Computational [Principles of] Psychology*

Shimon Edelman
Department of Psychology
Cornell University
Ithaca, NY 14853
`http://kybele.psych.cornell.edu/~edelman`

VERSION 2.0 — March 2, 2017

Abstract

This course states, motivates, and offers detailed support for the observation that cognition is fundamentally a computational process [30]. Students are introduced to a number of conceptual tools for thinking about natural behavior and the cognitive information processing that underlies it, including statistical learning from experience and the use of patterns distilled from past experience in guiding future actions. The application of these tools to the understanding of natural minds and to advancing the goals of artificial intelligence is illustrated on selected examples drawn from the domains of perception, memory, motor control, language, action planning, problem solving, decision making, reasoning, and creativity.

The material is conceptually advanced and moderately technical. It is aimed at advanced undergraduate students, as well as graduate students from psychology, neurobiology, computer science, and other cognitive sciences.

The recommended textbook is *Computing the Mind: How the Mind Really Works* (Oxford University Press, 2008). Additional readings (a zipped collection of PDFs) are available on the course Blackboard site.

How to use this syllabus

- For each week, there's a list of references. All these references, and then some, are also listed at the end of the syllabus, alphabetically by first author.
- An alphabetical roster of key ideas and topics begins on p. 15.
- Don't panic! Only some of the 100 or so papers listed in the references are required reading (these are clearly marked).

Important dates

- First prelim: Thursday, 2/23, in class.
- Second prelim: Thursday, 3/30, in class.
- Final exam: Monday, May 15, 7:30–9:00pm (set by the college).

*Psych 3140/6140. Spring semester, Tuesdays and Thursdays.

Week 1 (1/24; 1/26): introduction and overview

1.1 [No meeting — the semester hasn't started yet.]

1.2 The subject matter of psychology. The fundamentality of computation [30].

Examples:

- perception — *lightness*, or estimating surface reflectance from images;
- thinking — *planning*, or estimating the actions needed to attain a goal);
- action — *motor control*, or estimating the signals that need to be sent to the muscles to execute the desired motion.

The nature of behavior:

- Dewey on the "reflex arc" concept [26];
- Thurstone on the "stimulus-response" fallacy [96].

A quick overview of computation:

- dynamical systems [52, sections 1,2];
- Turing Machines [5].

Primary readings

[30] S. Edelman. *Computing the mind: how the mind really works*. Oxford University Press, New York, NY, 2008, chapters 1,2.

[26] J. Dewey. The reflex arc concept in psychology. *Psychological Review*, 3:357–370, 1896.

[96] L. L. Thurstone. The stimulus-response fallacy in psychology. *Psychological Review*, 30:354–369, 1923.

Other readings

[101] E. P. Wigner. The unreasonable effectiveness of mathematics in the natural sciences. *Comm. Pure Appl. Math.*, XIII:1–14, 1960.

[5] D. Barker-Plummer. Turing machines. In E. N. Zalta, editor, *The Stanford Encyclopedia of Philosophy*. 2007. Available online at http://plato.stanford.edu/archives/win2007/entries/turing-machine/.

[52] S. Hotton and J. Yoshimi. Extending dynamical systems theory to model embodied cognition. *Cognitive Science*, 35:444–479, 2010, sections 1,2.

Week 2 (1/31; 2/2): methodology

2.1 The blind men and the elephant.

Four case studies:

- A theory [2];
- A single-cell electrophysiology study [86].
- An imaging study [48];
- A computational hack [68] (to be revisited later in the semester; take note!).

2.2 Open your eyes!

- The general methodology: the Marr-Poggio program for neurosciences [30, ch.4].
- A worked-out example: sound localization in the barn owl [59].

Primary readings

[30] S. Edelman. *Computing the mind: how the mind really works*. Oxford University Press, New York, NY, 2008, chapters 3,4.

[67] D. Marr and T. Poggio. From understanding computation to understanding neural circuitry. *Neurosciences Res. Prog. Bull.*, 15:470–488, 1977.

[59] P. X. Joris, P. H. Smith, and T. C. T. Yin. Coincidence detection in the auditory system: 50 years after Jeffress. *Neuron*, 21:1235–1238, 1998.

Other readings

[2] J. R. Anderson. ACT: A simple theory of complex cognition. *American Psychologist*, 51:355–365, 1996.

[68] V. Mnih, K. Kavukcuoglu, D. Silver, A. A. Rusu, J. Veness, M. G. Bellemare, A. Graves, M. Riedmiller, A. K. Fidjeland, G. Ostrovski, S. Petersen, C. Beattie, A. Sadik, I. Antonoglou, H. King, D. Kumaran, D. Wierstra, S. Legg, and D. Hassabis. Human-level control through deep reinforcement learning. *Nature*, 518:529–533, 2015.

[35] S. Edelman. The minority report: some common assumptions to reconsider in the modeling of the brain and behavior. *Journal of Experimental and Theoretical Artificial Intelligence*, 27:1–26, 2015. doi: 10.1080/0952813X.2015.1042534.

[32] S. Edelman. Vision, reanimated and reimagined. *Perception*, 41:1116–1127, 2012. Special issue on Marr's *Vision*.

[48] J. V. Haxby, M. I. Gobbini, M. L. Furey, A. Ishai, J. L. Schouten, and P. Pietrini. Distributed and overlapping representations of faces and objects in ventral temporal cortex. *Science*, 293:2425–2430, 2001.

[86] C. D. Salzman, K. H. Britten, and W. T. Newsome. Cortical microstimulation influences perceptual judgements of motion direction. *Nature*, 346:174–177, 1990.

Week 3 (2/7; 2/9): universal tools, I: measurement and representation

3.1 Perceptual measurement [30, ch.5]. Channel coding and hyperacuity [91]. The fundamental uncertainty [88].

3.2 Representation spaces [30, ch.5]. The face space [41, 58].

Primary readings

[30] S. Edelman. *Computing the mind: how the mind really works*. Oxford University Press, New York, NY, 2008, chapter 5.

[58] F. Jiang, V. Blanz, and A. J. O'Toole. Probing the visual representation of faces with adaptation: A view from the other side of the mean. *Psychological Science*, 17:493–500, 2006.

Other readings

[1] R. A. Altes. Ubiquity of hyperacuity. *J. Acoust. Soc. Am.*, 85:943–952, 1988.

[91] H. P. Snippe and J. J. Koenderink. Discrimination thresholds for channel-coded systems. *Biological Cybernetics*, 66:543–551, 1992.

[41] S. Eifuku, W. C. De Souza, R. Tamura, H. Nishijo, and T. Ono. Neuronal correlates of face identification in the monkey anterior temporal cortical areas. *J. of Neurophysiology*, 91:358–371, 2004.

[88] R. Shahbazi, R. Raizada, and S. Edelman. Similarity, kernels, and the fundamental constraints on cognition. *Journal of Mathematical Psychology*, 70:21–34, 2016.

Week 4 (2/14; 2/16): February break; universal tools, II: probability and the Ace of Bayes

4.1 A probabilistic formulation of cognition [18]. The Bayesian framework [46].

4.2 The Bayesian approach, applied to lightness perception [10].

Primary readings

[18] N. Chater, J. B. Tenenbaum, and A. Yuille. Probabilistic models of cognition: Conceptual foundations. *Trends in Cognitive Sciences*, 10:287–291, 2006.

[46] T. L. Griffiths and A. Yuille. Technical introduction: A primer on probabilistic inference. *Trends in Cognitive Sciences*, 10, 2006. Supplementary material. DOI: 10.1016/j.tics.2006.05.007.

[30] S. Edelman. *Computing the mind: how the mind really works*. Oxford University Press, New York, NY, 2008, Appendix A.

[30] S. Edelman. *Computing the mind: how the mind really works*. Oxford University Press, New York, NY, 2008, chapter 5.

[53] A. Hurlbert. Colour constancy. *Current Biology*, 17:R906–R907, 2007.

[10] D. H. Brainard and W. T. Freeman. Bayesian color constancy. *J. Opt. Soc. Am. A*, 14:1393–1411, 1997.

Other readings

[100] M. Welling. Are ML and statistics complementary?, 2015. Roundtable discussion at the 6th IMS-ISBA meeting on Data Science in the next 50 years.

[44] Z. Ghahramani. Probabilistic machine learning and artificial intelligence. *Nature*, 521:452–459, 2015.

[61] D. Knill and W. Richards, editors. *Perception as Bayesian Inference*. Cambridge University Press, Cambridge, 1996.

[42] T. Evgeniou, M. Pontil, and T. Poggio. Regularization networks and support vector machines. *Advances in Computational Mathematics*, 13:1–50, 2000.

[89] R. N. Shepard. Toward a universal law of generalization for psychological science. *Science*, 237: 1317–1323, 1987.

[28] S. Edelman. Representation is representation of similarity. *Behavioral and Brain Sciences*, 21:449–498, 1998.

[94] J. B. Tenenbaum and T. L. Griffiths. Generalization, similarity, and Bayesian inference. *Behavioral and Brain Sciences*, 24:629–641, 2001.

Week 5 (2/21; 2/23): Bayes, applied; 1st prelim

5.1 [no meeting — February break]

5.2 **First prelim.**

Primary readings

- None [take a break, or read ahead, or read the short, non-technical stuff listed below]

Other readings

[33] S. Edelman. *The Happiness of Pursuit*. Basic Books, New York, NY, 2012

[34] S. Edelman. *Beginnings*. BookBaby, 2014. Available electronically via Amazon (click here) and iTunes (here)

Week 6 (2/28; 3/2): universal tools, III: similarity; learning and generalization

6.1 Learning, similarity, and generalization [30, ch.5]. Shepard's Law [89].

6.2 Learning from examples as function approximation. Object recognition [80, 12, 22]. The challenges of function approximation. Ill-posedness and regularization [6].

Primary readings

[81] T. Poggio, M. Fahle, and S. Edelman. Fast perceptual learning in visual hyperacuity. *Science*, 256: 1018–1021, 1992.

[80] T. Poggio and S. Edelman. A network that learns to recognize three-dimensional objects. *Nature*, 343:263–266, 1990.

[12] H. H. Bülthoff and S. Edelman. Psychophysical support for a 2-D view interpolation theory of object recognition. *Proceedings of the National Academy of Science*, 89:60–64, 1992.

[22] F. Cutzu and S. Edelman. Faithful representation of similarities among three-dimensional shapes in human vision. *Proceedings of the National Academy of Science*, 93:12046–12050, 1996.

[39] S. Edelman and R. Shahbazi. Renewing the respect for similarity. *Frontiers in Computational Neuroscience*, 6:45, 2012.

Other readings

[6] M. Bertero, T. Poggio, and V. Torre. Ill-posed problems in early vision. *Proceedings of the IEEE*, 76: 869–889, 1988.

[60] D. Kersten, P. Mamassian, and A. Yuille. Object perception as Bayesian inference. *Annual Review of Psychology*, 55:271–304, 2004.

[28] S. Edelman. Representation is representation of similarity. *Behavioral and Brain Sciences*, 21:449–498, 1998.

[36] S. Edelman. Varieties of perceptual truth and their possible evolutionary roots. *Psychonomic Bulletin and Review*, 22:1519–1522, 2015. doi: 10.3758/s13423-014-0741-z

Week 7 (3/7; 3/9): universal tools, IV: dimensionality; memory

7.1 Managing dimensionality. Dimensionality reduction by random projections [29, 39].

7.2 Managing the storage and retrieval of examples. Associative memory [30, ch.6]. Locality-sensitive hashing [4, 39].

Primary readings

[30] S. Edelman. *Computing the mind: how the mind really works*. Oxford University Press, New York, NY, 2008, chapter 6.

[39] S. Edelman and R. Shahbazi. Renewing the respect for similarity. *Frontiers in Computational Neuroscience*, 6:45, 2012.

[4] A. Andoni and P. Indyk. Near-optimal hashing algorithms for approximate nearest neighbor in high dimensions. *Communications of the ACM*, 51:117–122, 2008.

Other readings

[29] S. Edelman. *Representation and recognition in vision*. MIT Press, Cambridge, MA, 1999.

Week 8 (3/14; 3/16): actions and consequences

8.1 Managing action. The basics of motor control [102]. Bayesian motor decision making [62].

8.2 Action and reward. Reinforcement learning (RL) [104] and its relationship to Bayes. Hierarchical RL [8, 9].

Primary readings

[102] S. P. Wise and R. Shadmehr. Motor control. In V. S. Ramachandran, editor, *Encyclopedia of the Human Brain*, volume 3, pages 137–157. Academic Press, San Diego, CA, 2002.

[62] K. P. Körding and D. M. Wolpert. Bayesian decision theory in sensorimotor control. *Trends in Cognitive Sciences*, 10:319–326, 2006.

[104] D. M. Wolpert, J. Diedrichsen, and J. R. Flanagan. Principles of sensorimotor learning. *Nature Reviews Neuroscience*, 12:739–751, 2011.

[8] M. M. Botvinick. Hierarchical models of behavior and prefrontal function. *Trends in Cognitive Sciences*, 12:201–208, 2008.

[9] M. M. Botvinick, Y. Niv, and A. C. Barto. Hierarchically organized behavior and its neural foundations: A reinforcement learning perspective. *Cognition*, 113:262–280, 2009.

Other readings

[30] S. Edelman. *Computing the mind: how the mind really works*. Oxford University Press, New York, NY, 2008, chapter 6.

[103] W. Woergoetter and B. Porr. Reinforcement learning. *Scholarpedia*, 3(3):1448, 2007.

[23] P. Dayan. How to set the switches on this thing. *Current Opinion in Neurobiology*, 22:1068–1074, 2012.

[16] N. Chater. Rational and mechanistic perspectives on reinforcement learning. *Cognition*, 113:350–364, 2009.

[44] Z. Ghahramani. Probabilistic machine learning and artificial intelligence. *Nature*, 521:452–459, 2015.

Week 9 (3/21; 3/23): higher cognition

9.1 Graphical models (Bayesian Networks) and reasoning [30, ch.8].

9.2 Induction [95]. General intelligence and IQ [70].

Primary readings

[30] S. Edelman. *Computing the mind: how the mind really works*. Oxford University Press, New York, NY, 2008, chapter 8.

[95] J. B. Tenenbaum, C. Kemp, T. L. Griffiths, and N. D. Goodman. How to grow a mind: statistics, structure, and abstraction. *Science*, 331:1279–1285, 2011.

[70] R. E. Nisbett, J. Aronson, C. Blair, W. Dickens, J. Flynn, D. F. Halpern, and E. Turkheimer. Intelligence: new findings and theoretical developments. *American Psychologist*, 2012.

Other readings

[75] J. Pearl. Structural counterfactuals: A brief introduction. *Cognitive Science*, 37:977–985, 2013.

Week 10 (3/28; 3/30): higher cognition (cont.); 2nd prelim

10.1 Problem solving [30, ch.8]. Analogy [50] and creativity [49, 7].

10.2 **Second prelim.**

Primary readings

[30] S. Edelman. *Computing the mind: how the mind really works*. Oxford University Press, New York, NY, 2008, chapter 8.

[49] D. R. Hofstadter. Variations on a theme as the crux of creativity. In *Metamagical Themas*, chapter 12, pages 232–259. Viking, Harmondsworth, England, 1985.

[50] D. R. Hofstadter. Analogy as the core of cognition. In D. Gentner, K. J. Holyoak, and B. N. Kokinov, editors, *The Analogical Mind: Perspectives from Cognitive Science*, pages 499–538. MIT Press, Cambridge MA, 2001.

Other readings

[7] M. A. Boden. Precis of "the creative mind: Myths and mechanisms". *Behavioral and Brain Sciences*, 17:519–570, 1994. Available online at http://psych.utoronto.ca/users/reingold/courses/ai/cache/bbs.boden.html.

Week 11 (4/11; 4/13): intro to neural computation; neurons, I

11.1 Introduction to neural computation. Brains and neurons. Cortical receptive fields (RFs), maps, and hierarchies [30, ch.2,3].

11.2 What do neurons do?

Projection:

- neurons as readout devices [14];
- neurons and projection pursuit [20].
- neurons, random projections, and similarity [39, 65].

Kernels:

- landmarks in representation spaces, similarity, and kernels [88].

Primary readings

[30] S. Edelman. *Computing the mind: how the mind really works*. Oxford University Press, New York, NY, 2008, chapters 2,3.

[14] G. Buzsáki. Neural syntax: cell assemblies, synapsembles, and readers. *Neuron*, 68:362–385, 2010.

[20] L. N. Cooper and M. F. Bear. The BCM theory of synapse modification at 30: interaction of theory with experiment. *Nature Reviews Neuroscience*, 13:798–810, 2012.

Other readings

[97] W. Truccolo, L. R Hochberg, and J. P Donoghue. Collective dynamics in human and monkey sensorimotor cortex: predicting single neuron spikes. *Nature Neuroscience*, 13:105–111, 2009. doi: 10.1038/nn.2455.

[65] T. P. Lillicrap, D. Cownden, C. J. Akerman, and D. B. Tweed. Multilayer controllers can learn from random feedback weights. In *Proc. Symp. on Translational and Computational Motor Control (TCMC)*, pages 83–84, 2013. Satellite to the annual Society for Neuroscience meeting.

[88] R. Shahbazi, R. Raizada, and S. Edelman. Similarity, kernels, and the fundamental constraints on cognition. *Journal of Mathematical Psychology*, 70:21–34, 2016.

Week 12 (4/18; 4/20): neurons, II

12.1 What do neurons do?

Time-dependent dynamic learning:

- spike timing dependent plasticity (STDP), and Hebbian learning (see the Scholarpedia article [90]);
- history-dependent learning with the BCM rule [20];
- learning dynamic embodied control modes [25].

12.2 What do neurons do?

Population dynamics:

- neural trajectories and classification [13].

Primary readings

[20] L. N. Cooper and M. F. Bear. The BCM theory of synapse modification at 30: interaction of theory with experiment. *Nature Reviews Neuroscience*, 13:798–810, 2012.

[25] R. Der and G. Martius. Novel plasticity rule can explain the development of sensorimotor intelligence. *Proceedings of the National Academy of Science*, pages E6224–E6232, 2015.

[90] J. Sjöström and W. Gerstner. Spike-timing dependent plasticity. *Scholarpedia*, 5(2):1362, 2010.

[13] D. V. Buonomano and W. Maass. State-dependent computations: spatiotemporal processing in cortical networks. *Nature Reviews Neuroscience*, 10:113–125, 2009.

Other readings

[55] E. M. Izhikevich and N. S. Desai. Relating STDP to BCM. *Neural Computation*, 15:1511–1523, 2003

[24] S. Deneve. Bayesian spiking neurons I: Inference. *Neural Computation*, 20:91–117, 2008.

[54] E. M. Izhikevich. Solving the distal reward problem through linkage of STDP and dopamine signaling. *Cerebral Cortex*, 17:2443–2452, 2007.

[72] B. J. Fischer J. L. Peña. Owl's behavior and neural representation predicted by Bayesian inference. *Nature Neuroscience*, 14:1061–1067, 2011.

Week 13 (4/25; 4/27): neurons, III; brains

13.1 What do neurons do?

Population dynamics (cont.):

- ongoing dynamics and chaotic itinerancy [84].

13.2 What does the brain do? Prediction [71, 19].

Primary readings

[84] M. I. Rabinovich, A. N. Simmons, and P. Varona. Dynamical bridge between brain and mind. *Trends in Cognitive Sciences*, 19:453–461, 2015.

[43] K. J. Friston. The free-energy principle: a unified brain theory? *Nature Neuroscience*, 11:127–138, 2010.

[71] H.-J. Park and K. J. Friston. Structural and functional brain networks: from connections to cognition. *Science*, 342:1238411, 2013. doi: 10.1126/science.1238411.

Other readings

[19] A. Clark. Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences*, 36:181–204, 2013.

[11] J. Bruineberg, J. Kiverstein, and E. Rietveld. The anticipating brain is not a scientist: the free-energy principle from an ecological-enactive perspective. *Synthese*, 2016. doi: 10.1007/s11229-016-1239-1.

Week 14 (5/2; 5/4): advanced topics

14.1 Bayes and the real world [87]

14.2 Real behavior and real brains [68, 35]

Primary readings

[87] A. N. Sanborn and N. Chater. Bayesian brains without probabilities. *Trends in Cognitive Sciences*, 20:883–893, 2016. doi: 10.1016/j.tics.2016.10.003.

[68] V. Mnih, K. Kavukcuoglu, D. Silver, A. A. Rusu, J. Veness, M. G. Bellemare, A. Graves, M. Riedmiller, A. K. Fidjeland, G. Ostrovski, S. Petersen, C. Beattie, A. Sadik, I. Antonoglou, H. King, D. Kumaran, D. Wierstra, S. Legg, and D. Hassabis. Human-level control through deep reinforcement learning. *Nature*, 518:529–533, 2015.

[35] S. Edelman. The minority report: some common assumptions to reconsider in the modeling of the brain and behavior. *Journal of Experimental and Theoretical Artificial Intelligence*, 27:1–26, 2015. doi: 10.1080/0952813X.2015.1042534.

Other readings

[64] Y. LeCun, Y. Bengio, and G. Hinton. Deep learning. *Nature*, 521:436–444, 2015.

[63] N. Kriegeskorte. Deep neural networks: a new framework for modelling biological vision and brain information processing. *Annual Reviews of Vision Science*, 2017. In press.

Week 15 (5/9): wrap-up

15.1 Summary [30, ch.11]. Being human.

15.2 [the semester's over!]

Computational concepts, principles, and methods: an abecedary[1]

ANALOGY — a comparison that involves a structure mapping between complex entities, situations, or domains [50, 51]. Analogy is central to general cognitive function (general fluid intelligence, often referred to as *gF* or IQ [92]) and has been hypothesized to underlie structural LEARNING in vision and in language.

BAYES THEOREM — a direct consequence of the definition of conditional probability; the basis for the so-called rational theories of cognition. The Bayes Theorem prescribes a way of integrating prior beliefs with new data, in a way that proves useful in all domains, from perception, through thinking and language, to motor control [61, 105, 38, 94, 77, 62].

CHANNEL CODING — measuring a stimulus with a set of graded, overlapping filters (receptive fields or "channels") supports a high degree of resolution, or hyperacuity, that cannot be achieved through dense sampling by pointlike filters [91]. This principle is at work throughout cognition [30, p.90].

DIMENSIONALITY REDUCTION — A high-dimensional perceptual measurement space is advantageous because it may capture more of the useful structure of the problems that a cognitive system needs to deal with, such as categorization. LEARNING by "tiling" the representation space with examples is, however, infeasible in a high-dimensional space, because the number of required examples grows exponentially with dimensionality [37]. This necessitates dimensionality reduction prior to learning, which, moreover, needs to be done so as to lose as little as possible of the useful information.

EMBODIMENT AND SITUATEDNESS — EVOLUTION fine-tunes the computations carried out by natural cognitive systems to the mechanics of the bodies that they control and the ecological niche in which they are situated [3].

FUNCTION APPROXIMATION — LEARNING from examples and generalization to new queries is equivalent to function approximation, a problem in which the values of an unknown function are given at a number of points in its domain and are used to form an estimate that can then support generalization [79].

GRAPHICAL MODELS — The relationships among a set of variables of interest to a cognitive system can be conveniently represented in the form of a directed graph, in which the vertices stand for variables (of which some may be observable and others hidden, corresponding to the properties of the world that need to be inferred from sensory data) and the edges — for probabilistic dependencies between pairs of variables [73]. One type of such model is the BAYES Network [76]. Graphical models map naturally onto the architecture of the brain [66].

HOLISM — The PATTERN of causal dependencies in a system of knowledge about the natural world is such that any two items may be potentially interdependent; in this sense, rich cognitive representations are holistic [83]. This property of world knowledge gives rise to serious algorithmic challenges in truth

[1] The number of important principles in cognitive science is somewhat larger than the number of letters in the alphabet. To make up for this inconvenience, I have highlighted in SMALL CAPITALS every concept of interest, whether or not it has its own entry.

maintenance systems, where facts newly acquired through LEARNING can potentially interact with, and cause the revision of, the entire knowledge base.

ILL-POSEDNESS — Problems arising in perception, thinking, and action planning are typically ill-posed in the sense that they do not possess a unique solution (e.g., 6). FUNCTION APPROXIMATION, which is central to LEARNING, is ill-posed because an infinite number of mappings may be consistent with a given set of input-output pairs, and so is probability density estimation. Such problems can be made well-posed by REGULARIZATION.

JOINT PROBABILITY — The most that can be learned about the world by observing or tracking a set of variables of interest is an approximation to their joint probability density (note that the problem of probability estimation is ILL-POSED). To learn the causal dependencies among the variables, one must go beyond mere observation and intervene on variables of interest [74].

KERNELS — A family of mathematical methods that arise from measurements of SIMILARITY of two vectors and that are widely applicable in modeling cognition [56, 57, 88]. Formally, a positive definite kernel is a function of two arguments that represents an inner product (dot product) in some feature space.

LEARNING AND LEARNABILITY — Most of the detailed knowledge about how the world works that animals with advanced cognitive systems need to master cannot be "squeezed" through the genomic bottleneck and must therefore be learned from experience. The field of machine learning has amassed a wealth of insights into the computational nature of this process, including constraints and limitations on learning and learnability (e.g., 98, 99).

MINUMUM DESCRIPTION LENGTH (MDL) PRINCIPLE — The fundamental principle of the computational theory of learning, due to Solomonoff [93], is that LEARNING is learning of regularities. It is derived from the observation that learning is only useful insofar as it supports generalization and that generalization is only possible if regularities are discovered in the observed data. A modern operationalization of this idea is the Minimum Description Length Principle of Rissanen [85], according to which regularities in the data are best captured by a representation that minimizes the sum of the description lengths of the code and of the training data under that code [47]. A related principle is that of SIMPLICITY [17].

NAVIGATION — Finding a route through a representation space, subject to certain CONSTRAINTS, is a paradigm for all sequential behaviors. Thus, in foraging, for instance, the SEARCH space represents the terrain in which the animal is situated; in planning, it may be a graph representing the space of possible solutions to the problem at hand; in language production, the graph would be a representation of the speaker's knowledge of language [31].

OPTIMIZATION — A wide range of tasks in cognition, including perception, thinking (e.g., problem solving and decision making), and motor control reduce to SEARCHING a space of possible solutions for an optimal one [30]. Optimality in this context is imposed by various CONSTRAINTS, which may stem from the nature of the problem, from implementational considerations, from EVOLUTIONARY pressure, or from general requirements of tractability and uniqueness (as in REGULARIZATION).

PREDICTION — A true understanding of the world (e.g., one that takes the form of a CAUSAL PROBABILISTIC model) should allow the cognitive system to exercise FORESIGHT: to predict impending events and the consequences of its own actions [30]. Such capacity for prediction turns out to be a very general explanatory principle in cognition [19], which can be linked to other general principles, such as BAYESIAN probability theory.

QUANTUM PROBABILITY — Animal behaviors that involve probabilistic assessment of cues and outcomes are often strongly context- and order-dependent. Understanding such behaviors may require positing individual states that are superpositions (i.e., are impossible to associate with specific values), as well as composite systems that are entangled (i.e., that cannot be decomposed into their subsystems). The relevant theories are best expressed in terms of quantum probability postulates [82].

REGULARIZATION — A problem that is formally ILL-POSED in that it has no unique solution can be turned into a well-posed one by imposing external CONSTRAINTS on the solution space. One class of such constraints, which has a profound grounding in LEARNING theory, is regularization through smoothing, which is related to BAYESIAN probability, to statistical learning theory, and to the MAXIMUM LIKELIHOOD idea [42].

SIMILARITY — the most important ultimate use to which sensory data could be put involves estimating the similarity between two stimuli, which constitutes the only principled basis for GENERALIZATION of response from one stimulus to another, and therefore for any non-trivial LEARNING from experience [89, 28, 39].

TUNING — Neurons in animal brains, and neuron-like units in artificial distributed cognitive systems, are typically tuned to various features of the perceptual world, of motor behavior, or of the animal's internal representational states [30, ch.3]. Graded, shallow tuning, with a high degree of overlap between the profiles of adjacent units, is behind perceptual filters or CHANNELS that support hyperacuity. Because a tuned unit effectively represents the SIMILARITY between the actual and optimal stimuli, graded tuning also underlies VERIDICALITY [29].

UNCERTAINTY — "The percept is always a wager. Thus uncertainty enters at two levels, not merely one: the configuration may or may not indicate an object, and the cue may or may not be utilized at its true indicative value" [45]. The fundamental uncertainty in dealing with the world is the central motivation for the use of PROBABILISTIC representations and processes by cognitive systems.

VERIDICALITY — Perceptual representations based on CHANNEL CODING are provably veridical in that they faithfully reflect the SIMILARITY relationships of the represented items, such as visual objects [29].

WEIGHT LEARNING — In computational systems composed of simple, neuron-like elements, LEARNING typically proceeds by adjusting the weights of the synaptic connections, although the threshold of the nonlinear transfer function that is part of the standard "formal neuron" can also be adjusted [21]. The rule for experience-based weight adjustment proposed by Hebb in 1949, according to which "neurons that fire together, wire together," has now been recast and widely accepted as spike timing dependent plasticity, or STDP [15].

MA**X**IMUM LIKELIHOOD ESTIMATION (MLE) — According to the MLE principle, the parameters of a PROBABILISTIC model that is intended to reproduce a set of observations should be tuned so as to make the actual observed data set most likely [69].

S**Y**NTAX — The system of PATTERNS and CONSTRAINTS that governs the composition of utterances in a natural language [78, 40].

DOB**Z**HANSKY — "Nothing in biology makes sense except in the light of EVOLUTION" [27].

References

[1] R. A. Altes. Ubiquity of hyperacuity. *J. Acoust. Soc. Am.*, 85:943–952, 1988.

[2] J. R. Anderson. ACT: A simple theory of complex cognition. *American Psychologist*, 51:355–365, 1996.

[3] M. L. Anderson. Embodied cognition: A field guide. *Artificial Intelligence*, 149:91–130, 2003.

[4] A. Andoni and P. Indyk. Near-optimal hashing algorithms for approximate nearest neighbor in high dimensions. *Communications of the ACM*, 51:117–122, 2008.

[5] D. Barker-Plummer. Turing machines. In E. N. Zalta, editor, *The Stanford Encyclopedia of Philosophy*. 2007. Available online at http://plato.stanford.edu/archives/win2007/entries/turing-machine/.

[6] M. Bertero, T. Poggio, and V. Torre. Ill-posed problems in early vision. *Proceedings of the IEEE*, 76: 869–889, 1988.

[7] M. A. Boden. Precis of "the creative mind: Myths and mechanisms". *Behavioral and Brain Sciences*, 17:519–570, 1994. Available online at http://psych.utoronto.ca/users/reingold/courses/ai/cache/bbs.boden.html.

[8] M. M. Botvinick. Hierarchical models of behavior and prefrontal function. *Trends in Cognitive Sciences*, 12:201–208, 2008.

[9] M. M. Botvinick, Y. Niv, and A. C. Barto. Hierarchically organized behavior and its neural foundations: A reinforcement learning perspective. *Cognition*, 113:262–280, 2009.

[10] D. H. Brainard and W. T. Freeman. Bayesian color constancy. *J. Opt. Soc. Am. A*, 14:1393–1411, 1997.

[11] J. Bruineberg, J. Kiverstein, and E. Rietveld. The anticipating brain is not a scientist: the free-energy principle from an ecological-enactive perspective. *Synthese*, 2016. doi: 10.1007/s11229-016-1239-1.

[12] H. H. Bülthoff and S. Edelman. Psychophysical support for a 2-D view interpolation theory of object recognition. *Proceedings of the National Academy of Science*, 89:60–64, 1992.

[13] D. V. Buonomano and W. Maass. State-dependent computations: spatiotemporal processing in cortical networks. *Nature Reviews Neuroscience*, 10:113–125, 2009.

[14] G. Buzsáki. Neural syntax: cell assemblies, synapsembles, and readers. *Neuron*, 68:362–385, 2010.

[15] N. Caporale and Y. Dan. Spike timing-dependent plasticity: A Hebbian learning rule. *Annual Review of Neuroscience*, 31:25–46, 2008.

[16] N. Chater. Rational and mechanistic perspectives on reinforcement learning. *Cognition*, 113:350–364, 2009.

[17] N. Chater and P. Vitányi. Simplicity: a unifying principle in cognitive science? *Trends in Cognitive Sciences*, 7:19–22, 2003.

[18] N. Chater, J. B. Tenenbaum, and A. Yuille. Probabilistic models of cognition: Conceptual foundations. *Trends in Cognitive Sciences*, 10:287–291, 2006.

[19] A. Clark. Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences*, 36:181–204, 2013.

[20] L. N. Cooper and M. F. Bear. The BCM theory of synapse modification at 30: interaction of theory with experiment. *Nature Reviews Neuroscience*, 13:798–810, 2012.

[21] L. N. Cooper, N. Intrator, B. S. Blais, and H. Z. Shouval. *Theory of Cortical Plasticity*. World Scientific, Singapore, 2004.

[22] F. Cutzu and S. Edelman. Faithful representation of similarities among three-dimensional shapes in human vision. *Proceedings of the National Academy of Science*, 93:12046–12050, 1996.

[23] P. Dayan. How to set the switches on this thing. *Current Opinion in Neurobiology*, 22:1068–1074, 2012.

[24] S. Deneve. Bayesian spiking neurons I: Inference. *Neural Computation*, 20:91–117, 2008.

[25] R. Der and G. Martius. Novel plasticity rule can explain the development of sensorimotor intelligence. *Proceedings of the National Academy of Science*, pages E6224–E6232, 2015.

[26] J. Dewey. The reflex arc concept in psychology. *Psychological Review*, 3:357–370, 1896.

[27] T. Dobzhansky. Nothing in biology makes sense except in the light of evolution. *The American Biology Teacher*, 35:125–129, 1973.

[28] S. Edelman. Representation is representation of similarity. *Behavioral and Brain Sciences*, 21:449–498, 1998.

[29] S. Edelman. *Representation and recognition in vision*. MIT Press, Cambridge, MA, 1999.

[30] S. Edelman. *Computing the mind: how the mind really works*. Oxford University Press, New York, NY, 2008.

[31] S. Edelman. On look-ahead in language: navigating a multitude of familiar paths. In M. Bar, editor, *Prediction in the Brain*, chapter 14, pages 170–189. Oxford University Press, New York, 2011.

[32] S. Edelman. Vision, reanimated and reimagined. *Perception*, 41:1116–1127, 2012. Special issue on Marr's *Vision*.

[33] S. Edelman. *The Happiness of Pursuit*. Basic Books, New York, NY, 2012.

[34] S. Edelman. *Beginnings*. BookBaby, 2014. Available electronically via Amazon (click here) and iTunes (here).

[35] S. Edelman. The minority report: some common assumptions to reconsider in the modeling of the brain and behavior. *Journal of Experimental and Theoretical Artificial Intelligence*, 27:1–26, 2015. doi: 10.1080/0952813X.2015.1042534.

[36] S. Edelman. Varieties of perceptual truth and their possible evolutionary roots. *Psychonomic Bulletin and Review*, 22:1519–1522, 2015. doi: 10.3758/s13423-014-0741-z.

[37] S. Edelman and N. Intrator. Learning as extraction of low-dimensional representations. In D. Medin, R. Goldstone, and P. Schyns, editors, *Mechanisms of Perceptual Learning*, pages 353–380. Academic Press, 1997.

[38] S. Edelman and R. Shahbazi. Survival in a world of probable objects. *Behavioral and Brain Sciences*, 34:197–198, 2011. A commentary on *Bayesian Fundamentalism or Enlightenment? On the Explanatory Status and Theoretical Contributions of Bayesian Models of Cognition* by Jones & Love.

[39] S. Edelman and R. Shahbazi. Renewing the respect for similarity. *Frontiers in Computational Neuroscience*, 6:45, 2012.

[40] S. Edelman and H. R. Waterfall. Behavioral and computational aspects of language and its acquisition. *Physics of Life Reviews*, 4:253–277, 2007.

[41] S. Eifuku, W. C. De Souza, R. Tamura, H. Nishijo, and T. Ono. Neuronal correlates of face identification in the monkey anterior temporal cortical areas. *J. of Neurophysiology*, 91:358–371, 2004.

[42] T. Evgeniou, M. Pontil, and T. Poggio. Regularization networks and support vector machines. *Advances in Computational Mathematics*, 13:1–50, 2000.

[43] K. J. Friston. The free-energy principle: a unified brain theory? *Nature Neuroscience*, 11:127–138, 2010.

[44] Z. Ghahramani. Probabilistic machine learning and artificial intelligence. *Nature*, 521:452–459, 2015.

[45] J. J. Gibson. Survival in a world of probable objects. *Contemporary Psychology*, 2:33–35, 1957.

[46] T. L. Griffiths and A. Yuille. Technical introduction: A primer on probabilistic inference. *Trends in Cognitive Sciences*, 10, 2006. Supplementary material. DOI: 10.1016/j.tics.2006.05.007.

[47] P. Grunwald. Introducing the minimum description length principle. In *Advances in minimum description length: theory and applications*, pages 3–22. MIT, Cambridge, MA, 2005.

[48] J. V. Haxby, M. I. Gobbini, M. L. Furey, A. Ishai, J. L. Schouten, and P. Pietrini. Distributed and overlapping representations of faces and objects in ventral temporal cortex. *Science*, 293:2425–2430, 2001.

[49] D. R. Hofstadter. Variations on a theme as the crux of creativity. In *Metamagical Themas*, chapter 12, pages 232–259. Viking, Harmondsworth, England, 1985.

[50] D. R. Hofstadter. Analogy as the core of cognition. In D. Gentner, K. J. Holyoak, and B. N. Kokinov, editors, *The Analogical Mind: Perspectives from Cognitive Science*, pages 499–538. MIT Press, Cambridge MA, 2001.

[51] D. R. Hofstadter and M. Mitchell. The Copycat project: a model of mental fluidity and analogy-making. In D. R. Hofstadter, editor, *Fluid Concepts and Creative Analogies*, chapter 5, pages 205–265. Basic Books, NY, 1995.

[52] S. Hotton and J. Yoshimi. Extending dynamical systems theory to model embodied cognition. *Cognitive Science*, 35:444–479, 2010.

[53] A. Hurlbert. Colour constancy. *Current Biology*, 17:R906–R907, 2007.

[54] E. M. Izhikevich. Solving the distal reward problem through linkage of STDP and dopamine signaling. *Cerebral Cortex*, 17:2443–2452, 2007.

[55] E. M. Izhikevich and N. S. Desai. Relating STDP to BCM. *Neural Computation*, 15:1511–1523, 2003.

[56] F. Jäkel, B. Schölkopf, and F. A. Wichmann. Generalization and similarity in exemplar models of categorization: Insights from machine learning. *Psychonomic Bulletin & Review*, 15:256–271, 2008.

[57] F. Jäkel, B. Schölkopf, and F. A. Wichmann. Does cognitive science need kernels? *Trends in Cognitive Sciences*, 13:381–388, 2009.

[58] F. Jiang, V. Blanz, and A. J. O'Toole. Probing the visual representation of faces with adaptation: A view from the other side of the mean. *Psychological Science*, 17:493–500, 2006.

[59] P. X. Joris, P. H. Smith, and T. C. T. Yin. Coincidence detection in the auditory system: 50 years after Jeffress. *Neuron*, 21:1235–1238, 1998.

[60] D. Kersten, P. Mamassian, and A. Yuille. Object perception as Bayesian inference. *Annual Review of Psychology*, 55:271–304, 2004.

[61] D. Knill and W. Richards, editors. *Perception as Bayesian Inference*. Cambridge University Press, Cambridge, 1996.

[62] K. P. Körding and D. M. Wolpert. Bayesian decision theory in sensorimotor control. *Trends in Cognitive Sciences*, 10:319–326, 2006.

[63] N. Kriegeskorte. Deep neural networks: a new framework for modelling biological vision and brain information processing. *Annual Reviews of Vision Science*, 2017. In press.

[64] Y. LeCun, Y. Bengio, and G. Hinton. Deep learning. *Nature*, 521:436–444, 2015.

[65] T. P. Lillicrap, D. Cownden, C. J. Akerman, and D. B. Tweed. Multilayer controllers can learn from random feedback weights. In *Proc. Symp. on Translational and Computational Motor Control (TCMC)*, pages 83–84, 2013. Satellite to the annual Society for Neuroscience meeting.

[66] S. Litvak and S. Ullman. Cortical circuitry implementing graphical models. *Neural Computation*, 21:3010–3056, 2009.

[67] D. Marr and T. Poggio. From understanding computation to understanding neural circuitry. *Neurosciences Res. Prog. Bull.*, 15:470–488, 1977.

[68] V. Mnih, K. Kavukcuoglu, D. Silver, A. A. Rusu, J. Veness, M. G. Bellemare, A. Graves, M. Riedmiller, A. K. Fidjeland, G. Ostrovski, S. Petersen, C. Beattie, A. Sadik, I. Antonoglou, H. King, D. Kumaran, D. Wierstra, S. Legg, and D. Hassabis. Human-level control through deep reinforcement learning. *Nature*, 518:529–533, 2015.

[69] I. J. Myung. Tutorial on maximum likelihood estimation. *Journal of Mathematical Psychology*, 47: 90–100, 2003.

[70] R. E. Nisbett, J. Aronson, C. Blair, W. Dickens, J. Flynn, D. F. Halpern, and E. Turkheimer. Intelligence: new findings and theoretical developments. *American Psychologist*, 2012.

[71] H.-J. Park and K. J. Friston. Structural and functional brain networks: from connections to cognition. *Science*, 342:1238411, 2013. doi: 10.1126/science.1238411.

[72] B. J. Fischer J. L. Peña. Owl's behavior and neural representation predicted by Bayesian inference. *Nature Neuroscience*, 14:1061–1067, 2011.

[73] J. Pearl. Causal inference in statistics: A gentle introduction. In *Computing Science and Statistics: Proceedings of Interface '01*, volume 33, 2001. Avaliable online at http://ftp.cs.ucla.edu/pub/stat_ser/R289.pdf.

[74] J. Pearl. Causal inference in statistics: an overview. *Statistics Surveys*, 3:96–146, 2009.

[75] J. Pearl. Structural counterfactuals: A brief introduction. *Cognitive Science*, 37:977–985, 2013.

[76] J. Pearl and S. Russel. Bayesian networks. In M. A. Arbib, editor, *Handbook of Brain Theory and Neural Networks*, pages 157–160. MIT Press, Cambridge, MA, 2003. Available online at http://ftp.cs.ucla.edu/pub/stat_ser/R277.pdf.

[77] A. Perfors, J. B. Tenenbaum, and T. Regier. Poverty of the stimulus? A rational approach. In R. Sun and N. Miyake, editors, *Proceedings of the 28th Annual Conference of the Cognitive Science Society*, pages 663–668, Austin, TX, 2006. Cognitive Science Society.

[78] C. Phillips. Syntax. In L. Nadel, editor, *Encyclopedia of Cognitive Science*, volume 4, pages 319–329. Macmillan, London, 2003.

[79] T. Poggio. A theory of how the brain might work. *Cold Spring Harbor Symposia on Quantitative Biology*, LV:899–910, 1990.

[80] T. Poggio and S. Edelman. A network that learns to recognize three-dimensional objects. *Nature*, 343:263–266, 1990.

[81] T. Poggio, M. Fahle, and S. Edelman. Fast perceptual learning in visual hyperacuity. *Science*, 256: 1018–1021, 1992.

[82] E. M. Pothos and J. R. Busemeyer. Can quantum probability provide a new direction for cognitive modeling? *Behavioral and Brain Sciences*, 36:255–327, 2013.

[83] W. V. O. Quine. Two dogmas of Empiricism. In *From a Logical Point of View*, pages 20–46. Harvard University Press, Cambridge, 1953.

[84] M. I. Rabinovich, A. N. Simmons, and P. Varona. Dynamical bridge between brain and mind. *Trends in Cognitive Sciences*, 19:453–461, 2015.

[85] J. Rissanen. Minimum description length principle. In S. Kotz and N. L. Johnson, editors, *Encyclopedia of Statistic Sciences*, volume 5, pages 523–527. J. Wiley and Sons, 1987.

[86] C. D. Salzman, K. H. Britten, and W. T. Newsome. Cortical microstimulation influences perceptual judgements of motion direction. *Nature*, 346:174–177, 1990.

[87] A. N. Sanborn and N. Chater. Bayesian brains without probabilities. *Trends in Cognitive Sciences*, 20:883–893, 2016. doi: 10.1016/j.tics.2016.10.003.

[88] R. Shahbazi, R. Raizada, and S. Edelman. Similarity, kernels, and the fundamental constraints on cognition. *Journal of Mathematical Psychology*, 70:21–34, 2016.

[89] R. N. Shepard. Toward a universal law of generalization for psychological science. *Science*, 237: 1317–1323, 1987.

[90] J. Sjöström and W. Gerstner. Spike-timing dependent plasticity. *Scholarpedia*, 5(2):1362, 2010.

[91] H. P. Snippe and J. J. Koenderink. Discrimination thresholds for channel-coded systems. *Biological Cybernetics*, 66:543–551, 1992.

[92] R. E. Snow, P. C. Kyllonen, and B. Marshalek. The topography of ability and learning correlates. In R. J. Sternberg, editor, *Advances in the psychology of human intelligence*, volume 2, pages 47–104. Erlbaum, Hillsdale, NJ, 1984.

[93] R. J. Solomonoff. A formal theory of inductive inference, parts A and B. *Information and Control*, 7:1–22, 224–254, 1964.

[94] J. B. Tenenbaum and T. L. Griffiths. Generalization, similarity, and Bayesian inference. *Behavioral and Brain Sciences*, 24:629–641, 2001.

[95] J. B. Tenenbaum, C. Kemp, T. L. Griffiths, and N. D. Goodman. How to grow a mind: statistics, structure, and abstraction. *Science*, 331:1279–1285, 2011.

[96] L. L. Thurstone. The stimulus-response fallacy in psychology. *Psychological Review*, 30:354–369, 1923.

[97] W. Truccolo, L. R Hochberg, and J. P Donoghue. Collective dynamics in human and monkey sensorimotor cortex: predicting single neuron spikes. *Nature Neuroscience*, 13:105–111, 2009. doi: 10.1038/nn.2455.

[98] L. G. Valiant. A theory of the learnable. *Communications of the ACM*, 27:1134–1142, 1984.

[99] V. N. Vapnik. An overview of statistical learning theory. *IEEE Transactions on Neural Networks*, 10: 988–999, 1999.

[100] M. Welling. Are ML and statistics complementary?, 2015. Roundtable discussion at the 6th IMS-ISBA meeting on Data Science in the next 50 years.

[101] E. P. Wigner. The unreasonable effectiveness of mathematics in the natural sciences. *Comm. Pure Appl. Math.*, XIII:1–14, 1960.

[102] S. P. Wise and R. Shadmehr. Motor control. In V. S. Ramachandran, editor, *Encyclopedia of the Human Brain*, volume 3, pages 137–157. Academic Press, San Diego, CA, 2002.

[103] W. Woergoetter and B. Porr. Reinforcement learning. *Scholarpedia*, 3(3):1448, 2007.

[104] D. M. Wolpert, J. Diedrichsen, and J. R. Flanagan. Principles of sensorimotor learning. *Nature Reviews Neuroscience*, 12:739–751, 2011.

[105] A. Yuille and D. Kersten. Vision as Bayesian inference: analysis by synthesis? *Trends in Cognitive Sciences*, 10:301–308, 2006.