Reinforcement Learning in Graph Theory

Master's Thesis Proposal

Joran Michiels

June 11, 2018

Machine learning seem to have only been used scarcely in theoretical mathematics. However, particular learning models can be used for certain mathematical problems. In a previous semester's thesis, we successfully applied reinforcement learning in online graph exploration. We achieved results close to previous theoretical results. In the master's thesis, we would like to improve the latter results and apply reinforcement learning to similar problems in graph theory.

The general idea of employing computers in pure mathematics is attributed to the field of *experimental mathematics* "in which computation is used to investigate mathematical structures and identify their fundamental properties and patterns" [1]. A well-known and pioneering work is the computer assisted proof of the four color theorem in 1976 (although it was not corrected until 1995) [2]. In the book *Mathematics by experiment: Plausible reasoning in the 21st century* [3] a list of directions are given. We are primarily interested in "gaining insight and intuition; discovering new patterns and relationships" and "testing ... conjectures" [3].

Automated theorem proving (ATP), arguably the supreme goal of experimental mathematics, has (successfully [4]) utilized machine learning. In Komendantskaya et al. [4] a selection of papers is provided. Although being very profitable in verifying software and integrated circuits [5], it is safe to say that most of the current mathematical theorems are not proven by ATP. Our interest is in these more complex theorems. To our knowledge, machine learning has not been used in this context.

More precisely we are interested in theorems concerning a supremum of a quantitative property p over a certain family of mathematical objects $\mathcal{O}$. If one finds an example object $O_1 \in \mathcal{O}$ with $p_1 = p(O_1)$, it is evident that p_1 is a valid lower bound on the supremum. This is particularly interesting if p_1 is higher than any previously achieved lower bound, and even more when p_1 is equal to a previously computed upper bound (on the supremum). In the later case the supremum would be equal to p_1. Such situations seems to occur frequently in graph theory, making it the study of our choice. A very recent example is the discovery that the chromatic number of the plane is at least 5 [6].

The research done in the semester's thesis [7] was towards a similar goal. Without going into details, we intelligently searched the space of graphs for one (such as O_1) with a high competitive ratio (the property p). With "intelligently" we mean "employing machine learning". Unfortunately we were not able to surpass theoretical bounds established in prior research, but it was considered close, i.e. we found a graph with a competitive ratio of ≈ 1.69 for the greedy exploration algorithm, whereas the corresponding theoretical bound was 1.75 [8].

A second important constraint on the scope of mathematical problems is that they should be transformable to a game with actions and rewards. This arises from the reinforcement learning approach where an agent is trained to execute the right actions by interacting with the environment via rewards [9].

In the master's thesis we will first try to improve the lower bound achieved in the semester's thesis by further exploration of the hyper parameter space and optimization of the training process. Secondly, another problem from graph theory should be transformed to a game, and with reinforcement learning a possibly better lower bound, as explained above, should be acquired. Lastly, if time allows, the trained agents in the aforementioned game could be investigated to gain insight in the corresponding theorems.

References

[1] Eric W. Weisstein. Experimental Mathematics. From MathWorld—A Wolfram Web Resource. Last visited on 13/6/2018.

[2] Jonathan Borwein and David H. Bailey. Will computers replace humans in mathematics? Last visited on 13/6/2018.

[3] Jonathan Borwein and David Bailey. *Mathematics by experiment: Plausible reasoning in the 21st century.* AK Peters/CRC Press, 2008.

[4] Ekaterina Komendantskaya, Jónathan Heras, and Gudmund Grov. Machine learning in proof general: Interfacing interfaces. *arXiv preprint arXiv:1212.3618*, 2012.

[5] James P Bridge. Machine learning and automated theorem proving. Technical report, University of Cambridge, Computer Laboratory, 2010.

[6] A. D. N. J. de Grey. The chromatic number of the plane is at least 5. *ArXiv e-prints*, April 2018.

[7] Joran Michiels. Online Graph Construction with Reinforcement Learning. Master's thesis, ETH Zürich, 2017.

[8] Cor A.J. Hurkens and Gerhard J. Woeginger. On the nearest neighbor rule for the traveling salesman problem. *Operations Research Letters*, 32(1):1 – 4, 2004.

[9] Tom M Mitchell. *Machine Learning*. McGraw-Hill, 1997.