Incorporation of Natural Language Processing and a Generative System - an Interactive System that Constructs Topological Models from Spatial Descriptions in Natural Language

Haruyuki Fujii

Izumi Research Institute
Institute of Technology
Shimizu Corporation
JAPAN

The natural language processing technique and the spatial reasoning technique are incorporated to create a computational model representing the process of updating and maintaining the knowledge about spatial relations. An algorithm for the spatial reasoning is proposed. An interactive system that understands sentences describing spatial relations is implemented. The system determines the reference of an anaphoric or deictic expression from the literal meaning of the input and the implicit meaning derived from the literal meaning. The consistency of the spatial relations is maintained. The correct topological representations of the spatial relations are generated from well-formed descriptions.

Keywords: natural language processing, discourse analysis, artificial intelligence, architecture, CAD

1 Introduction

The architectural design involves more than one participants. For the purpose of collaboration, they communicate with each other to exchange the information embedded in a design and have common information to some degree. The communication is mediated by a diversity of the representation formalism of the information: drawings, sketches, diagrams, natural language, etc. Recent development of the CAAD systems and electronic devices enable the participants to exchange the information through computer networks. A lot of representation formalisms containing as much information have been introduced. The formalisms are precise enough to represent a design in a certain level of abstraction, but they are cumbersome to manipulate. On the other hand, natural language is easy to handle for human and can represent the multiple levels of abstraction, even though it often generates vague or ambiguous expressions.

We focus on the expressive power of natural language as a common language to talk about an architectural design between human and a computer. If we find the mapping between a verbal expression describing an architectural design and the properties of the design, the expression becomes another form of the representation of an architectural design. One of the objectives of this work is to represent a process of interpreting natural language descriptions of an architectural design. This work offers the potential of developing a natural language user interface for CAAD systems to update and maintain the information concerning an architectural design. I have created a computational model of the process of interpreting spatial descriptions based on the achievements of the linguistic research, and have implemented an interactive system acquiring the knowledge about the planning design of an apartment from English sentences. The system takes a sequence of sentences describing the planning design of an apartment. It incrementally constructs the model(s) of the layout. The information handled by the current system is limited to the topological representation of the spatial relation among rectangular rooms in a rectangular apartment.

This paper is organized into seven sections. Section 2 reviews the related works. Section 3 describes the problems that is going to be solved, in this paper, to implement a system understanding a spatial description. Section 4 explains the reasoning algorithms and heuristics to solve the problems. Section 5 explains the system architecture, functionalities, and behavior of the system that understands a spatial description of the planning layout of an apartment. The paper ends with the further direction of this work and a brief summary.

2 Background and motivation

The mapping between the natural language expression and the image of an architectural design is not straightforward. This implies that it is hard to understand what a discourse says only by analyzing the syntactic structure of each sentence in the discourse. Since the mapping is determined by the interaction between the context of the expression and the knowledge about the things involved in the expression, the incorporation of a natural language processing system and a knowledge-based CAAD system provides a starting-point for the development of a system that acquires and maintains knowledge incrementally from a sequence of sentences.

Interactive systems that acquire and maintain the knowledge about a particular domain have been introduced. Cullingford developed Script Applier Mechanism (SAM), an interactive system that represents the world described in a story (Schank 1981). SAM parses natural language texts about a story, understands the contents, and answers questions about the story. The formalism based on Schanks Conceptual Dependency (CD) theory is incorporated with the classification of primitive action to represent what the story means. CD is an attractive theory on which we can rely when we focus on the representation of action and the transition of a world by the action. Winograd (1973) developed SHURDLU, an interactive system that maintains the knowledge about the world consisting of blocks. SHURDLU understands a natural language sentence and responses it. It moves the blocks in the world when it is ordered to do so, answers the questions about the action that the system made as well as the world. Akin designed Architectural Inference Maker (AIM) to perform the building-puzzle task (Akin 1986). The system samples information about the general properties of an object in question by asking the user questions, and identifies the object. The inference mechanism contains the rewriting rules to acquire the higher level of specific information concerning architecture from the lower level of general information.

Pereira & Pollack (1991) focused on the influence of context on interpretation of natural language sentences. They introduced the representation formalism for incremental interpretation of a discourse. Their idea has been implemented in Candide, a system for interactive acquisition of procedural knowledge. There are two aspects of incremental interpretation of a discourse. The combinatorial aspect makes explicit the interpretation alternatives that arise from the interaction between phrase interpretations and contextual information. The inferential aspect concerns the means by which a language interpreter chooses amongst combinatorial alternatives. The domain knowledge about architectural space is essential for the understanding of what a spatial description means. Mani & Johnson-Laird (1982) hypothesized the existence of a mental model consisting of richer information than a set of propositions corresponding to each sentence in the descriptions. They acquired experimental data supporting their hypothesis that the mental model representation constructed by interpreting a spatial description is not solely propositional presentation. Byrne & Johnson-Laird (1989) have confirmed the hypothesis that a Mental Model constructed form a spatial description contains richer information than a set of the propositions corresponding to each sentence in the spatial description. They verified that people are capable of making two-dimensional spatial inferences. As a set of formalisms to represent an architectural design, Mitchell (1990) introduced the conceptions of the design world" and "the critical language" based on the model-theoretic semantics. The design world is a conceptual world depicting the properties of an architectural design. The critical language is a formal representation, first order logic, of natural language descriptions concerning the properties of the architectural space. The design world is some sort of language or a symbolic data structure consisting of graphic tokens such as points, lines, and polygons, the other tokens representing the properties of the graphic tokens, relations among the graphic tokens, functions mapping the graphic tokens onto the other tokens, axioms defining necessary relationships between shapes (collections of primitives) within the design world, and operators transforming a state in the design world to another state. The semantics of the critical language is defined with respect to the design world. A truth value assignment between the critical language and the design world defines an interpretation of the critical language. These works imply that a model-based formalism plays an important role in reasoning of the spatial relations in an architectural design.

3 Features of a spatial description

For the present, it may be useful to look closely at some of important features of a spatial description. There are two features to be taken into consideration to make a computational model of a spatial description understanding. One is that the spatial description is context-dependent, the other is that the expression often contains inconsistent, implicit, and/or incomplete descriptions. The first feature is articulated by two questions, Q1 and Q2, while the second is articulated by the other two questions, Q3 and Q4.

(Q1) How to resolve the reference of anaphoric expressions?
(Q2) How to resolve the reference of projective prepositions?
(Q3) How to detect and eliminate inconsistent information?
(Q4) How to understand implicit and/or incomplete descriptions?

We introduce two spatial descriptions containing the features described above. The first discourse (G1-G6) is describing the layout of an apartment from a point of view that is outside of the apartment, while the second discourse (W1-W5) is given from a point of view that is inside of the apartment. Suppose that the descriptions are given in the context of explaining the layout of an apartment. The discourses are ones of the examples that our system can handle.

(G1) The hall faces south.
(G2) The bedroom is on the west of the hall.
(G3) Another bedroom is on the west of the hall.
(G4) The kitchen is on the north of the hall.
(G5) The bedroom is on the south of the (other) bedroom.
(G6) The kitchen is on the north of the bedroom.

An addressee can easily interpret GI. Even though there is no object to which "the hall" is referring, one believes that the sentence presupposes the existence of the reference of the hall from the context. G2, G3, and G4 are also easy enough to understand, except that the addressee doesn't know the relation between the two bedrooms. At the time when G5 is given, the addressee is somewhat confused because the references of the two uses of the bedroom' are not explicit. Yet, it is possible to maintain a model(s) of the layout. It is hard to understand G6 based on only the literal meaning of the previous discourse since we are confused by the reference of "the bedroom'. However, if the addressee has constructed the model(s) that represents the layout of an apartment described by the discourse, one can infer that there is a unique bedroom to which 'the bedroom refers. The reference is the bedroom on the north of the other bedroom. To understand the discourse, the addressee assumes that every sentence in the discourse should be true and that the ambiguous sentence is interpreted from the implicit meaning of the discourse as well as the explicit meaning literally expressed. The next description explains a part of the same layout from a different point of view. These are also examples that our system can handle.

(W1) The hail opens south.
(W2) You enter the hall.
(W3) The bedroom is to the left.
(W4) You enter the bedroom.
(W5) Another bedroom is to the right.

The view point is moving as like the observer is walking through the apartment. If the addressee assumes that the addresser is leading a walking tour, the addressee can understand W3, even though the reference object of the projective projection is not mentioned and the word "1eft" is relative to the perspective of the participants. Grice's maxims of conversation (Grice 1975) specify what participants have to do to converse in a maximally effect, rational, cooperative way (Levinson 1983). Some maxims relevant to our project are the following.

The maxim of quality says that participants should try to make their conversation one that is true, specifically, not to say what they believe to be false.

The maxim of quantity says that participants should make their contribution as informative as is required for the current purpose of exchange.

These maxims encourage us to make the assumptions as follows.

(A1) Every sentence should be true unless the addresser misunderstand the contents.

(A2) Every sentence should assert the new information. It isn't redundant.

4 Knowledge-based interpretation of a spatial description

We propose to incorporate knowledge-based reasoning systems with a method for natural language processing to answer the questions shown above. We represent the knowledge used to interpret a spatial description in some forms, namely, the literal context, the taxonomic knowledge, the model, and the observing agent. The information acquired from a discourse is kept in the form of a model(s) constructed from a discourse as well as a set of propositions corresponding to each sentence in the discourse. The propositions are used to interpret and represent the literal meaning of natural language sentences from the combinatorial aspect and stored as so called the literal context. The model(s) is used to interpret and represent the sentence from the inferential aspect. The observing agent represents the knowledge concerning the state of the effective point of observation. The effective point of observation is a point where the observation of the spatial relations is made. This is essential for the interpretation of a prepositional phrase (PP). Incomplete and/or implicit description is complemented based on the knowledge about spatial relations and the taxonomic knowledge about architecture. Consistency of the spatial descriptions is maintained with respect to the information represented in each form. When all of the knowledge fail to give the unique interpretation of a sentence, the interaction between system and the user helps the system to interpret an ambiguous or vague expression.

4.1 Literal context

The literal context plays an important role especially in the noun phrase instantiation. The enumeration of possible instances of a noun phrase (NP) is accomplished by the interaction of the semantics of the phrase and the context. The verification of the instances in a whole sentence is accomplished with respect to the literal meanings of the context. Since we have assumed that every sentence in a spatial description should be true, the entity to which a definite NP refers should exist even if it has not been introduced in the previous discourse of spatial descriptions. If a definite noun phrase appears and if there is no reference of the definite NP in the previous context, the noun phrase presupposes the existence of the reference. An indefinite NP always introduces a new instance in the domain of spatial descriptions. An indefinite determiner is usually treated as an existential quantifier, but, our algorithm generates another instance and assign it to the current indefinite NP since we have assumed that every sentence is not redundant but informative under the purpose of the conversation. A pronoun refers to any instance compatible, with preference given to more recent occurrences. A proper name refers to its unique referent so that the proper name can be used as a symbol of an instance. The algorithms to compute these are shown in Figure 1. We suppose that the literal context consists of Immediate Context, Current Context, and Global Context (Pereira and Pollack 1991). Current Context contains the information on the previous sentence, Immediate Context contains the information on the currently processed sentence, and Global Context contains the explicit information about a whole discourse.

4.2 Taxonomic knowledge

The taxonomic knowledge about the domain identifies the instance of a vague expression when the context is not informative enough to determine the meaning of a sentence. The taxonomy of nouns involved in a spatial description classifies an NP and tries to find the instance whose class is subsumed by the class of the NP. When there exist two or more entities that instantiate an NP based on the literal context, the entities are filtered by using the taxonomic knowledge of the domain. For example, in a context that contains 'bedroom 1" and "closet 2", the NP "the room" may refer to one of "bedroom-l" when the class "bedroom" is subsumed by the class "room", but the class "closet" isn't.

[Definite NP Instantiation]

Step-1 Look at the Current Context.

Step-2 Find all instances of the class that is either the class of the noun phrase or one of subclasses of it.

Step-2a If at least one instance is found then RETURN a list of the instances and store the list in the immediate context else look at the Global Context and: .

Step-3 Find all instances of the class that is either the class of the noun phrase or one of subclasses of it.

Step-3a If at least one instance is found then RETURN a list of the instances and store the list in th immediate context else:

Step-4 Generate an instance whose class is the same as that of the noun and RETURN it and store it in the immediate context.

[Indefinite NP Instantiation]

Step-1 Generate an instance whose class is the same as that of the noun and RETURN it and store it in the immediate context.

[Pronoun Instantiation]

Step-1 Look at the Current Context.

Step-2 Look for the instance whose class is subsumed by the class of the pronoun.

Step-2a If at least one instance is found then RETURN a list of the instances and store the list in the immediate context else look at the Global Context and:

Step-3 Look for the instance whose class is subsumed by the class of the pronoun.

Step-3a If at least one instance is found then RETURN a list of the instances and store the list in th immediate context else RETURN an empty list indicating that there is no reference of the pronoun in the context.

[Propername Instantiation]

Step-1 Return its referent.

Figure 1; Algorithms for Noun Phrase Instantiation

4.3 Model-based reasoning

In the case that the meaning of a sentence is not determined by either the literal meaning of the context or the taxonomy of the nouns involved in a spatial description, the next maneuver is the model-based reasoning. The verification of the instances in terms of consistency of the interpretation containing the instance is accomplished with respect to the knowledge acquired from the context. The idea is to update a model(s) which renders every sentences in a spatial description is true and find the instances of a spatial expression in the mode (s). When there exist two or more entities that instantiate a spatial expression even though the taxonomic knowledge of the domain eliminated some interpretations, each interpretation is verified whether it involves one of the entities can be rendered true with respect to the models updated from the spatial description. A reasoning algorithm that maintains the model(s) representing spatial relations described in a spatial description is essential to this approach. We introduce the algorithm implemented in our system. We are inspired by the seven functions for maintaining a mental model of discourse (Johnson-Laird 1988) . The major difference between Johnson-Laird's functions and our algorithm is that this algorithm does not aim at representing the mental process while his functions are introduced as a representation of the mental process to understand spatial descriptions. In addition, it is assumed, in his functions, that the unique interpretation of a sentence is already given so that the functions does not care how a sentence such that it has more than one interpretations is interpreted. Our algorithm can handle the sentence that involves more than one interpretation.

Step-0 of the algorithm initializes the model. When a new discourse begins, Step-i updates the knowledge based on the information concerning an event, while Step-3 through 6 maintain the spatial relation model(s) based on the spatial descriptions. Step-3 and 4 are the filter that passes the sentence that conveys new information with respect to the literal context. Step-5 and 6 verify whether the current sentence can update the model(s) constructed from the previous context.

Step-0 Set the model and the state of the domain to neutral in the sense that there exist only constraints that are intrinsic to the domain of discourse.

Step-1 **For** a sentence that describes the event in the domain, **if** the event can occur in the current state of the domain **then** REVISE the state by DELETING a subset of the current state changed by the event and ADDING the result of the event **else** do nothing.

Step-2 **For** a sentence that describes a certain thing concerning the model of the domain, ENUMERATE all interpretations of the sentence with respect to the references of each constituent of the sentence and make a list of the interpretations.

Step-3 **For** each interpretation in the list created by Step-2, VERIFY whether the truth conditions of the interpretation are satisfied by a set of propositions collected from the previous sentences in the discourse, and:

- **Step-3a** **If** the interpretation is true with respect to the set of the propositions **then** MARK the interpretation with redundant.
- **Step-3b** **If** the interpretation is false with respect to the set of the propositions **then** MARK the interpretation with inconsistent.

Step-4 ELIMINATE all interpretations marked with redundant or inconsistent in Step-3 from the list of interpretations. **If** the list is empty **then** go to Step-1 and wait for the next sentence.

Step-5 **For** each model in the set of the models that renders the previous discourse true, **For** each interpretation remaining in the list after Step-4, VERIFY whether the truth condition of the interpretation is satisfied by the model, and:

- **Step-5a** **If** the model is independent of the truth condition of the interpretation **then** USE the interpretation to add to the set of propositions the properties or relations expressed by the interpretation.
- **Step-5b** **If** the interpretation is not false in the model, i.e., true or uncertain **then** REVISE the model, if possible, to construct the models that are consistent with the previous discourse and render the interpretation true. **If** this task is possible **then** the interpretation is a valid deduction from the previous discourse or conveys new information: store the revised models. If the task is impossible **then** ELIMINATE the model.
- **Step-5c** **If** the sentence is false in the model **then** ELIMINATE the model.
- **Step-5d** **If** all models are eliminated during Step-5 **then** MARK the interpretation with inconsistent.

Step-6 ELIMINATE the interpretation marked with inconsistent in Step-4, and:

- **Step-6a** **If** the list of the interpretations is empty **then** MARK the sentence with invalid: restore the models that render the previous discourse true.
- **Step-6b** **Else if** the list of the interpretations contains exactly one interpretation **then** MARK the sentence and the interpretation with valid: store the proposition that is corresponding to the interpretation.
- **Step-6c** **Else if** the list of the interpretations contains at least two interpretations **then** ask the utterer of the sentence which interpretation is intended. **For** the interpretation that is intended to be true do **then** MARK the sentence and the interpretation with valid: store the proposition that is corresponding to the interpretation. **If** every interpretation is not intended to be false by the utterer **then** MARK the sentence with invalid: restore the models that render the previous discourse true.

Step-7 Go to Step-1 and wait for the next sentence.

Figure 2: Algorithm for Natural Language Processing Incorporated with Model -based Reasoning

4.4 Observing agent

It is efficient to keep track of the location and perspective of the effective point of observation in the discourse to instantiate projective prepositions. The effective point of observation is a point where the spatial relations are observed. When a spatial description is given from a point of view that is outside of the space described in a discourse like G1G6, the viewpoint is set to the point of someone who is looking at an imaginary drawing of the layout of the rooms in the discourse. When a spatial description is given from a point of view that is inside of the space described in the discourse like W1-W5, the view point moves (or stays) corresponding to each sentence in the discourse and a projective preposition such as "at the back' has different instances dependent upon the view point. The mapping between the projective prepositions and the reference is determined by two types of situations among the object which is described mainly, the reference object and the observer (Herskovits 1986) . The coincidence situation is the situation that the location of the reference object and that of the observer are the same. The encounter situation is the situation that the location of the reference and that of the observer are different from each other and that the reference object and the observer are facing each other. In a coincidence situation, the direction to which a projective preposition refers is the same to the direction of the vector from the location of the observer. In an encounter situation, the direction of FRONT and BACK is the other way around. Figure 3 shows the mapping.

4.5 Human-computer interaction

When all of the knowledge representations mentioned above fail to give the unique interpretation of a sentence, the interaction between the addressee and the addresser happens to interpret an ambiguous or vague expression. The questions to confirm the information are used.

5 A computational system of decoding

We look at the features of the computational system implemented based on the maneuvers mentioned above. The system understands a spatial description consisting of projective prepositions and constructs the model(s) representing the knowledge acquired from the description. We say that the system "understands a spatial description in the sense that it

(1) parses an input sentence,
(2) verifies the consistency with respect to the literal context as well as a spatial model(s),
(3) reasons about the implicit spatial relations, and
(4) responds to the input sentences.

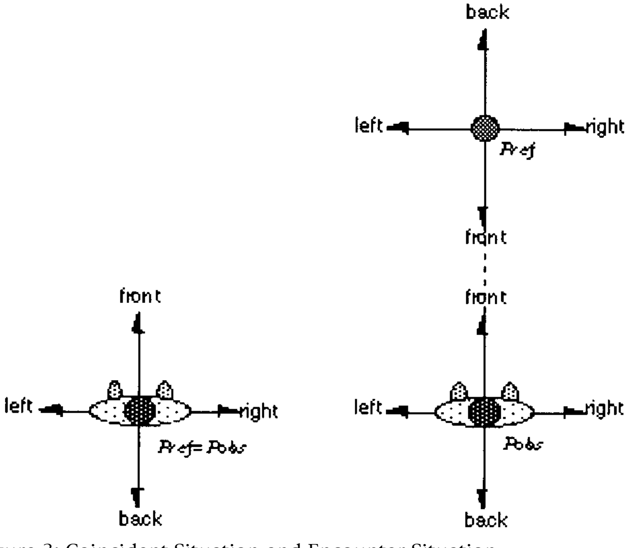

Figure 3: Coincident Situation and Encounter Situation

A spatial description given from a point of view of both the inside and the outside of the space in question is understood. The current system is limited to the domain of the spatial relations of rectangular rooms in a residential house.

The system provides the following four functionalities.

(1) It facilitates the interaction with the user to understand the meaning and reference of an ambiguous or vague sentence. It generates a question to confirm what the user refers to in the case when the system comes across an ambiguous or vague sentence. It also generates a question to eliminate the inconsistent information as well as to narrow the search space for the spatial relations.

(2) The system resolves the reference of an anaphoric or deictic expression by keeping track of the effective point of observation and its perspective as well as by referring to the model(s) constructed previously. A virtual agent called 'observer" walks through the model(s).

(3) The system assembles the consistent knowledge about the apartment in question and constructs the model(s) of the layout. It enumerates every possibility of updating the current model(s) that is consistent with new information. Consistency of the knowledge is maintained by incorporating model-theoretic semantics with a generative grammar of planning design.

(4) The system complements an incomplete or implicit expression by referring to the domain knowledge concerning architecture and the model(s). The system contains a subsystem of knowledge representation using "frame" to update and maintain the taxonomic

information concerning the objects in the discourse. The information that is not explicitly provided is inferred from the frame and the model(s).

5.1 System architecture

The system consists of five modules, namely, Incremental Interpreter (II), Spatial Reasoner (SR), View Point Tracer (VPT), Human-Computer Interface (HCI), and Knowledge-Base (KB). Figure 4 shows the architecture of the system. HCI takes a natural language sentence from a console and sends it as the sequence of the words to II. II parses and interprets the sequence of the words incrementally. If II succeeds in the parsing and interpretation, it returns the formal representation of the interpretation. Otherwise, it sends the error code indicating that the sentence is not accepted either syntactically or semantically to HCI. HCI shows the natural language sentence corresponding to the error code on an console. The formal representation of the interpretation is sent to SR in the case that the sentence is a spatial description and to VPT in the case that the sentence shows the action of the observer of the space, respectively. SR updates the model(s) which has been constructed from the previous discourse. VPT updates the effective point of observation. When the updates are done, HCI asks the user to input the next sentence.

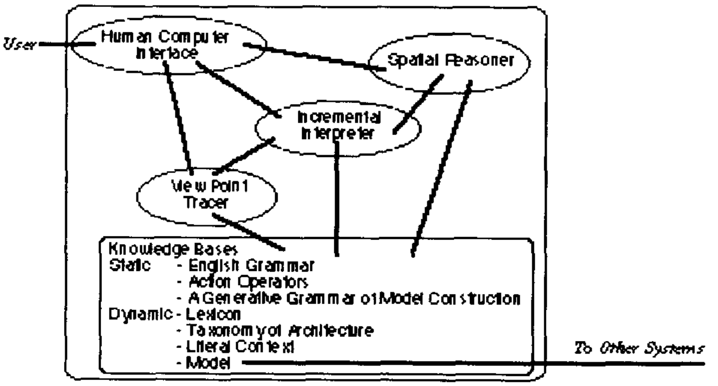

Figure 4: System Architecture

II understands anaphoric expression and transition of the effective point of observation and its perspective. It parses an input sentence and translates it into the semantic representation. It verifies whether the sentence has inconsistent or redundant interpretation with respect to the literal meanings of the previous sentences in the same discourse. If the sentence is inconsistent or redundant interpretations, it is not accepted. II accesses the taxonomic knowledge representing the relation among rooms which are consistent with the previous sentences. The new instance is stored as an instance of the most specific class in which the instance can be classified. Figure 5 shows a part of the syntax of a natural language sentence implemented in II. Where, S stands for a sentence and its syntactic structures are derived from the rewriting rules. Copular is rewritten by the terminal symbol "is". P, V, Det, N, ProNoun, and ProperName are rewritten by a preposition, a verb, a determiner, a noun, a pronoun, and a proper name, respectively. The grammar is represented as a definite clause grammar (DCG).

S -> NP Copular PP | PP Copular NP | [There] Copular NP PP | PP [There] Copular NP | NP VP
PP -> P NP
VP -> V | V NP
NP -> Det N | ProNoun | ProperName
N -> N PP

Figure 5: Syntactic Structure of Input Sentences

The semantic representation for II is an enhanced version of the formalism for Candide system (Pereira & Pollack 1991) . The representation mainly concerns the combinatorial aspects of incremental interpretation. The semantics of a phrase is represented by the conditional interpretation formalism that consists of the type, assumptions, and sense of the phrase. The sense is the first order logical representation of a sentence. It participates compositionally in the interpretation of larger phrases. The assumptions represent the constraints on how the sense is connected to its context. They contain the information about the restriction for the instantiation of the object. When a sentence is vague with respect to NP instantiation, the assumptions keep the information to narrow down the candidates. If the NP instantiation has not been done, the argument position corresponding to the NP is filled with a variable bounded by the restrictions in the assumptions. The type represents a type of constituent of a sentence. It is composed from the basic types , 'Bool', "Place", and "Action". The type system is added to the original formalism to distinguish different kinds of sentences from each other. In particular, it is used to distinguish a sentence asserting the information about the apartment in question from the other types of sentences, such as a sentence describing action. Figure 6 shows how the sentence "The kitchen is on the south of the bedroom." is translated into the semantic representation. The syntactic category, the type, the assumptions, and the sense are represented in each node of the tree.

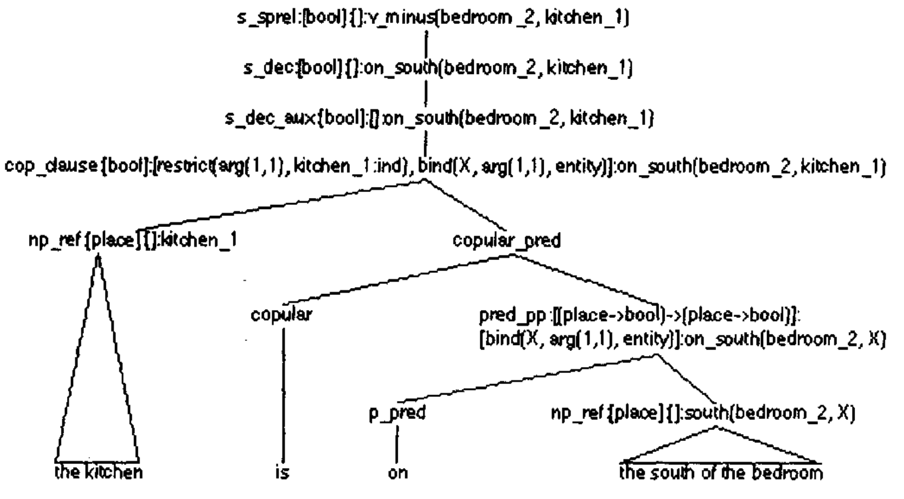

Figure 6: A Parse Tree of a Copular Clause

SR takes the interpretation(s) of an input sentence, verifies whether it is consistent, and updates the model(s) when the interpretation is consistent. The module assembles the topologically consistent information about an apartment layout planning and constructs the model(s) of the layout. It constructs a model(s) from the interpretation of each sentence that is consistent with respect to the current model(s) constructed from the previous context. When the literal context gives more than one interpretation in the sense that there are two or more entities that instantiate the NP, SR verifies whether each of the interpretations can update the current model(s) or not. If the interpretation can update the model(s), it is considered to be a consistent interpretation. When there are more than one consistent interpretations, SR sends HCI the error code to let the system ask the user which interpretation is meant. If no interpretation is consistent, SR does not update the model(s).Every model is represented as a graph whose vertices represent a space and whose edges represent the wall between two spaces. Not only the edges that are literally mentioned but also the edges inferred from the discourse construct a model(s). The status of each edge is also stored to distinguish literally mentioned edges and inferred edges. The inference rules in current SR are an enhanced version of the generative rules introduced by Flemming (1978). Figure 7 shows a graphical representation one of spatial model(s) constructed from the following discourse. 'The hall faces south. The bedroom is to the west of the hail. The kitchen is to the west of the hail. The bathroom is to the west of the hall. The living room is to the north of the hail." The solid edges represent the relations literally mentioned or the default relations, while the hatched edged represent the relations inferred by the Spatial Reasoner.

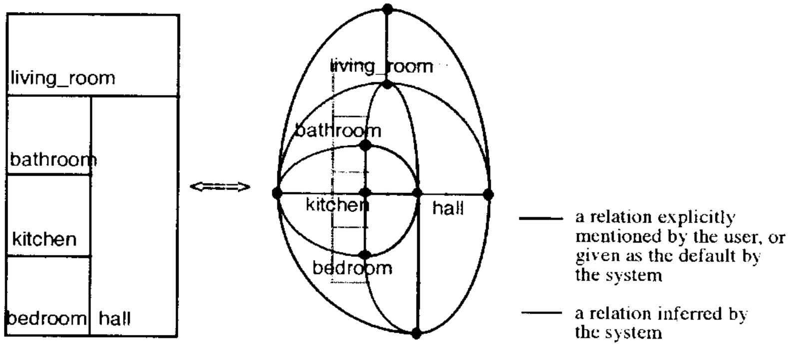

Figure 7 Graph Representation of a Planning Layout

The observing agent is implemented as VPT. VPT keeps track of the state of the world consisting of the model(s) and the observing agent and maintains the location and the perspective of the effective point of observation of the observing agent. The module is triggered by the interpretation referring to an action in an imaginary walking tour. An operator changing the state is instantiated corresponding to the interpretation of the input sentence. This module is important to interpret a sentence given from a point of view walking through the inside of an apartment. For example, the location of the picture mentioned in the sentence 'At the back, there is a picture of mountains." depends on the location and the perspective of the observer, as described in Figure 8. The system gets the information about where the picture is when it knows the location and the perspective or when it gets more information that specifies the location of the picture.

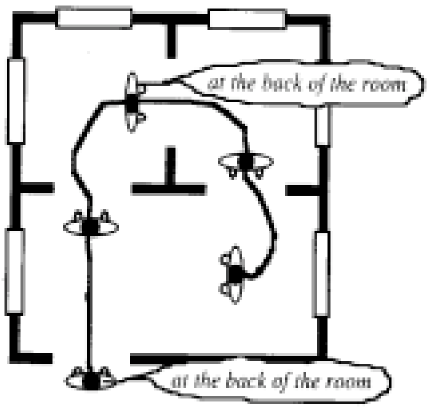

Figure 8: Referents of Prepositional Phrases

VPT can interpret a sentence containing a projective preposition lacking the reference object if the effective point of observation is known. For example, "to the right" in the sentences 'You enter the hall. The bedroom is to the right.' is understood as "to the right of the hall (the place where you are)". "Right' is not explicit unless the orientation of the perspective of 'you' becomes explicit, i.e., when we know that you are facing north, "right" refers to west. HCI enables the system and the user to interact with each other. The module has two functionalities. One is the conversion of the sentences inputted by the user. When the user types an English sentence, HCI converts the sentence into a list of words. The other functionality is the generation of English sentences. HCI generates a question to confirm what the user refers to if the system comes across an ambiguous or vague sentence. It also generates a question to eliminate inconsistent sentences with regard to the spatial relations in the model(s) as well as to narrow the search space for the spatial relations. The sentences are generated from the feedback information from the modules in the system by using the templates. KB maintains two types of knowledge, namely the static knowledge and the dynamic knowledge. The knowledge about the grammar, a taxonomy of architecture, rules to keep track of the view point, rules to construct and maintain the spatial models, and templates for the response from the system belong to the static knowledge. The knowledge about the lexicon, the instances in a discourse, the literal meaning of each sentence, the effective point of observation, and the model(s) being updated through the discourse belong to the dynamic knowledge.

5.2 System behavior

The system behavior is shown in Figures 9 and 10. It is an interaction between the system and the user. It describes the process that the system incrementally acquires the information about the planning layout of an apartment in question. The system is evaluated with respect to the following criteria.

(1) Is the consistency of a discourse maintained with respect to the spatial relations? Yes. The system verifies whether an input sentence is consistent, inconsistent, or redundant with respect to the literal meanings as well as spatial relations inferred from the literal meanings.

(2) Does the interface between the natural language processing unit and the knowledge representation concerning the spatial relations in an apartment layout work well? The answer is "yes" in the sense that at least the system incorporated a natural language processing system with a model-based reasoning system. The system demonstrates that the consistent interpretation of a spatial description is maintained by referring to the knowledge basis. The current interface only accepts the spatial relations expressed by using projective prepositional phrases. The rich information included in 'on', "in", "at", and "to" is not handled with the current system.

(3) Is a correct spatial model constructed from well-formed spatial descriptions? Yes. The system can construct a correct spatial model described by a discourse consisting of well-formed spatial descriptions. The system acquires the consistent information about the spatial relations and construct the model(s) from a spatial description, even though the spatial relations are limited to those among rectangular rooms in a rectangular apartment.

6 Future direction

Concerning to the current system, it is the first problem to design the interface for the generation of natural language from the knowledge base about architecture and spatial relations. The current system can translate natural language sentences into the spatial representations. The new task is to translate the information that the spatial representations have into natural language. It might be possible to represent the domain knowledge about architecture based on the model-theoretic semantics by Barwise and Perry (Barwise & Perry 1983; Barwise 1989), and to implement the semantics with the typed feature structures (Carpenter 1992) . The idea is to give each constituent in natural language the features concerning space and action, and to interpret a sentence incrementally by unifying the sentence and the previous context. Concerning this idea, we might want to verify whether our spatial reasoning method can be reduced into the unification or the integration of the unification and complementary operations. Intuitively, the process in which our spatial model(s) is constructed is similar to the process in which the information is specified by the unification of the partial information. It might be not so difficult to let our model acquire the higher knowledge from the literal meanings by using the typed feature structures. The crucial problem is how to represent the reasoning process to acquire the implicit information from the explicit partial information whose functionality is implemented in current SR. The current system accepts only consistent interpretations, but the system can be refined to correct inconsistent interpretations. The users do not always address the correct sentence. Their belief sometimes makes them generate a incorrect scene. The unification of the typed feature structures might enable us to check the consistency of a discourse and detect the information that is inconsistent with other part of the information. To make the system practical with respect to the representation of architectural space, the higher knowledge and the properties other than spatial relations have to be represented in the system.

```
YOUR INPUT>> the hall opens south
  RESPONSE>>  I understand the input sentence as hall_1 faces south
  RESPONSE>>  There is one model for the relation if it is true that hall_1 faces south .
  RESPONSE>>  I renewed the model.
YOUR INPUT>> you enter the hall
  RESPONSE>>  I am entering hall_1 from south and facing north .
  RESPONSE>>  I entered hall_1 .
  RESPONSE>>  This information is not used to renew the model.
YOUR INPUT>> to the left is the bedroom
  RESPONSE>>  I understand the input sentence as bedroom_2 is to the west of hall_1
  RESPONSE>>  There is one model for the relation if it is true that hall_1 is to the east
              of bedroom_2 .
  RESPONSE>>  I renewed the model.
YOUR INPUT>> the kitchen is to the left
  RESPONSE>>  I understand the input sentence as kitchen_3 is to the west of hall_1
  RESPONSE>>  There are 2 kinds of models for the relation if it is true that hall_1 is to
              the east of kitchen_3 .
  RESPONSE>>  I constructed 2 kinds of models from your description.
              I remember them all.
YOUR INPUT>> you enter the kitchen
  RESPONSE>>  I am entering kitchen_3 from hall_1 and facing west .
  RESPONSE>>  I entered kitchen_3 .
  RESPONSE>>  This information is not used to renew the model.
YOUR INPUT>> there is a dining room to the right
  RESPONSE>>  I do not understand since the sentence is too difficult.
  RESPONSE>>  I can not understand what you mean.
  RESPONSE>>  After all , I could not renew the model.
YOUR INPUT>> there is a dining_room to the right
  RESPONSE>>  I understand the input sentence as dining_room_4 is to the north of kitchen_3
  RESPONSE>>  There are 4 kinds of models for the relation if it is true that kitchen_3 is
              to the south of dining_room_4 .
  RESPONSE>>  I constructed 4 kinds of models from your description.
              I remember them all.
YOUR INPUT>> you enter the dining_room
  RESPONSE>>  I am entering dining_room_4 from kitchen_3 and facing north .
  RESPONSE>>  I entered dining_room_4 .
  RESPONSE>>  This information is not used to renew the model.
YOUR INPUT>> the hall is to the right
  RESPONSE>>  I understand the input sentence as hall_1 is to the east of dining_room_4
  RESPONSE>>  There are 2 kinds of models for the relation if it is true that hall_1 is to
              the east of dining_room_4 .
  RESPONSE>>  I constructed 2 kinds of models from your description.
              I remember them all.
YOUR INPUT>> you enter the hall
  RESPONSE>>  I am entering hall_1 from dining_room_4 and facing east .
  RESPONSE>>  I entered hall_1 .
  RESPONSE>>  This information is not used to renew the model.
YOUR INPUT>> another bedroom is to the right
  RESPONSE>>  I understand the input sentence as bedroom_5 is to the south of hall_1
  RESPONSE>>  The interpretation seems to be inconsistent with earlier ones.
              bedroom_5 should not be to the south of hall_1 .
  RESPONSE>>  After all , I could not renew the model.
YOUR INPUT>> you turn left
  RESPONSE>>  I am in hall_1 and turning to the north .
  RESPONSE>>  Now, I am in hall_1 and facing north .
  RESPONSE>>  This information is not used to renew the model.
YOUR INPUT>> the bedroom is to the right
  RESPONSE>>  I assume that the reference of the word < bedroom > is one of the
              following : [bedroom_2,bedroom_5] .
  RESPONSE>>  The interpretation is inconsistent with what you said before.
              bedroom_2 should be to the west of hall_1 .
  RESPONSE>>  There are 2 kinds of models for the relation if it is true that
              bedroom_5 is to the east of hall_1 .
  RESPONSE>>  I constructed 2 kinds of models from your description.
              I remember them all.
YOUR INPUT>> you enter the bedroom
  RESPONSE>>  I am entering bedroom_5 from hall_1 and facing east .
  RESPONSE>>  I entered bedroom_5 .
  RESPONSE>>  This information is not used to renew the model.
YOUR INPUT>> you enter the bedroom
  RESPONSE>>  I am entering bedroom_5 from hall_1 and facing east .
  RESPONSE>>  I entered bedroom_5 .
  RESPONSE>>  This information is not used to renew the model.
```

Figure 9: System Behavior (1)

```
YOUR INPUT>> to the left is a bathroom
    RESPONSE>>   I understand the input sentence as bathroom_6 is to the north of bedroom_5
    RESPONSE>>   There are 4 kinds of models for the relation if it is true that bedroom_5
                 is to the south of bathroom_6 .
    RESPONSE>>   I constructed 4 kinds of models from your description.
                 I remember them all.
YOUR INPUT>> you enter the hall
    RESPONSE>>   I am entering hall_1 from bedroom_5 and facing east .
    RESPONSE>>   I entered hall_1
    RESPONSE>>   This information is not used to renew the model.
YOUR INPUT>>  you turn right
    RESPONSE>>    I am in hall_1 and turning to the north .
    RESPONSE>>    Now, I am in hall_1 and facing north .
    RESPONSE>>    This information is not used to renew the model.
YOUR INPUT>> the living_room is to the front
    RESPONSE>>   I understand the input sentence as living_room_7 is to the north of hall_1
    RESPONSE>>   There are 10 kinds of models for the relation if it is true that hall_1 is
                 to the south of living_room_7
    RESPONSE>>   I constructed 10 kinds of models from your description.
                 I remember them all.
YOUR INPUT>> the dining_room is on the front of the living_room
    RESPONSE>>   I understand the input sentence as dining_room_4 is to the south of
                 living_room_7
    RESPONSE>>   There is one model for the relation if it is true that dining_room_4 is to
                 the south of living_room_7 .
    RESPONSE>>   I renewed the model
YOUR INPUT>> the patio is on the south of the bedroom
    RESPONSE>>   I assume that the reference of the word < bedroom > is one of the following
                 : [bedroom_2,bedroom_5] .
    RESPONSE>>   There is one model for the relation if it is true that patio_8 is to the
                 south of bedroom_2 .
    RESPONSE>>   There is one model for the relation if it is true that patio_8 is to the
                 south of bedroom_5
    RESPONSE>>   Do you mean that patio_8 is to the south of bedroom_2 ? (yes or no)
          >> YOUR INPUT>> no
    RESPONSE>>   Do you mean that patio_8 is to the south of bedroom_5 ? (yes or no)
          >> YOUR INPUT>> yes
    RESPONSE>>   I renewed the model.
YOUR INPUT>> describe the layout
    RESPONSE>>   bathroom_6 faces east
                 hall_1 is to the east of bedroom_2
                 bedroom_5 faces east
                 hall_1 is to the east of dining_room_4
                 bathroom_6 is to the east of hall_1
                 bedroom_5 is to the east of hall_1
                 patio_8 is to the east of hall_1
                 hall_1 is to the east of kitchen_3
                 bathroom_6 is to the east of living_room_7
                 patio_8 faces east
                 bedroom_2 faces east
                 dining_room_4 faces west
                 kitchen_3 faces east
                 living_room_7 faces west
                 bedroom_5 is to the south of bathroom_6
                 bedroom_2 faces south
                 patio_8 is to the south of bedroom_5
                 kitchen_3 is to the south of dining_room_4
                 hall_1 faces south
                 bedroom_2 is to the south of kitchen_3
                 dining_room_4 is to the south of living_room_7
                 hall_1 is to the south of living_room_7
                 bathroom_6 faces north
                 living_room_7 faces north
                 patio_8 faces south
```

Figure 10: System Behaviour (continued)

7 Conclusion

An interactive system that understands natural language sentences describing spatial relations has been proposed. This work is evaluated in terms of

(1) whether the system determines the reference of an anaphoric or deictic expression from the literal meaning of the input and the implicit meaning derived from the literal meaning,

(2) whether the consistency of the spatial relations is maintained, and

(3) whether the correct topological representations of the layout are generated from well-formed descriptions.

At least concerning spatial relations, the system provides the natural language descriptions of architectural space with its symbolic representation in a non-deterministic way, and the integration enables us to verify whether natural language descriptions of architectural space are consistent or not with respect to a certain model of architectural space. Needless to say, spatial relations are not enough to describe the properties of architectural space, however, our system convinces us that it is possible to implement such a system that other information is handled based on a framework of the current system.

8 Acknowledgments

This work could not have been completed without the encouragement and careful eye for detail of Prof Bob Carpenter, Prof David Evans, and Prof Omer Akin. I wish to express my great thanks to them.

9 References

Akin, 0. Psychology of Architectural Design. (Pion 1986).

Barwise, J. The Situation in Logic. (Stanford: Center for the Study of Language and Information, 1989).

Barwise, J. & Perry, J. Situations and Attitudes. (The MIT Press, 1983).

Byrne, R. M. J. and Johnson-Laird, P. N. Spatial Reasoning. In Journal of Memory and Language, 28. (1989): pp. 564-575.

Carpenter, B. The Logic of Typed Feature Structures (Cambridge University Press, 1992).

Ehrlich, K. & Johnson-Laird, P. N. Spatial Description and Referential Continuity. In Journal of Verbal Learning and

Verbal Behavior, 21. (1982): pp. 296-306.

Flemming, U. Wall Representation of rectangular dissections and their use in automated space allocation. In Environment and Planning B, 5. (1978): pp. 215-232.

Grice, H. P. Logic and Conversation. In The Philosophy of Language, Edited by Martinich. (New York: Oxford University Press, 1975).

Herskovits, A. Language and Spatial Cognition. (Cambridge University Press, 1986).

Johnson-Laird, P. N. How is Meaning Mentally Represented? In Meaning and Mental Representation, (ed) U. Eco, M. Santambrogio and P. Violi. (Bloomington: Indiana University Press, 1988).

Levinson, S. C. Pragmatics. (New York: Cambridge University Press, 1983).

Mani, K. & Johnson-Laird, P. N. The Mental Representation of Spatial Description. In Memory & Cognition, 10(2). (1982): pp. 181-187.

Mitchell, W. J. The Logic of Architecture. (Cambridge: The MIT Press, 1990).

Pereira, F. C. N. & Pollack, M. E. Incremental Interpretation. In Artificial Intelligence, 50. (1991): pp. 37-82.

Schank, R. & Riesbeck, C. (ed). Inside Computer Understanding. (Yale University, 1981).

Winograd, T. A Procedural Model of Language Understanding. In Computer Models of Thought and Language, (ed) R. C. Schank and K. M. Colby. (San Francisco: W. H. Freeman and Company, 1973): pp. 152- 186.