Classification of Japanese Spatial Nouns

Tokunaga Takenobu   **Koyama Tomofumi**
Saito Suguru   **Nakajima Masayuki**

Department of Computer Science
Tokyo Institute of Technology
Tokyo Meguro Ôokayama 2-12-1, Japan
take@cl.cs.titech.ac.jp

Abstract

We have already proposed a framework to represent a location in terms of both symbolic and numeric aspects. In order to deal with vague linguistic expressions of a location, the representation adopts a potential function mapping a location to its plausibility. This paper proposes classification of Japanese spatial nouns and potential functions corresponding to each class. We focused on a 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 location with reference to the desk. This expression were collected from corpora, and spatial nouns appearing in the Y position were classified into two major classes; designating a part of the reference object and designating a location apart from the reference object . And the latter class were further classified into two subclasses; direction-oriented and distance-oriented. For each class, a potential function were designed for providing meaning of spatial nouns.

1 Introduction

Research of 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 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 terms of a three-dimensional animation.

In such situation, it is indispensable to deal with 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 goal location which the agent walks to. 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 (Olivier et al., 1994; Horswill, 1995; Herskovits, 1986; Retsz-Schmidt, 1988; Levelt, 1989; Levinson, 2003). But there are few dealing with spatial vagueness in dynamic situation as described above. For this purpose, we have already proposed a framework to deal with spatial vagueness by introducing two level planning and a bilateral representation of locations interfacing two planners (Tokunaga et al., 2003). In this framework, a location is represented in terms of both symbolic and numeric representations. The numeric representation adopts a potential function which maps a location to its plausibility. The potential function plays a key role in dealing with spatial vagueness in our framework.

In this paper, we discuss how Japanese spatial expressions, particularly spatial nouns designating a location, can be represented in terms of potential functions. We collect spatial nouns from corpora, and classify them into several classes. Then, we explore the relations between these classes and the type of corresponding potential functions which give the meaning of spatial nouns. Section 2 describes a procedure to collect spatial nouns from corpora, section 3 describes the classification of the collected spatial nouns. Relations between this classification and potential functions are discussed in section 4. Section 5 concludes the paper and looks at the future work.

2 Collecting Japanese Spatial Nouns

There is no general way to enumerate all spatial expressions. In this paper, we focus on a common Japanese spatial expression "*X no Y* (Y of X)" where X is a reference object and Y is a spatial noun. For example, in an 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.

To start with, we collected the expression "*X no Y*" from five year worth of newspaper articles. However, Japanese expression "*X no Y*" has various meaning other than denoting a location. In order to filter out irrelevant expressions, the following constraints were used:

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

- The expression "*X no Y*" should be followed by a postposition "*he* (to)" which strongly suggests a direction (Tanaka and Matumoto, 1997). Thus we can assume the expression preceding "*he*" designates a certain location. There are other postpositions suggesting locations such as "*de* (at)", "*ni* (to)", "*kara* (from)", but they are more ambiguous than "*he* (to)". We aim to gain precision at the cost of recall by using only "*he* (to)".

As a result, we obtained 269 nouns in type appearing in position Y. After manual inspection, 59 nouns left. Interestingly, there were several missing antonyms in this list. For example, we had "*X no zenbu* (the front part of X)" but not "*X no koubu* (the rear part of X)". We added them manually to have 70 spatial nouns. We do not claim at all that these nouns make a thorough list of Japanese spatial nouns. As a first step, we will analyze these nouns in what follows.

3   Classification of Spatial Nouns

One of the obvious classification viewpoints is if a noun suggests a part of the reference object denoted by X. For example, "*tukue no hasi* (the end of a desk)" designates a certain part of the desk. On the other hand, "*tukue no tikaku* (near the desk)" designates a certain location near and apart from the desk.

There are ambiguous cases such as "*terebi no syômen* (in front of/the front of the TV set)". This expression designates a part of the TV set, its front surface, when being 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 being used as "*terebi no syômen ni tatu* (to stand in front of the TV set)".

The main verbs provide clues to resolve this kind of ambiguity as in the above example; "*fuku* (wipe)" vs. "*tatu* (stand)". Another potential clue is *Kanzi* characters composing spatial nouns. A *Kanzi* character is a ideogram representing a certain meaning by itself. This characteristic provides rich word formation ability of Japanese. For example, "部/*bu* (part)" has a meaning of "part", thus spatial nouns ending with this character suggest a part of the reference objects, such as "前部/*zen-bu* (front part)" and "後部/*kou-bu* (rear part)". On the contrary, "方/*hou* (toward)" denotes a direction, thus the nouns with this character suggest a certain location apart from the reference objects toward a certain direction, such as "前方/*zen-pou* (forward)" and "後方/*kou-hou* (backward)". In this paper, we focus on classifying spatial nouns and their corresponding potential functions, and do not deal with disambiguation issues.

The nouns designating a location apart from the reference object can be further classified into two subclasses; distance-oriented and direction-oriented nouns. The distance-oriented nouns concern a distance of a location from the reference object. The distance could be measured from a centroid of the reference object, but this is not always the case. There are cases in which the shape of the reference object should be also 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 by the distance from the contour of the desk rather than its centroid. The direction-oriented 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 summarize, we could classify spatial nouns into two major classes; designating a part of the reference object (*part*) and designating a location apart from the reference object (*apart*). And the *apart* class are further classified into two subclasses; *direction-oriented* and *distance-oriented*. Table 1 shows examples of extracted spatial nouns and their classification.

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

Table 1: Example of Japanese spatial nouns

There are several notable characteristics in Table 1. First, comparing to the *part/apart* contraposition, nouns are distinguished clearly by the *direction/distance* contraposition. No noun is classified into both *direction* and *distance* classes as far as our data is concerned. On the contrary, there are many nouns classified into both *part* and *apart* classes. This tendency was particularly prominent in direction-oriented nouns.

Many of direction-oriented 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-oriented 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 right columns indicate. Therefore, they appear in the separate rows in the first column. For example, "*uti* (in)" and "*soto* (out)" are antonym each other, but they have different classification, that is, "*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 into this classification. The expression "*X no kage* (behind X)" designates a location which is hidden by the reference object. To find the location designated by this expression, we need to search for an area where a view is actually blocked by the object. We could not explain this location in terms of only distance and direction.

Another case is "*ato* (trace)" which designates a location where the reference object used to occupy but not now. We need information about changes over time in this case.

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

4   Type of Potential Functions

Each class of these 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 necessary to find the maximal value representing the most plausible location.

Second, the function value ranges between 0 and 1 inclusive, representing plausibility of the location. Here, greater value suggests more plausible location.

The potential function with these conditions enable us to translate logical relations between locations into arithmetic calculation with corresponding potential functions.

Logical conjunction of two locations is translated into the product of corresponding potential functions. For example, the plausibility of a location x expressed by "*A no migi katu B no hidari* (the right of A AND the left of B)" can be calculated by the product of two functions, right_of_A$(x) \times$ left_of_B(x), where right_of_A and left_of_B are potential functions corresponding to the expressions "*A no migi* (the right of A)" and "*B no hidari* (the left of B)" respectively.

The plausibility of logical negation can be calculated by $(1 - f)$ where f is the original 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 which monotonically decreases according to the distance from the base as described in (1).

$$f = e^{-al^2}, \tag{1}$$

where a is a constant.

The base is geometrically represented by (a) a point, (b) a line or (c) a plane, depending on the characteristic of each noun.

In plausibility calculation of "*sumi* (corner)", the distance from the corner vertex is used as the parameter. For the plausibility of "*huti* (edge)", the distance from the edge line is used as the parameter. For the plausibility of "*zenbu* (front part)", the distance from the front plane of the reference object is used. Figure 1 shows the potential fields of the above cases. The meshed faces represent iso-surfaces of the potential value and the base are marked by the bold dotted lines.
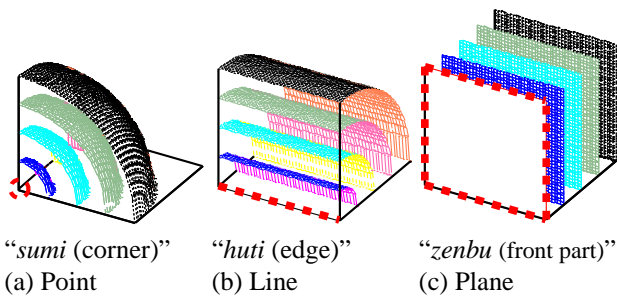

"*sumi* (corner)"   "*huti* (edge)"   "*zenbu* (front part)"
(a) Point   (b) Line   (c) Plane

Figure 1: The base of potential functions of *part* nouns

4.2 Distance-oriented Noun

The potential function of distance-oriented nouns is defined as a Gaussian function with a parameter of distance from

Figure 2: Potential field of distance-oriented noun "*soba* (close)"

the reference object as shown in (2).

$$f = e^{-b(l-d)^2}, \tag{2}$$

where b and d are constants.

Constant d depends on the characteristic of a noun and defines the peak of the function. For example, d would be 0 for the potential function of "*soba* (close)", and d would be a certain positive value for "*tikaku* (near)".

Distance l is measured from the closest surface of the convex hull around the reference object. Approximating the reference object by the convex hull improve the calculation speed. In addition, this approximation is allowable in our target domain.

Figure 2 shows the contour representation of the potential field of spatial noun "*soba* (close)", which is projected onto a horizontal plane. The potential value decreases monotonically according to the distance from the convex hull of the reference object.

4.3 Direction-oriented Noun

The potential function of direction-oriented nouns is designed based on the following two conditions. First, a plausible direction is identical to one of six semi-axes of the coordinate system. Second, the values of the neighboring potential functions becomes identical on the planes bisecting two semi-axes. The potential functions satisfying these conditions is depicted in Figure 3. In this figure, the centrally located box is an approximation of the reference object, and the value of the neighboring functions are the same on the translucent planes.
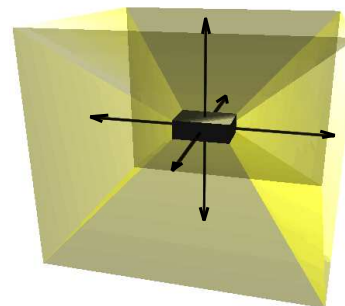

Figure 3: Potential field of direction-oriented nouns

As a potential function satisfying these two conditions, we designed a function by using an oval function as shown

in (3).

$$f = e^{-((\frac{x}{A+z-C})^2+(\frac{y}{B+z-C})^2)} \cdot g(z) \qquad (3)$$

$$g(z) = \begin{cases} 1 & (z \geq C) \\ e^{-c(z-C)^2} & (z < C) \end{cases},$$

where A, B, C are one half of the edge lengths of the rectangular prism approximating the reference object, and c is a constant (see Figure 4). The function $g(z)$ truncates the first factor of f in the area behind the plane $z = C$.
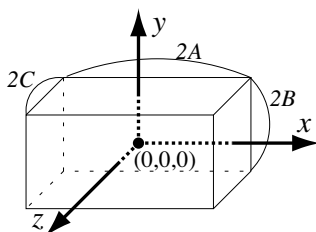

Figure 4: Coordinate system of Equation (3)

Figure 5 shows the contour representation on the horizontal plane ($y = 0$) of a potential field defined by the above function. In this case, we concern the direction of the z axis. This function can be used to represent such as "*X no mae* (in front of X)".
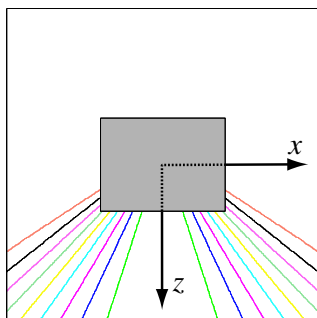

Figure 5: Potential field of a direction-oriented noun

Note that we assume that the base semi-axis is already given, although it is generally difficult to identify the indented coordinate system. This issue has been studied in various research field in the name of "reference frame" (Levinson, 2003). We do not deal with this problem in this paper.

5 Concluding Remark

In this paper, we collected Japanese spatial nouns from corpora and made classification consisting of two major classes, *part* nouns and *apart* nouns. The latter were further classified into two subclasses, *distance-oriented* and *direction-oriented* nouns. Then, we sketched out potential function types corresponding to each of these classes. With respect to the potential function type, we found that the potential function of *part* nouns could be further classified into three types based on their base part, that is, a base point, a base line and a base plane.

In the course of analysis, we assumed that the reference frame of a spatial expression is already given, and potential function can be determined independently of the reference frame. But this is still a open question, and we need to further investigate the interaction between the reference frame and potential functions.

As described in section 3, there are still exceptions which can not be classified by our current classification. Another research issue would be collecting more examples of spatial expressions and extend the current classification.

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.

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).

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.

P. Olivier, T. Maeda, and J. Tsujii. 1994. Automatic depiction of spatial descriptions. In *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.