been invoked to explain. In Karmiloff-Smith's (1992) Representational Redescription Hypothesis, a layer of representations can be accessed at will [see multiple book review of "Beyond Modularity" *BBS* 17(4) 1994]. In our account, a layer of *mechanisms* can be accessed at will. A standard example of representational redescription is when data stored as a stack are redescribed as a list. (Items in a stack can only be accessed from the top, but items in a list can be accessed in any order.) Suppose that at some point in our evolution we could only access items or perform actions in a stack-like sequence, and later we learned to access or perform them in any sequence. Now we *could* say that we learned to represent the world differently, and that this in turn gave us better access to it, but it is simpler to say that we learned to access it more efficiently. This is a story about mechanisms, not representations.

The story sheds light on abstraction. If by "abstraction" we mean "getting at the general features," then a system of layered mechanisms would give us this ability. Early in our evolution we focussed on the most salient features of the environment (edges and corners, high and low pitches of sound, etc.), but we gradually developed mechanisms for accessing it in more detail. We accordingly need only revert to our earlier mechanisms to pull out the most general features of the environment. Rather than doing a search for general features, we revert to our lowest-level detectors and filters. So abstraction, our most prized intellectual possession, may not be what it seems.

This story avoids the problem of the representational veil. But we can still have representations and maps. The condition is that we must be able to "get behind them," to check up on the information they are trying to organise. We might bring multiple maps to experience, but we need fast ways of checking them out. The driver who only looks at the map is not an evolutionary success. Thus, on the one hand we have mechanisms that "let the information in," and on the other we have manipulable, controllable structures that enable us to reduce the complexity of subsequent search. And we have a sequence: filters and feature detectors first, maps and representations later. It is reasonable to assume that the latter exploited the former. But how?

We know that filters and feature detectors are more efficient if they are accompanied by domain knowledge: the more we know about the domain, the less search we have to do. For example, it is easier to understand a sentence from which letters have been deleted if we know that the sentence is a proverb. Representations and maps also depend on accompanying knowledge, more so as the domain becomes more abstract. This ability to shift the cognitive load from specific filters and feature detectors to general knowledge about the domain must have been a major step forward. So was the ability to have accompanying knowledge, not about the domain ("this is a proverb"), but *about the filter or feature detector itself* ("this only picks out proverbs . . . or corners . . . or edges"). Such knowledge enables us (a) to choose between filters within a domain and (b) to use them across domains, thus overcoming domain-specific constraints. This is a tall order, but it brings us closer to C&T's multiple maps and representations.

Reducing problem complexity by analogical transfer

Peter F. Dominey
INSERM U94 69500 Bron; CNRS EP-100 69008 Lyon, France.
dominey@lyon151.inserm.fr

Abstract: Analogical transfer in sequence learning is presented as an example of how the type-2 problem of learning an unbounded number of isomorphic sequences is reduced to the type-1 problem of learning a small finite set of sequences. The commentary illustrates how the difficult problem of appropriate analogical filter creation and selection is addressed while avoiding the trap of strong nativism, and it provides theoretical and experimental evidence for the existence of dissociable mechanisms for type-1 learning and type-2 recoding.

Clark & Thornton (C&T) cite analogical reasoning as an example of how previously learned concepts can be used to filter out the salient regularities in novel situations in order to reduce type-2 problems to type-1 status. This commentary addresses the important open issue of how such filters might be developed and chosen while avoiding the trap of strong nativism. The trap is that one might require a specific mechanism for representing each different concept/filter, and thus be no better off than without analogical reasoning. What would be more appropriate is a single mechanism that could provide a general capacity for analogical reasoning.

In this context, it has been proposed that the recognition of structural isomorphisms is one of the primary means by which parts of two analogs can be placed in correspondence with each other (Thagard et al. 1990). For example, consider the two sequences, ABCBAC and DEFEDF. While they share no common surface structure, these isomorphic sequences share the abstract relational structure "*u, u, u, n*-2, *n*-4, *n*-3," where *u* indicates unique or nonrepeated (unpredictable), and *n*-2 indicates a repetition predictable of the element 2 places behind, and so on. The ability to store, recognize, and use this kind of structural isomorphism should contribute to a general mechanism for analogical reasoning in a profitable tradeoff between nativistic predefined functions and robust generalized behavior.

In order to study such a mechanism we developed a test of analogical transfer in sequence learning (ATSL). The test is based on the serial reaction time (SRT) task, in which learning is demonstrated by a reduction in reaction times for stimuli that appear in a repeating sequence versus stimuli that appear in a random series (Nissen & Bullemer 1987). Sequence learning can occur in uninformed or implicit conditions, that is, the statistical regularities in the sequence can be extracted by an uninformed type-1 mechanism. In the ATSL task, however, the same sequence is never repeated. Instead, a number of isomorphic sequences are successively presented. This is a type-2 problem in that the statistical regularities of the potentially unbounded number of sequences become visible only when they are recoded in terms of their shared relational structure.

We have recently observed that normal human subjects are capable of this kind of recoding in the ATSL task, that is, they display learning in the form of monotonically decreasing reaction times for predictable stimuli in a series of isomorphic sequences (Dominey et al. 1995b). It is interesting to note that this type-2 learning is only observed in subjects who have been explicitly informed that such an abstract structure might exist. Implicit or noninformed subjects display no such learning, in striking contrast with their known capacity for type-1 learning in the SRT task (Nissen & Bullemer 1987).

To confirm the observation that this type-2 task cannot be performed by a type-1 system, we performed simulation studies using a model of type-1 sequence learning based on the neural architecture of the primate frontostriatal system (Dominey et al. 1995a; Dominey 1995). In the model, learning-related synaptic modifications generate increased activation of appropriate response units for stimuli in learned sequences, with a corresponding RT reduction. Due to this property, the model demonstrates type-1 SRT sequence learning when using a single repeating sequence (Dominey 1996). It fails, however, in the type-2 ATSL task, with the same lack of learning as observed in the implicit learning group (Dominey 1996; Dominey et al. 1995b).

For the model to exploit the abstract structure shared by isomorphic sequences like ABCBAC and DEFEDF, it must be capable of representing such sequences in terms of the structural relation sequence "*u, u, u, n*-2, *n*-4, *n*-2" that is common to them. This requires (1) a form of short term memory (STM) of the several previous elements in a sequence, and (2) the capacity to recognize whether the current element matches any of those in the STM, for example, that the second A in ABCBAC matches the element 4 places behind. Finally, this recoding of the sequences must be made available to the existing type-1 learning mechanism.

The type-1 sequence learning system can then be used to learn this abstract structure that serves as a filter for subsequent inputs. In the same way the type-1 system alone can predict specific elements in learned sequences, the modified type-2 system can predict repeated elements in learned classes of isomorphic sequences. Indeed, we observed that the modified type-2 model reproduces the performance of explicit subjects in the ATSL task (Dominey et al. 1995b). The type-2 mechanism is simultaneously capable of (1) applying "filters" learned from previous experience, that is, recognizing learned abstract structures in order to predict the repetitive structures of new isomorphic sequences, and (2) developing new "filters," thus learning new abstract structures for isomorphic sequences whose abstract structure has not previously been learned. The system achieves this by continuously exploring, in parallel, the space of possible abstract structures, recognizing learned structures and learning new ones as necessary. The filters (abstract structures), are stored as remembered sequences, and are selected by a type-1 process of sequence recognition. Note that this type-2 mechanism should generalize to the related problems of (1) exploiting several abstract structures in a body of input data, and (2) grammaticality judgment after letter set transfer in artificial grammar learning (Gomez & Schaveneveldt 1994).

From a neurophysiological perspective, it is of interest that type-1 SRT learning is impaired in Parkinson's disease (Jackson et al. 1995), indicating that the frontostriatal system may participate in this type-1 learning. In contrast, our recent studies of analogical transfer in Parkinson's patients have demonstrated that the impairment in type-1 SRT learning is not seen in type-2 ATSL learning in these patients (Dominey et al. 1996). This suggests a functional distinction in the role of the frontostriatal system in type-1 and type-2 learning.

In conclusion, analogical transfer in sequence learning is presented as an example of how the type-2 problem of learning an unbounded number of isomorphic sequences is reduced to the type-1 problem of learning a single or greatly reduced set of sequences, by continuously recoding sequences in terms of their relational structure. This example is of theoretical interest in that (1) it provides an explicit demonstration of how the potentially difficult problem of appropriate analogical filter creation and selection is addressed while avoiding the trap of strong nativism, and (2) it provides theoretical and experimental evidence for the existence of dissociable mechanisms for type-1 learning and type-2 re-coding.

ACKNOWLEDGMENT
The author is supported by the Fyssen Foundation (Paris).

Cognitive success and exam preparation

Matthew Elton
Department of Philosophy, University of Stirling, Stirling, Scotland, FK9 4LA.
matthew.elton@stirling.ac.uk
www.stir.ac.uk/philosophy/staff/elton

Abstract: Evolution is not like an exam in which pre-set problems need to be solved. Failing to recognise this point, Clark & Thornton misconstrue the type of explanation called for in species learning although, clearly, species that can trade spaces have more chances to discover novel beneficial behaviours. On the other hand, the trading spaces strategy might help to explain lifetime learning successes.

Clark & Thornton's (C&T's) target article is about the general principles of operation of cognitive mechanisms. More precisely, they are interested in operational principles that are likely (rather than those that necessarily must or, as it happens, are) to be found in mechanisms that have evolved through natural selection. Hence their paper is about cognitive science, or even evolutionary cognitive science, rather than, say, simply about artificial intelligence.

The principle of operation at issue is learning and at the heart of C&T's paper is a distinction between two sorts of learning problem. Statistical approaches to learning will find type-1 problems fairly easy, C&T explain, and, in the absence of suitable re-coding, type-2 problems very hard. One thing that is attractive about this central claim is the level at which it is pitched. Rather than simply undertaking empirical trials, comparing this connectionist system with that symbolic one, or that network architecture with another, they have also sought to identify a general principle that operates independently of a connectionist-symbolist dispute. While connectionist learning systems do have many advantages, they can no more work magic with a type-2 problem than any other form of statistical learning. This sort of thesis is much needed in cognitive science. If it holds up, then, at a stroke, it can transform empirical hunches about the power of this or that algorithm into firm theoretical results.

C&T describe an animat that starts life lacking the skill of "conditional approach." Its problem is to acquire this skill, but, because the acquisition problem is type-2, the animat is unlikely to stumble across the solution without trading spaces, even if given a great deal of help in the form of training. (And, if I understand C&T aright, it is given rather more help than an animal in a natural setting could expect.)

Although an effective illustration of the type-1/type-2 distinction, the example uncovers a curious inversion in C&T's thinking. Imagine the evolution by natural selection of a creature such as their animat – call it a "clarkton." At some stage, the clarkton is in the same state as the unschooled animat. Let us suppose the environment is putting it under increasing pressure and that the clarkton would greatly benefit from acquiring "conditional approach." Consider three possible futures: (i) the clarkton, by very lucky chance, or with less lack and the application of the trading spaces strategy, solves C&T's problem, acquires "conditional approach," and so improves its fitness; (ii) the clarkton solves another quite different type-2 problem, one which improves its fitness in such a way as to obviate the benefit of "conditional approach"; (iii) the clarkton fails to improve its fitness and dies. (Note that if we forget the high frequency of outcomes like [iii] we gain a false impression of the efficiency of natural selection.)

The only difference between (i) and (ii) is that in (i) the behavioural change is one in which C&T have a special interest. It is only with hindsight that the clarkton can be seen to have "solved" C&T's "problem," and this shows that talking about problems and searches for solutions is out of place here. Evolution is not like an exam, where the problems are set ahead of time. Rather, many different routes are tried out, and creatures stick with those that work. Creatures don't "aim" to acquire specific skills, though when they do acquire a new skill and survive, it is usually a skill worth having.

By contrast, lifetime learning can be rather like an exam. A creature has to keep tackling just one problem until it gets it right. Here the trading spaces idea might make a real contribution, explaining how a strategy of repeatedly recoding and trying again (or perhaps trying different codings in parallel) improves one's chances. Of course, chance is still involved in stumbling across the right kind of coding.

But not only chance. C&T argue that the sorts of problems that, say, human beings have become adept at solving are structured in such a way as to be amenable to our problem solving powers. Lifetime learning is like an exam, but the exam script consists of problems at which earlier candidates were successful. This analogy is too crude: the development of language, say, must surely have involved very many incremental cycles, with small dividends at each stage, in which complexity was slowly ratcheted up. But the route of development was such that each individual was well equipped to "solve" the relevant "learning problem" at each stage. To the extent, however, that this part of the argument is a success, it begins to lessen, though not eliminate, the need to explain type-2 problem solving powers.

By thinking of evolution as being like an exam, C&T create a spurious difficulty. It is only with hindsight that natural selection "solves" its design "problems." Any successful evolution of a