Journal of Visual Languages and Computing (1992) 3, 343–371

Qualitative Spatial Reasoning about Distances and Directions in Geographic Space

ANDREW U. FRANK*

National Center for Geographic Information and Analysis, Department of Surveying Engineering, University of Maine, Orono, Maine, U.S.A.

Received 2 May 1991 and accepted 6 May 1992

Most known methods for spatial reasoning translate a spatial problem into an analytical formulation in order to solve it quantitatively. This paper describes a method for formal, qualitative reasoning about distances and cardinal directions in geographic space. The problem addressed is how to deduce the distance and direction from point A to C, given the distance and direction from A to B and B to C.

We use an algebraic approach, discussing the manipulation of distance and direction symbols (e.g. 'N', 'E', 'S' and 'W', or 'Far' and 'Close') and define two operations, *composition* and *inverse*, applied to them. After a review of other approaches, the desirable properties of deduction rules for distance and direction values are analyzed. This includes an algebraic specification of the 'path' image schema, from which most of the properties of distance and direction manipulation follow. Specific systems for composition of distance are explored. For directions, a formalization of the well-known triangular concept of directions (here called cone-shaped directions) and an alternative projection-based concept are explored.

The algebraic approach leads to the completion of distance or direction symbols with an identity element, standing for the direction or distance from a point to itself. The so completed axiom system allows deductions, at least 'Euclidean-approximate', for any combination of input values.

1. Introduction

HUMANS reason in various ways and in various situations about space and spatial properties. The most common examples for reasoning about geographic or large-scale space [1] are navigational tasks in which the problem is to find a route between two given points given certain conditions [2] and decisions about the location of a resource, which translates to a mundane household question like 'where should the telephone be placed?', or the major problem of locating a nuclear waste facility. Military applications using spatial reasoning for terrain analysis, route selection in terrain, etc. are frequent [3, 4]. Indeed, spatial reasoning is so widespread and common that it is often not recognized as a special case of reasoning.

Spatial reasoning is a major requirement for a comprehensive geographic information system (GIS) and several research efforts are underway to address this need for a better understanding of spatial reasoning [5, 6]. A GIS must carry out spatial tasks which include specific inferences based on spatial properties in a manner similar to a human expert, and must then explain the conclusion to users in terms they can understand [7]. In current GISs, such spatial reasoning tasks are most often

* Current address: Department of Geo-Information (E 127), Technical University Vienna, Gusshausstrasse 27–29, A-1040 Vienna, Austria.

formalized by translating the situation to Euclidean geometry using an analytical treatment for finding a solution. Though Euclidean geometry is a convenient and sometimes the only known model of space available for rigorous analytical approaches, this is admittedly not an appropriate model for human reasoning [1] and thus does not lead to acceptable explanations. A similar problem was found in physics and a formalization of the physical laws used in everyday life, so called 'naive physics', was started [8] using more qualitative than quantitative approaches [9, 10].

This paper treats qualitative reasoning with cardinal directions and distances between point-like objects in a two-dimensional geographic space. Spatial reasoning is influenced by the concept of the underlying space and the task to be solved; here 'geographic' or large-scale space and people that have 'survey knowledge' of the points and their location, are assumed. Examples for locations would include cities, landmarks, etc. The reasoning task treated is not directly related to navigation, which is usually restricted to a network of navigable routes, but assumes a uniform space. Specifically, distances are thought of 'as the crow flies' and not measured along routes. Reasoning about cardinal direction is carried out in a globally fixed reference frame, and radial reference frames, as are customary, for instance in Hawaii [11], are excluded.

Rules for inferring certain direction relations from a set of directional data about some points can be established. These basic capabilities are necessary for solving the more complex spatial reasoning problems [12]. A previous paper with the terms 'qualitative reasoning' in its title [13] is essentially based on analytical geometry. In contrast, the treatment in this paper is entirely qualitative and uses Euclidean geometry only as a source of intuition in Sections 3 to 5 to determine the desirable properties of reasoning with cardinal directions and distances.

As an example, consider the following information:

Chicago is *far* and *north* of *St Louis*, *Los Angeles* is *near* and *south* of *San Francisco*, *St Louis* is *far* and *east* of *San Francisco*, *New Orleans* is *near* and *south* of *St Louis*, etc.

This is sufficient information to conclude that *Chicago* is *north* of *New Orleans*, using the reasoning rules presented here.

This work is part of a larger effort to understand how human beings describe and reason about space and spatial situations. Within the research initiative 2, 'Languages of Spatial Relations' of the National Center for Geographic Information and Analysis [6] a need for multiple formal descriptions of spatial reasoning—both quantitative–analytical and qualitative—became evident [14–17].

This paper is organized as follows: the remainder of this section reviews previously presented approaches to spatial reasoning, most of them using analytical geometry. Section 2 introduces the concept of qualitative reasoning and relates it to spatial reasoning using analytical geometry; a definition of 'Euclidean exact' qualitative reasoning based on a homomorphism is presented. In the following section an algebra for 'path', the relevant imaging schema [18] is defined. From this Sections 4 and 5 deduce the desirable properties of distance and direction reasoning systems. Section 6 compares a few different qualitative distance systems, whereas Section 7 discusses the direction system based on cone-shaped (triangular) areas and a direction system based on projections. The last section integrates the distance and direction reasoning systems into a single one and assess its deductive power. The conclusions give a summary of the results and lists topics for future research.

1.1. Previous Approaches

A standard approach to modeling human spatial reasoning is to use Euclidean geometry in the plane or three-dimensional space and represent the task using analytical geometry formulae. Many problems can be expressed as an optimization problem with a set of constraints, such as location of a resource and the shortest path in this framework.

Similarly, the important field of geographic reference frames in natural language [19] has mostly been treated using an analytical geometry approach. Typically, spatial positions are expressed relative to positions of other objects. Examples occur in every day speech in forms such as 'the church is west of the restaurant'. In the past these descriptions were translated into Cartesian coordinate space and the mathematical formulations were analyzed. A special problem is posed by the inherent undertainties in these descriptions and the translation of the uncertainties into an analytical format. McDermott [12] introduced a method using 'fuzz' and in Dutta [13, 20] fuzzy logic [21] is used to compose such approximately metric data.

An approach that is entirely qualitative, and thus similar to the thrust of this paper, is the work on symbolic projections. This approach translates exact metric information (primarily about objects in pictures) in a qualitative form [22, 23]. The order in which objects appear, projected vertically and horizontally, is encoded into two strings and spatial queries are executed as fast substring searches [24].

Hernández [25] discusses qualitative reasoning between extended objects, based on Allen [26], and uses topological relations (adjacent, overlap, etc.) between objects, but not distances—thus avoiding the problem of defining distance relations between extended objects. His approach also includes the composition of reference frames. It uses eight cone shaped directions (left, right, front, back and the intermediates left-front, etc.) and four steps of distance, expressed as topological relations. A similar approach is taken by Guesgen [27] and Munkerjee [28, 29].

2. Qualitative Approach

2.1. Qualitative Reasoning

This paper gives a set of qualitative deduction rules for a subset of spatial reasoning, namely reasoning with cardinal directions and qualitative distance descriptors, without relying on qualitative calculations (e.g. square roots, trigonometric functions) or analytical geometry. In qualitative reasoning a situation is characterized by variables which 'can only take a small, predetermined number of values' [30, p. 116] and the inference rules use these values and not numerical quantities approximating them. It is clear that the qualitative approach loses some precision, but simplifies reasoning and allows deductions when precise information is not available.

Without debating whether human reasoning follows the structure of propositional logic, there is some evidence that human thinking is at least partially symbolic and qualitative [18, 31, 32]. Formal, qualitative spatial reasoning is crucial for the design of flexible methods to represent spatial knowledge in GISs and for constructing usable GIS expert systems [12, 33]. Spatial knowledge is currently seldom included in expert systems and is considered 'difficult' [34].

The problem is, given a set of propositions about the relative positions of objects in a plane, how to deduce other spatial relationships [13]. The relations focused on are the cardinal directions and distances. In terms of the example given in the introduction, the following chain of reasoning deduces a direction from Chicago to New Oreleans.

1. Use '*Chicago* is *north* of *St Louis*' and '*New Orleans* is *south* of *St Louis*', two statements which establish a sequence of directions *Chicago–St Louis–New Orleans.*
2. Deduce '*St Louis* is *north* of *New Orleans*' from '*New Orleans* is *south* of *St Louis*'
3. Use a concept of transitivity: '*Chicago* is *north* of *St Louis*' and '*St Louis* is *north* of *New Orleans*' thus conclude '*Chicago* is *north* of *New Orleans*'.

This paper formalizes such rules and makes them available for inclusion in an expert system.

2.2. Advantage of Qualitative Reasoning

A qualitative approach can deal with imprecise data, and therefore yields less precise results than the quantitative one. This is highly desirable [6, 35], because:

- precision is not always desirable; and
- precise, quantitative data is not always available.

Qualitative reasoning has the advantage that it can deal with imprecise data and need not translate it into a quantitative form [36]. Verbal descriptions are typically not metrically precise, but sufficient for the task intended. Imprecise descriptions are necessary in query languages where one specifies some property that the requested data should have, for example a building about 3 miles from town. It is difficult to show this condition in a figure, because the figure necessarily over-specifies (Figure 1 also determines a direction) or is overly complex (Figure 2). Qualitative reasoning can also be used for query simplification to transform a query from the form in which it is posed to another equivalent one that is easier to execute. In other cases, the available data is in qualitative form, most often text documents. For example Tobler and Wineberg [37] tried to reconstruct spatial locations of historic places from scant descriptions in a few documents. Verbal information about locations of places can leave certain aspects imprecise and force humans to deduce information from such descriptions (for example, in order to analyze automatically a description of a location in natural science collections [38]).

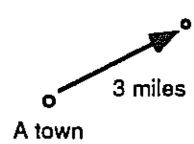

Figure 1. Overspecific visualization

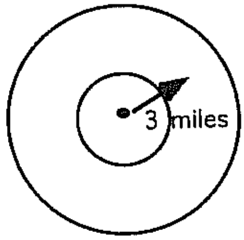

Figure 2. Complex visualization

2.3. Algebraic Methods for a Qualitative Approach

The approach proposed here discusses combinations of directional symbols in terms of an algebra, and introduces an identity symbol which makes reasoning more powerful. The description of directional relationships between points in the plane can be formulated as propositions, such as 'A is north of B' or 'north (A, B)'. From a set of propositions one can then deduce other relative positions as the induced set of spatial constraints [13]. In Hernández [25] the result of the typical composition—given an item of information about the relation of A and B and another one for B and C, conclude the relation from A to C—is a set of possible outcomes. If a position is determined by more than one chain of reasoning, one selects the intersection of all the results. This is a widely used convention, assuring that the result of a deduction chain is certainly containing the correct value. The underlying assumption is that the qualitative symbols describe sharply bounded intervals. This assumption contradicts the evidence from human reasoning, which points towards prototypes and radial categories [18]. The rules developed here result always in a single value (not a set of values), which may be only approximately correct. Other mechanisms than isolated reasoning with distances and directions are used to correct for potential errors.

Following an algebraic concept, in this paper rules are given for the manipulation of the directional or distance symbols when combined by operators, e.g. composition. In order to have rules that yield a single result for all possible combinations of input values, inference rules that are only Euclidean approximate (defined in the next subsection) may be necessary.

In this context, one would typically use the well known algebra of vectors. The operations, e.g. vector addition '$+_v$' or vector subtraction '$-_v$', are qualified by a subscript, to indicate the specific algebra (subscripts are used extensively throughout

Table 1. The components of an algebra [44].

An algebra consists of: • A set of symbols, D, called the domain of the algebra—comparable to the concept of data type in computer programming languages (e.g., $D = \{N, E, W, S\}$). • A set of operations over D, comparable to functions in a computer program (primarily operations to reverse and to compose directions). • A set of axioms that set forth the basic rules explaining what the operations do.

the paper to indicate the type to which the operation applies). The typical axioms for a vector algebra are:

$$a +_v 0_v = a \qquad 0_v \text{ is identity element} \tag{1}$$

$$a +_v b = b +_v a \qquad \text{commutative law} \tag{2}$$

$$a +_v (b +_v c) = (a +_v b) +_v c = a +_v b +_v c \qquad \text{associative law} \tag{3}$$

$$-_v(-_v a) = a \tag{4}$$

$$a +_v (0 -_v a) = 0_v \qquad \text{inverse for } +_v \tag{5}$$

2.4. Exact and Approximate Reasoning

The result of a qualitative reasoning rule is compared with the result obtained by translating the data to analytical geometry and applying the equivalent functions to them. If the results are always the same, i.e., a homomorphism (Figure 3), the qualitative rule is called *Euclidean exact*. If the qualitative rule produces results, at least for some data values, which are different from the ones obtained from analytical geometry, it is called *Euclidean approximate.*

This general definition applies, for example, to the operation to compose two directions and deduce the direction of the resultant (introduced in Section 4.3, see Figure 5). The vectors $\langle P_i, P_j \rangle$ are mapped into symbolic directions using a function dir (P_i, P_j). For reasoning with distance, the function dist (P_i, P_j) maps vectors into the distance symbols. In both cases, vector addition $(+_v)$, with the regular properties is carried to symbolic composition $(*_c)$ for direction and $(+_d)$ for distance.

DEFINITION 1: a rule for qualitative reasoning on directions is called *Euclidean exact* (or short *exact*), if dir (P_1, P_2) is a homomorphism.

$$\operatorname{dir}(P_1, P_2) *_c \operatorname{dir}(P_2, P_3) = \operatorname{dir}((P_1, P_2) +_v (P_2, P_3)) \tag{6}$$

DEFINITION 2: a rule for qualitative reasoning on distances is called *Euclidean exact* (or short *exact*), if dist (P_1, P_2) is a homomorphism.

$$\operatorname{dist}(P_1, P_2) +_d \operatorname{dist}(P_2, P_3) = \operatorname{dist}((P_1, P_2) +_v (P_2, P_3)) \tag{7}$$

2.5. Tolerance Geometry

Any qualitative spatial reasoning must map values that are slightly different to a single value. For example, the distance from a point to several other points will be judged 'close', even if the distances themselves are different and not exactly zero. But other

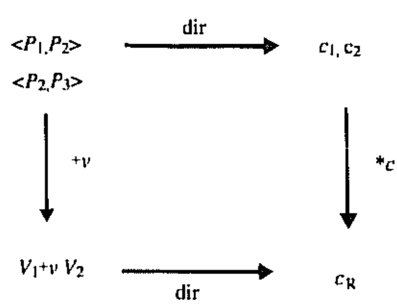

Figure 3. Homomorphism for directions

points, 'close' to these, will be considered 'far' from the first point (i.e. 'close to A' does not form an equivalence class).

$$\text{dist}(P_1, P_2) = 0 \text{ AND } \text{dist}(P_2, P_3) = 0 \not\Rightarrow \text{dist}(P_1, P_3) = 0 \tag{8}$$

A tolerance space [39, 40] is defined as a set (in this case the points P) and a tolerance relation, in this case $\text{dist}(P_1, P_2) = 0$. The tolerance relation relates objects which are so close, one can or need not differentiate between them. A tolerance relation is similar to an equality, except that it admits small differences. It is reflexive and symmetric, but not transitive (as an equality would be).

3. An Algebra of Paths

This discussion treats only the relative positions of points with respect to a fixed directional reference frame. This is primarily the metric information, i.e. relations that are not invariant under topological (homeomorph) transformations and includes the distance and direction relations [41]. One could form other reasoning rules—especially if one considers a more flexible description in first order predicates—for example, if A to B is parallel to C to D and to E to F one can deduce that C to D is parallel to E to F, or if A to B is at right angles of B to C and the direction from A to B is given, one can deduce the direction B to C. One could also make statements like point P is in the middle of A to B, etc. [42].

Rules for reasoning with distances and cardinal directions in large-scale space specifically the two operations studies here, follow from the properties of the 'path' imaging schema [43]. A directed path from a start point P_1 to an ending point P_2 is here called shortly a path, written as $\langle P_1, P_2 \rangle = s_1$. This section develops an algebra for paths and gives axioms for the two operations *inverse* and *compose*.

3.1. Reversing Direction of a Path

The inverse of a path from P_1 and P_2 gives the path from P_2 to P_1 (Figure 4). The symbol $-_p$ is used for this operation. Inverse has the property that the inverse of the inverse of the path is the original path, which also states that there are only two different directions to traverse a path, i.e. forward and back.

$$-_p \langle P_1, P_2 \rangle = \langle P_2, P_1 \rangle \tag{9}$$

$$-_p(-_p s) = s \tag{10}$$

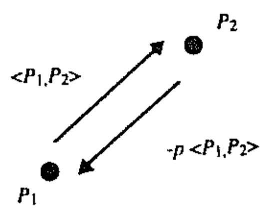

Figure 4. Inverse $s_1 = -_p(s_2)$

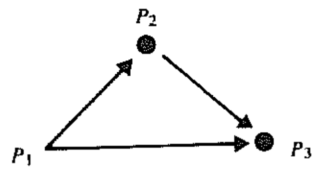

Figure 5. Composition $s_3 = s_1 *_p s_2$

3.2. Composition

Composition merges two contiguous paths, from P_1 to P_2 and P_2 to P_3, into a single path from P_1 to P_3 (Figure 5). The two paths are joined at the middle point and composition is *not commutative* [Equation (13)], indeed in the general case only one of $s_1 *_p s_2$ or $s_2 *_p s_1$ is a meaningful expression. Composition is thus different from vector addition ($+_v$). Vector addition is commutative [Equation (2)], as one assumes that vectors can be translated, but this is not assumed for paths. Combinations of more than two paths are independent of the order in which they are combined [*associative* law, Equation (14)] and parentheses are not needed [follows from Figure 6 or the definition of composition [Equation (11)].

$$\langle P_1, P_2 \rangle *_p \langle P_2, P_3 \rangle = \langle P_1, P_3 \rangle \tag{11}$$

$$\langle P_1, P_2 \rangle *_p \langle P_2, P_3 \rangle = \langle P_1, P_4 \rangle *_p \langle P_4, P_3 \rangle = \langle P_1, P_3 \rangle \tag{12}$$

$$s1 *_p s2 \neq s2 *_p s1 \tag{13}$$

$$a *_p (b *_p c) = (a *_p b) *_p c = a *_p b *_p c \tag{14}$$

Concatenation of a path and the composition of paths must be differentiated. If paths are concatenated, all intermediate points are retained and the resulting aggregate path is a zig-zag line. The result of the composition of paths is a single path, from the start to the end point, and the intermediate points are dropped [Equation (12)].

3.3. Identity

The special path from a point to itself is called the identity path and denoted by the symbol 0_p. It is an identify for composition: combining the identity path to any other path results in the same path [Equation (16)]; this follows from the definition of composition [Equation (11)]. The identity path is a concept which has no specific symbol in natural language, but for formal reasoning an explicit symbol is needed.

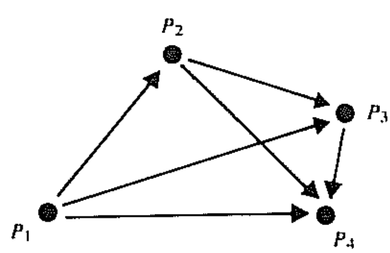

Figure 6. Associativity

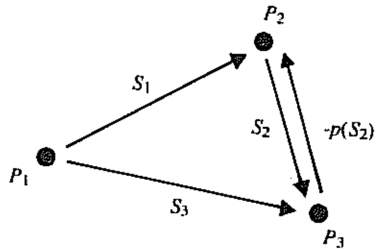

Figure 7. $s_1 = s_3 *_p -_p (s_2)$

Therefore:

$$\langle P_1, P_2 \rangle = 0_p \tag{15}$$

$$p *_p 0_p = 0_p *_p p = p \tag{16}$$

$$-_p 0_p = 0_p \tag{17}$$

3.4. Inverse to Composition

The inverse to a binary operation is defined such that a value, combined with its inverse, results in the identity value [Equation (18)]. Because composition is not commutative, two different inverses are possible, but if they both exist, they must be the same [44, p. 270]. The inverse operation defined previously [Equation (10)] has the required properties (Figure 7). Finally, a composition can be inverted by piece-wise inversion [Equation (10)]:

$$p *_p (-_p p) = 0_p \quad \text{and} \quad (-_p p) *_p p = 0_p \tag{18}$$

$$\text{dir}\,\langle P_1, P_2 \rangle\,_c\,\text{dir}\,\langle P_2, P_1 \rangle = \text{dir}\,\langle P_1, P_1 \rangle \tag{19}$$

for any a, b: path exist x: path

$$x *_p a = b \qquad \text{and} \quad a *_p x = b$$

$$-_p (a\,_c\,b) = -_p b\,_c -_p a \tag{20}$$

Table 2. Algebra for path.

Sort:	path		
Operations:	inverse ($-_p$):	path → path	
	compose ($*_p$):	path × path → path	
Constant:	identity: 0_p		
Axioms:	$-_p(-_p s) = s$		
	$a *_p (b *_p c) = (a *_p b) *_p c = a *_p b *_p c$		associative law
	$a *_p 0_p = a,\ 0_p *_p a = a$		identity element
	$-_p(a *_p b) = -_p b *_p -_p a$		piecewise inversion
	for any a, b: path exist x: path		
	$x *_p a = b$ and $a *_p x = b$		

4. Properties for Distances

In this section desirable properties for a qualitative distance function are examined, starting with the standard distance functions used in Euclidean geometry. Functions and properties that follow from the path algebra are introduced.

4.1. Standard Distance Functions in Geometry

The standard properties for the mathematical definition of distance capture the meaning of distance well. Distance is defined as a function which maps from a pair of points to the positive real numbers with the properties [Equations (21)–(23)]. These rules are compatible with the rules for composition of path, examining the axioms of the path algebra. (The symbols + and $\geq$ here stand for the regular addition and comparison of real numbers.)

$$\text{dist}\,(P_1, P_1) = 0 \tag{21}$$

$$\text{dist}\,(P_1, P_2) = \text{dist}\,(P_2, P_1) \tag{22}$$

$$\text{dist}\,(P_1, P_2) + \text{dist}\,(P_2, P_3) \geq \text{dist}\,(P_1, P_3) \tag{23}$$

4.2. Qualitative Distance Function

A qualitative distance function maps from two points in the plane onto a qualitative distance value. From the definition of path [Equation (9)] it follows that the distance function also maps a path to a distance value. The symbols are in the simplest case $D_2 = \{C, F\}$ for 'close' and 'far', or $D_4 = \{CC, C, M, F, FF\}$ for 'very close' to 'very far' or another finite set of distance symbols. In order to fulfill the distance properties [Equations (21)–(23)] an *addition* $(+_d)$ and *comparison* $(\leq_d)$ operation must be defined on these qualitative distance symbols. The values must be totally ordered by $\leq_d$, and the set must have a smallest element, written as 0_d, which is the distance of the identity path (C in D_2 and CC in D_4). From the finiteness of the set of values and the total order it follows that there exists a maximal element. From addition of reals we inherit commutativity of the distance addition [Equation (24)] and monotonic increase [(Equation (25)].

$$d_1 +_d d_2 = d_2 +_d d_1 \tag{24}$$

$$d_2 \geq_d d_3 \Rightarrow d_1 +_d d_2 \geq_d d_1 +_d d_3 \tag{25}$$

Distance values are summed along a concatenated path. Adding a distance to another increases the value or leaves it the same, but does not decrease it [Equation (27)].

$$\begin{aligned} &\text{dist: point} \times \text{point} \rightarrow D \quad \text{or} \quad \text{dist: path} \rightarrow D \\ &\text{dist}\,(0_p) = 0_d \end{aligned} \tag{26}$$

$$\begin{gathered} d_1 +_d d_2 \geq_d d_1 \quad \text{and} \quad d_1 +_d d_2 \geq_d d_2 \\ d_1 +_d d_2 \geq_d \max\,(d_1, d_2) \end{gathered} \quad \text{or short:} \tag{27}$$

4.3. The Protype Finite Distance Values

Despite the fact that many different sets of symbolic distance values can be thought of—every natural language has its own—providing they have the properties defined above they can all be mapped to a prototypical set, constructed from a subset of the integers. Different distance systems can have any desired number of intermediate steps as a subset of the integers from 0 to $n-1$. The 'meaningful' symbols are then mapped onto the numbers (by toInt and its inverse toSymbol). These mappings must be order preserving. For example, the set $\{C, F\}$ is mapped to $\{0, 1\}$, or $\{CC, C, M, F, FF\}$ is mapped to $\{0, 1, 2, 3, 4\}$.

$$\text{toInt:} \quad d \rightarrow \text{int}$$
$$\text{toSymbol:} \quad \text{int} \rightarrow d$$
$$\text{toSymbol toInt}(s)) = s$$
$$\text{toInt}(C) = 0 \quad \text{respective in } D_4 \qquad \text{toInt}(CC) = 0.$$
$$d_1 \leq_d d_2 \Leftrightarrow \text{toIntd}(d_1) \leq \text{toInt}(d_2).$$

4.4. Rules Following from the Path Algebra

The path algebra imposes some restrictions on how distance functions can work. The goal is to find a (nearly) homomorphic mapping, induced by dist, which maps composition of paths $*_p$ onto addition of distances $+_d$.

$$\text{dist}(p_1 *_p p_2) = \text{dist}(p_1) +_d \text{dist}(p_2) \tag{28}$$

The smallest distance value (0_d) must be the identity element for addition [Equation (29)] and the distance of the inverse of a path is the same as the distance of the path [Equation (30)]—there is no need for an inverse of distance addition [follows from Equation (22)]. From the associative law for path composition it follows that distance addition must be associative as well.

$$d_1 +_d 0_d = d_1 \quad \text{and} \quad 0_d +_d d_1 = d_1 \tag{29}$$
$$\text{dist}(-_p p) = \text{dist}(p) \tag{30}$$
$$(d_1 +_d d_2) +_d d_3 = d_1 +_d (d_2 +_d d_3) = d_1 +_d d_2 +_d d_3 \tag{31}$$

Proof for Equation (29): From Equation (16) $\text{dist}(p_1 *_p 0_p) = \text{dist}(p_1)$, therefore $\text{dist}(p_1) +_d \text{dist}(0_p) = \text{dist}(p1) + 0_d = \text{dist}(p_1)$.

From path composition and the triangular inequality, it follows that the distance of a composite path is less or equal to the sum of the distance of the components—and not strictly equal, as the homomorphism [Equation (28)] would require. The difficulty is seen drastically in the composition of a path with its own inverse [Equation (32)].

$$\text{dist}(p_1 *_p p2) \leq \text{dist}(p_1) +_d \text{dist}(p_2)$$
$$\text{dist}(p_1 *_p -_p(p_1)) = \text{dist}(0_p) = 0_d$$
$$\text{dist}(p_1 *_p -_p(p_1)) = \text{dist}(p_1) +_d (\text{dist}(p_1) \geq_d 0_d \tag{32}$$

Table 3. Desirable properties of distances.

$\text{dist}(0_p) = 0_d$
$\text{dist}(-_p p) = \text{dist}(p)$
$d_1 +_d 0_d = 0_d +_d d = d$
$d_1 +_d d_2 \geq \max(d_1, d_2)$
$d_1 +_d (d_2 +_d d_3) = (d_1 +_d d_2) +_d d_3 = d_1 +_d d_2 +_d d_3$

5. Properties of Cardinal Directions

In this section the properties of a direction system are discussed, starting with an analysis of the standard mathematical system. The properties which follow from the path algebra are deduced and extended with other desirable properties. Unfortunately, this set of properties contains contradictions and cannot be simulatenously fulfilled.

5.1. Standard Mathematical Definition of Directions

Usually the direction of a path is defined as its angle with a given fixed direction in the frame of reference, typically geographic north or the positive x or y axis. The direction is expressed as an angle, measured as a real number from 0 to 2π. Values are cyclically ordered such that an increment operation starting from 0 generates all values and returns to 0. Conventional mappings to combinations of cardinal directions and degrees as often seen as 'bearings' (e.g. N 15° 12′ E), are (isomorphic) variants of the same system.

5.2. Qualitative Directions

Qualitative direction is a function between two points in the plane that maps onto a symbolic direction or its equivalent, from a path onto a symbolic direction. The n different symbols available for describing the directions are given as a set C_n; they depend on the specific system of directions used, e.g. $C_4 = \{N, E, S, W\}$ or more extensively $C_8 = \{N, NE, E, SE, S, SW, W, NW\}$. The number of different qualitative direction symbols is finite and are cyclically ordered and equidistant. They can be mapped to the integers modulo $(n-1)$, for example $C_8^* = \{0, 1, 2, 3, 4, 5, 6, 7\}$. A cyclical positive (anti-clockwise) turn of $2\pi/n$, where n is the number of direction symbols, is useful and can be defined by a table, for example, $\text{turn}(N) = E$, $\text{turn}(E) = S$, $\text{turn}(S) = W$, $\text{turn}(W) = N$, or a full turnaround is $\text{turn}^{(n)}$. the identity operation.

$$\text{dir}: p \, x \, p \rightarrow C \quad \text{or} \quad \text{dir}: \text{path} \rightarrow C \tag{32}$$

It is often assumed that the two points between which the direction is determined must not be the same. To overcome this limitation, the identity element (the 'identity direction' 0_c) is introduced. The sets of direction symbols become then $C_4' = \{N, E, S, W, 0\}$ or more extensively $C_8' = \{N, NE, E, SE, S, SW, W, NW, 0\}$. Two operations, inverse ($-_c$) and composition ($*_c$) are applied to these symbols, with the rules that follow in the remainder of this section.

5.3. Rules Following from the Path Algebra

The path algebra imposes some restrictions on how direction functions can work. The goal is to find a (nearly) homomorphic mapping, induced by dir, which maps $*_p$ onto $*_c$ and $-_p$ onto $-_c$.

$$\text{dir}(-_p p) = -_c \, \text{dir}(p) \tag{33}$$

$$\text{dir}(p_1 *_p p_2) = \text{dir}(p_1) *_c \text{dir}(p_2) \tag{34}$$

Cardinal directions depend on the order of travel. If a direction is given for a line segment between points P_1 and P_2, the direction from P_2 to P_1 must be deducible.

Already Freeman [45] and Peuquet [46] have stressed the importance of this operation: 'Each direction is coupled with a semantic inverse' [46]. From path algebra Equation (10) immediately follows Equation (35), which is only possible if the number of definite symbols (i.e. $\neq 0_c$) is even. Then the inverse is equal to half of a full turnaround (5.5) and groups the direction symbols in pairs.

$$-_c(-_c(c)) = c \tag{35}$$

$$-_c(c) = \text{turn}^{(n-1)/2}(c) \tag{36}$$

The identity path must map to the identity directions [Equation (37)] and the identity direction must be the identity element for composition of directions [Equation (39)]. From the associative law for path composition it follows that direction composition must be associative as well.

$$\text{dir}(0_p) = 0_c \tag{37}$$

$$-_c 0_c = 0_c \tag{38}$$

$$0_c *_c c = c \quad \text{and} \quad c *_c 0_c = c \tag{39}$$

$$(c_1 *_c c_2) *_c c_3 = c_1 *_c (c_2 *_c c_3) = c_1 *_c c_2 *_c c_3 \tag{40}$$

Cancelation—A path and its inverse cancel and compose to 0_p [Equation (18)], which translates for directions to:

$$\text{dir}(a) *_c \text{dir}(-_p(a)) = \text{dir}(a) *_c -_c(\text{dir}(a)) = c_1 *_c -_c(c_1) = 0_c. \tag{41}$$

This introduces a problem in case only the two directions are given (c_1 and $c_2 = -_c(c_1)$), but their lengths are different (Figure 8). The result is not Euclidean exact and the degree of error depends on the definition of 0_c used and the difference in the size of the line segments—if they are the same the inference rule is exact. This represents a type of reasoning like Baltimore is east of San Francisco, San Francisco is west of Washington D.C., thus the distance from Baltimore to Washington is too close to determine a direction.

5.4. Transitivity for Cardinal Directions

For cardinal directions an additional transitivity rule holds. The composition of two line segments with the same direction maintains the same direction.

$$\text{dir}\langle P_1, P_2\rangle = \text{dir}\langle P_2, P_3\rangle = c \Rightarrow \text{dir}\langle P_1, P_3\rangle \tag{42}$$

or short: $c *_c c = c$, for any c.

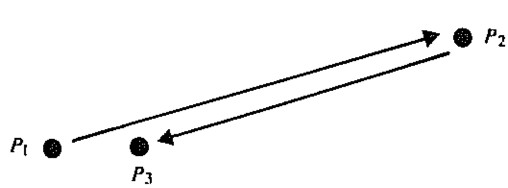

Figure 8. $d *_c -_c(d)$

Table 4. Desirable properties of directions.

$\mathrm{dir}(P_1, P_1) = \mathrm{dir}(0_p) = 0_c$
$-_c(-_c(c)) = c$
$\mathrm{dir}(P_1, P_2) = -_c(\mathrm{dir}(P_2, P_1))$ or $-_c(\mathrm{dir}(p)) = \mathrm{dir}(-_c(p))$
$c *_c 0_c = 0_c *_c c = c$
$c_1 *_c (c_2 *_c c_3) = (c_1 *_c c_2) *_c c_3$
$c *_c c = c$
$c *_c -_c c = 0_c$

The desired properties lead, unfortunately, to a contradiction and no algebraic system can fulfill all of them at once. The problem results from the demand for associativity and cancelation [Equation (40) and (41)]. Struss [47] discusses extensively 'problems of interval-based qualitative reasoning' and found that for somewhat different assumptions that 'there is no qualitative reasoning method with a finite number of values preserving associativity of arithmetic operators, except for the quantity space $\{-\infty, 0, \infty\}$'.

$$c *_c (c *_c -_c(c)) = c *_c 0 = c \quad \text{but} \tag{43}$$

$$(c *_c c) *_c -_c(c) = c *_c -_c(c) = 0. \tag{44}$$

Formalizing concepts from natural language, one should not be surprised to find formal inconsistencies [48]. The problem remaining to be solved is to find a method to cope with these differences between formal and natural language reasoning. This difficulty could be resolved by defining combination to result in a set of values and specificially $c *_c -_c c = \{c, 0_c, -_c c\}$, but this is not the path followed here.

6. Distance Systems

The previously defined properties of a distance can be fulfilled by different sets of symbols and specific rules for operations, similar to various distance functions that fulfill the requirements of the distance definition used in geometry. In this section we define a number of distance systems and assess their power. Given that we deal only with finite numbers of symbols, we can define operations, etc., usually by enumeration.

6.1. Two-step Distance

The first and most obvious concept of distance is a set of two symbols for far and close [49], where not far is close and not close is far. The symbols are $D_2 = \{C, F\}$ and map to integers $\{0, 1\}$, with C mapped to 0 and F to 1. The rules for addition [Equation (24)–(27) determine Equation (45):

$$C +_d C = C, F +_d C = C +_d F = F, F +_d F = F. \tag{45}$$

6.2. Three-step Distance

It is a valid criticism of the above system that 'not far' is not necessarily 'close', it is just 'intermediate' or 'normal' compared to other distances in the same space. This leads to a system with three symbols: close, medium, far; $D_3 = \{C, M, F\}$. C is the identity element and the order is $C < M$, $M < F$.

For the definition of addition, starting with Equation (45), all but $M +_d M$ are already determined by the rules for addition [Equation (24)–(27)] through (4.7): $M +_d C = M$, $M +_d F = F$. Only $M +_d M = M$ or $M +_d M = F$ remains to be selected. The following table shows the two variants.

Table 5. Additions for three-step distance symbols.

$+_d$	C	M	F
C	C	M	F
M	M	M or F	F
F	F	F	F

6.3. Multi-step Distance

The above concepts can be generalized to create any finite sequence of symbols of increasing distance symbols from C to F, with the previously defined mapping to a sequence of integers $D_n = \{0, 1, 2 \cdots n-1\}$. Rules (25) and (27) define all the 'border' cases of composition. The case for $n = 4$, $C_4 = \{C, c, f, F\}$ or $\{0, 1, 2, 3\}$ will reveal the general situation.

Table 6. Additions for multi-step distance.

$+_d$	C	c	f	F
C	C	c	f	F
c	c	*	*	F
f	f	*	*	F
F	F	F	F	F

The four compositions marked with * are not determined yet. The result of $c +_d f$ and $f +_d c$ can be selected to be either f or F. One also observes that $c +_d f = F$ and $f +_d f = f$ is inconsistent with Equation (25). Therefore, only three different distance systems are possible: Equation (50) using Equation (46) and Equation (48); Equation (51) using Equation (46) and Equation (49); and Equation (52) using Equation (47) and Equation (49).

$$\text{choice } A \quad c +_d f = f, \qquad f +_d c = f \tag{46}$$

$$\text{or} \quad c +_d f = F, \qquad f +_d c = F \tag{47}$$

$$\text{choice } B \quad c +_d c = c, f \qquad +_d f = f \tag{48}$$

$$\text{or} \quad c +_d c = f, \qquad f +_d f = F \tag{49}$$

$$\text{Distance system 1:} \quad x +_d y = \max(x, y) \tag{50}$$

$$c +_d c = c$$

$$c +_d f = f +_d c = f$$

$$f +_d f = f$$

$$\text{Distance system 2:} \quad x +_d y = \text{if } x = y \text{ then succ (x) else max } (x, y) \tag{51}$$
$$c +_d c = f$$
$$c +_d f = f +_d c = f$$
$$f +_d f = F$$

$$\text{Distance system 3:} \quad x +_d y = \text{succ}(\max(x, y)) \tag{52}$$
$$c +_d c = f$$
$$c +_d f = f +_d c = F$$
$$f +_d f = F$$

6.4. Geometric Interpretation

The analysis of these cases of distance systems by considering distance addition along a straight line yields some insight into the geometric meaning of these formulae. A number of assumptions are necessary for this interpretation:

- A system of distance symbols is described by n intervals. The set $[0, 1)$, $[1, 2)$, $[2, 3)$ and $[3, \infty)$ is arbitrarily chosen as a first example.
- All Euclidean distance values that fall in each interval are mapped to the corresponding symbol.
- The intervals form a partition of the positive real numbers.
- Distance addition is mapped to interval addition and the middle point of the result determines the resulting symbol.

Applying interval arithmetic as usual $C_1 +_d C_2 = [a_1, e_1) +_i [a_2, e_2) = [a_1 + a_2, e_1 + e_2)$, etc., yields the following table for addition, which represents distance system 3 (6.8).

Table 7. Additions in four-step distance system 3.

$+_i$	$[0, 1) \to C$	$[1, 2) \to c$	$[2, 3) \to f$	$[3, \infty) \to F$
$[0, 1) \to C$	$[0, 2) \to C$	$[1, 3) \to c$	$[2, 4) \to f$	$[3, \infty) \to F$
$[1, 2) \to c$	$[1, 3) \to c$	$[2, 4) \to f$	$[3, 5) \to F$	$[4, \infty) \to F$
$[2, 3) \to f$	$[2, 4) \to f$	$[3, 5) \to F$	$[4, 6) \to F$	$[5, \infty) \to F$
$[3, \infty) \to F$	$[3, \infty) \to F$	$[4, \infty) \to F$	$[5, \infty) \to F$	$[6, \infty) \to F$

What other selection of intervals would produce the other two cases [Equation (50) or (51)]?

1. If the intervals for the distances are selected not to overlap and not to leave any gaps, then the first distance system [Equation (50)] is not possible, because $c +_d c$ must be more than c.

Proof: Using the abbreviations $c = [a_1, e_1)$, etc. From $c +_d c = c$ [Equation (50)] follows $c +_d c = [a_1, e_1) +_i [a_1, e_1) = [2a_1, 2e_1)$. The middle point is $a_1 + e_1$, which is greater than a_2 whenever $a_2 = e_1$ (assumption that intervals do not overlap).

2. Similar reasoning leads to a st of intervals to achieve distance system 2 [Equation

(51)]. The critical condition is:

$$c +_d f = [a_1, e_1) + [a_2, e_2) = [a_1 + a_2, e_1 + e_2),$$

with the middle point $(a_1 + a_2 + e_1 + e_2)/2$, which must be less than e_2. With d_n for the width of the interval n ($d_1 = e_1 - a_1$ and $d_2 = e_2 - a_2$), one finds $2a_1 + d_1 < d_2$.

For example, see Table 8.

Table 8. Additions in four-step distance system 2.

$+_c$	$[0, 1) \to C$	$[1, 3) \to c$	$[3, 8) \to f$	$[8, \infty) \to F$
$[0, 1) \to C$	$[0, 2) \to C$	$[1, 4) \to c$	$[3, 9) \to f$	$[8, \infty) \to F$
$[1, 3) \to c$	$[1, 4) \to c$	$[2, 6) \to f$	$[4, 11) \to f$	$[9, \infty) \to F$
$[3, 8) \to f$	$[3, 9) \to f$	$[4, 11) \to f$	$[6, 16) \to F$	$[11, \infty) \to F$
$[8, \infty) \to F$	$[8, \infty) \to F$	$[9, \infty) \to F$	$[11, \infty) \to F$	$[16, \infty] \to F$

Assuming that the assignment of values to intervals is not based on a partition and using, for example, a fuzzy rule to assign qualitative distances to real distances, other solutions are possible {including distance system 1 [Equation (50)] above}.

6.5. Assessment

The rules developed so far are sufficient to resolve the distance information from our example and to deduce, for example, the distance from Chicago to New Orleans:

dist (*Chicago, St Louis*) = F and

dist (*New Orleans, St Louis*) = $c \Rightarrow$ dist (*St Louis, New Orleans*) = c.

Thus:

dist (*Chicago, New Orleans*) = dist (⟨*Chicago, St Louis*⟩ $*_p$ ⟨*St Louis, New Orleans*⟩) = $F +_d c = F$.

Reasoning with qualitative distances is necessarily Euclidean approximate. Considering only the case of distances along a straight line, the results are, however, quite good. For the four distance symbol systems described above and some plausible but simplifying assumptions, the rules for distance system 3 [Equation (52)] yield correct results in more than 80% of the cases. For the distance system 2 [Equation (51)], the approximation is even better, and the results are correct in 85% of the cases. Cases that produce incorrect results can be easily found in Table 8.

7. Direction Systems

7.1. Cardinal Directions as Cones

The most often used prototypical concept of cardinal directions is related to the angular direction between the observer's position and a destination point. This direction is rounded to the next established cardinal direction. The compass is usually divided into four cardinal directions, often with subdivisions for a total of eight or more directions. This results in cone-shaped areas for which a symbolic direction is applicable. The investigation here is limited to the case of four and eight directions.

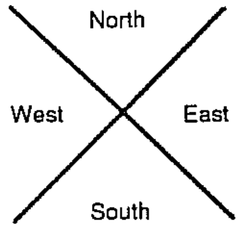

Figure 9. Cone-shaped directions

This model of cardinal directions has the property that 'the area of acceptance for any given direction increases with distance' [46, p. 66, with additional references] and is sometimes called 'triangular'.

7.1.1. Traditional Definitions with Four Directional Symbols and no Identity

Four cardinal directions are defined as cones (Figure 9), such that for every line segment, exactly one direction from the set of north, east, south or west applies ($C_4 = \{N, S, E, W\}$). This system allows inferences only for cases like $W +_c W = W$.

7.1.2. Four Directions with Identity

Introducing an identity element 0_c removes the restriction in the input values for the direction function:

$$\text{for all } P_1, P_2 \text{ exist } c\text{: dir}(P_1, P_2) = c, \text{ with } c \text{ in } C_4' = \{N, S, E, W, 0_c\}. \quad (53)$$

A quarter turn is defined as $q(N) = E$, $q(E) = S$, $q(S) = W$, $q(W) = N$, and $q(0_c) = 0_c$ and the inverse $-_c$ is defined by two quarter turns $-_c(c1) = q^2(c_1)$.

A result cannot be deduced for all combinations of input values of the direction system based on the set C_4' or its subset C_4. From the total of 25 different combinations, one can only infer 13 cases exact and four approximate; summarized in Table 9 (lower case letters indicate approximate reasoning):

Table 9. Composition for four cone-shaped directions.

	N	*E*	*S*	*W*	0
N	*N*		*o*		*N*
E		*E*		*o*	*E*
S	*o*		*S*		*S*
W		*o*		*W*	*W*
0	*N*	*E*	*S*	*W*	0

7.1.3. Directions with Eight Directional Symbols and Identity

The set of eight cardinal directions and an identity symbol $C_8' = \{N, NE, E, SE, S, SW, W, NW, 0_c\}$. A turn of an eighth is defined as $e(N) = NE$, $e(NE = E$, $e(E) = SE, \ldots, e(NW) = N$, $e(0_c) = o_c$. The inverse is equal to four eighth turns $-_c(d) = e^4(d)$.

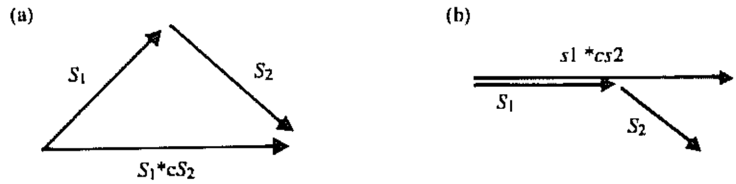

Figure 10. Two averaging rules

In addition to the rules from Table 4, two averaging rules [Equation (54) and (56)] are introduced (Figure 10), which allow composition of two directions that are each one eighth off. For example, SW combined with SE should result in S, or N combined with E should result in NE. Also, two directions are combined that are just one eighth turn apart, selecting one of the two (S combined with SE results in S, N combined with NW results in NW).

$$e(d) *_c -_e(d) = d \tag{54}$$

$$\text{with } -_e(d) = e^7(d), \text{ (one eight turn in the other direction)} \tag{55}$$

$$e(d) *_c d = d \quad \text{and} \quad d *_c e(d) = d \tag{56}$$

In this system, for all the 81 pairs of values (64 for the subset without 0_c), combinations can be inferred, but most of them only approximately. Only 24 cases (eight for the subset) can be inferred exactly, 25 result in a value of 0 and another 32 give approximate results. This is shown in Table 10, where lower case denotes Euclidean approximate inference.

Table 10. Composition for eight cone-shaped directions with averaging rules.

	N	*NE*	*E*	*SE*	*S*	*SW*	*W*	*NW*	0
N	*N*	*n*	*ne*	*o*	*o*	*o*	*nw*	*nw*	*N*
NE	*n*	*NE*	*ne*	*e*	*o*	*o*	*o*	*n*	*NE*
E	*ne*	*ne*	*E*	*e*	*se*	*o*	*o*	*o*	*E*
SE	*o*	*e*	*e*	*SE*	*se*	*s*	*o*	*o*	*SE*
S	*o*	*o*	*se*	*se*	*S*	*s*	*sw*	*o*	*S*
SW	*o*	*o*	*o*	*s*	*s*	*SW*	*sw*	*w*	*SW*
W	*nw*	*o*	*o*	*o*	*sw*	*sw*	*W*	*w*	*W*
NW	*nw*	*n*	*o*	*o*	*o*	*w*	*w*	*NW*	*NW*
0	*N*	*NE*	*E*	*SE*	*S*	*SW*	*W*	*NW*	0

7.2. Cardinal Directions Defined by Projections

7.2.1. Directions in Four Half-planes

Four directions can be defined, such that they are pair-wise opposites and each pair divides the plane into two half-planes (Figure 11). The direction operation assigns for each pair of points a composition of two directions, e.g. south *and* east, for a total of four different directions (Figure 12). This is an alternative semantic for the cardinal direction, which can be related to Jackendoff's principles of centrality, necessity and typicality [50, p. 121]. Peuquet pointed out that directions defined by half-planes are

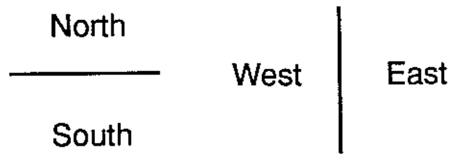

Figure 11. Two sets of half-planes

related to the necessary conditions, whereas the cone-shaped directions give the typical condition [17, p. 24].

Another justification for this type of reasoning is found in the structure geographic longitude and latitude imposes on the globe. Cone-shaped directions better represent the direction of 'going towards', whereas the 'half-planes' (or equivalent parts of the globe) better represents the relative position of points on the earth. However, the two coincide most of the time.

In this system, the two projections can be dealt with individually. They are labeled by the directions included, not the direction of the projection. Each of them has the exact same structure and the next subsection presents one case separately and then the combination is shown.

One projection and two directions

The N–S case, considered the prototype for the two cases E–W and N–S has the following axioms:

$$\text{for all } P_1, P_2 (P_1 \neq P_2): \text{dirs}_{ns}(P_1, P_2) = d_{ns} \text{ with } d_{ns} \text{ in } \{N, S\} \tag{57}$$

$$-_c N = S, -_c S = W \tag{58}$$

$$\text{for all } d \text{ in } \{N, S\}: d *_c d = d, \text{ which is } N *_c N = N, S *_c S = S) \tag{59}$$

and for E–W:

$$\text{for all } P_1, P_2 (P_1 \neq P_2): \text{dir}_{ew}(P_1, P_2) = \text{d}_{ew} \text{ with } d_{ew} \text{ in } \{E, W\} \tag{60}$$

$$-_c E = W, -_c W = E \tag{61}$$

$$\text{for all } d \text{ in } \{E, W\}: d *_c d = d, \text{ which is } E *_c E = E, W *_c W = W). \tag{62}$$

Combinations to four directions

The two projections in N–S and E–W form a single system, and for each path one of four combinations of directions from $D_4 = \{NE, NW, SE, SW\}$ is assigned. Two projections operations and their inverse combination operation lead to the definition

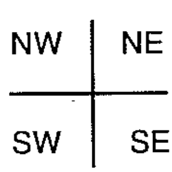

Figure 12. Directions defined by half-planes

of inverse and composition as the inverse or the composition of the components.

$$\text{projections:} \quad p_{ns}: d_4 \rightarrow d_{ns}, \qquad d_{ns} \text{ in } \{N, S\} \tag{63}$$

$$p_{ew}: d_4 \rightarrow d_{ew}, \qquad d_{ew} \text{ in } \{E, W\}$$

$$\text{combination:} \quad c: d_{ns} \text{ x } d_{ew} \rightarrow d_4 \tag{64}$$

$$c(p_{ns}(d), p_{ew}(d)) = d$$

$$\text{inverse:} \quad -_c(d) = c(-_c(p_{ns}(d)), -_c(p_{ew}(d))) \tag{65}$$

$$\text{composition:} \quad d_1 *_c d_2 = c(d_{ns}(d_1) *_c d_{ns}(d_2)\ d_{ew}(d_1) *_c d_{ew}(d_2)) \tag{66}$$

In this system, composition yields only a result in four cases and combinations, like NE $*_c$ NW which should approximately result in N, cannot be computed. This system is not very powerful, as only four of the 16 combinations can be inferred.

$$NE *_c NE = NE \qquad NW *_c NW = NW$$

$$SE *_c SE = SE \qquad SW *_c SW = SW$$

7.2.2. Directions with Neutral Zone

Here cardinal directions are defined such that points which are near to due north (or west, east, south) are not assigned a second direction, i.e. no decision is made if such a point is more east or west. This results in a division of the plane into nine regions, a central neutral area, four regions where only one direction letter applies and four regions where two are used. For the N–S projection the three values for direction are $d_{ns} = \{N, P, S\}$ and for the E–W direction they are $d_{ew} = \{E, Q, W\}$.

It is important to note that the width of the 'neural zone' is not determined explicitly. Its size is effectively fixed when the directional values are assigned and a decision is made that P_2 is north (not north-west or north-east) of P_1, assuming that these decisions are made consistently. (Similar arguments apply to the neural zone of cone-shaped directions.)

For each projection, a symbol is assigned. The inverse operations [Equations (61) and (62)] are extended by Equation (67), which follows from Equation (38). Projection and combination is defined analog to Equation (63) and (64) except for the

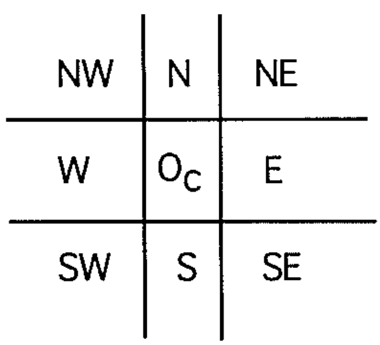

Figure 13. Directions with neutral zone

presence of an identity element.

$$\text{for all } P_1, P_2\text{: dir}_{ns}(P_1, P_2) = d_{ns} \text{ with } d_{ns} \text{ in } \{N, P, S\} \tag{66}$$
$$\text{for all } P_1, P_2\text{: dir}_{ew}(P_1, P_2) = d_{ew} \text{ with } d_{ew} \text{ in } \{E, Q, W\}$$

$$-_c(N) = S,\ -_c(S) = N,\ -_c(P) = P \tag{67}$$
$$-_c(E) = W,\ -_c(W) = E,\ -_c(Q) = Q$$

Composition and inverse are defined exactly as in Equations (65) and (66) as the combination of the composition (resp. the inverse) of the projections.

The composition of the two projections is such that for each path a pair of direction letters is assigned, with the abbreviations N for NQ, E for PE, 0_c for PQ, etc.

$$C_8' = d_{ns} \times d_{ew} = \{N, P, S\} \times \{E, Q, W\} = \{NE, NQ, NW, PE, PW, PQ, SE, SQ, SW\}$$
$$C_8'' = \{NE, N, NW, E, W, 0_c, SE, S, SW\}$$

With these symbols, the inverse and the composition operations are given as tables (again, lower case indicates approximate reasoning):

$d =$	NE	N	NW	E	W	0	SE	S	SW
$-_c d =$	SW	S	SE	W	E	0	NW	N	NE

Table 11. Inverse and composition for projection-based directions.

$*_d$	*N*	*NE*	*E*	*SE*	*S*	*SW*	*W*	*NW*	0
N	*N*	*N*	*NE*	*e*	*o*	*w*	*NW*	*NW*	*N*
NE	*NE*	*NE*	*NE*	*e*	*e*	*o*	*n*	*n*	*NE*
E	*NE*	*NE*	*E*	*SE*	*SE*	*s*	*o*	*n*	*E*
SE	*e*	*e*	*SE*	*SE*	*SE*	*s*	*s*	*o*	*SE*
S	*o*	*e*	*SE*	*SE*	*S*	*SW*	*SW*	*w*	*S*
SW	*w*	*o*	*s*	*s*	*SW*	*SW*	*SW*	*w*	*SW*
W	*NW*	*n*	*o*	*s*	*SW*	*SW*	*W*	*NW*	*W*
NW	*NW*	*n*	*n*	*o*	*w*	*w*	*NW*	*NW*	*NW*
0	*N*	*NE*	*E*	*SE*	*S*	*SW*	*W*	*NW*	0

In the projection-based system of directions with a neutral zone, we can deduce a value for all input values for composition (81 total), 56 cases are exact reasoning, resulting in a definite direction symbol, nine cases yield a value of 0_c and another 16 cases are approximate.

This system, however, is not associative, as discussed in subsection 5.3.

$$(N +_c N) +_c S = N +_c S = 0_c, \quad \text{but} \quad N +_c (N +_c S) = N +_c 0_c = N \tag{68}$$

7.3. Assessment

The powers of the four direction cone-shaped and the four half-plane directional systems are both very limited and are, as such, of not much interest. The eight direction cone-shaped and the four projection-based directional systems are comparable. Each system uses nine directional symbols, eight cone directions plus identity or the Cartesian product of three values (two directional symbols and one identity symbol) for each projection. There are less reasoning rules in the projection-based system, as each projection is handled separately with only two rules. The cone-shaped system uses two additional approximate rules which are then combined with the other ones.

For both the cone-shaped and the projection-based directions the subset without the identity element is very limited. The addition of an identity element increases the deductive power of the system significantly. Both systems with an identity element allow composition for any pair of input values (81 different pairs) at least approximately.

In comparison we find that the projection-based directions result in more cases of exact deductions than the cone-shaped (32 *vs.* 16) and that it less often yields the value 0_c (9 *vs.* 25). Considering the definite values (other than 0_c) deduced, we see small differences for eight pairs, depending on the definition of the approximate reasoning rules [Equation (54) and (56)]. These are cases in which the cone shaped system uses the approximate rule $d *_c e(d) = d$ and the projection-based deduces each projection separately and arrives at an exact result. Not counting this insignificant difference, we observe that both systems produce the same deductions if they both result in a definite value.

Both systems violate some of the desired properties, specifically associativity (as discussed in subsection 5.3), but this seems only to affect extreme cases (see deduction 4 from the example data at the end of this section).

An implementation of these rules and a comparison of the computed compositions with the exact values was done and confirms the theoretical results. Comparing all possible 10^6 combinations in a grid of 10 by 10 points (with a neutral zone of size 3 for the projection-based directions) shows that the results for the projection-based directions are correct in 50% of the cases and for cone-shaped directions in only 25%. The result 0_c is the outcome of 18% of all cases for the projection-based, but 61% for the cone-shaped directions. The direction-based system with an extended neutral zone produces in 2% of all cases a result that is a quarter turn off, otherwise the deviation from the correct result is never more than one-eighth of a turn, namely in 13% of all cases for cone-shaped and 26% for projection-based direction systems. In summary, the projection-based system of directions produces in 80% of all cases a result that is within 45° of the correct value.

With these definitions, the following deductions follow from the information in the introductory example:

1. *north* (*Chicago, St Louis*) $*_c$ *south* (*New Orleans, St Louis*) ⇒ *north* (*Chicago, New Orleans*) because
   $north *_c -_c (south) = north *_c north = north.$
2. *north* (*Chicago, St Louis*) $*_c$ *east* (*St Louis, San Francisco*) ⇒ *north-east* (*Chicago, San Francisco*).
3. *east* (*St Louis, San Francisco*) $*_c$ *south* (*Los Angles, San Francisco*) ⇒ *east* (*St Louis, San Francisco*) $*_c$ *north* (*San Francisco, Los Angeles*) ⇒ *north-east* (*St Louis, Los Angles*).
4. *north-east* (*Chicago, San Francisco*) $*_c$ *south* (*Los Angeles, San Francisco*) ⇒
   north-east (*Chicago, San Francisco*) $*_c$ *north* (*San Francisco, Los Angeles,*) ⇒
   north-east (*Chicago, Los Angeles*)
   or:
   north (*Chicago, St Louis*) $*_c$ *north-east* (*St Louis, Los Angeles*) ⇒ *north-east* (*Chicago, Los Angeles*).

8. Combinations of Distance and Directions

Reasoning about distances and directions between points can be considered, using the above described systems individually, or a set of rules is developed that takes into account the interactions between distance and direction reasoning. In this section these two alternatives are briefly sketched.

8.1. Separate Combination of Components

The distance and the direction components from a spatial relation can be handled separately, using the formulae previously developed. Any combination from the distance or direction systems previously described can be selected. For example, combining cone-shaped directions with four directions with two-step distances produces the symbols $M_8 = \{N, E, S, W, n, e, s, w\}$, where capitals stand for far south, far north, etc., and lower cases for close north, close east, etc.

For deduction, the direction and the distance value is first separated by two 'projection' operations, which yield the directions $\{N, E, S, W\}$ and the distances $\{C, F\}$. Then the composition for both distance and direction is calculated and the result combined into a single value, using the combination operation (which is the inverse of the projections).

$$\text{proj}_{\text{dist}}(m) = c, \text{ with } c \text{ in } \{N, E, S, W\}$$
$$\text{proj}_{\text{dir}}(m) = d, \text{ with } d \text{ in } \{C, F\}.$$
$$\text{combine } (c, d) = m$$
$$\text{such that the combine } (\text{pro}_{\text{dir}}(m), \text{proj}_{\text{dist}}(m)) = m. \tag{69}$$

If we compose the three-step distance with the projection-based directions with identity, we have a system with a total of 27 differentiated spatial relations [Equation (70)] and 729 different combinations. All combinations of inputs yield results for *inversion* and *composition* using appropriate projections and the inference tables in subsections 6.2 and 7.2.2.

$$M'_{24} = C'_8 \times S_3 = \{N, NE, E, SE, S, SW, W, NW, 0_c\} \times \{C, M, F\}. \tag{70}$$

8.2. Interaction Between Distance and Direction Reasoning

The reasoning about distances and directions does interact. Not all the combinations that the previous example constructs are meaningful:

1. The path 0_p has the direction 0_c [Equation (37)] and the distance 0_d [Equation (16)], and there are no other combinations. This reduces the number of differentiated (legal) distance/direction combinations in the last example to 17; 16 definite distances plus $0_d/0_c$. There are, therefore, only 289 combinations to infer.
2. Inferences, which result in a direction of 0_c but a definite distance (i.e. those different from 0_d) are not acceptable. This is typically the result of using the rule Equation (41)] $a *_c -_c a = 0_c$). No such problem exists with distance reasoning, because no composition except $0_d +_d 0_d$ yields 0_d.

In general, combinations where the distance values are the same are resolved quite well, whereas cases where the distance values are different may result in values that are not Euclidean exact. Roughly, three quarters of all combinations are resolved acceptably, whereas one quarter of the combinations may be quite wrong.

8.3. Further Integration

The integration of distance and direction inferences can be increased and precision of the results increased. Separating distance and direction information by 'projection' from the input values, applying independent rules for the inference of distance and direction and combining the results at the end, ignores the potential influence between distance and direction inference. It reveals the shortcomings of both distance and direction reasoning as outlined before: distance reasoning does (strictly) only deal with the case where the two paths are parallel; direction reasoning does only work (strictly) when the two paths have the same length.

It is clearly possible to construct tables for the addition of qualitative distance symbols under the assumption that the two paths are anti-parallel (i.e. parallel, but in the inverse direction). Such an operation could be seen as a 'subtraction' of distances (but would not necessarily be an inverse of the addition $+_d$). Other tables can be constructed if the two paths form a right angle, or 1 or 3 quarter turns. Distance reasoning then takes into account the angle between the two paths, computed as the (symbolic) difference between the qualitative directions of the path and the distance addition table selected accordingly. The same concept can be applied to the direction reasoning. Tables are modified to take into account the difference in length of the paths (computed from the qualitative distance information available).

9. Conclusions

This paper introduced a system for inference rules for completely symbolic, qualitative spatial reasoning with cardinal directions and distances. We have first stressed the need for symbolic, qualitative reasoning for spatial problems that does not translate the problem to analytical geometry. The systems investigated are capable of resolving any composition of directional inference using a few rules. Based on the example in the introduction, we cannot only assert that *Chicago* is *north* and *far* from *New Orleans*, but also that *Chicago* is *north-east* from *San Francisco* and *St Louis north-east* from *Los Angeles*, etc.

Geometric intuition based on the image-schema 'path' lead to the definition of an algebra of paths (Table 2), which contains most of the relevant properties one expects from distance and cardinal directions between point-like objects. Specifically, two operations, namely *inverse* and *composition*, were defined. The algebraic approach suggested the introduction of a zero path, i.e. the path from a point to itself, which acts as the identity element for composition. Starting with this common core of rules, distance and direction reasoning was examined and additional desirable rules found. The identity path leads to the definition of a zero distance and a zero direction. The set of rules are summarized in the Tables 3 and 4.

Several systems for distance and direction reasoning can be constructed on this base and a few consistent systems were explored and compared with coordinate-based, analytical computations. Because the rules are defined to yield a single value, not a set of possible values, the result is sometimes only approximate. The notion of 'Euclidean exact' and 'Euclidean approximate' as properties of a qualitative spatial reasoning system are defined, such that a deduction rule is called 'Euclidean exact' if it produces always the same results as Euclidean geometry operations would.

Systems for reasoning with distances with any number of steps were investigated and it was found that for each number of steps exactly three consistent variants exist. Tables for composition for the systems with three steps (close, medium, far) and four steps (very close, close, far, very far) are enclosed. All distance reasoning is Euclidean-approximate.

For cardinal directions, two different approaches were investigated. One is based on cone-shaped (or triangular) directions, the other deals with directions in two orthogonal projections. Both systems, if dealing with four cardinal directions are very limited and most combinations of values cannot be composed. The systems with eight direction symbols (no zero direction) are more powerful, but only the systems with an identity element (zero direction) can resolve all cases.

The most important result from the comparison of the two direction systems is that they do not differ substantially in their power or their inferences. There are a few cases the projection-based system can deal with and deduce Euclidean exact results, but the cone-based system produces only Euclidean approximate results. There are a few cases of approximate reasoning, where the cone based system deduces the direction 0_c whereas the projection-based system deduces a definite cardinal direction. Overall, the projection-based system with neutral zone is more powerful, produces better results and is also slightly simpler to define.

Combining distance and direction reasoning, one becomes aware of the limitations of separating distance and direction reasoning in individual chains of inference. We pointed out a method to improve the precision of the reasoning, without having to deal with the composition of all possible inputs (as suggested by Munkerjee [51]). Taking advantage of the invarience of composition of paths under rotation, the distance composition tables are additionally indexed for the difference of the directions (and vice versa). In lieu of a table for the results of all input values, only a small set of smaller tables are necessary.

9.1. Future work

There is little previous work on qualitative spatial reasoning about distances and directions in geographic space and several different directions for work remain open.

- *Hierarchical systems for qualitative reasoning*—a system for reasoning with distances differentiating only two or three steps of farness is quite limited. Depending on the circumstances a distance appears far or near compared to others. One could thus construct a system of hierarchically nested neighborhoods, wherein all points are about equally spaced. Such a system can be formalized and may quite adequately explain some forms of human spatial reasoning.
- *Directions of extended objects*—the discussion in this paper dealt exclusively with point-like objects. This is a severe limitation and avoided the difficult problem of explaining directions between extended objects. Peuquet [46] tried to find an algorithm that gives the same result as 'visual inspection'; however, visual inspection does not yield consistent results. It might be useful to see if sound rules, like the above developed ones, may be used to resolve some of the ambiguities.

- *What system of qualitative reasoning do humans use?*—We can also ask which one of the systems proposed humans use? For this, one has to see in which cases different systems produce different results and then test human subjects to see which one they employ. This may be difficult for the cone and projection-based direction system, as their deduction results are very similar. Care must be applied to control for the area of application, as we suspect that different types of problems suggest different types of spatial reasoning.

Acknowledgements

Comments from Matt McGranaghan and Max Egenhofer and the reviewers contributed to improve the presentation and I appreciate their help. Funding from NSF for the NCGIA under grant SES 88-10917, from Intergraph Corporation, and Digital Equipment Corporation is gratefully acknowledged.

References

1. B. Kuipers (1978) Modeling spatial knowledge. *Cognitive Science* **2**, 129–153.
2. J. McCarthy (1977) Epistomological problems of artificial intelligence. In: *IJCAI-77*. Cambridge, Massachusetts, pp. 1038–1044.
3. R. T. Antony (1990) A hybrid spatial/object-oriented DBMS to support automated spatial, hierarchical, and temporal reasoning. In: *Advances in Spatial Reasoning* (S.-s. Chen, ed.) Ablex, Norwood, New Jersey, pp. 63–132.
4. P. A. Piazza & F. Pessaro (1990) A cognitive model for a "smart" geographic information system. In *GIS World Annual Source Book*. GIS World.
5. R. Abler (1987) The National Science Foundation National Center for Geographic Information and Analysis. *International Journal of Geographical Information Systems* **1**, 303–326.
6. NCGIA (1989) The U.S. National Center for Geographic Information and Analysis: An overview of the agenda for research and education. *International Journal of Geographical Information Systems* 2(3), 117–136.
7. P. D. Try & J. R. Benton (1988) Critical investment strategy guidance for artificial intelligence development in support of the realistic battlefield (terrain/weather/obscurants)—a report of the workshop on Artificial Intelligence Research for Exploration of the Battlefield Environment. STC Technical report 3087. Science and Technology Corp., Hampton, Virginia.
8. P. J. Hayes (1985) The second naive physics manifesto. In: *Formal Theories of the Commonsense World* (J. R. Hobbs and R. C. Moore, eds) Ablex, Norwood, New Jersey, pp. 1–36.
9. J. Hobbs & R. C. Moore (eds) (1985) *Formal Theories of the Commonsense World* Ablex, Norwood, New Jersey.
10. D. S. Weld & J. de Kleer (eds) (1990) *Qualitative Reasoning about Physical Systems* Morgan Kaufmann, San Mateo, California.
11. J. A. Bier (1976) *Map of O'ahu: The Gathering Place* Univeristy of Hawaii Press, Honolulu.
12. D. McDermott & E. Davis (1984) Planning routes through uncertain territory. *Artificial Intelligence* **22**, 107–156.
13. S. Dutta (1990) Qualitative spatial reasoning: a semi-quantitative approach using fuzzy logic. In: *Design and Implementation of Large Spatial Databases* (A. Buchmann, O. Günther, T. R. Smith and Y.-F. Wang, eds) Springer Verlag, New York, pp. 345–364.
14. A. U. Frank (1990) Spatial concepts, geometric data models and data structures. In: *GIS Design Models and Functionality* (D. Maguire, ed.) Midlands Regional Research Laboratory, University of Leicester, Leicester, U.K.

15. A. U. Frank & D. M. Mark (1991) Language issues for geographical information systems. In: *Geographic Information Systems: Principles and Applications* (D. Maguire, D. Rhind and M. Goodchild, eds) Longman, London.
16. D. M. Mark & A. U. Frank (1990) Experimental and formal representations of geographic space and spatial relations. Technical report. NCGIA.
17. D. M. Mark, A. U. Frank, M. J. Egenhofer, S. M. Freudschuh, M. McGranaghan & R. M. White (1989) Languages of spatial relations: initiative two specialist meeting report. Technical report 89–2. National Center for Geographic Information and Analysis.
18. G. Lakoff. (1987) *Women, Fire, and Dangerous Things: What Categories Reveal About the Mind* University of Chicago Press, Chicago, Illinois.
19. D. M. Mark, S. Svorou & D. Zubin (1987) Spatial terms and spatial concepts: geographic, cognitive, and lingustic perspectives. In: *International Geographic Information Systems (IGIS) Symposium: The Research Agenda.* II-101–II-111. Arlington, Virginia, pp. II101–111.
20. S. Dutta (1988) Approximate spatial reasoning. In: *First International Conference on Industrial and Engineering Applications of Artificial Intelligence and Expert Systems.* Tullahoma, Tennessee, pp. 126–140.
21. L. A. Zadeh (1974) Fuzzy logic and its application to approximate reasoning. In: *Information Processing* North-Holland Publishing Company.
22. S. K. Chang, Q. Y. Shi and C. W. Yan (1987) Iconic indexing by 2-D string. *IEEE Transactions on Pattern Analysis and Machine Intelligence* PAMI **9**, 413–428.
23. S.-K. Chang, E. Jungert & Y. Li (1990) The design of pictorial databases based upon the theory of symbolic projection. In: *Design and Implementation of Large Spatial Databases* (A. Buchmann, O. Günther, T. R. Smith and Y.-F. Wang, eds) Springer Verlag, New York, pp. 303–324.
24. S. K. Chang, C. W. Yan, T. Arndt & D. Dimitroff (1988) An intelligent image database system. *IEEE Transactions on Software Engineering* May, 681–688.
25. D. Hernández (1990) Relative representation of spatial knowledge: the 2-D case. Report FKI-135-90. Technische Universität München, Munich, Germany.
26. J. F. Allen (1983) Maintaining knowledge about temporal intervals. *Communications of the ACM* **26**, 832–843.
27. H. W. Guesgen (1989) Spatial reasoning based on Allen's temporal logic. Technical report TR-89-049. Berkely, California International Computer Science Institute.
28. A. Munkerjee (1989) Representing spatial relations between arbitrary oriented objects. TAMU-89-019. Texas A & M University, Huston, Texas.
29. A. Munkerjee (1990) Representation for modelling functional knowledge in geometric structures. TAMU-90-004. Texas A & M University, Huston, Texas.
30. J. de Kleer & J. S. Brown (1985) A qualitative physics based on confluence. In: *Formal Theories of the Commonsense World* (J. R. Hobbs and R. C. Moore, eds) Ablex, Norwood, New Jersey, pp. 109–184.
31. S. M. Kosslyn (1980) *Image and Mind* Harvard University Press, Cambridge Massachusetts.
32. Z. Pylyshyn (1981) The imagery debate: analogue media versus tacit knowledge. *Psychological Review* **87**, 16–45.
33. L. Buisson (1990) Reasoning on space with object-centred knowledge representation. In: *Design and Implementation of Large Spatial Databases* (A. Buchmann, O. Günther, T. R. Smith and Y.-F. Wang, eds) Springer Verlag, New York, pp. 325–344.
34. D. G. Bobrow, S. Mittal & M. J. Stefik (1986) Expert systems: perils and promise. *Communications of the ACM* **29**, 880–894.
35. B. Kuipers (1983) The cognitive map: could it have been any other way? In: *Spatial Orientation* H. L. Pick and L. P. Acredolo, eds) Plenum Press, New York, pp. 345–359.
36. Freska (1991) Qualitative spatial reasoning. In: *Spatial Languages* (D. M. Mark and A. U. Frank, eds) Kluver, Amsterdam.
37. W. R. Tobler & S. Wineberg (1971) A Cappadocian speculation. *Nature* **231**, 39–42.
38. M. McGranaghan (1991) Schema and object matching as a basis for interpreting textual specifications of geographical location. In: *Spatial Languages* (D. M. Mark and A. U. Frank, eds) Kluver, Amsterdam.

39. E. C. Zeeman (1962) The topology of the brain and visual perception. In: *The Topology of 3-Manifolds* (M. K. Fort, ed.) Prentice Hall, Englewood Cliffs, New Jersey, pp. 240–256.
40. F. S. Robert (1973) Tolerance geometry. *Notre Dame Journal of Formal Logic* 14, 68–76.
41. D. V. Pullar & M. J. Egenhofer (1988) Toward the definition and use of topological relations among spatial objects. In: *Third International Symposium on Spatial Data Handling* (D. F. Marble, ed.) International Geographical Union, Commission on Geographical Data Sensing and Processing, Sydney, Australia, pp. 225–242.
42. W. Kuhn (1989) Interaktion mit raumbezogenen Informationssystemen—Vom Konstruieren zum Editieren geometrischer Modelle. Dissertation No. 8897 thesis, Swiss Federal Institute of Technology.
43. G. Lakoff & M. Johnson (1980) *Metaphors We Live By* University of Chicago Press, Chicago, Illinois.
44. A. Gill (1976) *Applied Algebra for the Computer Sciences* Prentice-Hall, Englewood Cliffs, New Jersey.
45. J. Freeman (1975) The modelling of spatial relations. *Computer Graphics and Image Processing* 4, 156–171.
46. D. Peuquet & C.-X. Zhan (1987) An algorithm to determine the directional relationship between arbitrarily-shaped polygons in a plane. *Pattern Recognition* 20, 65–74.
47. P. Struss (1990) Problems of interval-based qualitative reasoning. In: *Qualitative Reasoning About Physical Systems* (D. S. Weld and J. deKleer, eds) Morgan Kaufmann, Palo Alto, California, pp. 288–305.
48. A. Herskovits (1986) *Language and Spatial Cognition—An Interdisciplinary Study of the Propositions in English* Studies in Natural Language Processing Cambridge University Press, Cambridge.
49. V. Robinson & R. Wong (1987) Acquiring approximate spatial representations of spatial relations: theory and results of test. In: *Auto-Carto* 8 (N. R. Chrisman, ed.) ASPRS and ACSM, Baltimore, Massachusetts, pp. 604–622.
50. R. Jackendoff (1983) *Semantics and Cognition* MIT Press, Cambridge, Massachusetts.
51. A. Munkerjee (1990) A Qualitative model for space. TAMU-90-005. Texas A & M University, Huston, Texas.