PARALLEL DISTRIBUTED PROCESSING APPROACHES TO CREATIVE REASONING: TENSOR MODELS OF MEMORY AND ANALOGY

GRAEME S. HALFORD, JANET WILES, MICHAEL S. HUMPHREYS
Department of Psychology
University of Queensland
QLD 4072 Australia

and

WILLIAM H. WILSON
School of Computer Science and Engineering
University of New South Wales
PO Box 1
Kensington NSW 2033 Australia

Abstract. This paper presents a computational model, using a Parallel Distributed Processing architecture, of the role of memory retrieval and analogical reasoning in creativity. The memory model stores information as the tensor product of up to three vectors representing, for example, context, cue, and target. The model can retrieve information in the form in which it was stored, or it can use two cues to generate an item that has been separately stored with each of them, but has never been associated with both cues jointly. This means that the intersection of two sets can be computed, without having been stored, thereby providing for the generation of novelty. For analogical reasoning, predicate-argument-bindings are represented as the tensor product of vectors representing relations and their arguments. The model can simulate the transfer of relations from one domain to another, as occurs in the creative use of analogy.

Two properties appear to be essential to creativity. The first is novelty, which must include an element of surprise, predictable variation being insufficient. The second is effectiveness, which may include practical utility, but also means that a coherent set of relations is formed between formerly unrelated, or differently related, elements. In this paper we want to examine two psychological processes that are capable of producing creativity in this sense. The first of these is the retrieval from memory of ideas, or items of information, that relate formerly unrelated things. The second is the transfer of relations from one domain to an unrelated domain, a process which is normally identified with analogy. We will examine computational models of these processes based on parallel distributed processing (PDP) architectures.

The retrieval of an idea that relates formerly unrelated things can be illustrated using one of the practice items from the Remote Associates Test for creativity (Mednick, 1962). What word relates "rat", "blue" and "cottage"? (cheese). These words are not normally seen as related, but all have a common association with cheese. Another example is supplied by Rubin & Wallace, (1989) as analysed Humphreys, Wiles & Bain (1993): Name a mythical being that rhymes with post (ghost). Here the target element, ghost, is weakly associated with either the cue "mythical being" or the cue "rhymes with post", but is quite strongly retrievable to the two cues. These examples may not meet the effectiveness criterion, because the set of relations created by the retrieval of cheese, or ghost, is not particularly coherent. Nevertheless the memory processes entailed in these tasks probably have much in common with those used by a painter producing new juxtapositions of colors and forms, a poet composing images, or a scientist seeing a connection between apparently unrelated ideas.

Creativity in the sense of transferring a relation from one domain to another is illustrated by the Rutherford analogy between the structure of the hydrogen atom and the structure of the solar system, as analysed by Gentner (1983). Here the essential insight was recognizing that the structure of the atom entailed the principles of orbital motion, as exemplified in the solar system. This entails transferring the system of relations between solar and planetary bodies to the nucleus and electron. This example illustrates the role of analogy in creativity.

In this paper we will examine how the processes of memory retrieval and analogical reasoning can produce an output which is creative according to the definition above. Although there are numerous theories of both memory retrieval and analogy, we will focus on models in each domain which use the same cognitive architecture, based on parallel distributed representations.

Creativity and Memory Retrieval

The model of Humphreys, Bain and Pike (1989), integrates a number of previous memory models, and has been applied to creativity (Wiles, Halford, Stewart, Humphreys, Bain & Wilson, 1992). In this model, an item is represented as a vector, and bindings between items (e.g. context, cue, target) are represented as the tensor product of vectors. A memory is formed from the linear sum (or superposition) of tensors representing the bindings.

One type of memory retrieval uses the three-way association between items (e.g.cue, target, and context). For example, in the question "what did you have for breakfast on Sunday?", "what did you have for breakfast?" is the cue, "Sunday" is the context, and (say) "bacon and eggs" is the target. The binding of context, cue and target is achieved by computing the tensor (outer) product of the cue, context, and target vectors, as shown in Figure 1. The tensor product of two cues is computed and the inner product between this rank 2 tensor and the rank 3 tensor memory is computed. This process retrieves memories in exactly the form in which they were stored.

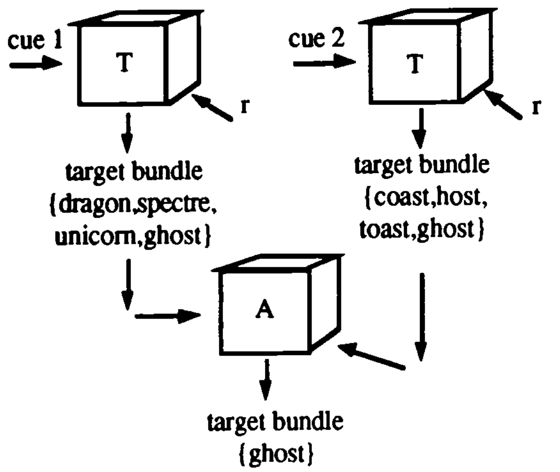

Fig. 1. Computation of intersection of sets retrieved using two cues.

A second type of retrieval also uses two cues, but in this case the target item has been associated with each cue separately, not in the joint manner that is represented by the use of a rank 3 tensor. It is exemplified by the problem above, namely "Name a mythical being that rhymes with post". There are two retrieval cues in this case, "mythical being" and "post" (rhymes are assumed to function as associations in the model), so two tensor memories are shown. Because the association is context independent, a fixed vector is used in place of the context vector. This effectively reduces the representation to two rank 2 tensors, one storing the association between "mythical being" and "ghost", and one storing the association between "post" and "ghost". There is no rank 3 tensor storing the 3-way association between "mythical being", "post" and "ghost".

Retrieval occurs by using the vector representing one cue (e.g. "mythical being") as input to the appropriate memory. The output is a vector representing possible targets (a "target bundle"), as shown in Figure 1. The second cue is also used in the same manner. "Mythical being" elicits {dragon, spectre, unicorn, ghost}, while "post" elicits {coast, host, toast, ghost}. These possible targets for each cue are superimposed, so the output to each cue is the linear sum of vectors representing the possible solutions.

"Ghost" is a weak associate of both cues, and so is weakly represented in both sets of targets. Its effective retrieval depends on accessing the intersection of the target sets. The computation of the intersection in the model is shown in Figure 1. It entails a second tensor memory A, that only stores item information. That is, A is an auto-associative memory, $A = \sum a_i \otimes a_i \otimes a_i$, where a_i includes{dragon, spectre, unicorn, ghost, post, coast, toast, host}. The input to A is the two target bundles comprising the possible solutions from the first stage. The output is the intersection of the bundles, "ghost".

It is assumed that the association between ghost and post, and between ghost and mythical being, have been learned in separate episodes, and the three-way association between post, mythical being, and ghost has not been stored. This means that the intersection cannot be directly retrieved, but has to be computed. Its computation amounts to generating a new representation. It has an element of novelty, and provides some degree of relationship between the formerly unrelated ideas "post" and "mythical being". In this sense it is creative.

Our next step is to see how the other major component of creativity, transfer of relations between domains, can be modelled by the same type of mechanism. Given that we want to integrate the analogy model with the memory model, we require a parallel distributed processing (PDP) model of analogy.

Creativity and Analogy

Gentner's (1983) analysis of analogy as a structure-preserving map from a base to a target has provided a basis for successful theories in the context of human cognition. In the Rutherford analogy, the solar system constitutes the base and the atom the target. The analogy consists of mapping base into target in such a way that correspondence is achieved between the respective structures.

Several computational models of human analogical reasoning have emerged in recent years. These include COPYCAT (Mitchell & Hofstadter, 1990) which solves problems of the form abc:abd::ijk:? (ijl). It is based on a slipnet which enables representations to be modified as structure mapping progresses. In a problem such as abc:abd::xyz:? the representation might be changed from "change last element to its successor" to "change first element to its predecessor". This yields "wyz" as the solution.

In the Structure Mapping Engine (SME, Falkenhainer, Forbus & Gentner, 1989) source and target are represented as predicate-argument bindings, coded in predicate calculus. It is a serial model in which match hypotheses are created, based on similarity of predicates and/or arguments in base and target. These are collected into global matches, selected according to overall consistency criteria.

In the Analogical Constraint Mapping Engine (ACME, Holyoak & Thagard, 1989), a parallel constraint satisfaction algorithm finds the mapping which maximizes the correspondence between base and target. This is achieved by inhibitory connections between rival mappings (tending to ensure uniqueness of mapping) and excitatory connections which tend to ensure that if a predicate is mapped, its arguments are mapped, and vice versa, thereby creating structural correspondence.

These models use local rather than distributed representations. However the Structured Tensor Analogical Reasoning model (STAR, Halford, Wilson, Guo, Gayler, Wiles & Stewart, in press) uses distributed representations, with predicates and arguments represented as vectors. It was motivated by the desire for a PDP model of analogical reasoning which would also offer a solution to the problem of defining processing capacity limitations in cognition and cognitive development (Halford, in press). However it also uses an architecture which is consistent with the memory retrieval model discussed in the previous section.

In the STAR model predicate-argument bindings are represented as tensor products, as shown in Figure 2. The proposition MOTHER-OF(woman,baby) is represented as a rank 3 tensor product. The three vectors represent the predicate MOTHER-OF and its arguments, "woman" and "baby". As with the memory model of Humphreys et al. (1989), representations are superimposed. Therefore the bindings MOTHER-OF(mare,foal), MOTHER-OF(cat,kitten), as well as LOVES(mother,baby), . . . , LARGER-THAN(mare,foal) are all superimposed on the same tensor product.

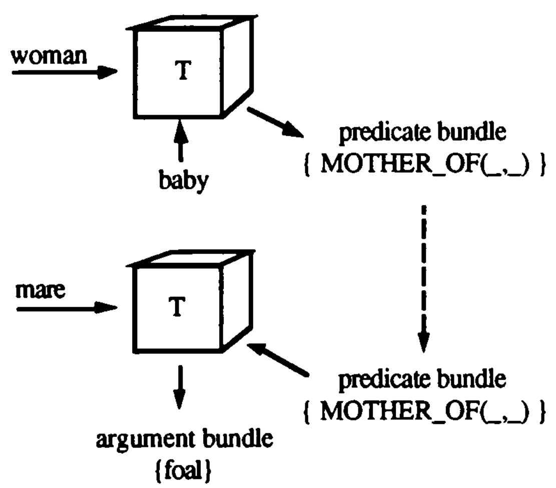

Fig. 2. Solution of simple proportional analogy

Simple proportional analogies, such as mother:baby::mare:? can be solved by entering the base arguments, "mother" and "baby" into the representation, and the output is a predicate bundle representing "MOTHER-OF", "FEEDS", "LARGER-THAN" etc. That is, it is a vector equivalent to the linear sum of vectors representing each of these outputs. In the next step this predicate bundle is the input, together with the first argument of the target, "mare", as shown in Figure 2. The output is an argument bundle which includes "foal". Possible solutions can be recognized by one or more cleanup processes, the simplest of which entails computing the inner product of vectors representing candidate solutions with this output.

This model can solve all the major classes of analogies (Halford et al., in press). Because of its relevance to creativity, we will consider how the model would solve a simplified and idealized version of the Rutherford analogy.

We will assume that the structure of the base, the solar system, is known completely, but that the structure of the atom is incompletely known. We will further assume that the atom was known to comprise two major components, a nucleus and an electron, that there was a mutual attraction relation between them, and a size difference (the nucleus was known to be of larger mass). Most important of all, we assume that the correspondence between the atom and the solar system had not been previously recognized. This recognition was at the core of the creative contribution. The simulation of the analogy in the STAR model begins with this information coded in a tensor product.

The discrepancies between this hypothetical situation and the actual state of Rutherford's knowledge at the time are less relevant than the idea of starting with an incomplete representation of a target domain, finding a base domain to which it is analogous, transferring relations from the base to the target, then making candidate inferences about the target. It is this creative use of analogy that want to simulate, the Rutherford analogy being a convenient and well-known example.

Recognition of the solar system as a potential base is a memory retrieval process in the model. The predicates ATTRACTS and DIFFERENT-SIZE in the representation of the target serve as retrieval cues. If these are used as input to the representation of the task, the output is a tensor product of vectors representing all sets of arguments of these predicates. For example, when ATTRACTS is used as input, the output will include the tensor product of vectors representing "sun" and "planet", plus other pairs of arguments of ATTRACTS.

Both ATTRACTS and DIFFERENT-SIZE are likely to be weak retrieval cues for "sun" and "planet". Many things are associated with ATTRACTS; e.g. lovers attract each other, lights attract moths, etc. Similarly, DIFFERENT-SIZE has many associates besides "sun" and "planet". In these circumstances, effective retrieval depends on computation of the intersection of the outputs from the two cues, in the same manner as the retrieval of ghost from "mythical being" and "post" described earlier. ATTRACTS would retrieve a target bundle, that is a vector equivalent to the linear sum of vectors representing sun-planet, lover-lover, light-moth etc. Similarly, DIFFERENT-SIZE would retrieve another target bundle representing sun-planet, elephant-horse, adult-child, etc. These target bundles would be used as inputs to the auto-associative memory, as shown in Figure 1. The output would represent sun-planet more strongly than the output to either cue alone. If further processing were required, this could include finding the intersection of this output with the output of another cue, such as "relevant to mechanics".

When arguments of a potential base are recognized, they can be used as input. If the input is "sun" and "planet" the output will include ATTRACTS, LESS-MASSIVE-THAN, ORBITS-AROUND, etc. These predicates become candidate inferences for the target, leading to the hypothesis that electron orbits around the nucleus.

This has been a highly simplified account of how analogical reasoning is simulated in the STAR model. It has been designed to illustrate the main argument of this section, that a computational process can produce an output which is inherently creative. In this case a computational model has been constructed which can retrieve a base, the solar system, which is a potential analog of the target. Relations in the base are transferred to the target, and become candidate inferences for the target.

Creativity, Memory Retrieval, and Analogy

Memory and analogy can be creative, and can be modelled by computational processes. The analogical reasoning model discussed here shares a lot of common architecture with a model that has been shown to capture many properties of human memory. However this architecture also has some of the power and flexibility that is characteristic of human reasoning. In part this derives from the fact that the outcome of a retrieval operation can consist of a set of items, or even a set of bindings, rather than a specific item. This set can then be used in a new retrieval operation, with additional cues where appropriate. The ability to query memory with different cues, including sets, provides flexibility that is important in higher cognitive processes. Equally important is the fact that memory can be queried in different ways. Notice that, in the tensor product representations in Figures 1 and 2, there is no fixed input or output. That is, a particular vector can be an input at one time, and an output at another time, in contrast to standard three-layered nets, in which one vector is always input and another is output. This is another property that provides some of the flexibility and power that is important in higher cognition. Nevertheless, all this occurs in an architecture that models basic memory operations, including creative retrieval of information linking formerly unrelated items. Perhaps therefore a beginning has been made in producing an integrated model of human creativity.

The creativity displayed by the computational processes outlined here is undoubtedly of a restricted kind, and it will take considerably more research before we know the limits of these processes. The important point however is that these computational processes have been shown to be creative in principle. In human beings, creativity depends heavily on memory retrieval and analogy, both of which are now the subject of intensive efforts at computational modelling. Memory is not merely reproductive, but has long been known to be generative, and memory retrieval clearly includes a construction process. When the construction forms a new relationship between formerly unrelated items of information, it can be said to be creative.

References

Falkenhainer, B., Forbus, K. D. and Gentner, D.: 1989, The structure-mapping engine: Algorithm and examples, *Artificial Intelligence*, 41, 1-63.

Gentner, D.: 1983, Structure-mapping: A theoretical framework for analogy, *Cognitive Science*, 7, 155-170.

Halford, G. S.: in press, *Children's understanding: the development of mental models*. Hillsdale, N. J.: Erlbaum.

Halford, G. S., Wilson, W. H., Guo, J., Gayler, R.W., Wiles, J. and Stewart, J.E.M: in press, *Connectionist implications for processing capacity limitations in analogies*. Norwood, NJ: Ablex.

Holyoak, K. J. and Thagard, P.: 1989, Analogical mapping by constraint satisfactions, *Cognitive Science*, 13,3, 295-355.

Humphreys, M. S., Bain, J. D. and Pike, R.: 1989, Different ways to cue a coherent memory system: A theory for episodic, semantic and procedural tasks, *Psychological Review*, 96,2, 208-233.

Humphreys, M. S., Wiles, J. and Bain, J. D.: 1993, in press, Memory retrieval with two cues: Think of intersecting sets, *in* D. E. Meyer and S. Kornblum (eds.), *Attention and performance XIV: A silver jubilee* , Hillsdale, NJ: Erlbaum,

Mednick, S. A.: 1962, The associative basis of the creative process, *Psychological Review*, 69, 220-232.

Mitchell, M. and Hofstadter, D. R.: 1990, The emergence of understanding in a computer model of concepts and analogy-making, *Physica D*, 42,1-3, 322-334.

Rubin, D. C. and Wallace, W. T.: 1989, Rhyme and reason: Analysis of dual retrieval cues, *Journal of Experimental Psychology: Learning, Memory and Cognition*, 15, 698-709.

Wiles, J., Halford, G. S., et al.: 1992, Tensor models: A creative basis for memory retrieval and analogical mapping, *in* T. Dartnall (ed.), ,