PhD proposal:

Towards bridging the gap between induction and deduction: the case of Reinforcement Learning*

Supervisor: Philippe Preux
Research group: Scool
Lab: CRIStAL & Inria
Location: Lille, France
philippe.preux@univ-lille.fr
https://philippe-preux.github.io

Application: now
Begins: October 1st 2021

Abstract

Machine learning relies on the notion of generalization from examples, that is induction. Mostly based on statistics, machine learning completely lacks any means of reasoning by deduction. Lying in two very different areas of artificial intelligence, this PhD ambitions to investigate how these two may principles be bridged.

We will consider this question in the field of reinforcement learning.

Keywords: reinforcement learning, induction learning, deduction, symbolic reasoning, artificial intelligence.

*The most up-to-date version of this proposal is available at https://philippe-preux.github.io/proposals/Bridging-ML-with-symbolic-reasoning.PhD-proposal.pdf.

1 Research project

1.1 Context

In the last 10 years at least, reinforcement learning (RL) has been a very active field of research. Aiming at designing agents able to learn to solve tasks by themselves, this is a fascinating scientific topic, originating from the search for understanding how animals adapt their behavior to their environement. RL also conveys immense hopes from companies to create effective agents able to optimize the resolution of a task, and able to adapt to an environment changing along time. Leveraging the deep learning technology, agents have been able to learn by themselves to play board games at an unprecented level of expertise, well beyond human experts: go successes have been celebrated and the algorithms underlying this success are now able to learn to play chess at a much higher level than Deep Blue chess player celebrated in its time to have defeated world chess champion G. Kasparov. Progress on go and the go "resolution" have been widely advertized[1]. Though this is a great achivement, it should be said that go is not such a formidable task to solve with regards to the tasks we would like to solve, in particular in real, industrial applications: go is a deterministic task, the state of the task is perfectly well defined at any moment, the number of plays is also quite limited to a few hundreds, and the rules are perfectly well known and fixed along time. It should also be reminded that this go player was obtained using massive computing power. This latter point means that to obtain an agent able to solve more complex tasks, even more computing power will be required, a computing power that is out of reach of all actors except but a very few. Moreover, such an unrestricted use of computing power is not compatible with the current climate change crisis that urges us to bifurcate towards energy consumption reduction and sustainability in all sectors of human activities.

At the same time, when one uses reinforcement learning to solve a task, one always feel like a lot of knowledge could be used to make the learning more efficient and effective. Most of the time, the only possibility to help the learning agent is to carefully design the problem representation; a lot of work has been done and is still done to try to help the agent while it is learning by "showing" it how it should behave, or guide it towards a good behavior, but this approach is not really working today and it is limited to some types of problems[2].

1.2 Objectives of the PhD

Building on the previous observations, one feels that the agent may benefit from knowledge about the task it is learning to solve, this knowledge being expressed with symbolic rules. Inspecting how an RL agent learns to solve a task, most of the time, we see that it takes a very long time to figure out what we know about the task. However, we do not have any means to provide this information.

This PhD aims at exploring how this paradoxical situation could be mended. Ultimately, but this is probably much too ambitious so please read this has a naive illustration of our goal, we would like to combine a reinforcement learner with a logic reasoning engine, say a hybrid between an actor-critic and a Prolog engine.

More concretely and witin a shorter-term, in the path towards this objective, we will at least examine two different points. The first point is how an RLearner may take advantage of rules to

[1] A former SequeL member made very important contributions on this topic around 2006.
[2] SequeL and Scool members have been and are still working on this topic.

improve its learning. The rules will not be learned, but given to the agent. The means (language) to express these rules is to be studied, as well as their use by the agent. Rules may translate into hard constraints ruling the behavior of the agent, they may tweek the return function or the function being optimized, or may be used in any other way to be studied. We will consider reliable rules and progressively relax the reliability expectations, that is rules that express constraints that are sometimes wrong, or even alwaysa approximate. The second point is how an RLearner may take advantage of basic symbolic reasoning capabilities, such as an order 0 logic-based reasoning engine.

Currently, this idea has been a bit explored mostly through the idea of sub-symbolism and neural-symbolic learning (see *e.g.* Besold and al. 2017; Dong et al. 2019; Manhaeve et al. 2018; Zimmer and al. 2021).

The impact of this research may be very important. Once we would be able to combine symbolic rules, symbolic deduction, and statistical based reinforcement learning, this would be a key tool that would be used in many if not all applications of reinforcement learning.

The outcome of this PhD is expected to be mainly publications in top journals and conferences in artificial intelligence and machine learning. Software prototypes may also be developed. In this case, they will be freely available for the scientific community.

2 Advisor and research environment

Philippe Preux has been doing research in various areas of artificial intelligence for almost 30 years. First working on bio-inspired algorithms (originally genetic algorithms), he moved to machine learning around 1999, and particularly to reinforcement learning.

He created SequeL at Inria back in 2006 which turned into Scool in 2020. Scool is dedicated to the study of the sequential decision problem under uncertainty, in particular bandit and reinforcement learning problems. Scool enjoys an international recognition on these topics. Scool is currently made of 6 permanent researchers, more than 20 PhD students, and a couple of post-docs (at the end of 2021, we should have between 5 and 10 post-docs in Scool), and 2 engineers (2 more to hire in 2021).

Ph. preux is currently supervising 9 PhD students, most of them being co-supervised. 2 of them will defend their PhD in 2021, many in 2022.

Scool is a very favorable environment to do research in reinforcement learning, attracting many collaborations worldwide. Our PhD students enjoy a quite unique opportunity to interact with each others and with other members of the group, helping each others, creating a very lively and stimulating environment for research. Scool members publish in the top ranked journals and conferences in machine learning and artificial intelligence (*e.g.* NIPS, ICML, IJCAI, ICLR).

Regarding this PhD proposal, it clearly builds on past work in Scool and it is clearly in line and complementary to many current activities of research Scool members. It will also take advantage of some of the collaborations and projects Scool is involved in. It fits very well with at least the following on-going projects:

- our participation to the Inria challenge HY_AIAI in which we study the combination of symbolic and numerical approaches of machine learning,
- our participation to the CausalXRL Chist-Era project (starting in 2021) where we will investigate causal reinforcement learning to explain what is learned,

- the application of reinforcement learning to agriculture we study with Cirad and CGIAR,

Other projects are under discussion that may provide relevant use cases for this PhD.

3 About the candidate

The candidate should hold a master preferably in computer science, with a very good knowledge of CS fundamentals (including but not only: algorithms, data structure, programming in either C, or C++, or Python, AI, combinatorial optimization) and either a very good background in machine learning, or in symbolic AI. A serious background in maths is a bonus.

Ability to work in English is necessary, as well as being able to clearly present one's work orally and in written accounts. It is important that the candidate is able to interact with the other members of the group, from a scientific point of view, but also from a social point of view.

We expect the candidate to be rather autonomous, seek information and solutions by him/her-self, and be pro-active.

4 Fit in the local research strategy

This proposal is perfectly in-line with the research activities in Région Hauts-de-France related to artificial intelligence, in particular in CRIStAL, Inria-Lille, and the HumAIn alliance.

This proposal is made as part of the AI chair Apprenf, co-funded by I-Site, MEL, Université de Lille, Inria, and ANR, chaired by Philippe Preux and Odalric-Ambrym Maillard (R-PILOTE-19-004-APPRENF).

References

[1] T.R. Besold and et al. *Neural-Symbolic Learning and Reasoning: A Survey and Interpretation.* arxiv 1711.03902. 2017.

[2] H. Dong et al. "Neural logic machines." In: *Proc. ICLR*. 2019.

[3] R. Manhaeve et al. "Neural probabilistic logic programming." In: *Proc. NIPS*. 2018.

[4] M. Zimmer and et al. *Differentiable Logic Machines.* arxiv 2102.11529. 2021.