



# Simple substrates for complex cognition

**Peter Dayan\***

*Gatsby Computational Neuroscience Unit, UCL, London, UK.*

Complex cognitive tasks present a range of computational and algorithmic challenges for neural accounts of both learning and inference. In particular, it is extremely hard to solve them using the sort of simple policies that have been extensively studied as solutions to elementary Markov decision problems. There has thus been recent interest in architectures for the instantiation and even learning of policies that are formally more complicated than these, involving operations such as gated working memory. However, the focus of these ideas and methods has largely been on what might best be considered as automatized, routine or, in the sense of animal conditioning, habitual, performance. Thus, they have yet to provide a route towards understanding the workings of rule-based control, which is critical for cognitively sophisticated competence. Here, we review a recent suggestion for a uniform architecture for habitual and rule-based execution, discuss some of the habitual mechanisms that underpin the use of rules, and consider a statistical relationship between rules and habits.

**Keywords:** prefrontal cortex, striatum, rules, actions, habits

**Edited by:**

Misha Tsodyks, Weizmann Institute of Science, Israel

**Reviewed by:**

Stefano Fusi, Columbia University, USA; ETH & University Zurich, Switzerland  
Florentin Wörgötter, University of Göttingen, Germany

**\* Correspondence:**

Peter Dayan is a Professor of Computational Neuroscience at UCL built mathematical and computational models of neural processing, with a particular emphasis on representation and learning. The main focus is on reinforcement learning and unsupervised learning, covering the ways that animals come to choose appropriate actions in the face of rewards and punishments, and the ways and goals of the process by which they come to form neural representations of the world. The models are informed and constrained by neurobiological, psychological and ethological data.  
dayan@gatsby.ucl.ac.uk

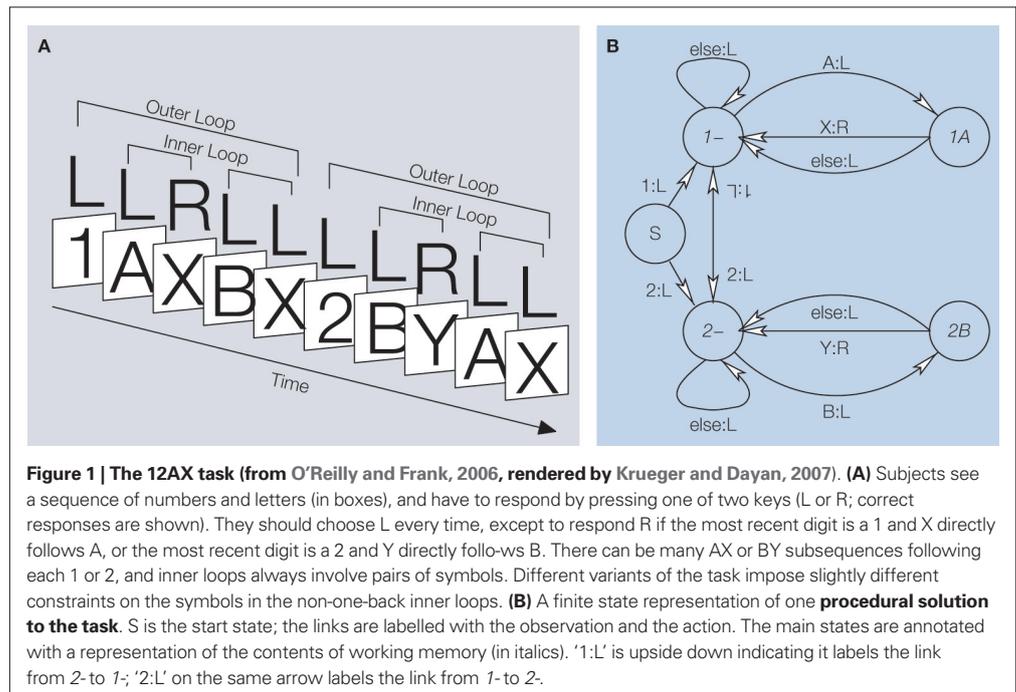
## INTRODUCTION

Figure 1A shows a cognitively complex task (called the 12AX task) which one might attempt to teach to a macaque or explain to a human subject (Frank et al., 2001). The task is called a conditional one-back task, since at its heart are two different one-back tasks. One requires a particular response to X when preceded by A; the other to Y preceded by B. Which of the two tasks is active depends on whether the most recent *digit* seen was a 1 or a 2. Thus the task has a hierarchical structure with the *outer* loop involving the storage of the most recent digit (the 1 or 2), and *inner* loops involving the two one-back tasks. The subjects are required to press one (L) of the two buttons as the particular response for the one-back task that is in operation, and otherwise to press the other button (R). The complexity of the task derives from the hierarchy, and also the long outer loops.

Sophisticated tasks of this sort (Badre et al., 2005; Frank et al., 2001; Fusi et al., 2007; Koechlin and Summerfield, 2007; Koechlin et al., 1999, 2003) pose a variety of challenges for neurally inspired

computational theories in terms of both acquisition and execution. However, despite their neuropsychological and neuroimaging importance, until rather recently, they fell between large cracks in computational cognitive science and neuroscience. They were too simple for the wing of the field interested in realizing artificial intelligence (AI) architectures in connectionist form (e.g. Derthick, 1990; Hinton, 1990; Plate, 2003; Shastri and Ajjanagadde, 1993; Smolensky, 1990; Touretzky and Hinton, 1988; Wermter and Sun, 2000), which largely seeks to represent and manipulate arbitrary symbolic structures. Conversely, they were too complicated to be readily accommodated by the more straightforward feedforward (Fukushima, 1980; Riesenhuber and Poggio, 2000) or recurrent (Hopfield, 1982) processing architectures popular in supervised, unsupervised, and reinforcement learning (Rumelhart and McClelland, 1986; Sutton and Barto, 1998).

However, developments in larger-scale architectures in behavioural neuroscience (Frank et al., 2001; O'Reilly and Frank, 2006; O'Reilly and Munakata, 2000), ideas about more sophisticated



### Procedural representation of a task solution

Crudely, a set of explicit instructions about what actions to take given what input (e.g. 'in state P press the left button; in state Q press the right button'). Standard, imperative, computer programming languages are procedural.

### Rule

A combination of a (typically simple) input–output mapping and a set of conditions defining this mapping's applicability. Rules require retrieval from memory, having their conditions be checked, and then their actions executed. The paper considers a uniform architecture involving habits and rules.

### Goal-directed action

In conditioning, an action that is executed because a subject has good reason to expect it to produce an outcome that is actually desired. This form of control has been related to model-based reinforcement learning, which involves a declarative representation of a task.

### Habit

In conditioning, a response that lacks the contingency or outcome sensitivity of a goal-directed action. This has been related to model-free or cached reinforcement learning. Here, we consider habits as learned, possibly complex, procedural, input–output mappings.

### Working memory

Many complex tasks require information from the recent past to be stored temporarily to influence ongoing choices, and, in reinforcement learning terms, to resolve ambiguities about state. In this paper, we borrow an idea about gated working memory, allowing explicit internally directed actions to control reading information in and out of such memories.

multiply-articulated models of decision-making (Daw et al., 2005), and indeed the new attention being paid to the neural basis of general-purpose cognitive architectures such as ACT-R (Anderson et al., 2008) have breathed new life into the solution of such tasks (Frank et al., 2001; Hazy et al., 2007; Rigotti and Fusi, 2006; Rigotti et al., 2007). Here, we review Dayan's (2007) nascent attempt to meld aspects of these traditions so that a plausible route towards solving the more complex, AI-inspired, tasks can be mapped.

The premise of that paper is that there are procedural approaches to the challenges at two different levels, which we describe as **rules** and **habits**. The actual rules of a task such as 12AX can be described in a few sentences of a natural language (as indeed in the first paragraph of this introduction). More formally, they could easily be rendered as the *productions* of production-system cognitive architectures such as SOAR (Newell, 1990) or ACT-R (Anderson and Lebiere, 1998), employing a blackboard-like **working memory** (e.g. Newell, 1990) to remember facts such as which digit was most recently presented. However, *using* productions of this sort poses particular computational difficulties, including *retrieving* a rule most appropriate to the current state, *checking* that its preconditions actually match the state of the input and the working memory, and then *executing* the rule in the sense of producing an appropriate output action and/or storing aspects of the input in working memory. This process is somewhat analogous

to running an interpreted language on a computer, involving substantial mediation between the rules themselves and their underlying execution.

The whole architecture of the policy for solving this task that underlies the hand-crafted solution in Frank et al. (2001) and the learned solution in Hazy et al. (2006) is completely different from this. It is a single, relatively complex, procedural entity, putatively realized by interactions between basal ganglia and the prefrontal cortex (PFC), and requiring neither retrieval nor checking. Rigotti and Fusi (2006) and Rigotti et al. (2007) considered a similarly complex procedural solution to a slightly different task, but in their case considering hidden units within PFC rather than interactions with the basal ganglia.

Borrowing a term in broad use, but that has been most clearly defined in the animal conditioning community (Dickinson and Balleine, 2002), we call such solutions **habits**, in the sense that they are forms of complex stimulus–response, or conditional input–output mapping, where the stimulus includes both the externally presented letter or digit and the contents of working memory, and the response can be both external, such as pressing a button, and internal, storing a stimulus in working memory. Like conventional habits, these are cognitively impenetrable and inflexible, in that parts of them cannot be edited separately. The main differences from the notion of a conventional habit is that they include the coordinated responses to many stimuli, rather than the

more conventional view which generally (though not exclusively; Balleine et al., 1995) concentrates more on single stimuli and single responses, and that they involve complex actions such as storing a stimulus in working memory. If the use of rules can be likened to running an interpreted, high-level computer language, then the use of a habit is more like the execution of assembly language. Indeed, we briefly discuss two notions of compilation below, also providing a link with popular notions of **automatization** (Anderson, 1982; Anderson and Lebiere, 1998; Anderson et al., 2004; Crossman, 1959; Fitts, 1964; Logan, 1988; Newell, 1990; Newell and Rosenbloom, 1981; Taatgen et al., 2008).

In conditioning, the normal opposites of habits are **goal-directed actions** (Dickinson and Balleine, 2002). However, rules are procedural, whereas goal-directed actions are *declarative*, a term with resonance in both neuroscience (Eichenbaum and Cohen, 2004) and computer science (as in programming languages such as PROLOG). Goal-directed methods are declarative since they specify a set of facts about a domain (understood by Daw et al., 2005 in terms of reinforcement learning as a form of forward model; Wolpert and Kawato, 1998; Wolpert and Miall, 1996) leaving out the hard deductive task of figuring out from these facts (e.g. searching the forward model) to determine what it is best to do. We consider issues surrounding declarative control towards the end of the paper. Nevertheless, like goal-directed control, the system of rules can be highly labile to new information about a task (by virtue of changing the contents of the store of rules). This is one of their most critical differences from habits.

Dayan's (2007) intent was to build a model that provides two explicit links between rules and habits. One concerns the relationship between the execution component of each. For tasks such as 12AX, rule execution involves exactly the same class of internal and external effects as a habit. It is just that the complexities of the stimulus-response mapping have largely been removed by the processes of rule matching and checking, leaving a simple stimulus-response map. In fact, depending on the exact form of the rule, the process of checking might obviate the need for any conditionality even on the stimulus. The commonality between rules and habits is important because it suggests that a rather simple architecture might implement parts of both, and also provides a route towards seeing how a set of verbal commands could ultimately translate into a set of activities in populations of neurons that actually execute them<sup>1</sup>.

<sup>1</sup>Dayan (2007) labels this process compilation, but does *not* implement it.

We consider a uniform architecture, in which rule execution indeed involves just the same operations as a habit. This then leads to a set of issues about the structural and statistical relationships between the execution components of rules and habits, and how one might emerge from the other. Again arguing by analogy, unsupervised learning methods applied to the natural statistics of sensory inputs build models that use basis representations to split up complex perceptual scenes into their separate pieces (Rao et al., 2002). Top-down effects in perception depend on the ability to use these models to synthesize the internal activities associated with new, imagined, scenes. We consider the possibility that there could be habitual basis functions in the space of policies that can be acquired in a similar manner, and that the equivalent of synthesizing a new scene is synthesizing the execution aspect of a rule.

The other link between rules and habits stems from the observation that the tasks involved in interpreting rules themselves involve the execution of some particular internally directed actions, including memory retrieval for the rule (which we consider to be associative) and condition checking. In our computer analogy, interpretation of a high-level language (the rules) requires execution of assembly language (the habits) associated with the interpreter. We thus consider some of the habits associated with rule use.

We first consider the habitual elements, crudely, the form of processor that underlies performance in such tasks. We then consider rules, and finally generalizations beyond the work in Dayan (2007).

## HABITS

The 12AX task (and indeed the other conditional input–output mapping task modelled in detail in Dayan, 2007) poses two particular computational problems associated with: (1) working memory (for stimuli 1, 2, A and B), and (2) the mappings from the current input to the required output conditional on the contents of the working memory. Both of these can be solved using particular forms of non-linear interactions based on simple neural representations of the input stimuli and working memory. **Figure 1B** shows a habitual solution to this task in a finite state representation. Working memory (annotated inside the nodes) is used to realize the states; at least two 'slots' are needed, one for 1 or 2, and one for A or B. The conditional input–output mappings realize the actions and the transitions. Frank et al. (2001), O'Reilly and Frank (2006), Rigotti and Fusi, (2006) and Rigotti et al. (2007) suggest rather different approaches to implementing

### Automatization

In conditioning, behaviour can transfer from being under goal-directed control to being habitual, and therefore automatic in a particular sense. Related and expanded notions of automatization in skill learning appear in many forms in the literature, including a transfer from declarative to procedural representations, and streamlined procedural forms.

such a solution, which were partially integrated in Dayan (2007).

Based on the analysis of Hochreiter and Schmidhuber (1997), Frank et al. (2001) and O'Reilly and Frank (2006) build a model of gated working memory called PBWM (prefrontal, basal ganglia working memory). The idea is that regions (stripes; Pucak et al., 1996) in the PFC are disposed to provide the sort of activity-based storage frequently observed in electrophysiological studies of neurons in this area (Fuster, 1997), for specific information available in inputs from higher order sensory areas. In our case, this on-line storage is the substrate for the representation of the slots, i.e., the persistent states associated with the inner and outer loops in the task. In this framework, control over working memory amounts to control over reading information into (and out of; Gers et al., 2000) the memory. O'Reilly and Frank (2006) suggest that this is implemented by plastic connections between the PFC and the basal ganglia that learn to implement appropriate gating operations at suitable points in the performance of the task. Other connections through the basal ganglia realize the mapping necessary to generate motor outputs. Frank et al. (2001) show that it is possible to wire up such a gated working memory system in a way that executes the 12AX task; O'Reilly and Frank (2006) provide a rather complex reinforcement learning mechanism that can acquire this ability from (substantial) experience. However, the slowness of learning may be at least partially ameliorated through structured *shaping* of competence, itself an underinvestigated topic (Krueger and Dayan, 2007).

Gating information into or out of working memory is a particular sort of non-linear computational operation. In PBWM, one can interpret the neurons in the basal ganglia pathway as hidden units with trainable inputs (to accommodate the different operations that are required) and fixed outputs (i.e. gating particular pathways from sensory cortices to PFC) associated with this non-linearity. By contrast, although Rigotti and Fusi (2006) and Rigotti et al. (2007) also consider working memory attractors as the core substrate of PFC, they employ hidden units that are intended to form part of PFC itself, with random, fixed, inputs, but trainable outputs. They do not consider gating operations as a particular, separate, class, with their own specific (striato-thalamo-cortical) mechanism, instead treating them like any other computational operation. Rigotti and Fusi (2006) and Rigotti et al. (2007) show first that many simple computational operations associated with conditional input-output mappings are not linearly-separable, and so can-

not be performed without such hidden units, and then that, given only a modest number of random, fixed, hidden units, an impressively large fraction of all possible single operations becomes viable.

Both these ideas can be thought of in terms of procedural primitives, out of which more general competence can be fashioned, for instance through learning. Dayan (2007) adopted the notion of gating from Frank et al. (2001), with explicit operations for reading into and out of working memory. However, rather than employing the random hidden units of Rigotti and Fusi (2006) and Rigotti et al. (2007), this model used a bilinear form (Koenderink and Van Doorn, 1997; Tenenbaum and Freeman, 2000), which is a rather simpler structure that is an abstraction of the sort of multidimensional basis functions that are frequently considered in neural terms (Poggio, 1990), particularly in the context of coordinate transformations in the parietal cortex (Olshausen et al., 1993; Poggio, 1990; Pouget, 1997). This overall habitual architecture (Figure 1A,B in Dayan, 2007, with the solution to the 12AX task being shown in Figure 7) provides a computational substrate that is rich enough to encompass a whole range of tasks. One important question for future work is whether this is formally broad enough, i.e., in some sense Turing equivalent (bar the infinite tape), possibly when coupled to an ability to create new representations of combined sensory input and internal states that are also coupled to their own working memory units.

The target article considered this part of the architecture as being habitual, because it realizes a direct mapping between internal and external stimuli and internal and external responses. In the literature from which this term was borrowed, the distinction is made between goal-directed actions, which are performed because (computationally, as a deductive consequence of the fact that) subjects have good reason to (1) expect that they will give rise to particular outcomes; and (2) value those outcomes, and habits, which form a more amorphous class of control, including stimulus-response mappings, that do not respect one or other of these precepts (Dickinson, 1985; Dickinson and Balleine, 2002). For instance, if a subject learns that a formerly valuable outcome is currently worthless (for instance having been poisoned), under goal-directed control, it should refuse to perform an action leading to that outcome. The policies realized by the bilinear architecture are not based in any way on the sort of reasoning processes inherent to goal-directed actions, and cannot be changed in the sort of on-line way implied by the poisoning experiment. This underlies their habitual moniker.

## RULES

Another way of viewing the habitual solution is that it is automatic, in the terms of Chaiken and Trope (1999), Dickinson (1985), Epstein (1994), Evans (2003), Kahneman and Frederick (2002), Schneider and Shiffrin (1977), Shiffrin and Schneider (1984), Sloman, (1996) and Stanovich and West (2002). These authors all consider dual process theories, setting up different, but inter-related dichotomies between different systems. One of these, which involves the equivalent of habits, is typically inflexible and slow to learn, but fast to execute and attentionally undemanding. The other, 'System 2' in the sense of Kahneman and Frederick (2002) and Stanovich and West (2002), is rule-based, and subject to something more akin to attentional or deliberate control<sup>2</sup>. We now turn to this flexible, but demanding and computationally expensive (and hence likely inaccurate, Daw et al., 2005) system.

In humans, rules for this controlled processor can be provided through verbal instructions, which can obviously only impact behaviour by affecting the underlying mechanisms of choice, putatively, exactly the same elements of working memory and conditional input–output control as above. We think of the mapping from precept to procedure as a problem of *compilation* or translation, an area that is ripe for further investigation. However, at a minimum, there must be a set of pre-existing execution competences that can be the targets of such mappings. In our uniform architecture, we consider these competences to be simple habits themselves, and suggest that simple habits form a basis set for the rule-governed execution of tasks.

The first step is to consider what rules might look like for a task such as 12AX. One view of a rule is that it captures an appropriate procedure for a restricted piece of the task, as in a (formalized) part of the description given at the start of this paper. Examples of this include requiring that 1 or 2 be stored in working memory when it occurs (rules 1 and 2 of Table 4 in Dayan, 2007), or that if 1A is stored and is followed by an X then the R button should be pressed, and the A (but not the 1) forgotten. This latter rule (rule 5 in the table) has a limited domain (i.e. only applying at a restricted set of steps of the trial). A whole collection of rules has to be combined to create performance that is appropriate overall. Different sets of rules can be computationally equivalent, just as different algorithms rendered in different programming language can realize the same

computation in different ways (Marr, 1982), posing greater or lesser loads on different aspects of neural competence.

Cast in this manner, the action part of a rule is really an isolated, basic procedure that can be implemented using the same mechanism that realises the complete, complex, habit for 12AX (that we discussed in Section 'Habits'). This uniform realization is very important, since it provides a natural means of melding rules and habits in the same structure. We can imagine a whole task such as 12AX ultimately becoming the execution component of a rule all on its own, illustrating a compositional route to more sophisticated functionality.

Of course, the rule itself can only be simple if the complexity of extracting it from the store of rules that collectively solve the task, and deciding which of this collection should be active at what point in the task, is addressed by some other mechanism. We argue that rule storage and extraction is the job for a content-addressable episodic memory; and testing the exact conditions for the match of a rule is a specialized, but bilinearly simple, procedure all of its own, based on matching conditions that form a critical part of the representation of the rule.

The rules describing a complex task, like the rules describing a board game, can readily exceed the limited explicit storage capacity of working memory (Halford et al., 2007). It therefore seems likely that they are stored in long-term memory, extracted by associative match with the sensory input and the contents of working memory. However, there is a critical computational difference between associative matching and rule matching that has to do with exclusion. Rules can have exclusion conditions – for instance firing only if some aspect of the input does *not* match a particular feature, whereas associative recall is generally more promiscuous, using any available information as a key to recall. It is straightforward to specify a bilinear matching process that tests inclusion and exclusion conditions associated with a rule; this can be seen as another specialized procedure or habit itself.

In sum, we have created all the elements for rule storage, recall, testing and execution. The execution component of rules are, just like habits, stimulus-response mappings, but with the conditionality that plays an important part in the full habitual representation of a whole task such as 12AX being replaced by an explicit test of a set of conditions that the input must meet, or must fail to meet. In Dayan (2007), these conditions are simple to enforce, since the representations of inputs and the contents of working memory

<sup>2</sup>Although, of course, it is not clear how similar are all these different dichotomies, which span different species and different computational problems.

are unary. We assume that the associative matching procedure can recall a number of candidate rules; the testing process winnows this down to the one that is appropriate; its associated habit is then executed in exactly the same (bilinear) manner as the overall habit for 12AX.

Rule-based execution of a task is more computationally involved than habitual execution, because of the sequential process of retrieving, testing and realizing the procedures associated with multiple rules. This computational challenge, and the resulting possibility of error, provides a spur towards automatization and the learning and acquisition of habitual competence.

### PROSPECTS

The target article Dayan (2007) contains an extensive discussion of the gaps in, and immediate extensions of, this work, notably to cases in which multiple rules may be appropriate at any point in a task, and to more complex matching conditions. However, two issues merit more notice, namely the relationships between habitual, rule-based, and declarative control, and the provenance of the basis sets of rules.

Our work was partly inspired by the analysis of the difference between goal-directed and habitual control (Dickinson and Balleine, 2002) in terms of model-based and model-free reinforcement learning (Daw et al., 2005). Both forms of reinforcement learning attempt to acquire optimal policies. Model-based mechanisms do it by building a model of the world, which is typically straightforward and can readily change with new information. However, using such a model requires its *inverse*. For instance, for the case of a Markov decision problem, this requires taking a description of the transition structure of a domain and outcome utilities, and performing an operation equivalent to dynamic programming to calculate the optimal actions (Sutton and Barto, 1998). This amounts to a form of search in the tree that comprises the transitions and outcomes produced by actions. Search is computationally expensive, taxing on working memory, and error-prone. Daw et al., (2005) suggested that these challenges motivated model-free mechanisms, which *learn* the inverse by substituting experience and sampling for search. This makes them slower to acquire optimal behaviour, and less adaptable, but also much easier to use in practice.

However, the requirement for tree search points out an important distinction between model-based and rule-based control. The most natural form of model is a forward or generative (Dayan et al., 1995; Wolpert and Miall, 1996) account of the domain that specifies

the transitions and outcomes that result from actions. However, this is a declarative form of representation of a problem, and hides the (search) complexities of how this knowledge can be used to work out what to do.

There are at least two straightforward analogies for this. One is with logic or theorem proving – the model is like a set of axioms, the task of working out what to do is like deducing a consequence that is true. Deciding which of the declarative facts to use in this deduction and in what order, is well known to be a tough computational problem. A second analogy is to generative (or synthesis) and recognition (or analysis) models in statistical unsupervised learning for sensory input (Dayan et al., 1995; Hinton and Zemel, 1994; Hinton et al., 1995; Kawato et al., 1993; Neisser, 1967). Consider the case of vision – the generative model is a form of graphics model that indicates what a scene would look like given that it contains certain objects, lit in particular ways and viewed from some angle. The recognition model takes a view of a scene, and reverse-engineers it to work out what objects and lighting generated it. This is also well known as a computationally challenging inverse problem; and indeed the recognition model is often seen as the (Bayesian statistical) inverse of the generative model.

For us, the rules of the task are actually themselves already an inverse, i.e., they already determine an effective procedure for executing the task successfully. One can also imagine a set of rules that is more **declarative** in form – a set of generative rules which describes the workings of a task without directly indicating its solution (that is, a generative model without an inverse). Indeed, by contrast with imperative programming languages such as BASIC, there are declarative languages like PROLOG which require inverting in a somewhat analogous manner, employing a similar form of tree search. Our architecture should generalize to cases in which tree search is required based on a declarative specification of a task; this would then combine goal-directed and rule-based aspects of cognitive competence.

Despite being procedural rather than declarative, the sort of rules that we have considered do share one important characteristic with goal-directed actions, namely that they can readily be changed. Depending on the flexibility of storage and controlled forgetting in the associative memory, it would be straightforward to add new rules, and erase old ones and have them work appropriately, without changing everything about the existing rule-base. This is impossible for habits. Of course, if sets of rules come to conflict as they change, behaviour will be confused – declarative

#### Declarative representation of a task solution

Crudely, this describes the contingencies of the task (e.g. 'pressing the left button in state P leads to state Q; pressing the right button in state Q leads to reward') leaving a search problem to work out the appropriate action. Programming languages such as PROLOG are declarative; in the text, we relate such representations to model-based control.

representations face a similar concern in the form of the infamous frame problem (McCarthy and Hayes, 1969).

In conditioning, behaviour that is initially goal-directed, emanating from the model-based system, often becomes habitual, controlled in a model-free manner, over the course of experience. This is related to the notions of automatization cited above. One difference is that the distinction in conditioning between goal-directed and habit systems has a somewhat sharper operationalization than for the other cases. There is an active debate in the field about the nature of the transfer from the former to the latter, for instance, whether it is for reasons of decreased perceived covariance between action and outcome arising from overly regular choices (Dickinson, 1985), or the relative uncertainties of the two systems arising from their different statistical and computational profiles (Daw et al., 2005). The characterizations of transfer in the other treatments of automatization again appear somewhat less formal than these. However, Dayan (2007) focussed on the representation of policies rather than their evolution over the course of skill learning; the latter is an important direction for the future.

Note that Anderson (1982); Taatgen et al. (2008), building on the notion of 'chunking' in SOAR (Newell and Rosenbloom, 1981), consider that one aspect of automatization is turning a declarative representation of a task that is solved (inverted) by general-purpose procedures into a rule-based representation in terms of productions. They call this knowledge or procedural compilation, which is slightly different from the compilation step that Dayan (2007) suggested as being involved in the translation from a verbal to a rule-based specification of the policy for a task.

Our account stresses an additional feature of the interaction between the systems, namely the dependence of rule-based control on habitual mechanisms. First, the action components of rules are seen as being just like simple habits. Second, rule matching can be performed by a bilinear computational substrate just like the one assumed for the execution of other habits. Rules influence other rules directly by changing the contents of working memory (which acts like a form of blackboard; Newell, 1990), and indirectly by causing actions in the world, which may then change (observable) state. In principle, the cognitive operations associated with rule-based control, such as influencing or manipulating associative matching, could themselves be rule-based or habitual. However, this recursion must in the end be habitually grounded. The operations

required to invert a declarative specification of a task would be yet more complicated.

We have argued that rules employ a basis set of habitual primitives. This poses an obvious question as to the origin and choice of these primitives. Much of the work on connectionist structure-sensitive rule and symbol processing (Hinton, 1990; Plate, 2003; Smolensky, 1990) has focussed on universality, i.e. on the implementation of a small set of powerful operations that can be combined and manipulated in a nearly arbitrary manner. By contrast, motivated by notions about structured visual representations (Dayan, 2006), we suggested considering this basis set in an hierarchical manner as the result of an unsupervised representational learning process (Rao et al., 2002) applied to the internal descriptions of habitual procedures, just like the internal descriptions in early sensory cortices of visual, auditory or other sensory input that arise in the paired generative and recognition models that we discussed above.

The representational learning process that creates these paired models involves taking an ensemble of patterns and determining and parameterizing their underlying statistical structure. Given this, (old and) new members of the ensemble can be represented according to their parameter values with respect to this structural decomposition. This representation is the product of the recognition or analytical model. Along with this model comes the statistical generative model that can synthesize new instances of the ensemble. Unsupervised learning applied to the internal representations of habits would decompose complex ones into their individual parts, and provide a mechanism by which new habits, and the execution components of new rules, can be synthesized from their elemental pieces. One can imagine basing learning of a sophisticated representational system for habits, which will be automatically appropriate for rules, on the acquisition over the course of a developmental period of a whole range of different tasks.

## CONCLUSIONS

The main goal of Dayan (2007) was to provide a uniform architecture into which habits and instructed, rule-based, competence both fit. Although the former provides the ultimate, automatic, means by which we and animals negotiate simple and complex tasks, instructed rule-based performance is very important, at least for humans. Further, the issues for instantiating rules, such as the notion of hierarchically specified basis sets of habits created by a representational learning procedure, are important even in the absence of instruction. The other point stressed

in the paper is the rather strong reliance of rule instantiation and execution on habits. Not only is the action component of a rule a habit, but also we have suggested that mechanisms such as matching that are associated with using rules, depend on habits too. At a most fundamental level, it is habits that are in control.

## ACKNOWLEDGEMENTS

I am very grateful to David Badre, Paul Burgess, Nick Chater, Mark D'Esposito, Stefano Fusi, Kai Krueger, Tim Shallice and Wako Yoshida for discussion, models and ideas associated with the original article. Funding was from the Gatsby Charitable Foundation.

## REFERENCES

- Anderson, J. R. (1982). Acquisition of cognitive skill. *Psychol. Rev.* 89, 369–406.
- Anderson, J. R., Bothell, D., Byrne, M. D., Douglass, S., Lebiere, C., and Qin, Y. (2004). An integrated theory of the mind. *Psychol. Rev.* 111, 1036–1060.
- Anderson, J. R., Fincham, J. M., Qin, Y., and Stocco, A. (2008). A central circuit of the mind. *Trends. Cogn. Sci.* 12, 136–143.
- Anderson, J. R., and Lebiere, C. (1998). The Atomic Components of Thought. Mahwah, NJ, Erlbaum.
- Badre, D., Poldrack, R. A., Paré-Blagoev, E. J., Insler, R. Z., and Wagner, A. D. (2005). Dissociable controlled retrieval and generalized selection mechanisms in ventrolateral prefrontal cortex. *Neuron* 47, 907–918.
- Balleine, B., Garner, C., Gonzalez, F., and Dickinson, A. (1995). Motivational control of heterogeneous instrumental chains. *J. Exp. Psychol.* 21, 203–217.
- Chaiken, S., and Trope, Y. (1999). Dual-Process Theories in Social Psychology. New York, NY, Guilford Press.
- Crossman, E. (1959). A theory of the acquisition of speed-skill. *Ergonomics* 2, 153–166.
- Daw, N. D., Niv, Y., and Dayan, P. (2005). Uncertainty-based competition between prefrontal and dorsolateral striatal system for behavioral control. *Nat. Neurosci.* 8, 1704–1711.
- Dayan, P. (2006). Images, frames, and connectionist hierarchies. *Neural Comput.* 18, 2293–2319.
- Dayan, P. (2007). Bilinearity, rules and prefrontal cortex. *Front. Comput. Neurosci.* 1, 1.
- Dayan, P., Hinton, G. E., Neal, R. M., and Zemel, R. S. (1995). The Helmholtz machine. *Neural Comput.* 7, 889–904.
- Derthick, M. (1990). Mundane Reasoning by Parallel Constraint Satisfaction. Ph.D. thesis, Dept of Computer Science, Carnegie-Mellon, Pittsburgh, PA.
- Dickinson, A. (1985). Actions and habits: the development of behavioural autonomy. *Philos. Trans. R. Soc. Lond., B, Biol. Sci.* 308, 67–78.
- Dickinson, A., and Balleine, B. (2002). The role of learning in motivation. In Stevens' Handbook of Experimental Psychology, Vol. 3, C. Gallistel, ed. (New York, NY, Wiley), pp. 497–533.
- Eichenbaum, H., and Cohen, N. J. (2004). From Conditioning to Conscious Recollection: Memory Systems of the Brain (Oxford Psychology Series). New York, NY, Oxford University Press.
- Epstein, S. (1994). Integration of the cognitive and the psychodynamic unconscious. *Am. Psychol.* 49, 709–724.
- Evans, J. S. B. T. (2003). In two minds: dual-process accounts of reasoning. *Trends. Cogn. Sci.* 7, 454–459.
- Fitts, P. (1964). Perceptual-motor skill learning. In Categories of Human Learning, A. Melton, ed. (New York, NY, Academic Press), pp. 243–285.
- Frank, M. J., Loughry, B., and O'Reilly, R. C. (2001). Interactions between frontal cortex and basal ganglia in working memory: a computational model. *Cogn. Affect. Behav. Neurosci.* 1, 137–160.
- Fukushima, K. (1980). Neocognition: a self organizing neural network model for a mechanism of pattern recognition unaffected by shift in position. *Biol. Cybern.* 36, 193–202.
- Fusi, S., Assad, W. F., Miller, E. K., and Wang, X. J. (2007). A neural circuit model of flexible sensorimotor mapping: learning and forgetting on multiple timescales. *Neuron* 54, 319–333.
- Fuster, J. M. (1997). The Prefrontal Cortex: Anatomy, Physiology, and Neuropsychology of the Frontal Lobe, 3rd edn. New York, NY, Raven Press.
- Gers, F. A., Schmidhuber, J., and Cummins, F. (2000). Learning to forget: continual prediction with LSTM. *Neural Comput.* 12, 2451–2471.
- Halford, G. S., Cowan, N., and Andrews, G. (2007). Separating cognitive capacity from knowledge: a new hypothesis. *Trends Cogn. Sci.* 11, 236–242.
- Hazy, T. E., Frank, M. J., and O'Reilly, R. C. (2006). Banishing the homunculus: Making working memory work. *Neuroscience* 139, 105–118.
- Hazy, T. E., Frank, M. J., and O'Reilly, R. C. (2007). Toward an executive without a homunculus: computational models of the prefrontal cortex/basal ganglia system. *Philos. Trans. R. Soc. Lond., B, Biol. Sci.* 362, 1601–1613.
- Hinton, G. (1990). Preface to the special issue on connectionist symbol processing. *Artif. Intell.* 46, 1–4.
- Hinton, G. E., Dayan, P., Frey, B. J., and Neal, R. M. (1995). The wake-sleep algorithm for unsupervised neural networks. *Science* 268, 1158–1161.
- Hinton, G. E., and Zemel, R. S. (1994). Autoencoders, minimum description length, and Helmholtz free energy. *Adv. Neural Inf. Process. Syst.* 6, 3–10.
- Hochreiter, S., and Schmidhuber, J. (1997). Long short-term memory. *Neural Comput.* 9, 1735–1780.
- Hopfield, J. J. (1982). Neural networks and physical systems with emergent collective computational abilities. *Proc. Natl. Acad. Sci. USA* 79, 2554–2558.
- Kahneman, D., and Frederick, S. (2002). Representativeness revisited: attribute substitution in intuitive judgment. In Heuristics of Intuitive Judgement: The Psychology of Intuitive Judgement, T. Gilovich, D. Griffin and D. Kahneman, eds (New York, NY, CUP), pp. 49–81.
- Kawato, M., Hayakawa, H., and Inui, T. (1993). A forward-inverse optics model of reciprocal connections between visual cortical areas. *Netw. Comput. Neural Syst.* 4, 415–422.
- Koehlin, E., Basso, G., Pietrini, P., Panzer, S., and Grafman, J. (1999). The role of the anterior prefrontal cortex in human cognition. *Nature* 399, 148–51.
- Koehlin, E., Ody, C., and Kouneiher, F. (2003). The architecture of cognitive control in the human prefrontal cortex. *Science* 302, 1181–1185.
- Koehlin, E., and Summerfield, C. (2007). An information theoretical approach to prefrontal executive function. *Trends Cogn. Sci.* 11, 229–235.
- Koenderink, J. J., and Van Doorn, A. J. (1997). The generic bilinear calibration-estimation problem. *Int. J. Comput. Vis.* 23, 217–234.
- Krueger, K. A., and Dayan, P. (2007). Flexible shaping: how learning in small steps helps. In Poster II-41. COSYNE 2007.
- Logan, G. (1988). Toward an instance theory of automatization. *Psychol. Rev.* 95, 492–527.
- Marr, D. (1982). *Vision*. New York, NY, Freeman.
- McCarthy, J., and Hayes, P. (1969). Some philosophical problems from the standpoint of artificial intelligence. In Machine Intelligence, Vol. 4, B. Meltzer and D. Michie, eds (Edinburgh, Edinburgh University Press), pp. 463–502.
- Neisser, U. (1967). *Cognitive Psychology*. New York, NY, Appleton-Century-Crofts.
- Newell, A. (1990). *Unified Theories of Cognition*. Cambridge, MA, Harvard University Press.
- Newell, A., and Rosenbloom, P. (1981). Mechanisms of skill acquisition and the law of practice. In *Cognitive Skills and their Acquisition*, J. Anderson, ed. (Hillsdale, NJ, Lawrence Erlbaum Associates), pp. 1–56.
- Olshausen, B. A., Anderson, C. H., and Van Essen, D. C. (1993). A neurobiological model of visual attention and invariant pattern recognition based on dynamic routing of information. *J. Neurosci.* 13, 4700–4719.
- O'Reilly, R. C., and Frank, M. J. (2006). Making working memory work: a computational model of learning in the prefrontal cortex and basal ganglia. *Neural Comput.* 18, 283–328.
- O'Reilly, R. C., and Munakata, Y. (2000). *Computational Explorations in Cognitive Neuroscience: Understanding the Mind by Simulating the Brain*. Cambridge, MA, MIT Press.
- Plate, T. (2003). *Holographic Reduced Representation: Distributed Representation for Cognitive Structures*. Stanford, CA, CSLI Publications.
- Poggio, T. (1990). A theory of how the brain might work. *Cold Spring Harb. Symp. Quant. Biol.* 55, 899–910.
- Pouget, A. (1997). Spatial transformations in the parietal cortex using basis functions. *J. Cogn. Neurosci.* 9, 222–237.
- Pucak, M. L., Levitt, J. B., Lund, J. S., and Lewis, D. A. (1996). Patterns of intrinsic and associational circuitry in monkey prefrontal cortex. *J. Comp. Neurol.* 376, 614–630.
- Rao, R., Olshausen, B., and Lewicki, M. (2002). Probabilistic Models of the Brain: Perception and Neural Function. Cambridge, MA, MIT Press (Bradford Books).
- Riesenhuber, M., and Poggio, T. (2000). Models of object recognition. *Nat. Neurosci.* 3(Suppl.), 1199–1204.
- Rigotti, M., Ben Dayan Rubin, D., Wang, X.-J., and Fusi, S. (2007).

- The importance of neural diversity in complex cognitive tasks. In Society for Neuroscience Annual Meeting. Washington, DC, Society for Neuroscience.
- Rigotti, M., and Fusi, S. (2006). Modeling event drives transitions between attractors of the neural dynamics. In Society for Neuroscience Annual Meeting. Washington, DC, Society for Neuroscience, p 569.14/LL5.
- Rumelhart, D., and McClelland, J. (1986). *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*. Cambridge, MA, MIT Press.
- Schneider, W., and Shiffrin, R. (1977). Controlled and Automatic Human Information Processing: I. Detection, search, and attention. *Psychol. Rev.* 84, 1–66.
- Shastri, L., and Ajjanagadde, V. (1993). From simple associations to systematic reasoning: a connectionist representation of rules, variables, and dynamic bindings using temporal synchrony. *Behav. Brain Sci.* 16, 417–494.
- Shiffrin, R. M., and Schneider, W. (1984). Automatic and controlled processing revisited. *Psychol. Rev.* 91, 269–276.
- Slooman, S. A. (1996). The empirical case for two systems of reasoning. *Psychol. Bull.* 119, 3–22.
- Smolensky, P. (1990). Tensor product variable binding and the representation of symbolic structures in connectionist systems. *Artif. Intell.* 46, 159–216.
- Stanovich, K., and West, R. (2002). Individual differences in reasoning: implications for the rationality debate? In *Heuristics and Biases: The Psychology of Intuitive Judgment*, T. Gilovich, D. Griffin and D. Kahneman, eds (New York, NY, Cambridge University Press), pp. 421–440.
- Sutton, R. S., and Barto, A. G. (1998). *Reinforcement Learning: An Introduction* (Adaptive Computation and Machine Learning). Cambridge, MA, MIT Press.
- Taatgen, N. A., Huss, D., Dickison, D., and Anderson, J. R. (2008). The acquisition of robust and flexible cognitive skills. *J. Exp. Psychol. Gen.* 137, 548–565.
- Tenenbaum, J. B., and Freeman, W. T. (2000). Separating style and content with bilinear models. *Neural Comput.* 12, 1247–1283.
- Touretzky, D., and Hinton, G. (1988). A distributed connectionist production system. *Cogn. Sci.* 12, 423–466.
- Wermter, S., and Sun, R. (2000). *Hybrid neural systems*. Heidelberg, Springer.
- Wolpert, D., and Kawato, M. (1998). Multiple paired forward and inverse models for motor control. *Neural Netw.* 11, 1317–1329.
- Wolpert, D., and Miall, R. (1996). Forward models for physiological motor control. *Neural Netw.* 9, 1265–1279.

**Conflict of Interest Statement:** The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 11 October 2008; accepted: 07 November 2008; published: 15 December 2008.

Citation: *Front. Neurosci.* (2008) 2, 2: 255–263. doi: 10.3389/neuro.01.031.2008

Copyright: © 2008 Dayan. This is an open-access article subject to an exclusive license agreement between the authors and the Frontiers Research Foundation, which permits unrestricted use, distribution, and reproduction in any medium, provided the original authors and source are credited.

