# RL: Basics

## The Markov Assumption

Marius Lindauer



AutoML.org
Automated
Machine Learning
Hannover

# Markov Assumption

- Information state: sufficient statistic of history
- State $s_t$ is Markov if and only if:

$$p(s_{t+1} \mid s_t, a_t) = p(s_{t+1} \mid h_t, a_t)$$

- with history $h_t = (a_1, s_1, r_1, \ldots, a_t, s_t, r_t)$

# Markov Assumption

- Information state: sufficient statistic of history
- State $s_t$ is Markov if and only if:

$$p(s_{t+1} \mid s_t, a_t) = p(s_{t+1} \mid h_t, a_t)$$

- with history $h_t = (a_1, s_1, r_1, \ldots, a_t, s_t, r_t)$
- Question: Hypertension control: let state be current blood pressure, and action be whether to take medication or not. Is this system Markov?
- Question: Website shopping: state is current product viewed by customer, and action is what other product to recommend. Is this system Markov?

# Why is Markov Assumption Popular?

- Can always be satisfied
  - Setting state as history always Markov: $s_t = h_t$
- In practice often assume most recent observation is sufficient statistic of history: $s_t = o_t$
- State representation has big implications for:
  - Computational complexity
  - Data required
  - Resulting performance

# Full Observable vs Partially Observable

- If we assume $s_t = o_t$, we have fully observable process

# Full Observable vs Partially Observable

- If we assume $s_t = o_t$, we have fully observable process
- If some information of $s_t$ is hidden s.t. the hidden is important for the Markov property, we have a partially observable process
  - Card game where you cannot see the cards of the other players

# Full Observable vs Partially Observable

- If we assume $s_t = o_t$, we have fully observable process
- If some information of $s_t$ is hidden s.t. the hidden is important for the Markov property, we have a partially observable process
  - Card game where you cannot see the cards of the other players
- In theory, we can transform any partially observable process into a fully observable by modelling non-deterministic state transitions and any possible outcome of the hidden information
  - In partice, often infeasible
  - special subfield of RL that deals with such POMDP problems

# Full Observable vs Partially Observable

- If we assume $s_t = o_t$, we have fully observable process
- If some information of $s_t$ is hidden s.t. the hidden is important for the Markov property, we have a partially observable process
  - Card game where you cannot see the cards of the other players
- In theory, we can transform any partially observable process into a fully observable by modelling non-deterministic state transitions and any possible outcome of the hidden information
  - In partice, often infeasible
  - special subfield of RL that deals with such POMDP problems
- We will see later in the course that even if the Markov asumption is violated and some information is hidden, we can nevertheless train well-performing agents