Tutorial: Recent Advances in Deep Learning

Matthew Botvinick (botvinick@google.com)
DeepMind, London, U.K.
Gatsby Computational Neuroscience Unit, University College London

Peter Battaglia (peterbattaglia@google.com)
DeepMind, London U.K.

Keywords: Deep learning, reinforcement learning, artificial intelligence

Overview

The past several years have seen a dramatic acceleration in artificial intelligence (AI) research, driven in large part by innovations in deep learning and reinforcement learning (RL) methods. The relevant developments, as showcased in a series of recent high-profile publications in *Nature* and elsewhere (e.g., Graves et al., 2016; Mnih et al., 2015; Silver et al., 2016), have generated intense interest in cognitive science, partially because they appear to have potentially far-reaching implications for understanding human intelligence. Unfortunately, the pace of innovation in AI has been so rapid that it is difficult for non-experts — and sometimes even for experts — to stay abreast of the latest developments.

The present tutorial brings together five front-line researchers in AI, each with dual credentials in neuroscience and/or cognitive science, to provide an accessible overview and update on the most important recent developments in deep learning and deep RL. The tutorial will be aimed at a broad audience, ranging from graduate students to senior investigators, and spanning specialties from cognitive and developmental psychology to psychiatry, human factors research, and systems neuroscience. The focus will be on fundamental concepts and principles, and a central goal will be to maximize accessibility, in line with the tutorial format.

Significance

Neural network modeling has played a pivotal role in cognitive science since at least the 1980's. Over the past decade or so, neural networks have been overshadowed to some extent by other techniques. Beginning around 2012, interest in neural network methods (often rebranded as 'deep learning') began to take off machine learning research, and have since then become the dominant approach in AI. In combination with RL methods, deep learning has enabled a series of breakthroughs in tasks ranging from image classification to game play (see Marblestone et al., 2016 for a review).

The implications of this tectonic shift for cognitive science are currently under intensive debate (Marblestone, 2016; Lake et al., 2016). It seems clear that AI innovations, including memory architectures, generative models, and deep RL techniques are likely to stimulate new hypotheses about human cognition. At the same time, it seems likely that AI research would benefit from richer input from cognitive science.

Ironically, the potential for exchange between the two fields has been hindered by the very pace of innovation in AI. (Emblematic is the fact that the developments we will review in the present tutorial have almost all emerged since the last time the Cognitive Science meeting featured a tutorial on neural networks, just two years ago.) Our aim in the present tutorial is to mitigate this problem by providing an accessible update on the most recent key developments in deep learning and deep RL.

Tutorial structure and activities

The tutorial will assume a half-day format, consisting of five tutorial lectures, each covering an area in which some of the most important recent innovations have arisen. As detailed below, all five lecturers are members of the research team at DeepMind in London (deepmind.com), with dual citizenship in AI and cognitive science. The material covered in each lecture will include recent work at DeepMind, but also related work from other groups.

Participant credentials

Matthew Botvinick (Organizer) is DeepMind's Director of Neuroscience Research and Honorary Professor in the Gatsby Computational Neuroscience Unit. He holds a Ph.D in Cognitive Neuroscience from CMU's Center for the Neural Basis of Cognition, and has done extensive research in the computational neuroscience of reinforcement learning and decision making. His research at DeepMind focuses in part on meta-learning.

Peter Battaglia (Organizer) is a senior Research Scientist at DeepMind. He holds a Ph.D. in Brain and Cognitive Sciences from MIT and has done extensive research in scene representation, intuitive physics, and probabilistic inference. His research at DeepMind focuses on novel architectures for structured inference, with an emphasis understanding physical systems.

Tim Lillicrap is a senior Research Scientist at DeepMind. He holds a Ph.D in Neuroscience from Queen's University, and has done high profile research on deep reinforcement learning and biologically plausible neural network learning algorithms. His research at DeepMind focuses on the interface between reinforcement learning and memory.

Greg Wayne, a senior Research Scientist at DeepMind, holds a Ph.D. in Neurobiology from Columbia University. He has conducted high profile work in hierarchical planning and deep RL. His research at DeepMind focuses on integrative architectures for artificial intelligence.

Daan Wierstra is a senior Research Scientist and research team leader at DeepMind. He holds a Ph.D. in Artificial Intelligence from IDSIA, and has conducted high-profile research in RL and neural networks. His research at DeepMind focuses in part on computational models of imagination and planning.

Presentations

As noted above, the tutorial will be comprised of five lectures, each covering a key area.

Deep Reinforcement Learning (Tim Lillicrap)

Advances in deep RL have driven some of the highest-profile recent work in AI, including differentiable neural computers (Graves et al., 2016) and superhuman play in the game of go (Silver et al., 2016). This talk will provide a tutorial review of the cutting edge in deep RL research.

Memory Architectures (Greg Wayne)

A major development over the past couple of years has been the incorporation of special modules for memory storage into deep learning AI systems (e.g., Graves et al. 2016; Santoro et al., 2016). This talk will review the state of the art, and consider the relationship with human episodic and working memory.

Structured Models for Structured Domains (Peter Battaglia)

Recent work in AI has introduced structure into deep learning architectures, which biases such systems toward particular forms of representation. Such measures have allowed dramatic advances in modeling physical systems and other structured domains (e.g., Battaglia et al., 2016). The present talk will review this approach and discuss its relation to the notion of compositional representation in cognitive science.

Deep Learning to Learn (Matthew Botvinick)

Recent work has explored the capacity of deep learning systems to 'learn how to learn,' leveraging previous experience to adapt more quickly to new challenges (e.g., Wang et al., 2016). This lecture will review recent progress toward endowing deep learning systems with this important capacity.

Deep Generative Models (Daan Wierstra)

One of the most exciting developments in recent deep learning research has been the rapid progress in building rich and flexible generative models, models that support operations like imagination and forecast-based planning (e.g., Gregor et al., 2015). This lecture will review the most recent techniques for building generative models, considering their many implications for cognitive science.

Related events

Yildirim and colleagues will present a half-day workshop focusing on the interface between recent AI advances and cognitive science. This can be considered a companion to the present tutorial, in that it will also cover techniques other than deep learning/RL, and will more deeply explore applications to cognitive and neuroscience.

References

Battaglia, P., Pascanu, R., Lai, M., & Rezende, D. J. (2016). Interaction networks for learning about objects, relations and physics. In *Advances in Neural Information Processing Systems* (pp. 4502-4510).

Graves, A., Wayne, G., Reynolds, M., Harley, T., Danihelka, I., Grabska-Barwińska, A., ... & Badia, A. P. (2016). Hybrid computing using a neural network with dynamic external memory. *Nature*, *538*(7626), 471-476.

Gregor, K., Danihelka, I., Graves, A., Rezende, D. J., & Wierstra, D. (2015). DRAW: A recurrent neural network for image generation. *arXiv preprint arXiv:1502.04623*.

Lake, B. M., Ullman, T. D., Tenenbaum, J. B., & Gershman, S. J. (2016). Building machines that learn and think like people. *arXiv preprint arXiv:1604.00289*.

Marblestone, A. H., Wayne, G., & Kording, K. P. (2016). Toward an integration of deep learning and neuroscience. *Frontiers in Computational Neuroscience*, *10*.

Mnih, V., Kavukcuoglu, K., Silver, D., Rusu, A. A., Veness, J., Bellemare, M. G., ... & Petersen, S. (2015). Human-level control through deep reinforcement learning. *Nature*, *518*(7540), 529-533.

Rezende, D. J., Mohamed, S., Danihelka, I., Gregor, K., & Wierstra, D. (2016). One-shot generalization in deep generative models. *arXiv preprint arXiv:1603.05106*.

Santoro, A., Bartunov, S., Botvinick, M., Wierstra, D., & Lillicrap, T. (2016). One-shot learning with memory-augmented neural networks. *arXiv preprint arXiv:1605.06065*.

Silver, D., Huang, A., Maddison, C. J., Guez, A., Sifre, L., Van Den Driessche, G., ... & Dieleman, S. (2016). Mastering the game of Go with deep neural networks and tree search. *Nature*, *529*(7587), 484-489.

Wang, J. X., Kurth-Nelson, Z., Tirumala, D., Soyer, H., Leibo, J. Z., Munos, R., ... & Botvinick, M. (2016). Learning to reinforcement learn. *arXiv preprint arXiv:1611.05763*.