Dual Deep-Structure to Associate a Shape with its Linguistic Description:  
A Hypothetical Framework and an Experimental Implementation

Haruyuki Fujii*1 and Yoshitsugu Aoki2

1 Associate Professor, Department of Architecture, Tokyo Institute of Technology, Japan  
2 Professor, Department of Architecture, Tokyo Institute of Technology, Japan

Abstract  
This paper discusses how formal language representation brings about some burdens in design support system after giving the outline of formal representation in design. Then, this paper proposes, without formal representation, a framework to represent the relation between shapes and natural language expression about the shapes, though the framework is not completed since it is a part of an on going project. Finally, an experimental implementation, which uses a connectionism model, of the proposed framework and the behavior of the implemented system are described.

Keywords: shape; linguistic description; semantic features; design; operation

1. Introduction  
A computer aided design (CAD) system helps a designer not only to manipulate on the description of an artifact but also to predict the performance and quality of the artifact. The system internally maintains information about the configuration of the artifact and modifies information in response to the manipulation of the system to change the configuration. A sophisticated system may maintain the consistency of information so as to let the configuration be realistic. The system owes formal language representation of an artifact. A language of predicate logic is usually used for the representation and some algorithms for design, such as spatial reasoning, reasoning about the domain of design, and analysis in design, have been proposed on the basis of predicate logic. However, the use of a language of predicate logic could provide burdens in design support by a computer system, too. Especially in the early stages in designing, we intuitively operate on the thing being designed. Some operations, such as an unconscious operation performed without thought using representation in a certain symbol system, are hard to be represented in formal language. Rather, designers are afraid that it loses important information that cannot be represented by language to force them to represent all information in formal language.

A designer uses natural language expressions to describe what she or he sees. Some operations on the things being designed are often expressed by natural language expression, too. For example, a designer may instruct her or his assistant to modify the shape of the things by saying, “It is better that the roof has lighter shape.” Even though a design product is required to have the concrete shape and dimension, vague expressions in natural language might be sufficient to determine the properties in some situations in designing.

We introduce a framework to represent a computational model of the visual and spatial reasoning behind such an observation in designing. Computation is just an approach to understand the visual and spatial reasoning in design. We are not claiming that the real mechanism of the reasoning in our brain and body is equivalent to the computational model. It is not easy to explain what we understand intuitively. However, if we try to convey what we understand in a scientific way, we cannot help relying on a formal method such as mathematics and logic. We admit that it is often easier for us as a designer and a planner to express what we understand spatially by showing sketches or by enumerating the fragments of natural language expressions associated with the understandings.

This paper discusses how formal language representation brings about some burdens in design support system after giving the outline of formal representation in design. Then, this paper proposes, without formal representation, a framework to represent the relation between shapes and natural language expression about the shapes. Finally, an experimental implementation, which uses a connectionism model, of the proposed framework and the behavior of the implemented system are described.

2. A Traditional Framework  
This chapter briefly reviews the outline of a traditional
framework to organizing representations in design. The framework described here does not have a particular referent but it is constructed on the basis of Mitchell (1990) and Eastman et al. (1997).

An artifact is represented from at least three aspects. The representation from each aspect is, for example, related to one another as Fig. 1 depicts. The framework contains three worlds, i.e., Shape World, Shape Representation World, and Linguistic Representation World. We closely look at the framework in the following sections.

2.1 Shape World

Shape World, or SW, focuses on the structure of an artifact. The notion of structure, here, refers to the totality of the shape and dimension of an object, which is the artifact designed or the objects composing the artifact, and the relations among the elements and the artifact. A design description, which is a product of designing, depicts the structure of an artifact in terms of graphic tokens, such as points, lines, and polygons. A group of graphic tokens is seen as a shape, which has dimension, of an object. A design description represents a state of the objects composing an artifact proposed as the result of designing.

In SW, an artifact is represented as a set of shapes and the relations among the shapes. The shapes play the role as the indices that let us associate with an artifact or some of the objects that are the parts of it. An object is usually composed of the other objects that are the parts of the artifact. Since an object is defined in terms of another object or a set of objects, if we admit that the shapes are the indices to the objects, the relations among the shapes that are equivalent to the relations among the objects would be expected to be represented.

Action in design changes the structure of an artifact. Designing is an activity to decide the structure of an artifact being capable of providing the functions and quality that are expected to be provided by the artifact (Gero, 1990). In a process of designing, the structure is synthesized or modified. The changes in a state of shapes in SW correspond to the changes by action in designing. In other words, the content of an operation on an object is represented as the changes in a state of shapes.

2.2 Shape Representation World

Shape Representation World, or SRW, focuses on both the structure of an artifact and the higher levels of information derived from the structure. It is typical to employ language of predicate logic to describe the entities and information in SRW.

A shape is represented not as a collection of graphic tokens but as a set of propositions involving the predicates related to the nature of the shape. For example, a state of shapes that there is one rectangle, which is distinguished from the other things by the name of a, is represented as follows. Where, rectangle is a predicate symbol indexed to the nature of rectangle and a is a constant symbol that indicate an individual.

\[
\text{rectangle}(a) \quad (1)
\]

In SRW, the content of action in designing is represented as the change from a set of propositions to another set of propositions. For example, the content of action where a designer changes the shape of object a from rectangular to circular could be represented as follows.

\[
\text{rectangle}(a) \rightarrow \text{circle}(a) \quad (2)
\]

It is required to define the mappings between SRW and SW so that an entity in SRW refers to a particular set of graphic tokens in SW. It is required to define the mappings between a set of operations on graphic tokens and the changes in the state of shapes in SW in order to reflect the changes in SW on the changes in SRW. It is also required to define the relations among the entities and information in SRW so as to maintain the consistency of the relations among the objects and concepts in a real world indicated by them. For example, a state of shapes, which is represented as a set of propositions shown below, cannot exist in SW or a real world.

\[
\text{rectangle}(a), \text{circle}(a) \quad (3)
\]

To avoid this inconsistency, a rule shown below is required.

\[
\text{for all } X \text{ rectangle}(X) \text{ and } \text{circle}(X) \text{ must not exist in a same set} \quad (4)
\]

2.3 Linguistic Representation World

Linguistic Representation World, or LRW, focuses on information that is hard to be directly depicted by a graphic token or shape. We talk about the functions of an artifact. We talk about the image of an artifact. We also express what we feel about an artifact, its configuration, and so on. These things are usually given as a sequence of natural language expressions. We describe, in natural language, an object not only by its structure but also by some other things with which the object is associated. For example, when we talk about the function of an artifact by a natural language expression, the referent of the expression is the consequences brought about by the artifact rather than the artifact itself.
In LRW, the meaning of a natural language expression is represented as a set of propositions, which are represented as formal language sentences, in SRW. A state that object \(a\) has the characteristic indicated by word \(p\) is represented as \(\{p(a)\}\). Again, it is required to define the mappings between LRW and SRW so that a set of propositions in LRW refers to a particular set of propositions in SRW. Through these mappings, a shape and a language expression of the image of the shape are represented formally. It is also required to define the relations among the objects and concepts in a real world indicated by them. It might be required to define the mappings between a set of operations on the changes in the state of shapes in SW and a linguistic expression of an operation in design. We often, especially in the earlier stages in designing, express how a shape should be changed in natural language. For example, we say, “Let’s make the shape of the hall bright and sharp.”

3. Burdens in Formal Representation in Design

We describe the major burdens brought about by a framework on the basis of formal language representation described above. We don’t intend to insult or blame the framework since we admit the advantages of the framework and employ the framework, i.e., Fujii (1994; 1998). We just focus on what the framework can hardly do.

3.1 Representation of Shapes and Operations on the Shapes

SRW needs a formal language that is used to represent states of shapes. A set of axioms needs to be defined explicitly to maintain the consistency of a state of shapes. A set of the contents of operations also has to be defined explicitly. The definitions require a lot of work. It is expected that the definitions be updated to meet the requirements to the expressive power of SRW.

In addition, explicit representation of shapes does not fit design thinking in the early stages in designing. In the early stage, the concepts associated with shapes rather than the shapes themselves are usually focused on. Design descriptions and sketches play the role of tags or indices that remind a designer of the concepts internally held by herself/himself. Or a relation that is not defined in the language cannot be dealt with in SRW. Suppose that a language is composed of \(N\) different sentences. The number of different states of shapes is at most the number of all proper sets of the set of sentences composing a language, i.e., \(2^N\). If the domain of designing needs to distinguish \(M\) states of shapes, the language has to consist of more than \(\log M\) sentences.

3.1.2 Description of Operations on a Design Description

The content of operations in design is represented, in SRW, as the change from a set of formal language propositions to another set of formal language propositions. A set of all propositions to describe a state of shapes has to be defined prior to the definition of the content of operations. Since it is hard to define all propositions, it is also hard to represent the contents of all operations.

It is possible to classify the contents of operations by organizing them on the basis of subsumption relation. This type of organization could be “operation ontology.” However, it is not easy task to define complete operation ontology.

3.2 Mapping between Shapes and Representation

It is hard to define one-to-one relation between elements in SW, i.e., shapes, and their representation in SRW. The mapping from SW to SRW could be one-to-many and the mapping from SRW to SW could also be one-to-many. If we were forced to define one-to-one mapping, the redundancy in representation would be lost.

3.2.1 From Shapes to Shape Depictions

One spatial relation is represented in a several ways. Suppose that a state of shapes that is composed of one rectangular and one circle and that the rectangle is to the left of the circle. This configuration can be represented in SRW as follows.

\[
\text{\{rectangle}(a), \text{circle}(b), \text{to the right of}(a, b)\} \quad (5)
\]

\[
\text{Fig. 2. A Circle is to the Right of a Rectangle}
\]

The same state of shapes is also represented as follows. This means that a configuration doesn’t have unique formal representation.

\[
\text{\{rectangle}(a), \text{circle}(b), \text{to the left of}(b, a)\} \quad (6)
\]

It is possible to express explicitly the implication that the two representations equivalent to each other by adding a proposition shown below. To do this, we need another class of proposition that is not direct translation of a configuration in SW. A set of these propositions could be called “domain ontology.” It could be done by defining sufficient amount of entities and their relations in a formal symbolic system. If we completely
understand a particular domain we could define the entities in the domain and the ontology that represent the nature of the domain. The ontology enables spatial reasoning, analysis of an artifact, and prediction of the consequence brought about by an artifact, within the range of the ontology.

\[
\text{for all } X, Y \text{ to\_the\_right\_of}(X, Y) \iff \text{to\_the\_left\_of}(Y, X) \quad (7)
\]

A mapping between a shape in SW and its formal representation in SRW is not unique. For example, the configuration shown in Fig. 3 is represented in more than two ways as follows.

\[
\{\text{rectangle} (a), \text{rectangle} (b), \text{overlap}(a, b)\} \quad (8)
\]

\[
\{\text{L\_shape}(a), \text{L\_shape} (b), \text{touch}(a, b)\} \quad (9)
\]

Fig. 3. Two Rectangles or Two L\_shape

3.2.2 From Shape Depictions to Shapes

A set of propositions in SRW does not necessarily refers to one and only one configuration. For example, the set of propositions shown below may refer to configuration (a) in Fig. 4 as well as configuration (b).

\[
\{\text{rectangle} (a), \text{circle} (b), \text{to\_the\_right\_of}(a, b)\} \quad (10)
\]

Fig. 4. What “A Circle is to the Right of A Rectangle” Refers

It is possible to let a set of propositions have one and only of referent, by adding other propositions that articulate the spatial relation in question. However, the addition of the other proposition requires extra work. We sometimes are interested only in a dominant relation among shapes. It is not sure how precise formal representation of the spatial relation should be.

3.3 Mapping between Shapes and Linguistic Expression

An interpretation function, which maps the words and phrases in LRW to the words and phrases in SRW, is required to associate the meaning of a linguistic expression with the shape referred by the expression. However, it is not easy to explicitly define the relations between SRW and LRW.

There are two reasons. One reason is that the mental state brought about by an object is expressed in LRW and the correspondence between such expression and a state of objects is vague, fuzzy, context dependent, and situated. So many studies to find the relation between a state of objects or and the linguistic expression describing the image associated with by the state have been done, but it is hard to say that a complete set of general laws about the relation has been invented. The other reason is that the world populated by what natural language can express is extremely larger than the world populated only by shapes and their relations. A linguistic expression of an artifact conveys diverse levels of abstract or concrete information even if it refers to a state of shapes in SRW. In addition, a linguistic expression refers to the other concepts related to a state of shapes.

Although a natural language phrase describes the size of a shape, the size referred to by the phrase might be vague. We would not capture how small the size of the shape expressed as “a small rectangle” is without other information. The referent of such a word is context dependent, too. Suppose that one small living room and one large bathroom are planned in a house design. Subsets of a description of the house may be represented as follows.

\[
\{\text{living\_room} (a), \text{rectangle}(a), \text{small} (a)\} \quad (11)
\]

\[
\{\text{bathroom} (b), \text{square}(b), \text{large} (b)\} \quad (12)
\]

If we focus on the size of the rooms and predicate small and large, the relation between the size of the living room and that of the bathroom may be derived as follows. However, the inferred relation, intuitively, sounds strange.

\[
\text{small}(a) \text{ and } \text{large}(b) \text{ therefore } \text{size}(a) < \text{size}(b) \quad (13)
\]

The observations mentioned above show that it is not easy to find a sufficient language, its semantics, and a set of ontology to define the domain of designing. In the next chapter, we try to give the relation between a shape and a set of linguistic expressions of the shape without giving formal language representation.

4. A Dual Deep-Structure Framework

4.1 The Framework

A framework to associate a linguistic expression about a state of shapes and the state is proposed. The proposed framework doesn’t use formal language to represent explicitly the relations between LRW and SRW. The framework tries to be free from the mappings and interpretation functions, which are required to be defined in a traditional system.
We hypothesize, following Karmiloff-Smith (1992), that a formal language representation of the relations between a state of shapes and its linguistic expression is a representation of explicit or tacit understanding about our internal mechanism that couples shapes and linguistic expression. On the basis of the hypothesis, we replace the explicit representation of the relations in formal language by other symbol systems that implicitly represent something to reproduce the relations.

Fig. 5 depicts the architecture of the proposed framework. The framework looks similar to the traditional framework shown in Fig. 1. The major difference between them is that the latter employs formal representation but the former does not use it as less as possible. Shape Feature Space, or SHFS, replaces SRW in the traditional framework. SHFS is a space that is populated by the representation of the features to determine a state of shapes. In addition, we introduce Semantic Feature Space, or SMFS, to the framework. SMFS is a space that is populated by the representation of the features of natural language expressions. In the traditional framework, SRW plays a role of coupling SW and LRW and brings about the burdens described above. SRW is regarded as a model of a linguistic representation in some sense. In the proposed framework, we decide not to let SRW play the role of a model of a linguistic representation. We do not define one and only one intermediate structure that is a model of representation in a shape world as well as that of a world described by natural language. Rather, we split the model into a group of two models loosely coupled, which we call Dual Deep-Structure. We let a linguistic expression have its own model, i.e., SMFS. We propose not to define a fixed interpretation function between SRW and LRW. Instead, we will give loose and fluid correspondence between SHFS and SMFS. The features in the both space are not represented in the form whose meaning could be interpreted intuitively.

4.2 Experimental Implementation

An experimental implementation of the proposed framework has been performed to confirm whether the framework makes sense. The details of the implementation are described in Aoki and Inage (2000). The relation between a state of shapes and a linguistic expression of the state is focused. The implemented system associates a state of shape with a linguistic expression, which is composed of adjectives - originally in Japanese language, of the image brought about by the state. The system translates a linguistic expression about how a state of shapes should be changed into the changes in the state.

4.2.1 Shape World and Shape Feature Space

There is only one object in SW. The object is a curved surface. The form of the curved surface is determined by bi-dimensional spline function on the basis of the coordinates of the \(N\) points on the curved surface (Fig. 6). If we assign and fix the value of the \(x\) and \(y\) coordinates of the \(N\) points in advance, it is a set of the values of the \(z\) coordinate of the \(N\) points that determines the form of the curved surface. We represent a set of the \(z\) coordinate values as a vector in an \(N\) dimensional space. Each element of the vector corresponds to the value of \(z\) coordinate of each of the \(N\) points on the curved surface. We call this vector and the space a shape feature vector and SHFS, respectively. Since the mapping between the forms of the curved surface and the shape feature vectors is one-to-one, i.e., a feature vector is peculiar to a form of the surface; it is fare to regard a shape feature vector as a representation of a curved surface form. Thus, every form in SW is represented in Shape Feature World. Modification of a form is represented as changes in the values of each element of a shape feature vector.
4.2.2 Linguistic Representation World and Semantic Feature Space

We suppose that Linguistic Representation World is populated by the entities corresponding to adjectives. We choose the adjectives to express images associated with a form of the curved surface. We also suppose that each entity has a value between 0 and 1. The value expresses the strength of characteristics corresponding each adjective. The higher the value of an adjective becomes, the more strongly the image is expressed by the adjective. Therefore, an image of a form is expressed as a set of the values assigned to each of the adjectives. Since some adjectives are dependent on one another, we introduce the principle factors to represent the feature of the image expressed by a set of the values. The number of the factors is at most the number of adjectives used. Let \( M \) the number of the principle factors, the feature of the image expressed by the selected adjectives is represented as a vector in \( M \) dimensional space. We call the vector and the space a linguistic feature vector and SMFS, respectively.

4.2.3 Coupling Shape Features and Semantic Features

As an example of implementation of the proposed framework, we employed a connectionism model that bridges SW and Linguistic Representation World to implement their hypothetical relations. A connectionism model is known as a model of neural network in a brain (Rich and Knight, 1991). Through this bridge, operations on a shape could be represented by changes in linguistic expression of the image of the shape. If we request our experimental system to make the current shape more stable then the system synthesizes a shape that would intuitively seem to be more stable than the current shape.

The connectionism model is trained so as to let the system operate on a shape in the described above. A three-layered back propagation model (BPM), which is one of the models for machine learning (Rich and Knight, 1991), is used. The input for the BPM is a set of values of the elements composing a semantic feature vector. The output of the BPM is a set of values of the elements composing a shape feature vector. The BPM is trained by twenty-eight pairs of the two kinds of vectors constructed from a set of empirical data. Twenty-eight shape feature vectors are constructed from twenty-eight different characteristic forms. We display each of the form of the curved surface and let some people express the image associated with the form in terms of the strength of each of the forty adjectives shown in Table 1. The strength of the aspect of the image associated with each adjective is voted by each person by the scale using the number from zero to one. On the basis of the sets of the values related to the twenty-eight forms, a set of 15 dimensional semantic feature vectors are constructed.

4.2.4 Behavior of the Implemented System

The system translates a linguistic instruction into the representation of its features in SMFS. The features of the expression related to the change from the current form of a curved surface to the expected form activate the connectionism model and let the model output the values of parameters in SHFS. The values indicate the features of the modified shape of the curved surface. The curved surface changes its form on the basis of the features.

Fig. 7 shows examples of how the form of the curved surface is transformed by a linguistic expression about the changes in a state of shapes. The upper picture shows that operation “let the current form be more stable” changes the form on the left to the form on the right through the form in the middle. Suppose that the leftmost form in the figure is shown to a designer using the system. When the form is expected to become more stable than the current form, the only thing for the designer to do is to give the system the instruction described above through the graphical user interface. The changes by operation “let the current form be lighter” are shown in the lower picture.

![Fig. 7. Examples of Linguistic Operations and the Changes in a Form](image)

5. Implications

Our experimental system demonstrates that we can operate the shape of a curved surface not by directly manipulating the graphic tokens composing the shape or their formal representation but by giving a linguistic instruction that indicates how the shape should be changed. If we define the operation to change the shape directly, we have to define explicitly how each graphic token should be modified. Even if we define the operation on SRW, we still have to write how each value of \( N \) parameters should be modified.

The experimental implementation and its behavior
have some implication. First, a set of operations to synthesize or modify a state of shapes is represented in less hard way than a traditional framework. We don’t have to define every correspondence between a collection of graphic tokens and an object. Second, a linguistic expression, which does not directly refer to a shape, a relation among shapes, or the dimension of a shape, is expected to be dealt with.

Further, it might be expected that the ontology of a domain be expressed in less hard way. It might be expected that spatial constraints be dealt with in a similar way. In a traditional system, reasoning concerning ontology and reasoning about spatial constraints relies on logic programming. If we find a good mechanism, which couples SHFS and SMFS and facilitates the transformation between representation and its features, the issues concerning ontology and spatial constraints are approached without logic programming. If we find a good mechanism, which couples SHFS and SMFS and facilitates the transformation between representation and its features, the issues concerning ontology and spatial constraints are approached without logic programming that requires many mappings. Of course, there is a trade-off between a traditional framework and the proposed framework. In the traditional framework, once the ontology that is sufficient to represent a certain domain, most of all operations and reasoning in the domain would be executed on a CAD system. However, if the ontology were not updated promptly in response to the changes in the domain, it would become out of date. It cannot be predicted what the ontology does not derive. On the other hand, the relations among shapes and linguistic expressions reproduced in the proposed system might not be as precise as the explicit representation in formal language. However, the relations are updated on the fly when a new relation between a state of shapes and a linguistic expression is discovered.

6. Conclusion
A framework to bridge shapes and its linguistic expression is proposed. The architecture and the behavior of an on-going experimental implementation of the proposed framework are given.

Acknowledgments
The authors express thank to Mr. Makoto Inage, who was a graduate student of Tokyo Institute of Technology, for his help on implementation of the experimental system and some experiments with the system. The base of this study is presented at the international workshop on Visual and Spatial Reasoning in Design II held in July 2001 in Italy. This paper is the enhanced version of the study.

A part of this particular research is partially funded by the Grants-in-Aid for Scientific Research JSPS, classified as the Priority Areas Research C (#13650667).

References
1) Aoki, Y. and Inage, M. (2000) Linguistic Operation System for Design of Architectural Form. In Design and Decision Support System in Architecture (pp.15-24). The Netherlands.
2) Eastman, C., Jeng, T. S. Chowdury, R. and Jacobsen, K. (1997) Integration of Design Applications with Building Models. In R. Junge (ed.), Computer Aided Architectural Design Futures 1997 (pp.45-59). Kluwer Academic Publishers, Dordrecht, The Netherlands.
3) Fujii, H. (1994) Understanding Spatial Descriptions, Master Project Report. Department of Philosophy, Carnegie Mellon, University, Pittsburgh, USA.
4) Fujii, H. (1998) A Framework of Multilingual Representation of a Design World on the basis of Philosophy of Language. In J. Gero et al. (eds.), Artificial Intelligence in Design ’98 (pp.385-404). Kluwer Academic Publishers, Dordrecht, The Netherlands.
5) Gero, J. S. (1990) Design Prototypes - A Knowledge Representation Schema for Design. AI Magazine, 11(4), 26-36.
6) Karmiloff-Smith, A. (1992) Beyond Modularity - A Developmental Perspective on Cognitive Science. The MIT Press, Cambridge, USA.
7) Mitchell, W. J. (1990) The Logic of Architecture. The MIT Press, Cambridge, USA.
8) Rich, E. and Knight, K. (1991) Artificial Intelligence - second edition, McGraw-Hill, New York, USA.
9) Yoon, K. B. (1992) A Constraint Model of Space Planning. Computational Mechanics Publications, Southampton, UK.