

Theories of Learning

Samuel Gershman

This course provides a tour of foundational topics in learning from a theoretical perspective. It covers a diversity of learning processes, aiming for breadth over depth (although it inevitably neglects several important forms of learning). Each meeting will consist of student-led presentations of two papers. Experience with computational modeling is not required. The course is aimed at graduate students and advanced undergraduates.

Meeting 1: Foundational concepts

- Jacobs, R.A., & Kruschke, J.K. (2011). Bayesian learning theory applied to human cognition. *Wiley Interdisciplinary Reviews: Cognitive Science*, 2, 8-21.
- Hinton, G. (1989). Connectionist learning procedures. *Artificial Intelligence*, 40, 185-234.

Meeting 2: Classical conditioning

- Rescorla, R., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: variations in the effectiveness of reinforcement and nonreinforcement. In A. Black & W. Prokasy (Eds.), *Classical Conditioning II: Current research and theory* (pp. 64–99). New York, NY: Appleton-Century-Crofts.
- Courville, A.C., Daw, N.D., & Touretzky, D.S. (2006). Bayesian theories of conditioning in a changing world. *Trends in Cognitive Sciences*, 10, 294-300.

Meeting 3: Attention in associative learning

- Pearce, J.M., & Hall, G. (1980). A model for Pavlovian learning: Variations in the effectiveness of conditioned but not of unconditioned stimuli. *Psychological Review*, 87, 532-552.
- Mackintosh, N. J. (1975). A theory of attention: Variations in the associability of stimuli with reinforcement. *Psychological Review*, 82, 276–298.

Meeting 4: Reinforcement learning

- Schultz, W., Dayan, P., & Montague, P.R. (1997). A neural substrate of prediction and reward. *Science*, 275, 1593-1599.
- Daw, N.D., Niv, Y., & Dayan, P. (2005). Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. *Nature Neuroscience*, 8, 1704-1711.

Meeting 5: Skill acquisition

- Newell, K. M. & Rosenbloom, P. S. Mechanisms of skill acquisition and the law of practice. In Anderson, J. R. (ed.) *Cognitive skills and their acquisition*, 1–55 (Lawrence Erlbaum Associates, 1981).
- Anderson, J. R. (1982). Acquisition of cognitive skill. *Psychological Review*, 89, 369-403.

Meeting 6: Category learning

- Kruschke, J. K. (1992). ALCOVE: An exemplar-based connectionist model of category learning. *Psychological Review*, 99, 22–44.
- Anderson, J.R. (1991). The adaptive nature of human categorization. *Psychological Review*, 98, 409-429.

Meeting 7: Rule-based concept learning

- Nosofsky, R.M., Palmeri, T.J., & McKinley, S.C. (1994). Rule-plus-exception model of classification learning. *Psychological Review*, 101, 53–79.
- Goodman, N. D., Tenenbaum, J. B., Feldman, J., & Griffiths, T.L. (2008). A rational analysis of rule-based concept learning. *Cognitive Science*, 32, 108–154.

Meeting 8: Causal induction

- Gopnik, A., Glymour, C., Sobel, D.M., Schulz, L.E., Kushnir, T., & Danks, D. (2004). A theory of causal learning in children: Causal maps and Bayes nets. *Psychological Review*, 111, 3-32.
- Cheng, P.W. (1997). From covariation to causation: a causal power theory. *Psychological Review*, 104, 367-405.

Meeting 9: Active learning

- Oaksford, M., & Chater, N. (1994). A rational analysis of the selection task as optimal data selection. *Psychological Review*, 101, 608-631.
- Markant, D.B., Settles, B., & Gureckis, T.M. (2016). Self-directed learning favors local, rather than global, uncertainty. *Cognitive Science*, 40, 1-21.

Meeting 10: Social learning

- Laland, K. N. 2004. Social learning strategies. *Learning and Behavior*, 32, 4-14.
- Shafto, P., Goodman, N.D., & Griffiths, T.L. (2014). A rational account of pedagogical reasoning: Teaching by, and learning from, examples. *Cognitive Psychology*, 71, 55-89.

Meeting 11: Generalization

- Shepard, R.N. (1987). Toward a universal law of generalization for psychological science. *Science*, 237, 1317-1323.
- Marcus, G.F. (1998). Rethinking eliminative connectionism. *Cognitive Psychology*, 37, 243-282.

Meeting 12: Language acquisition

- Xu, F., & Tenenbaum, J. (2007). Word learning as Bayesian inference. *Psychological Review*, 114, 245–272.
- Seidenberg, M.S., & MacDonald, M.C. (1999). A probabilistic constraints approach to language acquisition and processing. *Cognitive Science*, 23, 569–588.

Meeting 13: Inductive constraints

- Keil, F.C. (1981). Constraints on knowledge and cognitive development. *Psychological Review*, 88, 197-227.
- Hsu, A.S., & Chater, N. (2010). The logical problem of language acquisition: a probabilistic perspective. *Cognitive Science*, 34, 972-1016.

Meeting 14: Interactions between development and learning

- Elman, J.L. (1993). Learning and development in neural networks: the importance of starting small. *Cognition*, 48, 71-99.
- Quartz, S.R., & Sejnowski, T.J. (1997). The neural basis of cognitive development: a constructivist manifesto. *Behavioral and Brain Sciences*, 20, 537-596.

Meeting 15: Interactions between evolution and learning

- Hinton, G., & Nowlan, S. (1987). How learning can guide evolution. *Complex Systems*, 1, 495-502.
- Smith, K., Kirby, S., & Brighton, H. (2003). Iterated learning: A framework for the evolution of language. *Artificial Life*, 9, 371-386.

Meeting 16: Structure discovery

- Doumas, L.A.A., Hummel, J.E., & Sandhofer, C.M. (2008). A theory of the discovery and predication of relational concepts. *Psychological Review*, 115, 1-43.
- Kemp, C., Tenenbaum, J., Niyogi, S., & Griffiths, T. (2010). A probabilistic model of theory formation. *Cognition*, 114, 165-196.

Meeting 17: Neurobiology of learning

- Martin, S.J., Grimwood, P.D., & Morris R.G.M. (2000). Synaptic plasticity and memory: an evaluation of the hypothesis. *Annual Review of Neuroscience*, 23, 649-711.
- Gallistel, C.R., & Matzel, L.D. (2013). The neuroscience of learning: beyond the Hebbian synapse. *Annual Review of Neuroscience*, 64, 169-200.