THESIS PROPOSAL

ML
MACHINE LEARNING
DEPARTMENT

Neural Cognitive Architectures for Never-Ending Learning

Abstract:

Allen Newell argued that the human mind functions as a single system and proposed the notion of a unified theory of cognition (UTC). Most existing work on UTCs has focused on symbolic approaches, such as the Soar architecture (Laird, 2012) and the ACT-R (Anderson et al., 2004) system. However, such approaches limit a system's ability to perceive information of arbitrary modalities, require a significant amount of human input, and are restrictive in terms of the learning mechanisms they support (supervised learning, semi-supervised learning, reinforcement learning, etc.). For this reason, researchers in machine learning have recently shifted their focus towards subsymbolic processing with methods such as deep learning. Deep learning systems have become a standard for solving prediction problems in multiple application areas including computer vision, natural language processing, and robotics. However, many real-world problems require integrating multiple, distinct modalities of information (e.g., image, audio, language, etc.) in ways that machine learning models cannot currently handle well. Moreover, most deep learning approaches are not able to utilize information learned from solving one problem to directly help in solving another. They are also not capable of never-ending learning, failing on problems that are dynamic, ever-changing, and not fixed a priori, which is true of problems in the real world due to the dynamicity of nature. In this thesis, we aim to bridge the gap between UTCs, deep learning, and never-ending learning. To that end, we propose a neural cognitive architecture (NCA) that is inspired by human cognition and that can learn to continuously solve multiple problems that can grow in number over time, across multiple distinct perception and action modalities, and from multiple noisy sources of supervision combined with self-super vision. Furthermore, its experience from learning to solve past problems can be leveraged to learn to solve future ones. The problems the proposed NCA is learning to solve are ever evolving and can also be automatically generated by the system itself. In our NCA, reasoning is performed recursively in a subsymbolic latent space that is shared across all problems and modalities. The goal of this architecture is to take us a step closer towards general learning and intelligence. We have also designed, implemented, and plan to extend an artificial simulated world that allows us to test for all the aforementioned properties of the proposed archi-tecture, in a controllable manner. We propose to perform multiple case studies—within this simulated world and with real-world applications—that will allow us to evaluate our architecture.

Speaker:

Anthony Platanios

Thesis Committee:
Tom Mitchell (Chair)
Graham Neubig
Rich Caruana (Microsoft Research)
Eric Horvitz (Microsoft Research)

June 19, 2019
4:00pm
GHC 6501

Link to draft document:
https://www.dropbox.com/s/bamqghx36l34xan/proposal.pdf?dl=0