

In press, *Trends in Cognitive Sciences*

Hierarchical models of behavior and prefrontal function

Matthew M. Botvinick
Princeton Neuroscience Institute and
Department of Psychology
Princeton University

February 18, 2008

Running head: Hierarchical models

Word count: 3195

Figures: 4

Text Boxes: 3

Corresponding author:
Matthew Botvinick
Princeton University
Department of Psychology
Green Hall
Princeton, NJ 08540
(609)258-1280
matthewb@princeton.edu

Abstract

The recognition of hierarchical structure in human behavior was one of the founding insights of the cognitive revolution. Despite decades of research, however, the computational mechanisms underlying hierarchically organized behavior are still not fully understood. Recent findings from behavioral and neuroscientific research have fueled a resurgence of interest in the problem, inspiring a new generation of computational models. In addition to developing some classic proposals, these models also break fresh ground, teasing apart different forms of hierarchical structure, placing a new focus on the issue of learning, and addressing recent findings concerning the representation of behavioral hierarchies within the prefrontal cortex. While offering explanations for some key aspects of behavior and functional neuroanatomy, the latest models also pose new questions for empirical research.

In 1951, Karl Lashley¹ delivered an address that did much to catalyze the cognitive revolution in psychology. Lashley's key assertion was that sequential behavior cannot be understood as a chain of stimulus-response associations. Instead, he argued, behavior displays hierarchical structure, comprising nested subroutines. A few years later, another key event in the shift to cognitivism occurred, with the publication of *Plans and the Structure of Behavior*, by Miller, Galanter and Pribram.² Once again, the theme was hierarchical structure in human action.

While the issue of hierarchy was central to the birth of cognitivism, it was also central to the birth of computational modeling as a tool in psychology. A critical contribution of the Miller, Galanter and Pribram book was to propose one of the first computer-inspired models of cognition. In the decades since this pioneering work, a considerable number of computational models have been proposed to account for hierarchical structure in human behavior (Box 1). While interest in the topic has been steady among modelers, recent developments have begun to push hierarchy back toward center stage of empirical research, as well. In particular, a renewed focus on hierarchical structure in behavior has appeared in cognitive psychology,^{3, 4} developmental psychology,^{5, 6} neuropsychology,⁷ and perhaps most strikingly in neuroscience^{8,9}.

Amid this resurgence of interest, a new generation of computational models has appeared. As reviewed in what follows these models, as a group, develop some key earlier proposals, while at the same time addressing some important new considerations.

BOX 1 AROUND HERE

Two Kinds of Hierarchical Structure

Descriptions of human behavior as hierarchical have rarely been accompanied by precise technical or operational definitions. Nevertheless, the basic idea is clear enough: The sequences or streams of action that humans produce can be analyzed into coherent *subunits* or *parts*.¹⁰ Though it has not always been explicit, the nature of these parts, and

of the wholes into which they form, has been understood in two ways, depending on whether the focus is on the *instrumental* structure of behavior, or its *correlational* structure.

Human behavior displays *instrumental* structure in the sense that action sequences bring about valued or desired outcomes, and successive actions fit together through means-ends relationships, with earlier actions setting up the conditions necessary for later ones. As an illustration of how instrumental structure can give rise to hierarchy, consider the simple action sequence outlined in Figure 1a. The arrows in the figure indicate means-ends links, where one action is performed in order to allow performance of another. Actions in red accomplish components of the overall goal of the sequence. In figure 1b, the same sequence is redrawn to highlight the presence of part-whole structure. As indicated by the overlay, actions within the sequence can be grouped into two sets, each accomplishing a component of the overall goal. Smaller subunits can be identified within these large ones, as well, for example the sequence accomplishing the subgoal of opening the door. This nesting of parts defines a hierarchy, as diagrammed in Figure 1c.

FIGURE 1 AROUND HERE

Hierarchical organization can alternatively be seen as a function of action's *correlational* structure. Here, coherent subunits are defined not by means-ends relationships, but by statistical co-occurrence, both among actions and between actions and environmental contexts. To return to our example, the actions *get-key*, *unlock-door*, and *open-door* are liable to occur together not only in the context of placing money in a safe, but also in the contexts of removing money, depositing or removing other items, or checking the safe's contents. As a result, the actions within this subsequence share stronger statistical associations with one another than they do with actions outside the sequence's boundaries. The actions within the sequence are also bound by a common association with the condition *door-closed*.

Note that correlational relations, like instrumental relations, can be nested or recursive. Thus, when the door-opening sequence is preceded by picking up money, the two together predict placing the money inside the safe. The overall sequence from picking up the money to depositing it thus forms a coherent subunit, which subsumes the door-opening sequence. Once again, recursive part-whole relationships yield a hierarchical structure.

Clearly, there is a close relationship between correlational and instrumental organization. Nevertheless, the distinction is important. Empirical work indicates that both forms of structure separately impact event memory, parsing and imitation.^{5, 10-13} Despite such evidence, hierarchical models of action production have tended until recently to neglect the distinction between instrumental and correlational structure, a point that provides an important piece of the background for understanding recent developments and current challenges.

Computational Fundamentals

All existing models of hierarchically structured behavior share at least one general assumption, which is that the hierarchical, part-whole organization of human action is mirrored in the internal or neural representations underlying it. Specifically, the assumption is that there exist representations not only of low-level motor behaviors, but also separable representations of higher-level behavioral units. Thus, in the example from Figure 1, there would exist representations not only for the action *pick-up-key*, but also for the more complex actions *open-door*, *deposit-money* and *lock-money-in-safe* (see Figure 1c). Such representations have gone by many names, including *task*, *subtask*, and *goal* representations; *context* representations; *schemas*; *macros*; *chunks*; and *skills*. At a computational level, what unites all such constructs is the property of *temporal abstraction*, the use of a single representation to span and unite a sequence of events. Often, temporal abstraction also goes hand in hand with *policy abstraction*, the use of a single representation to cover an entire mapping from stimuli to responses. For example,

a representation for *turn-on-light* might cover cases where this involves flipping a switch, pushing a dimmer, or pulling a chain.

Of course, simply positing action representations at multiple hierarchical levels is not sufficient to explain the production of hierarchically structured behavior. Also needed is an account of how these representations are selected or activated at the appropriate time, and of how, once selected, they guide the orderly selection of lower-level action representations. Over the course of the relevant lower-level actions, activation of higher-level action representations must be maintained, and in this sense the control of hierarchically structured action requires working memory. O'Reilly and Frank¹⁴ have noted that, in this context, working memory requires three special properties: (1) robust maintenance, the ability for high-level action representations to remain stably active over the course of finer-grained events, (2) rapid updating, allowing for the immediate selection and de-selection of high-level action representations at the boundaries of subtasks, and (3) selective updating, allowing selection of action representations at different levels of hierarchical structure at different times.

While the above are the basic elements shared by existing models of hierarchical behavior, there is wide variation in the ways that these elements are cashed out in different models, as detailed in the following sections.

Recent Models of Hierarchically Structured Behavior

Perhaps the most extensively developed recent model of hierarchical behavior has been put forth by Cooper and Shallice.¹⁵⁻¹⁷ This model carries forward an approach with a long tradition (Box 1), according to which it is assumed that the processing system has a hierarchical architecture that maps directly onto the structure of the relevant task domain. Cooper and Shallice¹⁵⁻¹⁸ begin by conceptualizing hierarchically structured tasks in the form of tree structures like the one shown in Figure 1c, and then build models that contain a simple connectionist processing unit, or “schema node,” for every element in the tree. Figure 2a shows one such model, which implements the task of coffee-making.

The network contains units for simple operations like *pick-up*, higher-level procedures like *add-milk-from-carton*, and for the entire task *prepare-instant-coffee*. These units carry scalar activation values, which are influenced by excitatory input from higher-level units, inhibitory input from competing units at the same hierarchical level, and excitatory input from perceptual inputs. Top-down input to each unit is also gated by a “goal node,” which prevents top-down excitation from activating any action whose outcome has already occurred (for example, top-down excitation to the action *empty-spoon* is gated if the spoon is already empty). Performance of the task is simulated by assigning a positive activation to the top-level schema node and applying external inputs representing the initial state of the environment. Figure 2b illustrates the operation of the model by plotting the activation of each schema node over cycles of processing. In addition to providing a basic account of hierarchical action production, the Cooper and Shallice model has been used to account for detailed empirical findings concerning performance errors in apraxia^{7, 19} and everyday ‘slips of action’.¹⁵⁻¹⁷

FIGURE 2 AROUND HERE

An alternative computational account, which addresses some of the same data, was put forth by Botvinick and Plaut.²⁰⁻²² Picking up on earlier work by Elman²³ and Cleeremans²⁴ (Box 1), this model does not assume a constitutively hierarchical processing system. Instead, all that is assumed is a set of processing units representing perceptual inputs, a set of units representing actions, and an interposed set of hidden or internal units (Figure 2c). A critical feature of the model is that its internal units are interconnected. This allows information to be preserved and transformed over time, in the pattern of activation over these units. Botvinick and Plaut²¹ applied this model to the coffee-making task addressed by Cooper and Shallice,¹⁵ demonstrating its ability to negotiate this hierarchically structured domain. Rather than relying on an explicitly hierarchical architecture, the model learned to represent the task’s layers of structure within the distributed patterns of activation arising over its internal units (Figure 1d). The hierarchical structure of the task was encoded implicitly in its internal representations and dynamics, as shaped through domain-general learning mechanisms,

rather than being explicitly mirrored, in a one-to-one fashion, by the system's basic elements.

Comparing this account to their own, Cooper and Shallice¹⁸ criticized the absence of explicit goal representations in the Botvinick and Plaut model. With the inclusion of goal nodes, the Cooper and Shallice model explicitly encodes one component of instrumental structure, as defined earlier. The Botvinick and Plaut model, in contrast, encodes only correlational structure. Botvinick and Plaut²² argued, however, that this may not be a flaw. Both models, they noted, specifically address routine or habitual behavior, and there is reason to believe that such behavior can be triggered directly by environmental and sequential contexts, without the mediation of goal representations.²⁵ This assertion fits closely with research pointing to the existence of two systems for action control, a habit system based on context-response associations and a goal-directed system operating through the anticipation of action outcomes.^{26, 27} Botvinick and Plaut²² suggest that, in humans, both of these systems may be capable of encoding hierarchical structure, but may encode it differently, with the habit system capturing correlational structure and the goal-directed system capturing instrumental structure.

Another important difference between the models proposed by Botvinick and Plaut and by Cooper and Shallice is that the former includes a fully implemented account of learning. While this is a relative strength, O'Reilly and Frank¹⁴ have called attention to slow rate at which hierarchical task structure can be learned within recurrent networks of kind studied by Botvinick and Plaut. In particular, although such networks can support the robust maintenance necessary for hierarchical behavior, this functionality is acquired only with great difficulty. As an alternative, O'Reilly and Frank¹⁴ put forth a biologically-inspired model that learns, through reinforcement-based mechanisms, to gate context information into a dedicated working-memory buffer. Importantly, this buffer, which models the role of prefrontal cortex, contains multiple independently gated subunits or "stripes," allowing the model to negotiate tasks with multiple, hierarchical levels of structure (see Figure 4d).

Hierarchical Reinforcement Learning

The work of O'Reilly and Frank¹⁴ is representative of an emerging focus, within research on hierarchically structured behavior, on the issue of learning. In an interesting parallel development, the potential role of hierarchy has taken on increasing interest within the field of machine learning, and in particular in research on reinforcement learning. As explained in Box 2, hierarchical methods for reinforcement learning provide a powerful computational framework for understanding how abstract action representations might develop through experience, and also call attention to the role that such representations might play in *supporting* learning in novel task domains. As recently explored by Botvinick, Niv and Barto,²⁸ and further discussed in Box 2, hierarchical reinforcement learning may also shed light on the neural mechanisms underlying hierarchically structured behavior in humans.

BOX 2 AROUND HERE

Hierarchical Structure in Prefrontal Cortex

The neural mechanisms underlying the production of hierarchically organized behavior have long been considered to reside, at least in part, within the dorsolateral prefrontal cortex (DLPFC). Based on neurophysiological and neuropsychological findings, Fuster^{29, 30} proposed that the DLPFC plays a key role in the temporal integration of behavior, serving to maintain context or goal information at multiple, hierarchically nested levels of task structure. In connection with this function, Fuster also noted the position of the DLPFC at the apex of an anatomical hierarchy of cortical areas (Figure 4a). Recent research has introduced an important extension to this idea, by suggesting that there may exist a topographical organization *within* frontal cortex and the DLPFC, according to which progressively higher levels of behavioral structure are represented as one moves rostrally^{8, 9, 31-35} (Figure 4b).

The discovery of this topographic organization places a new constraint on computational models of hierarchical behavior, and models addressing the relevant findings are now beginning to appear. In one such effort, Botvinick³⁶ reimplemented the recurrent neural network model from Botvinick and Plaut,²¹ introducing a structural hierarchy resembling the hierarchy of cortical areas described by Fuster³⁰ (Figure 4c). When the resulting network was trained on a hierarchically structured task, the processing units at the apex of the structural hierarchy spontaneously came to code selectively for temporal context information, while units lying lower in the hierarchy, nearer to the network's input and output layers, came to code more strongly for current stimuli and response information. These simulations thus showed how a functional-representational gradient like the one observed in the cerebral cortex could emerge spontaneously through learning, given only an initial architectural constraint.

A simulation study by Reynolds and O'Reilly³⁷ took a closely related approach, but arrived at interestingly different conclusions. This work began with the gating model from O'Reilly and Frank¹⁴ (see Figure 4d), introducing hierarchical functional relationships among the model's prefrontal stripes. The introduction of such structure was expected to give rise, through learning, to the kind of topographic division of labor described in recent empirical studies of DLPFC, with different levels of temporal or task structure sorting out in a topographic fashion (see Figure 4B). However, this result was not in fact obtained — an outcome that invites scrutiny of the differences between this model and the one tested by Botvinick.³⁶ In a second simulation, Reynolds and O'Reilly³⁷ did obtain a topographic division of labor when learning was permitted in different stripes at different phases in training. This, along with relevant neurodevelopmental findings, led to the proposal that the functional-topographic organization observed in DLPFC may arise from a maturational gradient, according to which rostral areas come on line later than more caudal areas.

FIGURE 4 AROUND HERE

According to Reynolds and O'Reilly,³⁷ the region at the apex of the architectural hierarchy within the DLPFC, frontopolar cortex (FPC; Brodmann area 10), would bear responsibility for representing task context at the highest levels of hierarchical structure. A different computational account of FPC was recently offered by Koechlin and Hyafil.³⁸ In their model, while nested levels of task context are represented by multiple areas posterior to FPC, the FPC itself enables representation of a *pending* task to be maintained, awaiting completion of another, currently running task (Figure 4e). On this account, rather than representing the uppermost element in a single hierarchy, the FPC effectively allows switching among independent hierarchies.

Despite the differences among the models of prefrontal cortex we have mentioned, it is natural, in each of them, to consider *timescale* as the key parameter governing the layout of cortical representations: As one moves rostrally over the cerebral cortex, and within the DLPFC, one encounters representations that guide behavior over progressively longer temporal spans.³⁹ From this point of view, the prefrontal hierarchy is understood as involving levels of increasing 'temporal abstraction,' as defined earlier. While this is consistent with much empirical evidence, there is also data suggesting that the prefrontal hierarchy may also align with other forms of abstraction. For example, Badre and D'Esposito⁴⁰ found that frontal activation shifted rostrally as successive layers of conditional dependency were added to a behavioral task. Their results suggest that the representations at each level of the prefrontal hierarchy may express task structure in terms of the representations at the level below. Here, it is not temporal abstraction that seems to be most relevant, but rather policy abstraction — the representation of a task, at any given moment of performance, as a set of choices over lower-level tasks.

Still other data point to the potential relevance of another form of abstraction. Christoff and colleagues⁴¹ observed a rostral shift in prefrontal activation as subjects performed a task that required them to process words naming increasingly abstract concepts. Such a finding may appear incommensurable with the notion that the prefrontal hierarchy is organized according to levels of temporal or policy abstraction. However, it has been noted in work on computational reinforcement learning that there is a natural relationship

between temporal or policy abstraction and *state abstraction*, the treatment of non-identical stimuli or situations as equivalent.^{42, 43} It is often the case that the successful performance of temporally extended tasks requires strict distinctions concerning some aspects of current inputs (e.g., in reading, what are the letters?) while allowing one to ignore or ‘abstract over’ others (what is the font?).⁴⁴ State abstraction, in this sense, bears a close logical relationship with category representation,⁴⁵ a point that may provide a clue as to why evidence has emerged for prefrontal organization both in terms of temporal grain and in terms of semantic category.

Conclusion

Computational modeling of hierarchically structured behavior, once at the center of the cognitive revolution, has been re-energized thanks to a new focus on hierarchy in behavioral, developmental and neuroscientific research. As reviewed here, recent models have elaborated on a number of earlier ideas but also added some new ones. In particular, there is an emerging focus on how hierarchical action representations are learned and on how they affect learning; a growing cognizance of the distinction between correlational and instrumental structure, and of the parallel between this distinction and the one between habitual and goal-directed forms of action control; and, finally, a new effort to provide a computational account for the role of prefrontal cortex in hierarchically structured behavior. While the latest crop of models provides new insights, it also poses new or refined questions for empirical research, including questions about how abstract action representations emerge through learning, how they interact with different modes of action control, and how they sort out within the prefrontal cortex (Box 3).

Text Boxes

Box 1: The Evolution of Hierarchical Models

In one of the first attempts to explicate hierarchically structured behavior in computational terms, Miller, Galanter and Pribram² introduced what they called the “test-operate-test-exit” or TOTE unit. When selected, a TOTE unit would test for a specific environmental state, and if this condition were not met, the TOTE unit would execute a specific action until the condition became true. Critically, the action executed by the TOTE unit could be to select another TOTE unit, allowing a recursive or hierarchical processing structure. A related approach was developed by Estes,⁴⁶ who posited a hierarchy of “control elements,” which activated units at the level below. Ordering of the lower units depended on lateral inhibitory connections, running from elements intended to fire earlier in the sequence to later elements. After some period of activity, elements were understood to enter a refractory period, allowing the next element in the sequence to fire. This same basic scheme was later implemented in a computer simulation by Rumelhart and Norman,⁴⁷ with a focus on typing behavior. Models proposed since this pioneering work have introduced a number of innovations. Norman and Shallice⁴⁸ discussed how schema activation might be influenced by environmental events; MacKay⁴⁹ introduced nodes serving to represent abstract sequencing constraints (see also Dell, Berger, & Svec⁵⁰); Dehaene and Changeux⁵¹ built in mechanisms for means-ends analysis and backtracking; and Grossberg⁵² and Houghton⁵³ introduced methods for giving schema nodes an evolving internal state and explored the consequences of allowing top-down connections to vary in weight. Despite these developments, all of these models retain the basic idea, present in Estes’ model, that the architecture of the processing system maps directly onto the underlying task, with a discrete processing element corresponding to every node in the task hierarchy.

A different approach to encoding hierarchical structure was introduced by Elman²³ and Cleeremans²⁴ (see also Reynolds, Zacks and Braver⁵⁴). Instead of assuming a hierarchy of processing elements, these models assumed an undifferentiated, recurrently connected

group of units mediating between stimulus inputs and action outputs. Applying this architecture to prediction tasks, both Elman and Cleeremans showed that its internal units learned to encode hierarchical sequential structure in a distributed fashion. Different levels of temporal context came to be represented along different dimensions within a high-dimensional space defined by the vector of internal-unit activations (see Figure 2c-d).

Box 2: Hierarchical Reinforcement Learning

Reinforcement learning (RL) problems are defined in terms of a set of world states, a set of available actions, a reward function attaching immediate rewards to specific states and actions, and a transition function encoding action consequences.⁵⁵ The learning objective is to arrive at a policy (a mapping from states to actions) that maximizes long-term cumulative reward. In hierarchical reinforcement learning (HRL),^{28, 56-59} the set of actions is expanded to include temporally abstract actions, referred to in one influential implementation⁵⁹ as *options*. Options are selected just as primitive actions, but once selected they remain active until an option-specific termination state is reached. While active, an option imposes its own option-specific policy, which can lead to selection of any combination of primitive actions and other options (Figure 3a). HRL furnishes algorithms for learning an optimal policy given a set of options, as well as algorithms for learning optimal option-specific policies.

FIGURE 3 AROUND HERE

The primary motivation for adding temporally abstract actions to RL is that they can dramatically facilitate learning. Figure 3b provides an illustration from Botvinick, Niv and Barto,²⁸ based on work by Sutton, Precup and Singh.⁵⁹ Here, learning to reach a goal in an environment divided into four rooms occurs faster when a basic set of primitive, single-step actions is supplemented with a set of four options, each of which defines a sub-policy for reaching one of the four doors. Critically, as shown by Botvinick, Niv and

Barto,²⁸ such a payoff accrues only if there is a good fit between the options available and the current learning task; an ill-suited set of options can actually slow learning. Thus, in HRL, a current challenge is to understand how learning across multiple problems might give rise to a set of options that is likely to be applicable in a wide range of future tasks.⁵⁷

Extensive evidence suggests that standard RL algorithms may be relevant to understanding learning mechanisms in the brain.^{60, 61} In view of this, it is intriguing to consider whether HRL might have relevance to understanding the neural basis of hierarchical learning. In order to explore this, Botvinick, Niv and Barto²⁸ considered how existing theories concerning RL and the brain would need to be modified in order to accommodate HRL mechanisms (Figure 3c-d). The results of this analysis suggested how the requisite extensions might map onto the functionality of dorsolateral and orbital prefrontal cortex, as revealed by recent single-unit recording studies (Figure 3d).

Box 3: Outstanding Questions

1. Mounting evidence points to the existence of two action control mechanisms: a habit system that operates based on context-response associations and a goal-directed system that plans based on predicted action outcomes. Such systems would appear naturally suited to encode correlational and instrumental hierarchical form, respectively. In modeling hierarchically structured behavior, should we have a dual-system account in mind?
2. How are hierarchical representations of behavior learned? How might this yield skills that are useful across tasks, supporting later learning? What are the relevant neural mechanisms? Are computational techniques for hierarchical reinforcement learning potentially relevant in addressing these questions?
3. Recent findings suggest that hierarchical action representations may map topographically onto frontal cortex. Why purpose might this serve, from a computational

perspective? What are the computational factors driving the development of this topographic organization? What are the roles of temporal, policy and state abstraction in structuring this functional-anatomic hierarchy, and how might these different forms of abstraction interrelate or align?

References

1. Lashley, K.S. (1951) The problem of serial order in behavior. In *Cerebral mechanisms in behavior: The Hixon symposium* (Jeffress, L.A., ed), 112-136, Wiley
2. Miller, G.A., *et al.* (1960) *Plans and the structure of behavior*. Holt, Rinehart & Winston
3. Schneider, D.W., and Logan, G.D. (2006) Hierarchical control of cognitive processes: switching tasks in sequences. *Journal of Experimental Psychology: General* 135, 623-640
4. Kurby, C.A., and Zacks, J.M. (2008) Segmentation in the perception and memory of events. *Trends in Cognitive Sciences* 12, 72-79
5. Saffran, J.R., and Wilson, D.P. (2003) From syllables to syntax: multilevel statistical learning by 12-month-old infants. *Infancy* 4, 273-284
6. Whiten, A., *et al.* (2006) Imitation of hierarchical action structure by young children. *Developmental Science* 9, 574-582
7. Schwartz, M. (2006) The cognitive neuropsychology of everyday action and planning. *Cognitive Neuropsychology* 23, 202-221
8. Badre, D. (in press) *Trends in Cognitive Sciences*
9. Koechlin, E. (in press) The cognitive architecture of the human lateral prefrontal cortex. In *Attention and Performance*, MIT Press
10. Zacks, J.M., *et al.* (2007) Event perception: a mind-brain perspective. *Psychological Bulletin* 133, 273-293
11. Avrahami, J., and Kareev, Y. (1994) The emergence of events. *Cognition* 53, 239-261
12. Baldwin, D.A., *et al.* (2001) Infants parse dynamic action. *Child Development* 72, 708-717
13. Want, S., and Harris, P. (2001) Learning from other people's mistakes: causal understanding in learning to use a tool. *Child Development* 72, 431-443
14. O'Reilly, R.C., and Frank, M.J. (2006) Making working memory work: a computational model of learning in prefrontal cortex and basal ganglia. *Neural Computation* 18, 283-328

15. Cooper, R., and Shallice, T. (2000) Contention scheduling and the control of routine activities. *Cognitive Neuropsychology* 17, 297-338
16. Cooper, R.P. (2007) Tool use and related errors in Ideational Apraxia: The quantitative simulation of patient error profiles. *Cortex* 43, 319-337
17. Cooper, R.P., *et al.* (2005) The simulation of action disorganization in complex activities of daily living. *Cognitive Neuropsychology* 22, 959-1004
18. Cooper, R.P., and Shallice, T. (2006) Hierarchical schemas and goals in the control of sequential behavior. *Psychological Review*, in press
19. Schwartz, M.F. (2006) The cognitive neuropsychology of everyday action and planning. *Cognitive Neuropsychology* 23, 202-221
20. Botvinick, M., and Plaut, D.C. (2002) Representing task context: proposals based on a connectionist model of action. *Psychological Research* 66, 298-311
21. Botvinick, M., and Plaut, D.C. (2004) Doing without schema hierarchies: a recurrent connectionist approach to normal and impaired routine sequential action. *Psychological Review* 111, 395-429
22. Botvinick, M., and Plaut, D.C. (2006) Such stuff as habits are made on: A reply to Cooper and Shallice (2006). *Psychological Review*, under review
23. Elman, G. (1990) Finding structure in time. *Cognitive Science* 14, 179-211
24. Cleeremans, A. (1993) *Mechanisms of implicit learning: connectionist models of sequence processing*. MIT Press
25. Wood, W., and Neal, D.T. (2007) A new look at habits and the habit-goal interface. *Psychological Review* 114, 843-863
26. Daw, N.D., *et al.* (2005) Uncertainty-based competition between prefrontal and striatal systems for behavioral control. *Nature Neuroscience* 8, 1704-1711
27. Balleine, B.W., and Dickinson, A. (1998) Goal-directed instrumental action: contingency and incentive learning and their cortical substrates. *Neuropharmacology* 37, 407-419
28. Botvinick, M.M., *et al.* (in press) Hierarchically organized behavior and its neural foundations: a reinforcement-learning perspective. *Cognition*
29. Fuster, J.M. (1990) Prefrontal cortex and the bridging of temporal gaps in the perception-action cycle. *Annals of the New York Academy of Sciences* 608, 318-329
30. Fuster, J.M. (2004) Upper processing stages of the perception-action cycle. *Trends in Cognitive Sciences* 8, 143-145

31. Koechlin, E., and Jubault, T. (2006) Broca's area and the hierarchical organization of behavior. *Neuron* 50, 963-974
32. Koechlin, E., *et al.* (2003) The architecture of cognitive control in the human prefrontal cortex. *Science* 302, 1181-1185
33. Wood, J.N., and Grafman, J. (2003) Human prefrontal cortex: processing and representational perspectives. *Nature Reviews Neuroscience* 4, 139-147
34. Christoff, K., and Gabrieli, J.D.E. (2000) The frontopolar cortex and human cognition: evidence for a rostrocaudal hierarchical organization within the human prefrontal cortex. *Psychobiology* 28, 168-186
35. Hamilton, A.F.d.C., and Grafton, S.T. (2007) The motor hierarchy: from kinematics to goals to intentions. In *Attention and Performance XXII* (Rosetti, Y., *et al.*, eds), Oxford University Press
36. Botvinick, M.M. (2007) Multilevel structure in behaviour and the brain: a model of Fuster's hierarchy. *Philosophical Transactions of the Royal Society (London), Series B* 362, 1615-1626
37. Reynolds, J.R., and O'Reilly, R.C. (in press) Developing PFC representations using reinforcement learning. *Cognition*
38. Koechlin, E., and Hyafil, A. (2007) Anterior prefrontal function and the limits of human decision-making. *Science* 318, 594-598
39. Koechlin, E., and Summerfield, C. (2007) An information theoretical approach to prefrontal executive function. *Trends in Cognitive Sciences* 11, 229-235
40. Badre, D., and D'Esposito, M. (2007) Functional magnetic resonance imaging evidence for a hierarchical organization of the prefrontal cortex. *Journal of Cognitive Neuroscience* 19, 2082-2099
41. Christoff, K., and Keramatian, K. (2007) Abstraction of mental representations: Theoretical considerations and neuroscientific evidence. In *Perspectives on Rule-Guided Behavior* (Bunge, S.A., and Wallis, J.D., eds), Oxford University Press
42. Jonsson, A., and Barto, A. (2001) Automated State Abstraction for Options using the U-Tree Algorithm. In *Advances in Neural Information Processing Systems 13*, 1054-1060, MIT Press
43. Dietterich, T.G. (2000) Hierarchical reinforcement learning with the maxq value function decomposition *Journal of Artificial Intelligence Research* 13, 227-303
44. Rougier, N.P., *et al.* (2005) Prefrontal cortex and flexible cognitive control: Rules without symbols. *Proceedings of the National Academy of Sciences* 102, 7338-7343

45. Rogers, T.T., and McClelland, J.L. (2004) *Semantic Cognition*. MIT Press
46. Estes, W.K. (1972) An associative basis for coding and organization in memory. In *Coding processes in human memory* (Melton, A.W., and Martin, E., eds), 161-190, V. H. Winston & Sons
47. Rumelhart, D., and Norman, D.A. (1982) Simulating a skilled typist: a study of skilled cognitive-motor performance. *Cognitive Science* 6, 1-36
48. Norman, D.A., and Shallice, T. (1986) Attention to action: Willed and automatic control of behavior. In *Consciousness and self-regulation: Advances in research and theory* (Davidson, R.J., et al., eds), Plenum Press
49. MacKay, D.G. (1987) *The organization of perception and action: a theory for language and other cognitive skills*. Springer-Verlag
50. Dell, G.S., et al. (1997) Language production and serial order. *Psychological Review* 104, 123-147
51. Dehaene, S., and Changeux, J.-P. (1997) A hierarchical neuronal network for planning behavior. *Proceedings of the National Academy of Sciences* 94, 13293-13298
52. Grossberg, S. (1986) The adaptive self-organization of serial order in behavior: Speech, language, and motor control. In *Pattern Recognition by Humans and Machines, vol. 1: Speech Perception* (Schwab, E.C., and Nusbaum, H.C., eds), 187-294, Academic Press
53. Houghton, G. (1990) The problem of serial order: A neural network model of sequence learning and recall. In *Current Research in Natural Language Generation* (Dale, R., et al., eds), 287-318, Academic Press
54. Reynolds, J.R., et al. (2007) A computational model of event segmentation from perceptual prediction. *Cognitive Science* 31, 613-643
55. Sutton, R.S., and Barto, A.G. (1998) *Reinforcement Learning: An Introduction*. MIT Press
56. Dietterich, T.G. (1998) The MAXQ method for hierarchical reinforcement learning. *Proceedings of the International Conference on Machine Learning*
57. Barto, A.G., and Mahadevan, S. (2003) Recent advances in hierarchical reinforcement learning. *Discrete Event Dynamic Systems: Theory and Applications* 13, 343-379
58. Parr, R., and Russell, S. (1998) Reinforcement learning with hierarchies of machines. *Advances in Neural Information Processing Systems* 10, 1043-1049

59. Sutton, R.S., *et al.* (1999) Between MDPs and semi-MDPs: A framework for temporal abstraction in reinforcement learning. *Artificial Intelligence* 112, 181-211
60. Schultz, W., *et al.* (1997) A neural substrate of prediction and reward. *Science* 275, 1593-1599
61. Joel, D., *et al.* (2002) Actor-critic models of the basal ganglia: New anatomical and computational perspectives. *Neural Networks* 15, 535-547
62. Fuster, J.M. (2001) The prefrontal cortex--An Update: Time is of the essence. *Neuron* 30, 319-333

Figure Captions

Figure 1. An illustration of hierarchical instrumental structure. a) An action sequence for locking money in a safe. Arrows denote means-ends relationships. Red indicates that the action accomplishes one component of the goal (money in the safe with the door closed and locked). b) The sequence in panel a, redrawn. Blue fields indicate coherent parts and subparts of the action sequence. At the coarsest level, the sequence breaks down into two parts, one organized around the subgoal of depositing the money, the other around the subgoal of locking the safe door. The action *pick-up-key* subserves both goals. Also indicated is a subordinate or nested sequence, organized around opening the safe door, c) One way of representing the sequence as a schema-, subtask- or subgoal hierarchy. Temporally abstract actions appear in blue.

Figure 2. a) Schematic of the coffee-making model from Cooper and Shallice.¹⁵ Filled circles: schema nodes. Bold labels: goal nodes. b) Activation of the schema nodes in the model from panel a, over the course of one task-completion episode. c) Schematic of the model from Botvinick and Plaut,²¹ showing only a subset of the units in each layer. Arrows indicate all-to-all connections. d) A two-dimensional representation of a series of internal representations arising in the Botvinick and Plaut model, generated using multi-dimensional scaling. Each point corresponds to a 50-dimensional pattern of activation across the network's hidden units. Both traces are based on patterns arising during performance of the sugar-adding subtask (o = first action, *locate-sugar*; x = final action, *stir*). The solid trajectory shows patterns arising when the sequence was performed as part of coffee-making, the dashed trajectory when it was performed as part of another task: tea-making. The resemblance between the two trajectories reflects the fact that the sugar-adding subtask involves the same sequence of stimuli and responses, across the two contexts. The difference between trajectories reflects the fact that the model's internal units maintain information about the overall task context, throughout the course of this subtask.

Figure 3. a) A schematic of action selection in the options framework. On the first time-

step, a primitive action, AI , is selected. On time-step two, an option, OI , is selected, and this option's policy leads to selection of a primitive action, $A2$, followed by selection of another option, $O2$. The policy for $O2$, in turn, selects primitive actions $A3$ and $A4$. The options then terminate, and another primitive action, $A5$, is selected at the topmost level.

b) Inset: The rooms domain from Sutton, Precup and Singh,⁵⁹ as implemented by Botvinick, Niv and Barto.²⁸ S : start. G : goal. Primitive actions include single-step moves in the eight cardinal directions. Options contain policies to reach each door. Arrows show a sample trajectory, involving selection of two options (red and blue arrows) and three primitive actions (black). The plot shows the mean number of steps required to reach the goal, over learning episodes with and without inclusion of the door options.

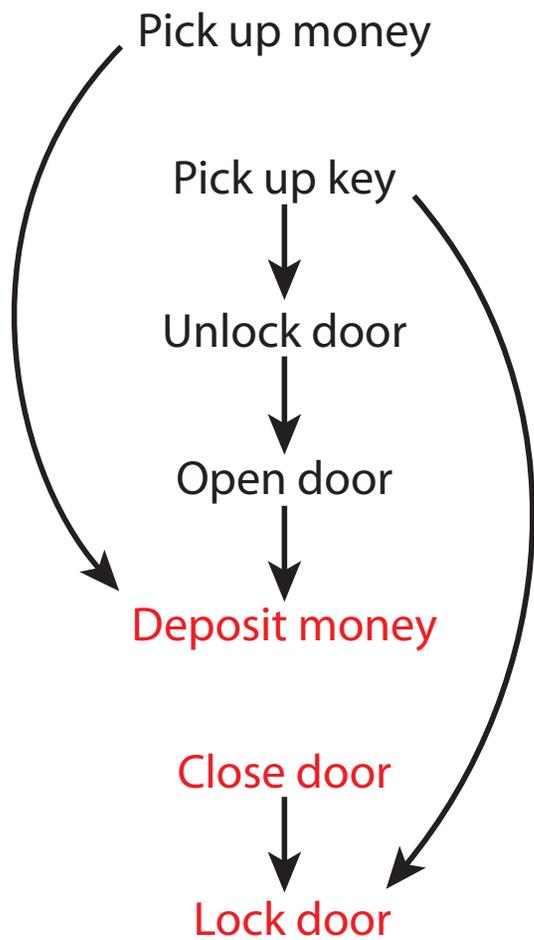
c) An actor-critic implementation of HRL, from Botvinick, Niv and Barto.²⁸ Standard elements are the actor, which implements the policy (π), and the critic, which stores state-values (V), monitors rewards (R), computes reward-prediction errors (δ) and drives learning (see Sutton & Barto⁵⁵). To these is added a new component serving to represent the currently active option (o), which impacts the operation of both actor and critic.

d) Neural correlates of the elements in (c), as proposed by Botvinick, Niv and Barto.²⁸ DA : dopamine; $DLPFC+$: dorsolateral prefrontal cortex, plus other frontal structures potentially including premotor, supplementary motor and presupplementary motor cortices; DLS : dorsolateral striatum; $HT+$: hypothalamus and other structures, potentially including the habenula, the pedunculopontine nucleus, and the superior colliculus; OFC : orbitofrontal cortex; VS : ventral striatum.

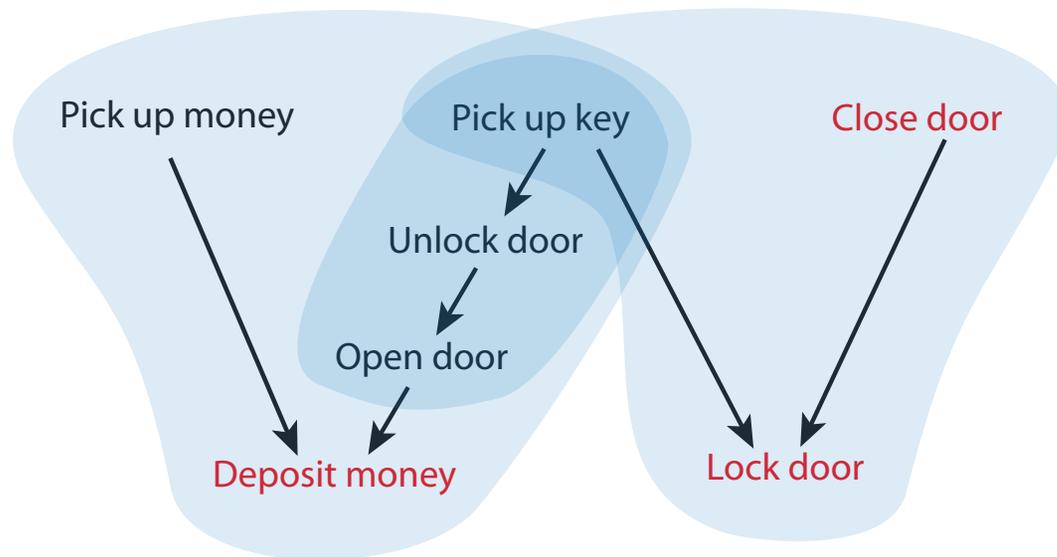
Figure 4. a) The position of the DLPFC within a hierarchy of cortical areas, as described by Fuster.⁶² b) Levels of control represented in different sectors of frontal cortex, according to Koehlin.⁹ Representations become progressively more abstract as one moves rostrally. c) The hierarchically structured network studied by Botvinick,³⁶ showing only a subset of units in each layer. Arrows indicate all-to-all connections. When trained on a hierarchically structured task, units in the apical group spontaneously came to represent context information more strongly than did groups further down the hierarchy. d) Schematic of the gating model proposed by O'Reilly and Frank,¹⁴ during performance of a task requiring maintenance of the stimuli I and A in working memory.

At the point shown, a *I* has already occurred, and has been gated into a prefrontal (*PFC*) stripe via a pathway through the striatum, substantia nigra (*SNr*) and thalamus (*thal*). At the moment shown, an *A* stimulus occurs (*Stim*) and is gated into another *PFC* stripe. Two levels of context are thus represented. e) Koechlin's³⁸ model of FPC function. Orbitofrontal cortex (*Ofc*) encodes the incentive value of various tasks. When two tasks are both associated with a high incentive value, the one with the highest value is selected within lateral prefrontal cortex (*Lpc*) for execution, while the runner-up is held in a pending state by the frontopolar cortex (*Fpc*).

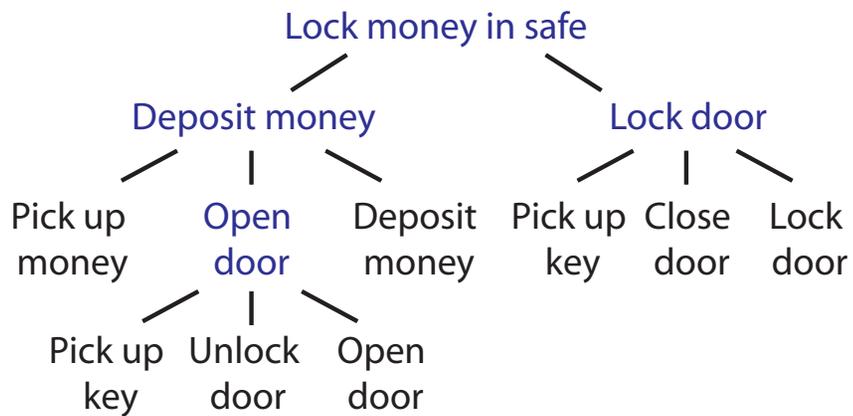
a

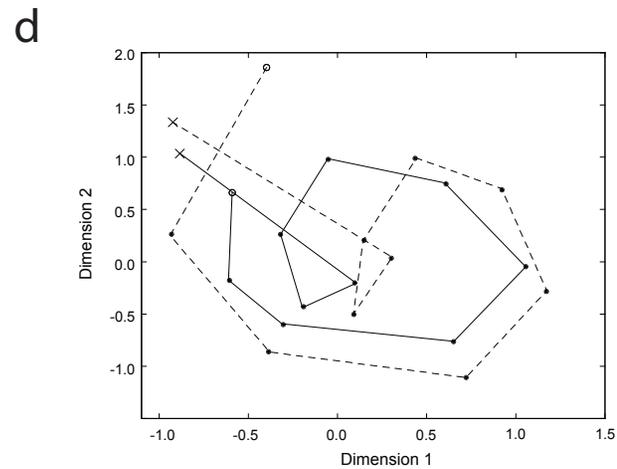
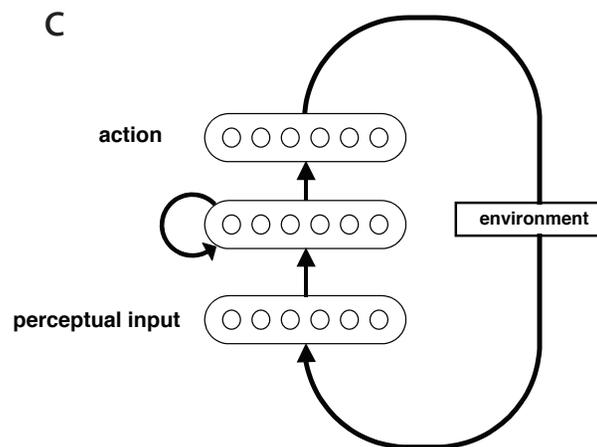
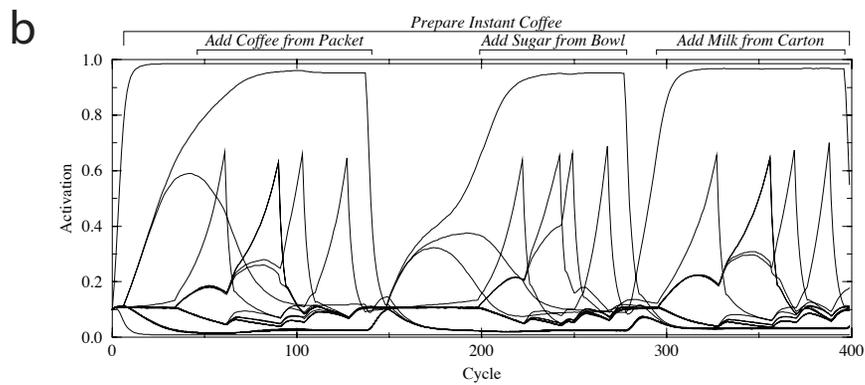
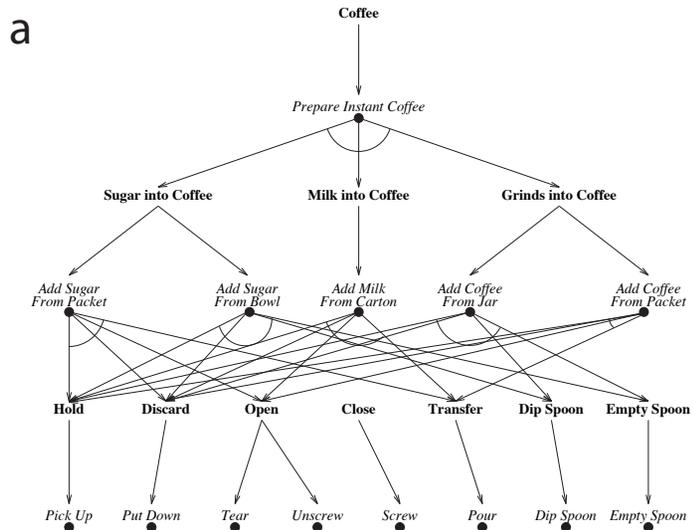


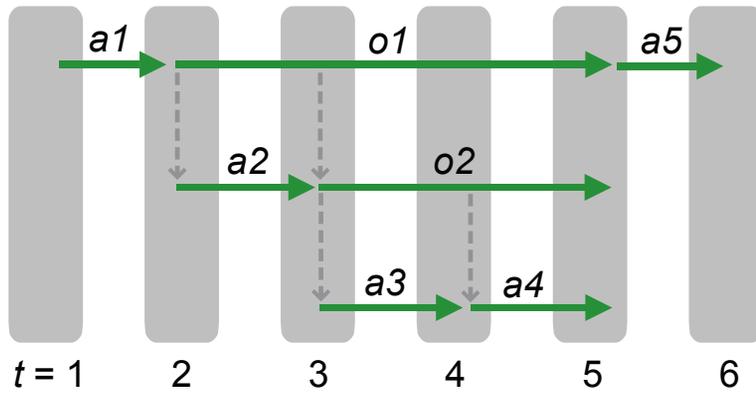
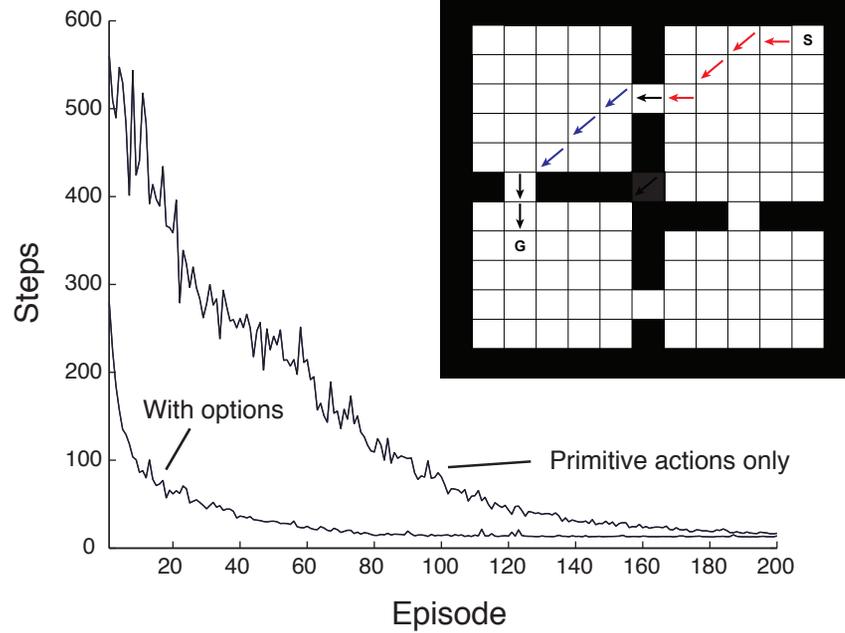
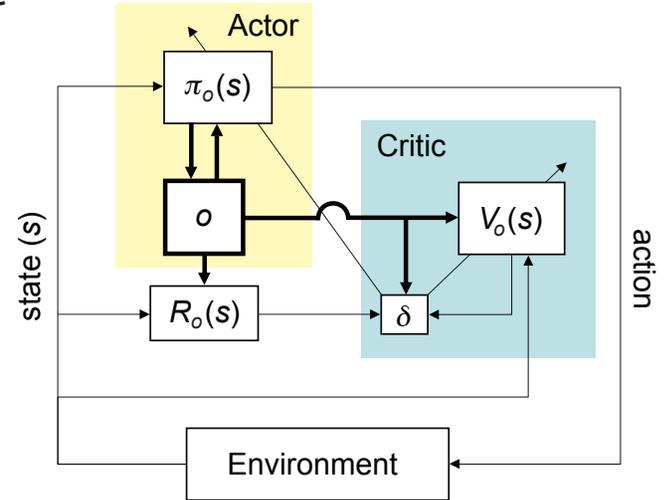
b



c





a**b****c****d**