Tutorial Workshop on Contemporary Deep Neural Network Models

James L. McClelland (mcclelland@stanford.edu)
Steven Hansen (sshansen@stanford.edu)
Andrew Saxe (asaxe@fas.harvard.edu)
Department of Psychology, Stanford University
Stanford, CA, 95305, USA

Keywords: deep learning; neural networks; unsupervised learning; computational modeling; reinforcement learning, simulation tools

Contemporary Deep Neural Networks

Neural network models were exploited in the late 80's and 90's to model human cognition, based on developments such as the back propagation learning algorithm. Over the last 10 years, massive data sets, enhanced computational resources, and new developments in algorithms have led to explosive growth in use of such models in machine learning and artificial intelligence. We propose a workshop to bring the tools, methods, and insights from this research into contemporary cognitive science and cognitive neuroscience.

Here we mention three relevant recent architectures: Deep Neural Networks (DNNs), Deep Reinforcement Learning Networks (DRNs), and Recurrent Neural Networks with Long-Short-Term Memory Units (RNN-LSTMs). Each is a rich architectural framework rather than a single fixed model and each has been used on a different class of problem. DNNs (including convolutional neural networks, CNNs) are trained using a combination of supervised and unsupervised learning methods, and are used in vision, object recognition and other tasks including numerosity judgment; DRNs learn optimal action policies in video games and other action selection settings; RNN-LSTMs have been used for cross-domain mapping, either between different spoken languages or (combined with DNNs) between video and spoken language and for creating provocative new cognitive models such as the Neural Turing Machine (NTM). Together the architectures incorporate many key tools and concepts, including convolution, unsupervised learning and unsupervised pre-training, enhanced methods for reinforcement learning, contemporary regularization methods (to prevent overfitting), and the LSTM mechanism—a mechanism that learns to gate information into and out of buffers to maintain context over long sequences. These and related architectures, applications, tools, and concepts will be the focus of our workshop.

Presenters and their Contributions

Literature citations highlighting the work of the presenters in the areas described above are contained in the bibliography below. Several of the authors have released preprints on *arXiv* of papers first presented at top Machine Learning conferences (NIPS, CVCPR) prior to publication in journals due to the fast-moving nature of research in this field. Here we mention briefly each presenter's expertise and background.

Marco Zorzi, Professor, University of Padova (PhD in Cognitive Science, University of Trieste, 1999). Has applied deep networks to modeling human numerosity judgment and reading and has developed tools for efficient implementation of these models.
Nikolaus Kriegeskorte, Program Leader, Memory and Perception Group, MRC-CBU. Has applied a deep convolutional neural network architecture introduced in 2012 to model human voxel-level activity patterns in different layers of visual cortex.
Timothy Lillicrap, Senior Research Scientist, Google DeepMind, London UK (PhD in Systems Neuroscience, Queen's University, 2012). Leading the effort to extend Deep Reinforcement Learning approaches that allow simulated agents to learn sophisticated continuous motor control policies.
Greg Wayne, Research Scientist, Google DeepMind, London, UK (PhD in Neuroscience, Columbia University, 2013). One of the creators of the Neural Turing Machine, a Deep Learning Model that relies on the LSTM mechanism for the storage and retrieval of information in memory.
Alberto Testolin, Post-Doctoral Fellow, University of Padova. Testolin is a co-author with Zorzi on several deep learning papers including tutorial articles.

Presenter Contributions

Organizer McClelland will provide historical context and an overall perspective on the role of deep neural networks in contemporary cognitive science, and Zorzi, Kriegeskorte, Lillicrap, and Wayne will each lead tutorials on unsupervised DNN's; CNN's; DQN's and other reinforcement learning methods; and RNN's including LSTM's and other methods that allow networks to span very long distance dependencies. These tutorials will cover the basic network paradigm and architectures, algorithms, applications, and best practices.

Testolin, Kriegeskorte, Lillicrap, and Wayne will each host breakouts in the afternoon in which participants will have to opportunity to run example neural networks of each type. During these sessions participants will try out existing networks and begin to develop their own plans guided by the tutorial leaders. Advanced preparation for this will be coordinated by co-Organizers Hansen, a PhD student in

McClelland's lab who has spent 4 months as an intern at DeepMind, and Saxe, formerly of McClelland's lab and now a post-doc in computational neuroscience at Harvard. Hansen and Saxe will ensure that the relevant software tools are available for remote use and/or download to Workshop Participant's laptops. Their expertise will also allow them to contribute to helping students formulate specific proposals for their own models and to configure software for their implementation.

Workshop Structure

The morning session will begin with a brief introduction by McClelland (10 min), followed by 40 minute presentations by Zorzi, Kriegeskorte, Lillicrap, and Wayne, with a 10 minute break after Kriegeskorte's presentation. The afternoon session will begin with a brief overview of available software tools by Lillicrap with input from others. This will be followed by a description of the specific software tools and demonstration simulations prepared for the workshop, and short descriptions of the focus of each of the four breakout sessions. During the breakouts, workshop participants will discuss with presenters/breakout leaders possible applications they might develop themselves and will begin to work on implementation. Starting at about 3pm, a subset of participants from each group will describe their ideas to all the workshop participants, allowing for all presenters to offer feedback and for broad sharing of ideas. A final Q&A session with the presenters as a panel will round out the session, which will end at 4 pm.

Breakouts will be organized in different corners of the allocated conference room. If there is sufficient interest and available space, part of afternoon session might use a second room; in that case, one room might be allocated to DNN's and CNN's, and the other to reinforcement learning and recurrent networks, depending on the interests of participants. We hope to coordinate on this with the Cognitive Science meeting program committee as participant interest becomes better defined.

Budget

We request $1,200 toward support for registration fees (including Workshop registration). Although registration cost information is not currently available, we expect that this amount would cover approximate 4 participants who can document inability to cover costs of registration.

Other Funding and Coordination Information

Some funding for this tutorial workshop will be provided by the Rumelhart Emergent Cognitive Functions fund created by the proceeds from McClelland's Rumelhart Prize. This will be used to defray presenter's registration for those presenters who will not otherwise be participating in the Cognitive Science Society Meeting, as well as one hotel night for presenters.

We also note that McClelland, with Stefan Frank of Radboud University, Netherlands, and CogSci 2016 program committee member Daniel Mirman will be hosting a related longer co-located workshop (Neural Computation and Psychology – NCPW 2016) at Drexel University in Philadelpia on the 8th and 9th of August, just prior to this Cogsci Tutorial workshop. The NCPW workshop, supported by the W.K. and K. W. Estes Fund, the Rumelhart Fund, and a gift from Google DeepMind, will cover the travel and accommodation for all of the listed presenters, who will also be participating in the earlier workshop. Their presentations in the NCPW workshop will focus on concepts and research results; their presentations in this CogSci tutorial workshop will focus on algorithms, best practices, and implementation.

References

Graves, A., Wayne, G., & Danihelka, I. (2014). Neural Turing Machines. *arXiv preprint arXiv:1410.5401*.

Heess, N., Wayne, G., Silver, D., Lillicrap, T., Erez, T., & Tassa, Y. (2015). Learning continuous control policies by stochastic value gradients. In *Advances in Neural Information Processing Systems* (pp. 2926-2934).

Khaligh-Razavi, S. M., & Kriegeskorte, N. (2014). Deep Supervised, but Not Unsupervised, Models May Explain IT Cortical Representation. PLoS Computational Biology, 10(11). E1003915.

Lillicrap, T. P., Hunt, J. J., Pritzel, A., Heess, N., Erez, T., Tassa, Y., & Wierstra, D. (2015). Continuous control with deep reinforcement learning. *arXiv preprint arXiv:1509.02971*.

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.

Saxe, A. M., McClelland, J. L., & Ganguli, S. (2013). Exact solutions to the nonlinear dynamics of learning in deep linear neural networks. *arXiv preprint arXiv:1312.6120*.

Silver, D, Huang, A., Maddison, C. J., … Lillicrap, T., … & Haassabis, D. (2016). Mastering the game of Go with deep neural networks and tree search. *Nature*, 529, 484-489.

Stoianov, I., & Zorzi, M. (2012). Emergence of a'visual number sense'in hierarchical generative models. *Nature Neuroscience, 15(2),* 194-196.

Zorzi, M., Testolin, A., & Stoianov, I. P. (2013). Modeling language and cognition with deep unsupervised learning: a tutorial overview. *Frontiers in psychology, 4*.