Reinforcement learning - Contemporary Challenges

Machine learning divides traditionally into three major fields supervised learning (SL), unsupervised learning (UL) and reinforcement learning (RL). Yann LeCun, one fathers of the current AI/ML revolution, famously called RL to be "cherry on the cake". While LeCun claims that we know "how to do the cheery" in fact a lot of fundamental ideas are still to be understood and found. For example, recent impressive breakthroughs in games (DQN, GO, Dota2) still operate in rather sterile computer environments and require enormous computational power. This proposal aims to investigate some of these bottlenecks. Our intention is to establish a strong group on the University of Warsaw performing RL research.

Research project objectives/Research hypothesis. We plan to research the following topics: **hierarchical learning, curriculum learning, meta-learning /memory, model-free and model-based methods.** All these topics are core areas of contemporary research in reinforcement learning.

Hierarchical learning is still a fairly unexplored area. In a recent paper we present a novel architecture for parametrized hierarchical learning, using this as a basis we aim to investigate two important issues: *compatibility of macro-actions* and *macro-actions discovery*. The first task aims in developing a robust training procedure in applicable for broad number of problems. In the second task we propose an algorithm that automize the discovery of macro-actions.

Curriculum learning We propose a novel method, which we call *on-hype curriculum generation* method. Assume that that there are a spectrum of goals of a similar nature but different levels of difficulty. At a given stage of training one goal is chosen to be the principal one on which the focus is put. Our method chooses the major goal which is hopefully the most beneficial for the speed of the learning process. Training other related goals enhances the process by creating useful synergies. Our research objective is to implement and evaluate this method.

Meta-learning /memory Our research objective is to evaluate usefulness of various architectures used in machine learning in meta-learning. We will focus on estimating capacity to memorize details of experiment and store execution plans for fast RL algorithms.

Model-free and model-based methods We would like to construct an algorithm bridging between model-free and model based methods. The task divides into three stages. The first one is devoted to automatic discovery of small dimensional, locally valid models of RL environments. The second goal is to develop a version of the MCTS algorithm which is well-suited for such models. We suspect that avoiding visiting the same states will be the crucial ingredient, Investigating such questions itself is an interesting research topic. The final goal is to connect this two techniques into one end-to-end method.

Research project methodology. The usual methodology in the contemporary AI research is performing experiments. In our project we propose a few novel methods along with experiments to evaluate them. In order to successfully execute them we plan to buy computers with GPUs and set up a research team.

Expected impact of the research project on the development of science. Reinforcement learning is a cutting-edge area of the now-a-days artificial intelligence research. We believe that the topics proposed above are pioneering and can impact the discipline.

Hierarchical learning The hierarchical approach to machine learning is conjectured to be a crucial ingredient of truly intelligent systems. The task *compatibility of macro-actions* may solve an important technical issue leading to easier application of hierarchal methods. The task *Macro-actions discovery* touches the very core is devoted to automatization of the discovery of the hierarchical structure. This field is basically unexplored and virtually any contribution to the field may have significant impact.

Curriculum learning Devising a general method of curriculum generation would be a major step in development of RL techniques. Our algorithm may contribute to this field.

Meta-leaning/memory Meta-learning is based on the idea that a class of problems have some common rules. Automatic discovery of such rules and being able to use them may greatly improve the sample efficiency of RL algorithms when applied to "common" tasks. As such it might have a significant impact on application of RL techniques in real world problems which often have some underlying a-priori structure.

Model-free and model-based methods Bridging model-free and model-based method might have a major impact on reinforcement learning. Important partial result in the field in the case of robotic learning (the so-called GPS algorithm). We would to repeat this success in case of discrete domains It might significantly reduce the sample complexity and thus training time making the RL techniques possible to use for broader spectrum of applications. Achieving any of sub-goals listed above will itself be a valuable discovery in the field.