Supporting Perspective Changes in Graphic Composition

Junya Morita† **Yukari Nagai**† **Toshiharu Taura**‡
†School of Knowlege Science
Japan Advanced Institute of Science and Technology
‡Graduate School of Engineering Kobe University
{j-morita,ynagai}@jaist.ac.jp, t-taura@kobe.ac.jp

Abstract

One of the basic skills in design work is the composition ability that controls the attributes of objects in constructing wholly consistent figures. One way to acquire this skill is through learning by examples, which is widely adopted in design or art classrooms. However, it is necessary for an effective learning to help a learner gain an adequate perspective of an example. In this paper, we describe a method of extracting lerners' perspectives from their works. Our method is used in the task where leaners are presented with an example and compose a novel visual space. Once the learner completes the task, the system automatically calculates the scores of similarity between the space composed by the learner and the example presented by the system. Then, it gives the scores to the learner as part of the learning feedback. This feedback helps the learner extract abstract principles from the presented example.

Keywords: Design education, analogical reasoning, composition ability

1 Introduction

Creative thinking can be considered an activity that generates novel combinations of known entities in a given domain. This view is contrasted by the idea that creative thinking is accomplished by generating novel entities out of nothing. Compared to the later view, the former one is commonly accepted in many fields. For example, linguistic researchers often think that the creative aspect of language use comes from the infinite use of finite elements[1].

Cognitive science research also investigates creative thinking by conducting experiments in which novel ideas are generated through the combination of a variety of object parts[2]. Furthermore, Nagai and Taura analyzed the process of concept design through the study of the primitive operations of synthesizing individual concepts[3].

In addition to the above fields, the combinations of elements are important in the domains that involve visible entities, such as art or graphic design. This is especially true in design education, to cultivate the composition ability that controls the attributes of objects in constructing wholly consistent figures.

In this paper, we propose a method of cultivating composition ability through the development of a learning support system. Basically, our method follows the traditional pedagogy. We employ the "learning by examples" concept, which is widely adopted in design or art classrooms. For example, students in art class often recreate famous works. Observational learning is also widely considered an effective way for assisting learners in design education. As with the proverb, "a picture is worth a thousand words", it might be difficult to communicate visual information with words[4].

However, there is a problem with learning by examples. If the proverb were true, there would be a thousand features in a visual example. In this type of situation, the question that would arise would be, which features should the learner focus on? Obviously, it is necessary for an effective learning support system to help a learner gain an adequate perspective of an example.

Our proposal is motivated by the above problem. In the next section, we propose a hypothetical solution to the problem, by reviewing some of the studies on human analogical reasoning. In third section we describe the learning support system. The fourth section presents a preliminary experiment that evaluates the system. The final section presents the proposals implications and future studies are discussed.

2 Hypothesis

As noted earlier, learning by examples has a problem with the perspective settings. In this section the problem is considered from the viewpoint of analogical reasoning. Analogical reasoning is an activity constructing a mapping from the known base domain to the novel target domain[5]. This type of activity has been intensively investigated in the field of cognitive science. The main problem of focus has been on the constraints imposed on analogical reasoning. Since numerous commonalities can be found between any two domains, it is impossible to construct a mapping without constraints that can determine the similarities between two domains.

It can be assumed that the same problems underlie in analogical reasoning and learning by example. Therefore, we propose a method of learning by examples applying the findings of analogical reasoning. Past studies have revealed that several constraints are involved in analogical reasoning, including the following two types of similarities[6; 7; 8].

- Surface similarity: the number of attributes shared between the base and the target.
- Structure similarity: the commonality of relational structure of the base and the target.

Psychological experiments have confirmed that two types of similarities influence different aspects of human cognition. That is, people tend to immediately notice surface features, whereas they prefer structure similarity to surface similarity in limited situations[8; 7]. More precisely, surface similarity mainly influences the lower-level cognition, such as for instant perception, retrieving, or associating the examples from memory. On the other hand, structure similarity involves in higher-order cognition, such as for problem solving or scientific discovery. We believe that the distinction between the two types of similarity is also important in the field of graphic design. Specifically, we have the following hypotheses.

- Learners have a tendency to focus on the surface features of objects, such as color or shapes.
- The important features in composition are not the surface features but the structure ones, such as distance, balance, or the combinations of shapes.

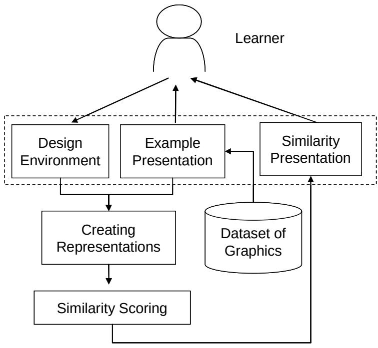


Figure 1. The leaning support system.

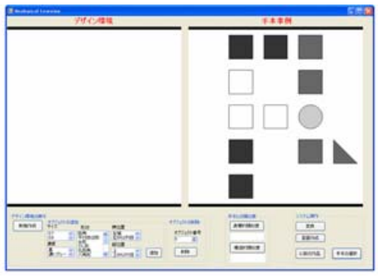


Figure 2. User interface of the leaning support system.

These hypotheses are consistent with the definition of composition ability. The combination of elements in graphic design can be considered the structural features in analogical reasoning. The present study develops the method of extracting larner's perspectives based on these hypotheses. This method incorporates an analogical reasoning model that computes the surface and structure similarities of graphics.

3 Learning support system

In this section, we describe a learning system for composition ability, one that presents the learner

Figure 3. A work of composition.

with an example and prompts him/her to compose a novel graphic. Once the learner completes the task, the system automatically calculates the scores of similarity between the graphic composed by the learner and the example presented by the system. Then, it gives the scores to the learner as part of the learning feedback. This feedback helps the learner extract the abstract principles of graphic compositions from the presented example.

Figures 1 and 2 present the system overview and its user interface, respectively. The system is composed of four main components: (1) the design environment, (2) the example presentation window, (3) the creating the propositional representation, and (4) the similarity scoring. In this section, we describe these components.

3.1 Task environment

The system provides the design environment to the learner. The learners in this environment compose a graphic on a space that consists of a 5 by 5 matrix by placing several objects on the grids. The system provides scrolling menus for assigning the following attributes to the objects.

- Location:
  The user can choose the objects locations on the horizontal (x) and vertical (y) axes but cannot place two objects in the same location.

- Size:
  The size of the objects can be chosen from five values ranging from 175 to 250 pt.

- Density:
  The density of the objects can be chosen from five values ranging from white to black.

- Shape:
  The shape of the objects can be chosen from sixteen types, including oval, rectangle, rectangular triangle, or parallelogram.

This environment was inspired by the composition tasks used in real-world design education. Figure 3 presents a work from a composition task, which was created in a ceramic art school. It was created by locating several geometric shapes.

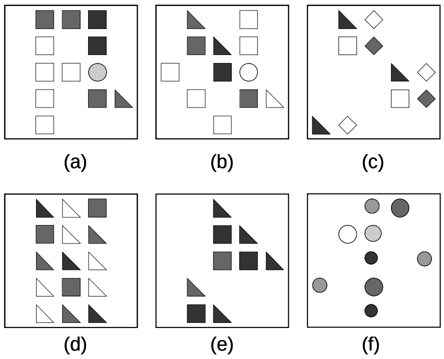


Figure 4. Examples

3.2 Examples

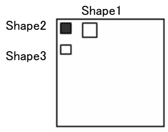
The learner considers an example presented through the example presentation window. It is chosen from a dataset including the graphics presented in Figure 4. The ceramic artist who is the author of the work in Figure 3 composed these graphics.

3.3 Representations

The system computes two types of similarities between an example and a graphic composed by learners. The calculation of the two types of similarities uses a common scheme that represents a visual scene as propositions (predicates-calculus). The representation is developed through the following two steps.

1. Representing attributes:
   The system creates propositions that describe the five dimensions of attributes; *Location on x-axis, Location on y-axis, Density, Size*, and *Shape*. These attributes are obtained directly from the task environment.

2. Representing relations:
   The system creates propositions that describe the five dimensions of relations; *Distance, Direction, Density difference, Size difference*, and *Shape difference*. These relations are obtained from the attributes of two objects. *Distance* and *Direction* are obtained from the *Locations on x- and y-axis*. *Density difference, Size difference*, and *Shape difference* are obtained from each corresponding attribute of *Density, Size*, and *Shape*.




```
;;Propositions of Attributes
;Attributes of Shape1
(2x Shape1)
(1y Shape1)
(4size Shape1)
(5darkness Shape1)
(rectangle Shape1)
;Attributes of Shape2
(1x Shape2)
(1y Shape2)
(1size Shape2)
(1darkness Shape2)
(rectangle Shape2)
;Attributes of Shape3
(1x Shape3)
(2y Shape3)
(1size Shape3)
(5darkness Shape3)
(rectangle Shape3)

;; Propositions of Relations
;Relations between Shape1 and Shape2
(horizontal Shape1 Shape2)
(1distance Shape1 Shape2)
(4dens_difference Shape1 Shape2)
(3size_difference Shape1 Shape2)
(same_shape  Shape1 Shape2)
;Relations between Shape1 and Shape3
(right_up  Shape1 Shape3)
(1distance Shape3 Shape1)
(0dens_difference Shape3 Shape1)
(3size_difference Shape1 Shape3)
(same_shape  Shape3 Shape1)
;Relations between Shape2 and Shape3
(vertical Shape2 Shape3)
(1distance Shape3 Shape2)
(4dens_difference Shape3 Shape2)
(0size_difference Shape3 Shape2)
(same_shape  Shape3 Shape2)
```

Figure 5. Example of representations.

Figure 5 shows an example of the developed representations. The graphic in the figure contains three objects: *Shape1*, *Shape2*, and *Shape3*. The attributes of each object are represented in "Propositions of attributes." The relations of every possible combination of objects are represented in "Propositions of relations."

As shown in the figure, in our scheme the predicates do not take a specific value as an augment (e.g., *(x Shape1 2)*, *(distance Shape1 Shape2 1)*) but instead indicate the values directly. For example, the proposition *(2x Shape1)* means that the object *Shape2* is located in the second column of the graphic. The proposition *(1distance Shape1 Shape2)* means that *Shape1* and *Shape2* are placed in adjacent locations. Thus, the propositions of attributes are always described with a one-place predicate, whereas the propositions of relations are described with a two-place predicate.

3.4 Similarity Scoring

The two types of similarity scores are calculated by using the above representation. The calculation of the surface similarity uses the representations of the attribute. The calculation of the structure similarity uses the representations of the relation.

3.4.1 Surface Similarity

Surface similarity is computed through a process that is slightly modified from the method presented by Forbus et. al[6]. They use a feature vector that specifies which predicates were used in that representation and the number of times they occurred. According to them, this type of vector is a flat summary of the knowledge structures. They use the dot product of these vectors as a rough estimate of overlap between two representations. They indicated that this value is positively correlated with retrievability measured in studies of the human memory system.

Our system differs from theirs in the types of predicates specified in the vectors. Our method restricts a type of predicate in the vectors to object attributes. For example, the following feature vector is created from the representation in Figure 5:

((2x 2) (1x 2) (1y 2) (2y 1) (4size 1) (1size 2) (5darkness 2) (1darkness 1) (rectangle 3))

We use the dot product of these vectors to estimate the score of the surface similarity, which reflects the overlap of attribute features between two representations.[1]

3.4.2 Structure similarity

Structure similarity is computed as the common relational structures of the base and the target. The commonality of their relational structure is calculated by estimating the maximum mapping from the base to the target. Gentner proposed the following structural consistency constraints that guide the process of mapping[5].

- Parallel connectivity:
  If two predicates are placed into correspondence then the arguments to these predicates are also placed into correspondence, and vice versa.

- One-to-one mapping:
  Each item in the base maps to at most one item in the target, and vice versa.

Similar to the models in the past studies [9; 10], our system follows the above constraint. The algorithm for computing the structure similarity consists of the following three steps.

[1] In the experiment described later in this paper, the feature vectors were normalized to the unit vectors.

Step 1: P-match Construction The system first constructs the correspondences of propositions. This step is similar to the algorithm shown by Falkenhainer et al[9]. These correspondences (P-matches) are created 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 augments (O-matches). For example, if the base contains the proposition *(vertical Shape1 Shape3)* and the target contain the proposition *(vertical Shape2 Shape3)*, the following P-match is created.

((vertical vertical) (Shape1 Shape2) (Shape3 Shape3))

This process is applied to every possible combination of propositions, and a list that contains every constructed P-match is created.

Step 2: Weighting P-matches In the second stage, in order to sort the list of P-match, the system assigns a weight to each P-match. The assigned weight is a product of the following two sub-weights.

- Frequencies of O-match:
  This weight reflects the frequencies of the O-matches in the list of P-matches. This is calculated as a summation of the individual O-match frequency (*freq (O-match$_i$)*, *freq(O-match$_j$)*) and co-occurrence frequency of two O-matches (*freq (O-match$_i$, O-match$_j$)*).

- Frequencies of Pre-match:
  This weight reflects the dimensions of the Pre-match. The number of occurrences of each dimension is counted from the entire P-match list. Then the P-match is assigned the number of corresponding dimensions as the weight.

This weighting is intended to extract the designer's focus from their works. In the composition task, a designer would focus on specific sets of objects or dimensions. We considered that the estimated mapping should contain these focused elements. This assumption is based on the results of studies on similarity judgment. Spencer-Smith and Goldstone report that human subjects tend to estimate higher similarity when the mapping is concentrated on specific objects or dimensions[11].

Step 3: Global-map construction In the third stage, the system constructs a global map that is a set of consistent P-matches. This step is similar to the algorithm presented by Forbus et al[10]. The basis for this process 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 differ from one another. Beginning with the P-match that has the highest weight, the system 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.

The global mapping process is illustrated schematically in Figure 6, where descriptions are represented as propositional networks. The networks in the top and middle parts of the figure represent the base and the 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 links: solid links connecting a predicate with its first augment and dashed links connecting a predicate with its second augment. If the predicate is commutative, there are no distinctions between these two types of links.

The bottom network in the figure represents a global map from the base to the target. It does not contain pairs of predicates whose types are different or pairs of predicates whose augments are not placed in corresponding positions. Consistent with our intuition on the graphics, this global map mainly consists of the correspondence of *Direction*, *Size difference*, and *Shape difference*.

The score of structure similarity is quantified as the number of elements in the global map. In the case of Figure 6, the score is 13, which is shown in the lower right-hand corner of Figure 6. The value in parentheses is the normalized score of structure similarity, which is the size of the global map as a fraction of the size of the target structure.

4 Experiment

A preliminary experiment was conducted, in which the participants were presented with examples (bases) and then drew their own original

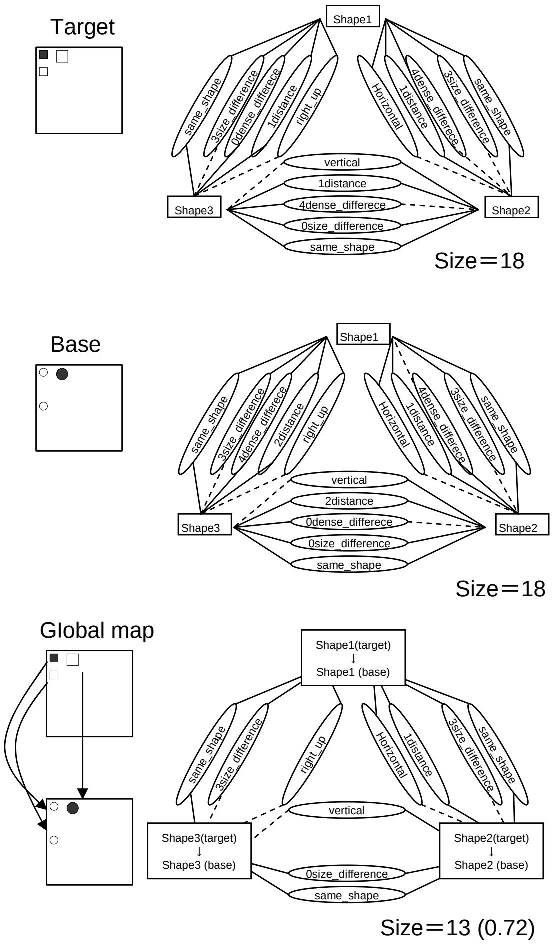


Figure 6. Global map.

graphics (targets). It was not directly designed to provide statistical evidence on the hypotheses discussed in the section two. Instead, we tried to present qualitative examples demonstrating how our system can extract lerners' perspectives from their works.

4.1 Method

4.1.1 Participants

Three graduate students (Participant A, Participant B, and Participant C) in Japan Advanced Institute of Science and Technology voluntary participated in the experiment. All of them had some knowledge of graphic composition.

4.1.2 Materials

The experiment was conducted using the system described in the above section. Two examples of the graphics were chosen from Figure 4, e and f. In the following section, we call them *Base 1* (Figure 4e) and *Base 2* (Figure 4f).

4.1.3 Procedure

Participants individually took part in this experiment, which comprised two drawing sessions. In each of the sessions, one of the two graphics was presented in the right-hand panel of the task environment, and then the participant was asked to compose a graphic in the left-hand panel of the environment. They were prompted to use the features in the example in their graphics and to make their graphics as creative as they could. All the participants were shown the two graphics in the same order. First, Base 1, then Base 2. Each session lasted for thirty minutes.

4.2 Results and Discussion

All of the participants were able to compose their graphics within the allotted thirty minutes. Figure 7 presents the graphics prepared by each participant. As can be seen, all of the graphics share sufficient commonalities with the examples (Base). In addition, there are a variety of different features seen between each of the graphics. These impressions indicate that the task adopted in this study were suitable for cultivating creative minds.

Additionally, the system computed the two types of similarity scores for each graphic (Figure 8). The distinction of the two types of similarity was investigated by using Pearson's coefficient correlations between the two scores. A high positive correlation was obtained for the scores computed with Base 1, $r(3) = 0.99$, $p < .05$, while a negative correlation was obtained for the scores computed with Base 2, $r(3) = -0.94$, $n.s.$

The high correlation for the scores with Base 1 indicates that the drawings by the subjects were influenced equally by the surface and structure features of this base. In other words, they did not distinguish two types of feature in their drawing. This interpretation is consistent with literatures that indicates a preponderance of mundane *literal similarity* based on both the surface and structure commonalities in human memory retrieval [6].

The low correlation for the scores with Base 2 is interesting, indicating differences between participants with regard to the ratio of the two types of similarity. That is, Participant C was weakly influenced by the surface features compared to the other participants, whereas participant A and B was weakly influenced by the

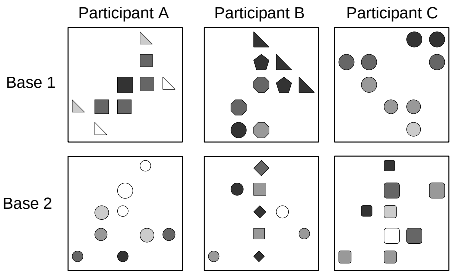


Figure 7. Works composed in the experiment.

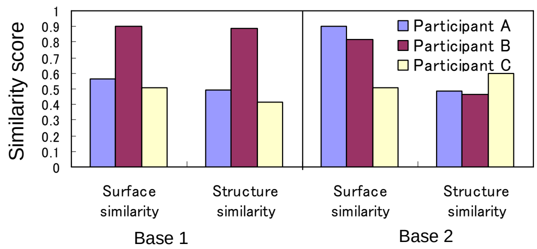


Figure 8. Similarity scores for graphics.

structure features. The result indicates the differences of their perspectives of the examples.

4.2.1 Dividing features into dimensions

To obtain more detailed information on the subjects' perspectives, we computed the scores of similarity after dividing the features into the dimensions. That is, we separately computed dot products of feature vectors in each dimension of *Location on x- and y-axis, Size, Density* or *Shape type*. Similarly we separately counted the number of elements in global maps in each dimensions of *Direction, Distance, Size difference, Density difference,* or *Shape difference*. Because of space limitations, here we present only the similarity scores obtained with Base 2, with which the lower correlation was observed.

The star plots in Figures 9 presents the scores of surface and structure similarity for each dimension. Large differences between the participants can be observed for the dimensions related to the shape of objects (*Shape type* and *Shape difference*), but the pattern differs between the two scores. Participant C obtained the lowest score of the surface similarity for this dimension, whereas Participant B obtained the lowest score of the structure similarity for this dimension.

These patterns can be interpreted by observing the graphics in Figures 4 and 7. As can be seen in Figure 4, Base 2 (Figure 4f) consists of

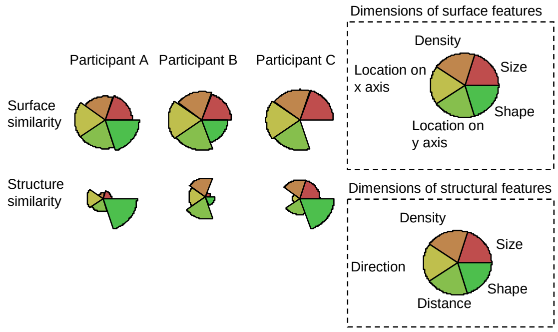


Figure 9. Star plots indicating the scores of surface and structure similarity for each dimension.

the *same-shaped objects* (the 9 circles). All the participants except Participant B used this structure feature in their graphic (see Figure 7). This is the reason why Participant B obtained the lowest score of the structure similarity. The reason for the lowest surface similarity score obtained by Participant C is in the change of the shape type. As can be seen in Figure 7, the graphic drawn by Participant C consists of the 9 rounded rectangular.

From the viewpoint of graphic composition, we consider that making the graphic with the same-shaped objects is an important structure feature. Participant C used a strategy that emphasized this feature, whereas the strategy participant B used did not. Although it is not clear which of the strategy is more creative in graphic composition, the example above demonstrates that our system is effective for extracting leaners' perspectives in this task.

5 Conclusion

In the above experiment, we successfully extracts lerners' perspectives from their works. The two types of similarity were correlated in one session, whereas they were discriminated in the other session. Further investigation revealed that this result was not caused by trivial matters of the algorithm, but was related to the important features of the drawings created by the participants. We considered that the presentations of the types of similarity would help a learner learns structure features from the example.

Our system is characterized as an application of the model of analogical reasoning. There have been many attempts to implement the analogical reasoning model in a design support system[12;

13; 14]. Most of them use the model to retrieve examples from a database. For example, in the system proposed by Forbus et al.,[12] the model receives a designer's design solution, and then retrieves examples from the database. The designer receives the matching results and inference produced by the model.

Our study can be distinguished from these studies, because our system does not retrieve examples. Instead the system computes the two scores of similarity between a presented example and a work created by a learner. The scores are used to evaluate the learner's perspectives on the example. We believe that by using these similarities, it could be possible to prompt a learner's reflective thinking.

However, this paper does not provide any evidence on the learning effect of these feedbacks. In future studies, the hypothesis must be investigated. Also, we must carefully consider the learning goal, shifting a leaner's perspectives from surface to structure features. The goal may be efficient in the elemental stages of learning; but it may be insufficient for cultivating creativity. In the later stages of learning, it might be necessary to cultivate an ability to discover a variety of structures in the example. Therefore, it is necessary for the design learning system to shift the learner's perspective not only to the biggest structure but also to the varieties of structures in an example.

References

[1] N. Chomsky. *Aspects of the theory of syntax*. the MIT Press, Cambridge, 1965.

[2] R. A. Finke, T. B. Ward, and S. M. Smith. *Creative Cognition: Theory, Research, and Applications*. The MIT Press, Cambridge, 1992.

[3] Y. Nagai and T. Taura. Formal description of concept-synthesizing process for creative design. In *Proceedings of Second International Conference on Design Computing and Cognition*, pages 443–460, 2006.

[4] J. H. Larkin and H. A. Simon. Why a diagram is (sometimes) worth 10,000 word. *Cognitive Science*, 11:65–99, 1987.

[5] D. Gentner. A theoretical framework for analogy. *Cognitive Science*, 7:155–170, 1983.

[6] K. Forbus, D. Gentner, and K. Law. Mac/fac: A model of similarity-based retrieval. *Cognitive Science*, 19:141–205, 1995.

[7] J. Morita, K. Miwa, T. Kitasaka, K. Mori, Y. Suenaga, S. Iwano, M. Ikeda, and T. Ishigaki. Chance discovery in image diagnosis: Analysis of perceptual cycles. In *Proceedings of the 1st European Workshop on Chance Discovery (ECAI 2004)*, pages 162–171, 2004.

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

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

[10] K. Forbus and D. Oblinger. Making sme greedy and pragmatic. In *Proceedings of the 12th Annual Conference of the Cognitive Science Society*, pages 61–68. Lawrence Erlbaum, 1990.

[11] J. Spencer-Smith and R. L. Goldstone. The dynamics of similarity. *Bulletin of the Japanese Cognitive Science Society*, 4:38–56, 1997.

[12] K. Forbus, P. Whalley, J. Everett, L. Ureel, J. Baher, and S. Kuehne. Cyclepad: An articulate virtual laboratory for engineering thermodynamics. *Artificial Intelligence*, 114:297–347, 1999.

[13] J. Kulinski and J. S. Gero. Constructive representation in situated analogy in design. In *Proceedings of the CAAD Futures 2001*, pages 507–520. Kluwer Academic Publishers, 2001.

[14] H. Takeda, H. Sasaki, M. Nomaguchi, Y. Yoshioka, Y. Shimomura, and T. Tomiyama. Universal abuduction studio -proposal of a design support environment for creative thinking in design-. In *Proceedings of the International Conference on Engineering Design 2003*, 2003.