PhD Topics --- Daniel Kudenko

Background

RL is a machine learning methodology for decision making of an intelligent agent embedded in a physical or virtual environment (e.g. autonomous robots or trading agents). RL techniques are based on the agent receiving a numerical reward (or punishment) at each time step, and compute the action that will maximize the expected sum of future rewards based on past experience. While RL has been at the core of many recent AI successes (e.g. Google DeepMind), there are still obstacles applying them to real-world tasks: (1) RL methods do not scale well to the complexity of real-world tasks, and (2) there are concerns about safety guarantees, especially in safety-critical applications. These difficulties are even bigger for multi-agent reinforcement learning, where a (potentially large) number of agents are sharing an environment and need to learn to compete and cooperate. PhD projects 1-3 will contribute towards making RL and MARL methods applicable in complex real-world tasks.

1. Multi-Agent Knowledge Based Reinforcement Learning

In past work [1,2], we have shown how high-level heuristic knowledge can be used to guide and speed up the reinforcement learning process for a single agent. This research project will take the work one step further and investigate the use of multi-agent coordination knowledge to improve MARL. The work will involve the following aspects (step 3 is optional but nevertheless interesting):

1. Finding a suitable representation of multi-agent knowledge and developing algorithms to transform this into a reward function that encourages the exploration of behaviours the knowledge suggests.
2. Adding a human to the loop, to observe the agents' behaviour and modify the knowledge base (KB) to discourage undesirable behaviour emerging or exploit new knowledge learnt from observation.
3. Enabling agents to identify when the guiding knowledge and their experiences differ (suggesting invalid or inaccurate knowledge in the KB), formalise such observations as revisions to the KB and either apply automatically or suggest to the “human-in-the-loop”.
4. Evaluation of the techniques on simulated multi-robot tasks. This will show how incorporating domain knowledge into MARL can speed up learning and increase the scale of tasks learnable.

[1] M. Grzes, D. Kudenko (2008): "Plan-based reward shaping for reinforcement learning", *Fourth IEEE International Conference on Intelligent Systems (IS)*.
[2] K. Efthymiadis, D. Kudenko (2015): "Knowledge Revision for Reinforcement Learning with Abstract MDPs", *Fourteenth International Conference on Autonomous Agents and Multi-Agent Systems (AAMAS-15)*

2. Curriculum Learning for Multi-Agent Reinforcement Learning

It has been shown [3] that the presentation of subtasks and their sequence (the so-called curriculum) can improve reinforcement learning for individual agents. How such subtasks and subtask sequences can be computed automatically is still subject to ongoing research. This project will study curriculum learning in the context of multi-agent RL. Specifically, the research will:

1. Explore what subtasks are useful in a multi-agent context (e.g. should they involve only some of the agents or all, what kind of interactions they should focus on, etc.)
2. Test the impact of different subtask sequences on the learning performance.
3. Explore methods for the automated generation of subtasks and their sequences.
4. Evaluate the techniques on simulated multi-robot tasks.

[3] Sanmit Narvekar, Jivko Sinapov, Matteo Leonetti, and Peter Stone (2016): "Source Task Creation for Curriculum Learning". *Fifteenth International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*.

3. Assured Multi-Agent Reinforcement Learning

In past work [4], we have developed methods to combine reinforcement learning with quantitative verification algorithms that enable a designer to specify constraints on the learned agent behaviour and provide guarantees that these constraints will be satisfied during and after learning. This research project will take the results one step further and develop a method that allows a designer to specify safety constraints based on multi-agent interactions and ensure the satisfaction of these constraints during learning and in the final agent behaviour. Specifically, the research will:
1. Identify/design a representation formalism for safety constraints.
2. Extend and improve the techniques presented in [4] for multi-agent reinforcement learning.
3. Evaluate the techniques on various multi-agent tasks.
Note: the past work [4] was a DSTL funded PhD project. The student, George Mason, did a placement in connection with this work at DSTL, and DSTL showed considerable interest to continue work in this area.

[4] G. Mason, R. Calinescu, D. Kudenko, A. Banks (2017): "Assured Reinforcement Learning with Formally Verified Abstract Polocies", *Ninth International Conference on Agents and Artificial Intelligence (ICAART).* [Best Student Paper Award].