Army research leads to more effective training model for robots

December 29 2020

Credit: The Army Research Laboratory

Multi-domain operations, the Army's future operating concept, requires autonomous agents with learning components to operate alongside the warfighter. New Army research reduces the unpredictability of current training reinforcement learning policies so that they are more practically

applicable to physical systems, especially ground robots.

These learning components will permit [autonomous agents](#) to reason and adapt to changing battlefield conditions, said Army researcher Dr. Alec Koppel from the U.S. Army Combat Capabilities Development Command, now known as DEVCOM, Army Research Laboratory.

The underlying adaptation and re-planning mechanism consists of [reinforcement](#) learning-based policies. Making these policies efficiently obtainable is critical to making the MDO operating concept a reality, he said.

According to Koppel, policy gradient methods in reinforcement learning are the foundation for scalable algorithms for continuous spaces, but existing techniques cannot incorporate broader decision-making goals such as risk sensitivity, safety constraints, exploration and divergence to a prior.

Designing autonomous behaviors when the relationship between dynamics and goals are complex may be addressed with reinforcement learning, which has gained attention recently for solving previously intractable tasks such as strategy games like go, chess and videogames such as Atari and Starcraft II, Koppel said.

Prevailing practice, unfortunately, demands astronomical sample complexity, such as thousands of years of simulated gameplay, he said. This sample complexity renders many common training mechanisms inapplicable to data-starved settings required by MDO context for the Next-Generation Combat Vehicle, or NGCV.

"To facilitate reinforcement learning for MDO and NGCV, training mechanisms must improve sample efficiency and reliability in continuous spaces," Koppel said. "Through the generalization of existing

policy search schemes to general utilities, we take a step towards breaking existing sample efficiency barriers of prevailing practice in reinforcement learning."

Koppel and his research team developed new policy search schemes for general utilities, whose sample complexity is also established. They observed that the resulting policy search schemes reduce the volatility of reward accumulation, yield efficient exploration of an unknown domains and a mechanism for incorporating prior experience.

"This research contributes an augmentation of the classical Policy Gradient Theorem in reinforcement learning," Koppel said. "It presents new policy search schemes for general utilities, whose sample complexity is also established. These innovations are impactful to the U.S. Army through their enabling of reinforcement learning objectives beyond the standard cumulative return, such as risk sensitivity, safety constraints, exploration and divergence to a prior."

Notably, in the context of ground robots, he said, data is costly to acquire.

"Reducing the volatility of reward accumulation, ensuring one explores an unknown domain in an efficient manner, or incorporating prior experience, all contribute towards breaking existing sample efficiency barriers of prevailing practice in reinforcement learning by alleviating the amount of random sampling one requires in order to complete policy optimization," Koppel said.

The future of this research is very bright, and Koppel has dedicated his efforts towards making his findings applicable for innovative technology for Soldiers on the battlefield.

"I am optimistic that reinforcement-learning equipped autonomous

robots will be able to assist the warfighter in exploration, reconnaissance and risk assessment on the future battlefield," Koppel said. "That this vision is made a reality is essential to what motivates which research problems I dedicate my efforts."

The next step for this research is to incorporate the broader decision-making goals enabled by general utilities in reinforcement learning into multi-agent settings and investigate how interactive settings between reinforcement learning agents give rise to synergistic and antagonistic reasoning among teams.

According to Koppel, the technology that results from this research will be capable of reasoning under uncertainty in team scenarios.

Provided by The Army Research Laboratory

Citation: Army research leads to more effective training model for robots (2020, December 29) retrieved 24 April 2024 from https://techxplore.com/news/2020-12-army-effective-robots.html