Solving Arithmetic Problems on a Checkered Paper*

Adrián Csiszárik[1], Beatrix Benkő[2], Gergely Stomfai[1], and Milán Vásárhelyi[1]

[1] Alfréd Rényi Institute of Mathematics
[2] Eötvös Loránd University

1 Introduction

> In [3] Clark and Chalmers highlight a conversation taken verbatim from Gleik's (1992) biography of the famous scientist, Richard Feynman and the historian, Charles Weiner. In this conversation, Weiner mentions that Feynman's notes are "a record of the day-to-day work". Feynman retorts: "I actually did the work on paper". Weiner then suggests: "The work was done in your head, but the record of it is still here" to which Feynman counters: "No, it's not a record, not really. It's working. You have to work on paper and this is the paper. Okay?".

Our ongoing project explores how simple arithmetic reasoning tasks can be carried out by an artificial neural network that operates on a checkered paper, without utilizing any external prover. Our general setup can be formulated with the notions of reinforcement learning as follows.

- The environment is a checkered paper represented by an $N \times M$ grid, with at most one symbol in each cell from a fixed finite symbol set S.
- The starting state is a paper with an arithmetic problem written on it with the symbol set S. (E.g., with `"-6.1213 + 2543.073?"` or `"44342.23412 * -534.24?"` written on the first line of the paper.)
- The action the agent can take in each step is to write a symbol in a cell.
- The agent receives a reward when the correct solution is written on the paper (again with the symbol set S) at a prescribed position, followed by a special symbol ■ to "hand in" the paper for evaluation.

Why simple arithmetic problems? Harnessing artificial neural networks to solve arithmetic problems is a long pursued goal for the field at large [2, 4, 5, 6, 7]. The motivation behind this ambition is the idea that solving tasks of elementary school difficulty with artificial neural networks could bring us closer to understanding how one can approach the general goal of human-level intelligent behavior with machines. The subject of elementary school mathematics serves as a particularly suitable testbed in this regard: the necessary theory required to resolve such problems is narrow, one has exact solutions with solid reasoning steps, and perhaps most importantly, one can easily generate synthetic datasets to train the models. Despite the ubiquitousness of such tasks, handling arithmetic problems utilizing neural networks still seem to be challenging. (See, e.g., [7], where utilizing LSTMs or even a powerful Transformer model still fails to solve several seemingly easy tasks of e.g., multiplication. State-of-the-art learning systems utilizing theorem provers struggle with these tasks, [8] provides several experimental results.)

*With an artificial neural network, but without an external prover.

Why on a checkered paper? The starting point of our project is considering the semiotic perspective implicitly appearing in our motto behind the words of Feynman. Indeed, utilizing a (checkered) paper for solving arithmetic problems is a rather natural approach for humans. A human being mechanically executes straightforward steps of algorithms engrained in his mind, which have been learned and memorized over several years in elementary school. In the process, the arrangement of symbols in a spatial structure plays a key role. Our working hypothesis for this project is that more complex mathematical reasoning also relies on the same fundamental approach of pattern matching (of course with more convoluted notions).

2 Preliminary results and project plan

Our preliminary experiments demonstrated that utilizing a supervised training scheme with a simple attention augmented convolutional network [1] performs surprisingly well on the task of predicting the next symbol to be written on the paper by an algorithm that is utilized by humans to solve that problem. (E.g., predicting the next symbol to be written on the paper when utilizing the classic multiplication algorithm taught for children in elementary school.)

Based on this, our goal is to solve various simple mathematical problems by learning to carry out step-by-step reasoning procedures with our proposed approach. In particular, as the next step, we plan to implement step-by-step solutions for the problem set presented in [7].

In our talk, we will present the general idea, the above mentioned preliminary results, and report our progress on how far we can reach with this approach.

References

[1] Irwan Bello, Barret Zoph, Ashish Vaswani, Jonathon Shlens, and Quoc V. Le. Attention augmented convolutional networks. *CoRR*, abs/1904.09925, 2019.

[2] Kaiyu Chen, Yihan Dong, Xipeng Qiu, and Zitian Chen. Neural arithmetic expression calculator. *ArXiv*, abs/1809.08590, 2018.

[3] Andy Clark and David Chalmers. The extended mind. *Analysis*, 58(1):7–19, 1998.

[4] Sepp Hochreiter and Jürgen Schmidhuber. Long short-term memory. *Neural computation*, 9(8):1735–1780, 1997.

[5] Armand Joulin and Tomas Mikolov. Inferring algorithmic patterns with stack-augmented recurrent nets. In *NIPS*, 2015.

[6] Lukasz Kaiser and Ilya Sutskever. Neural gpus learn algorithms. In *4th International Conference on Learning Representations, ICLR 2016, San Juan, Puerto Rico, May 2-4, 2016, Conference Track Proceedings*, 2016.

[7] David Saxton, Edward Grefenstette, Felix Hill, and Pushmeet Kohli. Analysing mathematical reasoning abilities of neural models. In *International Conference on Learning Representations*, 2019.

[8] Zsolt Zombori, Adrián Csiszárik, Henryk Michalewski, Cezary Kaliszyk, and Josef Urban. Towards finding longer proofs. *CoRR*, abs/1905.13100, 2019.