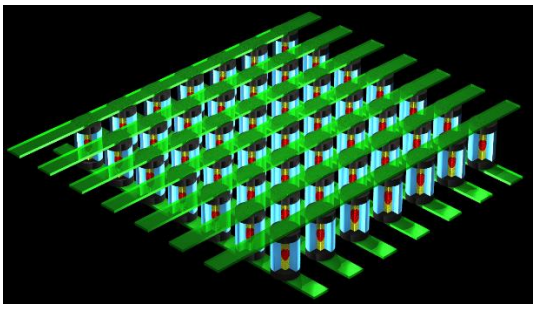
IBM Research –Zurich

Machine learning based on optimal transport using in-memory computing



For decades, conventional computers based on the von Neumann architecture have performed computation by repeatedly transferring data between their processing and their memory units, which are physically separated. As computation becomes increasingly data-centric and as the scalability limits in terms of performance and power are being reached, alternative computing paradigms are searched for in which computation and storage are collocated. A fascinating new approach is that of computational memory where the physics of nanoscale memory devices are used to perform certain computational tasks within the memory unit in a non-von Neumann manner.

Computational memory is finding applications in a variety of areas such as machine learning and signal processing [1]. Most importantly, it is very appealing for making energy-efficient inference hardware, where the relevant information would be encoded in crossbar arrays of memory devices [2][3]. Optimal Transport (OT) [4] has recently received significant attention in the machine learning community, especially because of the computational advantages of entropic regularization. However, there are several challenges that need to be overcome at both hardware and algorithmic levels to realize reliable and accurate learning engines based on OT using in-memory computing.

We are inviting applications from students to conduct their Master's thesis work at IBM Research – Zurich on this exciting new topic. The work performed could span low-level hardware experiments on phase-change memory chips comprising more than 1 million devices to high-level algorithmic development in a high-performance computing framework. It also involves interactions with several researchers across IBM research focusing on various aspects of the project. The ideal candidate should have a multi-disciplinary background, strong mathematical aptitude and programming skills. Prior knowledge on emerging memory technologies such as phase-change memory is a bonus but not necessary.

If you are interested in this challenging position on an exciting new topic, please send your most recent curriculum vitae including a transcript of grades by email to:

Dr. Giovanni Cherubini (cbi@zurich.ibm.com) and Dr. Abu Sebastian (ase@zurich.ibm.com)

[1] A. Sebastian, M. Le Gallo, R. Khaddam-Aljameh *et al.* Memory devices and applications for in-memory computing. *Nature Nanotechnology* (2020). https://doi.org/10.1038/s41565-020-0655-z

[2] V. Joshi, M. Le Gallo, S. Haefeli *et al.* Accurate deep learning inference using computational phase-change memory. *Nature Communications* (2020). Pre-print at https://arxiv.org/abs/1906.03138

[3] G. Karunaratne, M. Le Gallo, G. Cherubini, L. Benini, A. Rahimi, A. Sebastian. In-memory hyperdimensional computing. *Nature Electronics* (2020). Pre-print at https://arxiv.org/abs/1906.01548

[4] A. Alaoui-Belghiti, S. Chevallier, E. Monacelli *et al.* Semi-Supervised Optimal Transport Methods for Detecting Anomalies. In 2020 *IEEE International Conference on Acoustics, Speech and Signal Processing* (ICASSP), pp. 2997-3001.