Docket #: S15-261

Reinforcement learning with bootstrapped value function randomization

Stanford researchers have developed a new algorithm for reinforcement learning, which can learn to take good actions with potentially long term consequences in a general unknown complex system. Unlike previous approaches, this method is able to combine complex nonlinear machine learning techniques (such as deep neural networks) with efficient experimentation in a computationally tractable manner. This method balances exploration with exploitation in way that can lead to exponential improvements over current state of the art. The system incorporates a machine learning model that receives observations of the environment and takes actions to optimize cumulative rewards. The system balances the needs for exploration with exploitation through the use of randomized value function estimates to incentivize policies that are poorly understood. This provides a balanced algorithm for automating decision and learning systems that can act and learn efficiently in terms of data, computation, and observed performance.

Visualizing Uncertainty:

Applications

- **Internet systems:** to optimize long term customer interactions with a website service
- **Advertising:** to enhance ad serving over customer lifetime
- **Healthcare:** to direct experimental and personalized medical trials
- **Agriculture:** to dynamically manage crop treatment
- **Robotics:** to learn new behaviors not specifically pre-programmed

Advantages

- **Generic and Adaptable:**
  - Non-parametric algorithm design can be applied in conjunction with any new machine learning algorithm.
  - Allows for efficient experimentation scheme to be combined with complex models
- **Demonstrably Efficient:**
  - Benchmark improvements over currently-used competitors
  - Scalable, computationally efficient implementation
  - Can provide exponential improvements over current state of the art

Publications

- Ian Osband, Charles Blundell, Alexander Pritzel, Benjamin Van Roy, "Deep Exploration via Bootstrapped DQN," *Advances in Neural Information Processing Systems* 29 (NIPS 2016).
- Published Patent Application: US20170032245A1

Patents

- Published Application: WO2017004626
- Published Application: 20170032245
- Published Application: 20200065672
- Issued: 12,536,444 (USA)

Innovators

- Ian Osband
- Benjamin Van Roy

Licensing Contact

Imelda Oropeza

Senior Licensing Manager, Physical Sciences

Email