P}(s', r|s, a)$ when queried



Motivation — How to solve MDPs?

Online *Planning*

- we have access to a generative model:
  - ↳ yields samples of $s', r \sim \mathbb{P}(s', r|s, a)$ when queried



Motivation — How to solve MDPs?

Online *Planning*

- we have access to a generative model:
  - ↳ yields samples of $s', r \sim \mathbb{P}(s', r|s, a)$ when queried