Computational Analysis of Graphic Generation: Effects of surface and structural similarity

Junya Morita (j-morita@jaist.ac.jp)
School of Knowledge Science, Japan Advanced Institute of Science and Technology
1-1 Asahidai, Nomi, Ishikawa 923-1292, Japan

Abstract

In studies of analogical reasoning, the distinction between surface and structural similarity has been repeatedly investigated. However, this distinction has not been investigated in a generative analogy where target representations are not provided in advance. This study uses computational methods to analyze how this distinction is involved in generative analogy. In the experiment, participants were given an example and asked to generate an original graphic by modifying the example. The procedure was repeated twice. The result indicated that the first graphics that the participants generated were similar in both surface and structural features to the presented example. However, the second graphics that they generated varied in the degree of surface and structural similarity. This result may indicate characteristics of generative analogy but is also compatible with the results of previous studies of analogical reasoning.

Keywords: Similarity; Analogical Reasoning; Creation; Learning by example

Introduction

Similarity is a central component of intelligent processing in both human thinking and computational systems. In particular, analogical reasoning, which infers a *target* situation by applying a known *base* instance, is known to play a central role in human cognition. There have been numerous psychological and computational studies of the nature of analogical reasoning.

Recently, computational modeling on this topic extended the scope of its field of application to include more complex, large-scale situations. One such situation is visual analogy. For example, Forbus and Usher (2002) constructed a model that shows mapping between two freehand drawings and Ferguson (2007) proposed a method that flexibly constructs mapping between non-isomorphic pictures. These studies offer a new direction for studies of analogical reasoning by using recent developments in computational resources.

In contrast, there has been little extension of the research framework for psychological studies of analogical reasoning. Indeed, although recent experimental research on analogical reasoning controls details factors based on a computational model and uses complex stimuli (e.g., Gentner & Sagi, 2006), the tasks used in these studies are similar to those in studies conducted 10 years ago. The main task used in studies of analogical reasoning has always been similarity judgment or mapping between a base and target.

The authors consider that the tasks used in these previous studies involve analogical reasoning that defines target representations in advance. In the real world, analogical reasoning is applied to a task where humans freely generate a target representation after observing a base. In this paper, we call such activity *generative analogy*. A creation of visual art is an example of generative analogy, where artists explore novel representations (target) based on observations of a natural object as a motif (base). Although some researchers have dealt with generative analogy in psychological studies (Ishibashi & Okada, 2004; Okada, Yokochi, Ishibashi, Namba, & Ueda, 2007), there are no studies presenting a formal analysis that links generative analogy with computational models of analogical reasoning.

Therefore, in the present study, we attempted to apply formal computational analysis to a generative analogy. To achieve this goal, we used computational models of analogical reasoning to score work produced in a generative analogy task. Based on the computed scores, we investigated the hypotheses described below.

Hypotheses

The focus of studies on analogical reasoning is the distinction between two types of similarity: similarity of surface features and similarity of structural features. Surface features indicate the attributes of objects, and structural features indicate the relations between objects. The structure mapping theory (Gentner, 1983) used predicate calculus to represent this distinction: surface similarity is a matching of predicates that have a single argument. Structural similarity is constructed based on matching predicates that have several arguments. This distinction was implemented in a computational model constructed by Falkenhainer, Forbus, and Gentner (1989).

The structure mapping theory assumes that useful analogical reasoning is achieved by using a base that shares only structural features with a target. However, many psychological studies have indicated a difficulty in using structural similarity in analogical reasoning (Holyoak & Koh, 1987). People usually prefer *literal similarity* in which a base shares both surface and structural features with a target (Gentner, Rattermann, & Forbus, 1993). If there is no surface similarity, they take little notice of structural similarity. The exceptions are experts in the target task domain (Novick, 1988) or those who have received sufficient training on the target task (Markman & Gentner, 2000).

Following the above studies, the present study investigates the distinctions between surface and structural similarity in a generative analogy task. Specifically, we examined the following two hypotheses:

- Our *literal similarity hypothesis* predicted that participants would be influenced by both the surface and structural features of a presented base when they performed the target task first.

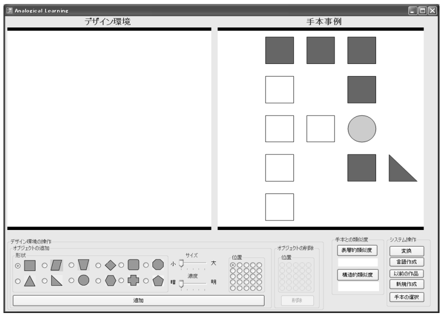

Figure 1: User interface of task environment.

- Our *relational shift hypothesis* predicted that participant would be strongly influenced by the structural features of presented example after they practiced the target task.

Task

In this study, we formulated a generative analogy task af ter investigating basic education strategies in visual art. A Ishibashi and Okada (2004) indicate, learning by example i a popular learning method in this field. In our study, partic ipants generated original graphics by modifying an example This task is similar to the type of analogical modification tha Okada et al. (2007) investigated.

The type of graphic that participants generated was *graphic composition*, which requires learners to create an attractive and original layout by arranging simple geometric shapes. In basic education on visual art, learners work on this task to develop the ability to control attributes of an individual object with the aim of constructing a wholly consistent graphic. This learning goal of graphic composition seems to be compatible with the assumptions of the structure mapping theory.

We developed a task environment in which participants observed an example of a graphic composition and then produced an original composition (Figure 1). The user interface consisted of two panels: the right panel presented an example and the left panel provided a space in which the participants created their composition. Each panel consisted of a 5 by 5 grid and the graphics were created by arranging several objects on the grid. The environment provided menus for assigning the values of the x-axis, y-axis, darkness, size, and shape.

Method of analysis

For the above task environment, we developed a method of computing the surface similarity and structural similarity of the graphics generated by a participant (target) and the graphic presented to the participant (base). We assumed that if a target shared structural features with a base, the participant who had generated the target was influenced by the structural features of the graphic. Similarly, if a target shared common surface features with a base, we assumed that the surface features of the base influenced the creation of the graphic. Our method of computation, which was a modification of previous computational models of analogical reasoning, is explained as follows.

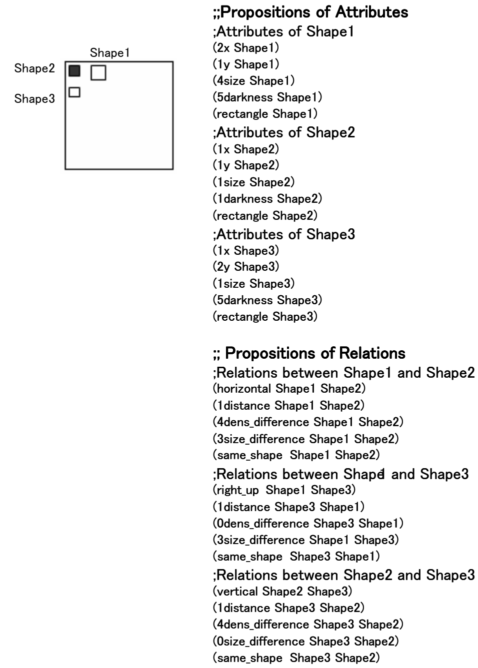

Figure 2: Example of coding.

Coding The computation of both types of similarities uses a coding scheme that represents a graphic as propositions. An example of coding is presented in Figure 2. This coding contains the following two types of propositions.

- *Propositions of attributes* describe values selected by a participant when the graphic is generated. Specifically, the propositions indicate the values of five dimensions of attributes: *location on x-axis, location on y-axis, lightness, size*, and *shape*. These values are obtained directly from the task environment.
- *Propositions of relations* indicates the relations between pairs of objects. These relations are described according to five dimensions: *distance, direction, lightness difference, size difference*, and *shape difference*. These values are obtained from the attributes of the two objects. *Distance* and *direction* are obtained from the *locations on the x- and y-axis*. *lightness difference, size difference*, and *shape difference* are obtained from each corresponding attribute of *lightness, size*, and *shape*.

Surface similarity

Surface similarity is computed as shared attributes between a base and a target. We used content vector matching (Forbus, Gentner, & Law, 1995) in our computations of surface similarity. Our method creates content vectors for each dimension of attribute (k) by counting each value. For example, a vector *((2x 1) (1x 2))* is created as a content vector of *x-axis* in Figure 2. For all five dimensions (d), a dot product of a content vector of the target (t_k) and a content vector of the base (b_k) is calculated. Finally an average of the dot products is obtained as follows[1].

$$Surface\ similarity\ score = \sum_{k=1}^{d} \frac{t_k \cdot b_k}{d}$$

Structural similarity

Structural similarity was computed as the common relational structures of a base and target. The commonality of their structure was computed by estimating the maximum mapping from the base to the target while satisfying two constraints: parallel connectivity and one-to-one mapping (Gentner, 1983). The former means that if two predicates are placed into correspondence then the arguments to these predicates are also placed into correspondence, and vice versa. The latter indicates that each item in the base maps to, at most, one item in the target, and vice versa.

An example of a common relational structure is illustrated schematically in Figure 3, where descriptions are represented as graph structures. The top and middle graphs in the figure represent the base and target structures constructed from the graphics on the left side of the figure. The oval nodes represent predicates and the boxed nodes represent objects. There are two types of edges: solid edges connecting a predicate with its first argument and dashed edges connecting a predicate with its second argument. If the predicate is commutative, there are no distinctions between these two types of edges.

The bottom network in the figure represents a common structure of the base and the target. The structural similarity score is quantified as the number of predicates in the global map. In the case of Figure 3, the score is 10, which is shown in the lower right-hand corner of the figure. The value in parentheses is the size of the global map as a fraction of the size of the target structure.

Generally, the common structure of two graphs can be extracted using graph-matching algorithms. The present study used an algorithm modified from the SME (Structure-Mapping Engine; Falkenhainer et al., 1989), which includes the following two steps.

Local match construction First, the correspondences of propositions (P-match) are constructed by comparing the predicates in the base proposition with the predicates in the target proposition. If the two propositions have a predicate that is the same, a P-match is created. Each of the P-matches consists of a pair of predicates (a Pre-match) and pairs of arguments (O-matches). This process is applied to every possible combination of propositions, and a list that contains every constructed P-match is created.

Global map construction In the second step, a global map that is a set of consistent P-matches is created. This step uses a ranking algorithm, as Forbus and Oblinger (1990) did. In this study, we used a weight that was a simple summation of two individual O-match frequencies and the co-occurrence frequency of two O-matches for ranking.

The basis for connecting a P-match is an investigation of whether a pair of P-matches conflict with each other. Conflicts are defined as situations where several P-matches share the same object but have a different O-match that includes this object. Beginning with the P-match that has the highest weight, this method sequentially chooses one of the P-matches and deletes the P-matches conflicting with it. The process results in a set of consistent P-matches.

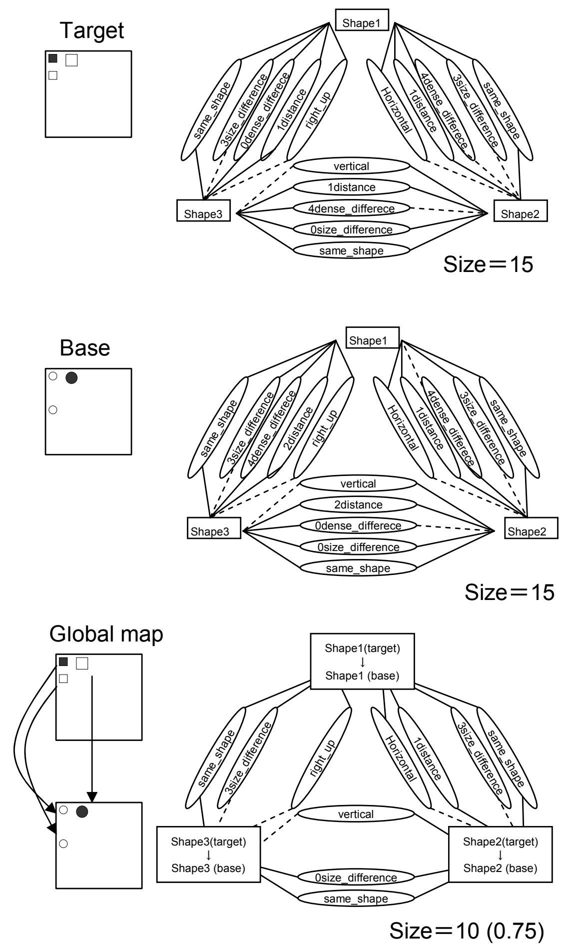

Figure 3: Example of common relational structure.

Experiment

We conducted an experiment to test our literal similarity and relational shift hypotheses. To test the hypotheses, we controlled the degree of experience by presenting participants with two examples separately. The participants generated their original graphics by modifying each graphic. It can be assumed that the generation of the second graphic was a more familiar situation for the participants who had no previous experience in graphic composition.

Participants Nineteen graduate student volunteers from the Japan Advanced Institute of Science and Technology par-

[1] To produce surface similarity scores ranging from 0 to 1, feature vectors were normalized to unit vectors.

ticipated in the experiment. Most of them had no basic knowledge of graphic composition.

Materials The examples used in this experiment are shown in Figure 4. A ceramic artist who graduated from an art university and who has a master's degree in architecture created the examples.

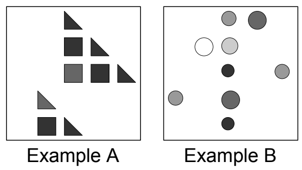

Figure 4: Examples used in experiment.

Procedures Participants took part individually in the experiment, which included the following steps.

Instruction: The participants were told the goal of the experiment was "an investigation of learning by example in graphic composition", and they were asked to create original graphics by modifying the presented example. In addition, the experimenter strongly prompted the participants to use the features in the example in their graphics and to make their graphics as creative as they could.

First session: The experimenter placed one of the examples in the task environment. Ten participants were presented with Example A, and the other nine participants were presented with Example B. The participants generated their original graphics while observing the example.

Second session: After the participants finished generating their graphic in the first session, they received an example that had not been provided in that session. As in the first session, they generated their graphics by modifying the example for a period of thirty minutes. In the following analysis, we investigated the difference between the two sessions by counterbalancing the effect of graphic types.

Results

To investigate the distinctions between surface and structural similarity in the generative analogy task, we computed two similarity scores for the generated graphics. Using the computed scores, we then tested our two hypotheses.

Mean scores for surface and structural similarity Figure 5 presents the means of the surface and structural similarity scores for the two sessions. In addition to the scores that were computed for the presented example, this figure shows the scores computed for the example that was not presented in the session. The former is represented by the bars labeled "Influence" and the latter is indicated by the bars labeled "Control." For example, when the graphics were generated by modifying Example A, we call the similarity score computed for Example A "Influence", and the similarity score computed for Example B "Control." To judge whether the participants were influenced by the features of the presented example, we compared "Influence" with "Control."

We conducted a 2 (order: first-second) $\times$ 2 (base type: influence-control) analysis of variance on the surface and structural similarity scores. As a result, significant main effects of the base type were obtained for both the surface similarity score [$F(1,18) = 29.07$, $p < .01$] and the structural similarity score [$F(1,18) = 7.74$, $p < .05$]. We could not detect significant main effects of the order [surface similarity: $F(1,18) = 0.00$, *n.s.* structural similarity: $F(1,18) = 0.02$, *n.s.*] and interactions of the base type and order [surface similarity: $F(1,18) = 0.09$, *n.s.* structural similarity: $F(1,18) = 0.11$, *n.s.*].

The results support the literal similarity hypothesis by indicating the influence of the surface and structural features on the generative analogy task. However, we could not find evidence to support the relational shift hypothesis because no effects concerning the order difference were observed.

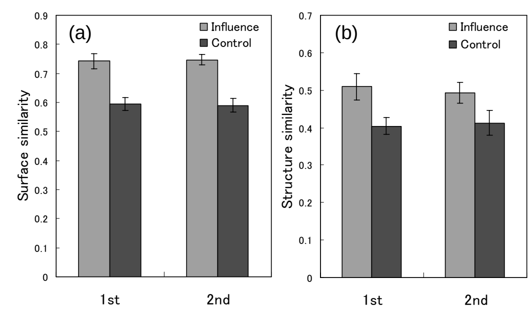

Figure 5: (a) Mean scores for surface similarity, (b) Mean scores for structural similarity. Error bars indicate standard errors of means.

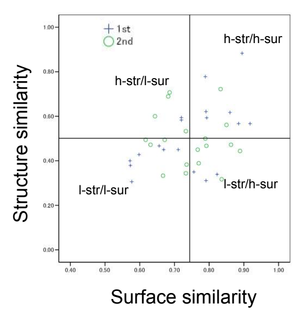

Figure 6: Scatter plot of graphics.

Correlations of the two similarity scores Although the above results did not indicate a relational shift, we could not conclude that there were no differences in the influence of surface and structural features on participants' generation of graphics between the first and second sessions. Since Figure 5 presents the means of the two similarity scores independently, we were unable to observe how the two scores related to each other.

To explore the relations between the two similarity scores, we drew a scatter plot whose two axes corresponded with the two similarity scores (Figure 7), and computed the correlation coefficients of the two scores. The results showed a significant correlation for the first session [$r(17) = .561$, $p < .05$] but no significant correlation for the second session [$r(17) = -.118, n.s$]. The difference between the correlations was statistically compared using Fisher's Z-transformation. The results showed there was a significant difference between the two coefficients [$\chi(1) = 4.33$, $p < 01$].

These results are consistent with the literal similarity hypothesis because the two similarity scores were correlated for the first session. However, the same pattern did not apply to the second session where the two similarity scores were discriminated. It could be considered that in the second session two types of graphics were generated: graphics sharing only the structural features of the presented example, and graphics sharing only the surface features of the presented example. The following analysis presents cases in which the two similarity scores were discriminated.

Case studies Figure 6 can be divided into four areas by the means of surface and structural similarity (the two lines in the figure). We defined representative cases of each area as cases that had the biggest normalized distance from the two means of the similarity scores (the cases labeled with *h-str/h-sur*, *h-str/l-sur*, *l-str/h-sur*, *l-str/l-sur* in the figure).

Figure 7 presents screenshots of our analysis system showing details of the computations for each case. The left part of each screen presents two graphics that are inputs of similarity computations. The lines on the graphics represent O-matches in the computed structural similarity. The size of the circles on the objects corresponds with the number of Pre-matches that the O-match is included in. The right part of each screen presents stacked bar graphs where the degree of similarity of the two scores was divided into each dimension. In the graphs, the colors of the bars are associated with the dimensions as follows: light blue - *x-axis/distance*, yellow - *y-axis/direction*, red - *size/size difference*, green - *lightness/lightness difference*, dark blue - *shape/shape difference*.

From the outputs of our computations, we can speculate on situations in which two similarity scores were discriminated. In the cases labeled with *h-str/h-sur* and *l-str/l-sur*, the ratios of the degree of similarity in five dimensions were almost the same as in the two similarity scores. In contrast to the above cases, in the cases of *h-str/l-sur* and *l-str/h-sur*, there are differences in the ratios of the corresponding dimensions between the two similarity scores. That is, in the case of *h-str/h-sur*, the surface similarity score of the *x-axis* was very low, but the structural similarity scores for *direction* and *distance*, which are relations constructed from the values of the *x-axis* and *y-axis*, were high. Similarly, in the case of *l-str/h-sur*, the surface similarity scores for *lightness* and *size* were high, but the structural similarity scores for *lightness difference* and *shape difference* were low.

That is, in cases in which the two similarity scores were separated, the scores were different in the degree of similarity in corresponding dimensions. This difference can be explained by the method used for computation. Computation of the surface similarity score used the absolute reference values of the attributes, but the structural similarity score was computed using the relative reference values of the attributes. Therefore, the structural similarity score was robust in cases where the locations of objects moved in parallel, such as *h-str/l-sur*. Also, the surface similarity score was computed as the means of dimensions, but the structural similarity score was computed as the size of the single global mapping. Therefore, the structural similarity score was naturally low when there were many conflicting O-matches in a set of local matches, as *l-str/h-sur* indicates.

Discussion

Our results supported the literal similarity hypothesis in a generative analogy task; the means of the two similarity scores exceeded those of the control conditions in the first session, and there were significant correlations between the two similarity scores. Thus, our study successfully extended the field of application of analogical reasoning to include generative analogy.

However, our findings did not support the relational shift hypothesis. Instead, we observed that the two similarity scores were discriminated in the second sessions.

A possible interpretation of this discrimination is that after the first session, participants definitely noticed structural features in the graphic composition. However, some of them did not use the structural features of the example in the second session because they did not consider that a simple copy of the structure would lead to an original graphic. This interpretation does not seriously contradict previous findings on analogical reasoning. Rather, we consider that this result may indicate characteristics of generative analogy tasks.

Of course, our results could be interpreted on the basis of experimental manipulation. In our experiment, participants practiced the task of generating a graphic composition only twice. It could be considered that this amount of experience was insufficient to produce a relational shift. If they had more experience, they might have created graphics that shared only structural features with the example. However, this interpretation does not explain the changes in the correlation between the two similarity scores. Therefore we consider that the first interpretation is more feasible.

However, we must note that the results depend on the implementation of similarity computation, as our case studies indicate. Our method constructs structural similarity only from first-order relations, although many researchers, including Gentner (1983), have pointed out the importance of higher-order relations (relations of relations) in analogical reasoning. According to previous studies, structural similarities constructed from first-order relations, which our study used, are relatively easy to notice, even by novices.

We did not include higher-order relations in our method of

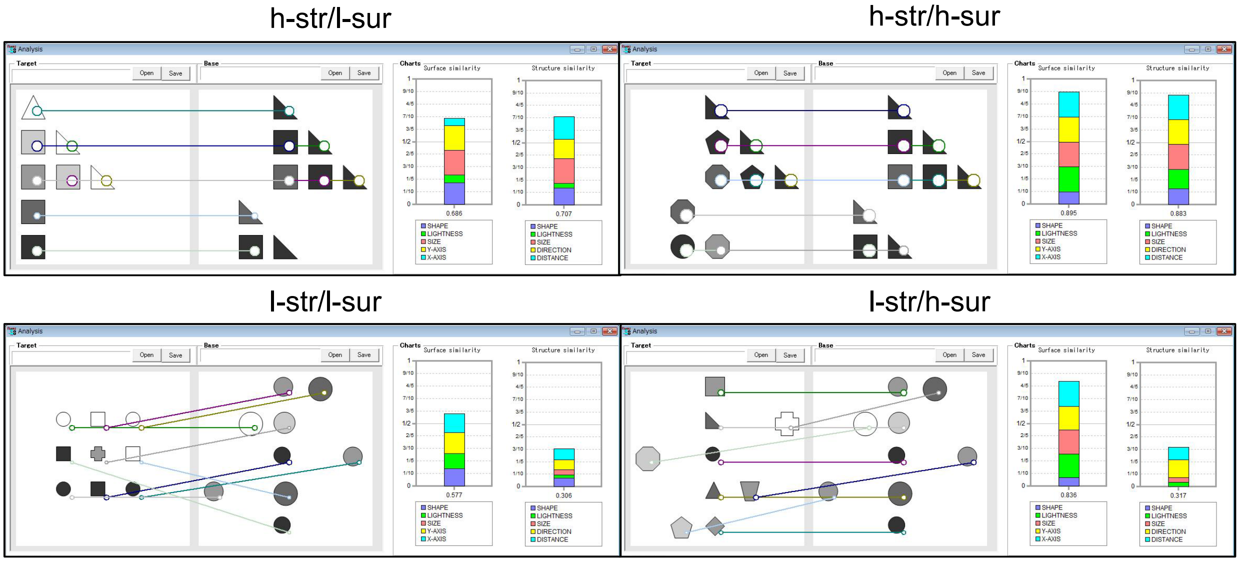

Figure 7: Case studies. Our analysis system indicates details of similarity computations.

analysis because we could not find a method to clearly describe these relations. To construct mapping of higher-order relations in computations of visual analogy, previous studies have proposed such methods as describing the creator's intention (Forbus & Usher, 2002), chunking similar objects (Ferguson, 2007), and using incremental mapping (Tomai, Lovett, Forbus, & Usher, 2005). However, we could not apply these methods to our task without making arbitrary assumptions. Our method is a simple one, but it is fully automated. Therefore, we consider it could play a role in providing objective criteria for examining links between human free thought and cognitive theories.

Acknowledgement

This study is supported in part by the Grant-in-Aid for Young Scientists (B) from the Ministry of Education, Culture, Sports, Science and Technology of Japan (No. 19700243, 2007).

References

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

Ferguson, R. W. (2007). On the mediation of levels of similarity in analogical comparisons. In *Analogies: Integrating multiple cognitive abilities* (pp. 15–20). Universitat Osnabrueck.

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

Forbus, K., & Oblinger, D. (1990). Making sme greedy and pragmatic. In *Proceedings of the 12th annual conference of the cognitive science society* (pp. 61–68). Lawrence Erlbaum.

Forbus, K., & Usher, J. (2002). Sketching for knowledge capture: A progress report. In *Proceedings of the 7th international conference on intelligent user interfaces* (pp. 71–77). ACM Press.

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

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

Gentner, D., & Sagi, E. (2006). Does "different" imply a difference? In *Proceedings of the twentyeighth annual meeting of the cognitive science society* (pp. 47–266).

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

Ishibashi, K., & Okada, T. (2004). How copying artwork affects students' artistic creativity. In *Proceedings of the 26th annual conference of the cognitive science society* (pp. 618–623).

Markman, A. B., & Gentner, D. (2000). Structure-mapping in the comparison process. *American Journal of Psychology*, *113*, 501–538.

Novick, L. R. (1988). Analogical transfer, problem similarity, and expertise. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *14*, 510–520.

Okada, Y., Yokochi, S., Ishibashi, K., Namba, K., & Ueda, K. (2007). Analogies: Integrating multiple cognitive abilities. In *Analogies: Integrating multiple cognitive abilities* (pp. 47–50). Universitat Osnabrueck.

Tomai, T., Lovett, A., Forbus, K., & Usher, J. (2005). A structure mapping model for solving geometric analogy problems. In *Proceedings of the 27th annual conference of the cognitive science society* (pp. 2190–2195).