Reinforcement Learning

Uczenie przez ocenę zachowań w środowisku

Piotr Bielak (218 137)

Wydział Informatyki i Zarządzania
Politechnika Wrocławska
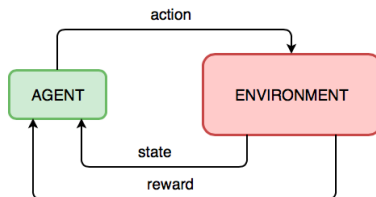
26 lutego 2018r.

Politechnika
Wrocławska

Reinforcement Learning

- definition
- basic concepts:
  - agent,
  - environment,
  - states,
  - actions,
  - rewards,
  - etc.
- machine learning branches:
  - RL vs supervised learning,
  - RL vs unsupervised

Theory

Theoretical basics:

- MDP (Markov Decision Process),
- exploitation vs exploration,
- discounted rewards,
- Bellman equation,
- essential approaches:
  - dynamic programming,
  - Monte Carlo methods,
  - temporal difference (TD) learning,
- value functions,
- policy functions,
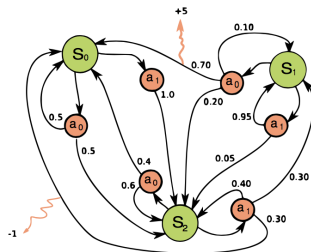
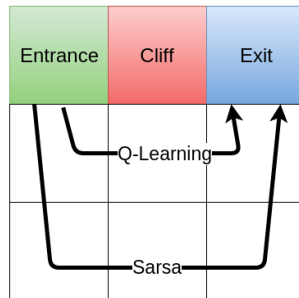*https://en.wikipedia.org/wiki/Markov_decision_process*

Algorithms

Most popular algorithms:

- Q-learning,
- SARSA,
- on-policy and off-policy methods,
- actor-critic methods,
- Deep Reinforcement Learning:
  - DQN (Deep Q-Networks),
  - Deep SARSA,
  - tabular approaches vs function approximators,

https://i.stack.imgur.com/tj9DV.png

Materials

Books:

- Richard S. Sutton and Andrew G. Barto, *Reinforcement Learning: An Introduction*, 2017
- Csaba Szepesvari, *Algorithms for Reinforcement Learning*, 2009

Papers:

- L. P. Kaelbling, M. L. Littman and A. W. Moore, *Reinforcement Learning: A Survey*, 1995
- S. S. Kerrthi and B. Ravindran, *A Tutorial Survey of Reinforcement Learning*, 1994
- Volodymyr Mnih et al., *Human-level control through deep reinforcement learning*, 2015
- H. van Hasselt, A. Guez and D. Silver, *Deep Reinforcement Learning with Double Q-learning*, 2015
- E. Morales and J. Zaragoza, *An Introduction to Reinforcement Learning*, 2011

Materials

University courses:

- David Silver, *Reinforcement Learning*, University College London
- Stefan Riezler, *Reinforcement Learning*, Universitat Heidelberg

Presentations:

- Szymon Zaręba, *Reinforcement Learning*, Politechnika Wrocławska, 2017
- Witoldi Paluszyński, *Reinforcement Learning*, Politechnika Wrocławska

Reinforcement Learning

Uczenie przez ocenę zachowań w środowisku

Piotr Bielak (218 137)

Wydział Informatyki i Zarządzania
Politechnika Wrocławska

26 lutego 2018r.