RL: Basics

The Markov Process

Marius Lindauer

AutoML.org
Automated
Machine Learning
Hannover

Reminder: 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 Process or Markov Chain

- Memoryless random process (/walk)
  - $\rightsquigarrow$ Sequence of random states with Markov propery
- Definition of Markov Process $M = (S, P)$
  - S is a (finite) set of states ($s \in S$)
  - P is dynamics/transition model that specifies $p(s_{t+1} = s' \mid s_t = s)$
- Note: no rewards, no actions
- If finite number (N) of states, can express P as a matrix

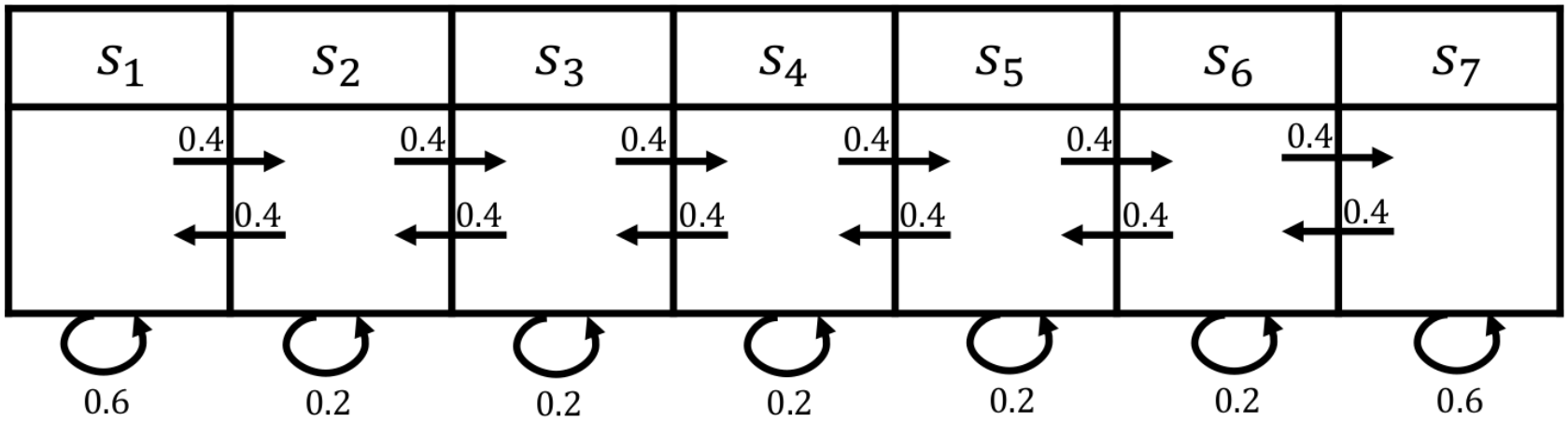
$$P_{i,j} = P(s_i \mid s_j)$$

Mars Rover

s_1	s_2	s_3	s_4	s_5	s_6	s_7
						

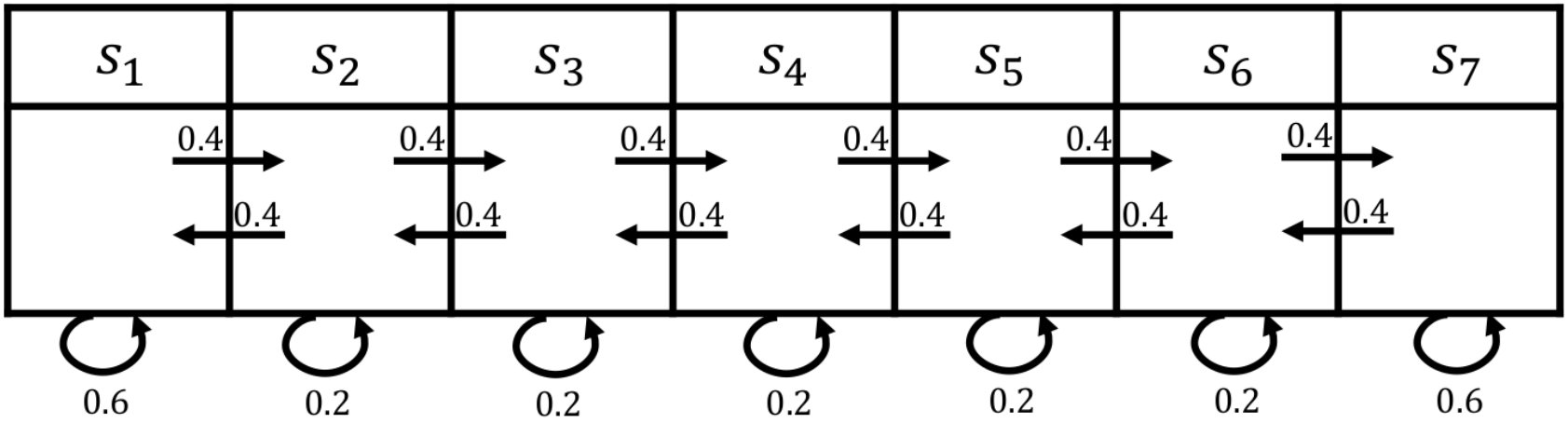
- States: Location of rover $(s_1, \ldots, s_7)$

Mars Rover as Markov Process



$$P = \begin{pmatrix} 0.6 & 0.4 & 0 & 0 & 0 & 0 & 0 \\ 0.4 & 0.2 & 0.4 & 0 & 0 & 0 & 0 \\ 0 & 0.4 & 0.2 & 0.4 & 0 & 0 & 0 \\ 0 & 0 & 0.4 & 0.2 & 0.4 & 0 & 0 \\ 0 & 0 & 0 & 0.4 & 0.2 & 0.4 & 0 \\ 0 & 0 & 0 & 0 & 0.4 & 0.2 & 0.4 \\ 0 & 0 & 0 & 0 & 0 & 0.4 & 0.6 \end{pmatrix}$$

Mars Rover as Markov Process (cont'd)



Exemplary episodes:

- $s_4, s_5, s_6, s_7, s_7, s_7, \ldots$
- $s_4, s_4, s_5, s_4, s_5, s_6, \ldots$
- $s_4, s_3, s_2, s_1, \ldots$

Applications

- Google's page rank with originally based on a random walk
  - Follow links on homepages
  - Rank websites based on probability to discover this page

Applications

- Google's page rank with originally based on a random walk
  - Follow links on homepages
  - Rank websites based on probability to discover this page
- Modeling of navigation behavior on websites

Applications

- Google's page rank with originally based on a random walk
  - Follow links on homepages
  - Rank websites based on probability to discover this page
- Modeling of navigation behavior on websites
- Early versions of text completion based on Markov processes

Applications

- Google's page rank with originally based on a random walk
  - Follow links on homepages
  - Rank websites based on probability to discover this page
- Modeling of navigation behavior on websites
- Early versions of text completion based on Markov processes
- Used in forecasting of trends, e.g., prices and wind power