Similarity by feature creation: Reexamination of the asymmetry of similarity

Hitoshi Ohnishi
Dept. of Systems Science
Tokyo Institute of Technology
Midori-ku, Yokohama 227, Japan
ohnishi@sys.titech.ac.jp

Hiroaki Suzuki
Dept. of Education
Aoyama Gakuin University
Shibuya-ku, Tokyo 150, Japan
susan@ri.aoyama.ac.jp

Kazuo Shigemasu
Dept. of Systems Science
Tokyo Institute of Technology
Midori-ku, Yokohama 227, Japan
kshige@sys.titech.ac.jp

Abstract

We developed a computational model of similarity judgment in problem-solving contexts. The model first attempts to transform an object to another using the knowledge of the domain, the strategy, and the goal. If the transformation succeeds, new feature about transformability is created. A similarity of an object to another is computed, based on the created features. If the model fails to create a new feature, it computes a similarity by feature comparison in the same way as the contrast model. An important prediction of the model is that the asymmetry of similarity judgments is caused by the directionality of the problem-solving skills. We examined the model's prediction. The material was the Tower of Hanoi puzzle. Subjects were required to rate the similarities of one state to the goal as well as those of the goal to a state. In Experiment 1, we taught one group of subjects the 'move-pattern strategy' that induced learners to acquire highly directional skills, and compared their judgments with those by naive subjects. The asymmetry was observed only in the judgments by the trained subjects. The second experiment showed that the results of the experiment 1 could not be attributed to the 'prototypicality' of the goal.

Introduction

People have an ability to deal with situations to which they are not familiar. This ability is mainly based on the analogical use of past experiences. That is, people retrieve a similar past experience and make use of, or adapt, its solution to the current situation. Various models for analogy have been proposed (Falkenhainer, Gentner, & Forbus, 1989; Hammond, 1990; Holyoak et al. 1989).

There is one important problem concerning analogy. As mentioned above, people have to retrieve a similar experience to the current situation in analogy. However, in what sense is a retrieved experience *similar* to the current situation? More generally, how is similarity defined in problem-solving and learning contexts? Feature-based models of similarity such as the contrast model (Tversky, 1977) do not seem to be quite successful for this problem. It is often the case that two objects which are superficially dissimilar to each other are sometimes judged to be similar in a problem-solving context. This is because features important in a problem-solving context are not necessarily salient in the judgment of object-level similarity. For example, a feature such as non-Japanese is far from important if you are an American and live in the United States. However, this feature becomes crucial if you move to Japan and need public services such as the National Health Service.

This suggests that two types of similarities should be distinguished — deep (goal-related) and shallow (superficial). Gentner and Forbus (1991) proposed a model, called MAC/FAC, that computes similarities of both deep and shallow levels. The MAC/FAC model consists of two stages. While in the MAC stage computationally cheap matchers act on content vectors of items in LTM, structural examinations are made in order to compute deep similarity in the FAC stage.

Although the MAC/FAC model captures important aspects of human similarity judgment, further steps should be taken to model judgments of similarity. It is well-known that goal and knowledge of the domain play crucial roles in problem-solving and learning. However, these are not taken into account by their model.

Suzuki, Ohnishi, & Shigemasu (1992) found that people's judgments of similarity are greatly affected by their recognition of the task goal and by knowledge of the domain, using the Tower of Hanoi Puzzle. When subjects did not know the puzzle and were asked to judge the similarity between a state and the goal state, their judgments were based on superficial features shared with both states. In contrast, experts' judgments were dependent on the distance between a state and the goal. For example, while similarity between the states 4 and the goal in Figure 1 was rated very high by naive subjects, experts' ratings were very low. When given a stimulus set shown in Figure 1(a) and 1(d), patterns of rated similarity were reversed between naives and experts. What happened if similarity of the states is judged by non-experts who only know the rules of the puzzle? They relied on the distance in some cases, and the number of shared features in other cases. When a given state was easily transformed to the goal, their judgments were based on the distance. However, they relied on shared features when it was difficult for them to transform a given state to the goal.

These results showed that we should incorporate the goal recognition mechanism and domain knowledge into the model of similarity judgment. Furthermore, these suggest a challenging problem to theories of similarity. That is, not every feature exists prior to the judgment of similarity. Rather some features are created by the

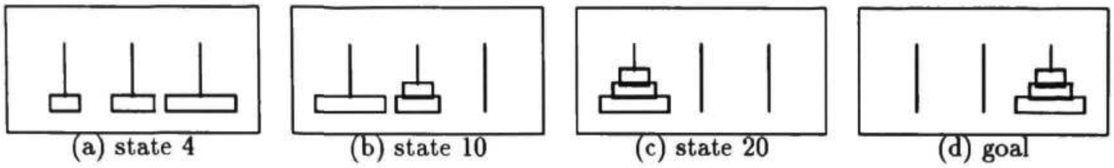

Figure 1: States of the Tower of Hanoi

recognition of the goal and the knowledge of the domain. Although most of the previous theories have assumed that all the features necessary for the computation of similarity be provided externally, [1] the assumption is highly dubious especially if one considers similarity judgments in problem-solving contexts. For example, it is not reasonable to suppose that there are features such as small disk's being movable to the peg C, medium disk's not being movable to the peg B, etc. Rather, these features are created or inferred in the judgment process. Murphy & Medin (1985) addressed the same issue in the context of categorization. They gave a good example of the feature creation. When we observe that someone jumps into a swimming pool with his clothes on, we naturally judge that he is drunk. Since the categorization is based on the matching between observed features and concept's attributes, the concept "intoxicated" should have the attribute such as "jumping into a swimming pool with one's clothes on." However, it is very unlikely that such an attribute is associated with the concept "intoxicated" in advance of the judgment. Thus, it is reasonable to assume that the attribute is created internally in the course of categorization, and then matched to the observed feature. Studies of expert-novice differences also give evidence of feature creation. Chi, Feltovitch, & Glaser (1981) showed that novices' sortings of physics problems were based on superficial similarity between the problems, while experts' ratings were dependent on physics principles. These principle-related features seem to be created by experts' knowledge. Novices do not have the necessary knowledge to create these features. That is why their sortings were based on superficial similarities between the problems.

Therefore, a model should be developed that creates features by using the goal and the domain knowledge. We have already outlined the model in the previous study (Suzuki, Ohnishi, & Shigemasu, 1992). In this study, we introduce a computational model of Similarity by Feature Creation (*SFC*), and examine the model's psychological validity in terms of the asymmetry of similarity.

SFC: The Two-Stage Model of Similarity Judgment

Figure 2 depicts the processes of *SFC*. *SFC* consists of two stages: in the first stage, the new feature is created, and in the second stage similarity is computed. The input is a physical description of a pair to be compared (A and B). In the first stage, *SFC* assumes that a query for the similarity of A to B is that for the transformability of A to B. In other words, B is treated as the goal, while A is treated as another state in the problem space. *SFC* first tries to detect differences between A and B. If the differences are detected, the model tests whether A can be transformed to B using the domain knowledge. The knowledge of the domain consists of the operators, task constraints, and strategies. The strategy consists of a list of subgoals and their dependency. If it is difficult to directly transform A to B, another subgoal is retrieved from the subgoal list. This procedure is repeated recursively until the test succeeds. If the test succeeds, a new feature 'transformable(Goal, Distance)' is created, where 'Goal' is B or the subgoal, and 'Distance' is the number of operators required to transform A to 'Goal.' If the goal is not recognized, or a relevant subgoal is not retrieved, the test fails, and any new feature is not created.

In the second stage, the model computes a similarity between a state A and the goal B based on a newly created feature or superficial features. If a new feature created, a deep similarity between A and the B is computed in the *d-sim* process, based on the newly created feature and the goal dependency.

If 'Goal' is in the higher branch of the goal dependency tree, the similarity between A and B are rated high. The similarity is rated low, if 'Distance' is greater.

If a new feature is not created, *SFC* shifts to computing a shallow similarity in the *s-sim* process. In this process, similarity is computed in the same way as Tversky's contrast model, using the physical description of the objects.

We take the Tower of Hanoi as an example to illustrate the behavior of *SFC*. Suppose that the inputs are Figure 1(a) (we represent it as [[1],[2],[3]]) and Figure 1(d) ([[],[],[1,2,3]]), and that the subgoal list consists of [[[],[],[1,2,3]], [[-],[-],[2,3]], [[-],[-],[3]]] where '-' represents 'wild card', and the order of elements represents the goal/subgoal hierarchy. Through the feature creation phase, a feature 'transformable([[-],[-],[2,3]], 1)' is created, because the state already achieves the subgoal [[-],[-],[3]], and can be transformed to [[-],[-],[2,3]] in one step. Now suppose Figure 1(b) ([[3],[1,2],[]]) is given instead of Figure 1(a). Through the feature creation stage, a feature 'transformable([[-],[-],[3]], 1)' is created. In the d-sim process, Figure 1(a) is judged to be more similar to Figure 1(d) than the state Figure 1(b) is, because [[-],[-],[2,3]] is located in a higher branch of the goal dependency tree than [[-],[-],[3]] is. If feature creation fails, similarity is calculated by comparison of superficial features in the s-sim process.

[1] Few exceptions are models based on the case-based Reasoning, such as Kolodner (1989) and Leak (1991).

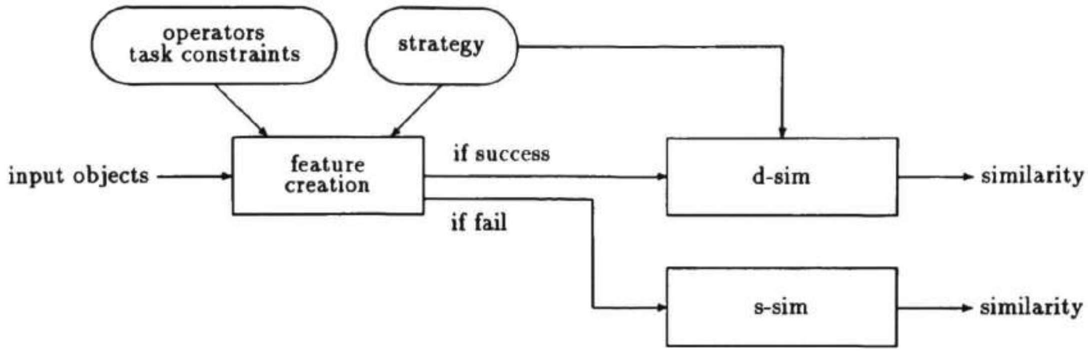

Figure 2: *SFC* model

Asymmetry in Similarity Judgments

There is an important issue with which *SFC* has not yet dealt. Tversky (1977) showed that similarity is asymmetric. For example, the judged similarity of North Korea to Red China is greater than that of Red China to North Korea. His model, the contrast model, was proposed to explain the asymmetry of similarity. In his model, similarity between object A and B is given by

$$s(A,B) = \theta f(A \cap B) - \alpha f(A - B) - \beta f(B - A)$$

where f is measure a function of the feature salience, and θ, α and β are weighting factors. According to the model, the asymmetry of similarity is explained by differential salience and differential weighting of distinctive features. However, his model cannot easily be extended to the similarity judgment in the problem-solving context, since a number of important features are created and inferred in the processes, rather than given in advance. Moreover, which features are created is dependent on problem-solver's knowledge.

Although we did not suggest any cognitive mechanism for the asymmetry of similarity in the previous study, the asymmetry may be observed in the problem-solving context. For example, translating a foreign language to one's mother language is sometimes very different from translating the latter to the former. It is due to the directionality of skills. At the early stages of learning, learners are usually forced to acquire skills to achieve a specific goal. In consequence, they may well acquire a set of specialized skills that transform a certain set of fixed initial states to the goal. It is not likely that learners are required to learn reverse operators. In this sense, problem-solving skills are directional.

If the skills are directional, a judged similarity should sometimes be asymmetric. *SFC* explains this in terms of the subgoal and the domain knowledge. Suppose that an expert has been trained in a highly routinized way, where only the transformation of the fixed initial state to the fixed goal is required. His acquired knowledge may well be specialized only to transform the initial state to the goal. If he is asked to rate the similarity of the initial state to the goal, his judgment should be a function of distance. Suppose that he is asked to rate the similarity of the goal to the initial state. In this case, he recognizes the initial state as the goal and the goal as the initial state. He can also detect the difference between the two states, and tries to transform the goal to the initial state. However, he may not be able to transform the goal to the initial state, because the strategy that used to be relevant is no longer relevant in this case. Thus, the similarity is computed in the s-sim stage. In this case, the rated similarity of the initial state to the goal and that of the goal to the initial state are sometimes quite different. We suppose that this is the main source of asymmetry of similarity in the problem-solving contexts.

What happens if a naive subject is asked to rate the similarity? *SFC* predicts that the asymmetry of similarity is not observed in such a case. Since the subjects do not recognize the goal or have any available skills for transformation, their judgments should be based on the number of shared features. In addition, for those subjects neither of the states is more prototypical. Therefore, judged similarities should not be asymmetric.

In order to explore this hypothesis, we conducted an experiment using the Tower of Hanoi puzzle. Subjects were asked to rate a similarity of a state to the goal and that of the goal to the state. In the training condition, subjects were taught 'move-pattern strategy' (Simon, 1975). This strategy can be described as follows: On odd-numbered moves, move smallest disk; On even-numbered moves, move another disk; The smallest disk is always moved from the left to the right to the center to the left peg, and so on. Since this strategy is rather mechanical or rote in a sense that this strategy does not require people to recognize the task structure, it is likely that the subjects acquire routinized skills about the puzzle. Thus, teaching this strategy is likely to enhance the acquisition of highly directional skills of the puzzle.

Experiment 1

Method

Subjects Subjects were 21 undergraduate students. They were randomly assigned to one of the two conditions: the training and control conditions. None of the subjects in the control condition had any prior experience with the Tower of Hanoi puzzle.

Procedure Subjects in the training condition first read instructions that described the rule of the Tower of Hanoi puzzle and the strategy to solve it. The strategy taught to subjects was 'move-pattern strategy', as described earlier. Next, subjects proceeded to the training phase. In this phase, they were given the three-disk puzzle with a fixed initial state where all disks were on the leftmost peg and asked to move all the disks to the rightmost peg. After subjects could solve it in six successive sessions, they proceeded to the rating phase. In this phase, subjects were asked to judge the similarity between the goal and the other states. Of the total of 52 pairs, a half were used to rate the similarity of the goal to another state and a remaining half were used to rate that of a state to the goal. Subjects were asked to circle '7' if the pairs were very similar, '1' if they were very dissimilar, and other numbers for the intermediary degrees of similarity. Subjects in the control condition skipped the instructions and training phase, and proceeded directly to the rating phase.

Results and Discussion

We first examined whether the similarity judgments of the subjects were based on both domain knowledge and the goal by regression analysis. The mean regression coefficients between distance and the rated similarity were -0.31 in the training condition and 0.05 in the control condition. The difference between group was significant ($t(11.5) = 4.75; p < 0.01$). This confirmed that the subjects in the training condition incorporated the goal and knowledge of the domain into their judging similarities.

The degree of the asymmetry was defined as $|sim(X, G) - sim(G, X)|$, where G and X represented the goal and one of the states, respectively. The mean degree of asymmetry in the training condition was 0.89, while that in the control condition was 0.41 ($t(19) = -3.50; p < 0.01$).

Our hypothesis was supported by the result that the degree of asymmetry in the training condition was greater than that in the control condition. The reason the subjects in the training condition judged similarities asymmetrically is due to the fact that these subjects acquired the directional skills specialized to achieve the fixed goal. In other words, they could create relevant features and use them in the judgments when the goal identical to the one in the training session was involved in the judgment task. However, they could not do so when the goal state was different.

Although we concluded the difference between the two conditions was attributed to whether subjects recognized the goal and had appropriate operators, there might be an alternative interpretation. Subjects in the training condition frequently observed the goal state in the training session. Every time they practiced, they had to keep the goal state in mind. This might leads features of the goal to be more salient. In other words, the goal might be a 'prototype' state, just as China is more prototypical than North Korea in a Tversky's example. If so, observed differences between the two conditions could not be attributed to the recognition of the goal and domain knowledge.

Experiment 2

In order to examine whether the observed differences between the two conditions merely reflect the salience of the features of the goal state, we conducted the second experiment.

One way to examine this possibility is to compare the ratings of the subjects in the training condition in Experiment 1 with those of other subjects who are trained to acquire flexible (non-directional) skills. If two groups of subjects receive the same amount of training but the degree of asymmetry is different, the results cannot be attributed by the prototypicality of the goal states. There are several different strategies to solve the Tower of Hanoi puzzle (Simon, 1975). One of the strategies is 'perceptual strategy.' It can be described as follows: to construct the tower of disks on the right peg, the largest disk must be placed on the right peg first, the next largest, and so on; to move the largest disk on the right peg, the others must be placed on the center peg. Unlike the move-pattern strategy in the Experiment 1, this strategy requires subjects to construct subgoals. This leads subjects to understand the task structure better and enables them to deal with unfamiliar situations flexibly. In this sense, the skills acquired through the training of this strategy are expected to be less directional than those of the move pattern strategy. Subjects who learn this strategy may create relevant features and use them in the similarity judgment even when different goals and states are involved. If so, the degree of asymmetry should be less than those observed in the Experiment 1.

Method

Subjects Subjects were 15 undergraduate students. They were randomly assigned to one of the two conditions, the MPS (move-pattern strategy) or the PS (perceptual strategy) conditions.

Procedure The procedure and the material for the MPS condition were identical to the training condition in the Experiment 1. Except that the perceptual strategy was taught to the subjects in the PS condition, the procedure and the material for the PS condition were identical to the MPS condition. Subjects in the both conditions were asked to solve the puzzle six times in the training session. The initial state in this session was fixed so that all the disks were placed in the leftmost peg. The goal was also fixed so that all the disks were placed in the rightmost peg.

Results

The mean degree of asymmetry in the MPS condition was 0.84, while that in the PS condition was 0.74. The difference of the asymmetry between the two conditions was not significant ($t(10.8) = 0.8173; p = 0.43$). Subjects in the both conditions judged the similarity asymmetrically. This might reflect that the degree of expertise of the subjects in the PS condition was less than we initially expected.

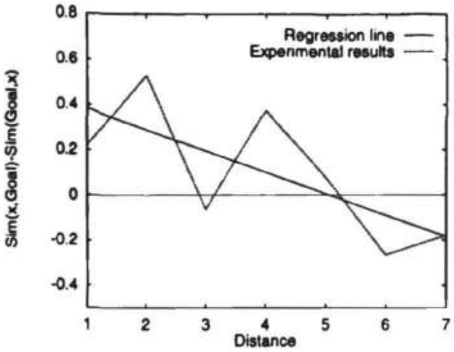

Figure 3: The degree of asymmetry in MPS condition

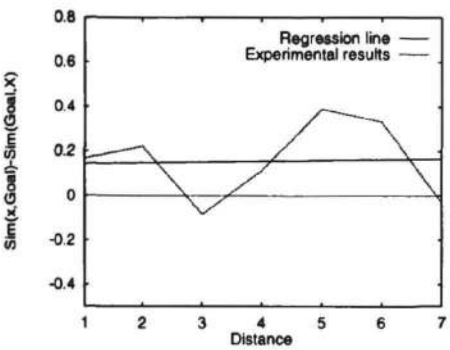

Figure 4: The degree of asymmetry in PS condition

We restrict our analysis to those pairs where each state was on the optimal solution path. Since the subjects in the PS condition rarely turned off the optimal solution path in the training session, they might acquire non-directional skills to transform states at least on this path. If so, these subjects might judge the similarity symmetrically.

The mean degree of asymmetry of the pairs on the optimal solution path in the MPS condition was 1.03, while that in the PS condition was 0.43. The difference between the two conditions was significant ($t(13) = -2.70; p < 0.05$).

Finally, we analyzed whether skills acquired through the training were actually directional or not. According to our hypothesis, when the subjects in the MPS condition judge the similarity of a state to the goal (StoG), their judgment should be based on the distance between the two. In contrast, their similarity judgments of the goal to a state (GtoS) should not be based on the distance because their skills cannot be applied to this kind of stimulus pairs. Since the number of matched features of each state in each distance is approximately equal, the degree of the asymmetry should be a decreasing function of distance in the MPS condition. In contrast, the degree of the asymmetry should not be a function of the distance in the PS condition. It is because their judgments are based on the distance, whether StoG or GtoS is asked.

General Discussion

We hypothesized that some features were created by the goal and knowledge, and that the asymmetry of similarity was attributed to the directionality of acquired skills. The results of the experiments supported the hypothesis. The asymmetry of similarity took place in a problem-solving context. When a subject recognizes the goal, has routinized skills, and can apply them, judged similarities are a function of the number of steps to achieve the goal. In other cases, the judgments are based on shared features. In addition, the Experiment 2 showed that the phenomenon could not be reduced to the prototypicality of the goal state. Even when the goal state was observed more frequently than the others, subjects with a flexible strategy did not make asymmetric judgments.

Routinized skills has directionality in the sense that it can be applied only to a small subset of possible states. For these states, subjects with such skills can create features relevant to the goal and make use of them in the similarity judgments. However, for other states, they cannot create the features and consequently rely on physical features available to them. This is the source of the asymmetry of similarity in the problem-solving context.

By incorporating the goal recognition and domain knowledge as the model's crucial components, *SFC* can explain these findings. The model can create features relevant to the goal achievement from the input (the physical description of the states), the prestored strategies, and the domain knowledge. In the *d-sim* phase, the created feature is of primary importance to compute a deep similarity. If the model fails to create the feature due to the lack of relevant strategies, it shifts to computing a shallow similarity. Therefore, according to the model, the difference of features used in the judgment of similarity is the main source of the asymmetry. In addition, *SFC* can explain the results obtained from the Experiment 2. According to the model, whether new features can be created is dependent on the strategies. The model, if equipped with a flexible strategy such as the perceptual one, succeeds in creating a feature even in unfamiliar situations. In this case, both the similarity of the goal to a state and that of a state to the goal are computed in the *d-sim* phase, which makes little difference in the rated similarities between the two.

SFC is in a sense similar to the 'transformation structure model' (Imai, 1977). According to his model, people judge that the pattern A is similar to the pattern B when A can be transformed to B. A degree of similarity is defined as the number of operators that transform one item to another. However, his model does not provide any explanation for the asymmetry of similarity as well as the expert–novice differences found in the previous study (Suzuki, Ohnishi, & Shigemasu, 1992).

SFC has several implications for the study of learning and analogy. It is well-known that experts attend to the structural aspect of problems, whereas novices do to the superficial one. According to *SFC*, this difference is due to whether one can create relevant features. Novices usually lack appropriate knowledge of the domain that enables them to create features structurally organized by the principles of the domain. This leads them to the reliance on superficial features.

When making an analogy, one has to retrieve a source analog that is structurally similar to the target. To do so, a deep similarity must be computed. In this sense, there is a possibility for *SFC* to be extended to a model of analog retrieval, although there are a lot of things to be done for its extension. One obvious advantage of *SFC* to other models such as ARCS (Thagard et al., 1990) is that it is not necessary for the model to be given a sheer description of a target. In making an analogy, it is often the case that only the superficial information of the target is available and important one has to be inferred. *SFC* can create important features from a partial description of the target and utilize them in computing a deep similarity.

References

Chi, M. T. H., Feltovich, P. J., & Glaser, R. (1981). Categorization and representation of physics problems by experts and novices. *Cognitive Science*, 5, 121–152.

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

Gentner, D. & Forbus, K. D. (1991). MAC/FAC: A model of similarity-based retrieval. In *Proceedings of the Thirteenth Annual Conference of the Cognitive Science Society*. Hillsdale, NJ: Larence Erlbaum Associates, 1991.

Hammond, K. J. (1990). Case-based planning: A framework for planning from experience. *Cognitive Science*, 14, 385–443.

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

Imai, S. (1977). Pattern similarity and cognitive transformations. *Acta Psychologica*, 41, 433–447.

Kolodner, J. (1989) Selecting the best case for a case-based reasoner. In *Proceedings of Eleventh Annual Conference of Cognitive Science Society*. Hillsdale, NJ: Larence Erlbaum Associates.

Leak, D. B. (1992). Constructive similarity assessment: Using stored cases to define new situations. In *Proceedings of the Fourteenth Annual Conference of Cognitive Science Society*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Murphy, G. L., & Medin, D. L. (1985) The role of theories in conceptual coherence. *Psychological Review*, 92, 289–316.

Simon, H. A. (1975). The functional equivalence of problem solving skills. *Cognitive Psychology*, 7, 268–288.

Suzuki, H., Ohnishi, H. & Shigemasu, K. (1992). Goal-Directed Processes in Similarity Judgements. *Proceedings of the Fourteenth Annual Conference of the Cognitive Science Society*. (pp. 343–348). Hillsdale, NJ: Lawrence Erlbaum Associates.

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

Tversky, A. (1977). Features of Similarity. *Psychological Review*, 84, 327–352.