Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Behavioral Sciences

Reinforcement learning, efficient coding, and the statistics of natural tasks

Matthew Botvinick[1,2,3], Ari Weinstein[1,2], Alec Solway[4] and Andrew Barto[5]

The application of ideas from computational reinforcement learning has recently enabled dramatic advances in behavioral and neuroscientific research. For the most part, these advances have involved insights concerning the algorithms underlying learning and decision making. In the present article, we call attention to the equally important but relatively neglected question of how problems in learning and decision making are internally represented. To articulate the significance of representation for reinforcement learning we draw on the concept of efficient coding, as developed in perception research. The resulting perspective exposes a range of novel goals for behavioral and neuroscientific research, highlighting in particular the need for research into the statistical structure of naturalistic tasks.

Addresses
[1] Princeton Neuroscience Institute, Princeton University, Princeton, NJ, United States
[2] Department of Psychology, Princeton University, Princeton, NJ, United States
[3] Google DeepMind, London, United Kingdom
[4] Virginia Tech Carilion Research Institute, Roanoke, VA, United States
[5] College of Information and Computer Sciences, University of Massachusetts, Amherst, MA, United States

Corresponding author: Botvinick, Matthew (matthewb@princeton.edu)

Current Opinion in Behavioral Sciences 2015, **5**:71–77

This review comes from a themed issue on **Neuroeconomics**

Edited by **John P O'Doherty** and **Colin C Camerer**

For a complete overview see the Issue and the Editorial

Available online 28th August 2015

http://dx.doi.org/10.1016/j.cobeha.2015.08.009

Over recent years, the field of computational reinforcement learning (RL) has served as an indispensable guide for behavioral and neuroscientific research on learning and decision making. Since the first empirical links between RL and biology were forged [1,2], scientific applications of RL theory have become steadily more detailed and sophisticated, distinguishing a variety of specific learning procedures and shedding increasing light on relevant neural substrates [3].

These advances certainly deserve to be celebrated. However, our purpose here is to argue that if the pace of progress is to be maintained, it will be necessary to widen the scope of inquiry. To date, most applications of RL in behavioral and neural science have focused heavily on identifying the *algorithms* that underlie learning and decision making. If we are to complete the entire picture, it will be critical to answer a second, complementary question: how are problems in learning and decision making internally *represented*?

The issue of representation has long been an intensive focus in artificial intelligence and machine learning, including basic RL research (see, e.g., [4–7,8•]). By contrast, the topic has received relatively sparse attention in behavioral and neuroscientific applications of RL. Our purpose here is to lay out an agenda to fill this gap, articulating why the issue of representation is critical and specifying the research objectives to which it gives rise.

A key theme in our proposal will be the notion of efficient coding, drawn from research on perception. We begin by reviewing the development of the efficient coding construct in that field, and then move on to consider its relevance to RL. As we shall argue, the idea of efficient coding opens up important new directions for RL-informed research in behavioral and neural science, highlighting in particular the need to examine the structure of naturalistic tasks.

Efficient coding in sensation and perception

Beginning in the 1950s, a series of influential papers in the perception literature introduced what has come to be known as the *efficient coding hypothesis* [9••,10,11••,12–14]. Stated simply, this proposes that the brain represents sensory inputs in the most compact or economical fashion possible. This idea is formalized in terms drawn from information theory [15]. Efficient codes are understood as *compressing* sensory data, representing them with the fewest bits or action potentials. This is accomplished both by 'whitening' the data, representing it in terms of statistically independent features, and by assigning short codes — representations using few bits or action potentials — to frequently encountered stimuli [9••]. The resulting compression is said to filter out redundancy from incoming sensory signals. In simple terms, this means that when one feature of the input (e.g., the shape of a banana) predicts some other (yellow color), a single representational element can be used to signal both.

Note that the notion of redundancy directly implies the existence of structure in the data, in the form of feature correlations. This in turn implies a non-uniform probability distribution over stimuli; of all possible stimuli (orange bananas, red bananas, blue bananas, *etc.*), only a small subset will commonly occur. Efficient coding is, in this sense, closely tied to probability density estimation. As Ming and Holt [16] have expressed it, efficient codes ''should match the statistics of the signals they represent'' (p. 1312).

One central motivation for the efficient coding hypothesis derives from the potential payoffs an efficient code would confer. These include metabolic savings proceeding from the frugal use of action potentials [9••]. However, efficient coding also has natural payoffs at the level of information processing [14]. These derive from the fact, just discussed, that efficient coding necessarily reflects insights into domain structure. This property in turn allows efficient coding to support flexible generalization, rapid learning, and accurate prediction and inference.

Since its original introduction, the efficient coding theory has been widely applied in perception research. Some of the relevant work, following the pioneering ideas of Barlow [9••], has focused on neural coding (e.g., [17,18]), while other research has focused on the cognitive level (e.g., [19]), following ideas initially advanced by Attneave [11••]. The core ideas of efficient coding theory have also been increasingly transposed into other fields, including research on memory [20,21] and language [22]. Over recent years, there have been sporadic but increasing efforts to apply the notion of efficient coding to RL models of learning and decision making [6,23–25,26••]. We turn now to a consideration of this enterprise, beginning with some background on RL itself.

The reinforcement learning problem

As a matter of basic definition, RL addresses the problem of selecting actions so as to maximize long-term expected reward [27]. The task faced by an RL agent is typically specified as a Markov decision problem, comprising a set of situations or states (S), a set of candidate actions (A), a transition function (P) specifying action-outcome contingencies, and a reward function (R) encoding the agent's preferences. The agent's objective is to approximate an optimal (reward maximizing) action policy (π), typically on the basis of a value function (Q) which associates pairings of action and state with estimates of long-term expected reward.

The objective of discovering an optimal policy can be pursued either through direct interaction with the environment or through planning by way of an internal model of the environment [28]. However, in both cases, the undertaking centers on a process of exploration or *search*. An important difficulty, better appreciated in computer science than in behavioral and neuroscience research, is that this search process quickly becomes infeasible as the size of the RL problem grows [29]. As an illustration, imagine a video game in which the goal is to find a treasure by making a series of twenty choices between pairs of distinctively decorated doors, with each door leading to a unique pair of successors. Assuming one played this game non-stop, choosing a new series of doors every waking minute, discovering the treasure would take an average of one and a half years. If the next level of the game presented three doors at each step, completing this level would demand more than fifty lifetimes.

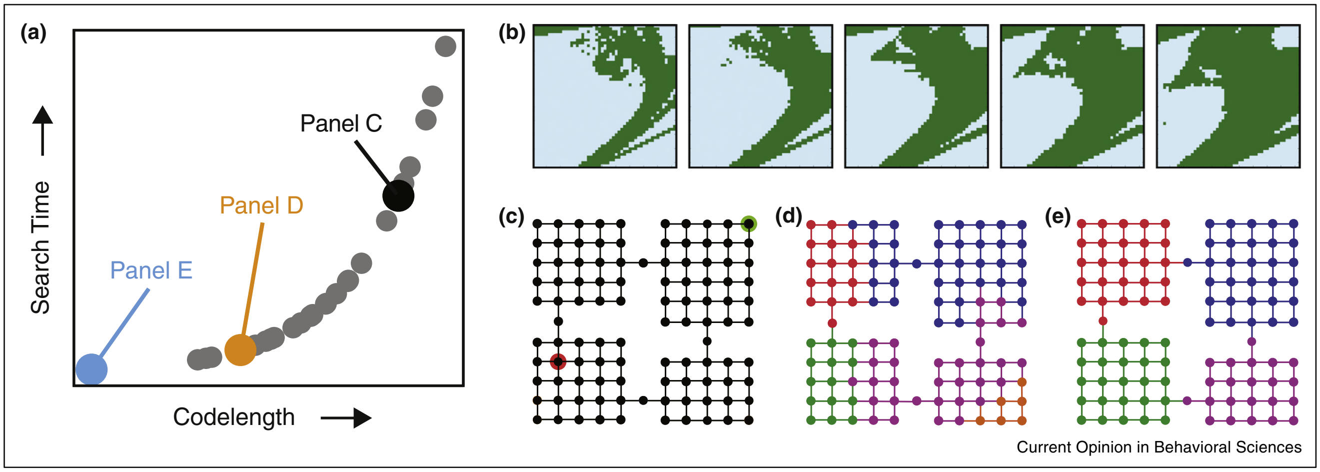
Given this kind of intractability, how can it be possible for humans and other animals to employ RL-like procedures to learn and make decisions? An answer becomes evident if one compares the video game from the example against the kinds of decision problems that more typically arise in everyday life. The video game, by design, represents a combinatorial worst-case scenario, devoid of any redundancy and therefore resistant to any kind of compact representation. In contrast, the world we inhabit in our everyday lives, and the tasks we face within it, are quite saturated with structure. As it turns out, this characteristic of real-world decision problems holds the key to overcoming the scaling problem in RL, precisely by providing the necessary conditions for efficient coding.

Efficient coding for RL

In order to make clear what it means for an RL problem to contain structure or redundancy, it may be useful to return briefly to the case of visual perception. There, the presence of redundancy meant that only a fraction of all possible patterns of retinal stimulation ever actually occur. A necessary consequence of this fact is that those patterns or images that do occur will share certain (possibly quite abstract) characteristics. This shared structure is what the term redundancy denotes.

Where might such shared structure inhere in RL? Clearly, one locus might be in the set of perceptual inputs that form part of an RL task, that is, the state set S. However, RL provides numerous other potential repositories for structure. For example, structure can reside within the set of optimal policies that represent behavioral solutions to a task. As a concrete example, consider the diagrams presented in Figure 1B. These visualize optimal policies from a simple but canonical task from the RL literature, the so-called 'mountain car' task (see [27]). Here, an underpowered car occupies a valley between two peaks, and must use its limited thrust to get over one of them. There are two actions: drive forward and drive in reverse. Each panel in the figure displays the optimal policy for a car with a particular level of thrust, with the forward

Figure 1



(a) Results from Solway *et al.*, *PLoS Computat Biol*, 2014. Each point relates to a particular agent hierarchy, corresponding to a specific graph partition. The *x*-axis indexes the number of bits needed, under a given hierarchy, to encode the set of target behaviors for the agent, that is, the set of all shortest paths within the graph (range 99,673–170,620.) The *y*-axis indexes the expected number of trial-and-error attempts needed to discover the solution to a randomly selected task (geometric mean; range 632–62,344). Note that the reference to panel C is intended to denote a trivial partition, with all nodes belonging to a single component. **(b)** Optimal policies for the mountain car task. Each panel shows the optimal policy for a car with a particular degree of thrust, with thrust increasing in a linear fashion from left to right. Within each plot, the *x*-axis indexes the car's velocity, and the *y*-axis indexes position along the mountain road, with the valley floor at center and the two mountain peaks at the upper and lower margins. Green pixels mark points in the state space where the optimal action is *drive forward*, and blue pixels where it is *drive in reverse*. **(c)** The rooms graph used as an example action domain by Solway *et al.* (2014). The agent was assumed to solve a series of shortest-path problems in the graph. For example, in one such task the agent would begin at the node marked in green, with the task of navigating by a shortest path to the node marked in red. **(d and e)** Two partitions of the rooms graph shown in panel C.

action shown in green and reverse in blue, and with velocity on the *x*-axis and spatial position on the *y*-axis. As is immediately obvious from the figure, the policy at each thrust level is highly structured or organized, as is the variation in the optimal policy across thrust levels. This is precisely the kind of structure that permits efficient coding. An agent solving the mountain car task could, in principle, represent the target policy using only a few parameters, projecting the pixels in the panels of Figure 1B into a lower-dimensional embedding.

Why would such compression be of benefit to an RL agent? As intimated above, it can provide a powerful antidote to the scaling problem. More specifically, if tasks can be represented compactly, in a way that captures their internal and shared structure, there are two interrelated payoffs: first, the agent can engage in more strategic and efficient search, discovering optimal policies more quickly and second, information acquired during exploration can be generalized more extensively, effectively allowing the agent to learn more from each hard-earned observation. The following section provides a concrete illustration of these benefits.

A proof of concept: hierarchical reinforcement learning

One form of structure that has been frequently remarked in human activity is hierarchy [30]. Hierarchical organization in behavior has been understood to reflect internal representations of action that are themselves hierarchical [31,32]. In recent work, Solway *et al.* [26••] discussed how such hierarchical coding of action can be viewed as a form of efficient coding.

This work drew on the computational framework of hierarchical reinforcement learning (HRL), a variation on RL in which the action set includes temporally extended behaviors or subroutines [5,32]. Solway *et al.* [26••] considered an HRL agent tasked with reaching a series of goal locations by moving from node to node in a floor-plan-like graph, shown in Figure 1C. The agent was assumed to 'carve up' the graph into a set of regions, and to build a corresponding action hierarchy containing subroutines designed to move the agent from one region to another. The most intuitive way for the agent to represent its environment hierarchically is indicated in Figure 1E. But there exist an astronomical number of other hierarchies the agent could build (see, e.g., Figure 1D). Solway *et al.* [26••] considered whether any particular hierarchical representation might be considered justifiably superior to all the others. They viewed this question from an efficient coding perspective. Specifically, they focused on the set of behaviors called for by the tasks posed to the HRL agent, and calculated how many bits would be required to represent those target behaviors under each candidate hierarchy. The resulting scores for several different hierarchies are plotted in Figure 1a. As suggested by intuition, the partition in Figure 1e achieves the optimal compression. The reason for this is readily evident: the paths the agent must take in navigating through the graph contain frequently recurring subsequences, each of which culminates at one of the door-like bottlenecks in the graph. Representing actions at the level of these subsequences rather than at the level of single edge traversals — rather like encoding written language at the level of words rather than individual letters — yields a maximally efficient code (see also [33•]).

Figure 1A also conveys a second, complementary observation from Solway *et al.* [26••]: The hierarchy that attains the most efficient code also allows the agent to solve new navigation problems in the smallest expected number of steps. This is because the agent's compact action representations structure its search. Rather than wandering aimlessly within any given 'room' in the graph, the agent's action hierarchy will cause it to march deliberately from room to room, as it must do to efficiently tackle new problems. In this case and in general, efficient coding results in efficient search.

Varieties of structure in RL

The results depicted in Figure 1A, like the example from Figure 1B, pertain to structure at the level of action policies. While action presents a particularly important domain for efficient coding (see e.g., [25,34]), we have already noted that RL contains a number of potential repositories for structure. At the broadest level, structure can inhere in each of the ingredient constructs enumerated earlier: S, A, P, R, Q and π, as well as in the mappings between these. Indeed, one can cull from the behavioral and neuroscience literatures a variety of studies examining structure in each of these domains, which together can be regarded as a nascent literature on efficient coding in RL.

Structure at the level of the action set, for example, has been considered in work on motor primitives, which seeks the 'alphabet' out of which complex gestures are composed [35–39] (for related work in engineering see [40–42]). Structure in the transition function has been considered in work on motor learning, where behavioral evidence shows that human learners can abstract complex patterns of action-outcome contingency, using such knowledge to support new learning [43•]. Also relevant here is recent work pursuing the idea that action-outcome associations are ascribed to inferred latent causes, providing an explanation for empirically observed patterns of generalization [44,45••] (for related work in engineering see [7,46]). Structure at the level of the value function has been considered in several recent lines of work (for engineering precedents, see [6,47]). Gershman et al. [48] studied behavior and brain activation in a task where the value function can be decomposed into a set of simpler representations. Wimmer *et al.* [24] applied similar methods to a task where redundancy in the value function supports generalization. And Niv et al. [49] looked at behavior and neural activity in a setting where feature-based attention enables a compact representation of the value function.

The statistics of natural tasks

Thus far, we have considered what efficient coding might mean in the context of RL, where the relevant forms of structure might reside, and what kinds of payoffs could be expected. However, in order to develop the full implications of an efficient coding account, there is another crucial question that must be broached: what specific forms of redundancy are actually encountered in everyday tasks? Without an answer to this question, efficient coding remains a purely abstract proposition. In order to apply the theory to the interpretation of specific empirical findings, and in order to make specific predictions, one must gather data on the statistical structure of the relevant domain. This enterprise has been central to the efficient coding literature in sensation and perception, where there have been extensive efforts to characterize the statistical properties of naturalistic images (e.g., [12,13,50]). A critical aspect of this enterprise is its focus on ecological validity. The structure that matters for efficient coding is the structure that the organism copes with in everyday existence (and over evolutionary time).

When ported to behavioral and neuroscientific research on RL, these considerations reveal a neglected imperative: in order to fully understand the critical role of representation in RL, we must attend to the statistics of natural tasks [51–53]. In practice, this means searching for redundancy in the structures we have been discussing (S, A, P, R, Q and π), but doing so in the context of everyday activities.

In undertaking this enterprise, some guidance might be drawn from existing work on the perception of action events. A key observation here has been that naturalistic activities are punctuated by discrete boundaries between subtask sequences, boundaries to which the brain appears to be highly sensitive [30,54–56]. This suggests an intriguing analogy to the importance of edges in image statistics (see [11••,50]). In this connection, it is perhaps no coincidence that the work that we have reviewed on action hierarchy [26••], which deals with the placement of boundaries between subtasks, rests upon the same mathematical foundations as recent work on object identification in vision [19], or that the same formal tools can be used to segment both Markov decision problems [57,58] and images [59].

In addition to work on event perception, useful ideas may also derive from past research on story understanding. In classic work in this area, Schank and Abelson [60] emphasized the ubiquity of structural overlap between naturalistic activities (see also [31,42]). On the basis of this, they proposed that naturalistic action might be represented by a set of general schemas capturing commonalities (e.g., a 'script' for going to a restaurant), supplemented by modifiers specifying exceptions to the general rules (e.g., the fact that in a fast-food restaurant one pays before eating) — a proposal that is strikingly similar to one made in Attneave's ([11••], p. 190) pioneering work on efficient coding in perception.

In developing a detailed picture of the statistics of natural tasks, it will be well to consider the pervasive structuring role played by basic physics. In image statistics, the simple fact of gravity results in a tendency for moveable objects to appear low in the field of view. In action, physics will clearly play an important role in structuring S, A and P. For example, informal observation suggests that action-outcome effects are generally sparse; any action tends to alter only one or a small set of features of the environment. Turning on one's laptop, for instance, alters the state of the laptop but leaves almost everything else (one's location, the clothes one is wearing, the weather, *etc.*) unchanged. Furthermore, in most cases, the effects of an action tend to depend only on a very small number of environmental features: under ordinary conditions, the effect of turning on the computer depends on whether the battery is charged, but not on much else. Such causal sparsity gives rise to ample opportunities for efficient coding [49,61].

Beyond basic physics, pervasive patterns of structure must also arise from the form of the human body, as well as from the organization of the cognitive faculties such as attention and memory [51]. A clear illustration comes from gaze- and motion-tracking work by Land, Hayhoe and colleagues [62,63], which has revealed strong regularities in patterns of visual fixation and eye–hand coordination during naturalistic action.

While these points bear on quite generic forms of structure, making good on the notion of efficient coding will also require understanding how task statistics differ in detailed ways between specific behavioral domains. Just as image statistics differ between indoor and outdoor scenes, task statistics will differ between cooking, shopping, sports and other activity classes.

How, practically speaking, can we get at such structure? One challenge relates to simply gathering the appropriate data. Here, recent advances in motion capture and 'life-logging' technology, as well as the availability of large video corpora [64,65•] may provide important leverage. Another challenge lies in finding methods of analysis that can help uncover important forms of redundancy in the relevant data. Here, among the many candidate tools to consider, neural networks [8•,31] and graph theoretic analysis [26••,58,66,67] appear particularly promising.

Finally, in pursuing insight into the statistics of natural tasks, it is worth considering that neuroscientific data may be informative. If, as we have advocated, the brain discovers efficient codes for everyday behaviors, examining these neural codes may provide clues as to the forms of redundancy upon which the brain is capitalizing. Intriguingly, analyses of neural activity in prefrontal cortex and hippocampus suggest that these regions represent dissociable features of complex tasks along orthogonal dimensions, consistent with the kind of 'whitening' discussed in efficient coding theory [68,69]. Examining such orthogonalization in the context of more naturalistic behaviors may provide a window on the structure of those behaviors.

We began the present article by arguing for the need to consider representation, in addition to algorithm, when considering how RL might relate to learning and decision making in humans and other animals. One consequence arising from the project of understanding task statistics is that it naturally reunites the topic of representation with that of algorithm. Just as task statistics will determine what form an efficient code will take, they also determine what class of planning or learning algorithm will be most efficient [70]. If we can build a detailed understanding of how these three pieces — algorithm, representation, and task statistics — fit together, the result will be a truly satisfying and comprehensive account of adaptive behavior.

Conflict of interest

None declared.

Author note

The present work was supported in part by the National Science Foundation (#1207833) and the John Templeton Foundation. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

References and recommended reading

Papers of particular interest, published within the period of review, have been highlighted as:

- • of special interest
- •• of outstanding interest

1. Schultz W, Dayan P, Montague P: **A neural substrate of prediction and reward**. *Science* 1997.
2. Barto AG: **Adaptive critics and the basal ganglia**. In *Models of information processing in the basal ganglia.* Edited by Houk JC, Davis J, Beiser D. MIT Press; 1995:215-232.
3. O'Doherty JP, Lee SW, McNamee D: **The structure of reinforcement-learning mechanisms in the human brain**. *Curr Opin Behav Sci* 2015, **1**:94-100.
4. Amarel S: **On representations of problems of reasoning about actions**. *Mach Intell* 1968, **3**:131-171.

5. Barto A, Mahadevan S: **Recent advances in hierarchical reinforcement learning**. *Discrete Event Dyn Syst* 2003, **13**:341-379.

6. Foster D, Dayan P: **Structure in the space of value functions**. *Mach Learn* 2002, **49**:325-346.

7. Konidaris G, Kaelbling LP, Lozano-Perez T: **Constructing symbolic representations for high-level planning**. In *Proceedings of the twenty-eighth conference on artificial intelligence*. 2014:1932-1940.

8. • Mnih V, Kavukcuoglu K, Silver D, Rusu AA, Veness J, Bellemare MG, Graves A, Riedmiller M, Fidjeland AK, Ostrovski G: **Human-level control through deep reinforcement learning**. *Nature* 2015, **518**:529-533.

Representation learning has played an increasing critical role in engineering applications of reinforcement learning. This recent paper provides a vivid illustration of that trend. It reports work using neural networks to learn representations supporting reinforcement learning in complex video-game environments.

9. •• Barlow HB: *Possible principles underlying the transformations of sensory messages*. 1961.

One of the classic papers introducing the notion of efficient coding into neuroscience research.

10. Barlow H: **Redundancy reduction revisited**. *Netw: Comput Neural Syst* 2001, **12**:241-253.

11. •• Attneave F: **Some informational aspects of visual perception**. *Psychol Rev* 1954, **61**:183.

A seminal paper exploring the potential role of efficient coding in understanding phenomena in the field of perception.

12. Simoncelli EP, Olshausen BA: **Natural image statistics and neural representation**. *Ann Rev Neurosci* 2001, **24**:1193-1216.

13. Geisler WS: **Visual perception and the statistical properties of natural scenes**. *Annu Rev Psychol* 2008, **59**:167-192.

14. Atick JJ: **Could information theory provide an ecological theory of sensory processing?** *Netw: Comput Neural Syst* 1992, **3**:213-251.

15. Cover TM, Thomas JA: *Elements of information theory*. John Wiley & Sons; 2012.

16. Ming VL, Holt LL: **Efficient coding in human auditory perception**. *J Acoust Soc Am* 2009, **126**:1312-1320.

17. Olshausen BA: **Emergence of simple-cell receptive field properties by learning a sparse code for natural images**. *Nature* 1996, **381**:607-609.

18. Ganguli S, Sompolinsky H: **Compressed sensing, sparsity, and dimensionality in neuronal information processing and data analysis**. *Ann Rev Neurosci* 2012, **35**:485-508.

19. Orbán G, Fiser J, Aslin RN, Lengyel M: **Bayesian learning of visual chunks by human observers**. *Proc Natl Acad Sci U S A* 2008, **105**:2745-2750.

20. Orhan AE, Sims CR, Jacobs RA, Knill DC: **The adaptive nature of visual working memory**. *Curr Direct Psychol Sci* 2014, **23**:164-170.

21. Botvinick M: **Effects of domain-specific knowledge on memory for serial order**. *Cognition* 2005, **97**:135-151.

22. Gibson E, Piantadosi ST, Brink K, Bergen L, Lim E, Saxe R: **A noisy-channel account of crosslinguistic word-order variation**. *Psychol Sci* 2013. 0956797612463705.

23. Veness J, Bellemare MG, Hutter M, Chua A, Desjardins G: *Compress and control*. 2014arXiv:1411.5326.

24. Wimmer GE, Daw ND, Shohamy D: **Generalization of value in reinforcement learning by humans**. *Eur J Neurosci* 2012, **35**:1092-1104.

25. Van Dijk SG, Polani D, Nehaniv CL: **Hierarchical behaviours: getting the most bang for your bit**. *Advances in artificial life: darwin meets von Neumann*. Springer; 2011:342-349.

26. •• Solway A, Diuk C, Córdova N, Yee D, Barto AG, Niv Y, Botvinick MM: **Optimal behavioral hierarchy**. *PLoS Comput Biol* 2014, **10**:e1003779.

Presents an efficient-coding perspective on hierarchical representation of goal-directed action. Theoretical results are tested in a series of human behavioral studies.

27. Sutton RS, Barto AG: *Reinforcement learning: an introduction*. Cambridge, MA: MIT Press; 1998.

28. Daw ND, Niv Y, Dayan P: **Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control**. *Nat Neurosci* 2005, **8**:1704-1711.

29. Kaelbling LP, Littman ML, Moore AW: **Reinforcement learning: a survey**. *J Artif Intell Res* 1996:237-285.

30. Zacks JM, Speer NK, Swallow KM, Braver TS, Reynolds JR: **Event perception: a mind-brain perspective**. *Psychol Bull* 2007, **133**:273-293.

31. Botvinick M, Plaut DC: **Doing without schema hierarchies: a recurrent connectionist approach to normal and impaired routine sequential action**. *Psychol Rev* 2004, **111**:395-429.

32. Botvinick MM, Niv Y, Barto AC: **Hierarchically organized behavior and its neural foundations: a reinforcement-learning perspective**. *Cognition* 2009, **113**:262-280.

33. • Tishby N, Polani D: **Information theory of decisions and actions**. *Perception-action cycle*. Springer; 2011:601-636.

An integrative theoretical framework bridging between reinforcement learning and information theory.

34. Huys QJ, Lally N, Faulkner P, Eshel N, Seifritz E, Gershman SJ, Dayan P, Roiser JP: **Interplay of approximate planning strategies**. *Proc Natl Acad Sci U S A* 2015, **112**:3098-3103.

35. Thoroughman KA, Shadmehr R: **Learning of action through adaptive combination of motor primitives**. *Nature* 2000, **407**:742-747.

36. d'Avella A, Saltiel P, Bizzi E: **Combinations of muscle synergies in the construction of a natural motor behavior**. *Nat Neurosci* 2003, **6**:300-308.

37. Graziano M: **The organization of behavioral repertoire in motor cortex**. *Annu Rev Neurosci* 2006, **29**:105-134.

38. Graziano MS: **Ethologically relevant movements mapped on the motor cortex**. *Primate Neuroethol* 2010:454.

39. Todorov E, Jordan MI: **Optimal feedback control as a theory of motor coordination**. *Nat Neurosci* 2002, **5**:1226-1235.

40. Minsky M: **Steps toward artificial intelligence**. *Proc IRE* 1961, **49**:8-30.

41. Ijspeert AJ, Nakanishi J, Hoffmann H, Pastor P, Schaal S: **Dynamical movement primitives: learning attractor models for motor behaviors**. *Neural Comput* 2013, **25**:328-373.

42. da Silva BC, Baldassarre G, Konidaris G, Barto A: **Learning parameterized motor skills on a humanoid robot**. *Robotics and automation (ICRA), 2014 IEEE international conference on IEEE*. 2014:5239-5244.

43. • Braun DA, Mehring C, Wolpert DM: **Structure learning in action**. *Behav Brain Res* 2010, **206**:157-165.

Describes behavioral findings indicating that human learners are capable of identifying shared structure across a series of task environments, allowing rapid adaptation to new environments displaying related structure. Effectively a case study in efficient coding in the setting of motor control.

44. Soto FA, Gershman SJ, Niv Y: **Explaining compound generalization in associative and causal learning through rational principles of dimensional generalization**. *Psychol Rev* 2014, **121**:526.

45. •• Gershman SJ, Niv Y: **Learning latent structure: carving nature at its joints**. *Curr Opin Neurobiol* 2010, **20**:251-256.

An excellent discussion of experimental and computational research examining the role of structure learning in adaptive behavior.

46. Ravindran B, Barto AG: **Model minimization in hierarchical reinforcement learning**. *Abstraction, reformulation, and approximation*. Springer; 2002:196-211.

47. Mahadevan S, Maggioni M: **Proto-value functions: a Laplacian framework for learning representation and control in Markov decision processes**. *J Mach Learn Res* 2007, **8**:2169-2231.

48. Gershman SJ, Pesaran B, Daw ND: **Human reinforcement learning subdivides structured action spaces by learning effector-specific values**. *J Neurosci* 2009, **29**:13524-13531.

49. Niv Y, Daniel R, Geana A, Gershman SJ, Leong YC, Radulescu A, Wilson RC: **Reinforcement learning in multidimensional environments relies on attention mechanisms**. *J Neurosci* 2015. [to be published].

50. Sigman M, Cecchi GA, Gilbert CD, Magnasco MO: **On a common circle: natural scenes and Gestalt rules**. *Proc Natl Acad Sci U S A* 2001, **98**:1935-1940.

51. Botvinick MM, Cohen JD: **The computational and neural basis of cognitive control: charted territory and new frontiers**. *Cognit Sci* 2014, **38**:1249-1285.

52. Barker RG: *Ecological psychology: concepts and methods for studying the environment of human behavior*. Stanford University Press; 1968.

53. Gibson JJ: *The theory of affordances*. USA: Hilldale; 1977.

54. Agarwal A, Triggs B: **Tracking articulated motion with piecewise learned dynamical models**. *European conference on computer vision*. 2004:54-65.

55. Fujii N, Graybiel AM: **Representation of action sequence boundaries by macaque prefrontal cortical neurons**. *Science* 2003, **301**:1246-1249.

56. Farooqui AA, Mitchell D, Thompson R, Duncan J: **Hierarchical organization of cognition reflected in distributed frontoparietal activity**. *J Neurosci* 2012, **32**:17373-17381.

57. Stachenfeld KL, Botvinick M, Gershman SJ: **Design principles of the hippocampal cognitive map**. *Advances in neural information processing systems*. 2014:2528-2536.

58. Schapiro A, Cordova N, Turk-Browne N, Rogers TT, Botvinick MM: **Neural representations of events arise from temporal community structure**. *Nat Neurosci* 2013, **16**:486-492.

59. Belkin M, Niyogi P: **Laplacian eigenmaps for dimensionality reduction and data representation**. *Neural Comput* 2003, **15**:1373-1396.

60. Schank RC, Abelson RP: *Scripts, plans, goals and understanding*. Hillsdale, NJ: Erlbaum; 1977.

61. Wolfe AP: *Paying attention to what matters: observation abstraction in partially observable environments*. Amherst: University of Massachusetts; 2010.

62. Hayhoe MM, Ballard D: **Eye movements in natural behavior**. *Trends Cognit Sci* 2005, **9**:188-194.

63. Land MF, Hayhoe MM: **In what ways do eye movements contribute to everyday activities?** *Vision Res* 2001, **41**:3559-3565.

64. Malmaud J, Huang J, Rathod V, Johnston N, Rabinovich A, Murphy K: **What's cookin'? Interpreting cooking videos using text**. *Speech Vision* 2015arXiv:1503.01558.

65. • Abbott WW, Thomik AA, Faisal AA: **Sparse encoding of complex action sequences**. *Cosyne (computational and systems neuroscience)*. 2015.

Describes a recent effort to identify low-dimensional structure in naturalistic human behavior, leveraging recent advances in motion-capture technology.

66. Simsek O, Barto AC: **Skill characterization based on betweenness**. *Adv Neural Inform Process Syst* 2008.

67. Zhang Y, Sreekumar V, Belkin M, Dennis S: **The network properties of episodic graphs. Member abstract**. In *Proceedings of the 32nd Annual Conference of the Cognitive Science Society*. 2010.

68. Sigala N, Kusunoki M, Nimmo-Smith I, Gaffan D, Duncan J: **Hierarchical coding for sequential task events in the monkey prefrontal cortex**. *Proc Natl Acad Sci U S A* 2008, **105**:11969-11974.

69. McKenzie S, Frank AJ, Kinsky NR, Porter B, Rivière PD, Eichenbaum H: **Hippocampal representation of related and opposing memories develop within distinct, hierarchically organized neural schemas**. *Neuron* 2014, **83**:202-215.

70. Simon DA, Daw ND: **Environmental statistics and the trade-off between model-based and TD learning in humans**. *Adv Neural Inform Process Syst* 2011:127-135.