Meaning of Japanese Spatial Nouns

Tokunaga Takenobu
Dept. of Computer Science
Tokyo Institute of Technology
Tokyo, 152-8552, Japan
take@cl.cs.titech.ac.jp

Koyama Tomofumi
Dept. of Computer Science
Tokyo Institute of Technology
Tokyo, 152-8552, Japan
tomoshit@img.cs.titech.ac.jp

Saito Suguru
Dept. of Computer Science
Tokyo Institute of Technology
Tokyo, 152-8552, Japan
suguru@img.cs.titech.ac.jp

Abstract

We have already proposed a framework to represent a location in terms of both symbolic and numeric aspects. To deal with vague linguistic expressions of a location, our representation adopts a potential function mapping a location to its plausibility. This paper proposes a classification of Japanese spatial nouns and potential functions corresponding to each class. We focus on the common Japanese spatial expression "*X no Y* (Y of X)" where X is a reference object and Y is a spatial noun. For example, "*tukue no migi* (the right of the desk)" denotes a certain location with reference to the desk. Expressions were collected from corpora, and the spatial nouns appearing in the Y position were divided into two major classes: designating a part of the reference object, and designating a location apart from the reference object. The latter class was further divided into two subclasses: *direction-type* and *distance-type*. For each class, a potential function was designed to provide the meaning of the corresponding spatial nouns.

1 Introduction

Research on animated agents capable of interacting with humans through natural language has drawn much attention in recent years (Badler et al., 1999; Bindinganavale et al., 2000; Cassell et al., 2000). We are also developing a dialogue system $\mathcal{K}3$ in which a user can command animated agents through speech input to manipulate various objects in a virtual world. The current system accepts simple Japanese speech inputs, including fragmental ones, such as "*Tukue made aruke* (Walk to the desk.)", "*Motto mae* (Further forward.)", and so on. The agent's behavior and the subsequent changes in the virtual world are presented to the user in the form of a three-dimensional animation (Tokunaga et al., 2004). In such situations, it is necessary to deal with the vagueness of language in order to generate an appropriate animation. In the previous example, generating an animation in response to "Walk to the desk." requires determining an exact target location to which the agent walks. Among various kinds of vagueness, we particularly focus on vagueness in spatial expressions as in this example.

There have been numerous studies on spatial knowledge and inference in various research fields, including artificial intelligence, cognitive science, and cognitive linguistics (Olivier et al., 1994; Horswill, 1995; Herskovits, 1986; Retsz-Schmidt, 1988; Levelt, 1989; Levinson, 2003). There are few studies, however, dealing with spatial vagueness in the interactive situation described above. For this purpose, we have already proposed a framework for dealing with spatial vagueness by introducing a two-level planning architecture and a hybrid representation of locations interfacing the two planners (Tokunaga et al., 2003). In this architecture, one planner deals with symbolic calculation, while the other deals with numeric calculation. A location is thus represented by a hybrid representation including both symbolic and numeric information. The numeric representation adopts a potential function that maps a location to its plausibility. In our framework, the potential function plays a key role in dealing with spatial vagueness .

In this paper, we discuss how Japanese spatial expressions, particularly spatial nouns designating a location, can be represented in terms of potential functions. We first collect spatial nouns from corpora and divide them into several classes. Then, we explore the relations between these classes and the corresponding potential functions, which provide the meaning of spatial nouns.

2 Japanese Spatial Nouns

In English, many spatial relations are represented by prepositions such as "in", "above", and so on. In contrast, in Japanese they are represented by spatial nouns and postpositional particles, or by postpositional particles alone. Table 1 lists examples of Japanese translations of English spatial prepositions.

English	Japanese
"in"	*naka de*, *naka ni*
"on"	*ue de*, *ue ni*
"at"	*de*, *ni*
"to"	*he*, *ni*

Table 1: Japanese translations of English spatial prepositions

In the first two cases, a spatial noun (*naka*, *ue*) and a postpositional particle (*de*, *ni*) are delimited by a space for readability, although usually no delimiter is put between words in Japanese. The primary role of Japanese postpositional particles is marking cases. For example, in sentence (1), the particles *ga*, *ni*, and *wo* mark the nominative, dative, and accusative cases, respectively[1].

(1) *Taro ga Hanako ni hon wo ageta*.
Taro NOM Hanako DAT book ACC gave.
(Taro gave Hanako a book.)

The particle *ni* often marks the locative case as well, as shown in sentence (2).

(2) *Ringo ga kago no naka ni aru*.
apple NOM basket of in LOC is.
(An apple is in the basket.)

[1]This is not a precise mapping, since the Japanese case system is different from that of the Romance languages. The mapping is shown for the convenience for non-Japanese speakers.

Comparing the structures of these Japanese sentences with their English translations, we find that there is no straightforward correspondence between English prepositions and Japanese postpositional particles. That is, from the viewpoint of case marking, *ni* takes the noun phrase *"kago no naka"* (inside of basket) as a complement and marks it as having the locative case. The structure will be *((kago no naka) ni)* in this case. Meanwhile, *"naka ni"* corresponds to "in", as shown in Table 1, from the viewpoint of spatial relations. Hence, the structure will be *((kago no) naka ni)* in this case. In general, spatial nouns rather than postpositional particles tend to bear the meaning of spatial relations in Japanese (Tanaka and Matumoto, 1997)[2].

In summary, spatial nouns play a central role in representing spatial relations in Japanese, just like prepositions do in English. Therefore, in this paper we analyze Japanese spatial nouns.

To collect spatial nouns for analysis, we first collected common Japanese spatial expressions of the form "*X no Y* (Y of X)", where X is a reference object and Y is a spatial noun. For example, in the expression "*tukue no migi* (the right of the desk)", "*tukue* (desk)" is a reference object and "*migi* (right)" is a spatial noun. This expression designates a certain location with reference to the location of the desk.

We extracted expressions of the form "*X no Y*" from five years worth of newspaper articles (Mainichi Simbun 1991–1995). However, the Japanese expression "*X no Y*" has various meanings besides denoting a location. To filter out irrelevant cases, we applied the following constraints:

- The word in position Y should be classified into the semantic class SPACE in a thesaurus. As a thesaurus, we used *Nihongo Goi Taikei* (Ikehara et al., 1997), in which the semantic class SPACE (2610) is classified under abstract relations and includes 2,761 words.

- The expression "*X no Y*" should be followed by the postpositional particle "*he* (to)", which

[2]Some English prepositions designate a path rather than a location, such as "to swim *across* the river" and "to go *through* the bush". In Japanese, path information about movement tends to be connoted in verbs and not explicitly expressed by postpositional particles or spatial nouns. Thus, we do not deal with path information in this paper.

strongly suggests a direction (Tanaka and Matumoto, 1997). We presume the expression preceding "*he*" designates a certain location. There are other postpositional particles suggesting locations, such as "*de* (at)", "*ni* (at/to)", and "*kara* (from)", but they are more ambiguous than "*he* (to)". We thus aim to gain precision at the cost of recall by using only "*he* (to)".

From this extraction process, we obtained 269 noun types appearing in position Y. After manual inspection, 59 nouns remained. The cases filtered out by the manual inspection included nouns that are too abstract, e.g., "*kûkan* (space)", and domain-specific nouns, e.g., "*naikaku* (inside corner)" in baseball and "*gôru* (goal)" in football. There are also cases in which a reference object immediately precedes the particle "*he* (to)", such as "*iriguti he* (to the entrance)".

Interestingly, this list missed several antonyms, which were manually added to obtain 70 spatial nouns in total. For example, we included "*X no zenbu* (the front part of X)" but not "*X no koubu* (the rear part of X)".

We do not claim at all that these nouns constitute a thorough list of Japanese spatial nouns. As a first step, however, we analyzed these nouns, as described in the following section.

3 Classification of Spatial Nouns

In this section, we discuss the classification of spatial nouns. This classification forms the basis of designing potential functions that provide the meaning of spatial nouns.

One notable classification viewpoint is the idea of whether a noun suggests a part of the reference object denoted by X. For example, "*tukue no hasi* (the end of the desk)" designates a certain part of the desk. On the other hand, "*tukue no tikaku* (near the desk)" designates a certain location near the desk but apart from it.

There are also ambiguous cases such as "*terebi no syômen* (the front of / in front of the TV set)". This expression designates a part of the TV set, its front surface, when used as "*terebi no syômen wo fuku* (to wipe the front of the TV set)", while it designates a location apart from the TV set when used as "*terebi no syômen ni tatu* (to stand in front of the TV set)".

The main verb and case marking provide clues to resolve this kind of ambiguity, as in the above example: "*-wo fuku* (ACC wipe)" vs. "*-ni tatu* (LOC stand)".

The nouns designating a location apart from the reference object can be further divided into two subclasses: *distance-type* and *direction-type*. *Distance-type* nouns concern the distance of a location from the reference object. The distance might be measured from the centroid of the reference object, but this is not always appropriate. There are cases in which the shape of the reference object should also be taken into account. For example, "*tukue no tikaku* (near the desk)" designates a certain area near the desk. The distance from the desk can be more appropriately measured from the contour of the desk rather than from its centroid.

Direction-type nouns concern the direction of a location with respect to the reference object, such as "*X no migi* (the right of X)", "*X no temae* (in front of X)", and so on. To identify a location designated by a spatial expression with *direction-type* nouns, a proper reference frame should be determined in the first place (Levinson, 2003). We assume that a proper reference frame can be selected independently of the spatial noun classes.

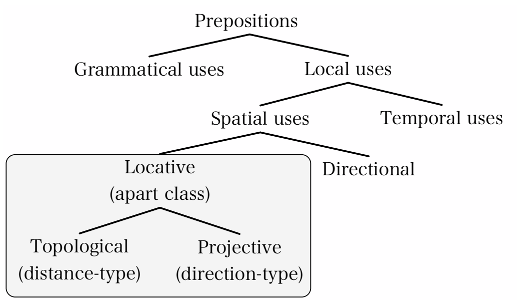


Figure 1: Classification of English prepositions

To summarize, we can divide spatial nouns into two major classes; designating a part of the reference object (*part*), and designating a location apart from the reference object (*apart*). The *apart* class is then further divided into two subclasses: *direction-type* and *distance-type*. Compared to an existing classification of English prepositions shown in Figure 1 (Lyons, 1968; Coventry and Garrod, 2004), the *apart* class and its subclasses *distance-type* and *direction-type* roughly correspond to "Locative", "Topological" and "Projective" prepositions, respectively.

Noun	Antonym	*part*	*apart*	
			dir.	*dist.*
zenbu (front part)	*koubu* (rear part)	o		
uti (in)	[*soto* (out)]	o		
tyûô (center)	[*mawari* (around)]	o		
sumi (corner)		o		
huti (edge)		o		
mae (front)	*usiro* (rear)	o	o	
migi (right)	*hidari* (left)	o	o	
syômen (front side)	*haimen* (back side)	o	o	
zenpou (forward)	*kouhou* (backward)	o	o	
higasi (east)	*nisi* (west)	o	o	
saki (tip, ahead)		o		o
soto (out)	[*uti* (in)]	o		o
mawari (around)	[*tyûô* (center)]	o		o
tikaku (near)				o
soba (close)				o
asimoto (below)				o

Table 2: Examples of Japanese spatial nouns

Table 2 lists examples of extracted spatial nouns and their classifications. Several notable characteristics can be observed in Table 2. First, compared to the *part*-*apart* opposition, nouns are clearly distinguished with respect to the *direction*-*distance* opposition. No noun is classified into both the *direction* and *distance* classes as far as our data is concerned. In contrast, many nouns classified into both the *part* and *apart* classes. This tendency is particularly prominent for *direction-type* nouns.

Many of the *direction-type* nouns and their antonyms behave similarly, that is, they are in the same row of the table, but this is not the case for *distance-type* nouns. The square bracketed words in the "antonym" column indicate that they are antonyms of the words in the first column but not classified as the row indicates. Therefore, they appear in separate rows in the first column. For example, "*uti* (in)" and "*soto* (out)" are antonyms of each other, but they have different classifications: "*uti* (in)" designates a part of the reference object, while "*soto* (out)" designates somewhere outside the reference object as well as a part of it.

There are two cases which do not fit well into this classification. The expression "*X no kage* (behind *X*)" designates a location hidden by the reference object *X*. With respect to the *part-apart* opposition,"*kage*" would be classified into the *apart* class, since its location is obviously somewhere apart from the reference object. However, we cannot explain this location in terms of only the *distance-direction* opposition. We need to take into account the constraint that a view is blocked by the reference object.

Another problematic case is "*ato* (trace)" which designates a place that the reference object used to occupy but does not now. Again, "*ato*" would be classified into the *apart* class, but we need to take into account changes in the situation over time in this case.

In the next section, we discuss the relation between spatial nouns and the corresponding potential functions on the basis of this classification.

4 Potential Functions

Each class of spatial nouns has a corresponding potential function type. Every function is designed to satisfy the following two conditions. First, it is differentiable throughout the domain. This condition is adopted to calculate the maximum value representing the most plausible location by using the steepest descent method (SDM).

Second, the function value ranges between 0 and 1 inclusive, representing the plausibility of a location, where a greater value suggests a more plausible location. Applying this condition enable us to translate logical relations between locations into arithmetic calculation of the corresponding potential functions.

A logical conjunction of two locations is translated into the product of the corresponding potential functions. For example, the plausibility of a location p expressed by "*Dai no migi de dai no soba* (the right of the table *and* close to the table)" can be calculated by the product of two functions: $F_{\text{right\_of}}(table_1, p) \times F_{\text{close\_to}}(table_1, p)$, where $F_{\text{right\_of}}$ and $F_{\text{close\_to}}$ are the potential functions corresponding to "*migi* (right)" and "*soba* (close)", respectively, and $table_1$ is the reference object. Figure 2 illustrates this example, with the z axis representing the plausibility of a location on the x-y plane.

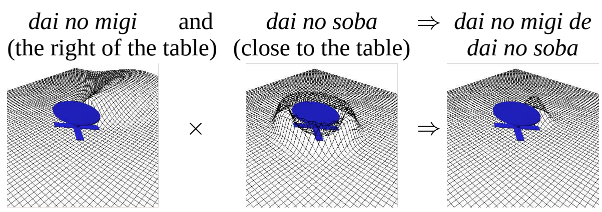


Figure 2: Conjunction of potential functions

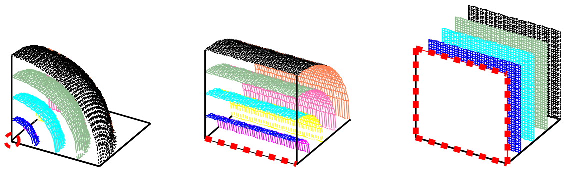
The plausibility of logical negation can be calculated by $(1 - F)$, where F is the potential function to be negated. The plausibility of logical disjunction can be derived from the combination of conjunction and negation (Tokunaga et al., 2003).

4.1 *Part* Nouns

The potential function of a *part* noun has a parameter of distance l from the base of the reference object. The function is defined as a Gaussian function that monotonically decreases according to the distance from the base, as described in (1):

$$f_{part}(l) = e^{-Al^2}, \qquad (1)$$

where A is an attenuation coefficient.



"*sumi* (corner)" (a) Base point — "*huti* (edge)" (b) Base line — "*zenbu* (front part)" (c) Base plane

Figure 3: The bases of potential functions for *part* nouns

The base is geometrically represented by (a) a point, (b) a line, or (c) a plane, depending on the characteristic of each noun. These bases should be described in the lexical entries of reference objects.

For example, the most plausible location designated by "*hako no sumi* (the corner of the box)" would be the corner vertexes of the box, as shown in Figure 3 (a). Therefore, the base of "*sumi* (corner)" is the vertex point of the box. Moving apart from this base point, the plausibility of the location designated by "*hako no sumi* (the corner of the box)" decreases. Likewise, "*huti* (edge)" and "*zenbu* (front part)" have a base line (Figure 3 (b)) and a base plane (Figure 3 (c)), respectively. In Figure 3, the meshed planes represent isometric surfaces of the potential values, and the bases are marked by the bold dotted lines.

4.2 Distance-type Nouns

The potential function of a *distance-type* noun is defined as a Gaussian function with a parameter of distance from the reference object:

$$f_{dist}(l) = e^{-A(l-O)^2}, \qquad (2)$$

where A and O are constants.

The constant A is an attenuation coefficient, and O (the offset) depends on the characteristics of a noun and defines locations apart from the reference object at which the potential function gives a maximum value. For example, O would be 0 for the potential function of "*soba* (close)", and a certain positive value for "*tikaku* (near)". The actual value of O for "*tikaku*" depends on the size of the reference object. In the current implementation, we set O at the height of the reference object.

The distance l is measured from the closest surface of the convex hull around the reference object.

The middle picture in Figure 2 shows the potential field of the spatial noun "*soba* (close)". The potential value decreases monotonically according to the distance from the convex hull of the reference object.

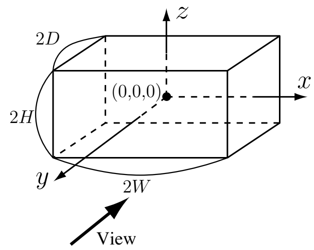


Figure 4: Coordinate system used in Equation (3)

4.3 Direction-type Nouns

The potential function of *direction-type* nouns is designed according to the following two considerations. First, a plausible direction is identical to one of six half-axes of the coordinate system (Herskovits, 1986). This consideration implicitly assumes that the shape of any reference object can be approximated as a cuboid. Note, however, that the half-axes of the cuboid are not always aligned to those of the reference object. The setting of the cuboid should be determined in consideration of the context and the reference frame (Tokunaga et al., 2003). Once the setting is determined, the proper coordinate system given by the reference frame is aligned to the cuboid, as illustrated in Figure 4. As mentioned before, we assume that a proper reference frame is selected before calculating the potential field.

Second, the values of the neighboring potential functions become identical on the planes bisecting two half-axes. These bisecting planes define the boundaries of neighboring directions. According to these considerations, we designed the following potential function:

$$f_{dir}(x,y,z) = e^{-A((\frac{y}{D+x-W})^2+(\frac{z}{H+x-W})^2)} \cdot g(x)$$
$$g(x) = \begin{cases} 1 & (x \geq W) \\ e^{-(x-W)^2} & (x < W) \end{cases}, \quad (3)$$

where A is an attenuation coefficient, and W, D, H are each one half of the edge lengths (width, depth, and height) of the cuboid approximating the reference object, as shown in Figure 4. The function $g(x)$ truncates the first factor of f_{dir} in the internal area of the cuboid. Figure 5 illustrates the potential field of the *direction-type* noun "*migi* (right)", where each cone represents an isometric surface of the potential value.

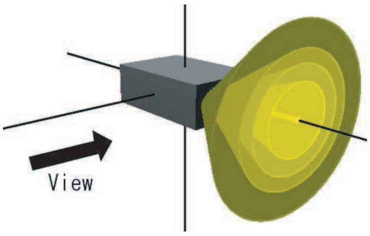
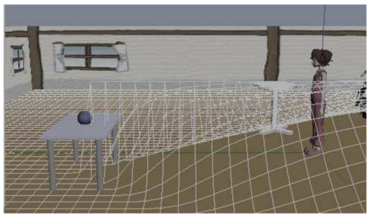
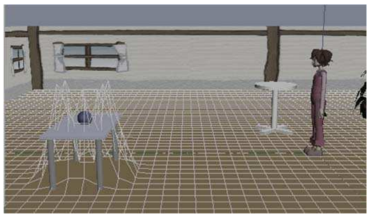


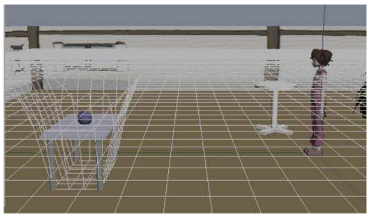
Figure 5: Potential field of a *direction-type* noun ("*migi* (right)")



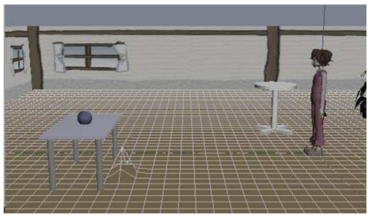
(a) Potential field of "in front of the table"



(b) Potential field of "reachable by hand"



(c) Potential field of "not occupied by the table"



(d) Composition of (a), (b) and (c)

Figure 6: Composition of potential fields

4.4 Using Potential Functions

As described in section 1, we intend to use the potential functions in an animated virtual agent system. We can illustrate the usage of the potential functions through the example shown in Figure 6. Suppose that the agent is instructed to take the ball on the table. She performs a conventional symbolic planning to generate a sequence of basic movements, such as "walk to the table" and "grasp the ball". Here, we focus on calculating the target location of the movement "walk". The location should satisfy at least three constraints: it is somewhere in front of the table, the agent can reach the ball from there, and it is not occupied by the table. The potential fields corresponding to these constraints are illustrated in Figures 6 (a), (b), and (c). By multiplying the value of these potential functions at each point, we obtain the potential field satisfying the three constraints, as shown in Figure 6 (d).

Note that the potential fields shown in Figures 6 (a) and (c) can be defined as proposed in this section, but the one shown in (b) cannot. The constraint corresponding to "reachable by hand" is not a spatial relation designated by spatial nouns but a functional relation. Our model has the possibility to deal with functional aspects of spatial relations, as shown in this example. However, the treatment of functional aspects in the current system is still *ad* *hoc*.

Although we have defined 3-dimensional potential functions, as in equations (1), (2) and (3), the potential function in this example is 2-dimensional. In the actual use of potential functions, we can derive 2-dimensional potential functions by projecting 3-dimensional ones on a proper plane where necessary. In this example, the functions are projected on the x-y plane.

5 Related Work

Using potential functions to represent spatial vagueness is not our original idea. Yamada *et al.* used potential functions to depict the deployment of buildings and landmarks on a map, given a text describing the spatial relations among them (Yamada et al., 1988). They divided spatial constraints among objects into three classes, namely, constraints on direction, distance and area, and they used potential functions to represent these constraints. However, they treated each object as a point and did not take into account the volume of an object. In addition, they dealt with only 2-dimensional potential fields.

Kalita and Badler conceptualized the space around a reference object as six orthogonal rectangular projected regions. Directions are discretely defined as being aligned to one of the six rectangular regions (Kalita and Badler, 1991).

Olivier and Tsujii also proposed a similar idea to give the semantics of projective prepositions of English by using potential fields. They used 2-dimentional potential fields to depict the furniture arrangement in a room as described by a text.

None of the above works dealt with the vagueness of a part of the reference object, that is, the *part* noun cases in our classification. This is because the main focus of these works is giving the semantics of prepositions that rarely represent a part of a reference object[3].

Coyne and Sprot reported a system named WordsEye that depicts a 3-dimensional picture based on a text describing spatial relations among objects (Coyne and Sprot, 2001). They dealt with 3-dimensional space by considering the volumes of objects. WordsEye avoids the problem of spatial vagueness by assigning tags representing spatial relations to each part of an object. Thus, WordsEye can deal with a part of an object. Annotating such tags, however, is labor-intensive and expensive. It is also generally difficult to enumerate all the necessary tags beforehand.

We designed the potential functions of spatial nouns so as to translate logical operations into numeric calculation of potential values. This feature enables us to deal with spatial vagueness by considering both logical and numerical constraints at the same time. In our prototype system, this is achieved by interfacing two types of planners and using a hybrid representation of a location (Tokunaga et al., 2003). We have also proposed a potential function corresponding to *part* nouns, which have attracted less attention in past research on English prepositions.

6 Concluding Remarks

For the work described in this paper, we collected Japanese spatial nouns from corpora and divided them into two major classes: *part* nouns and *apart* nouns. The latter were further classified into two subclasses: *distance-type* and *direction-type* nouns. Then, we sketched out potential functions corresponding to each of these classes. The potential functions provide the meaning of spatial nouns.

With respect to the potential functions, we found that *part* nouns could be further classified into three types regarding their bases, that is, nouns with base points, base lines, and base planes. The potential function for these nouns was designed so that the plausibility of a location decreases when moving apart from the base.

In the course of our analysis, we assumed that a proper reference frame for a spatial expression is already given, and that the potential function can be determined independently of the reference frame. This is still an open question, however, and we need to further investigate the interaction between the reference frame and the potential function.

As Coventry and Garrod have pointed out, interpretation of spatial prepositions sometimes requires functional relations, as well as geometric relations (Coventry and Garrod, 2004). We aim to use the proposed framework for planning agent behavior in a virtual world. The functional aspects are important in interpreting spatial expressions in this

[3] There are, of course, exceptions such as “inside” and “around”.

application. Suppose we would like to command an agent to open a door. The agent first needs to move to a certain location close enough to the door. In calculating the agent's target location, we need to take into account the fact that the agent must be able to reach the door knob from there. Furthermore, we might need to consider the type of door (a sliding door vs. a hinged door). Our current system incorporates such factors as potential functions, but in a very *ad hoc* manner, as described in 4.4. Further research is necessary to integrate such functional aspects into the proposed framework in a consistent manner.

The proposed potential functions include several parameters, which should be set according to the characteristics of the spatial noun and the reference object. Psychological experiments would be necessary to find the optimal values of these parameters. We have been working in collaboration with psycholinguists to determine the appropriate parameters for the model (Kojima and Kusumi, 2004). In addition, we are planning to evaluate our model through psychological experiments.

As described in section 3, there are still exceptions that cannot be classified well with our current classification scheme. Another research issue is collecting more examples of spatial expressions and extending the current classification.

Acknowledgments

We thank the anonymous peer reviewers for their valuable comments on a previous version of the paper. This work was partially supported by the "Grant-in-Aid for Creative Scientific Research" program 13NP0301 of Japan's Ministry of Education, Culture, Sports, Science and Technology.

References

N. I. Badler, M. S. Palmer, and R. Bindinganavale. 1999. Animation control for realtime visual humans. *Communication of the ACM*, 42(8):65–73.

R. Bindinganavale, W. Schuler, J. Allbeck, N. Badler, A. Joshi, and M. Palmer. 2000. Dynamically altering agent behaviors using natural language instructions. In *Autonomous Agents 2000*, pages 293–300.

J. Cassell, J. Sullivan, S. Prevost, and E. Churchill, editors. 2000. *Embodied Conversational Agents*. The MIT Press.

K. R. Coventry and S. C. Garrod. 2004. *Saying, Seeing, and Acting: The Psychological Semantics of Spatial Prepositions*. Psychology Press.

B. Coyne and R. Sprot. 2001. Wordseye: An automatic text-to-scene conversion system. In *SIGRAPH2001*, pages 487–496.

A. Herskovits. 1986. *Language and Spatial Cognition. An Interdisciplinary Study of the Prepositions in English*. Cambridge University Press.

I. D. Horswill. 1995. Visual routines and visual search. In *Proceedings of the 14th International Joint Conference on Artificial Intelligence*, August.

S. Ikehara, M. Miyazaki, A. Yokoo, S. Shirai, H. Nakaiwa, K. Ogura, Y. Ooyama, and Y. Hayashi. 1997. *Nihongo Goi Taikei – A Japanese Lexicon*. Iwanami Syoten. 5 volumes. (In Japanese).

J. K. Kalita and N. I. Badler. 1991. Interpreting prepositions physically. *AAAI-91 Proceedings Ninth National Conference on Artifical Intelligence, Vol. 1*, pages 105–110.

T. Kojima and T. Kusumi. 2004. The relationship between Japanese spatial terms and visual factors in three-dimensional virtual space. In *Proceedings of the Annual Conference of the Cognitive Science Society*.

W.J.M. Levelt. 1989. *Speaking: From Intention to Articulation*. The MIT Press.

S. C. Levinson. 2003. *Space in Language and Cognition*. Cambridge University Press.

J. Lyons. 1968. *Introduction to theoretical linguistics*. Cambridge University Press.

P. Olivier, T. Maeda, and J. Tsujii. 1994. Automatic depiction of spatial descriptions. In *Proceedings of AAAI 94*, pages 1405–1410.

G. Retsz-Schmidt. 1988. Various views on spatial prepositions. *AI Magazine*, 9(2):95–105.

S. Tanaka and Y. Matumoto. 1997. *Kûkan to Idou no Hyôgen (Expressions of Space and Movement*. Kenkyûsya. (in Japanese).

T. Tokunaga, T. Koyama, S. Saitô, and M. Okumura. 2003. Bridging the gap between language and action. In *Intelligent Virtual Agent - 4th International Workshop IVA 2003*, volume 2792 of *LNAI*, pages 127–135. Springer.

T. Tokunaga, K. Funakoshi, and H. Tanaka. 2004. K2: Animated agents that understand speech commands and perform actions. In *8th Pasific Rim International Conference on Artificial Intelligence (PRICAI 2004)*, pages 635–643.

A. Yamada, T. Nishida, and S. Doshita. 1988. Figuring out most plausible interpretation from spatial description. In *the 12th International Conference on Computational Linguistics (COLING)*, pages 764–769.