Finding Analogies Within Systems: Constraints on Unsegmented Mapping

David F. Baldwin and **Robert L. Goldstone**
Department of Cognitive Science - Indiana University
{baldwind, rgoldsto}@indiana.edu

Introduction

The complex structure and organization of knowledge in the human mind is one of the key facets of thought. One of the fundamental cognitive processes that operates over that structure is analogy. A typical computational model of analogy might juxtapose a source domain and a target domain, such as the solar system and the Bohr-Rutherford (BR) model of an atom (Gentner, 1983). The goal is to find a correspondence between these two domains - a mapping between the elements in each that preserves relevant relational commonalities between aligned entities, such as 'sun' to 'nucleus'. This is essentially a problem of finding the structural commonalities between the source and target domains. By finding these commonalities, it is possible to generate inferences and generalize knowledge to novel situations.

Several theories have been able to account for analogies such as the solar system model. Some of these include Gentner's SME (1983), Hummel and Holyoak's LISA (2003) and Kokinov's AMBR (2001). While these models differ in many ways, they share a common fundamental premise. Given two conceptual networks - typically represented by a series of symbolic entities (e.g., 'boy' and 'toy truck') and propositions (e.g., 'Likes(x,y)') - these models are built to find the maximal relational overlap between these two networks. Determining optimal correspondence is a computationally difficult problem and the equivalent graph-theoretic problem - subgraph isomorphism - is NP-hard.

Determining a mapping between the source and target domains of a non-trivial size would be intractable without a set of constraints to restrict the set of correspondences that are considered by a human reasoner. The most obvious constraint comes from the relational structure - matching similar relations to each other. Almost all theories also employ additional constraints such as systematicity (Gentner,1983) - a preference for finding relationally consistent substructures - and parallel connectivity - arguments of corresponding predicates are also placed into correspondence.

Less explicitly, the postulation of a source and target domain also puts substantial constraints on correspondence. Two distinct representations are given and any correspondence must be between these two domains (i.e., Entity x in the source domain cannot correspond to Entity y in source domain). Two elements in the same domain are related to one another through propositions, but are not placed in correspondence with one another. Most models of analogy consider placing "Sun" from the solar system domain into correspondence with either "nucleus" or "electron" from the atom domain, but do not consider placing "Sun" into correspondence with "Earth" because they come from the same domain. It is thus crucial in models of analogy to distinguish between two kinds of links between entities. Relational links occur between elements of the same domain, and correspondence links occur between elements of different domains. In the same way that authorship links connect papers to authors, but not authors to authors, or papers to papers, correspondence links connect only entities in one domain to those in another domain.

Another constraint imposed by virtually all analogical reasoning models is a bias against (Holyoak & Thagard, 1989; Hummel & Holyoak, 2004) or strict exclusion of (Falkenhainer et. al., 1989) two-to-one mappings between two domains. This constraint prevents or discourages Entity x in one domain from mapping to both Entities y and z in the other domain. The presence of this bipartite constraint on correspondence links, however, makes the implicit assumption that the elements of the two domains have been *a priori* identified and grouped. However, reasoners are rarely given the luxury of being told when and if corresponding structures exist. The average reasoner has an immense quantity of interconnected semantic knowledge that might be appropriate for a given task and determining the information that is to be part of a source and target domain can greatly influence the resulting correspondence. Revisiting the BR analogy, unless we are explicitly told to create an analogy between the solar system and atom, how are we to, for example, ignore the fact that the solar system is composed of innumerable atoms? That is, relational links connect the nucleus to the sun.

This problem poses a serious issue for analogical reasoning and other processes in high-level cognition. If any chunk of semantic knowledge can, in principle, be related to any other, how does a reasoner locate the task-relevant information. Clearly, finding correspondences becomes intractable if we cannot select a subset of this knowledge. Namely, unless we can postulate a mechanism by which the information in source and target domains is selected, any story of how correspondences are determined is incomplete.

Several models of analogical reasoning have gone beyond a single source and target domain to include analogy as part of a larger cognitive architecture. Most of these generalizations focus on the integration between retrieval of a source domain from long term, semantic memory and correspondence mappings. MAC/FAC (Forbus, 1995) uses an efficient similarity calculation to narrow the space of possible analogs chosen from memory before performing the full analogical correspondence using SME. Kokinov's AMBR (Kokinov and Petrov, 2001) has the additional ability to create 'blends' between representations in the multiple source domains and a single target domain. These sorts of approaches, however, still assume the prior specification of segmented domains. In these models, target instances stored in long term memory have no cross-domain predicate structure nor can stored analogs be composed arbitrarily. That is, they maintain the assumption that knowledge comes packaged into cleanly delimited domains.

If the presence of prespecified domains and restrictions such as one-to-one correspondence serve as constraints on analogical mapping, it is important to understand how these constraints might be realized in a single network of interconnected knowledge that comprises a reasoner's conceptual system. One possible way to study this is to examine how humans might transform an unsegmented predicate structure into segmented domains that serve as input to an analogy process such as SME.

This, however, may be the incorrect approach. If a reasoner has a general schema for 'Orbiting', it may map to many different subsets of their semantic knowledge (e.g., planets in the BR analogy, a tetherball game, etc.). Domains that are segmented from each other should have a one-to-one mapping constraint, but segmentation may produce more than two domains. A single entity in Domain D1 (e.g., a satellite rotating around the earth) might correspond to entities in two disjoint domains D2 and D3 (e.g. the mechanics of an atom and the motion of planets), Therefore, in many cases, a many-to-many correspondence may be needed. These systematic many-to-many mappings are generalization of the constraint found in traditional analogy engines.

It may also be the case that correspondence influences segmentation during the mapping process. In other words, if we are given an explicit segmentation, discovering a mapping - even a partial one - between a base and a target domain may cause us to reevaluate how the entire system that contains the base and target is segmented. For example, placing the 'Sun' in correspondence with 'Nucleus' might increase one's propensity to attempt to include 'Pluto' as part of the solar system domain.

An Alternate Approach : Unsegmented Mapping

Rather than focus on the processes that might give rise to disjoint domains, we present a set of constraints that produces analogical-like correspondences without explicitly segmenting knowledge structures - a process we will call *Unsegmented Mapping*. The advantage of this approach is that we are able to avoid the assumption that knowledge is partitioned into disjoint subsets. Instead, we specify a set of constraints that allow us assume that all of a person's knowledge is interconnected and still find analogies within this conceptual network. Our model, COWARD (COrrespondence Without A Priori Representation of Domains) is implemented as the interaction between three different constraints - relational consistency, semantic connectivity and transitivity.

The first constraint, relational consistency, is a crucial aspect of any analogy model. Thus, the discovery of common relational structure is fundamental regardless of whether explicit domains are used to guide the process. However, it is necessary to elaborate on the remaining two constraints in detail.

The second constraint - semantic connectivity - is necessary to have a measure by which entities compete for correspondence. Traditional models have a very constrained criteria - all pairs of entities in the one domain compete for correspondence with entities in the other domain. In order to find structural consistencies in unsegmented representations, it is necessary to generalize this constraint. One way of doing so is to assume that each pair of entities is assigned a weight indicating a measure of 'coherence' - the degree to which the pair of entities belong to the same domain or system. Current models assign maximal weight to pairs of entities between the two domains, while assigning zero weight to pairs of entities within the same domain. Our generalization of this is to establish a graded notion of coherence such that any two entities in a system may have several paths connecting them, but the more numerous and the shorter the paths are, the more coherent the pair of entities is.

It would be possible to assign two sets of weights on relational structure. The first, used to align the graphs, would come from the predicate structure. The second would be a weight indicative of the coherence of entities. However, we treat these weights identically. That is, coherence is a function of the semantic weights represented by a weight vector assigned to the edges in a graph. There is substantial evidence for semantic coherence in studies of category learning and causality. Rehder and Ross (2001) showed evidence for 'abstract coherent categories' - categories defined only by their relational structure. Ahn (1998) found strong correlations between the causal status of category features and similarity ratings.

If high relational connectivity is indicative of the presence of a domain-like structure, then to be consistent with a model such as SME, a model must do two things. Two entities that are relationally well connected should be inhibited from corresponding to each other. Less obvious, however, is that these two entities should also in-

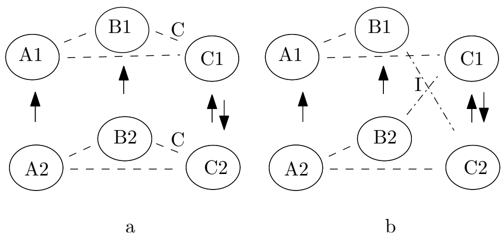

Figure 1: A simple example of correspondence without domains. Solid lines with arrows represent a directed relation between vertices. Dashed lines represent a bidirectional correspondence mapping (x corresponds to y). a) A consistent, transitively closed correspondence mapping b) A correspondence mapping without transitive closure. The two crossed correspondence links in b) labeled I lead to a correspondence in which every entity corresponds to every other entity via transitivity. Alternatively, the edges labeled C lead to a transitively closed system of two types of vertices (1's and 2's) in a)
.

hibit each others' correspondence to other entities. Figure 1a shows this graphically. For example, the correspondence between A1 and B1 should inhibit the correspondence between A1 and B2.

It is in this fashion that we might implement a generalized and liberalized one-to-one constraint given a single unsegmented representation. That is, the relative strength of relations between entities serves to constrain which entities compete for correspondence. Multiple alignable systems might be discovered within a single knowledge structure. If this is the case, it is then desirable to find many-to-many correspondences.

The third constraint involves the influence of transitivity on correspondence. COWARD scales to allow for correspondence between three or more domains. In the case that a reasoner finds more than two systems, it is possible to create relationally consistent correspondences between pairs of systems while still maintaining a higher-order inconsistency. This can occur when there are symmetries in the representation that allow for multiple mappings. This inconsistency is essentially a problem of transitivity. If the network in Figure 1b converged on the correspondence as shown, then every entity could correspond to every other entity if we consider correspondence to be transitive. We consider this to be the incorrect solution - an analogical mapping can be seen as classifying each entity as a role in each system (e.g., agent of Loves(x,y)). These inconsistencies would lead a reasoner to see the agent and the patient in our simple example as indistinguishable. The correspondence labeled C in Figure 1a are both consistent and transitively closed. That is, the set of '1s' (A1, B1, C1) and only that set, will be placed in correspondence with each other and that no further transitive inferences are possible.

Theoretically, this constraint is closely related to schema induction in analogical transfer (Gick and Holyoak, 1983) and pressures the model to find the consistent commonalities in multiple systems. Our model, however, uses different principles. Namely, schemas arise out of a set of multiple correspondences. The model is able to establish correspondences across three or more systems such that if A in System 1 corresponds to B in System 2, and B corresponds to C in System 3, then A should also be also placed in correspondence with C. According to virtually all existing models of analogical reasoning, placing A into correspondence with both B and C would violate a 1-to-1 mapping constraint, but COWARD views these correspondences as consistent because they cohere together without contradiction as long as A, B, and C do not come from the same system. Additionally, unlike other models, COWARD does not require full analogical mappings to create correspondences. The influence of transitivity in the model occurs well before correspondences converges, so a partial mapping is sufficient to initiate a correspondence between S1 and S3. Recent empirical evidence has shown evidence that a full correspondence mapping is not necessary for transfer (Ripolli, Brude, and Coulon, 2003).

The implementation of COWARD most resembles an iterative constraint satisfaction network (Goldstone and Rogosky, 2002; Larkey and Love, 2002). These three constraints interact to form correspondence mappings and it is in this fashion that commonalities between multiple systems might be found. Thus, it is possible to sufficiently constrain the problem of finding correspondence mappings without representation of domains or a typical one-to-one correspondence restriction. In the next section, we show empirical evidence for these constraints.

Empirical Evidence for COWARD

In this section, we describe two experiments demonstrating the effects of semantic connectivity and transitive closure. Our general approach is to train participants with a series of relational facts ("X is Y's brother") with a particular structure and then ask them to make a correspondence mapping across all entities.

In all experiments, we discuss pairwise correspondence. We will denote the asymmetric correspondence between X and Y by $X \rightarrow Y$, where X has been placed into correspondence with Y, but not necessarily vice versa. When correspondence is symmetric, it will be denoted $X \leftrightarrow Y$. Here, we use a stimuli based on social roles, as suggested by Rehder and Ross (2001).

Experiment 1

In Experiment 1, we created a simple example of how correspondence mappings might be affected by connectivity. Rather than create an explicit and complex relational structure, we used the simplest possible setup - three entities that should correspond to each other based on instructed roles. Participants were told that each entity was a President of a country and then asked to make correspondence mappings. Additionally, some subjects were told a relational fact connecting two of the presidents. In this experiment, we aimed to show that adding a relation between Presidents B and C decreases correspondence from A to at least one of B or C, and also decreases correspondence between B and C.

Participants

100 Indiana University undergraduates participated for partial course credit.

Materials and Procedure

The experiment consisted of two conditions. In the *added-relation* condition, an additional directed relation between B and C was presented. In the *no-relation condition*, no extra relation was present.

Participants received a two page packet. The first page consisted of the following instructions "You will be asked to answer a series of questions about presidents from different countries. Below are several facts. Please read them and then answer the questions on the next page." The facts were as follows "The President of the Ukraine is Leonid Kuchma," "The President of Peru is Alan Perez," and "The President of Chad is Idriss Deby." Participants in the edge condition received the following additional relation "The President of Peru put trade embargoes on Chad based on the president of Chad's failure to comply with his wishes."

Correspondence questions were prefaced with a short instructional paragraph consisting of the following instructions "Please answer the following questions about the roles of each of these presidents. The questions are in the form 'Which of the following people play the same role as X?' For example, if 'Mary loves John,' 'Fred loves Jill' and 'Larry loves Brenda,' one might say that Mary, Fred and Larry play the same role of the 'lover." Participants were asked this question for each president. Below each question were four check boxes, of which the subject was to select one. The checkboxes consisted of all possible correspondence conjunctions between the president in question. For example, if asked "Which of the following people play the same role as A?" the checkboxes consisted of 'None,' 'B,' 'C,' and 'B *and* C.'

Results

Across all correspondences, the mean proportion of mappings in the added-relation (M=.521) and no-relation (M=0.831) conditions were significantly different ($p < 0.0001$). All pairs of correspondences averaged across symmetries were also significantly different ($A \leftrightarrow B$, no-relation M: = .85,added-relation M = .57, t(98) = 3.64, $A \leftrightarrow C$, no-relation M: = .85, edge M = .52, t(98) = 4.04, $B \leftrightarrow C$, no-relation M: = .79, added-relation M = .49, t(98) = 3.4, all $p < 10^{-3}$).

Of the 50 people in the no-relation condition, 35 produced the mapping of all correspondences possible, 3 produced no correspondences and the remaining 18 produced other mappings. Of the 50 people in the added-relation condition, 18 produced all correspondences and 15 produced no correspondences. The remaining 17 produced other mappings. The distribution of full/no-mappings was significant (χ^2 (1,N = 71) = 13.166, $p < 10^{-3}$).

Of the participants that produced other mappings, 3 of the 18 participants in the no-relation condition produced only a single, exclusive correspondence between $A \rightarrow B$ or $A \rightarrow C$. Eleven of the 17 participants in the added-relation condition produced an exclusive correspondence (9 of which were $A \rightarrow B$). This difference is significant by Fisher Exact Test ($p < 0.01$). Similarly, the difference between mapping $A \leftrightarrow B$ was not significantly different between conditions by Fisher Exact Test ($p > .5$). However, the mapping $A \leftrightarrow C$ was significantly different between the two conditions ($p < 0.05$).

Discussion

The presence of an additional relation between B and C clearly influences the correspondences participants are willing to make between two entities in the above experiment. In the no-relation condition, participants overwhelmingly chose to make each president correspond to all other presidents. The presence of an additional relation between B and C, however, had two effects. First, it caused participants to cease making correspondences between B and C. It also influenced participants propensity to make correspondences between A and either B or C. Specifically, participants were more likely (relative to the no-relation condition) to make either no correspondences or a single correspondence between B or C. This shows that connectivity can decrease correspondence between related items. When B and C have some relation, not only do B and C's correspondence decline, so does that of A to B and A to C. That is, B and C are treated as a connected system to be matched with A.

Experiment 2

We have shown in the previous experiment the effects of connectivity on correspondence mappings. In this experiment we explore the conditions under which a transitive closure-like property might be used by subjects. More specifically, we wished to explore whether subjects make transitively inconsistent correspondences, and if they were willing to add a transitive relation when implied by the other correspondences (if $A \leftrightarrow B$, $B \leftrightarrow C$, then $A \leftrightarrow C$).

Participants

90 Indiana University undergraduates participated for partial course credit.

Methods

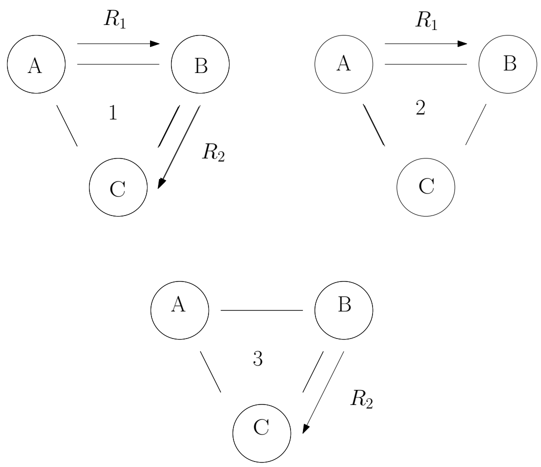

Figure 2: Abstract structure of Experiment 2

	Coherent Stimuli	Incoherent Stimuli
R_1	A1 scolds A2	A1 is jealous of A2's wealth
R_2	A2 picks on A3	A2 hates A3

Table 1: Experiment 2 Stimuli

The instructions preceding the correspondence questions (and the president example) were identical to the experiment above. Questions took the form "Which brothers play the same role as X" for every actor X in the system. Next to each question, participants were instructed to write as many characters as appropriate (including none).

Participants were randomly assigned to one of three conditions. In conditions CS (coherent stimuli) and IS (incoherent stimuli), participants received the coherent and incoherent stimuli, respectively (Table 1). In CS+I (coherent stimuli, plus instruction), participants received an additional instructional manipulation before the correspondence instructions intended to emphasize the brothers as a system. "Friends of these families have noted that the families are quite similar - each has three brothers, all of which are boys and they may have other commonalities as well." This additional instruction was intended to emphasize both the within family relations as well as the between family correspondences.

Results

Results were coded by the number of transitive mappings participants made between the actors in the system. Thus, each subject received a score between 0 and 3 indicating whether they had made a transitive mapping of AX, BX and CX for each X. Pairwise t-tests showed no difference between the CS+I (M=1.36) and CS (M=0.9667) conditions(t(58) = 1.0685, $p > .2$), but both were different from the IS (M=.0.2) condition (CS+I vs IS: t(58) = 3.8, $p < 10^{-3}$, CS vs IS: t(58) = 2.6, $p < .02$). The analysis included transitive mappings that were inconsistent with the relational structure, but instances of this were extremely rare. Only a single instance of inconsistent transitive mapping was found in the CS+I condition.

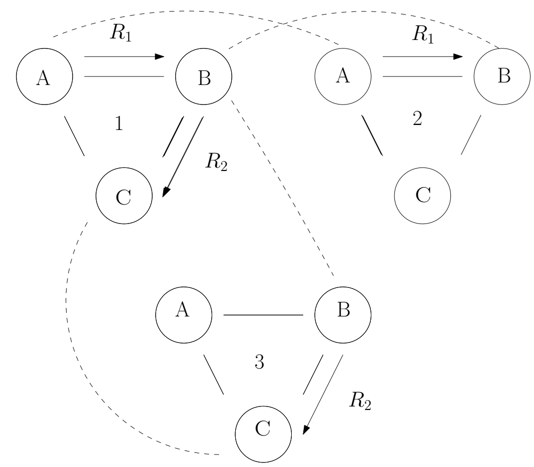

Figure 3: Locally consistent mappings seen in IS condition.

The vast majority of participants either produced the full transitive mapping (score of 3) or few (score of 0 or 1). 13/30 and 9/30 produced the full transitive correspondence exclusively in the CS+I and CS conditions, while only 2/30 produced this in the IS condition. Participants who did not produce any transitive correspondence tended to produce mappings that were locally consistent - $A1 \leftrightarrow A2$, $B1 \leftrightarrow B3$, $C1 \leftrightarrow C3$ and $B1 \leftrightarrow B2$ (see figure 3). In the IS, CS and CS+I conditions, 20/30, 8/30 and 9/30 participants produced this mapping. The remaining 8/30, 13/30 and 8/30 (IS, CS and CS+I) conditions were not classifiable in this manner.

The distribution of participants producing the full transitive mapping to those that did not was not significant between CS and CS+I by Fisher Exact Test ($p > 0.3$) but CS vs. IS as well as CS+I vs. IS were significant ($p < 0.02$, $p < 0.002$, respectively).

Discussion

First, it is clear that under certain conditions, participants are willing to produce many-to-many correspondences. However, most participants appear to do so in a systematic way. Rather than simply choosing all possible relationally consistent matches (e.g., all brothers corresponding to each other), participants tend to make correspondences based on more global relational consistencies. The degree of global consistency appears to be a function of the coherence of each domain. The qualitative pattern of matching is as follows. In the IS condition, participants made consistent mappings only when the local (relative to a single entity) relational structure was consistent (the local edge structure was identical, see Figure 3).

In both the CS and CS+I conditions, participants were more likely to make a more global, transitively closed set of mappings (connecting all A's, B's and C's) . It appears that participants are only willing to make the full mapping when the inference seems plausible. Our model produces the same pattern of matching by manipulating the strength of the brother relations. This may seem surprising given the non-diagnosticity of the brother relation within a family, but it is actually precisely the effect we might expect given the utility of the brother relation in demarcating families. Namely, when emphasized in CS or CS+I, the brother relations provide a degree of coherence between the members of each family, thereby increasing the relational commonality between the three groups. When the relations are less coherent as in the IS condition this leads to less transfer. According to our model, then, coherence highlights relational commonalities and allows for inconsistent or missing relations to be inferred. Surprisingly, only one participant in both CS and CS+I conditions produced correspondences that were transitively inconsistent, indicating that participants tend to avoid the higher order inconsistencies that such a mapping produces.

Conclusion

In sum, the stipulation of pre-specified domains has been a key feature of most analogical reasoning systems. However, few studies have actually defined how domains are realized in cognitive agents, and in fact there is good reason to think that knowledge is isotropic (Fodor,1983) - anything we believe could be connected to anything else we believe. As such, there is a strong motivation for developing algorithms that find similarities among pieces of knowledge even though all of these pieces of knowledge may be connected together in a connected web. Analogical reasoning systems to date assume that domains to be compared are easily separable and delineable in order to generalize knowledge by aligning representations. Our claim is that domains are not necessary for correspondence. We have presented a model, COWARD, which finds mappings without segmentation and we have provided initial empirical evidence for the constraints used to do so.

However, the idea of the domain still holds great importance for analogical reasoning. Domains allow for rapid retrieval of associated knowledge and support processes such as categorization and similarity judgments. In our account a domain might be a well connected cluster of relations, sensitive to changes in the weights of relational structure. Rather than discarding the idea of domain, our goal in is to show that the current use of a domain structure is not only too strict, but in many cases, not necessary to form an analogical mapping within the representation of a single system. In doing so, we have dissociated correspondence from domain structures. Future examinations might focus on the importance of the domain construct beyond that of a simple grouping mechanism. In fact, it may turn out that the act of finding correspondences is instrumental in creating separable domains, a reversal of the standard assumption that the delineation of domains is the critical first step on the way to finding correspondences.

Acknowledgements

Support for this work was provided by an NSF Graduate Fellowship and by Department of Education, Institute of Education Sciences grant R305H050116 and NSF REC grant 0527920.

References

Ahn, W. (1998). Why are different features central for natural kinds and artifacts? The role of causal status in determining feature centrality. *Cognition*, *69*(2), 135–178.

Falkenhainer, B., Forbus, K., & Gentner, D. (1989). The Structure-Mapping Engine: Algorithm and Examples. *Artificial Intelligence*, *41*(1), 1–63.

Forbus, K., Gentner, D., & Law, K. (1995). MAC/FAC: A Model of Similarity-Based Retrieval. *Cognitive Science*, *19*(2), 141–205.

Gentner, D. (1983, April–June). Structure-mapping: A theoretical framework for analogy. *Cognitive Science*, *7*(2), 155–170.

Gentner, D., & Kurtz, K. (2005). Relational categories. *Categorization Inside and Outside the Laboratory*, 151–175.

Gick, M., & Holyoak, K. (1983). Schema induction and analogical transfer. *Cognitive Psychology*, *15*(1), 1–38.

Grinberg, M., & Kokinov, B. (2003). *Analogy-Based Episode Blending in AMBR.* Constructive Memory. Sofia: NBU Press.

Holyoak, K., & Thagard, P. (1989). Analogical Mapping by Constraint Satisfaction. *Cognitive Science*, *13*(3), 295–355.

Hummel, J., & Holyoak, K. (2003). A symbolic-connectionist theory of relational inference and generalization. *Psychological Review*, *110*(2), 220–264.

Kokinov, B., & Petrov, A. (2001). Integration of memory and reasoning in analogy making: The AMBR model. *The analogical mind.*

Rehder, B., & Ross, B. (2001). Abstract coherent categories. *J Exp Psychol Learn Mem Cogn*, *27*(5), 1261–75.

Ripoll, T., Brude, T., & Coulon, D. (2003). Does analogical transfer involve a term-to-term alignment? *Memory & Cognition*, *31*(2), 221–230.

Ross, B. (1990). Reminding-Based Category Learning. *Cognitive Psychology*, *22*(4), p460–92.