Speaking with spatial relations

Lukasz Wawrzyniak*, Dennis Nikitenko
and Pascal Matsakis

Department of Computing and Information Science,
University of Guelph, Guelph, Ontario Canada, N1G 2W1
Fax: (519) 837-0323 E-mail: lwawrzyn@uoguelph.ca
E-mail: dnikiten@uoguelph.ca E-mail: matsakis@cis.uoguelph.ca
*Corresponding author

Abstract: Natural language descriptions are an important step in bridging the gap between numerical representations of spatial data and the human user. In this work, we present a system for generating linguistic descriptions of the spatial relationships between two-dimensional objects. The most pertinent relations for the description are chosen based on a fuzzification of the set relations DISJOINT, OVERLAP, SUBSET, SUBSETi and EQUAL. A handful of relevant Allen relations is then selected and their Allen F-histograms are analysed to extract further topological and directional information. The approach is validated using several sets of real and synthetic data.

Keywords: spatial relationships; topological relations; Allen relations; set relations; directional relations; natural language; linguistic descriptions; F-histograms; fuzzy set theory; scene understanding; computer vision; geographical information systems; GIS.

Reference to this paper should be made as follows: Wawrzyniak, L., Nikitenko, D. and Matsakis, P. (2006) 'Speaking with spatial relations', *Int. J. Intelligent Systems Technologies and Applications*, Vol. 1, Nos. 3/4, pp.280–300.

Biographical notes: Lukasz Wawrzyniak is a MSc student at the Department of Computing and Information Science, University of Guelph, Ontario, Canada. He received a BSc Degree in Applied Computer Science from Ryerson University, Toronto, Ontario. His research interests include spatial and spatio-temporal analysis, spatial relationship representation models, fuzzy set theory and fuzzy logic, as well as image analysis and computer graphics.

Dennis Nikitenko is a PhD student at the University of Guelph. He received his MSc in computer science from the University of Guelph, Guelph, Canada, and BSc in computer science from Ryerson University, Toronto, Canada. He worked in a research group at Princess Margaret Hospital's Ontario Cancer Institute (Toronto, Canada) developing software for ultrasound imaging. His research interests lie in natural computation and its applications of image and signal analysis. He specialises in medical imaging applications and his current research focuses on applications of digital imaging to mammography screening.

Pascal Matsakis is an Associate Professor of Computing and Information Science at the University of Guelph, Ontario, Canada, since August 2002. He was member of the Computational Intelligence Research Laboratory (CIRL) at the University of Missouri-Columbia, Missouri, USA, from

September 1999 to July 2002. He received the PhD in Computer Science (1998) from Paul Sabatier University, Toulouse, France. His interests include computer vision, pattern recognition, fuzzy set theory and fuzzy logic. His particular field is in modelling and utilising the spatial relations between multidimensional objects. He is an Associate Editor of the *IEEE Transactions on Fuzzy Systems*.

1 Introduction

In recent years, the problem of expressing the spatial relationships between two-dimensional objects using natural language has received considerable attention in many areas of computer science. The ability to accurately and succinctly describe the spatial configurations of objects can enhance applications such as mobile robot navigation, target recognition and acquisition, digital medical image diagnostics, landmark navigation, image database search and retrieval and numerous GIS applications.

In this paper, we propose a method for generating linguistic descriptions of the topological and directional relationships between two objects. The descriptions are based on the fuzzification of the set relations DISJOINT, OVERLAP, SUBSET, SUBSETi and EQUAL, as well as on information extracted from Allen F-histograms (Matsakis and Nikitenko, 2005). The descriptions convey the primary topological relationships between the objects and provide directional estimates of where these relations hold true. The paper is organised as follows. In Section 2, we survey existing methods of associating linguistic expressions with spatial configurations of objects and review the notion of Allen F-histograms, which play an important role in the proposed approach. Section 3 deals with how the linguistic descriptions are generated. Experimental results can be found in Section 4. A conclusion and directions of future work appear in Section 5.

2 Related work

2.1 Spatial relationships and natural language

Several methods have been recently proposed to derive linguistic expressions from numerical or symbolic representations of the spatial relationships between objects. For instance, a fuzzy rule-based system was proposed by Keller and Wang (2000). Based on a method described in Wang and Keller (1997, 1999), some numeric parameters characterising spatial relationships were extracted from a segmented scene. A system of fuzzy rules then associated natural language descriptions with these numeric values, resulting in descriptions such as “There are five missile launchers (1,2,3,6,8). They surround a centre vehicle (4) …” or “The roof is right of the tree. The wall is right of the tree …” Abella and Kender (1999) represented spatial relationships as fuzzy predicates and used them to produce natural language statements about location and space. Their system consisted of an image processing module, a fuzzy semantic representation module and natural language generation modules. The semantic representation module integrated visual information (extracted from the image) with linguistic information (gathered for the test set from human test subjects). The final system was tested on landmark navigation in spatial images (the example is 'set' in Disney World), generating descriptions such as

> "First, find the Chinese Pavilion which is the leftmost one near the German Pavilion. Then, find the Norwegian Pavilion, which is the leftmost one next to the Chinese Pavilion. The Mexican Pavilion is the topmost one below the Norwegian Pavilion."

The system was also tested on radiograph images to find abnormal densities, generating descriptions such as "The right kidney contains a density which may represent a right rental stone". A system that could analyse spatial relationships in a visual scene and connect them to appropriate linguistic descriptions was proposed by Kopp (1994). This system extracted the objects from the scene, classified them using a pattern recognition neural network and then used a correlation matrix to connect the text to the outputs of the network. It produced linguistic descriptions such as "frog under fly", "frog left of cat" or "cat right of frog under fly over dog".

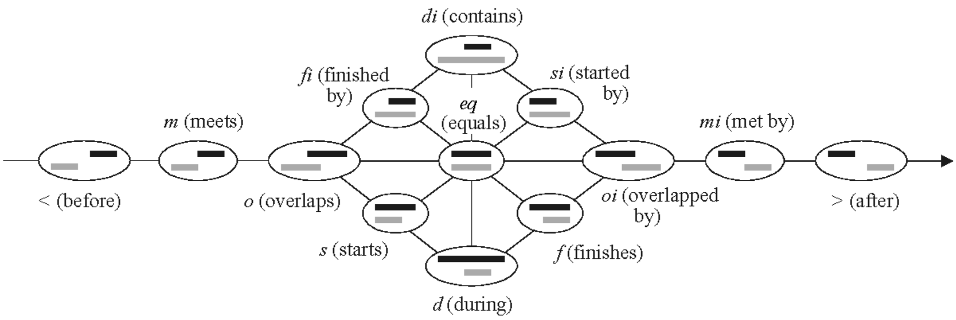
Skubic et al. (2002) generated linguistic spatial descriptions from an evidence grid map in the context of human-robot dialog. The map was used to represent objects in the robot's environment, and the spatial descriptions were generated using histograms of forces (Matsakis, 1998; Matsakis and Wendling, 1999) and a method introduced in Matsakis et al. (2001). A linguistic description output by the robot consisted of three parts: the primary component (e.g., "the object is in front of me"), the secondary component (e.g., "but somewhat to my left") and the self-assessment component (e.g., "the description is satisfactory"). Using R-histograms (Wang and Makedon, 2003), Wang et al. (2004) generated fuzzy semantic metadata describing spatial relationships. The image was segmented, and R-histograms provided a quantitative representation of spatial relationships between image regions. A fuzzy classifier was used for each of the four spatial relation pairs 'left of/right of', 'above/below', 'near/far' and 'inside/outside'. The system stopped short of generating complete natural language descriptions of the scene. A method that provided linguistic descriptions of topological relationships was given by Zhan (2001, 2002). Each object was represented as a fuzzy set (an object with fuzzy boundaries). As in Zhan (1997, 1998), the fuzzy membership value of a given topological relation between two fuzzy regions was computed based on the 9-Intersection model (Egenhofer and Herring, 1994). The results were demonstrated using a single topological relation, 'covers', and included descriptions such as "Region Q covers Region R a little bit" or "Region Q nearly completely covers Region R".

A comprehensive system for generating linguistic descriptions of the topological relationships between 2-D objects was presented in Matsakis et al. (submitted). The system generated descriptions based on the 13 Allen F-histograms (Matsakis and Nikitenko, 2005). Although satisfactory descriptions were generally obtained, some counter-intuitive descriptions were also produced. The descriptions were based on the Allen relations present along a direction of major object interaction. Ultimately, only a single direction was considered, and in some cases, important topological information was not taken into account. Wawrzyniak et al. (2005) proposed some simple solutions to eliminate the counter-intuitive descriptions. In this paper, we revisit the work presented in Matsakis et al. (submitted) and Wawrzyniak et al. (2005), and propose a comprehensive system which corrects the problems encountered.

2.2 Allen F-histograms

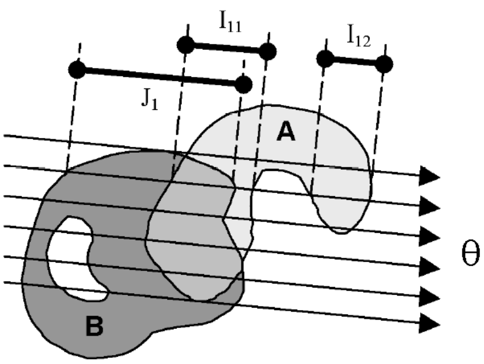
Allen (1983) introduced a set A of 13 mutually exclusive and collectively exhaustive relations to represent knowledge about time intervals: $\mathcal{A} = \{<, m, o, s, fi, d, eq, di, si, f, oi, mi, >\}$. The relations are illustrated on a conceptual neighbourhood graph in Figure 1. One possible way to extend Allen relations into the spatial domain was suggested by Matsakis and Nikitenko (2005) in the form of Allen F-histograms.

Figure 1 Conceptual neighbourhood graph of Allen relations



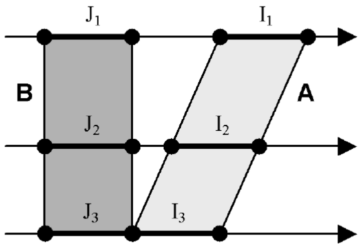
The notion of the F-histogram was first introduced in Matsakis (1998). F-histograms include force histograms (Matsakis, 1998; Matsakis and Wendling, 1999) and Allen F-histograms (Matsakis and Nikitenko, 2005). Consider a pair of objects (A,B) and an Allen relation *r*. The Allen F-histogram F_r^{AB} is one possible representation of the position of A (the *argument* object) with regard to B (the *referent* object). It is a numeric function. For any direction θ, the value $F_r^{AB}(\theta)$ is a weight attached to the proposition "A *r* B in direction θ". This weight is computed by decomposing A and B into pairs of aligned *longitudinal sections* (Figure 2). Thus, the handling of the object pair (A,B) comes down to the handling of pairs of longitudinal sections, and the handling of each pair of longitudinal sections comes down to the handling of pairs of segments for which the Allen relation *r* can be readily assessed.

Figure 2 For each direction, the plane is partitioned into a set of parallel lines. The intersection of a line and an object is a longitudinal section of that object. J_1, for instance, is a longitudinal section of B. It consists of one segment. I_1 is a longitudinal section of A and consists of two disjoint segments: I_{11} and I_{12}



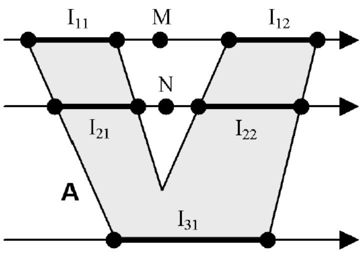
The Allen relations are fuzzified such that any relation r in $\mathcal{A}$ is a continuous function onto [0,1]. The idea is that if an object is moved slightly, then the Allen F-histograms should change equally slightly. Moreover, for any pair (I, J) of segments on an oriented line, $\sum_{r\in\mathcal{A}} r(\mathrm{I}, \mathrm{J}) = 1$. For any pair (I, J) and any r_1 and r_2 in $\mathcal{A}$, if $r_1(\mathrm{I}, \mathrm{J}) \neq 0$ and $r_2(\mathrm{I}, \mathrm{J}) \neq 0$ then r_1 and r_2 are direct neighbours in the graph of Figure 1. An illustrative example of the fuzzification of Allen relations is presented in Figure 3.

Figure 3 Fuzzy Allen relations. Here, $>(\mathrm{I}_1,\mathrm{J}_1)$ is 1 and $mi(\mathrm{I}_1,\mathrm{J}_1)$ is 0; $>(\mathrm{I}_2,\mathrm{J}_2)$ and $mi(\mathrm{I}_2,\mathrm{J}_2)$ are greater than 0 and less than 1 and $>(\mathrm{I}_3,\mathrm{J}_3)$ is 0 and $mi(\mathrm{I}_3,\mathrm{J}_3)$ is 1



The longitudinal sections are also fuzzified. Thus, if an object is deformed slightly, the Allen F-histograms change equally slightly. Fuzzification is achieved such that closer the two segments of a crisp longitudinal section, the more the space in between belongs to the fuzzified section. When sufficiently close, the two segments are seen, to a certain degree, as a single segment. An illustrative example of the fuzzification of longitudinal sections is presented in Figure 4. The processing of a pair $(\mathrm{I}_k,\mathrm{J}_k)$ of longitudinal sections, i.e., the computation of $r(\mathrm{I}_k,\mathrm{J}_k)$, is achieved by processing the α-cuts $\mathrm{I}_\mathrm{k}^\alpha$ and $\mathrm{J}_\mathrm{k}^\alpha$ and blending the $r(\mathrm{I}_{ki}^\alpha,\mathrm{J}_{kj}^\alpha)$ values appropriately for all α, i and j.

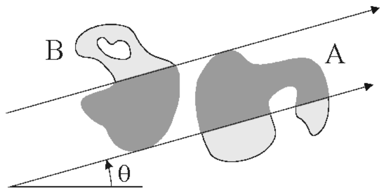
Figure 4 Fuzzy longitudinal sections. Before fuzzification, the point M does not belong at all to $\mathrm{I}_1 = \mathrm{I}_{11} \cup \mathrm{I}_{12}$ and N does not belong at all to $\mathrm{I}_2 = \mathrm{I}_{21} \cup \mathrm{I}_{22}$. After fuzzification of I_1 and I_2, the point M belongs to I_1 to some extent, and N belongs to I_2 more than M belongs to I_1



Finally, $\mathrm{F}_r^{\mathrm{AB}}(\theta)$ is a weighted sum of the $r(\mathrm{I}_k, \mathrm{J}_k)$ values for all k. It is such that $\sum_{r\in\mathcal{A}} \mathrm{F}_r^{\mathrm{AB}}(\theta)$ measures the object interaction in direction θ. To put it simply, $\sum_{r\in\mathcal{A}} \mathrm{F}_r^{\mathrm{AB}}(\theta)$ is the total area of the regions of A and B that are 'facing' each other in direction θ (Figure 5). Again, a slight change in the object configuration results in a correspondingly slight change in the histograms. Continuity is satisfied and, hence, robustness is achieved. Should the metric information be judged unimportant, the histograms can be normalised:

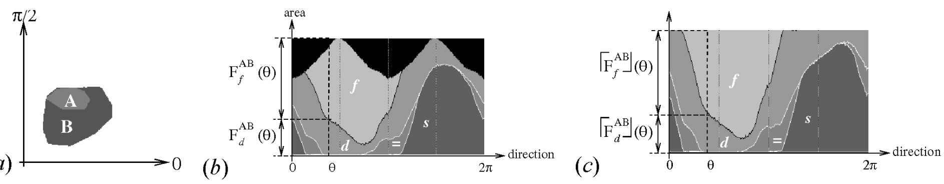
$$\forall \theta, \lceil F_r^{AB} \rfloor (\theta) = F_r^{AB} (\theta) / \sum_{\rho \in \mathcal{A}} F_\rho^{AB} (\theta).$$

Figure 5 $\sum_{r \in \mathcal{A}} F_r^{AB}(\theta)$ measures to what extent A and B are involved in some spatial relationship along direction θ. It is the total area of the dark grey regions



Examples of Allen F-histograms and their normalised counterparts are shown in Figure 6. All details pertaining to the computation of Allen F-histograms are presented in (Matsakis and Nikitenko, 2005).

Figure 6 A pair of objects (a) and their corresponding histograms: non-normalised (b) and normalised (c). In both (b) and (c), the 13 Allen F-histograms are 'piled up' on top of each other



Many numerical values can be extracted from a given Allen F-histogram F_r^{AB}. Here, we focus on two of them: the *primary direction* θ_r and the *directional acuteness* a_r (Matsakis et al., submitted; Wawrzyniak et al., 2005). Put simply, θ_r is the direction where r is most prominent. It is the 'average' direction where the relation r occurs. The value a_r conveys the *acuteness* of the directional relationship. If a_r is high, then the relation r occurs on a small range of directions near θ_r. If it is low, then r occurs on a wide range of directions, or it occurs along contradictory directions. θ_r and a_r will allow us to generate *directional estimates* as part of the linguistic description (Section 3.2.2).

3 Linguistic descriptions

The descriptions are generated using fuzzy models of set relationships between objects. A handful of relevant Allen relations may be selected to convey more detailed topological information, and directional estimates of where these relations are most prominent may also be provided.

3.1 Fuzzy set relations between 2-D regions

Numerous models have been developed to represent the topological relationships between two-dimensional regions. In GIS applications, boundary/area intersection models have gained a wide following. In the context of image analysis, however, it is often the case that no explicit boundary information is available. Image regions are typically defined as sets of pixels. Such raster objects can take arbitrarily complex shapes, and it may be difficult or impractical to algorithmically define a region boundary. Here, we focus on five set relations between 2-D objects, where each object is defined as a non-empty, finite set of pixels. These relations are defined in Table 1.

Table 1 Five mutually exclusive and collectively exhaustive set relations

Relation	*Abbr.*	*Symmetric*	$A \cap B \neq \varnothing$	$A - B \neq \varnothing$	$B - A \neq \varnothing$
DISJOINT(A,B)	DIS	Yes	0	1	1
OVERLAP(A,B)	OVE	Yes	1	1	1
SUBSET(A,B)	SUB	No	1	0	1
SUBSET(B,A)	SUBi	No	1	1	0
EQUAL(A,B)	EQ	Yes	1	0	0

Note that, in this table, the relation SUBSET represents *proper* subsethood. Let $\mathcal{R}$ = {DIS,OVE,SUB,SUBi,EQ}. The set relations in $\mathcal{R}$ are mutually exclusive and collectively exhaustive. It is quite clear, however, that crisp definitions of these relations are not very practical. Image acquisition and segmentation algorithms are susceptible to errors, often resulting in missing or superfluous pixels attributed to some object. Suppose we have two objects A and B which are, in reality, equal (EQ(A,B) = 1, SUB(A,B) = 0 and SUBi(A,B) = 0). Owing to an acquisition error, a single missing pixel in A would change the relation from EQ to SUB (EQ(A,B) = 0, SUB(A,B) = 1, and SUBi(A,B) = 0). On the other hand, a superfluous pixel in A would change the relation from EQ to SUBi (EQ(A,B) = 0, SUB(A,B) = 0 and SUBi(A,B) = 1). Minor errors can affect the crisp relations between 2-D regions drastically. To robustly handle real-world data, fuzzification of the relations is necessary. By defining the relations in $\mathcal{R}$ as continuous functions onto [0,1], the fuzzification proposed here achieves continuous transitions between the relations and allows us to discern between various degrees to which a certain relation is present. For instance, we could have two objects which are 'mostly' DISJOINT, but which also 'slightly' OVERLAP.

Let $x = |A \cap B|/|A|$ be the degree of subsethood of A in B and $y = |A \cap B|/|B|$ be the degree of subsethood of B in A. In order to fuzzify the relations in $\mathcal{R}$, we propose to use the following membership functions:

$$\text{DIS(A,B)} = \min[lo(x), lo(y)]$$

$$\text{OVE(A,B)} = \max\{\min[mid(x), 1 - hi(y)], \min[mid(y), 1 - hi(x)]\}$$

$$\text{EQ(A,B)} = \min[hi(x), hi(y)]$$

$$\text{SUB(A,B)} = hi(x) - \min[hi(x), hi(y)]$$

$$\text{SUBi(A,B)} = hi(y) - \min[hi(x), hi(y)]$$

These membership functions are illustrated in Figure 7. A contour map at 0.5 is shown in Figure 8. The functions *lo*, *mid* and *hi* are as in Figure 9.

Figure 7 Fuzzified set relations between 2-D regions. (a): DIS(A,B); (b): SUB(A,B); (c): SUBi(A,B); (d): EQ(A,B); (e) and (f): two views of OVE(A,B)

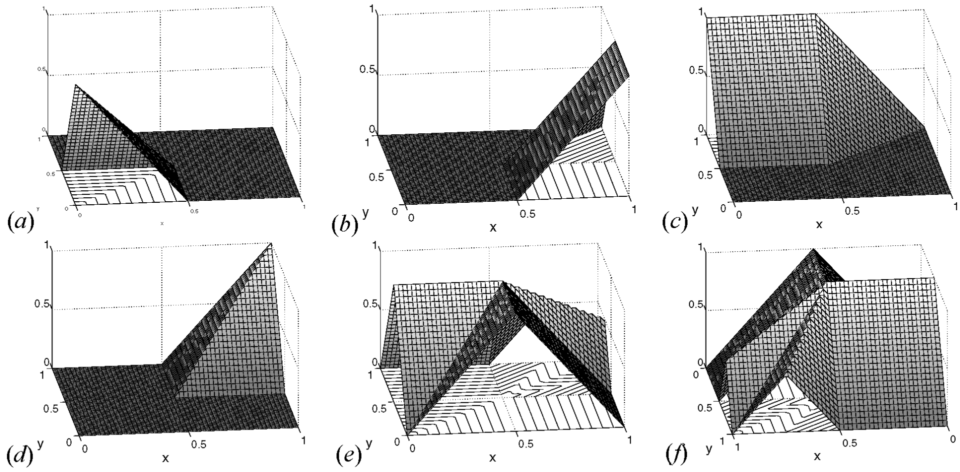


Figure 8 A contour map of the membership functions from Figure 7 at 0.5, showing where each function dominates

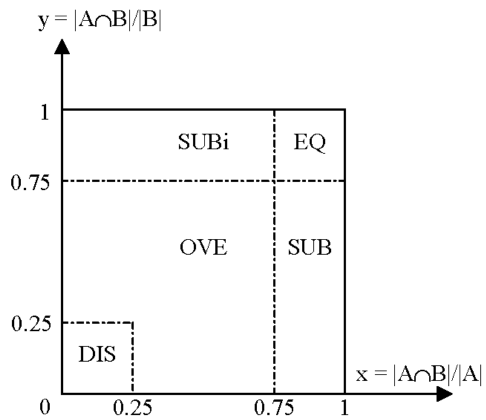
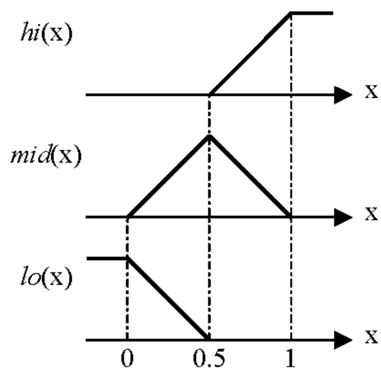
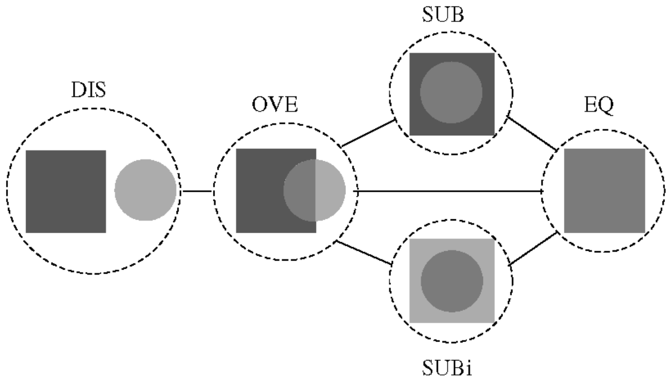


Figure 9 The functions *lo*, *mid* and *hi*



Consider a pair of objects (A,B). For any set relation r in $\mathcal{R}$, we have $r(\text{A,B}) \in [0,1]$. Boundary conditions are preserved, i.e., it is possible to have $r(\text{A,B}) = 0$, and it is possible to have $r(\text{A,B}) = 1$. Also, $\sum_{r\in\mathcal{R}} r(\text{A,B}) = 1$. Furthermore, for any r_1 and r_2 in $\mathcal{R}$, if $r_1(\text{A,B}) \neq 0$ and $r_2(\text{A,B}) \neq 0$, then r_1 and r_2 are direct neighbours on the conceptual neighbourhood graph shown in Figure 10. Let $\text{V}^{\text{AB}} = \{r\in\mathcal{R} \mid r(\text{A,B}) > 0\}$. The set V^{AB} contains at least one and at most three relations. In fact, there are 13 possible values for V^{AB}: {DIS}, {OVE}, {SUB}, {SUBi}, {EQ}, {DIS,OVE}, {OVE,SUB}, {OVE,SUBi}, {OVE,EQ}, {SUB,EQ}, {SUBi,EQ}, {OVE,SUB,EQ} and {OVE,SUBi,EQ}.

Figure 10 Conceptual neighbourhood graph of R



Given the relations in V^{AB}, we can begin to think about generating a simple linguistic description of the spatial relationships between the two objects A and B. For instance, if V^{AB} = {SUB}, we could say "A is perfectly contained by B". If V^{AB} = {DIS,OVE}, with DIS(A,B) = 0.7 and OVE(A,B) = 0.3, we could say "A is mostly disjoint from but somewhat overlaps B". In some applications, such descriptions might be considered perfectly sufficient. In other applications, however, the user might benefit from more detailed linguistic descriptions. Consider the first example, "A is perfectly contained by B". This statement tells the user nothing about the inner-adjacency of A to B. If A lies near the eastern border of B, a statement such as "A starts B to the east" could be included to supplement the description. This information can be obtained by extracting the values θ_s and a_s from the F-histogram of the Allen relation s (Section 2.2). The second example, "A is mostly disjoint from but somewhat overlaps B", could be augmented with directional estimates such as "A is before B to the north" and "A overlaps B to the northeast". Again, this information can be readily obtained from the F-histograms of the Allen relations < and o.

The general idea is that depending on the relations found in V^{AB}, we can seek further information in the relevant Allen F-histograms. The exact approach taken to generate a complete linguistic description varies with the contents of V^{AB}. In Section 3.3, we examine the different approaches on a case-by-case basis. First, however, we shall introduce several concepts that will be useful in generating the descriptions.

3.2 Elements of a linguistic description

3.2.1 Predominance and prominence

Let A and B be two objects. In cases where a single relation r is sufficient to describe the spatial relationships between A and B, the system generates a statement of the form:

"A perfectly r B",

e.g., "A perfectly contains B" or "A is perfectly equal to B". However, it is often the case that we need to convey the relative proportion of two set relations using natural language. Suppose that we have two relations r_1, $r_2 \in \mathcal{R}$ ordered such that $r_1(A,B) \geq r_2(A,B) > 0$. The *predominance* of r_1 over r_2 is given by:

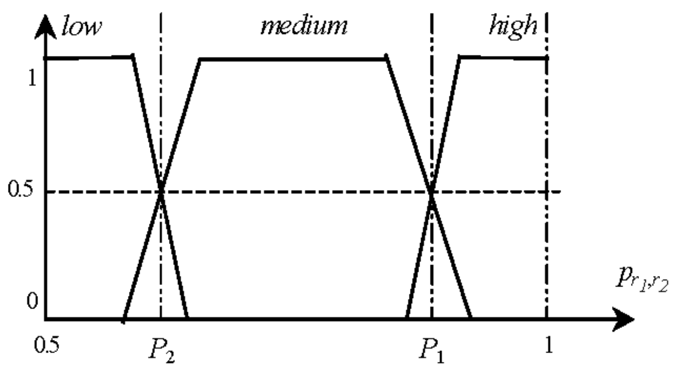
$$p_{r_1,r_2} = r_1(\mathrm{A,B})/[r_1(\mathrm{A,B}) + r_2(\mathrm{A,B})].$$

A statement about the relative proportion of r_1 and r_2 takes the form:

"A [⟨*adverb*$_1$⟩] r_1 ⟨*connective*⟩ [⟨*adverb*$_2$⟩] r_2 B",

e.g., "A contains but is marginally equal to B" or "A is disjoint from as much as it overlaps B". The choice of adverbs and connectives is made based on the value of p_{r_1,r_2}. One mapping between p_{r_1,r_2} and some linguistic terms is shown in Figure 11 as a simple fuzzy rule base.

Figure 11 Rules for generating statements about the relative proportion of two relations. In our experiments, $P_2 = 0.63$ and $P_1 = 0.88$

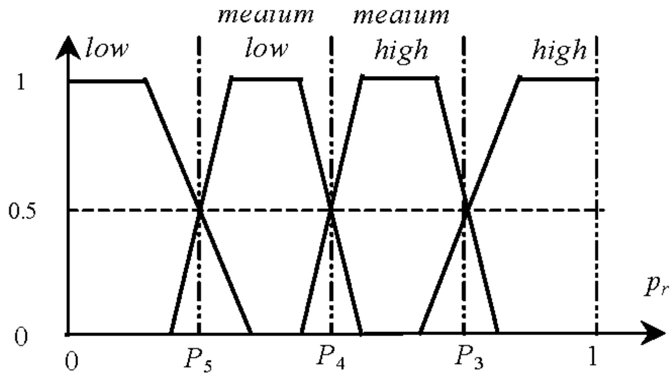


	p_{r_1,r_2}		
	Low	*Medium*	*High*
Adverb$_1$	–	Mostly	–
Connective	As much as	But	But
Adverb$_2$	–	Somewhat	Marginally

In some cases, we may wish to express how prominent a particular Allen relation r is, independently of any other relation. We will see in Section 3.3 how the *degree of prominence* p_r of r is computed. Suppose, for example, that $p_m = 0.35$. According to

Figure 12, which shows one possible mapping from the interval [0,1] onto a set of linguistic terms, we could say that "A *slightly* meets B".

Figure 12 Rules for generating statements about prominence. In our experiments, $P_5 = 0.25$, $P_4 = 0.5$ and $P_3 = 0.75$



	p_r			
	Low	*Medium low*	*Medium high*	*High*
Linguistic term	Very slightly	Slightly	Strongly	Very strongly

3.2.2 Directional estimates

A directional estimate of some Allen relation r is a linguistic statement describing where r is most prominent. Such an estimate is generated based on the directional acuteness a_r (Section 2.2) and a linguistic quantisation of the primary direction θ_r. Directional estimates take the form:

"A r B [⟨*adverb*$_r$⟩] to the $\theta_r^{\star}$",

where the adverb is chosen based on the value of a_r. A possible mapping from a_r to the appropriate adverb is shown in Figure 13. Generally, whenever a_r is *low*, the system will either output "A r B in multiple directions" or omit the directional estimate altogether. The symbol $\theta_r^{\star}$ denotes one of the 16 compass points N, S, E, W, NE, NW, SE, SW, NNE, NNW, SSE, SSW, ENE, ESE, WNW and WSW. It is the compass point that the angle θ_r coincides with best. A complete directional estimate might be "A meets B to the west" or "A overlaps B loosely to the northeast".

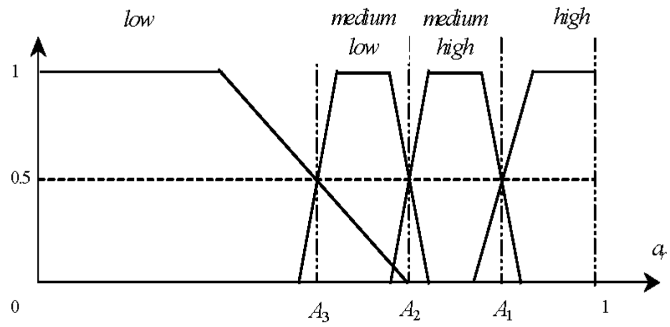
3.3 Generating descriptions on a case-by-case basis

In this subsection, we present the approaches to generating complete linguistic descriptions based on all possible values of the set V^{AB}.

3.3.1 Perfect equality ($V^{AB} = \{EQ\}$)

This is the trivial case. The statement "A is perfectly equal to B" is generated, and nothing remains unsaid.

Figure 13 Rules for generating directional estimates. In our experiments, A_3 is 0.5, A_2 is 0.67 and A_1 is 0.83



	a_r			
	*Low**	*Medium low*	*Medium high*	*High*
Adverb	N/a	Loosely	Primarily	–

*Generally, a special action is taken if the acuteness drops to low (Section 3.2.2).

3.3.2 *Perfect proper subsethood (V^{AB} = {SUB} or V^{AB} = {SUBi})*

If V^{AB} = {SUB}, then A is a proper subset of B and the description takes one of three possible forms, depending on the prominence of the Allen relation *s* (*starts*) and the directional acuteness a_s. The prominence of *s* can be measured with the quantity p_s defined as:

$$p_s = \max_\theta \, [\mathrm{F}_s^{\mathrm{AB}}(\theta) / \max_{\theta'} \sum_{r \in \mathcal{A}} \mathrm{F}_s^{\mathrm{AB}}(\theta')].$$

p_s takes a value on [0,1] and measures the degree to which the Allen relation *s* dominates along a direction of major object interaction.

Case 1: $p_s = 0$ or $a_s < A_3$

The relation *s* is not present at all or its directional acuteness drops below a reasonable value (see parameter A_3 in Figure 13). The description "A is perfectly contained by B" is generated.

Case 2: $p_s \in (0,1)$ and $a_s \geq A_3$

The relation *s* is present to some intermediate degree and a directional estimate is reasonable. The description "A is perfectly contained by and ⟨*adverb*⟩ starts B" is generated, followed by a directional estimate of *s*. The ⟨*adverb*⟩ is chosen based on the value p_s and the mapping in Figure 12.

Case 3: $p_s = 1$ and $a_s \geq A_3$

The relation *s* completely dominates the spatial relationships along a direction of maximum object interaction and a directional estimate is reasonable. The description "A perfectly starts B" is generated, followed by a directional estimate of *s*.

In cases where $V^{AB} = \{SUBi\}$, the descriptions are identical in format, but the relations *contained by* and *starts* are replaced with *contains* and *started by*, respectively. In the remainder of this paper, similar substitutions are assumed for cases involving the set relations SUB and SUBi.

3.3.3 Perfect overlap ($V^{AB} = \{OVE\}$)

First, the statement "A perfectly overlaps B" is generated. If a directional estimate of the Allen relation *o* is reasonable ($a_o \geq A_3$), it follows the opening statement.

3.3.4 Perfect disjointness ($V^{AB} = \{DIS\}$)

Similarly to cases of perfect subsethood (Section 3.3.2), the description takes one of three possible forms, depending on the prominence of the Allen relation *m* (*meets*). The prominence of *m* is defined as:

$$p_m = \max_\theta [F_m^{AB}(\theta) / \max_{\theta'} (F_<^{AB}(\theta') + F_m^{AB}(\theta'))].$$

Case 1: $p_m = 0$

The relation *m* is not present at all. The description "A is perfectly disjoint from B" is generated, followed by a directional estimate of the Allen relation < (*before*).

Case 2: $p_m \in (0,1)$

The relation *m* is present to some intermediate degree. The description "A is perfectly disjoint from and ⟨*adverb*⟩ meets B" is generated, followed by directional estimates of < and *m*. The ⟨*adverb*⟩ is chosen based on the value p_m and the mapping in Figure 12.

Case 3: $p_m = 1$

The relation *m* completely dominates the spatial relationships along a direction of maximum object interaction. The description "A perfectly meets B" is generated, followed by a directional estimate of *m*.

3.3.5 Overlap and disjointness ($V^{AB} = \{OVE,DIS\}$)

These are intermediate cases between perfect disjointness and perfect overlap (the objects overlap 'slightly'). First, a statement about the relative proportion of the relations DIS and OVE is generated (Section 3.2.1). To provide more detailed information, three Allen relations are of interest: <, *m* and *o*. The simplest way to handle these cases is to provide all three directional estimates. In the interest of brevity, however, it is not always necessary to do so. The condition of overlap is of primary interest here, and it seems reasonable to always give a directional estimate for the Allen relation *o*. Directional estimates for the relations < and *m*, however, need only be given if they actually provide additional information, i.e., the estimates differ from the estimate of *o*. Some examples of where these estimates are redundant are shown in Table 3 (c,d,e). Examples of where these estimates are useful are shown in Table 3 (f,g,h).

3.3.6 *Subsethood and equality (V^{AB} = {SUB,EQ} or V^{AB} = {SUBi,EQ})*

These cases occur whenever one of the objects is fully contained by the other and the objects are nearly equal. If V^{AB} = {SUB,EQ}, a statement about the relative proportion of the relations SUB and EQ is generated. As with cases of perfect proper subsethood, the prominence of the Allen relation *s* is examined to determine whether inner-adjacency exists. Here, p_s is given by:

$$p_s = \max_\theta [F_s^{AB}(\theta) / \max_{\theta'} \sum_{r \in \mathcal{A}-\{eq\}} F_r^{AB}(\theta')].$$

If $p_s > 0$ and $a_s \geq A_3$, the directional estimate of *s* follows as the final statement of the description.

3.3.7 *Subsethood and overlap (V^{AB} = {SUB,OVE}, V^{AB} = {SUBi,OVE}, or V^{AB} = {EQ,OVE})*

These cases occur whenever one object is partially contained by the other, but 'spills out' beyond the boundary of its container. A statement about the relative proportion of the set relations (SUB | SUBi | EQ)[1] and OVE is generated. If a directional estimate of the Allen relation *o* is reasonable ($a_o \geq A_3$), it follows the opening statement.

3.3.8 *Subsethood, equality and overlap (V^{AB} = {SUB,EQ,OVE} or V^{AB} = {SUBi,EQ,OVE})*

As above, these cases occur whenever one object is partially contained by the other, but 'spills out' beyond the boundary of its container. The difference is that here the objects are also somewhat equal. A statement describing the relative proportion of (SUB | SUBi) and OVE is generated, followed by the statement "The objects are ⟨*adverb*⟩ equal". The ⟨*adverb*⟩ is chosen based on the value EQ(A,B) and the rules in Figure 12. A directional estimate of the Allen relation *o*, if reasonable, follows as the last statement of the description.

4 Experimental results

Six series of experiments were performed to illustrate some important characteristics of the descriptions. The results can be found in Tables (2–7) and include images of the object configurations and the corresponding descriptions. In all the images, the argument object (A) is shown in light grey and the referent object (B) is shown in dark grey. Areas of overlap appear in medium grey.

Table 2 deals with cases of disjointness and adjacency. Configurations (a–d) show how the directional assessment changes with the acuteness of the Allen relation < (*before*). As the argument object begins to surround the referent, the directional estimates become less and less specific. Configurations (e–j) illustrate a gradual transition from perfect disjointness to perfect adjacency.

Table 2 Disjointness and adjacency dataset

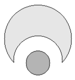
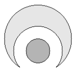
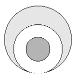
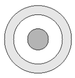
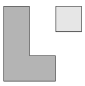
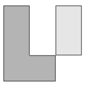
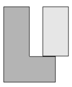
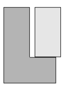
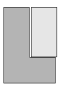
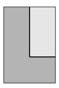
(a)		A is perfectly disjoint from B A is before B to the north
(b)		A is perfectly disjoint from B A is before B primarily to the north
(c)		A is perfectly disjoint from B A is before B loosely to the north
(d)		A is perfectly disjoint from B A is before B in multiple directions
(e)		A is perfectly disjoint from B A is before B to the northeast
(f)		A is perfectly disjoint from and very slightly meets B A is before B to the east-northeast A meets B from the northeast
(g)		A is perfectly disjoint from and slightly meets B A is before B to the east-northeast A meets B from the north-northeast
(h)		A is perfectly disjoint from and strongly meets B A is before B to the east A meets B from the northeast
(i)		A is perfectly disjoint from and very strongly meets B A is before B to the east A meets B from the northeast
(j)		A perfectly meets B A meets B from the east-northeast

Table 3 focuses on overlap, adjacency and disjointness. Configurations (a) and (b) show perfectly overlapping object pairs. According to the definition of the set relation OVE (Section 3.1), in both cases we have OVE(A,B) = 1. In the former, a directional estimate for the Allen relation *o* can be readily generated. In the latter, however, it is unclear from which direction A overlaps B, and no directional estimate is produced. Configurations (c–h) contain objects whose spatial relationships fall somewhere between perfect disjointness and perfect overlap (V^{AB} = {DIS,OVE}). For configurations (c–e), the only directional estimate given is for the Allen relation *o*. Although the relations < and *m* are present to some degree, they occur along the same primary direction as *o*, and their directional estimates would be somewhat redundant. The statement "A overlaps B to the east" says it all. Configurations (f–h), on the other hand, are cases where additional directional estimates give useful information. If these estimates were omitted, important topological and directional information would have been lost.

Table 3 Overlap, adjacency and disjointness dataset

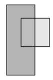
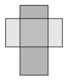
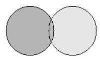
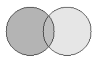
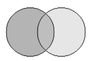
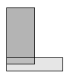
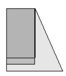
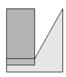
(a)		A perfectly overlaps B A overlaps B from the east
(b)		A perfectly overlaps B
(c)		A is disjoint from but marginally overlaps B A overlaps B from the east
(d)		A is mostly disjoint from but somewhat overlaps B A overlaps B from the east
(e)		A is disjoint from as much as it overlaps B A overlaps B from the east
(f)		A overlaps as much as it is disjoint from B A overlaps B from the south-southeast A is before B to the southeast
(g)		A is mostly disjoint from but somewhat overlaps B A meets B from the east A overlaps B from the south-southeast
(h)		A is mostly disjoint from but somewhat overlaps B A is before B to the east A meets B from the southeast A overlaps B from the south-southeast

Table 4 shows cases of proper subsethood. Configurations (a–f) represent a transition between a state of no inner-adjacency to a state of perfect inner-adjacency. Configurations (g) and (h) are cases where the argument object is inner-adjacent to the referent in multiple places. Since no reasonable directional estimates for the Allen relation *s* can be given, the estimates are suppressed (Section 3.3.2). Whether this constitutes a loss of important spatial information or not is, of course, a rather subjective matter. The method presented here is only one possible way to describe such configurations. Alternatively, a description of the form "A is perfectly contained by B and starts B in multiple directions" could be generated.

Table 4 Proper subsethood dataset

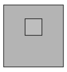
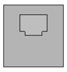
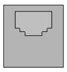
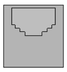
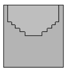
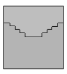
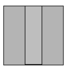
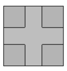
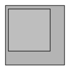
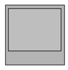
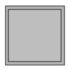
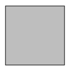
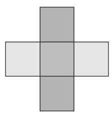
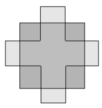
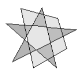
(a)		A is perfectly contained by B
(b)		A is perfectly contained by and very slightly starts B A starts B to the north
(c)		A is perfectly contained by and slightly starts B A starts B to the north
(d)		A is perfectly contained by and strongly starts B A starts B to the north
(e)		A is perfectly contained by and very strongly starts B A starts B to the north
(f)		A perfectly starts B A starts B to the north
(g)		A is perfectly contained by B
(h)		A is perfectly contained by B

Table 5 depicts the first seven cases from a dataset dealing with varying degrees of subsethood, equality and overlap. Configurations (a–d) are cases where the set relation EQ gradually becomes predominant over the set relation SUB. In (a) and (b), information regarding inner-adjacency is given in the form of directional estimates of the Allen relation *s*. In (c) and (d), however, this is not applicable. Configurations (e–g) are a strange mix. Their common characteristic is that their spatial relationships contain a fair amount of the set relation OVE, yet no reasonable directional estimate of the Allen relation *o* can be given. Table 6 shows a few more configurations from the subsethood, equality and overlap dataset.

Table 5 Subsethood, equality and overlap dataset (Part 1)

(a)		A is contained by but is marginally equal to B A starts B to the northwest
(b)		A is mostly contained by but is somewhat equal to B A starts B primarily to the north
(c)		A is contained by as much as it is equal to B
(d)		A is perfectly equal to B
(e)		A mostly overlaps but is somewhat disjoint from B
(f)		A overlaps but is marginally contained by B the objects are very slightly equal
(g)		A overlaps but marginally contains B

Finally, Table 7 contains a mayfly mating sequence captured on Doppler radar on June 26, 2001 over St. Clair County, Michigan, USA. The sequence was captured by the National Weather Service Detroit/Pontiac Doppler radar site. The objects were hand segmented from images obtained at http://www.crh.noaa.gov/dtx/mayfly.htm. This dataset was used to test the viability of the proposed approach on real-world data. The descriptions generated seem reasonable, despite the highly irregular (and often disconnected) shapes of the mayfly swarm.

Table 6 Subsethood, equality and overlap dataset (Part 2)

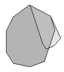
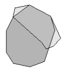
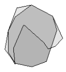
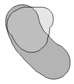
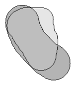
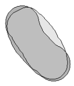
(a)		A is mostly contained by but somewhat overlaps B A overlaps B from the east
(b)		A is mostly contained by but somewhat overlaps B A overlaps B from the northeast
(c)		A is mostly contained by but somewhat overlaps B The objects are very slightly equal A overlaps B loosely from the north
(d)		A is contained by as much as it overlaps B A overlaps B from the northeast
(e)		A mostly overlaps but is somewhat contained by B the objects are slightly equal A overlaps B from the northeast
(f)		A overlaps but is marginally contained by B the objects are very strongly equal A overlaps B from the northeast

Table 7 Doppler radar dataset

(a)		the swarm is perfectly disjoint from St. Clair County the swarm is before the county to the southwest
(b)		the swarm is disjoint from but marginally overlaps St. Clair County the swarm is before the county to the southwest the swarm overlaps the county from the west-southwest
(c)		the swarm mostly overlaps but is somewhat contained by St. Clair County the swarm overlaps the county from the west-southwest
(d)		the swarm is mostly contained by but somewhat overlaps St. Clair County the swarm overlaps the county from the southwest
(e)		the swarm is contained by but marginally overlaps St. Clair County the swarm overlaps the county primarily from the east-southeast
(f)		the swarm is mostly disjoint from but somewhat overlaps St. Clair County the swarm meets the county from the south-southwest the swarm overlaps the county from the southwest

5 Concluding remarks

The system presented here can be used to describe the spatial relationships between two-dimensional objects using natural language. The linguistic descriptions include set, topological and directional relations. The system is quite flexible in the sense that it would be easy to substitute a custom dictionary of adverbs and hedges to suit a particular user's needs. For this approach to work, all that is needed is a set of Allen F-histograms and the values of the fuzzified set relations DIS, OVE, SUB, SUBi and EQ. The ultimate goal of this research is to one day be able to succinctly describe the spatial relationships between two arbitrary objects. In future work, we will investigate the use of relations such as *between* and *surrounds*, and we will incorporate distance information into the descriptions.

Acknowledgements

The authors would like to express their gratitude for the support from the Natural Science and Engineering Research Council of Canada (NSERC), grant 045638.

References

Abella, A. and Kender, J.R. (1999) 'From images to sentences via spatial relations', *Proceedings of ICCV'99 Workshop on the Integration of Image and Speech Understanding,* Corfu, Greece.

Allen, J.F. (1983) 'Maintaining knowledge about temporal intervals', *Communications of the ACM*, Vol. 26, No. 11, pp.832–843.

Egenhofer, M.J. and Herring, J.R. (1994) 'Categorizing topological spatial relations between point, line, and area objects', in Egenhofer, M.J., Mark, D.M. and Herring, J.R. (Eds.): *The 9-Intersection: Formalism and its Use For Natural-Language Spatial Predicates*, National Center for Geographic Information and Analysis, Report 94-1, Santa Barbara, CA.

Keller, J.M. and Wang, X.A. (2000) 'Fuzzy rule-based approach to scene description involving spatial relationships', *Computer Vision and Image Understanding*, Vol. 80, No. 1, pp.21–41.

Kopp, L. (1994) 'A neural network for spatial relations: connecting visual scenes to linguistic descriptions', *Lund University Cognitive Studies*, LUCS 32.

Matsakis, P. (1998) *Relations Spatiales Structurelles et Interprétation d'images*, PhD Thesis, Institut de Recherche en Informatique de Toulouse, France.

Matsakis, P. and Nikitenko, D. (2005) 'Combined extraction of directional and topological relationship information from 2D concave objects', in Cobb, M., Petry, F. and Robinson, V. (Eds.): *Fuzzy Modeling with Spatial Information for Geographic Problems*, Springer-Verlag, pp.15–40.

Matsakis, P. and Wendling, L. (1999) 'A new way to represent the relative position of areal objects', *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 21, No. 7, pp.634–643.

Matsakis, P., Keller, J., Wendling, L., Marjamaa, J. and Sjahputera, O. (2001) 'Linguistic description of relative positions in images', *IEEE Trans. on Systems, Man and Cybernetics, Part B*, Vol. 31, No. 4, pp.573–588.

Matsakis, P., Wawrzyniak, L. and Ni, J. (submitted) 'Topological relationships described in words', *Int. Journal of Geographical Information Science*.

Skubic, M., Perzanowski, D., Schultz, A. and Adams, W. (2002) 'Using spatial language in a human-robot dialog', *Proceedings of IEEE International Conference on Robotics and Automation*, May 11–15, Washington DC, USA, IEEE, pp.4143–4148.

Wang, X. and Keller, J.M. (1997) 'Fuzzy surroundedness', *Proceedings of the 6th IEEE International Conference of Fuzzy Systems, FUZZ-IEEE'97*, July, Barcelona, Spain, pp.1173-1178.

Wang, X. and Keller, J.M. (1999) 'Human-based spatial relationship generalization through neural/fuzzy approaches', *Fuzzy Sets and Systems*, Vol. 101, No. 1, pp.5–20.

Wang., Y. and Makedon, F. (2003) 'R-histogram: quantitative representation of spatial relations for similarity-based image retrieval', *Proceeding of 11th ACM International Conference on Multimedia*, November 2–8, Berkeley CA, USA, pp.323–326.

Wang., Y., Makedon, F., Ford, J., Shen, L. and Goldion, D. (2004) 'Generating fuzzy semantic metadata describing spatial relations from images using the R-histogram', *Proceedings of 2004 Joint ACM/IEEE Conference on Digital Libraries*, 7–11 June, Tucson AZ, USA, pp.202–211.

Wawrzyniak, L., Nikitenko, D. and Matsakis, P. (2005) 'Describing topological relationships in words: refinements', *Proceedings of the IEEE Int. Conf. on Fuzzy Systems (FUZZ-IEEE 2005)*, Reno, Nevada, USA, May, pp.743–748.

Zhan, F.B. (1997) 'Approximation of topological relations between fuzzy regions satisfying a linguistically described query (extended abstract)', in Hirtle, S.C. and Frank, A.U. (Eds.): *Proceedings of International Conference COSIT'97*, 15–18 October, Springer, Pennsylvania, USA, pp.509–510.

Zhan, F.B. (1998) 'Approximate analysis of topological relations between geographic regions with indeterminate boundaries', *Soft Computing*, Vol. 2, No. 2, pp.28–34.

Zhan, F.B. (2001) 'Approximate characterization of topological relations between fuzzy regions and their corresponding linguistic descriptions', *Systems Analysis Modeling Simulation*, Vol. 41, No. 1, pp.1–17.

Zhan, F.B. (2002) 'A fuzzy set model of approximate linguistic terms in descriptions of binary topological relations between simple regions', in Matsakis, P. and Sztandera, L. (Eds.): *Soft Computing in Defining Spatial Relations*, Physica-Verlag, Heidelberg, Germany, pp.179–202.

Note

[1]The notation (SUB | SUBi | EQ) means that only one of the three relations is applicable, depending on the object configuration in hand.