Recommender Systems and Reinforcement Learning

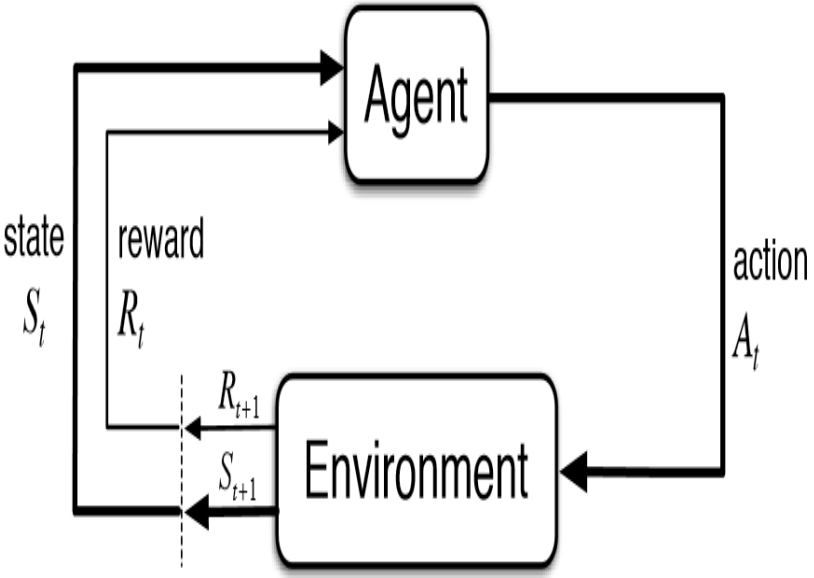
Vikas K. Sinha

Recommender Systems using Reinforcement Learning

- What is Reinforcement Learning
- Applicability of RL for Recommender Systems
- An example of Recommender System using RL
- References

What is Reinforcement Learning

- Algorithms to determine actions in an environment so as to maximize cumulative reward
- Environment is usually modeled as a Markov Decision Process
  - Next State depends only on Current State
- Goal-Oriented Algorithms for learning a complex goal over many steps
- Applications in game theory, operations research, statistics, information theory, …
- Successfully applied in robot control, telecom, backgammon, checkers, Go, ...



Applicability: When to consider RL for Recommender Systems

- Immediate feedback from user is insufficient for modeling the value of an item.
  - E.g. a recommendation may be well-received in the short term but lead to less user retention over long term.
- Long-term sequence of actions taken by user is more accurate indicator of rating.
  - Difference between immediate response and long-term utility.
- RL models long-term cumulative rewards.
  - Optimization over longer sequence of responses
  - “Richer” modeling of users

RL and Recommender Systems: An Example

- RL was used for Youtube video recommendation (see Reference for full presentation on REINFORCE recommender).
- Several challenges found in adapting RL paradigm to Youtube recommendation use case, e.g.
  - “Random choices” based on RL will not be well received
  - Exploration is expensive
  - Action space is large (number of videos)
- Paper describes results of algorithm on simulation and live environment
    - Amount of viewing time increased slightly and number of videos seen decreased slightly due to algorithm

References

1. https://en.wikipedia.org/wiki/Reinforcement_learning
2. https://skymind.ai/wiki/deep-reinforcement-learning
3. Reinforcement Learning for Recommender Systems: A Case Study on Youtube," by Minmin Chen. https://www.youtube.com/watch?v=HEqQ2_1XRTs
4. Top-K Off-Policy Correction for a REINFORCE Recommender System. https://arxiv.org/abs/1812.02353