Monte Carlo Simulation for Plausible Interpretation of Natural-Language Spatial Descriptions

Dan Tappan
College of Engineering, Idaho State University, Pocatello, ID, USA

Abstract - *This system generates plausible three-dimensional visualizations of basic English descriptions of a zoo environment (mainly animals and plants) by means of a Monte Carlo simulation. It combines a semantic network and an inheritance-based knowledge base as representations for explicit and implicit spatial information, respectively. Its linguistically motivated aspects address underspecification, vagueness, uncertainty, and context, as well as intrinsic and deictic frames of spatial reference. The underlying reasoning formalism is probability-based geometric fields, which are used for qualitative constraint satisfaction.*

Keywords: spatial reasoning, scene generation, simulation

1 Introduction

For any simple English description of a physical scene, like *a dog is in front of a cat and near a tree*, anyone can easily formulate a corresponding mental image or model. The description itself explicitly contributes only a tiny fraction of the details that the image contains. In fact, most of the content comes from an implicit, commonsense understanding of its objects and how they can and cannot be plausibly depicted.

Spatial reasoning of this sort, like most intelligent processes, is a difficult computational task to emulate, despite its apparent, intuitive simplicity for humans [1]. What makes the problem especially troublesome is that computers lack our intangible knowledge of the world and powerful abilities to reason intelligently over it. This work addresses the primary aspects of these issues in terms of what to represent and how to represent it. It uses a simple representation of a description in conjunction with a relatively simple knowledge base of relevant spatial details to define the declarative form of a valid solution. A constraint satisfaction algorithm running in a Monte Carlo simulation then generates any number of corresponding interpretations with plausible positions and orientations for the objects. Such solutions can directly support many applications that use, or could benefit from, natural language, like text understanding, machine translation, question-and-answer systems, query and search engines, and so on [2].

2 Background

The knowledge representation for explicit and implicit details in this work addresses overlapping issues in language and the spatial world. Language plays a key role because it closely reflects human perception and understanding, which are the basis of plausibility [3,4]. In particular, four spatial issues are the emphasis here. First, *underspecification*, or the lack of complete details in a description, requires background, world knowledge to supply non-explicit information. Second, *vagueness*, or the imprecise nature of descriptions, requires knowledge that defines a range of plausible interpretations. Third, *uncertainty*, or the lack of commitment to a particular interpretation, requires knowledge of tendencies or preferences over this range. And fourth, *context*, or the different interpretation of objects in certain combinations, requires knowledge to identify such patterns and define the differences.

These linguistic issues map to the primary goal: generating valid and preferred spatial interpretations of the objects in a description, specifically their positions and orientations with respect to three contextually determined frames of spatial reference [1,5,6]. The *intrinsic* (or object-centered) frame generally applies to objects that have an accepted, or canonical, front; e.g., *in front of the dog* means some position in line outward from its face. The *extrinsic* (or environment-centered) and *deictic* (or viewer-centered) frames are generally the opposite case for objects without a canonical front; e.g., *in front of the tree* means in line outward from it to another position in the world that establishes a virtual front region. In the extrinsic frame, this reference position is arbitrary; e.g., *in front of the tree as seen from the lake*. In the deictic frame, which is a specialized case of the extrinsic frame, it is the (usually implicit) position of the viewer; e.g., *in front of the tree (as seen by the viewer in the north looking south)*. For space reasons, this paper discusses only the intrinsic and deictic frames, although this work accommodates all three.

The underspecified, vague, uncertain nature of typical descriptions lacks the preciseness that a quantitative, or absolute, numerical approach to spatial reasoning would require [7]; e.g., *the cat is 3.2 meters and 45.0 degrees clockwise from the dog located at coordinate (25,15)*. This work, like most linguistically motivated work, adopts a qualitative approach that reasons in terms of more natural, relative constraints [8]; e.g., *the cat is to the front-right of the dog and near it*.

Despite the potential of such research, relatively few contemporary systems exist [9]. CarSim [10] focuses on graphically rendering the results of vehicle collisions based on accident reports. WordsEye [11], the closest to this work, focuses on depicting appropriate static poses for actions. Although both address text understanding and employ various degrees of knowledge representation, they focus more on producing the graphical results and less on investigating the underlying linguistics and knowledge processing. Other approaches adopt purely geometric solutions, with superficial consideration of relevant spatial knowledge [12-14].

The approach to inferring spatial knowledge in this work loosely extends [15-18] for scene interpretation. Tversky [19] covers in comprehensive detail many of the confounding spatial issues. And [1,5,6], in particular, form the basis for defining and interpreting spatial frames of reference. More recent work, especially in Geographic Information Systems, attempts to account for contextual information [20-27].

3 Natural-Language Descriptions

A description in this work consists of nouns, adjectives, prepositions, and several support words like determiners, conjunctions, and the verb *to be*. The nouns refer to concrete, physical objects within a zoo scenario. Aside from obvious visual appeal, animals and plants exhibit a variety of interesting spatial characteristics. This paper does not address the adjectives, which play a role in the contextually appropriate determination of size. The prepositions are the relations in Table 1, with determiners and conjunctions for readability, and without the hyphens; e.g., *in front and left of* and *at the fringe of*.

As in most related systems (except [10]), descriptions are manually fabricated rather than automatically acquired from real-world sources. This approach eliminates troublesome parsing issues that are outside the scope of investigation. They must also refer to static scenes only, which is a common limitation due to the complexities of verb interpretation, movement, time dependencies, the frame problem, etc. [28,29,2,11].

4 Explicit Knowledge Representation

Stage 1 represents the explicit knowledge in a description with a straightforward semantic network of object nodes, attribute nodes, and directed relation arcs, which map closely to nouns, adjectives, and prepositions, respectively [29]. Each object node refers to a single, unique, unambiguous object. Each arc specifies a binary relation that implies a constraint from its source object to its target object, as well as a context between them. For example, Figure 1 depicts the semantic network for the following description:

1. *Loki is a retriever.*

2a. *The tree is north of Loki.*
2b. *Loki is facing the tree.*

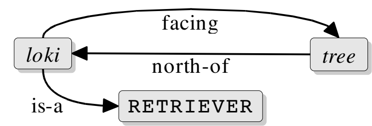

Figure 1: Semantic Network

5 Implicit Knowledge Representation

Stage 2 involves deriving the unstated properties and rules that implicitly describe the spatial relations of each object. This form of inference is static in the sense that it considers each object in isolation, not in context with other objects [21].

Despite its name, the semantic network from Stage 1 explicitly represents only the syntax (or structure) of the configuration without any consideration of its real-world semantics (or meaning). To understand the semantics even superficially requires deeper analysis into what the objects are and how their rules apply to them [30].

The source of the implicit, commonsense background knowledge for this analysis is a simple knowledge base that is similar to an inheritance hierarchy in object-oriented programming [31]. It currently contains 108 prototypical concepts, each of which either inherits its properties and rules from its ancestors (via single inheritance), or it defines/overrides them itself. A simplified example appears in Figure 2.

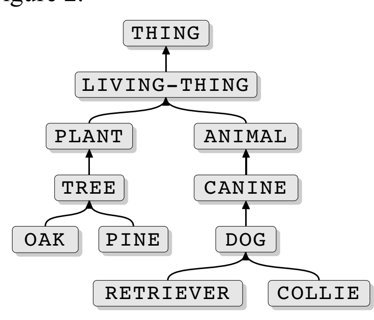

Figure 2: Knowledge Base

5.1 Spatial Properties

A property defines whether a concept exhibits a particular spatial characteristic. The only one found necessary within the scope of this paper specifies whether a concept has a canonical front. Since objects are not articulated, any head is always fixed in line with the body. This simplification eliminates the need to account for endless variation in the configuration of body parts; e.g., the body of the dog is oriented north, but it is looking east.

5.2 Spatial Relations

A relation is a qualitative spatial context and constraint between two objects. Table 1 defines 34 position, distance, and orientation relations; to conserve space, position and orientation omit 27 additional variants prefixed with `direct`, which specify a narrower interpretation of the same general meaning. Each relation *R* is of the binary form *xRy*, where *x* and *y* are objects. Most static, spatial prepositions in English fall into these classes [1, 32-36]. An additional 19 in other classes are beyond the scope of this paper.

Table 1: Spatial Relations

Class	Relations	
Position	`in-front-of`	`in-front-left-of`
	`in-back-of`	`in-front-right-of`
	`left-of`	`in-back-left-of`
	`right-of`	`in-back-right-of`
	`north-of`	`northeast-of`
	`south-of`	`northwest-of`
	`east-of`	`southeast-of`
	`west-of`	`southwest-of`
	`between`	
Distance	`inside`	`midrange-from`
	`outside`	`far-from`
	`adjacent-to`	`at-fringe-of`
	`near`	
Orientation	`facing-north`	`facing-northeast`
	`facing-south`	`facing-northwest`
	`facing-east`	`facing-southeast`
	`facing-west`	`facing-southwest`
	`facing`	
	`facing-away-from`	

5.3 Spatial Rules

A rule specifies when a particular relation, like `near`, applies from one object to another. It uses a formalism of geometric fields that describe a collection of cells in a two-dimensional, top-view, polar projection centered around the source object [6,13,37-39]. Experimentation suggests that 32 sectors and 100 rings are sufficient for the current domain of concepts and relations. Although any combination of selected cells among the available 3,200 is valid, in practice, only variations of two types define all spatial relations in this work: *wedges* apply to position and orientation relations, and *rings* to distance relations. Figures 3a and 3b show respective examples of the fields for relations `front-of` and `far-from`, where the object in the center is facing the direction of the arrow.[1]

1 For clarity, not all rings are shown.

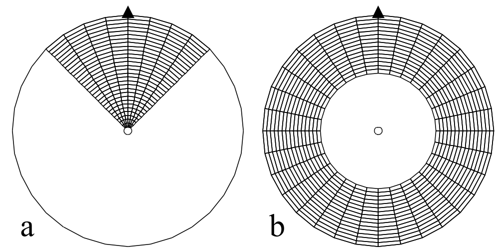

Figure 3: Geometry of Wedge and Ring Fields

Each field definition consists of two components. The first specifies its *geometry*, which constrains where other objects must appear with respect to the relation. The second specifies its *topography*, which overlays a commonsense probability distribution onto the geometry. The distribution is normally a variation of a bell curve, but it is adjustable. Figure 4 shows which positions in Figure 3 are favored. This formalism reflects the "scruffy" nature of spatial relations due to vagueness and uncertainty: positions in the center of perceptual focus are more probable than those at the periphery [8,3]. For the purpose of spatial inference, the geometry of a field sanctions the positions that are legal, and the topography recommends a subset that are contextually preferred [30].

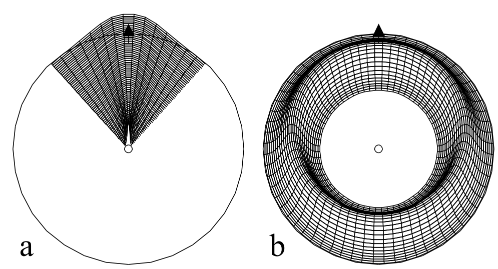

Figure 4: Topography of Wedge and Ring Fields

The relations in many descriptions interact to constrain the interpretation of objects further [1]. Fields accommodate such compositional behavior through logical operations over the geometry and topography. Figure 5 illustrates for the `front` and `far-from` fields the intersection that corresponds to the prepositional phrase *in front of and far from*. In the same way, it also supports union (*in front of or far from*), symmetric difference (*either in front of or far from, but not both*), and complement (*not in front of*).

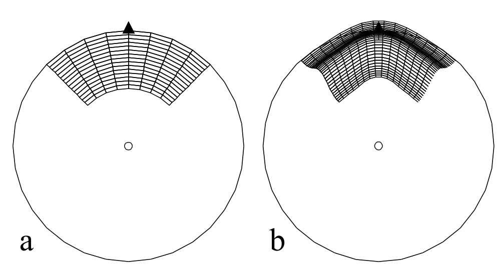

Figure 5: Intersection of `front` and `far-from` Fields

Two important factors play a role in the contextual application of fields. The first is frame of spatial reference. For any concept with a canonical front, the default frame of reference is intrinsic; i.e., anything in front of it is in line with the direction it is facing. A field reflects this spatial behavior by rotating itself so that its arrow aligns with the orientation of its object. Thus, its front field aligns with this direction, and its back, left, and right fields respectively align 180 degrees, 90 degrees counterclockwise, and 90 degrees clockwise from it. On the other hand, for any concept without a canonical front, the default reference frame is deictic; i.e., anything in front of it is in line between itself and the viewer. In this case, the arrow aligns to the position of the viewer. Finally, all concepts support compass directions for relations like `north-of`, `south-of`, etc. In this case, the arrow always aligns to north, or top center.

The second factor is scale. It receives only passing mention because it relies on the dimensions of objects, which this paper does not address. In general, the contextual interpretation of distance depends on the size of the reference object at the center of its field [40,6,41]. For example, what is *near* for a `giraffe` is *far* for a `rabbit`. The relative diameter of fields reflects this relationship.

5.4 Combined Representation

Stage 3 combines the explicitly stated information from the semantic network with the implicitly inferred background knowledge from the knowledge base. Figure 6 depicts a simplified example of this process: objects *Loki* and *tree* link to concepts `RETRIEVER` and `TREE`, respectively. Through inheritance, *Loki* derives the rules about his ancestor concepts `DOG`, `CANINE`, `ANIMAL`, `LIVING-THING`, and `THING`. The same process holds for *tree*. It is important to note the distinction between an *object*, which is a unique instance in the description, and a *concept*, which is a set of properties and rules that all instances of it must have in common. For clarity, this distinction is rendered typographically through italics and capitalized typewriter font, respectively.

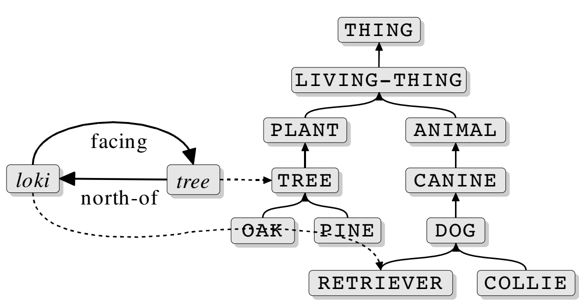

Figure 6: Semantic Network Linked to Knowledge Base

6 Constraint Satisfaction

The knowledge base addresses the problem of context by conditionally applying definitions for default and non-default interpretations. A default interpretation occurs when an object node is either not part of a relationship, or no other objects in any of its relationships affect its prototypical, spatial behavior. For example, *there is a hippo* instantiates a particular hippo that has no justification to differ from a standard, "generic" hippo. Similarly, *the hippo is in the zoo* states an inert relationship that generally imparts no different interpretation on this hippo than it would on any other. Thus, a default interpretation is independent of context and reflects the *semantics* of a concept pairing.

A non-default interpretation is the complementary case. For example, *the hippo is in the corral* implies that its body is on the surface of world, whereas *the hippo is in the lake* implies that it is below the surface. The appropriate vertical interpretation is critical and certainly not arbitrary or interchangeable for a `hippo`. On the other hand, either is acceptable for *the duck is in the lake*. Thus, a non-default interpretation is dependent on context and reflects the *pragmatics* of a concept pairing.

This work employs two mechanisms to identify such contextual patterns for any concept pairing. The first is by *association*, which triggers on specific target concepts in a relationship. The specification can be *extensional* by exhaustively listing all the concepts that have the same spatial effect on the source concept; e.g., `lake`, `pond`, and `pool`. It can also be *intensional* by indicating the branch of the hierarchy that subsumes the individual concepts; e.g., `body-of-water`. This form eliminates the need to enumerate all concepts that are equivalent in a certain respect. It also simplifies maintenance and expansion of the knowledge base because the list does not require updating if new, spatially equivalent concepts are added; e.g., `river`, `stream`.

The second mechanism is by *conditional dependency*, which triggers on specific properties inside the definitions of other concepts in a relationship. This frame-based formalism uses a traditional slot-filler structure to associate values with properties arbitrarily [29]. The most common is the boolean `has-canonical-front`.

6.1 Generation

Stage 4 uses a constraint propagator to find valid solutions for the various spatial behaviors that fields specify. It does so by calculating random values over the field topographies for the position and orientation of every object, so that all their values simultaneously satisfy all their field constraints. This nondeterministic behavior addresses uncertainty because there are an infinite number of valid interpretations for any description [2]. The geometries guarantee that any solution is valid, and the topographies attempt to bias them toward more plausible (or less controversial) interpretations.

The semantic network serves as a dependency graph that defines how the objects constrain each other. Objects that are neither directly nor indirectly interconnected form disjoint semantic sub-networks that cannot interfere with each other (except for placement collisions, which are prevented); e.g., *the dog is near the cat and the giraffe is facing the lake*. The constraint propagator can therefore solve these constraints independently.

The next step involves a greedy strategy to solve the constraints in each sub-network. It recursively processes every pair of objects in a relationship with the following (simplified) heuristics, which are based on whether their positions and orientations are set:

1. If neither object is set, then solve the one with the most constraints first, then the other.
2. If one is set, then solve the other.
3. If both are set and satisfy all constraints between them, then they are done.
4. If both are set and either violates a constraint between them, then unset them and start over at a previously solved pair. The restart mechanism uses backtracking to re-solve the previous pair first, then returns to this pair. If this pair fails again, it repeats this process a selectable number of times before abandoning it for the previous pair of the previous pair, and so on, recursively. If no solution can be generated after a selectably large number of attempts, the description is assumed to be inconsistent, and therefore unsatisfiable.

6.2 Simulation and Distillation

The basis of the Monte Carlo simulation is to execute multiple independent runs and record their individual solutions for aggregate analysis. The number of runs is selectable, and 30 to 70 generally produce a reasonable sample size. After a certain point (which varies depending on the description), returns diminish. The simulation time currently precludes any use of this approach in real-time applications.

Analysis is termed *distillation* here because it extracts the essence of the aggregate results into disjoint clusters of similar interpretation. Commonality is determined by averaging the positions and orientations of the objects. For example, Figure 7a distills 12 similar positions into a single representative one. Likewise, Figure 7b distills 10 positions of object *D* into two clusters, one to each side of object *C*. Each cluster can then assume a single representative position, which would be considered the most plausible.

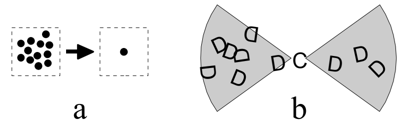

Figure 7: Distillation Clustering

7 Results and Future Work

The results of the simulation feed into a three-dimensional, graphical rendering engine, which dynamically visualizes the distilled (or individual) solutions from any vantage point. Figure 8 shows the solution for *the dog is south of the tree and near the panther; the panther is to the right of the dog; and the elk is near the maple tree and midrange from and facing away from the pond*. This presentation was found to be the easiest way to determine empirically whether ordinary people accept the results. It is intended to support a formal survey and subsequent statistical analysis of plausibility in future work.

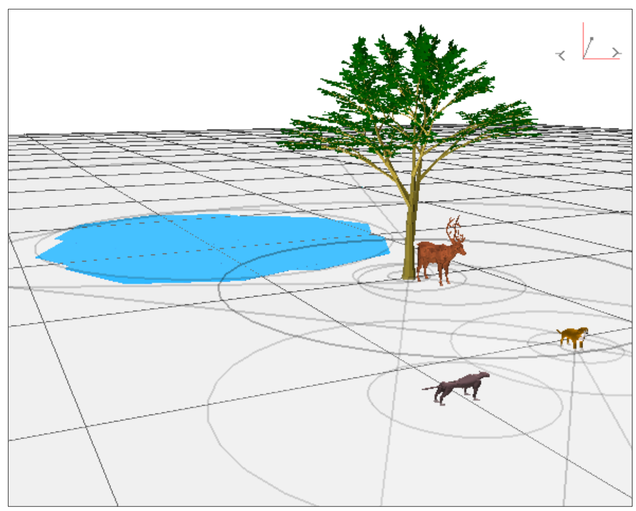

Figure 8: Sample Visualization

Another future application of the simulation framework is to test spatial hypotheses [42]. For instance, does *A near B* imply *B near A*, does *A north of B* and *B north of C* imply *A north of C*, and does *A facing north* and *B facing south* imply *B facing A*? All three of these are actually false, but proving so is not necessarily trivial.

8 References

[1] A. Herskovits. "Language and Spatial Cognition: An interdisciplinary Study of the Prepositions in English". Cambridge University Press, 1986.

[2] R. Srihari. "Computational Models for Integrating Linguistic and Visual Information: A Survey";

Artificial Intelligence Review, Vol. 8, pp. 349-369, 1994.

[3] P. Johnson-Laird. "Mental Models". Harvard University Press, 1983.

[4] R. Langacker. "Foundations of Cognitive Grammar, Volume 1, Theoretical Prerequisites". Stanford University Press, 1987.

[5] B. Claus, K. Eyferth, C. Gips, R. Hörnig, U. Schmid, S. Wiebrock, and F. Wysotzki. "Reference Frames for Spatial Inference in Text Understanding"; in C. Freksa, C. Habel, and K. Wender, eds. Spatial Cognition–An interdisciplinary approach to representing and processing spatial knowledge, No. 1404, pp. 214-226, 1988.

[6] P. Olivier and J. Tsujii. "A computational view of the cognitive semantics of spatial prepositions"; in Proceedings of 32nd Annual Meeting of the Association for Computational Linguistics, Las Cruces, New Mexico, 1994.

[7] B. Kuipers. "Modeling Spatial Knowledge. Cognitive Science"; Vol. 2, pp. 129-153, 1978.

[8] A. Mukerjee. "Neat vs Scruffy: A Survey of Computational Models for Spatial Expressions"; in P. Olivier and K. Gapp, eds. Computational Representation and Processing of Spatial Expressions, 1998.

[9] W. Wahlster. "Text and Images"; in R. Cole, J. Mariana, H. Uszkoreit, A. Zaenen, and V. Zue, eds. Survey of the State of the Art in Human Language Technology. Kluwer, 1996.

[10] S. Dupuy, A. Egges, V. Legendre, and P. Nugues. "Generating a 3D Simulation of a Car Accident from a Written Description in Natural Language: the CarSim System"; in Proceedings of the Workshop on Temporal and Spatial Information Processing, pp. 1-8. Toulouse, France, 2001.

[11] B. Coyne and R. Sproat. "WordsEye: An Automatic Text-to-Scene Conversion System"; in Proceedings of SIGGRAPH-01, pp. 487-496. Los Angeles, CA, 2001.

[12] K. Xu, J. Stewart, and E. Fiume. "Constraint-Based Automatic Placement for Scene Composition"; in Proceedings of the Conference on Human-Computer Interaction and Computer Graphics, pp. 25-34, Calgary, Canada, 2002.

[13] A. Yamada. "Studies on Spatial Description Understanding Based on Geometric Constraints Satisfaction"; Ph.D. dissertation, University of Kyoto, 1993.

[14] L. Seversky and L. Yin. "Realtime Automatic 3D Scene Generation from Natural Language Voice and Text Descriptions"; in Proceedings of 14th annual ACM international conference on Multimedia, Santa Barbara, CA, 2006.

[15] B. Neumann. "Natural Language Description of Time-Varying Scenes"; in D. L. Waltz, ed., Semantic Structures: Advances in Natural Language Processing, pp. 167-207, Lawrence Erlbaum, 1989.

[16] I. Walter, P. Lockemann, and H.-H. Nagel. "Database Support for Knowledge-Based Image Evaluation"; in P. M. Stocker, W. Kent, R. Hammersley, eds., Proceedings of the 13th Conference on Very Large Databases, pp. 3-11, Brighton, UK, 1988.

[17] H. Koller, N. Heinze, H.-H. Nagel. "Algorithmic Characterization of Vehicle Trajectories from Image Sequences by Motion Verbs"; in Proceedings of IEEE Conference. on Computer Vision and Pattern Recognition, pp. 90-95, Maui, Hawaii, 1992.

[18] J. Tsotsos. "Knowledge Organization and its Role in Representation and Interpretation for Time-Varying Data: the ALVEN System"; Computational Intelligence, Vol. 1, pp. 16-32, 1985.

[19] B. Tversky. "Levels and structure of spatial knowledge"; in Cognitive Mapping: Past, present and future, R. Kitchin and S. Freundshuh, eds., Routledge, 2000.

[20] S. Peters and H. Shrobe. "Using Semantic Networks for Knowledge Representation in an Intelligent Environment"; in 1st Annual IEEE International Conference on Pervasive Computing and Communications, Ft. Worth, TX, 2003.

[21] E. Davis. "Representations of Commonsense Knowledge". Morgan Kaufmann, 1990.

[22] M. Egenhofer and R. Franzosa. "Point-Set Topological Spatial Relations"; International Journal of Geographical Information Systems, Vol. 5, No. 2, pp. 161-174, 1991.

[23] A. Frank. "Qualitative Reasoning about Distances and Directions in Geographic Space"; Journal of Visual Languages and Computing, Vol. 3, No. 4, pp. 343-371, 1992.

[24] A. Frank. "Qualitative Spatial Reasoning: Cardinal Directions as an Example"; International Journal of Geographical Information Systems, Vol. 10, No. 3, pp. 269-290, 1996.

[25] D. Hernández, E. Clementini, and P. Di Felice. "Qualitative Distances"; in A. Frank and W. Kuhn, eds., Third European Conference on Spatial Information Theory, pp. 45-58, Semmering, Austria, 1995.

[26] D. Randell, Z. Cui, and A. Cohn. "A Spatial Logic based on Regions and Connection"; in Proceedings of 3rd International Conference on Knowledge Representation and Reasoning, pp. 165-176, San Mateo, CA, 1992.

[27] D. Tappan. "Knowledge-Based Spatial Constraint Satisfaction"; in Proceedings of Florida Artificial Intelligence Research Society International

Conference, Miami Beach, FL, 2004.

[28] G. Adorni, M. Di Manzo, and F. Giunchiglia. "Natural Language Driven Image Generation"; in Proceedings of COLING-84, pp. 495-500, Stanford, CA, 1984.

[29] J. Sowa, ed. "Principles of Semantic Networks: Explorations in the Representation of Knowledge by Computers". Academic Press, 1991.

[30] R. Davis, H. Shrobe, and P. Szolovits. "What is Knowledge Representation?"; AI Magazine, Vol. 14, pp. 17-33, 1993.

[31] K. Mahesh. "Ontology Development for Machine Translation: Ideology and Methodology"; Technical Report MCCS-96-292, Computing Research Laboratory, New Mexico State University, 1996.

[32] J. Freeman. "The Modeling of Spatial Relations"; Computer Graphics and Image Processing, Vol. 4, pp. 156-171, 1975.

[33] D. Bennet. "Spatial and Temporal Uses of English Prepositions". An Essay in Stratificational Semantics. Longman, 1975.

[34] C. Hill. "Up/down, front/back, left/right. A contrastive study of Hausa and English"; in J. Weissenborn and W. Klein, eds. Here and There: Cross-Linguistic Studies of Deixis and Demonstration, John Benjamins, 1982.

[35] L. Talmy. "How Language Structures Space"; in H. Pick and L. Acredolo, eds. Spatial Orientation: Theory, Research, and Application, Plenum Press, 1983.

[36] B. Hawkins. "The Semantics of English Spatial Prepositions"; Ph.D. dissertation, University of California, San Diego, 1984.

[37] A. Yamada, T. Yamamoto, H. Ikeda, T. Nishida, and S. Doshita. "Reconstructing Spatial Image from Natural Language Texts"; in Proceedings of COLING-92, pp. 1279-1283, Grenoble, France, 1992.

[38] K. Gapp. "Basic Meanings of Spatial Relations: Computation and Evaluation in 3D Space"; in Proceedings of AAAI-94, pp. 1393-1398, Seattle, WA, 1994.

[39] C. Freska. "Using Orientation Information for Qualitative Spatial Reasoning"; in A. Frank, I. Campari, and U. Formentini, eds. Theories and Methods of Spatio-Temporal Reasoning in Geographic Space, LNCS 639, Springer-Verlag, 1992.

[40] D. Hernández. "Qualitative Representation of Spatial Knowledge". Springer-Verlag, 1994.

[41] A. Stevens and P. Coupe. "Distortions in Judged Spatial Relations"; Cognitive Psychology, Vol. 13, pp. 422-437, 1978.

[42] D. Tappan. "Knowledge-Based Spatial Reasoning for Automated Scene Generation from Text Descriptions"; Ph.D. dissertation, New Mexico State University, 2004.