Figuring out Most Plausible Interpretation from Spatial Descriptions

Atsushi Yamada,Toyoaki Nishida and Shuji Doshita
Department of Information Science
Kyoto University
Sakyo-ku, Kyoto 606, Japan
phone: 81-75-751-2111 ext. 5396
email: yamada or nishida%doshita.kuis.kyoto-u.junet%japan@relay.cs.net

Abstract

The problem we want to handle in this paper is vagueness. A notion of space, which we basically have, plays an important part in the faculty of thinking and speech. In this paper, we concentrate on a particular class of spatial descriptions, namely descriptions about positional relations on a two-dimensional space. A theoretical device we present in this paper is called the *potential model*. The potential model provides a means for accumulating from fragmentary information. It is possible to derive maximally plausible interpretation from a chunk of information accumulated in the model. When new information is given, the potential model is modefied so that that new information is taken into account. As a result, the interpretations with maximal plausibility may change. A program called SPRINT(SPatial Relation INTerpreter) reflecting our theory is in the way of construction.

1 Introduction

Natural language description is vague in many ways. The real world described with natural language has continuous expanse and transition, although the natural language itself is a descrete symbolic system. Vagueness plays an important role in our communication in that it allows us to transfer partial information. Suppose a situation in which a boy is looking around for his toy. Even if we cannot tell exactly where it was if we know it was somewhere around my desk, we can transfer him this partial information by telling that his toy is *around* my desk. It would be nice if we can communicate with our robot in the same way. We also use vague expression in the case of thinking it useless to give more detailed information to the hearer.

A theoretical device we present in this paper for the interpretation of such vague information is called *potential model*. The potential model employs both continuous and discontinuous functions to represent spatial relations, so that the probability changes either continuously or discontinously, depending on the nature of a given description. Currently, we are concentrating on a particular class of spatial relations, namely positional relations on a two-dimensional space, although the potential model is more general. We assume objects to be sizeless.

A program called SPRINT (SPatial Relation INTerpreter) reflecting our theory is in the way of construction.

2 The Potential Model

At the center of potential model is a *potential* function, which gives a value indicating the cost for accepting the relation to hold among a given set of arguments. The lower is the value provided by a potential function, the more plausible is the corresponding relation. We allow the value of potential functions to range from 0 to

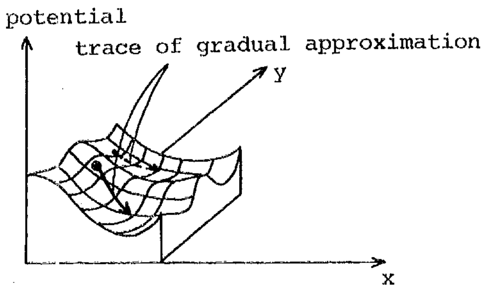


Figure 1: Potential Model and Gradual Approximation

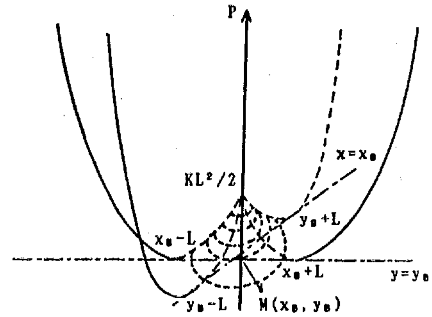


Figure 2: Distance Potential

$+\infty$. A potential function may give a minimal value for more than one combination of arguments. Such case may be taken as an existence of ambiguity.

A primitive potential function is defined for each spatial relation. A potential function for overall situation is constructed by adding primitive potential functions for spatial relations involved.

When a potential function is formulated from a given set of information, the system will seek for a combination of arguments which may minimize the value of potential function. We use a gradual approximation method to obtain an approximate solution. Starting from an appropriate combination of arguments, the system changes the current set of values by a small amount proportional to a virtual force obtained by differentiating the potential function. This process will be repeated until the magnitude of virtual force becomes less than a certain threshold. Figure 1 illustrates those idea.

Unfortunately, using the gradual approximation may not find a combination which makes a given potential function minimum. When there are some locally minimal solutions, this method will terminate with a combination appropriately near one of them. Which minimal solution is chosen depends on the initial set of arguments. We assume there exists some heuristic which predicts a sufficiently good initial values and the above approximation process works rather as an adjustment than as a means for finding solution.

2.1 The Spring Model

We use an imaginary, virtual mechanical spring between constrained objects to represent constraint on distance. If the distance between the two objects is equal to the natural length of the spring, the relative position is most plausible. The more extended or compressed the spring, the more (virtual) force is required to maintain the position, corresponding to the interpretation being less plausible.

An integration of the force needed either to extend or compress the spring is called an elastic potential. The spring model, subclass of the potential model, takes an elastic potential as a potential function. Let the positions of two objects connected by a spring of natural length L and elastic constant K be (x_0, y_0) and (x_1, y_1), respectively. Then the potential is given by the following formula:

$$P(x_0, y_0, x_1, y_1) = \frac{K(\sqrt{(x_1 - x_0)^2 + (y_1 - y_0)^2} - L)^2}{2}.$$

See figure 2 for the shape of this function.

2.2 Inhibited Region and Inhibited Half Plane

Unlike other primitive potential functions introduced so far, inhibited region and half plane pose a discontinuous constraint on the possible region of position. By inhibited region and half plane we mean a certain region and half plane is inhibited for an object to enter, respectively. Inhibited regions and half planes are not global in the sense that each is defined only for some particular object. Inhibited region is less basic concept because it can be represented by a logical combination of inhibited half plane.

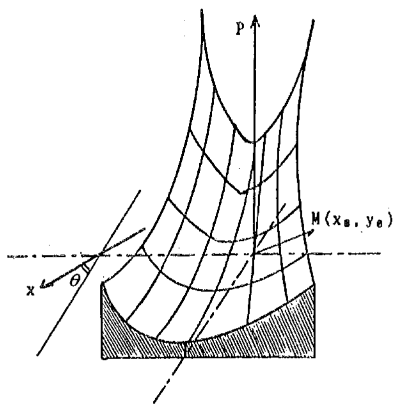


Figure 3: Directional Potential

An inhibited half-plane is characterized by its directed boundary line. A directed boundary line in turn is characterized by the orientation θ (measured counter-clockwise from the orientation of x-axis) and a location (X, Y) of a point (referred to as a *characteristic point*) on it. The inhibited half plane is the right hand side of the directed boundary.

2.3 Directional Potential

Suppose we want to represent a constraint that an object B is to the direction θ of another object A (measured counter-clockwise). Let the position of A and B be (x_0, y_0) and (x_1, y_1), respectively. We use the following potential function to represent the constraint:

$$P(x_0, y_0, x_1, y_1) = \frac{K_1(-(x_1 - x_0)\sin\theta + (y_1 - y_0)\cos\theta)^2 + K_2}{(x_1 - x_0)\cos\theta + (y_1 - y_0)\sin\theta + 1/\delta}.$$

When viewed horizontally from A, this function represents a hyperbola. If this function is cut vertically to the intended direction, this represent a parabola (upside down). See figure 3 for the shape of this function. Note that the notion of direction defined here denotes that in everyday life, which is not very rigid.

Since the value of the potential function defined above P jumps from $+\infty$ to $-\infty$ if one proceeds for the $-\theta$ direction.

We add inhibited half planes in the $-\theta$ direction, so that it is impossible to put the object in this region.

3 A Method of Gradual Approximation

A maximally plausible position is obtained by revising a tentative solution repeatedly.

The move $\Delta = (\Delta_x, \Delta_y)$ at each step is given as follows:

$$\Delta = (\Delta_x, \Delta_y) = K \cdot (\partial P/\partial x, \partial P/\partial y),$$

where K is a positive constant.

This basic move may be complicated by taking inhibited regions into account. The following subsection explain how it is done.

3.1 Avoiding to Place Objects within its Inhibited Half Plane

An algorithm for escaping from inhibited half plane is applied when an object is placed within its inhibited half plane. If such a situation is detected, the algorithm defined below will push the object out of an inhibited half plane in n steps. At this time, any influences from other constraints are taken into account. Thus, the move $d = (d_x, d_y)$ of the object at each step is the sum of $d_V = (d_{V_x}, d_{V_y})$ (a component vertical to the boundary) and $d_P = (d_{P_x}, d_{P_y})$ (a component in parallel to the boundary). Suppose the initial position of an object is (x_0, y_0), then each of which is defined as follows:

$$d_{V_x} = -L\sin\theta/n$$

$$d_{V_y} = L\cos\theta/n$$

where,$L = |(x_0 - Y)\sin\theta - (y_0 - Y)\cos\theta|$ represents the distance from the initial position to the boundary of the inhibited half plane. Note that the inhibited half plane is characterized by its directed boundary with a characteristic point (X, Y) and the orientation θ.

$$d_{P_x} = C(f_x\cos^2\theta + f_y\sin\theta\cos\theta)$$

$$d_{P_y} = C(f_x\sin\theta\cos\theta + f_y\sin^2\theta)$$

where, C is a positive constant, and $f = (f_x, f_y)$ is a virtual force from other constraints. Figure 4 illustrates how this works.

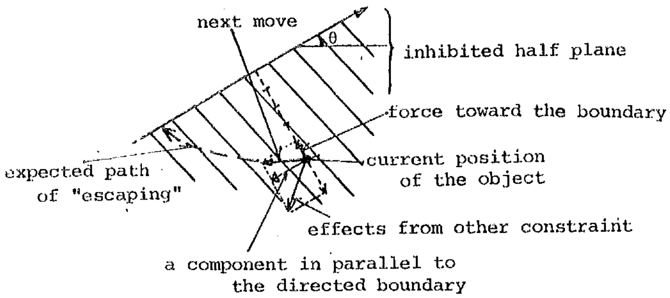


Figure 4: Pushing an Object Out of an Inhibited Region

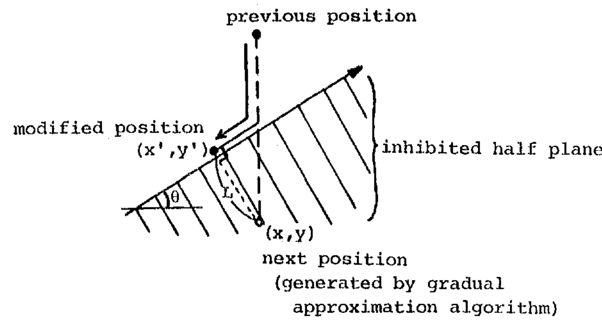


Figure 5: Avoiding to Push an Object into an Inhibited Region

Once an object has been put out of an inhibited half plane, one must want it not to have it re-enter the same inhibited half plane. However, the gradual approximation algorithm may try to push the object there again. An algorithm for avoiding to push objects into it watches out for such situation. If it detects, it will recourse the gradual move.

Suppose an inhibited half plane is characterized by θ and (X, Y) on the boundary. Suppose also that the next position suggested by the gradual approximation algorithm is (x, y). If

$$L = x\sin\theta - y\cos\theta - X\sin\theta + Y\cos\theta > 0$$

then, the next position will be forced into the inhibited half plane. In such a case, the move is modified and the new destination becomes:

$$(x', y') = (x - (1+\epsilon)L\sin\theta, y + (1+\epsilon)L\cos\theta)$$

where, ϵ is a positive infinitesimal.

See figure 5.

3.2 Dependency

It would require a great amount of computation, if the position of all objects have to be determined at once. Fortunately, human-human communication is not so nasty as this is the case; natural language sentences contain many cues which help the hearer understand the input. For example, in normal conversations, the utterance

Kyoto University is to the north of Kyoto Station

is given in the context in which the speaker has already given the position of *Kyoto Station*, or s/he can safely assume the hearer knows that fact. If such a cue is carefully recognized, the amount of computation must be significantly reduced.

Dependency is one such cue. By dependency we mean a partial order according to which position of objects are determined. SPRINT is designed so that it can take advantage of it. Instead of computing everything at once, the spatial reasoner can determine the position of objects one by one. An object whose position does not depend on any other objects is chosen as the origin of local coordinate. SPRINT determines the temporary position of objects from the root of the dependency network. The position of an object will be determined if the position of all of its predecessors is determined. Figure 6 shows how SPRINT does this.

This algorithm has three problems:

1. the total plausibility may not be maximal.
2. in the worst case, the above may result in contradiction.
3. objects may be underconstrained.

Currently, we compromise with the first problem. More adequate solution may be to have an adjustment stage after initial configulation of objects are obtained. The second problem will be addressed in the next subsection. The third remains as a future problem.

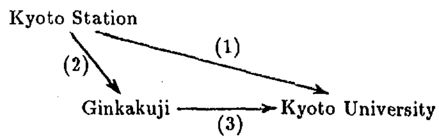
3.3 Resolving Contradiction

Adding new information may result in inconsistency. In order to focus an attention to this

GIVEN TEXT:

(1) Kyoto University is to the north of Kyoto Station.
(2) Ginkakuji(temple) is to the northeast of Kyoto Station.
(3) Kyoto University is to the west of Ginkakuji(temple).

DEPENDENCY:



RESULT(PRODUCED BY SPRINT):

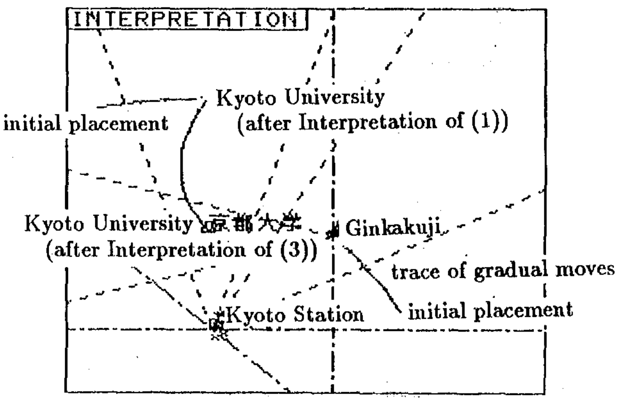
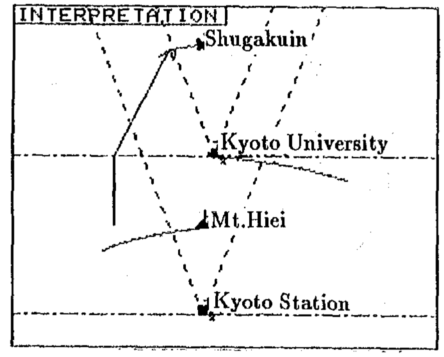


Figure 6: Positioning using Dependency

GIVEN TEXT:

(1) Mt.Hiei is to the north of Kyoto Station.
(2) Kyoto University is to the north of Kyoto Station.
(3) Shugakuin is to the north of Kyoto University.
(4) Shugakuin is to the south of Mt.Hiei.

BEFORE INTERPRETATION OF (4):



AFTER INTERPRETATION OF (4):

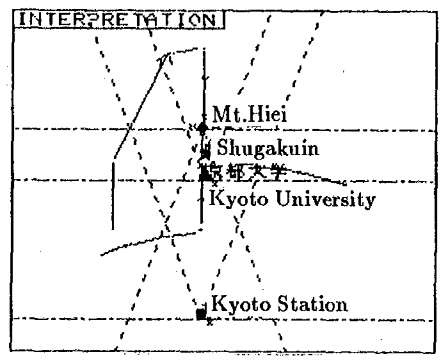


Figure 7: Resolving contradiction.

problem, let us temporalily restrict the spatial coordinate as one-dimensional. Suppose an object is given a maximally plausible position x_0. Suppose also that a new inhibited region (interval I in a one dimensional world) is given as a new constraint. Then the position of the object is recomputed so as to take this new constraint into account. If the interval I accidentally involves x_0, then the object may be moved out of interval I. This is the situation in which the object tends to move to the position x_0 but cannot due to the inhibited half space. In this case, the parent node in the dependency is tried to move in the reverse direction to resolve this situation.

A situation is worse than the above if the inhibited region (or interval) is too wide to fit in a space. This problem rises especially when we take size into account. Suppose the position of two objects A and B are already given maximally plausible positions x_0 and $x_1 (x_0 < x_1)$, respectively. Suppose now the third object C with width being wider than $x_1 - x_0$ is declared to exist between A and B. This causes a failure because there is no space to place C.

The solution to this problem comprises in two stages. First, the reason of the failure is analyzed. Then, parents of the current objects are moved gradually so that the inconsistency can be removed. Figure 7 illustrates how this works.

4 Related Work

The problem of vagueness has not been studied widely in spatial reasoning[Kui78,Lav77, NJ77,Wal81]. Work by Drew McDermott and Ernie Davis is among few exceptions. They addressed the vagueness of spatial concept and they introduced a theoretical device called *fuzz box*[Dav81]. A fuzz box denotes a region in which a given object may possibly exist. Possibility of the existence is uniformly positive in a fuzz box, and zero outside the box.

This formulation has a couple of drawbacks. First, the shape of the region must be rectangular or circle. Second, Davis had to have the boundary of a fuzz box sharp, due to computational reasons.

Thus their approach has a significant difficulty in modeling various spatial concepts. For example, the meaning of *aroundness* is hard to represent with fuzz box, since it is still hard to draw an exact boundary to distinguish the region which is around something from that that is not.

5 Concluding Remarks

Vagueness of natural language description contains lots of hard issues, only a fragment of which has been addressed in this paper. Regarding the language understanding process as a reconstruction of the described world, we have studied on spatial descriptions only declaring existence of objects.

A couple of major problems related to this work are mentioned below.

1. A systematic method should be developed to determine actual values of the model. Currently, we should confess that parameter values are determined rather subjectively so that sample problems may be solved. In future, we want to apply adaptive learning mechanism.

2. Although the current program is forced to figure out most plausible configuration from given information, there do exist cases in which things are so underconstrained that figuring out temporary configuration is useless or rather harmful.

Of course, other problems are remain unsolved. The model should be extended as that objects can have size and shape; initial placement heuristic should be incorporated; *etc.* Those seem to be less hard. In fact, the initial implementation of SPRINT is being extended.

References

[Dav81] E. Davis. *Organizing Spatial Knowledge*. Yale University, 1981.

[Kui78] B. Kuipers. Modeling spatial knowledge. *Cognitive Science*, 2(2), 1978.

[Lav77] M.A. Lavin. *Computer Analysis of Scenes from a Moving Viewing Point*. Massachusetts Institute of Technology, 1977.

[NJ77] G.S. Novak Jr. Representations of knowledge in a program for solving physics problems. In *Proceedings IJCAI-77*, 1977.

[Wal81] D.L. Waltz. Towards a detailed model of processing for language describing the physical world. In *Proceedings IJCAI-81*, pages 1–6, 1981.