

# CNS\*2020 notes

Ankur Sinha

Monday 27<sup>th</sup> July, 2020

- Conference website: <https://www.cnsorg.org>
- Full program, with links to recordings: <https://cns2020online.sched.com/> (Password: CNS\*2020!)

## 1 Keynotes

### 1.1 Matt Botvinick: Deep reinforcement learning and its neuroscientific implications

The recording can be found on YouTube here: <https://youtu.be/H4Sux8Q89cg>.

Talk is based closely on the paper:

Matthew Botvinick et al. "Deep Reinforcement Learning and Its Neuroscientific Implications". In: *Neuron* (July 2020). DOI: [10.1016/j.neuron.2020.06.014](https://doi.org/10.1016/j.neuron.2020.06.014).

Notes:

- Deep learning techniques have been around for a long time. We're just learning how to scale them, and more and more what they're capable of.
- However, *Deep Reinforcement Learning* is new and does not get the same amount of focus as Deep Learning.
- In a lot of RL work, rather impoverished representations of the environment have been used.
  - state representation, sometimes referred to as "tabular", consists of the world carved up into states,
  - but these states tend to be orthogonal to each other and do not have representational similarities to each other,
  - no shared features, no similarity space that can be exploited to generalise to new but related state or to transfer to problems that haven't been encountered before but which may share features with tasks that you've performed in the past.

- Deep learning at its root is just representation learning, and RL is just motivated behaviour. Thus, strongly linked to neuroscience.
- Deep Reinforcement Learning (DRL) combines both. It is greater than the sum of the two individual deep and reinforcement learning parts.
  - Meta-reinforcement learning [Wan+18]: instead of one task, train on a family of tasks and then at test time, disable RL and give the system a brand new task.
  - Given that learning has been disabled and thus cannot learn about the new task, it should do poorly.
  - However, turns out, it does quite well.
  - The RL that adjusts connection weights in this system also has the effect of tuning the activation dynamics in the system. These activation dynamics, over time, come to implement their own separate RL. So, in an emergent way, the system learns to solve problems using “working memory” rather than weight changes.
  - Meta-RLs are Bayesian learners [Ort+19].
  - In Ortega et al. [Ort+19], they showed that Meta-RLs could reproduce tasks used in neuroscience studies as emergent properties:
    - \* Probability matching [Tsu+16]. Also matched statistics from neural recordings.
    - \* Changes in learning rate in the brain depending on underlying reward probability [Beh+07] even though the Meta-RL system has a fixed learning rate.
    - \* Flipping in dopamine response [Bro+10].
  - In “Distributional TD learning” [Dab+20], they split the learning rate into two components such that they can be independently scaled for positive and negative reward prediction errors.
    - \* This follows that dopamine neurons should have different scaling factors for positive and negative reward prediction errors, and the relative scales of these parameters should be different across dopamine neuron. They observed this in mice.
  - Unsupervised DL [Hig+20]:  $\beta$ -VAE: generate disentangled representations. The brain also seems to do this type of disentangling.
    - \* Reconstruction: they can take as few as 12 neurons from the monkey’s brain, and using the parameters from the  $\beta$ -VAE, they could re-construct what the monkey was looking at.
- Model based RL [Sil+17; Sil+16] (Alpha-Go).

- Two step task: animals can show insight about the causal structure of a domain by observing the outcome of one action and modifying probability of a different action [Daw+11].
- Seen in rodents, and seen in their meta-RL agent which does not use model-based RL to learn, so it emerges from model-free learning.
- Memory and DRL [Gra+16].
- Motor control and DRL [Mer+19].
- Social cognition and DRL [McK+20].
- Connectomics and DRL [TRT18].
- DRL: caveats and challenges:
  - It's engineering, not neuroscience, so more work needs to be done to translate findings to neuroscience.
  - Back-propagation remains controversial.
  - It cannot handle continuous learning.
  - Do not do well in long term temporary credit assignment problems,
  - Do not currently match humans in flexibility and sampling efficiency.

Work cited in the talk:

- DRL:
  - Alex Krizhevsky, Ilya Sutskever, and Geoffrey E. Hinton. "Imagenet classification with deep convolutional neural networks". In: *Advances in neural information processing systems*. 2012, pp. 1097–1105.
  - Valerio Mante et al. "Context-dependent computation by recurrent dynamics in prefrontal cortex". In: *Nature* 503.7474 (Nov. 2013), pp. 78–84. DOI: [10.1038/nature12742](https://doi.org/10.1038/nature12742).
  - David Zipser. "Recurrent Network Model of the Neural Mechanism of Short-Term Active Memory". In: *Neural Computation* 3.2 (1991). PMID: 31167307, pp. 179–193. DOI: [10.1162/neco.1991.3.2.179](https://doi.org/10.1162/neco.1991.3.2.179). eprint: <https://doi.org/10.1162/neco.1991.3.2.179>. URL: <https://doi.org/10.1162/neco.1991.3.2.179>.
  - Richard S. Sutton and Andrew G. Barto. *Reinforcement learning: An introduction*. MIT press, 2018.

- Gerald Tesauro. "TD-Gammon: A Self-Teaching Backgammon Program". In: *Applications of Neural Networks*. Ed. by Alan F. Murray. Boston, MA: Springer US, 1995, pp. 267–285. ISBN: 978-1-4757-2379-3. DOI: [10.1007/978-1-4757-2379-3\\_11](https://doi.org/10.1007/978-1-4757-2379-3_11). URL: [https://doi.org/10.1007/978-1-4757-2379-3\\_11](https://doi.org/10.1007/978-1-4757-2379-3_11).
- A. G. Barto. "Temporal difference learning". In: *Scholarpedia* 2.11 (2007). revision #186927, p. 1604. DOI: [10.4249/scholarpedia.1604](https://doi.org/10.4249/scholarpedia.1604).
- Volodymyr Mnih et al. "Human-level control through deep reinforcement learning". In: *Nature* 518.7540 (Feb. 2015), pp. 529–533. DOI: [10.1038/nature14236](https://doi.org/10.1038/nature14236).
- Oriol Vinyals et al. "Grandmaster level in StarCraft II using multi-agent reinforcement learning". In: *Nature* 575.7782 (Oct. 2019), pp. 350–354. DOI: [10.1038/s41586-019-1724-z](https://doi.org/10.1038/s41586-019-1724-z).
- Max Jaderberg et al. "Human-level performance in 3D multiplayer games with population-based reinforcement learning". In: *Science* 364.6443 (May 2019), pp. 859–865. DOI: [10.1126/science.aau6249](https://doi.org/10.1126/science.aau6249).
- W. Schultz, P. Dayan, and P. R. Montague. "A Neural Substrate of Prediction and Reward". In: *Science* 275.5306 (Mar. 1997), pp. 1593–1599. DOI: [10.1126/science.275.5306.1593](https://doi.org/10.1126/science.275.5306.1593).
- Jane X. Wang et al. "Prefrontal cortex as a meta-reinforcement learning system". In: *Nature Neuroscience* 21.6 (May 2018), pp. 860–868. DOI: [10.1038/s41593-018-0147-8](https://doi.org/10.1038/s41593-018-0147-8).
- Pedro A. Ortega et al. *Meta-learning of Sequential Strategies*. 2019. arXiv: [1905.03030 \[cs.LG\]](https://arxiv.org/abs/1905.03030).
- Jens Volkmann, Christine Daniels, and Karsten Witt. "Neuropsychiatric effects of subthalamic neurostimulation in Parkinson disease". In: *Nature Reviews Neurology* 6.9 (Aug. 2010), pp. 487–498. DOI: [10.1038/nrneuro.2010.111](https://doi.org/10.1038/nrneuro.2010.111).
- Ken-Ichiro Tsutsui et al. "A dynamic code for economic object valuation in prefrontal cortex neurons". In: *Nature Communications* 7.1 (Sept. 2016). DOI: [10.1038/ncomms12554](https://doi.org/10.1038/ncomms12554).
- Timothy E. J. Behrens et al. "Learning the value of information in an uncertain world". In: *Nature Neuroscience* 10.9 (Aug. 2007), pp. 1214–1221. DOI: [10.1038/nn1954](https://doi.org/10.1038/nn1954).
- Will Dabney et al. "A distributional code for value in dopamine-based reinforcement learning". In: *Nature* 577.7792 (Jan. 2020), pp. 671–675. DOI: [10.1038/s41586-019-1924-6](https://doi.org/10.1038/s41586-019-1924-6).

- Ethan S. Bromberg-Martin et al. "A Pallidus-Habenula-Dopamine Pathway Signals Inferred Stimulus Values". In: *Journal of Neurophysiology* 104.2 (Aug. 2010), pp. 1068–1076. DOI: [10.1152/jn.00158.2010](https://doi.org/10.1152/jn.00158.2010).
- Irina Higgins et al. *Unsupervised deep learning identifies semantic disentanglement in single inferotemporal neurons*. 2020. arXiv: [2006.14304 \[q-bio.NC\]](https://arxiv.org/abs/2006.14304).
- David Silver et al. "Mastering the game of Go with deep neural networks and tree search". In: *Nature* 529.7587 (Jan. 2016), pp. 484–489. DOI: [10.1038/nature16961](https://doi.org/10.1038/nature16961),
- David Silver et al. "Mastering the game of Go without human knowledge". In: *Nature* 550.7676 (Oct. 2017), pp. 354–359. DOI: [10.1038/nature24270](https://doi.org/10.1038/nature24270).
- Kevin J. Miller, Matthew M. Botvinick, and Carlos D. Brody. "Dorsal hippocampus contributes to model-based planning". In: *Nature Neuroscience* 20.9 (July 2017), pp. 1269–1276. DOI: [10.1038/nn.4613](https://doi.org/10.1038/nn.4613).
- Nathaniel D. Daw et al. "Model-Based Influences on Humans' Choices and Striatal Prediction Errors". In: *Neuron* 69.6 (Mar. 2011), pp. 1204–1215. DOI: [10.1016/j.neuron.2011.02.027](https://doi.org/10.1016/j.neuron.2011.02.027).
- Alex Graves et al. "Hybrid computing using a neural network with dynamic external memory". In: *Nature* 538.7626 (Oct. 2016), pp. 471–476. DOI: [10.1038/nature20101](https://doi.org/10.1038/nature20101).
- Josh Merel et al. *Deep neuroethology of a virtual rodent*. 2019. arXiv: [1911.09451 \[q-bio.NC\]](https://arxiv.org/abs/1911.09451).
- Kevin R. McKee et al. "Social Diversity and Social Preferences in Mixed-Motive Reinforcement Learning". In: *arXiv preprint arXiv:2002.02325* (2020).
- Fabian David Tschopp, Michael B. Reiser, and Srinivas C. Turaga. *A Connectome Based Hexagonal Lattice Convolutional Network Model of the Drosophila Visual System*. 2018. arXiv: [1806.04793 \[q-bio.NC\]](https://arxiv.org/abs/1806.04793).

## References

- [Bar07] A. G. Barto. "Temporal difference learning". In: *Scholarpedia* 2.11 (2007). revision #186927, p. 1604. DOI: [10.4249/scholarpedia.1604](https://doi.org/10.4249/scholarpedia.1604) (cit. on p. 4).
- [Beh+07] Timothy E. J. Behrens et al. "Learning the value of information in an uncertain world". In: *Nature Neuroscience* 10.9 (Aug. 2007), pp. 1214–1221. DOI: [10.1038/nn1954](https://doi.org/10.1038/nn1954) (cit. on pp. 2, 4).

- [Bot+20] Matthew Botvinick et al. “Deep Reinforcement Learning and Its Neuroscientific Implications”. In: *Neuron* (July 2020). DOI: [10.1016/j.neuron.2020.06.014](https://doi.org/10.1016/j.neuron.2020.06.014) (cit. on p. 1).
- [Bro+10] Ethan S. Bromberg-Martin et al. “A Pallidus-Habenula-Dopamine Pathway Signals Inferred Stimulus Values”. In: *Journal of Neurophysiology* 104.2 (Aug. 2010), pp. 1068–1076. DOI: [10.1152/jn.00158.2010](https://doi.org/10.1152/jn.00158.2010) (cit. on pp. 2, 5).
- [Dab+20] Will Dabney et al. “A distributional code for value in dopamine-based reinforcement learning”. In: *Nature* 577.7792 (Jan. 2020), pp. 671–675. DOI: [10.1038/s41586-019-1924-6](https://doi.org/10.1038/s41586-019-1924-6) (cit. on pp. 2, 4).
- [Daw+11] Nathaniel D. Daw et al. “Model-Based Influences on Humans’ Choices and Striatal Prediction Errors”. In: *Neuron* 69.6 (Mar. 2011), pp. 1204–1215. DOI: [10.1016/j.neuron.2011.02.027](https://doi.org/10.1016/j.neuron.2011.02.027) (cit. on pp. 3, 5).
- [Gra+16] Alex Graves et al. “Hybrid computing using a neural network with dynamic external memory”. In: *Nature* 538.7626 (Oct. 2016), pp. 471–476. DOI: [10.1038/nature20101](https://doi.org/10.1038/nature20101) (cit. on pp. 3, 5).
- [Hig+20] Irina Higgins et al. *Unsupervised deep learning identifies semantic disentanglement in single inferotemporal neurons*. 2020. arXiv: [2006.14304](https://arxiv.org/abs/2006.14304) [q-bio.NC] (cit. on pp. 2, 5).
- [Jad+19] Max Jaderberg et al. “Human-level performance in 3D multi-player games with population-based reinforcement learning”. In: *Science* 364.6443 (May 2019), pp. 859–865. DOI: [10.1126/science.aau6249](https://doi.org/10.1126/science.aau6249) (cit. on p. 4).
- [KSH12] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E. Hinton. “Imagenet classification with deep convolutional neural networks”. In: *Advances in neural information processing systems*. 2012, pp. 1097–1105 (cit. on p. 3).
- [Man+13] Valerio Mante et al. “Context-dependent computation by recurrent dynamics in prefrontal cortex”. In: *Nature* 503.7474 (Nov. 2013), pp. 78–84. DOI: [10.1038/nature12742](https://doi.org/10.1038/nature12742) (cit. on p. 3).
- [McK+20] Kevin R. McKee et al. “Social Diversity and Social Preferences in Mixed-Motive Reinforcement Learning”. In: *arXiv preprint arXiv:2002.02325* (2020) (cit. on pp. 3, 5).
- [Mer+19] Josh Merel et al. *Deep neuroethology of a virtual rodent*. 2019. arXiv: [1911.09451](https://arxiv.org/abs/1911.09451) [q-bio.NC] (cit. on pp. 3, 5).
- [MBB17] Kevin J. Miller, Matthew M. Botvinick, and Carlos D. Brody. “Dorsal hippocampus contributes to model-based planning”. In: *Nature Neuroscience* 20.9 (July 2017), pp. 1269–1276. DOI: [10.1038/nn.4613](https://doi.org/10.1038/nn.4613) (cit. on p. 5).

- [Mni+15] Volodymyr Mnih et al. "Human-level control through deep reinforcement learning". In: *Nature* 518.7540 (Feb. 2015), pp. 529–533. DOI: [10.1038/nature14236](https://doi.org/10.1038/nature14236) (cit. on p. 4).
- [Ort+19] Pedro A. Ortega et al. *Meta-learning of Sequential Strategies*. 2019. arXiv: [1905.03030](https://arxiv.org/abs/1905.03030) [cs.LG] (cit. on pp. 2, 4).
- [SDM97] W. Schultz, P. Dayan, and P. R. Montague. "A Neural Substrate of Prediction and Reward". In: *Science* 275.5306 (Mar. 1997), pp. 1593–1599. DOI: [10.1126/science.275.5306.1593](https://doi.org/10.1126/science.275.5306.1593) (cit. on p. 4).
- [Sil+16] David Silver et al. "Mastering the game of Go with deep neural networks and tree search". In: *Nature* 529.7587 (Jan. 2016), pp. 484–489. DOI: [10.1038/nature16961](https://doi.org/10.1038/nature16961) (cit. on pp. 2, 5).
- [Sil+17] David Silver et al. "Mastering the game of Go without human knowledge". In: *Nature* 550.7676 (Oct. 2017), pp. 354–359. DOI: [10.1038/nature24270](https://doi.org/10.1038/nature24270) (cit. on pp. 2, 5).
- [SB18] Richard S. Sutton and Andrew G. Barto. *Reinforcement learning: An introduction*. MIT press, 2018 (cit. on p. 3).
- [Tes95] Gerald Tesauro. "TD-Gammon: A Self-Teaching Backgammon Program". In: *Applications of Neural Networks*. Ed. by Alan F. Murray. Boston, MA: Springer US, 1995, pp. 267–285. ISBN: 978-1-4757-2379-3. DOI: [10.1007/978-1-4757-2379-3\\_11](https://doi.org/10.1007/978-1-4757-2379-3_11). URL: [https://doi.org/10.1007/978-1-4757-2379-3\\_11](https://doi.org/10.1007/978-1-4757-2379-3_11) (cit. on p. 4).
- [TRT18] Fabian David Tschoep, Michael B. Reiser, and Srinivas C. Turaga. *A Connectome Based Hexagonal Lattice Convolutional Network Model of the Drosophila Visual System*. 2018. arXiv: [1806.04793](https://arxiv.org/abs/1806.04793) [q-bio.NC] (cit. on pp. 3, 5).
- [Tsu+16] Ken-Ichiro Tsutsui et al. "A dynamic code for economic object valuation in prefrontal cortex neurons". In: *Nature Communications* 7.1 (Sept. 2016). DOI: [10.1038/ncomms12554](https://doi.org/10.1038/ncomms12554) (cit. on pp. 2, 4).
- [Vin+19] Oriol Vinyals et al. "Grandmaster level in StarCraft II using multi-agent reinforcement learning". In: *Nature* 575.7782 (Oct. 2019), pp. 350–354. DOI: [10.1038/s41586-019-1724-z](https://doi.org/10.1038/s41586-019-1724-z) (cit. on p. 4).
- [VDW10] Jens Volkmann, Christine Daniels, and Karsten Witt. "Neuropsychiatric effects of subthalamic neurostimulation in Parkinson disease". In: *Nature Reviews Neurology* 6.9 (Aug. 2010), pp. 487–498. DOI: [10.1038/nrneuro.2010.111](https://doi.org/10.1038/nrneuro.2010.111) (cit. on p. 4).
- [Wan+18] Jane X. Wang et al. "Prefrontal cortex as a meta-reinforcement learning system". In: *Nature Neuroscience* 21.6 (May 2018), pp. 860–868. DOI: [10.1038/s41593-018-0147-8](https://doi.org/10.1038/s41593-018-0147-8) (cit. on pp. 2, 4).

- [Zip91] David Zipser. "Recurrent Network Model of the Neural Mechanism of Short-Term Active Memory". In: *Neural Computation* 3.2 (1991). PMID: 31167307, pp. 179–193. DOI: [10.1162/neco.1991.3.2.179](https://doi.org/10.1162/neco.1991.3.2.179). eprint: <https://doi.org/10.1162/neco.1991.3.2.179>. URL: <https://doi.org/10.1162/neco.1991.3.2.179> (cit. on p. 3).