THE MECHANISMS OF EPISODE CONSTRUCTION AND BLENDING IN DUAL AND AMBR: INTERACTION BETWEEN MEMORY AND ANALOGY

Boicho Kokinov[1,2]
[1]Central and East European Center for Cognitive Science
[2]Institute of Mathematics and Informatics, Bulgarian Academy of Sciences
bkokinov@nbu.bg

ABSTRACT

This paper presents an attempt to build an integrated model of analogy-making and memory. In contrast to other models of analogy-making (Forbus et. al, 1995, Thagards et al, 1990, Hummel and Holyoak, 1997), AMBR, Associative Memory-Based Reasoning, (Kokinov, 1988, 1994a, Kokinov and Petrov, 2001) is trying to make realistic assumptions about human memory and take into account its vulnerability to distortions (Kokinov and Petrov, 2001). On the other hand, in contrast to typical models of memory, AMBR is not designed specifically as a model of "pure memory tasks" such as recall or recognition. It utilizes memory in more complex problem solving tasks. Considered as a model of memory AMBR is in its infancy - it has to be further developed in order to cover the full range of memory tasks. However, the very fact of integration of analogy-making and memory makes it possible to model and predict various interaction effects which have not been studied so far. More specifically, using AMBR we can explore the role analogy-making plays in memory – in the organization of the information, in changing the memory representation, in generalization, in producing memory distortions. On the other hand, it is clear that analogy is a memory-based task and thus AMBR has to explain the influence memory has on the analogical reasoning.

HEADING 1

AMBR is based on a general cognitive architecture, called DUAL (Kokinov, 1994 b,c). This is a kind of Society of Mind architecture, consisting of numerous simple micro-agents and built upon their local interactions. The agents are hybrid (symbolic/connectionist) devices and that is why DUAL implements a micro-level hybridization. The symbolic part of each agent represents some piece of knowledge, while the connectionist part of it represents its relevance to the current context and thus its availability for participation in the emergent global processing.

Heading 2

Compared to other memory models, DUAL is a multi-trace model, i.e. each "trace" is represented separately – there is not one single memory vector where all memory traces are superimposed like in CHARM (Eich, 1982; Metcalfe, 1990) and TODAM (Murdock, 1982). Compared to other multi-trace models like SAM (Raaijmakers and Shiffrin, 1981), REM (Shiffrin and Steyvers, M. 1997), and MINERVA2 (Hintzman, 1988), "memory traces" in DUAL are not considered as relatively stable structures, such as vectors, memory images, etc. They are rather dynamic constructions which emerge on the spot from the interactions between the agents. In that sense, representations in DUAL are similar to the distributed representations used in connectionist models – they are some patterns of activation over a set of agents, and these patterns are formed dynamically. Unlike connectionist network models (McClelland, 1995, Nystrom and McClelland, 1992), however, the building blocks are at a higher level - they are simple statements, rather than simple features. The difference is very important since using these higher-level structures DUAL is able to encode also relational knowledge rather than simple independent features only. This is crucial for analogy-making which is based on relational similarity rather than on feature overlap. However, this extends also the possibilities for exploring the role of structure in human memory in general, which has been often underestimated. Another difference with SAM, REM, MINERVA2 and some connectionist models is that DUAL do not rely on a fixed list of features over which all elements to be represented. This assumption of having a fixed list of features seems too restrictive to me.

Heading 2

A concept, an episode, or a single object are represented in DUAL as a coalition of agents in LTM. Each agent represents a single element, relation, property (objects, relations and properties are treated in the same way). Thus the coalition as a whole represents all we know about that concept, event, or object. However, what is currently accessible is only what is currently active from this coalition, that is what is represented by agents that have a level of activation above a threshold and thus by agents that are part of WM (WM is the active part of LTM). Thus in different moments of time and in different contexts we have different knowledge of the episode, concept, or object. In this way the "memory trace" of the episode, concept, or object is constructed dynamically and depends on the context.

The perceptual process that encodes an episode builds the coalition of agents corresponding to it. Depending on what we are focusing on various aspects, features and relations between elements are encoded in micro-agents, and also various strength of the links between the agents are assigned thus making the coalition stronger or looser. A "binding agent" is also constructed that represents the fact that all these agents represent various aspects of a single episode. The binding agent, however, does not know anything, in particular, it does not know its neighbors – the agents from the coalition. We do not want to have such a centralized representation of the episode and to have an agent that could activate all members of a given coalition. So, only the coalition members have links to the binding agent, but not vice versa.

HEADING 1

If the same or a very similar episode is being experienced (e.g. I enter for the second time the same kitchen, where the same table, the same chairs are there, although they might be in slightly different spatial relations) a new coalition is built consisting of new agents (no direct overlap of the coalitions is envisaged). However, when we recognize that we are seeing the same table in the new situation than a "c-coreference link" is established between the old agent and the new agent. This link represents semantically that these are two representations of the same object in reality and the symbolic processes can use this information, but also the link makes it possible to spread activation from the one agent to the other. Thus always when we face a new situation which involves old, well-known objects, the corresponding old situations (in which these objects participated) get some activation via these c-corefence links. If these old coalitions are strong enough than relatively big portion of them may get activated.

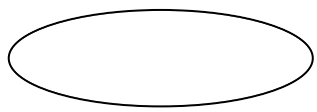

Figure 1. This is a figure.

Free recall is modeled in AMBR as a process of undirected spreading activation which is started from some cues given by perception (agents on the input list) or by internal motivation (agents on the goal list). The result may be the activation of any coalition or part of it. It may also happen that parts of two or more coalitions are simultaneously activated in WM and this will result in a potential blend of episodes (if the evaluation process does not eliminate it on some ground).

Cued recall, and more specifically question answering (like “What was there on the table when yesterday you entered the kitchen?”) is modeled in AMBR as a specific type of analogy-making (looking for literal similarity), i.e. given a target cue describing partially a situation we are looking for an old situation that will share not only the objects and features, but also the relations between them. The processes of analogical retrieval and mapping start to run in parallel and interact with each other in order to produce the best possible match. Now, if there are several old episodes that are similar to the situation at hand each of them is trying to map on it. There is a competition between various simple elements (represented by micro-agents) that potentially map onto the target elements (for example “a book” lying on the table from one old episode, and “a cup” being on the table in another one). This will be a direct competition between these two hypotheses. However, all hypotheses are related in a global constraint satisfaction network consisting of excitatory and inhibitory links. Normally, a whole episode wins the competition and we will get a consistent mapping between the target and that episode. However, there is no guarantee for this. If, for example, the episode is not a strong enough coalition that the support each hypothesis gets from the other will not be enough and thus it would be possible to have a mixed or blended result – various elements from the target are mapped onto elements from different coalitions. The presence and activation of the binding agents is an anti-blending mechanisms, but it is not a very strong one and cannot eliminate such situations.

Another paper at the workshop (Grinberg and Kokinov, 2003, this volume) presents simulation results with AMBR which demonstrate blending effects in cued recall. Moreover, such blended episodes are reconstructed not only when the two old episodes are almost identical (as in the misinformation paradigm of eyewitness testimony (Loftus, 1979)). It turned out that blends are also possible between episodes that are not so much superficially similar (different elements participate in them, they may have different properties), but which share relational structure and have been previously mapped in an analogy. Similar results have been obtained in a psychological experiment (when the X-rays used by the doctor for eliminating the tumor became laser beams, which was taken from an analogous story about a light bulb to be repaired).

The simulation predicted that blending should be possible even between episodes that do not share neither superficial, nor structural properties. It would be enough that the two episodes have participated in an analogy with a single other episode, each on its own grounds. This double analogy builds new links between the three coalitions and when one of them is activated the other two are also activated and if the coalitions are not very strong blending between them may happen. Such a result has been obtained in a simulation experiment (Grinberg and Kokinov, this volume).

Table 1. This is a table.

This was a very brave and counterintuitive prediction made by the model: blending even between dissimilar episodes is possible if they have participated in a double analogy. A series of experiments have been performed to test this prediction. The results have confirmed it (Kokinov and Zareva, 2001, Zareva and Kokinov, 2003, this volume). Moreover, it turned out that this blending is not a result of residual activation since it lasts for at least a week and thus must be based on permanent links established in LTM as result of the analogy. It also turned out that the links established by the analogy-making process are much stronger than the simple associative links established as a result of co-activation (thinking about the two episodes at the same time), thus in one experiment we found that double analogy produces blending while two single analogies involving the same episodes do not (Zareva and Kokinov, this volume).

As said in the beginning, AMBR is in its infancy as a memory model and a lot has to be further developed before we can claim that we have built a more general memory model. By implementing an integration between memory and analogy, this model predicted and demonstrated some nontrivial interactions between them and as result a new phenomenon "blending of dissimilar episodes" was established. Thus AMBR raised some new issues about the possible causes of blending and more generally, about the possible mechanisms of episode construction in memory, mechanisms that will respect relational structure as well as superficial similarity. It also raises the issue of the change of representation of old episodes and its dynamics.

REFERENCES

Eich, J. Metcalfe (1982). A Composite Holographic Associaltive Recall Model. Psychological Review, 89, 627-661.

Forbus K., Gentner D., & Law, K. (1995). MAC/FAC: A model of similarity-based retrieval. *Cognitive Science,* 19, 141-205.

Grinberg, M. & Kokinov, B. (2003). Simulating Episode Blending in the AMBR Model. In: *Proceedings of the European Cognitive Science Conference.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Hintzman, D. (1988). Judgements of frequency and recognition memory in a multiple-trace model. *Psychological Review*, *95*, 528-551.

Hummel, J. & Holyoak, K. (1997). Distributed representation of structure: A theory of analogical access and mapping. *Psychological Review, 104*, 427-466.

Kokinov, B. (1988). Associative memory-based reasoning: How to represent and retrieve cases. In T. O'Shea and V. Sgurev (Eds.), *Artificial intelligence III: Methodology, systems, applications*. Amsterdam: Elsevier.

Kokinov, B. (1994a). A hybrid model of reasoning by analogy. In K. Holyoak & J. Barnden (Eds.), *Advances in connectionist and neural computation theory: Vol. 2. Analogical connections* (pp. 247-318). Norwood, NJ: Ablex

Kokinov, B. (1994b). The DUAL cognitive architecture: A hybrid multi-agent approach. *Proceedings of the Eleventh European Conference of Artificial Intelligence*. London: John Wiley & Sons, Ltd.

Kokinov, B. (1994c). The context-sensitive cognitive architecture DUAL. *Proceedings of the Sixteenth Annual Conference of the Cognitive Science Society*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Kokinov, B. & Petrov, A. (2001). Integration of Memory and Reasoning in Analogy-Making: The AMBR Model. In: Gentner, D., Holyoak, K., Kokinov, B. (eds.) *The Analogical Mind: Perspectives from Cognitive Science*, Cambridge, MA: MIT Press.

Kokinov, B. & Zareva-Toncheva, N. (2001). Episode Blending as Result of Analogical Problem Solving. In: *Proceedings of the 23nd Annual Conference of the Cognitive Science Society*. Hillsdale, NJ: Erlbaum.

Loftus, E. (1979/1996). *Eyewitness testimony*. Cambridge, MA: Harvard University Press

McClelland, J. (1995). Constructive memory and memory distortions: A parallel distributed processing approach. In D.

Schacter (Ed.), *Memory distortions: How minds, brains, and societies reconstruct the past*. Cambridge, MA: Harvard University Press.

Metcalfe, J. (1990). Composite holographic associative recall model (CHARM) and blended memories in eyewitness testimony. *Journal of Experimental Psychology: General*, *119*, 145-160.

Murdock, B. (1982). A Theory for the Storage and Retrieval of Item and Associative Information. Psychological Review, 89, 609-626.

Nystrom, L. & McClelland, J. (1992). Trace synthesis in cued recall. *Journal of Memory and Language*,*31*:591-614

Raaijmakers, J., Shiffrin, R. (1981). Search for Associative Memory. Psychological Review, 88, 93-134.

Shiffrin, R., Steyvers, M. (1997). A Model for Recognition Memory: REM: Retrieving Effectively from Memory. Psychonomic Bulletin and Review, 4, 145-166.

Thagard, P., Holyoak, K., Nelson, G., & Gochfeld, D. (1990). Analog retrieval by constraint satisfaction. *Artificial Intelligence*, *46*, 259-310.

Zareva-Toncheva N. & Kokinov, B. (2003). Blending of Non-Similar Episodes as a Result of Analogical Mapping with a Third One. *Proceedings of the 25th Annual Conference of the Cognitive Science Society*. Hillsdale, NJ: Lawrence Erlbaum Associates.