Sequential models with long context

Supervisor
dr hab. Piotr Miłoś, prof. IMPAN (IDEAS NCBR)

Description

An enormous amount of training data used recently in language modelling (e.g. GPT) led to emerging properties (for example the models can handle a task, which they never encountered via prompting). We propose to study such models in area control (e.g. for robotic tasks). We speculate that by gathering in one model, a large number of skills can lead to more efficient learning on new tasks. The key questions to be studied during the project are: a) how to train a model capable of storing many tasks, b) how to query such a model efficiently, in order to learn new tasks faster, c) how to update such a model with a new task, while not forgetting the previous tasks. An instrumental to this goal will be endowing the transformer architecture with memory, e.g. [1]. Transformers have been extremely successful architectures in sequential modelling, however, they have a practical limitation of a relatively short context span due to the quadratic cost of the attention mechanism. The project aims to explore practical solutions to mitigate this problem by providing access to an external memory, which can be thought of as an external knowledge system. The aim is to factorise the reasoning capabilities, which could be stored in the weights of the transformer, from trivia facts, which can be stored in memory.
In recent work [2], we showed how the ideas of [1] can be extended to handle longer contexts. In this PhD. project we want to concentrate on applying novel unsupervised learning techniques to further extend the long-context capabilities.

Requirements
- MSc degree in computer science or related field
- Good knowledge of deep learning, including practical experience in Python and relevant libraries (TensorFlow, JaX, PyTorch, etc)
- Advanced skills in written and spoken English
- Prior research experience (e.g. publications in leading ML conferences) is a plus

References
[1] Y. Wu, M. Rabe, D. Hutchins, Ch. Szegedy. Memorizing Transformers, ICLR 2022
[2] Sz. Tworkowski, K. Staniszewski, M. Pacek, Y. Wu. H. Michalewski, P. Miłoś. Focused Transformer: Contrastive Training for Context Scaling, submitted to NeurIPS