To appear in D. Gentner, K. Holyoak, & B. Kokinov (Eds.) (2000).
***The analogical mind: Perspectives from Cognitive Science*, pp. 59-124.**
Cambridge, MA: MIT Press.

Integration of Memory and Reasoning in Analogy-Making: The AMBR Model

Boicho N. Kokinov **Alexander A. Petrov**

Central and East European Center for Cognitive Science
Department of Cognitive Science and Psychology
New Bulgarian University
21, Montevideo Str.
Sofia 1635, Bulgaria

and

Institute of Mathematics and Informatics
Bulgarian Academy of Science
Bl. 8 Acad. G. Bonchev Str.
Sofia 1113, Bulgaria

Running head: Memory and Reasoning in AMBR

Contact information:
`kokinov@cogs.nbu.acad.bg`
`http://www.nbu.bg/cogs/personal/kokinov.html`
`apetrov+@andrew.cmu.edu`
`http://www.andrew.cmu.edu/~apetrov/index.html`

Alexander Petrov is now at:
Department of Psychology
Carnegie Mellon University
Pittsburgh, PA
USA

Integration of Memory and Reasoning in Analogy-Making: The AMBR Model

Boicho N. Kokinov **Alexander A. Petrov**

1. Reuniting Memory and Reasoning Research: An Appeal for a Second Marriage after Their Divorce

Three blind men were exploring an elephant. The first of them, who happened to reach the leg, described the elephant as something like a tree trunk—high and of cylindrical shape. The second one grasped the ear and described the elephant as something like a blanket—flexible, thin, and covering a large surface. The third grasped the trunk and formed an image of a long and flexible pipe-shaped object like a hose. For a long time they argued about the right conception of the elephant.

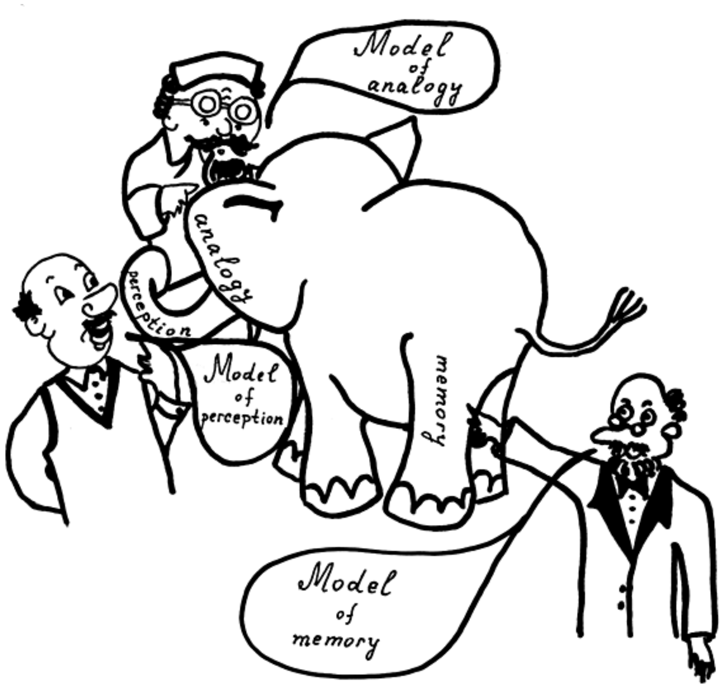


Figure 1. Cognitive scientists study human cognition in small fractions and often do not recognize its underlying unity.

We cognitive scientists are often in the role of those blind researchers trying to understand human cognition. Because it is a huge and complex object of study, each of us approaches it from a certain perspective and studies only a tiny bit of it. Although we do not misrepresent the whole of cognition with the particular object of study, say memory or analogy, we tend to think of mechanisms that could explain the tiny fraction we have focused on. To continue the elephant story, when "trunk specialists" observe the fluid that comes out when the trunk is cut, they tend to hypothesize that it is an olfactory secretion. "Leg specialists" also observe a fluid coming out when the leg is cut but have a very different hypothesis about it—it must be some filling of the leg. The fact that this fluid is one and the same in all cases (blood) and has the same function can be discovered only when these scientists come together and consider the elephant as a whole. They need to explore the interactions between various parts (e.g., that an infection in the leg might cause complications in the trunk) and to postulate general principles and systems (like the cardio vascular system).

There is nothing wrong with separating cognition into pieces and studying them. The practice of "carving nature at its joints" dates at least as far back as the dialogues of Plato. "Scientists try to study systems that are sufficiently closed to be predictable and sufficiently small to be understandable" (Hunt, 1999, p. 8). Big and complex systems are hardly manageable. Studies of isolated parts have led to very important achievements in understanding the mechanisms of human cognition and analogy-making in particular.

However, studies of components should be done with awareness of the fact that the separation of human cognition into various processes is just a convenient tool and not a reality. They should be complemented with explorations of the interactions among various cognitive processes, that is, instead of being carved, the "joints of nature" have to be studied.

Early philosophers like Aristotle considered thinking and memory in an integrated way. The doctrine of associationism explained human thinking by means of the content and organization of human memory. Later, as science developed and psychology became an experimental science, researchers tended to analyze simple and separate faculties of the human mind in order to be able to study them experimentally. Nowadays we have a huge pile of facts about both memory and reasoning (and analogical reasoning in particular). The problem is that these two research communities do not speak to each other often. As a result, facts established in one of the fields are often neglected and ignored in the other.

We feel the time has come to try to put the pieces back together. This chapter makes an attempt to re-integrate research on analogy-making with research on memory. Keith Holyoak and John Hummel (chapter 5, this volume) present another attempt in a similar direction—they integrate analogy with memory and learning. Kenneth Forbus (chapter 2, this volume) also appeals for integrating analogy models with models of large-scale cognitive processes. He presents an integrated model of commonsense thinking based on analogical reasoning and reasoning from first principles. Douglas Hofstadter (chapter 15, this volume) argues that analogy-making might be the core of many cognitive processes from perception to categorization to translation of poetry. Gilles Fauconnier (chapter 7, this volume) integrates analogy with conceptual blending. Paul Thagard and Cameron Shelley (chapter 10, this volume) integrate analogy with emotions. Arthur Markman and Page Morceau (chapter 11, this volume) integrate analogy-making with decision-making. These are all small but important steps in the direction of reintegrating our knowledge about human cognition. It seems that cognitive science has matured enough to pursue these steps.

Modeling has too many degrees of freedom. A phenomenon can often be modeled in several different ways and it is difficult to evaluate the model based on this single phenomenon alone. That is why it is important to restrict the space of possible models by bringing to bear as many constraints as possible. Several types of constraints can be exploited:

- Behavioral constraints—these come from psychological experiments and describe the behavior that should be generated by the model under different circumstances (the richer the set of circumstances the better)
- Biological constraints—these come from the neurosciences and describe the restrictions on the model arising from the known organization of the brain and body
- Evolutionary and developmental constraints—these come from developmental psychology and animal research and restrict the complexity and type of mechanisms as well as their evolution and development
- Architectural constraints—these come from theoretical considerations and require coherence among the mechanisms underlying human cognition so that they can function together and interact

In addition, we can differentiate between specific and general constraints. Typically when modeling a specific phenomenon we tend to concentrate on the constraints known to apply to that specific phenomenon. Thus when studying analogy we tend to collect data *with respect to analogy*. The utility of these data is clear, and we try to draw from as many sources as we can: psychological, neurological, evolutionary, and developmental. Very often, however, we ignore data that are not directly related to analogy but are nevertheless very useful because of their relation to other cognitive processes that in turn relate to analogy. If we consider analogy as an integrated phenomenon in the complex structure of human mind, we need to pay attention to these general constraints as well.

This is, of course, an overambitious task that is clearly beyond the scope of this chapter. However, it is an important motivation of the current work. This chapter describes only a few steps on the way toward integrating analogy back again into human cognition. Special emphasis is put on some general behavioral and architectural constraints and particularly on the integration of analogy-making and memory.

Section 2 presents a highly selective and biased review of the literature on memory. It concludes with a summary of the behavioral and architectural constraints on analogy models as seen by the authors. Section 3 reviews the AMBR research program. Finally, section 4 describes AMBR2—the current version of the AMBR model—which tries to bring memory and analogy back together.

2. Reconstructing the Dinosaur: The Dynamic and Constructive Nature of Human Memory

Is memory a storehouse or an action? There is no consensus on a single and unified theory of memory or even on a single general metaphor for memory (Roediger, 1980; Koriat & Goldsmith, 1996). The classical metaphor of memory describes it as a physical space where items are stored and later on searched for and retrieved. This metaphor has been very powerful and even dominant in the history of psychology. It uses some well-known source domains such as libraries, storehouses, and computers and thus helps us to transfer many inferences about memory. That is why the storehouse metaphor is so widespread. Even our terminology is influenced by it, so that we speak about storage and retrieval from memory.

On the other hand, starting with Sir Frederick Bartlett (1932), the spatial metaphor has been under continuous fire and a new dynamic and constructive view on human memory has emerged. One particularly notable new metaphor is due to Ulric Neisser (1967). He likens human memory to the constructive work of a paleontologist who uses a small set of bone fragments as well as general knowledge about dinosaurs and other similar animals in order to reconstruct and piece together the skeleton: "out of a few bone chips, we remember the dinosaur" (p. 285)[1].

According to the spatial metaphor, memory traces are "stable objects" or "information structures" placed in a store. The "retrieval" process then attempts to locate and select the appropriate ones given a probe. Once a particular memory trace has been retrieved, all the information stored in it is accessible. In other words, memory consists of static structures and active processes. The former simply lie there, possibly indexed and organized in some useful way, while the latter operate on them when necessary. The constructive view (Bartlett, 1932; Neisser, 1981; Barclay, 1986; Brewer, 1988; Metcalfe, 1990; Schacter, 1995; McClelland, 1995; Whittlesea, 1997) takes a different perspective. It does not separate structures from processes and considers memory as a constructive process. Memory traces are conceptualized as temporary states constructed on the spot rather than as "fortune cookies" cracked open to reveal the message contained in them.

There are no true and false metaphors, and each metaphor could be useful in certain contexts. The question is which metaphor would be more useful in the context of analogy-making and problem solving. The two schools of thought have been conducting experiments in different ways. The proponents of the first metaphor have experimented mostly with simple artificial material—lists of words, lists of numbers, sets of pictures, and so on. The dependent measure of main interest has been the success/failure ratio (or ***d'*** in more recent studies). In contrast, the protagonists of the second school have been studying memory in more natural settings[2]. They have been interested in autobiographical memory, in memory for complex events or stories (like a party or a witnessed burglary or car accident). Under these circumstances what is really interesting is not whether people remember the event or not, but rather what details they do remember and what types of errors they make. Focusing on the errors people make in recalling from memory became an important source of insights. Thus the main message sent across by the storehouse metaphor is that one may have trouble finding the book in the library or perhaps that the book might have been spoiled. However, one cannot find a book that does not exist in the library, one cannot find a modified (rewritten) book, and so forth. In contrast, the second metaphor easily communicates the message that because the paleontologist *reconstructs* the skeleton (even though constrained by the given fossils) the result might be quite different from the reality. It might even be the case that the reconstructed skeleton has not existed or even that it cannot exist. The reconstruction might also be a skeleton of a centaur—a nonexistent mixture of two or more kinds of animals. The paleontologist might make a second reconstruction that could

[1] This is actually a nice example of conceptual blending (Fauconnier, chapter 7, this volume)

[2] Dunbar (chapter 9, this volume) presents a nice example of naturalistic studies in analogy-making.

be different from the first one because something new was learned in between, or some fossils have disappeared, or new ones were found.

The empirical question is whether such phenomena happen with human memory, and the answer is yes. During the long history of the second school much evidence has been gathered for false and illusory memories, memory distortions, and so on (see Schacter, 1995, for a recent review). These constructive-memory effects are especially likely when the episode that is to be recalled is complex and agrees with commonsense knowledge. These are the exact characteristics of the sources for many analogies—past problem-solving episodes, familiar events, and real-world situations rather than lists of words. Therefore, we argue that the constructivist view of memory is highly relevant to analogy research and can bring important behavioral constraints for the modeling endeavor. The next section reviews some of the evidence supporting this position.

2.1. Human Memory: Sharp, Complete, and Fixed or Blurry, Partial, and Flexible?

Brown and Kulik (1977) suggested the existence of a special type of memory for important events in our life that they called *flashbulb memory*. They claimed that "it is very like a photograph that indiscriminately preserves the scene in which each of us found himself when the flashbulb was fired" (p. 74). They presented the results of a study which demonstrated that most Americans had a very vivid memory about the assassination of John F. Kennedy, including details about the place they were, the informant, the ongoing event, and so on. So, they supported Livington's idea for a special neurobiological mechanism called *Now print!* that is triggered when we evaluate an event as very important for us. The flashbulb memory theory has inspired a whole line of research and many controversial results have been obtained (Neisser & Harsch, 1992; Conway, 1995). What is clear nowadays is that there are differences in the degree of vividness and the details that we retain about different events. It is also clear that even "flashbulb memories" are partial and probably also distorted. For the sake of accuracy, we must point out that Brown and Kulik wrote in the same article that "a flashbulb memory is only somewhat indiscriminate and is very far from complete" (p. 75).

Now, if even flashbulb memories are not complete, what about our ordinary memories? Bartlett (1932) showed that people ignore many important details of a story. Nickerson and Adams (1979) tested the memory Americans have for a commonly used object such as a penny. It turned out that on average each element was omitted by 61% of the participants. Some elements, such as the text *Liberty*, were omitted by 90% of the participants. Others, such as *United States of America*, *E Pluribus Unum*, and even *one cent*, were omitted by about 50% of them. And, of course, each of us has had personal experiences when we could recall an episode but not some important aspects of it, such as the name of the person, the color of his or her eyes, or the place where we met.

Our inability to recall the details might mean that we have simply not attended and encoded them; in this case memory would not be responsible for the omissions. However, on a particular occasion in a specific context one might be able to recall these specific details. This means that the details are encoded, but one cannot *always* reproduce them. There is a huge number of studies of the effect context plays on our ability to recall or recognize objects and events (see

Davies & Thomson, 1988, for a review). These studies show that although some details can be recalled on one occasion, they may not be recalled on another. Thus Salaman (1982) and Spence (1988), in their reports of involuntary reminding, also claim that people are reminded about the same episode on different occasions at different level of detail, omitting various aspects of the event. Godden and Baddeley (1975) had divers study the material either on the shore or twenty feet under the sea. The divers were then asked to recall the material in either the same or a different environment. Participants clearly showed superior memory when they were asked to recall in the same context in which they studied. Similar environmental context effects on recall have been found in numerous experiments (for an overview see Smith, 1988). Human memory turned out to be mood-dependent as well (for a review see Guenther, 1988). Thus when in an elated mood participants tend to produce more "happy" memories, while when in a depressed mood they tend to produce more unhappy memories. Just having some cookies in the waiting room may influence them to produce more "positively colored life experiences" (Isen et al., 1978).

Many experiments have also demonstrated robust context effects on recognition. For example, Craik and Kirsner (1974) and Kolers and Ostry (1974) have shown that the same voice (vs. different) and same typography (vs. different) facilitate performance in a memory recognition test for words. Davies (1988) provides an exhaustive review of the experimental studies of memory for faces and places. The review shows that recognizing a face in a familiar context is much easier than recognizing it in an unusual one. Thus, for example, Thomson, Robertson, and Vogt (1982) manipulated systematically the setting in which a given person was observed, the activity this person was performing, and the clothing of the person. They found that all three factors had significant effects on a later face-recognition test.

Implicit memory has also been shown to be context-specific. Thus priming effects are decreasing with every single difference between study and test conditions (Tulving & Schacter, 1990; Roediger & Srinivas, 1993).

To summarize, people make many omissions and describe objects and events only partially, but they do so in a context-sensitive manner: different omissions on different occasions. There is an apparent hyperspecificity of human memory that leads us to think that all aspects of an episode are encoded and all of them facilitate our memory for that episode, but on any occasion only a very small part of them can be reproduced. The conclusion we draw is that memory representations are very flexible and context-dependent. This challenges the classic view of memory as consisting of stable representations of past episodes and objects. Spence (1988) also concluded that memories for episodes have "no clear boundaries"— neither in the details they describe, nor in the timing of the episode (when it starts and when it ends). He suggested that the "enabling context" which triggered the involuntary memory for the episode sets an "acceptance level," which is then used to filter out some aspects of the episode.

Barsalou has demonstrated that concepts also change their structure in different contexts. He suggested a context-sensitive representation of concepts—they are constructed on the spot rather than retrieved from memory (Barsalou, 1982; Barsalou & Medin, 1986; Barsalou, 1987; Barsalou, 1993). He studied the variability of the graded structure of concepts and demonstrated that it is highly context-sensitive. It varies substantially with changes in linguistic context and with changes in point of view. High variability occurs both within and between individuals

(Barsalou, 1987). Moreover, people can dynamically change their judgments of typicality when the context changes. In a related study Barsalou (1993) demonstrated context effects on the characterization of concepts. He came to the conclusion that "Invariant representations of categories do not exist in human cognitive systems. Instead, invariant representations of categories are analytic fictions created by those who study them" (Barsalou, 1987, p. 114). Furthermore, he claimed that "people have the ability to construct a wide range of concepts in working memory for the same category. Depending on the context, people incorporate different information from long-term memory into the current concept that they construct for a category" (p. 118).

The conclusion is that explaining the context-sensitive character of human memory for both episodes and concepts probably requires much more dynamic and flexible representations, which can be constructed on the spot rather than retrieved pre-packed from some static memory store.

2.2. Are There False Memories and Memory Illusions?

The extensive literature on this topic shows clearly that there is much evidence for false memories, that is, "memories" for aspects of events that did not occur. Moreover, in many cases people strongly believe in these false memories. False memories arise by two major means: either by blending two or more episodes, or by intrusions from some generic knowledge or schema. We will briefly review both aspects.

2.2.1. Blending of Episodes

The study of this phenomenon probably starts with the wave of research surrounding the interference theory of forgetting. Although the theory itself has long been forgotten, the experimental facts that were established remain important. Basically, these studies showed the interference between the traces of two learning events. The participants studied two different lists of items. Later on, at the test session, they mixed up items from the two lists. Just to mention one particular example out of many: Crowder (1976) has demonstrated an interference effect between pair-associations learned on two different occasions. A similar effect was observed by Deese (1959), who demonstrated false memories for non-studied but strongly associated items.

Loftus and her colleagues (Loftus, 1977, 1979; Loftus & Palmer, 1974; Loftus, Miller & Burns, 1978; Loftus, Feldman & Dashiel, 1995) developed a new paradigm for studying memory for complex real-world events such as crimes and accidents. These studies typically involve two sessions. On the first session the participants watch a slide show or a movie about some event, and on the second session they answer questions or listen to narratives describing the same event. The second session provides some misinformation about the event. It has been demonstrated that even though the context of learning and the sources were very different in the two sessions, there was blending between the two episodes in participants' memory. In a recent review, Loftus, Feldman and Dashiel (1995) report: "In some studies, the deficits in memory performance following exposure to misinformation have been dramatic, with performance difference exceeding 30%. With a little help from misinformation, subjects have recalled seeing stop signs when they were actually yield signs, hammers when they were actually screwdrivers, and curlyhaired culprits when they actually had straight hair" (p. 48). Moreover, the same authors have shown that in many cases people do believe they have really seen the mistaken element.

Neisser and Harsch (1992) have also demonstrated that people can have vivid memories and believe strongly in them though in fact they are false. They interviewed people immediately after the *Challenger* accident and asked them to write down a report of how they learned about the accident, what they were doing, where they were, and so on. One year later the experimenters asked the same subjects whether they still remember the accident and how they learned about it. People claimed they had very vivid ("flash-bulb") memories about every single detail. However, the stories they told on the second interview were often very different from the ones they had written on the previous one. Many participants were shocked when confronted with their original versions. Moreover, even in the face of this indisputable evidence (and what could be more convincing than an archive report written in one's own handwriting) some people still maintained that their second versions reflected better their memory of the accident. The exaggerated form of this memory distortion is called *confabulation* (Schacter, 1995b; Moscovitch, 1995). Neuropsychological patients with this symptom report their own biography in a very creative way. The misinformation effects of Loftus, the distorted *Challenger* reports told to Neisser and Harsch, and the confabulation of patients were attributed by Schacter (1995b) to the same possible cause: people's failure to distinguish between various sources of information about an event; that is to say from episode blending or *source confusion*. Because the pieces that are used in the memory-reconstruction process come from real (although different) episodes, the (false) memories constructed in this way can be very vivid and people can strongly believe they are real.

Blending of objects (as opposed to episodes) seems possible as well. Several experiments are particularly informative in this respect. McClelland and Mozer (1986) have shown that people can mix two items (words in this case) and produce a nonexistent item which is composed of phonetic elements from the original items (e.g. producing *land* out of *lane* and *sand*). Reinitz, Lammers, and Cochran (1992) presented people with human faces and asked them to learn them. Later on, on the test session, the participants were shown some novel faces that had not been presented before but were constructed out of elements of faces presented previously. This manipulation produced an illusion of memory for the novel faces (i.e., many participants "recognized" them as seen during the learning session). Finally, Nystrom and McClelland (1992) produced a blending of sentences which they called *synthesis errors*. About 10% of all errors were false recognitions of sentences in which one word came from one old sentence and another from a second one. The participants were asked to rate the confidence of their judgments, and 40% of the synthesis errors received the highest possible ranking. One particularly important observation that McClelland (1995) makes based on a simulation of these data is that "intrusions from the other sentence rush in when the most active trace provides no information" (p. 78).

2.2.2. Intrusions from Generic Knowledge

Another type of false memories comes from intrusions from generic knowledge. Thus Bartlett (1932) showed that episodes are remembered in terms of generic schemata and their representations are systematically shifted or changed in order to fit these schemata. He demonstrated, for example, the intrusions of expectations and rationalizations which were part of participant's schematic knowledge, but were not part of the real event (in this case a folktale). Research on autobiographical memory has also provided evidence that people use generic knowledge to fill in missing elements as well as to change existing elements in order to fit them into a schema (Barclay, 1986). It has also been shown that people systematically reconstruct their past in order to fit into their current self-image schema (Neisser, 1998; Neisser & Jopling, 1997).

Sulin and Dooling (1974) had their subjects read a brief paragraph about a wild and unruly girl. Then in one of the conditions they mentioned that the name of the girl was Helen Keller, whereas in the other condition they called her Carol Harris. Later on, they tested the rote memory of the participants for the sentences of the story. The test demonstrated robust false recognition of a completely novel sentence—"She was deaf, dumb, and blind"—in the first condition but not in the second. This intrusion obviously came from the generic knowledge the participants had about Helen Keller.

Loftus and Palmer (1974) demonstrated that subjects may claim they have seen broken glass in a car accident, whereas there was no broken glass in the slide show they had observed. Moreover, the percentage of subjects making this wrong reconstruction depended on the wording of the question (*smashed into* versus *hit*). In other words, the reconstructed episode contained intrusions from generic knowledge about car crashes. Similar results have been obtained in numerous other experiments summarized by Loftus, Feldman, and Dashiel (1995) as follows: "Subjects have also recalled non-existing items such as broken glass, tape recorders, and even something as large and conspicuous as a barn in a scene that contained no buildings at all" (p. 48).

Williams and Hollan (1981) used the think-aloud technique to study how people recollect the names of their classmates. They found that the participants in the experiment typically first looked for a specific context (e.g., a swimming pool or a specific trip), then searched this context to find the corresponding classmate(s) who were part of that context, and finally verified the information. Williams and Hollan described memory retrieval as a reconstructive and recursive process of problem solving. Partial information about a target item is used to construct a partial description of the item and this description is then used to recover new fragments. A new description is constructed and the process continues recursively. Obviously the result will depend on the starting point and in particular on the specific context in which the memory reconstruction takes place. Kolodner (1984) also found that people tend to construct details that they do not remember. The reconstruction is based on general schemata for similar events. Thus, for example, a person would say, "I must have gone to a hotel" and then possibly remember the specific hotel they were accommodated in.

Tulving (1983) also endorses the constructivist idea that memory traces result from a synthesis between stored information and current retrieval information. Schacter (1995b) provides additional data from brain studies and argues that the fact that many cortical areas are jointly involved in the recollection process suggests that information from various sources is being collected in order to reconstruct the episode.

Summarizing the results from this section, we may conclude that there are no clear-cut boundaries between episodes, or between episodes and generic knowledge. Episodes may become blended and elements of generic knowledge may be instantiated and implanted into an episode as if they had been part of the event. Which particular elements from other episodes or from generic knowledge will intrude depends on the context of recall.

2.3. Dynamics of Recollection and Order Effects

Recollecting an episode is not an instantaneous process. It takes time, which according to Anderson and Conway (1997) may run up to fifteen seconds in a laboratory experiment. Sometimes reminding is spontaneous, but recalling an episode may also be an effortful process. Even spontaneous memories come into our minds in portions.

As remembering is a slow and gradual process, we may be interested in the order in which various aspects of the event are being recalled. It turns out that this order may differ on different occasions (Salaman, 1982; Spence, 1988). The order in which the elements of the episode are recalled must have an effect on the mapping in analogy-making. We call these effects *memory order effects* (to contrast them with the order effects due to the timing of perceiving— see the end of section 2.4.3).

Ross and Sofka (1986), in an unpublished work, describe a protocol analysis they performed on remindings of old episodes. They presented subjects with problems and asked them which old problems they were reminded of. They found that reminding was slow and gradual rather than an instantaneous process, and that it runs in parallel and interacts with mapping. In particular, Ross and Sofka found that the subjects relied on the established mapping to recall details about the old episode. In other words, this study suggests that the mapping process (and, more broadly, reasoning) influences and guides the memory process.

Here is how Ross (1989) summarized these results: ”other work (Ross & Sofka, 1986) suggests the possibility that the retrieval may be greatly affected by the use. In particular, we found that subjects, whose task was to recall the details of an earlier example that the current test problem reminded them of, used the test problem not only as an initial reminder but throughout the recall. For instance, the test problem was used to probe for similar objects and relations and to prompt recall of particular numbers from the earlier example. The retrieval of the earlier example appeared to be interleaved with its use because subjects were setting up correspondences between the earlier example and the test problem during the retrieval” (p. 465).

This study was, however, performed in the context of a pure memory task. Subjects were not asked to solve the problems; they were rather asked to recall the problems they were reminded of. The next section looks at the complex interactions between memory, reasoning and perception in the context of problem solving.

2.4. Interplay between Memory, Reasoning, and Perception in Analogy-Making: Interaction Effects

Unfortunately, most of the research on memory has concentrated on deliberate and voluntary remembering. This applies both to the classical storehouse tradition and the constructive ecological tradition. The pure memory tasks, such as free recall, cued recall, and recognition tests, all have the drawback that they study memory in isolation. What we really need for understanding the complex interactions between memory and reasoning is the study of spontaneous remembering, that is, remindings that happen spontaneously in the context of a problem-solving activity. In particular, we are interested in spontaneous remindings of analogous situations and problems.

On the other side, the sparse research on memory within an analogy-making framework has ignored the constructive view on memory and has concentrated on how people select the most appropriate episode from the vast set of episodes in long-term memory. We will not review these studies in any detail because Hummel and Holyoak (1997) have done this already. We will only mention some basic findings. It has been established that the existence of similar story lines or similar objects (objects with similar properties) is a crucial factor for analogical reminding (Holyoak & Koh, 1987; Ross, 1989; Gentner, Rattermann, & Forbus, 1993). That is why remote analogies are very rare and difficult to achieve (Gick & Holyoak, 1980). However, Dunbar (chapter 9, this volume) demonstrates that people, both in natural settings and in the experimental laboratory, are able to produce remote analogies based on shared relations in both domains. Actually, the role of similarity between the relations in both domains has never been seriously studied. What has been studied and established is that structural correspondence (similar objects playing similar roles in similar relations) does not have much effect on reminding. It can possibly facilitate reminding under certain circumstances, but only when there is general similarity between the domains or story lines (Ross, 1989; Wharton, Holyoak, & Lange, 1996). Dunbar (chapter 9, this volume) and Ross and Bradshaw (1994) present evidence for encoding effects on remindings, that is, that reminding is facilitated when the subjects perform similar operations on the material at study and test, and when they focus on the same aspects (relations or properties) in both cases. Spencer and Weisberg (1986) have found context effects indicating that even the same or similar environmental context can facilitate reminding. Unfortunately, there is not much research on the dynamics of the process of reminding (or reconstructing), on the completeness and accuracy of the resulting descriptions of the old episodes, and on how these reconstructions depend on the target problem.

The following subsections review briefly some results obtained by the AMBR research group illustrating the possible effects reasoning can have on reminding, memory on reasoning, and perception on memory and reasoning.

2.4.1. Omissions, Blendings, and Intrusions in Spontaneous Remindings in Analogy-Making: Effects of Reasoning on Memory

A recent experiment looked at human memory in the context of analogical problem solving. It was designed as a replication of Holyoak and Koh's (1987) experiment 1. A think-aloud method was used, however, and the accuracy of the base story was measured as it was being recalled. The participants were college students taking an introductory cognitive science course. As part of the class on thinking, they discussed the radiation problem and its solution. Three to seven days later they were invited by different experimenters to participate in a problem-solving session in an experimental lab. They had to solve a version of the lightbulb problem. Almost all subjects (except one who turned out not to have attended the class discussing the tumor problem) constructed the convergence solution and explicitly (in most cases) or implicitly made analogies with the radiation problem. We were interested in how complete and accurate their spontaneous descriptions of the tumor problem story were.

It turned out that remembering the radiation problem was not an all-or-nothing event. Different statements from the story were recollected and used with varying frequency. Thus the application of several X-rays on the tumor was explicitly mentioned by 75% of the sixteen students participating in the experiment; the statement that high intensity rays will destroy the healthy tissue was mentioned by 66% of the subjects; and the statement that low-intensity rays will not destroy the tumor was mentioned by only 25%. Finally, no one mentioned that the patient would die if the tumor was not destroyed. All this demonstrates partial recall of the base. Our hypothesis is that the elements that tend to be reproduced are the ones that correspond to pragmatically important elements in the target. This hypothesis remains to be tested and corresponding experiments are under development.

On the other hand, there were some insertions, that is, "recollections" of statements that were never made explicit in the source domain description. Thus one subject said that the doctor was an oncologist, which was never explicated in the radiation problem description (nor should it be necessarily true). Another subject claimed that the tumor had to be burned off by the rays, which was also never formulated in that way in the problem description.

Finally, there were borrowings from other possible bases in memory. Thus one subject said that the tumor had to be "operated by laser beams" while in the base story an operation was actually forbidden. Such blendings were very frequent between the base and the target. Thus seven out of the eleven subjects who spontaneously retold the base (radiation) story mistakenly stated that the doctor used laser beams (instead of X-rays) to destroy the tumor. This blending evidently results from the correspondence established between the two elements and their high similarity.

In summary, the experiment has shown that remindings about the base story are not all-or-nothing events and that subjects make omissions, insertions, and blendings with other episodes influenced by the correspondences established with the target problem.

2.4.2. Priming: Effects of Memory on Reasoning

Memory in its turn, having its own life independent of reasoning, can influence the reasoning process. One example of this is the influence that our immediate or very recent past has on reasoning. Thus people are always in a particular memory state when they start solving a problem. This state is determined by what they have been doing and thinking about immediately before they switched to the new task. This state will typically be unrelated to the current problem but can nevertheless have an influence on how it is solved. This memory state is characterized by the person's currently active concepts, generic facts, rules, particular past episodes, goals, plans, and so on. In an attempt to partially control this memory state, Kokinov (1990, 1994a) carried subjects through a series of problem-solving tasks. The problems were chosen from a variety of domains (algebra, geometry, physics, commonsense, etc.), so that there were no apparent relations among them. The problems were presented to the subjects one by one and in different orders in the different experimental groups. Each presentation consisted of a series of ten problems, two of which were covertly related and hence anticipated to interact. The expected interaction was that the early problem would prime the other, that is, induce a memory state that would facilitate solving the later problem.

The experiment demonstrated that when the target problem was preceded by different priming problems subjects may solve it in different ways. The solution of the priming problem was known to the subjects in advance (it was a commonsense problem such as how to prepare tea in a mug). Therefore the only effect that this priming presentation had on the subjects was to make certain concepts, facts, rules, or episodes more accessible. This turned out to be crucial for the following problem-solving process, as the performance of the subjects in the task rose from 12% to 44%. In some cases we demonstrated that people can be influenced to find different solutions of the same problem depending on the specific priming provided. The experiment also studied the dynamics of the process by manipulating the length of the time interval between the priming and target problem (by making people solve distractor problems in between). The results showed that the priming effect decreased exponentially with the course of time and disappeared within about twenty-five minutes in this particular study. Thus immediately after priming the rate of successful performance was 44%, about five minutes later it declined to 29%, and after twenty-five minutes it was back at the control level of 12%. Schunn and Dunbar (1996) have also demonstrated priming effects on problem solving. Their results indicate that subjects were not aware of the priming effect.

Kokinov (1989) demonstrated that memory about general facts such as "which is the lightest chemical element?" is also sensitive to recent experience. The experiment demonstrated priming effects on recall of such general facts. Many experiments have demonstrated priming effects on particular concepts. For instance, studies in social psychology have demonstrated that a particular priming can affect the use of various prototypes in characterizing a person or person's behavior (see Bargh, 1994, for a review).

2.4.3. Context Effects: Effects of Perception on Reasoning

Based on a prediction derived from an earlier simulation of analogy-making (Kokinov, 1994a), the AMBR research group started to look for context effects, that is, how the perception of incidental elements of the environment during the problem-solving process can alter it. Thus Kokinov and Yoveva (1996) conducted an experiment on problem solving in which seemingly irrelevant elements of the problem solver's environment were manipulated. The manipulated material consisted of drawings accompanying other problems which happened to be printed on the same sheet of paper. There was no relation between the problems and the subjects did not have to solve the second problem on the sheet. However, these seemingly irrelevant pictures proved to play a role in the problem-solving process, as we obtained different results with the different drawings. We used Clement's (1988) spring problem as target:

> Two springs are made of the same steel wire and have the same number of coils. They differ only in the diameters of the coils. Which spring would stretch further down if we hang the same weights on both of them?

The problem description was accompanied by the picture in Figure 2.

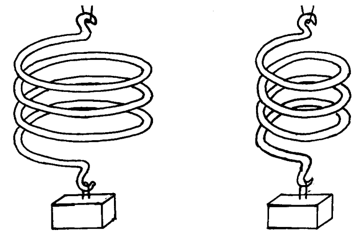


Figure 2. Illustration accompanying the target problem.

In different experimental conditions the drawings used to accompany a second unrelated problem on the same sheet of paper were different: a comb, a bent comb, and a beam (Figure 3).

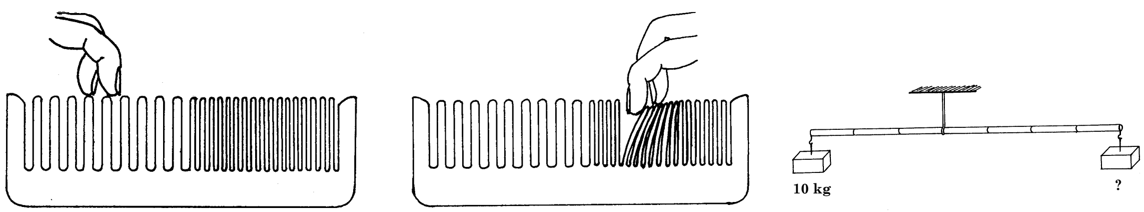


Figure 3. Illustrations accompanying the irrelevant problems in the various experimental conditions.

The results obtained in these experimental conditions differed significantly. In the control condition (no second picture on the same sheet of paper) about half of the subjects decided that the first spring will stretch more, the other half "voted" for the second one, and only a few said they will stretch equally. In the *comb* condition considerably more subjects suggested that the first spring will stretch more. In the *bent-comb* condition considerably more subjects preferred the second spring. Finally, in the *beam* condition more subjects than usual decided that both springs will stretch equally. Our interpretation is that the illustrations activate certain memory elements that, once activated, start to play a role in the problem-solving process. For example, the image of the bent comb probably activates concepts such as "bending" and facts such as "thicker teeth are more difficult to bend." This knowledge is then transferred (incorrectly in this case) by mapping teeth to springs, bending to stretching, and concluding that "thicker springs are more difficult to stretch."

Similar results, although not that dramatic, were obtained in the think-aloud experiment described in section 2.4.1. Subjects who had to solve the lightbulb problem were divided into two groups. In the control group there were no other problems on the sheet of paper, whereas in the context group the following problem was presented on the same sheet (Figure 4).

> The voting results from the parliamentary elections in a faraway country have been depicted in the following pie chart. Would it be possible for the largest and the smallest parties to form a coalition which will have more than 2/3 of the seats?

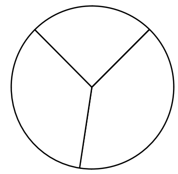


Figure 4. Illustration accompanying the context problem.

The results were the following: in the context group *all* seven subjects who produced the convergence solution to the lightbulb problem used *three* laser beams (7:0), while in the control group no one said three: two subjects said they would use *two or three* beams and the rest said they would use either *two* or *several* beams (2:5). The difference is significant at the 0.01 level.

Finally, Kokinov, Hadjiilieva, and Yoveva (1997) have demonstrated that subjects were not aware of the manipulations and the possible context effect of the second illustration. The context condition was contrasted with an explicit-hint condition in which subjects were invited to use the same picture during the problem-solving process. The results from the hint condition were significantly different. Moreover, in some cases when a hint was given to use the picture, subjects were less successful in solving the target problem compared to the control condition, while when they seemingly ignored the picture they were still influenced by it and showed better performance compared to the control.

The results from all the experiments described in this subsection demonstrate that sometimes perceiving small changes of a seemingly arbitrary element of the environment can radically change the outcomes of the problem-solving process (blocking it, or guiding it in a specific direction).

Another effect that perception can have on reasoning has been demonstrated by Keane, Ledgeway, and Duff (1994). They have shown that the specific order of perceiving the elements of the target can also influence the problem-solving process. We call these *perceptual order effects* to contrast with the *memory order effects* described in section 2.3. We hypothesize that the mapping process in its turn influences perception. For example, the currently established mapping may guide the attention and thus influence the selection of details to be perceived and their order. We do not have experimental support for this hypothesis yet. We call this potential influence *mapping effect on perception*.

The conclusion from this short review is that perception, memory, and reasoning strongly interact during the problem-solving process and must be studied and modeled together. The next sub-section attempts to summarize all these results and to describe the constraints they entail for models of analogy-making.

2.5. General and Specific Behavioral and Architectural Constraints on Models that Integrate Analogy and Memory

Let us briefly summarize the findings related to reminding of an analogical episode in a problem-solving context. The specific findings about remindings in analogy-making are reviewed by Hummel and Holyoak (1997). They are almost skipped in the present review inasmuch as they are well known; however, these findings are presented in Table 1. The foregoing review focused on more general characteristics of human memory that should be taken into account when modeling analogical remindings. These data, although well known as well, are often ignored in analogy models. They are also summarized in Table 1.

When modeling a cognitive process or subprocess we often focus on those data and characteristics that are highly specific for this process and we forget about features that cut across all cognitive processes. Because the focus of this chapter is on human analogy-making, we have to take into account both its specific and universal features. Moreover, we should not only be able to account for those universal features, but we should also model them in a unified way. Stated differently, our treatment of the universal features in models of analogy-making should allow equivalent treatment of the same features in models of other cognitive processes as well. This is analogous to the unified understanding of the role of blood in all parts of the elephant body presented in the introduction.

One such very important aspect of all human cognitive processes is their context-sensitivity, that is, their dynamic adaptation to the specific context. This property should be explained for memory, for reasoning, and for perception, in a unified way. Doing so requires that we build our models on a general cognitive architecture, and that this architecture provides basic mechanisms that ensure context-sensitivity of all cognitive processes.

Representations of episodes and generic knowledge should be appropriate not only for analogy-making, but for all possible cognitive processes that might need them. This does not mean that there should be unique and universal representations of episodes or concepts—on the contrary, people may well have several complementary representations of the same concept or the same episode. However, all representations should be accessible to all cognitive processes. Of course, some might be more suitable for one task than others. Table 2 summarizes the architectural constraints on analogy models.

Table 1. Behavioral constraints on modeling the interactions between analogy, memory, and perception

Type of Finding	Finding	Section in text
	similarity effect: semantic similarity between story lines, objects, properties, and possibly relations in both domains is crucial for analogical reminding	2.4.

Findings specific for analogy making	***structural effect:*** structural correspondence (similar objects playing similar roles) plays a very restricted role in analogical reminding and operates only when there is general similarity between the domains	2.4.
	encoding effect: similarity between encoding and test conditions (type of task and focus on similar aspects) plays a role in reminding	2.4.
	schema effect: the presence of generalizations of several analogous experiences from the past assists analogical reminding	2.4.
	familiarity effect: familiar analogs have advantage during reminding	2.4.
	memory order effect: the order of recalling the elements of the old episode influences the mapping	2.3.
	perceptual order effect: the order of perceiving the elements of the target influences the mapping	2.4.3.
	mapping effect on memory: the mapping process influences the recall of details of the old episode(s) and their order	2.3.
	mapping effect on perception: the mapping process influences the encoding of details of the target and their order (no experimental support for this potential effect)	2.4.3.
Findings about human memory in general	***omissions:*** details of the episodes are recalled selectively depending on the context	2.1. & 2.4.1.
	blending: episodes are blended; intrusions from other episodes take place, especially when important elements are not available in the dominant episode	2.2.1. & 2.4.1.
	schematization: intrusions from generic knowledge take place	2.2.2. 2.4.1.
	context-sensitive representation of episodes and objects (effects on reminding, recognition, priming)	2.1.
	context-sensitive representation of concepts	2.1.
	gradual recall and order of recall: episode elements may be recalled in different order	2.3.
	priming effects on episodes	2.4.2.
	priming effects on generic knowledge, including facts and concepts	2.4.2.
	environmental context effects: perception of accidental elements from the environment may play a role in reminding and mapping	2.4.3.

Table 2. Architectural constraints on analogy models.

Analogy models should be built on a general cognitive architecture.
Analogy models should be integrable with models of other cognitive processes.
Models of different processes and subprocesses should use unified representations.
A unified set of basic architectural mechanisms should support more complex mechanisms in models of different processes.
The cognitive architecture should ensure context-sensitivity of all cognitive processes.

Reviewing the existing models of analogy-making and especially those of them that involve reminding of an old episode—ARCS (Thagard, Holyoak, Nelson, & Gochfeld, 1990), MAC/FAC (Forbus, Gentner, & Law, 1995), AMBR1 (Kokinov, 1994a), and LISA (Hummel & Holyoak, 1997)—we will notice that they fail to incorporate most of the behavioral and architectural constraints described here[3]. Invariably these models use the storehouse metaphor of memory. Their long-term memory "stores" a collection of frozen representations of past episodes (prepared by the author of the model). One or more of these episodes are "retrieved" during the problem solving process and serve as a base for analogy. The very idea of having encapsulated centralized and frozen representations of base episodes is at least questionable, but it underlies most analogy-making models (Figure 5).

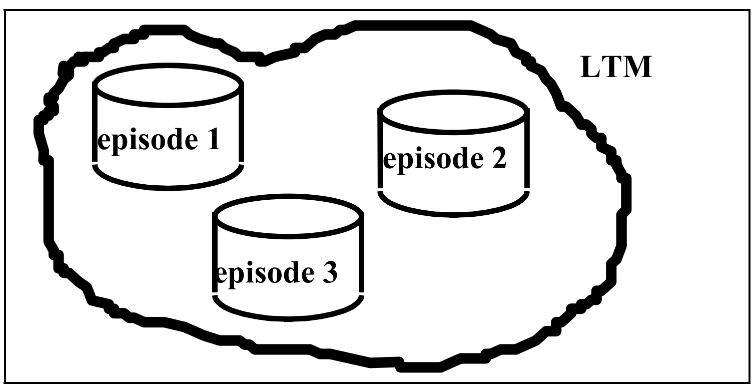


Figure 5. Centralized and frozen representations of episodes in long-term memory.

Both ARCS and MAC/FAC have centralized representations of past episodes, and the aim of the retrieval mechanism is to select the best one. The intactness and accuracy of the episode representation is taken for granted. Copycat (Hofstadter, 1984, 1995; Mitchell, 1993) and Tabletop (Hofstadter, 1995; French, 1995) lack episodic memory, but they do have more dynamic representation of concepts. The Metacat system (Marshall & Hofstadter, 1998; Marshall, 1999) stores problem-solving episodes in memory, but it also seems to do it in a very centralized way—

[3] At the same time, there are many pure-memory models that do try to capture some of the general behavioral findings listed in table 1. For example, Sparse Distributed Memory (Kanerva, 1988), MINERVA (Hintzman, 1988), CHARM (Metcalfe, 1990), and Trace Synthesis Model (McClelland, 1995). These models will not be discussed here because they do not address problem-solving issues.

by storing a package of variables. LISA is based on distributed representations, but only in working memory. The long-term memory consists of centralized localist representations of the episodes. Moreover, when retrieved in working memory all propositions of a given episode are switched from “dormant” to “active” state at one and the same moment. This implies that the system keeps for each episode a complete list of the propositions that participate in it. This amounts to a centralized and frozen representation. Thus even in this model, which relies on distributed representations, the episodes are static constructions—no omissions, no blending, no insertions are envisaged. However, we do believe that this model has the potential to be developed further to reflect these requirements, based on its ability for partial activation of memory elements. AMBR1 too is based on the storehouse metaphor and depends on stable and complete representations of episodes. Thus the current chapter presents the new version of the model—AMBR2—which has been developed further to meet these requirements.

3. Analogy-Making in a DUAListic Society: The AMBR View of Analogy

Associative Memory-Based Reasoning (AMBR) has been proposed as a model of human reasoning in problem solving, unifying analogy, deduction, and induction (Kokinov, 1988). Since its inception in 1988 the model has gradually been developed. The first fully implemented version that got up and running was reported by Kokinov (1994a). We will refer to it as AMBR1. Various simulation experiments on analogy-making and priming effects on problem solving were performed with it. The work on the model and the aspiration for generality and lack of ad hoc decisions led to the formulation of a core of general principles, representation scheme, and basic mechanisms that formed the general cognitive architecture DUAL (Kokinov, 1989, 1994b, 1994c, 1994d, 1997). Later on, an AMBR research group was established at the New Bulgarian University. The group developed a new portable implementation of both DUAL and AMBR. More importantly, it introduced many conceptual improvements and new mechanisms resulting into a new version of the model called here AMBR2 (Kokinov, 1998; Kokinov, Nikolov, & Petrov, 1996; Petrov, 1998; Petrov & Kokinov, 1998, 1999). In parallel with the modeling efforts, various psychological experiments tested some predictions of the model (Kokinov, 1990, 1992; Kokinov & Yoveva, 1996; Kokinov, Hadjiilieva, & Yoveva, 1997; Kokinov, 1998).

3.1. Basic Principles of the AMBR Research Program

The AMBR research program has always followed a number of methodological principles which have provided strategic guidance in our efforts to understand human cognition (Table 3). These principles set some very high requirements on the model design. Successive versions of DUAL and AMBR satisfied them to different degrees, often at very rudimentary levels. Many of the requirements are far from being completely satisfied yet. However, it is important to keep them in mind and to push the research closer and closer to their satisfaction. Or to put it differently, these principles make us aware of important limitations of our current models and specify the direction to look for better ones.

The first principle reflects our belief stated in the introduction that the time has come to reintegrate human cognition. This principle requires that analogy should be studied together with other forms of thinking, perception, memory, learning, and language. It is also very important to explore the interactions among these cognitive processes.

The second principle is a recursive application of the first one at the finer grain size of the various subprocesses of analogy-making. According to our current understanding these processes include representation-building of the target, analogical reminding, dynamic re-representation of the target and source, mapping, transfer, evaluation, and learning. The second principle dictates that all of them should be studied together and their interactions should be explored.

The third principle is an implication of the first two. It claims that in order to integrate analogy-making mechanisms and integrate human cognition as a whole we should not build small isolated models of separate "stages." We should rather combine the piecemeal models developed so far into bigger unified models based on a single cognitive architecture. This general architecture should ensure the compatibility of the models and their ability to interact. Moreover, it should bring harmony to the whole system of cognition, that is, it should ensure that the various models follow the same principles, use the same representations, and depend on a common set of basic mechanisms.

Apart from the methodological principles, the research program has followed certain principles which cannot be claimed to be the universal truth. These are decisions that the AMBR group has made in order to reflect some general behavioral constraints or particular philosophical views. We are fully aware that alternative principles can probably serve the same role, and that our selection reflects our personal views and choices. That is why we call them design principles.

The first design principle is based on our understanding that the dramatic context-sensitivity of human cognition as a whole and of human thinking in particular cannot be easily captured by models based on centralized control. Subtle changes in the environment or the memory state can result in abrupt changes in behavior. It is difficult to imagine a centralized system that accounts for that and does not fall prey to the frame problem. The central processor would have to go through all elements of the environment and assess their potential relevance to the problem at hand. Context-sensitivity seems to arise much more naturally within a distributed system where many small processors look for local changes in their respective elements of the environment and/or the memory state. The overall behavior of such system emerges from the local activities of the individual processors. We call a computation *emergent* when no explicit a priori specification of either what is computed or how it is computed exists in the system (Kokinov, Nikolov, & Petrov, 1996). Thus the first design principle calls for emergent context-sensitive computation.

The second design principle reflects the evidence presented in Section 2 that human memory does not consist of frozen stable representations of events and concepts. Much more dynamic, flexible, and context-sensitive representations are required. Thus the second principle proclaims the use of emergent context-sensitive representations. This means that the particular representation of the episode or concept used on particular occasion should emerge from the collective work of many smaller units and should reflect the context-relevant features and structures of the corresponding object of interest. Again it seems improbable that the representations of the many concepts and episodes needed on each particular occasion could be crafted by a centralized mechanism.

Finally, the third design principle reflects our belief in the need for complementary ways of describing human cognition. Such a complex object could hardly be explained by a simple and coherent set of principles or axioms. That is why we strongly believe that human cognition should be modeled by using two or more complementary approaches each reflecting certain aspects of the reality. So, we have adopted both symbolic and connectionist approaches (thus displeasing both camps). We have, however, integrated them at the microlevel, that is, at the level of small processing units, rather than at the level of cognitive processes. Having both symbolic and connectionist aspects at the micro-level in the underlying architecture makes both of them available for use by all cognitive processes.

Table 3. Methodological and design principles of AMBR and DUAL

Methodological Principles	integrating analogy-making with memory, perception, learning, reasoning, i.e., reintegrating human cognition
	integrating various subprocesses of analogy-making such as representation-building, analogical reminding, mapping, transfer, evaluation, learning, i.e., reintegrating analogy
	grounding the model of analogy-making in a general cognitive architecture
Design Principles	dynamic context-sensitive emergent computation
	dynamic context-sensitive emergent representations
	integrating symbolic and connectionist processing by microlevel hybridization

3.2. The DUAListic Society: A General Cognitive Architecture

Let us imagine that someone has the idea to establish an art museum in the capital of Utopia. The curator discusses it with friends, and some of them decide to join the project. These enthusiasts in turn solicit their friends or colleagues. Gradually a number of people get involved in the enterprise, each in a different way: some provide money, others expertise in a specific type of art, and so on. The level of participation also differs — some spend years on the project, others participate only incidentally; some donate a lot of money, others only give a small amount. The outcome of the whole project depends on so many people and circumstances that no one can foresee the result in advance.

Now, suppose the project was successful and the government of the neighboring country Antiutopia invites the same curator to build a similar art museum. Will the result be the same? Never! First of all, not all people who contributed to the first project will be interested in the second one for all sorts of reasons. But even if we imagine that exactly the same people carry out the second project, they will certainly build a different museum. The degree of their involvement will differ. Their experience with the first project will influence the choices they make on the second. Their resources and the timing of their contributions will differ as well. For example, if a philanthropist makes the same donation as before but does it a little earlier, the architect may start with a different budget and hence design a different building.

The DUAL cognitive architecture adopts a multi-agent approach to meet the design requirements listed in Table 3. Both computations and representations in the architecture are distributed over a big number of *microagents*. Each piece of knowledge is represented by a *coalition* of agents and each computation is carried out by a whole team of locally communicating agents. Moreover, these coalitions of agents are not fixed in advance. Instead, they are formed dynamically via communication among the agents, in a way that depends on the context. Thus in different contexts different groups of agents work on the same task (or slightly different groups but with different level of participation and with different timing), and may eventually produce different outcomes at the global level (Figure 6). This is how context effects on all cognitive processes are explained in DUAL (Kokinov, 1994b, 1994c).

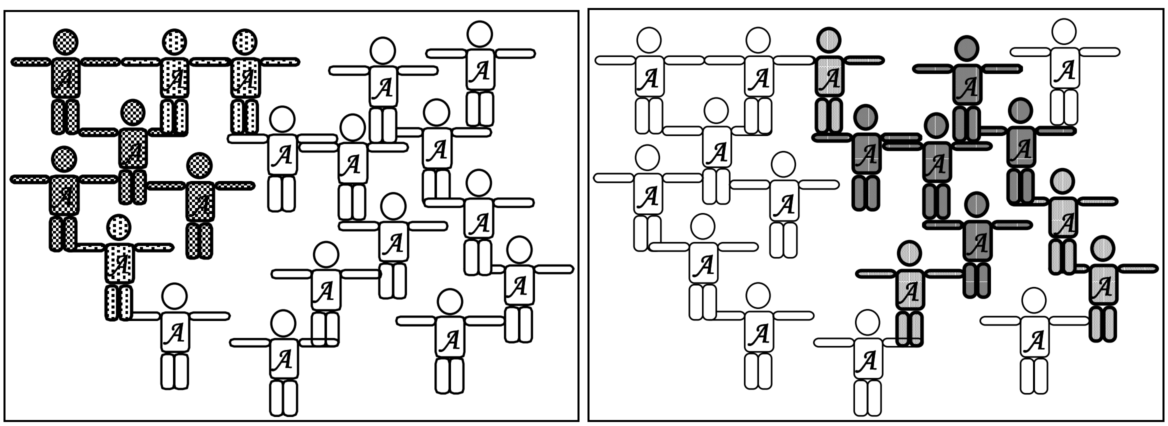


Figure 6. Different sets of agents are active and take part in the computation in different contexts. The filling pattern corresponds to the level of participation of the agent.

The DUAL agents are relatively simple and serve both representational and computational roles. A microagent might, for example, represent a simple proposition, or stand for a concept or a particular object. However, no agent possesses all the knowledge that the system has for that concept or object — it is distributed over several agents instead. The same agents carry out the information processing in the architecture. There is no central processor that operates on the agents; they do all the work themselves.

The participation of each agent in the whole process is graded. For example, the agent might loudly announce its knowledge so that all interested parties can use it. On another occasion the same agent might whisper so that only the closest and most attentive neighbors can hear it. The same principle of graded participation applies to the information-processing activities of the agents as well. An agent might be highly involved and work very fast on some tasks or be quite indifferent and work slowly on others. Even the same task may elicit different involvement in different contexts. The degree of participation of an agent depends on its motivational power. The motivational power reflects the relevance of the knowledge the agent has to the current task and context.

The microagents are hybrid. They consist of symbolic and connectionist aspects. The connectionist aspect calculates an activation level for each agent. This is how the "motivational power" suggested earlier is operationalized in DUAL. The activation level estimates the relevance of the agent to the current task and context. It is updated continuously according to connectionist rules.

Each agent has a symbolic aspect as well. It has a symbolic processor that can do simple symbol manipulations such as comparing two lists or sending a marker to another agent. Each agent interacts only with a few neighbors and any computation that spans over large populations of agents is carried out through massive exchange of messages. Communication is carried out through *links* between the agents: permanent or temporary. The same links are used both for connectionist and symbolic exchange—that is, for spreading activation and messages.

The activation level computed by the connectionist part is used to determine the speed of the symbolic processor. Active agents work quickly, moderately active agents work slowly, and the processors of inactive agents cannot run at all. This dualistic way of operation of the agents is very important. There are two separate but interdependent aspects of the computation—the connectionist aspect calculates context relevance while the symbolic aspect carries out the reasoning process. The two types of computation are done in parallel and influence each other. The context evolves continuously and provokes changes of the activation levels, which in turn alters the speed and availability of the symbolic processors, thus guiding the reasoning process. Reciprocally, the reasoning process sets new goals, shifts the attention to different aspects of the environment, and opens new lines for communication between agents. All this influences the activation levels calculated by the connectionist aspect.

Concepts, episodes, and objects are represented in a distributed way over a set of agents forming a coalition. The agents in a coalition are linked together so that when some members are active the remaining members tend to become active too. The weight of the link measures the strength of this coupling of the activation levels. Coalitions might be tight or weak depending on the weights of the respective links.

Finally, agents live in a big community that corresponds to the long-term memory of the system. Most agents are permanent but there are also temporary agents. There is a working-memory threshold. All agents, permanent or temporary, whose activation levels are above the threshold belong to the working memory. This active segment of the community is responsible for the outcome of all current computations. Most links within the community of agents are stable. They are established by the past experience of the system—something like old friendships or long-term business partnerships. The agents, however, can also establish new temporary connections. The possibility of establishing new temporary agents and links adds very important dynamism to the architecture. The topology of the network changes temporarily to adapt to the task. Table 4 outlines the meaning of some key DUAL terms.

DUAL has adapted the Society of Mind idea of Marvin Minsky (1986) as a basis for the cognitive architecture. The need for distributed and emergent computation and representation leads naturally to the idea that human cognition can be considered to be the product of the collective behavior of many simple microagents. Compared to Minsky's proposal, however, DUAL is more dynamic and less predetermined, inasmuch as new agents can be created on the fly and new links can be established between the agents. The emergent computation property of the DUAL system would also probably be at odds with some of Minsky's views.

It is closer to another recent implementation of the Society of Mind idea, namely the Copycat, Tabletop, and Metacat systems designed by Douglas Hofstadter and the Fluid Analogies Research Group (Hofstadter 1995, Mitchell 1993, French 1995, Marhsall 1999). These systems are also highly emergent and are based on the interaction between codelets that are very similar to DUAL agents. There are a number of significant differences between DUAL and these systems, however. While the working memory in DUAL is not a separate storage, Copycat and Tabletop maintain a separate storage area called Workspace where copies of the codelets run and construct representations. Another important difference is that DUAL is a deterministic system and the variability of its behavior derives from the ceaseless stream of influences from the environment and from the system's own recent internal states. In other words, the variations in context are responsible for the changes in behavior. In contrast, Hofstadter's systems are internally stochastic in nature and he believes that this is important for explaining creativity and human cognition in general.

Compared to a connectionist system, DUAL agents are more complicated and are not exact copies of each other, thus forming a heterogeneous system. Another difference is the dynamic re-organization of the network of agents described above. On the other hand, DUAL as it currently stands does not have learning abilities and its agents are predesigned by the programmer rather than evolving with experience. We would like to add learning capabilities to the future versions of the architecture.

Table 4. DUAL basic terms

DUAL term	Meaning
Agent (or microagent)	The basic computational unit in DUAL
Hybridization	Each agent has both symbolic and connectionist aspects
Communication	Via preestablished long-term links or via temporary links created on the spot. Both activation and symbolic structures are exchanged over the links.
Coalitions	Distributed representation of concepts, episodes, and objects
Large communities	Long term memory
Motivational power	Activation level as computed by the connectionist part of the agent; reflects the estimated relevance of the agent to the current context
Graded and variable participation	Variable individual speed of symbolic processing of each agent determined by its motivational power

3.3. The AMBR1 Model

The first version of the AMBR model (Kokinov, 1994a) integrated memory, mapping and transfer and simulated analogy-making in a commonsense domain—boiling water and preparing tea and coffee in the kitchen and in the forest. The most interesting example of analogy-making that this model addressed involved the following target problem.

> Suppose you are in the forest and you want to heat some water, but you have only a knife, an axe, and a match-box. You do not have a container of any kind. You can cut a vessel of wood, but it would burn in the fire. How can you heat the water in this wooden vessel?

This is not an easy problem for human beings. Only about 12-14% of the participants in several psychological experiments have been able to solve it (Kokinov 1990, 1994a). Solving this problem required that the participants recall a common situation involving heating tea in a plastic cup. All Bulgarian students participating in the experiments knew how to solve the latter problem using an immersion heater—an electric appliance that is put directly into the water and heats it without melting the plastic cup. This method of boiling water for tea is very popular in Bulgarian dormitories. Nonetheless, only 12% of the participants were reminded of this situation and were able to successfully make the analogy—to heat the knife and put it in the water. The reason is that the typical way of boiling water is by using a teapot on a hot plate. Most participants tried to use this source and failed to solve the problem, as the wooden vessel would burn in the fire. The priming studies described earlier used this same target problem, but as an experimental manipulation the subjects were primed with the plastic-cup problem in advance. The immediate priming raised the percentage of successful solutions to 44%. Four to five minutes after the priming the success rate dropped to 29%. Finally, after twenty-four minutes the priming disappeared and the results were at the base level of 12-14%. The simulation experiments with the AMBR1 model have replicated the qualitative trends of these data. Basically, without priming the model was not able to solve the problem. When primed with the immersion-heater situation it found the solution and the degree of this facilitation depended on the residual activation of the immersion heater situation.

The simulation experiments with AMBR1 have also made the prediction that if during the problem-solving process the subjects perceive a stone, they may use it instead of the knife for heating the water. This prediction was tested in a subsequent experiment (Kokinov & Yoveva, 1996). In this experiment an illustration of the situation in the forest has been added to the textual description and there were some stones to be seen by the river. The prediction was confirmed—the subjects who saw the illustration produced significantly more solutions involving stones than the subjects in the control condition (without illustration).

Thus AMBR1 has been successfully used in studying some interactions between memory (priming), perception (context effects), and reasoning (problem solving).

Remindings in AMBR1 are based on the connectionist mechanism of spreading activation. The sources of this activation are the perceived elements and the goals of the system. Mapping is a complex emergent process based on the local marker-passing and structure-comparison processes. Mapping is implemented by a form of constraint satisfaction network similar to ACME (Holyoak & Thagard, 1989). There are, however, a number of important differences that reflect our striving for psychological validity:

- The model has more realistic working-memory requirements because not all possible hypotheses are constructed, only those that seem plausible and relevant to the current context. Thus a hypothesis is constructed only when (and if) at least one agent finds a justification for it. The justification might be on the grounds of either semantic similarity or structural consistency.
- Mapping and memory processes run in parallel and thus can interact.
- The hypotheses are constructed dynamically. As different agents run at different speeds, some agents (the more relevant ones) establish their hypotheses earlier than others. This head start helps the early hypotheses gain activation.
- The constraint satisfaction network is constructed as part of the overall network of agents in the system. The activation can thus pass back and forth between the hypotheses and the representations of concepts and episodes. This allows for an interaction between memory and mapping tailored to the particular context.
- The semantic similarity is computed dynamically and is context dependent. The computations are done by a marker-passing process and the markers are guided, restricted, speeded up, or slowed down depending on the activation level of the agents which are processing the markers, that is, depending on the particular context.
- The structure-correspondence process is not limited by the **n**-ary restriction that was characteristic for all other models at that time (see Hummel & Holyoak, 1997; chapter 5 this volume). Once the semantic similarity between two relations has been detected, AMBR1 can map them even if they do not have the same number of arguments. This is because the marker passing mechanism disambiguates the correspondence between arguments of the two propositions. The disambiguation is based on the semantics of the arguments which is represented in the network of agents. LISA (Hummel & Holyoak, 1997) has recently solved the **n**-ary restriction in a similar way—the distributed representations of predicates capture the argument semantics.

The 1994 version of the AMBR model implemented only some of the AMBR principles as listed in table 3. AMBR1 is based on dynamic context-sensitive computation, but it has rigid and frozen representation of episodes. This is because there is an agent for each episode which points to all agents representing its various aspects. Thus the knowledge of the episode is distributed over a coalition of agents but this coalition is centralized – it has a leader which enumerates all the members of the group. This simplifies the mapping and transfer processes a lot because the system (and more specifically this agent) can use the list of mapped and unmapped propositions to guide the mapping. As we have argued in section 2, however, such a representation of episodes is psychologically implausible. This was one of the major reasons to develop a second version of the model.

3.4. The AMBR2 Model

The AMBR2 model described in more detail in the next section is a step further on the road delineated by the AMBR principles. Like its predecessor, it relies on emergent context-sensitive computations and implements them in an even more decentralized way. The big improvement, however, is that the episode representations are also emergent and context sensitive in AMBR2.

Concepts and objects are represented in the same way as in AMBR1—knowledge is distributed over a coalition of agents, but the coalition still has a leader which contains a list of the members (or more often of some of the members). The reason for having leaders of coalitions is that concepts and objects typically have names and thus these names are associated with the leaders. However, typically only part of the coalition becomes activated enough to become part of working memory, and thus we will use a partial context-dependent description of the concept or object as suggested by Barsalou (1993).

Episodes are, however, more complex and unique experiences and in most cases one cannot expect a name for an episode (other than using a general category name). Thus there is no need to have a leader of the coalition. For that reason in AMBR2 episodes are represented not only in a distributed but also a decentralized way. This means that no agent in the system knows all the agents of that coalition. Thus the coalitions become even more fuzzy and dynamic and even more susceptible to context influences.

Mapping and transfer are difficult to achieve in the absence of full lists of propositions on both sides. It is difficult to know what is mapped and what is not, when enough correspondences have been found, what remains to be transferred, and so on. “Difficult” does not mean “impossible,” however. Solutions to some of these problems have been found; for others they are still to be sought. The current version implements memory and mapping but not transfer. The simulations are in the same commonsense domain as AMBR1 but the knowledge base has been more than doubled. Both episodic and semantic knowledge has been added. These simulations explore the interplay between memory and mapping in various ways and demonstrate how most of the requirements listed in section 2 are fulfilled in AMBR2.

4. Integration of Memory and Reasoning in AMBR2

This section describes how the general architectural principles discussed earlier can be specified to produce a self-contained functional system—in this case a model of analogy-making. First the memory and reasoning mechanisms are described as emerging from the collective behavior of a set of agents. Then the interaction between memory and reasoning is explained. And finally, the section concludes with a brief description of several simulation experiments performed with the model.

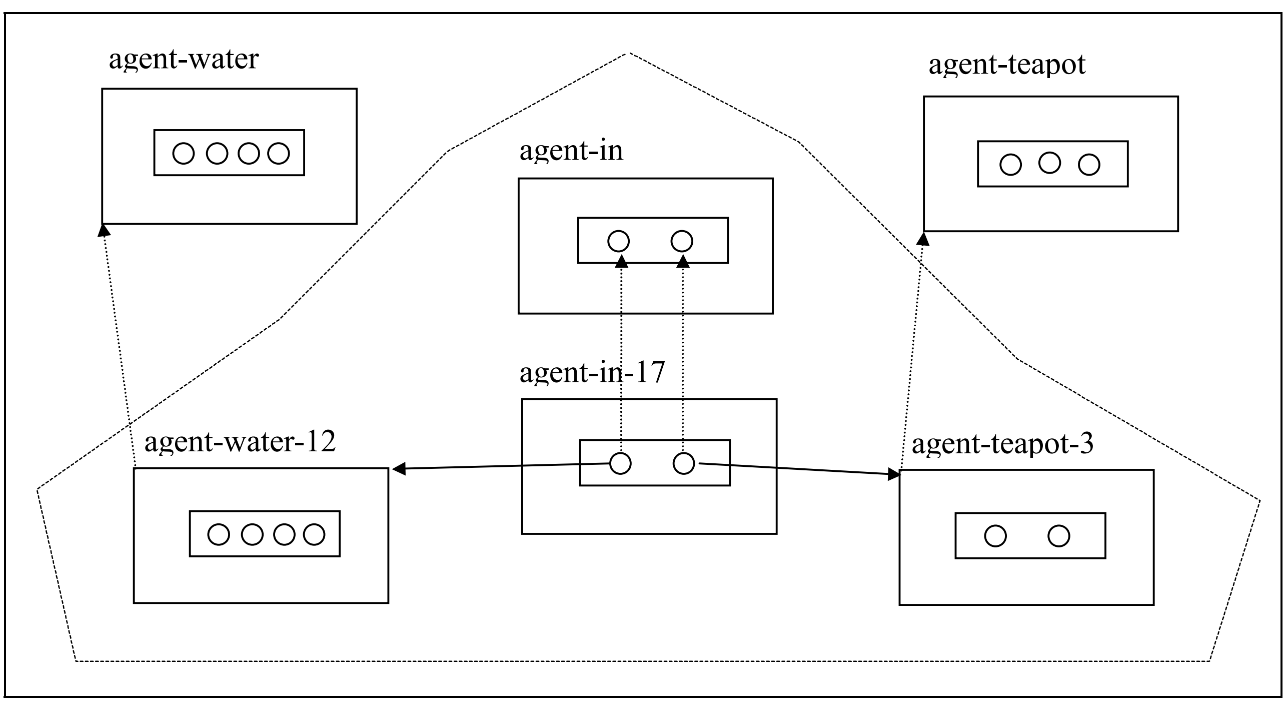
4.1. Collective Memory in AMBR2

Memory in AMBR is a collective phenomenon; just as in human society history is based on the memory of all members of the society. Each individual remembers a small piece of an entire event, a certain aspect of it from a certain point of view. Some individuals compare the versions held by others and draw conclusions about the relationships and correspondences. Thus the history of the event is gradually reconstructed and different individuals would offer different reconstructions. The global history emerges from all these local stories and is a collective product. Whenever a question about a certain event arises, the answer is constructed by the individuals who happened to be around with the partial knowledge they have. Thus there is never an ultimate truth about the event—each time the story is a bit different, but the stories also share a great deal. An interesting aspect of the AMBR view is that there are no historians—no special individuals write and keep the history. History is "written" and kept by the people who make it. Various individuals act in the social world. They communicate with each other and remember these communicative acts. Thus history is a by-product of acting.

4.1.1. Distributed and Decentralized Representations in AMBR

The representation scheme used in DUAL and AMBR is framelike, where the slot fillers are only pointers or lists of pointers to other agents (Kokinov, 1989). As a consequence the actual fillers are represented by separate agents. Thus even a simple proposition like "the water is in the teapot" will be represented by a small coalition of four agents (Figure 7). From a strictly connectionist point of view of this is a localist representation because it is symbolic. From a different perspective, however, it is also distributed because it is the whole coalition which represents the proposition and many different propositions will overlap their representations with this one, for example, "the teapot is green," "the water is hot," and so on. If it happens that only *agent-in*, *agent-in-17*, and *agent-water-12* are activated, the meaning will change, since this partial coalition will represent "the water is in something." This representation, although distributed, is highly centralized because there is a leader of the coalition (*agent-in-17*) that knows *all* coalition members.

Figure 7. Representation of the proposition "the water is in the teapot" by a coalition of four agents.



A particular object such as a stone would also be represented by a centralized coalition, with the leader standing for the object itself and each member standing for some of its properties or relations to other objects or classes of objects. In this case, however, the leader will not know directly all the members of the coalition; it might know only a few of them. Thus the leader would definitely not have a list of all agents representing the properties of the object, far less all its participations in various episodes.

Concepts are represented in the same way—distributed and centralized, with the leaders having only a partial list of the coalition members. Thus pieces of generic knowledge might be floating around in the space of coalitions and be associated with many coalitions but possibly not listed in any of them. For example, the fact that teapots are typically made of metal is a piece of generic knowledge that participates in several coalitions, such as the coalition representing the concept *teapot*, the coalition representing *metal*, and the coalition representing materials or *made-of* relations.

Another peculiar aspect of the representation scheme is the relationship between concepts and their instances. The leader of the coalition representing an object will probably always have a pointer to the concept (corresponding to the class of objects), but the leader of the coalition corresponding to the concept will only occasionally have a pointer to the object representation. The reason is that we consider it psychologically improbable to have complete list of all instances of a given category. Moreover, such a huge number of links from the concept to its instances would render these links useless, because the fan-out effect prevents any activation whatsoever to reach the instances. That is why a more flexible decision was taken, namely that such "top-down" links are established to a very restricted number of instances—the most familiar ones and the most recently used ones. As time passes by, different sets of instances will be pointed to because of the different recent usages or because new instances became familiar. This organization of the knowledge has an impact on the reminding process, because seeing a stone in the target situation will not automatically activate all stones and therefore all situations involving stones (an assumption that is true for models like MAC/FAC, ARCS, and LISA).

Finally, the episodes are represented in a distributed and decentralized way. They are represented by rather big coalitions than do not have leaders, that is, none of the members of the coalition has a list (even partial) of its members. There is a special member of the coalition which "stands for" the particular time and place location (it may be considered as a simple unique tag rather than a vector in some abstract space) and all members of the coalition point to it. This is the only way in which one can recognize that all these agents represent aspects of the same event. However, there are no pointers coming out of this special agent, that is, it does not list any of the coalition members.

Goals are represented as propositions that have the special tag of being a goal of the system. Thus whenever they get activated they are recognized as goals and put on the goal list. New goals can be established by the reasoning mechanisms or old goals can be reactivated.

4.1.2. Spreading Activation

The connectionist mechanism of spreading activation is the basic memory mechanism. Because the activation level of an agent determines its participation in both the representation and computation process, this activation mechanism has a pervasive influence on all other processes. It calculates a dynamic estimate of the relevance of each individual agent to the current context as defined by the goals, perception, and memory state. Based on this estimated relevance, it determines the motivational power and therefore the level of participation of each agent (i.e., its speed of processing and visibility to other agents). Because the outcomes depend on the participation level of all agents and its timing, we can describe AMBR functioning as context-guided emergent processing.

The connectionist processor of each agent computes the activation level and output activation from its inputs. There is spontaneous decay that forces each agent to lose activation according to an exponential law in the absence of external support. The particular activation function used is described by the following equation:

$$\left| \begin{array}{l} \dfrac{da}{dt} = -d \cdot a(t) + E \cdot net(t) \cdot [M - a(t)] \\ a(t_0) = a_0 \end{array} \right.$$

where ***a(t)*** is the activation level as a function of time, ***net(t)*** is the net input to the agent, ***M*** is the maximum activation level, ***d*** is the decay rate, and ***E*** is a parameter determining the excitation rate. In addition, there is a threshold (not explicated in the equation above) that clips small activation values back to zero. The sources of activation are the input and goal nodes. The input node is linked to all agents corresponding to elements of the environment that are currently perceived, and the goal node is linked to all the agents-leaders of coalitions that represent a currently active goal. Because the decay rate is low, there are significant amounts of residual activation. Thus the "previous" memory state influences the current one, giving rise to priming effects.

There are only excitatory links in the long-term memory. Inhibitory links are also built dynamically during processing, for example, in the constraint satisfaction network described in the next subsection. In this later case spreading activation is used for relaxation of the constraint satisfaction network.

4.2. Collective Reasoning in AMBR2

This subsection describes the mechanisms for mapping which result from the collective behavior of many agents in the system. Mapping is performed by gradually building and relaxing a constraint satisfaction network (CSN) similarly to ACME (Holyoak & Thagard, 1989, 1995). In sharp contrast to ACME, however, the network is built incrementally and in a distributed way by the independent operation of many agents that base their decisions only on local information. The CSN's function is to integrate the local opinions of the various agents and find a globally consistent mapping at the level of the coalition of hypothesis. It consists of temporary *hypothesis* *agents* and temporary excitatory and inhibitory links between them. In contrast to ACME, however, this net is tightly interconnected with the main network of permanent agents. Hypotheses receive activation from permanent agents and pass activation back to them. This feature ensures that the CSN works in harmony with the rest of the system and integrates this mechanism with others. Suppose, for example, that a particular concept is highly relevant in the current context. This is reflected by a high degree of activation of the corresponding agents in its coalition. This results in building more and stronger hypotheses based on that concept. And vice versa, if a particular hypothesis gains a lot of support and becomes very active, it activates the concepts and episodes that are linked to it and thus fosters the establishment of more and stronger hypotheses of a similar type (related to the same concept or episode).

Let us now describe briefly the main participants in the construction of the CSN. Although, it can be said that practically all active agents at a particular instance of time participate in the construction of the network, we can separate two main mechanisms for constructing new hypothesis agents: the *marker-passing mechanism* and the *structure-correspondence mechanism*. In addition, other mechanisms are responsible for synchronizing the network construction and avoiding duplication of hypotheses, inasmuch as they are built by decentralized local mechanisms. Next, mechanisms responsible for the promotion and selection of the winning hypotheses will be described. And finally, mechanisms for integrating generic knowledge in the mapping process will be presented.

4.2.1. Computing Semantic Similarity Dynamically by a Marker-Passing Mechanism

Each permanent agent in the system is capable of marker passing. Whenever it receives some markers it passes them over to its neighboring superclass agents with a speed proportional to its activation level. Whenever an agent that is the leader of a coalition representing an instance (object, property, or relation) enters the working memory, it emits a marker. This marker propagates upward through the superclasses hierarchy (there might be more than one superclass of a given class). It signals in this way indirectly to other agents the presence of an instance of that particular type. An intersection of two markers originating from two different instances (one from the target and another from permanent memory) means that these instances belong to the same class at a certain level of abstraction and thus are considered similar. This provides a justification for establishing a hypothesis that these two instances might correspond. The agent that detects the intersection constructs a new temporary agent representing such hypothesis. In this way semantic similarity between relations, properties or objects in both domains plays a role in the CSN construction. Moreover, because the speed of processing of markers depends on the relevance of the corresponding agents to the current context (estimated by their activation level), the similarity computed in this dynamic fashion is context-sensitive.

4.2.2. Ensuring Structural Consistency by a Local Structure Correspondence Process.

The structure correspondence mechanism is based on the ability of hypothesis agents to construct other hypothesis agents that will correspond to hypotheses consistent with the one they are standing for. There are both top-down and bottom-up hypothesis construction. Top-down construction is initiated when a hypothesis is established that two propositions correspond to each other. This should result in constructing hypotheses about the correspondence of their parts (e.g., arguments) as well as constructing excitatory links between them. Bottom-up construction is initiated when a hypothesis is established about the correspondence between instances of two concepts. This should result in establishing correspondences between the concepts themselves. If such a more general hypothesis is established this will facilitate the construction of more hypotheses at the instance level of the same type or will make them stronger. For example, in the preceding case when the two propositions are put into correspondence, this will result in the construction of a hypothesis about the corresponding relations of which they are instances. This will facilitate the later construction of other hypotheses about correspondences between propositions involving that same relations. All this work is performed locally by the hypothesis agents once they have been established. This mechanism ensures the emergence of global structural consistency in the winning hypotheses from the CSN as prescribed by the systematicity principle (Gentner, 1983).

4.2.3. Consolidating the CSN: Secretaries and Life Cycle of Hypothesis Agents

The fact that the hypotheses are established locally by individual agents complicates things, because it is perfectly possible that two independent agents find different justifications to establish one and the same correspondence (e.g., semantic similarity vs. structural consistency). This would result in establishing two different hypothesis agents standing for the same correspondence but competing with each other. To avoid this AMBR possesses certain mechanisms for merging such duplicate hypotheses. Instead of two agents with one justification each, the system ends up with a single hypothesis with two (and then three, etc.) justifications.

AMBR2 achieves all this by means of local interactions only. The so-called *secretaries* are instrumental in this respect. Each permanent agent keeps track of the hypothesis agents relating to it. To simplify the presentation we can assume that there is a secretary associated with each agent. (In the actual implementation each agent does all the bookkeeping itself.) All hypotheses are created as *embryo hypotheses*. Each embryo issues “registration requests” to the respective secretaries. The latter check their records and determine, locally, whether the hypothesis represents a unique correspondence or duplicates an existing one. In the former case the embryo is allowed to become a *mature hypothesis*. In the latter case the embryo resigns in favor of the established hypothesis that represents the same correspondence. The secretaries make sure they handle all links dealing with justifications, with non-identical but conflicting hypotheses, and so on. The net effect of their coordinated efforts is that the constraint satisfaction network is built gradually by decentralized addition of nodes (i.e., hypothesis agents) and links.

4.2.4. Dynamic Promotion and Selection of Winning Hypotheses

The phases of building the CSN and its relaxation are not separated in AMBR. The secretary of each object, relation, or concept maintains a current winner hypothesis at each point in time. This allows the transfer and evaluation processes to start in parallel with the mapping; they need not wait until it finishes. This opens the possibility for backward influences of the transfer and evaluation processes on the mapping. For example, it may turn out that the currently winning hypothesis is not interesting or not valid in the target domain and thus can be abandoned at a relatively early stage of the mapping. The process of selecting the best hypotheses is continuously running and is performed locally by the secretaries of the agents. Because they have registered all hypotheses that involve the current agent, they may decide which of these hypotheses is the most promising one. Of course, one would like to avoid a very early decision that cancels all the efforts by other agents to construct alternative hypotheses. On one hand, one would like early-established hypotheses to have some priority, because their early construction reflects the fact that the agents who constructed them have been highly active and therefore highly relevant to the context. On the other hand, hypotheses that arrive later might form a better and more consistent coalition that might provide a better global match. That is why the hypotheses are rated continuously by the secretaries, but promoted only gradually depending on many factors including the strength of their competitors and the duration of the time period in which they have led the competition. Thus if a hypothesis maintains its leading status long enough and is sufficiently ahead of its competitors (in terms of activation), it is promoted into a *winner hypothesis* and the evaluation and transfer mechanisms may use it as a starting point.

4.3. Interaction between Memory and Reasoning in AMBR2

This section describes several simulation experiments performed with AMBR2 that illustrate the interactions between memory and reasoning, and in some cases also perception, in the process of analogy-making. The experiments are of two types: case studies and aggregate statistics. The case studies track certain runs in detail, zooming into the specific mechanisms of the model. Aggregate statistics are collected over hundreds of runs of the system and disclose its overall tendency to produce certain solutions more readily than others. In the latter case we exploit the fact (described in section 4.1.1) that there could be only a restricted number of links from general concepts to their instances. Thus, one hundred variations of the knowledge base have been generated by randomly sampling which instances are connected and which are not. In addition, some associative links have also been established at random. Only about 4% of the approximately three thousand links in the overall long-term memory are changed from run to run, but as the results that follow will show, these changes are enough to produce a wide variety of solutions to identical target problems.

4.3.1. Perceptual Order Effects

Suppose a student reads the description of some problem from a textbook. The text is read sequentially and the internal representation of this text would tend to be constructed sequentially too. In the AMBR2 model this process can be crudely approximated by attaching the temporary agents representing the target sequentially to the activation sources of the system (i.e., the goal node and input node). In a more elaborated model these elements will be constructed by the perceptual mechanisms. When some target elements are perceived and/or declared as goals earlier than others, they start receiving activation earlier. This enables them in turn to activate their coalition partners in the network. These agents enter the working memory more vigorously than the agents related to the target elements that have not been perceived yet. Moreover, earlier elements establish hypotheses earlier, which in turn reinforces their advantage. The net result is that the order of presentation of the target problem will affect all subsequent work on the problem. Specifically, source analogs involving elements which are semantically similar to a given target element are used more frequently when this target element is presented earlier to the system.

A simulation experiment was designed to highlight this order effect. The experiment consisted of three conditions involving the same target problem:

> There is a teapot and some water in it. There is an egg in the water. The teapot is made of metal. The color of the egg is white. The temperature of the water is high. What will be the outcome of this state of affairs?

The long-term memory contained many episodes, three of which were most related to this particular target. Two episodes dealt with heating liquids and one with coloring Easter eggs. The target problem was run three times on the set of one hundred knowledge base variants, yielding a total of three hundred runs. In the control condition all target elements were presented simultaneously to the system at the beginning of the run. In the *hot water* condition the agents representing that the water was hot were presented first, followed after a certain delay by the agents representing the teapot and its material. The *color-of* relation was presented last. In the *colored egg* experimental condition the agents were presented in reverse order. The dependent variable was the frequency of activating and mapping the various source episodes.

The results were straightforward. In the control condition 48% of the runs were dominated by one of the two water-heating source analogs and 35% by the red-egg analog. When the target elements involving high temperatures were presented early (the *hot water* condition), these percentages changed to 74% and 5%, respectively. On the other hand, when the presentation began by the proposition that the color of the egg was white (the *colored egg* condition), the frequencies were 18% vs. 67%. Given that all runs involved exactly the same target problem and the same set of one hundred knowledge base variants, the experiment demonstrated clearly that AMBR2 was sensitive to the order in which target elements are presented to the system.

Thus the interaction of the subprocesses of perception, episode recall, and mapping in AMBR predicts *perceptual order effects* in analogy making. A psychological experiment testing this prediction is currently being carried out by the AMBR research group.

4.3.2. Influence of Mapping on Episode Recall

As stated throughout this chapter the various subprocesses of analogy-making in AMBR run in parallel and can interact. The interaction takes different forms, including influences that supposedly later “stages” exert on supposedly earlier ones. This subsection reviews a case study that focuses on the influence of mapping on episode recall. The full details of this simulation experiment are reported elsewhere (Petrov & Kokinov, 1998).

Such “backward" influences seem strange at first glance. How can a system map a source episode to the target if the source has not even been retrieved? The key here is that episodes are represented by decentralized coalitions in AMBR2 and thus can be brought to the working memory element by element. As soon as some members of a coalition become active, the mapping mechanisms can start constructing hypotheses relating these elements to various elements of the target. If these hypotheses do well in the constraint satisfaction network, their activation levels rise and part of this high activation propagates back to the LTM members that have generated them. In other words, if some (partially recalled) propositions from some source episode turn out to be structurally consistent with some target propositions, the source elements receive additional support from the constraint satisfaction network. This allows them to bring more of their coalition members above the working memory threshold. The latter then construct new hypotheses thus opening new opportunities to receive activation from the highly active target elements and so forth.

A simulation experiment was designed to highlight and test this sequence of mutual facilitation. It consisted of two experimental conditions, both of which solved the same target problem over exactly the same knowledge base. In the *parallel condition* the AMBR model operated in its normal manner—the mechanisms for mapping and memory worked in parallel. In the *serial condition* the mechanisms were artificially forced to work serially—first to activate episodes from memory, pick up the most active one, and only then map it to the target. The model produced different results in these two conditions. When all mechanisms worked in parallel, they succeeded in identifying a structurally isomorphic analog, activating it fully from LTM, and mapping it to the target problem. The serial condition resulted in activation of a superficially similar but structurally inappropriate base. (The relations that were crucial for successful transfer of the solution were cross-mapped.) This simulation not only explains the mapping effect of recall, but also sheds light on the mechanisms of the structural effect (Table 1). Other models (MAC/FAC, ARCS) have to incorporate patches which perform partial mapping in order to explain the structural effect. AMBR2 explains it just by the fact that both recall and mapping run in parallel and thus mapping can influence recall.

4.3.3. Blending of Episodes

More than 1,300 runs of the AMBR2 system have been performed on different target problems and with different concepts and episodes in LTM. A typical pattern in these simulations is that early during a run the spreading activation mechanism brings to the working memory an assortment of agents belonging to different episodes. These elements are recalled from LTM based solely on their semantic similarity to some target element. As more and more hypothesis agents are being constructed, however, the constraint satisfaction network begins to influence the pattern of activation over the entire community of agents. The dynamics of the CSN usually drives it into a state of minimum energy that corresponds to a consistent mapping between the target and one specific source episode.

Occasionally, however, the system produces blends in which two or more sources are partially mapped to the target. The exact conditions for the emergence of such blends are yet to be explored but the simulations so far have revealed that they are certainly possible, albeit rare. Blends tend to happen when none of the episodes in the long term memory matches the target well enough or when the appropriate episode is superseded by another one (e.g., as a result of a priming or context effect). Under these circumstances one of the sources maps to some fraction of the target and another source maps to the rest. This is possible in AMBR because the mapping is done element by element and the pressure to stay within the dominant source episode is soft (i.e., implemented via the constraint satisfaction mechanisms) rather than enforced in an all-or-none fashion.

4.3.4. Incorporating Generic Knowledge into Episode Representations: The Instantiation Mechanism

The instantiation mechanism extends the episode representations with elements derived from generic knowledge. This is a kind of re-representation of the episode performed during recall and under the pressure of mapping (Kokinov & Petrov, 2000). The instantiation mechanism thus exemplifies the interaction between memory and reasoning in one of its most sophisticated forms. Memory, deduction, and analogy meet together at this point. The episode representation is partially recalled from memory and partially inferred from generic knowledge, whereas the whole reconstructive process aims at aligning the episode with the current target.

The main ideas behind the instantiation mechanism are the following. The spreading activation typically brings agents belonging to various coalitions into working memory. Some of the agents belong to coalitions representing various episodes; other agents belong to coalitions representing generic knowledge. Each agent undertakes various actions whose ultimate goal is to establish a correspondence between the agent in question and some agent from the target problem. These actions include emission of markers, creation of hypotheses, and “acts of cooperation” within the coalition (e.g., sending activation to poor members). Not all aspirations of the agents can be satisfied, however, because the target agents act selectively (and thereby press for one-to-one mapping). This generates competition for the “valences” of the target problem. The epicenter of this competition is in the constraint-satisfaction network, but it reverberates throughout the working memory because the success of the hypotheses in the CSN depends on the support they receive from the other agents, and vice versa.

Two scenarios are possible at this point. The first happens when there is an episode that can use up all valences of the target, and in addition all members of the coalition representing this episode have been activated and held in working memory. Under these circumstances the hypotheses relating this episode to the target will form a complete and coherent set of pairwise correspondences and are likely to win the competition. Sometimes, however, the dominant episode cannot saturate all valences of the target. This leaves some target elements with no counterparts in the (active portion of the) dominant episode. These free valences then invite elements from other coalitions to intrude. If the intruders come from other episodes, we get blending. If the intruders represent pieces of generic knowledge, they become starting points for the instantiation mechanism.

Suppose, for example, that the target problem involves a bowl and it is explicitly represented that this bowl is made of wood. Suppose further that the episode that currently dominates the mapping involves a teapot but no information about the material of this teapot is available in the working memory. This might be either because this information has never been attended and encoded, or because it is represented by a loose part of the coalition and fails to reach the threshold. Finally, suppose the generic knowledge that teapots are typically made of metal has been activated (due to the salient *made-of* relation in the target). Under these circumstances the working memory contains agents (organized in small coalitions) representing the two propositions that, on one hand, teapots are generally made of metal and, on the other hand, the target bowl is made of wood. A hypothesis representing the tentative correspondence between these two propositions is established in the CSN. In the absence of any strong competitor from the dominating base episode, this hypothesis gains activation and hence comes on the top of the list maintained by the secretary of the *made-of* proposition in the target. The rating performed by this secretary detects that the top hypothesis involves a generic statement and triggers the instantiation mechanism by sending a message to the respective hypothesis-agent.

The instantiation process is carried out via a complicated sequence of messages exchanged between the agents. The net result of this process is that a *specific proposition* is generated to replace the *general proposition* currently mapped to the (specific) proposition in the target. In the example above, the new proposition states that the specific teapot in the base episode (rather than teapots in general) is made of metal. New temporary agents are constructed to represent this new proposition. In other words, the representation of the base episode is extended to include a statement inferred from generic knowledge. The new elements added to the episode representation can be both relations and objects. The instantiation mechanism tries to use existing agents from the old coalition whenever possible and generates new agents only upon necessity. In our example, the existing teapot will be used because it already corresponds to the bowl in the target. (This is the same bowl that is made of wood and that introduced *made-of* relations to begin with.)

Once the agents representing the new proposition are added to the working memory, they carry out the same activities that all permanent agents do upon entering WM. In other words, the mapping mechanism operates uniformly across all elements—it does not matter whether they are activated from LTM (gradually over time) or are constructed by instantiation (gradually over time). However, there is a built-in bias in favor of hypotheses about specific propositions over hypotheses about general ones. In addition, the new specific instances receive strong support from their coalition members because the episode overall has strong positions in the competition. Thus when the instantiation mechanism adds specific propositions to WM, the respective specific hypotheses tend to replace the hypotheses about general propositions even though the latter have appeared earlier in the constraint-satisfaction network.

In summary, the instantiation mechanism augments the description of an episode with objects and propositions that are specific instances of some generic concepts and propositions. On one hand, the specific propositions constructed in this way can be considered as deductions from generic knowledge. On the other hand, however, they are constructed only when needed to fill some free valences in the target, that is, guided by the analogy. That is why the instantiation process is a nice example of the interplay between deduction, analogy, and memory.

It is easy to see how the instantiation mechanism can run in the complementary direction too (although this feature is not implemented in the existing version of AMBR). The same basic sequence of events, with slight modifications, can be used to augment the description of the target so that it aligns better with the past episode that currently dominates the mapping. This constitutes a form of analogical transfer that is also backed up by generic knowledge and is yet another nice example of the interplay between deduction, analogy, and memory.

5. Conclusions

This chapter tries to draw a bridge between analogy and memory research. Based on the findings established in both areas we have presented the behavioral and architectural constraints that, in our view, realistic models of analogy-making should reflect. These constraints are summarized in tables 1 and 2. The AMBR research program was presented as a step-by-step attempt to build a model satisfying these constraints. Finally, the current version of the model—AMBR2—was described, along with a discussion of how it faces some of the challenges to cognitive models of analogy-making. The explanations provided by AMBR2 to these challenging phenomena are briefly summarized in table 5.

Table 5. Explanations provided by AMBR2 to the phenomena listed in table 1 as challenges to analogy models.

Findings	Explanation provided by AMBR
similarity effect: semantic similarity between story lines, objects, properties, and possibly relations in both domains is crucial for analogical reminding	Reminding is based on the spreading activation mechanism which is sensitive to similarity. There is no difference between properties and relations in that respect. The only requirement is that the element is encoded in the episode representation.
structural effect: structural correspondence (similar objects playing similar roles) plays a very restricted role in analogical reminding and operates only when there is general similarity between the domains	This effects is explained by the parallel work of mapping and memory and the backward influence of mapping on reminding as described in section 4.3.2.
encoding effect: similarity between encoding and test conditions (type of task and focus on similar aspects) plays a role in reminding	There are two reasons for this effect. First, as explained above, relations (or properties) have to be encoded; otherwise the spreading activation mechanism cannot activate them. Second, since agents represent both declarative and procedural knowledge, the operations performed by the agents, if the same in the two conditions, can facilitate processing.
schema effect: the presence of generalizations of several analogous experiences from the past assists analogical reminding	In this case activation needs to spread only in one direction—from instances "up" to class descriptions—and thus it avoids the insecure way "down." The way down is insecure because of a fan effect and because each AMBR concept has explicit links to only a few instances rather than all of them (section 4.1.1).
familiarity effect: familiar analogs have advantage during reminding	The more familiar an episode, the stronger the coalition, and the stronger the links to it (both "top-down" links from concepts and "lateral" links from other episodes).

perceptual order effect: the order of perceiving the elements of the target influences the mapping	Target elements that are encoded earlier can establish hypotheses earlier (section 4.3.1). Early hypotheses have a head start in the constraint satisfaction network.
memory order effect: the order of recalling the elements of the old episode influences the mapping	The earlier an element passes the working-memory threshold, the earlier it gets a chance to establish hypotheses and participate in the mapping. Early hypotheses have a head start in the constraint satisfaction network.
mapping effect on memory: the mapping process influences the recall of details of the old episode(s) and their order	This effect is explained by the parallel work and interaction between memory and mapping. The backward influence of mapping has been simulated as described in section 4.3.2.
mapping effect on perception: the mapping process influences the encoding of details of the target and their order	The current version of AMBR does not account for this effect yet because of its rudimentary perceptual capabilities. In a future version the perceptual subprocess will run in parallel with mapping (and with everything else) and will be influenced by it.
omissions: details of the episodes are recalled selectively depending on the context	Most episodes are represented by relatively loose coalitions. In such coalitions the activation of a few members does not necessarily bring the remaining members above the threshold.
blending: episodes are blended; intrusions from other episodes take place, especially when important elements are not available in the dominant episode	This is explained by coactivation of elements of several coalitions when none of them is really dominating (section 4.3.3). This is especially true when the more active coalition lacks important elements and thus leaves free valences to the competing episode.
schematization: intrusions from generic knowledge take place	The instantiation mechanism adds new elements to episodes by specializing generic facts and propositions (section 4.3.4). The instantiation mechanism is triggered and guided by the mapping.
context-sensitive representation of episodes and objects (effects on reminding, recognition, priming)	This is a direct consequence of the fact that context is represented by the whole state of activation over the memory elements and that the relevance of each element is estimated by its activation. Therefore the representations are always biased and influenced by the context.
context-sensitive representation of concepts	The same is true for the representation of concepts.
gradual recall and order of recall: episode elements may be recalled in different order	Episodes are represented in a distributed and decentralized way. They are recalled gradually as various elements pass the working memory threshold at different times.
priming effects on episodes	The priming effects are explained by residual activation from previously solved problems. The residual activation decays with time (section 4.1.2).
priming effects on generic knowledge, including facts and concepts	The same as above.

environmental context effects: perception of accidental elements from the environment may play a role in reminding and mapping	Perception activates certain memory elements which then take part in the computation. Thus even accidental elements, once activated by perception, participate in the process of reasoning and can influence it in various ways.

Finally, we are fully aware that all models are false, AMBR included. Some models are useful, however, and we hope AMBR might shed some light on the mysteries of analogy-making and on the role that dynamic context-sensitive emergent computations and representations may play in some of them. We also hope that the approach presented in this chapter will bring us one step further along the route toward seeing the elephant as a whole again.

Acknowledgements

This chapter reports the results of a project that started in 1988. It has been continuously supported by the Institute of Mathematics and Informatics at the Bulgarian Academy of Sciences (BAS) since its inception. During the last eight years it has also been supported by the Department of Cognitive Science and Psychology and the Central and East European Center for Cognitive Science at the New Bulgarian University (NBU). The project has also been supported by grants from the Bulgarian National Science Fund. This chapter was written during the stay of the first author at the Institute of Psychology at the Italian National Research Council (CNR) in Rome which was supported by a grant provided jointly by CNR and BAS. The second author worked on the paper while occupying a post-doctoral position at the Department of Psychology at Carnegie Mellon University, which was made possible by grant AFOSR F49620-99-10086 awarded to John R. Anderson.

We would like to thank first of all the AMBR research group at NBU for the long-lasting and fruitful collaborations and lively discussions. This group has involved the following people at various stages: Vassil Nikolov, Marina Yoveva, Kalina Hadjiilieva, Silvia Ivanova, Milena Leneva, Ivailo Milenkov, Radu Luchianov, Maurice Greenberg, Sonya Tancheva, Slavea Hristova, Iliana Haralanova, and Alexandrina Bahneva. We are also grateful to our colleagues and mentors Peter Barnev, Encho Gerganov, and Elena Andonova who supported the research over the years and from whom we have learned a lot. We would also like to thank Keith Holyoak, Dedre Gentner, Ken Forbus, Larry Barsalou, Brian Ross, Douglas Hofstadter, Pentti Kanerva, John Hummel, Robert French, John Anderson, Dario Salvucci, Cristiano Castelfranchi, and Maria Miceli for many insightful discussions and their comments on earlier drafts of this text which helped greatly to improve it. Finally, we would like to thank all participants of the Sofia Workshop on Analogy who provided cooperative and useful discussions of our work.

References

Anderson, S. & Conway, M. (1997). Representations of autobiographical memories. In M. Conway (Ed.), *Cognitive models of memory* (pp. 217-246). Hove, UK: Psychology Press.

Barclay, C. (1986). Schematization of autobiographical memory. In D. Rubin (Ed.), *Autobiographical memory* (pp. 82-99). Cambridge: Cambridge University Press.

Bargh, J. (1994). The four horsemen of automaticity: Awareness, intention, efficiency. and control in social cognition. In R. Wyer & Th. Srull (Eds.), *Handbook of social cognition: Vol. 1. Basic processes* (pp. 1-40). Hillsdale, NJ: Lawrence Erlbaum Associates.

Barsalou, L. (1982). Context-independent and context-dependent information in concepts. *Memory and Cognition, 10*, 82-93.

Barsalou, L. (1987). The instability of graded structure: Implications for the nature of concepts. In U. Neisser (Ed.), *Concept and conceptual development: Ecological and intellectual actors in categorization* (pp. 101-140). Cambridge: Cambridge University Press.

Barsalou, L. (1993). Flexibility, structure, and linguistic vagary in concepts: manifestations of a compositional system of perceptual symbols. In A. Collins, S. Gathercole, M. Conway, & P. Morris (Eds.), *Theories of memory* (pp. 29-101). Hillsdale, NJ: Lawrence Erlbaum Associates.

Barsalou, L. & Medin, D. (1986). Concepts: Fixed definitions or context-dependent representations? *Cahiers de Psychologie Cognitive, 6*, 187-202.

Bartlett, F. (1932). *Remembering*. Cambridge: Cambridge University Press.

Brewer, W. (1988). Memory for randomly sampled autobiographical events. In U. Neisser & E. Winograd (Eds.), *Remembering reconsidered: Ecological and traditional approaches to the study of memory* (pp. 21-90). Cambridge: Cambridge University Press.

Brown, R. & Kulik, J. (1977). Flashbulb memories. *Cognition, 5*, 73-99.

Clement, J. (1988). Observed methods for generating analogies in scientific problem solving. *Cognitive Science, 12*, 563-586.

Conway, M. (1995). *Flashbulb memories*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Conway, M. (1996). Failure of autobiographical remembering. In D. Herrmann, C. McEvoy, C. Hertzog, P. Hertel, & M. Johnson (Eds.), *Basic and applied memory research*, vol 1, *Theory in context* (pp. 295-315). Hillsdale, NJ: Lawrence Erlbaum Associates.

Craik, F. & Kirsner, K. (1974). The effect of speaker's voice on word recognition. *Journal of Experimental Psychology, 26*, 274-284

Crowder, R. (1976). *Principles of learning and memory*. Hillsdale, NJ: Lawrence Erlbaum Associates

Davies, G. (1988). Faces and places: Laboratory research on context and face recognition. In G. Davies & D. Thomson (Eds), *Memory in context: Context in memory* (pp. 35-53). Chichester, UK: John Wiley & Sons.

Davies, G. & Thomson, D. (1988). *Memory in context: Context in memory*. Chichester, UK: John Wiley & Sons.

Deese, J. (1959). On the prediction of occurrence of particular verbal intrusions in immediate recall. *Journal of Experimental Psychology, 58*, 17-22

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

French, R. (1995). *The subtlety of sameness: A theory and computer model of analogy-making*. Cambridge, MA: MIT Press.

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

Gentner, D. (1989). The mechanisms of analogical learning. In S. Vosniadou & A. Ortony (Eds.), *Similarity and analogical reasoning* (pp. 199-241) New York: Cambridge University Press.

Gentner, D. Rattermann, M., & Forbus, K. (1993). The roles of similarity in transfer: Separating retrievability from inferential soundness. *Cognitive Psychology, 25*, 524-575

Gick, M. & Holyoak, K. (1980). Analogical problem solving. *Cognitive Psychology, 12*, 306-355.

Godden, D.R. & Baddeley, A.D. (1975). Context-dependent memory in two natural environments: On land and under water. *British Journal of Psychology, 66*, 325-331.

Guenther, K. (1988). Mood and memory. In G. Davies & D. Thomson (Eds), *Memory in Context: Context in Memory* (pp 57-80) Chichester, UK: John Wiley & Sons.

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

Hofstadter, D. (1984). The Copycat project: An experiment in nondeterminism and creative analogies. AI Memo No. 775, Massachusetts Institute of Technology, Cambridge, MA.

Hofstadter, D., and the Fluid Analogies Research Group (1995). *Fluid concepts and creative analogies: Computer models of the fundamental mechanisms of thought*. New York: Basic Books.

Holyoak K. & Koh K. (1987). Surface and structural similarity in analogical transfer. *Memory and Cognition, 15* (4), 332-340.

Holyoak K. & Thagard P. (1989). Analogical mapping by constraint satisfaction. *Cognitive Science, 13*, 295-355.

Holyoak, K. & Thagard, P. (1995). *Mental leaps: Analogy in creative thought*. Cambridge, MA: MIT Press.

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

Hunt, E. (1999). What is a Theory of Thought? In R. Sternberg (Ed.), *The Nature of cognition* (pp. 3-49). Cambridge, MA: MIT Press.

Isen, A., Shalker, T., Clark, M., & Karp, L. (1978). Affect, accessibility of material in memory, and behavior: A cognitive loop? *Journal of Personality and Social Psychology, 36*, 1-12.

Kanerva, P. (1988). Sparse distributed memory. Cambridge, MA: MIT Press.

Keane, M., Ledgeway, K., & Duff, S. (1994). Constraints on analogical mapping: A comparison of three models. *Cognitive Science, 18*, 387-438.

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* (pp. 51-58). Amsterdam: Elsevier.

Kokinov, B. (1989). About modeling some aspects of human memory. In F. Klix, N. Streitz, Y. Waern, & H. Wandke (Eds.), *Man-computer interaction research MACINTER-II* (pp. 349-359). Amsterdam: Elsevier Science Publishers.

Kokinov, B. (1990). Associative memory-based reasoning: Some experimental results. *Proceedings of the Twelfth Annual Conference of the Cognitive Science Society* (pp. 741-749). Hillsdale, NJ: Lawrence Erlbaum Associates.

Kokinov, B. (1992). Inference evaluation in deductive, deductive, and analogical reasoning. *Proceedings of the Fourteenth Annual Conference of the Cognitive Science Society* (pp. 903-908). Hillsdale, NJ: Lawrence Erlbaum Associates.

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* (pp. 203-207). 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* (pp. 502-507). Hillsdale, NJ: Lawrence Erlbaum Associates.

Kokinov, B. (1994d). Flexibility versus efficiency: The DUAL answer. In Ph. Jorrand & V. Sgurev (Eds.), *Artificial intelligence: Methodology, systems, applications* (pp. 321-330). Singapore: World Scientific Publishers.

Kokinov, B. (1997). Micro-level hybridization in the cognitive architecture DUAL. In R. Sun & F. Alexander (Eds.), *Connectionist-symbolic integration: From unified to hybrid approaches* (pp. 197-208). Hillsdale, NJ: Lawrence Erlbaum Associates.

Kokinov, B. (1998). Analogy is like cognition: Dynamic, emergent, and context-sensitive. In K. Holyoak, D. Gentner, & B. Kokinov (Eds.), *Advances in analogy research: Integration of theory and data from the cognitive, computational, and neural sciences* (pp. 96-105). Sofia, Bulgaria: New Bulgarian University Press.

Kokinov, B., Hadjiilieva, K., & Yoveva, M. (1997). Explicit vs. implicit hint: Which one is more useful? In B. Kokinov (Ed.), *Perspectives on cognitive science: Vol.3* (pp. 176-183). Sofia, Bulgaria: New Bulgarian University Press.

Kokinov, B., Nikolov, V., & Petrov, A. (1996). Dynamics of emergent computation in DUAL. In A. Ramsay (Ed.), *Artificial intelligence: Methodology, systems, applications* (pp. 303-311). Amsterdam: IOS Press. [Available on-line: http://www.andrew.cmu.edu/~apetrov/publications/dyn_em_comp/abstract.html]

Kokinov, B., & Petrov, A. (2000). Dynamic extension of episode representation in analogy-making in AMBR. *Proceedings of the Twenty-Second Annual Conference of the Cognitive Science Society* (pp. 274-279). Hillsdale, NJ: Lawrence Erlbaum Associates. [Available on-line: http://www.andrew.cmu.edu/~apetrov/publications/rerepres/abstract.html]

Kokinov, B. & Yoveva, M. (1996). Context effects on problem solving. *Proceedings of the Eighteenth Annual Conference of the Cognitive Science Society* (pp. 586-590). Hillsdale, NJ: Lawrence Erlbaum Associates.

Kolodner, J. (1984). *Retrieval and organizational strategies in conceptual memory*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Kolers, P. & Ostry, D. (1974). Time course of loss of information regarding pattern analyzing operations. *Journal of Verbal Learning and Verbal Behavior, 13*, 599-612.

Koriat, A. & Goldsmith, M. (1996). Memory metaphors and the real-life/laboratory controversy: Correspondence versus storehouse conceptions of memory. *Behavioral and Brain Sciences, 19*, 167-228.

Loftus, E. (1977). Shifting human color memory. *Memory and Cognition, 5*, 696-699.

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

Loftus, E., Feldman, J., & Dashiell, R. (1995). The reality of illusory memories. In D. Schacter (ed.), *Memory distortions: How minds, brains, and societies reconstruct the past* (pp. 47-68). Cambridge, MA: Harvard University Press.

Loftus, E., Miller, D., Burns, H. (1978). Semantic integration of verbal information into a visual memory. *Journal of Experimental Psychology: Human Learning and Memory, 4*, 19-31.

Loftus, E. & Palmer, J. (1974). Reconstruction of automibile destruction: An example of interaction between language and memory. *Journal of Verbal Learning and Verbal Behavior, 13*, 585-589.

Marshall, J. (1999). *Metacat: A self-watching cognitive architecture for analogy-making and high-level perception*. Ph.D. dissertation, Indiana University at Bloomington.

Marshall, J. & Hofstadter, D. (1998). Making sense of analogies in Metacat. In K. Holyoak, D. Gentner, & B. Kokinov (Eds.), *Advances in analogy research: Integration of theory and data from the cognitive, computational, and neural sciences* (pp. 118-123). Sofia, Bulgaria: New Bulgarian 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* (pp. 69-90). Cambridge, MA: Harvard University Press.

McClelland, J. & Mozer, M. (1986). Perceptual interactions in two-word displays: Familiarity and similarity effects. *Journal of Experimental Psychology: Human Perception & Performance, 12* (1), 18-35.

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

Minsky, M. (1986). *The society of mind*. New York: Simon & Schuster.

Mitchell, M. (1993). *Analogy-making as perception: A computer model*. Cambridge, MA: MIT Press.

Moscovitch, M. (1995). Confabulation. In D. Schacter (Ed.), *Memory distortions: How minds, brains, and societies reconstruct the past* (pp. 226-251). Cambridge, MA: Harvard University Press.

Neisser, U. (1967). *Cognitive psychology*. New York: Appleton Century Crofts.

Neisser, U. (1981). John Dean's memory: A case study. *Cognition*, 9, 1-22.

Neisser, U. (1998). Stories, selves, and schemata: A review of ecological findings. In M. Conway, S. Gathercole, & C. Cornoldi (Eds.), *Theories of memory* (Vol. 2, pp. 171-186). Hove, UK: Psychology Press.

Neisser, U. & Harsch, N. (1992). Phantom flashbulbs: False recollections of hearing the news about *Challenger*. In E. Winograd & U. Neisser (Eds.), *Affect and accuracy in recall: Studies of "flashbulb memories"* (pp. 9-31). Cambridge: Cambridge University Press.

Neisser, U. & Jopling, D. (Eds.). (1997). *The conceptual self in context: Culture, experience, self-understanding*. New York: Cambridge University Press.

Nickerson, R. & Adams, M. (1979). Long-term memory for a common object. *Cognitive Psychology, 11*, 287-307

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

Petrov, A. (1998). *A dynamic emergent computational model of analogy-making based on decentralized representations*. Unpublished doctoral dissertation, New Bulgarian University, Sofia, Bulgaria.

Petrov, A. & Kokinov, B. (1998). Mapping and access in analogy-making: Independent or interactive? A simulation experiment with AMBR. In K. Holyoak, D. Gentner, & B. Kokinov (Eds.), *Advances in analogy research: Integration of theory and data from the cognitive, computational, and neural sciences* (pp. 124-134). Sofia, Bulgaria: New Bulgarian University Press. [Available on-line: http://www.andrew.cmu.edu/~apetrov/publications/ice_cube/abstract.html]

Petrov, A. & Kokinov, B. (1999). Processing symbols at variable speed in DUAL: Connectionist activation as power supply. In *Proceedings of the Sixteenth International Joint Conference on Artificial Intelligence* (vol 2., pp. 846-851). [Available on-line: http://www.andrew.cmu.edu/~apetrov/publications/energy/abstract.html]

Reinitz, M., Lammers, W., & Cochran, B. (1992). Memory-conjunction errors: miscombination of stored stimulus features can produce illusions of memory. *Memory & Cognition, 20*, 1-11.

Roediger, H. III (1980). Memory metaphors in cognitive psychology. *Memory & Cognition, 8*, 231-246.

Roediger, H. III & Srinivas, K. (1993). Specificity of operations in perceptual priming. In P. Graf & M. Masson (Eds.), *Implicit memory: New directions in cognition, development, and neuropsychology* (pp. 17-48). Hillsdale, NJ: Lawrence Erlbaum Associates.

Ross, B. (1989). Distinguishing types of superficial similarities: Different effects on the access and use of earlier problems. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15*, 456-468.

Ross, B. & Bradshaw, G. (1994). Encoding effects of remindings. *Memory and Cognition, 22*, 591-605.

Ross, B. & Sofka, M. (1986). Remindings: Noticing, remembering, and using specific knowledge of earlier problems. Unpublished manuscript.

Salaman, E. (1982). A collection of moments: A study of involuntary memories. In U. Neisser (Ed.), *Memory observed* (pp. 49-63). San Francisco: Freeman.

Schacter, D. (1995a). *Memory distortions: How minds, brains, and societies reconstruct the past*. Cambridge, MA: Harvard University Press.

Schacter, D. (1995b). Memory distortion: History and current state. In D. Schacter (Ed.), *Memory distortions: How minds, brains, and societies reconstruct the past* (pp. 1-43). Cambridge, MA: Harvard University Press.

Schunn, C. & Dunbar, K. (1996). Priming, analogy, and awareness in complex reasoning. *Memory and Cognition, 24*, 271-284.

Smith, S. (1988). Environmental context-dependent memory. In G. Davies & D. Thomson (Eds), *Memory in context: Context in memory* (pp. 13-34). Chichester, UK: John Wiley & Sons.

Spence, D. (1988). Passive remembering. In U. Neisser & E. Winograd (Eds.), *Remembering reconsidered: Ecological and traditional approaches to the study of memory* (pp. 311-325). Cambridge: Cambridge University Press.

Spencer, R. & Weisberg, R. (1986). Context-dependent effects on analogical transfer. *Memory and Cognition, 14*, 442-449.

Sulin, R. & Dooling, D. (1974). Intrusion of a thematic idea in retention of prose. *Journal of Experimental Psychology, 103*, 255-262.

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

Thomson, D., Robertson, S., & Vogt, R. (1982). Person recognition: The effect of context. *Human Learning, 1*, 137-154

Tulving, E. (1983). *Elements of episodic memory*. New York: Oxford University Press.

Tulving, E. & Schacter, D. (1990). Priming and human memory systems. *Science, 247*, 301-306.

Wharton, C., Holyoak, K., & Lange, T. (1996). Remote analogical reminding. *Memory and Cognition, 24* (5), 629-643.

Whittlesea, B. (1997). Production, evaluation, and preservation of experiences: Constructive processing in remembering and performance tasks. *The Psychology of Learning and Motivation, 37*, 211-264. New York: Academic Press.

Williams, D. & Hollan, J. (1981). The process of retrieval from very long-term memory. *Cognitive Science, 5*, 87-119.