Computational [Principles of] Psychology*

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

January 4, 2016

Abstract

This course states, motivates, and offers detailed support to the observation that cognition is fundamentally a computational process [28]. Students are introduced to a number of conceptual tools for thinking about cognitive information processing, 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! Not every one of the 90 or so papers listed in the references is obligatory reading.

Important dates

- First prelim: Thursday, 2/25, in class.

- Second prelim: Thursday, 3/24, in class.

- Final exam: TBA (set by the college).

*Psych 3140/6140. Spring semester, Tuesdays and Thursdays.
Week 1 (1/26; 1/28): introduction and overview

1.1 [no meeting (the semester hasn’t started yet)]

1.2 The subject matter of psychology. The fundamentality of computation [28].

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.

A quick overview of computation:

– dynamical systems [48] sections 1,2];
– Turing Machines [5].

Primary readings


Other readings


Week 2 (2/2; 2/4): methodology

2.1 The blind men and the elephant.

Three case studies:

– A theory [2];
– A single-cell electrophysiology study [81].
– An imaging study [44];

2.2 Open your eyes!

– The general methodology: the Marr-Poggio program for neurosciences [28, ch.4].
– A worked-out example: sound localization in the barn owl [54].

Primary readings


Other readings


Week 3 (2/9; 2/11): measurement and representation

3.1 Perceptual measurement \[28\] ch.5. Channel coding and hyperacuity \[85\]. The fundamental uncertainty \[82\].

3.2 Representation spaces \[28\] ch.5. The face space \[53\].

Primary readings


Other readings


Week 4 (2/16; 2/18): probability and the Ace of Bayes

4.1 A simple model of behavior [64] (cf. the generalized problem of behavior, to be discussed in Week 10 [31]). A probabilistic formulation of cognition [17]. The Bayesian framework [42].

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

Primary readings


Other readings


Week 5 (2/23; 2/25): break, exam

5.1 [no meeting (February break)]

5.2 First prelim.

Primary readings


Other readings


Week 6 (3/1; 3/3): universal tools, I

6.1 Learning, similarity, and generalization [28, ch.5]. Shepard’s Law [83].

6.2 Learning from examples as function approximation. Object recognition [11, 21, 76]. The challenges of function approximation. Ill-posedness and regularization [6].

Primary readings


Other readings


Week 7 (3/8; 3/10): universal tools, II

7.1 Managing dimensionality. Dimensionality reduction by random projections [27, 34].

7.2 Managing the storage and retrieval of examples. Associative memory [28, ch.6]. Locality-sensitive hashing [4].

Primary readings


Other readings


**Week 8 (3/15; 3/17): actions and consequences, I**

8.1 Managing action. The basics of motor control [96]. Bayesian motor decision making [59].

8.2 Action and reward. Reinforcement learning (RL) [98] and its relationship to Bayes. Hierarchical RL [8, 9].

**Primary readings**


**Other readings**


Week 9 (3/22; 3/24): actions and consequences, II; exam

9.1 Graphical models (Bayesian Networks) [28] ch.8. Goal-directed decision making [88].

9.2 Second prelim.

Primary readings


Other readings

Week 10 (4/5; 4/7): behavior; finally, brains and neurons, I

10.1 The generalized problem of behavior [31].

10.2 Brains and neurons. Receptive fields (RFs) and maps [28, ch.2,3]. What do neurons do?

Projection:
   – neurons as readout devices [13];
   – neurons and projection pursuit [19].

Primary readings


Other readings


Week 11 (4/12; 4/14): neurons, II

11.1 What do neurons do?

Time-dependent dynamic learning:

- spike timing dependent plasticity (STDP), and Hebbian learning (see the Scholarpedia article [84]);
- history-dependent learning with the BCM rule [19];
- learning dynamic embodied control modes [24].

11.2 What do neurons do?

Population dynamics:

- neural trajectories and classification [12].

Primary readings


Other readings


Week 12 (4/19; 4/21): neurons, III

12.1 Why do neurons do all these things? Similarity and kernels [82].

12.2 What does the brain do? Memory [63]. Prediction [18] [67].

Primary readings


Other readings


Week 13 (4/26; 4/28): back to the abstract / computational level

13.1 Advanced cognition [28, ch.8]. Reasoning [90]. General intelligence and IQ [66].

13.2 Even more advanced cognition [28, ch.8]. Analogy [46]. Creativity [7, 45].

Primary readings


Other readings


Week 14 (5/3; 5/5): sequential behaviors

14.1 The transition to language

– the problem of sequential behavior [60];
– the computational nature of language [28] ch.7.

14.2 Learning language

– general considerations [41];
– a process model [58].

Primary readings


Other readings


Computational concepts, principles, and methods: an abecedary

ANALOGY — a comparison that involves a structure mapping between complex entities, situations, or domains [46, 47]. Analogy is central to general cognitive function (general fluid intelligence, often referred to as $gF$ or IQ [86]) 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 [33, 56, 59, 73, 89, 99].

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 [85]. This principle is at work throughout cognition [28, 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 [32]. 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 [75].

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 [69]. One type of such model is the BAYES Network [72]. Graphical models map naturally onto the architecture of the brain [61].

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 [79]. This property of world knowledge gives rise to serious algorithmic challenges in truth.
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 [70].

**KERNELS** — A family of mathematical methods that arise from measurements of **SIMILARITY** of two vectors and that are widely applicable in modeling cognition [51, 52, 82]. 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., [92, 93]).

**MINIMUM DESCRIPTION LENGTH (MDL) PRINCIPLE** — The fundamental principle of the computational theory of learning, due to Solomonoff [87], 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 [80], 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 [43]. A related principle is that of **SIMPLICITY** [16].

**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 [29].

**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 [28]. 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 [28]. Such capacity for prediction turns out to be a very general explanatory principle in cognition [18], 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 [73].

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 [37].

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 [26, 34, 83].

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 [28, 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 [27].

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” [40]. 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 [27].

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 [20]. 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 [14].

Maximum 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 [65].

Syntax — The system of patterns and constraints that governs the composition of utterances in a natural language [35, 74].

Dobzhansky — “Nothing in biology makes sense except in the light of Evolution” [25].

References


