Assured Deep Reinforcement Learning for Safety-Critical Applications

Radu Calinescu and Daniel Kudenko
Department of Computer Science
University of York

This PhD project will develop a theoretical foundation and practical techniques enabling the adoption of deep reinforcement learning in the safety-critical domain.

Recently emerged through the integration of reinforcement learning and deep learning, *deep reinforcement learning* (DRL) is a powerful approach to developing software agents capable of human-level adaptive behaviour in scenarios characterised by incomplete knowledge and non-deterministic change.[1] Successful applications within non-critical domains such as computer games and board games bear witness to the effectiveness of the approach. Nevertheless, this success has so far eluded the important class of safety-critical applications, for which DRL is currently unable to guarantee compliance with safety requirements. The PhD project aims to address this limitation, and thus to enable the exploitation of deep reinforcement learning in the safety-critical domain.

To achieve its objective, the project will extend our recent advances in assured reinforcement learning,[2,3,4] and will develop and validate a three-pronged approach to providing formal assurance that DRL autonomous agents comply with their safety requirements:

- Runtime probabilistic model checking will be used to constrain the state space (i.e., options) explored by deep reinforcement learning to areas that are provably safe. This functionality will be supported through the continual analysis of mathematical models including Markov decision processes and continuous-time Markov chains augmented with costs/rewards. Domain-specific languages devised by the project will enable domain experts to provide initial versions of the models.
- Whenever permitted by the available compute and memory resources, runtime verification middleware will be embedded into the autonomous DRL agents, ensuring that these agents use similar mathematical models to verify their adaptation actions dynamically.
- Preliminary automated model transformations and decision-log analysis techniques will be used to generate humanly understandable justifications for the actions of the autonomous DRL agents.

---

[1] V. Mnih et al. (2015). Human-level control through deep reinforcement learning. *Nature* **518**:529-533.
[2] G. Mason, R. Calinescu, D. Kudenko, A. Banks (2017). Assured Reinforcement Learning with Formally Verified Abstract Policies. *9th International Conference on Agents and Artificial Intelligence*. (Best student paper award)
[3] G. Mason, R. Calinescu, D. Kudenko, A. Banks (2017). Assurance in Reinforcement Learning Using Quantitative Verification. In: I. Hatzilygeroudis, V. Palade (editors), Advances in Hybridization of Intelligent Methods, pp 71-96, Springer.
[4] G. Mason, R. Calinescu, D. Kudenko, A. Banks (2017). Assured Reinforcement Learning for Safety-Critical Applications. *Doctoral Consortium at 9th International Conference on Agents and Artificial Intelligence*. (Best PhD project award)