Spatial Relations Between 3D Objects:
The Association Between
Natural Language, Topology, and Metrics

Jennifer L. Leopold  Chaman L. Sabharwal  Katrina J. Ward
leopoldj@mst.edu  chaman@mst.edu  kjw26b@mst.edu
Computer Science Department
Missouri University of Science & Technology
Rolla, MO 65409 USA

Abstract—With the proliferation of 3D image data comes the
need for advances in automated spatial reasoning. One specific
challenge is the need for a practical mapping between spatial
reasoning and human cognition, where human cognition is
expressed through natural-language terminology. With respect to
human understanding, researchers have found that errors about
spatial relations typically tend to be metric rather than
topological; that is, errors tend to be made with respect to
quantitative differences in spatial features. However, topology
alone has been found to be insufficient for conveying spatial
knowledge in natural-language communication. Based on
previous work that has been done to define metrics for two lines
and a line and a 2D region in order to facilitate a mapping to
natural-language terminology, herein we define metrics
appropriate for 3D regions. These metrics extend the notions of
previously defined terms such as splitting, closeness, and
approximate alongness. The association between this collection of
metrics, 3D connectivity relations, and several English-language
spatial terms was tested in a human subject study. As spatial
queries tend to be in natural language, this study provides
preliminary insight into how 3D topological relations and metrics
correlate in distinguishing natural-language terms.

Keywords—Spatial Reasoning, Natural-Language Processing,
Region Connection Calculus, 3D Images

I. INTRODUCTION

In tandem with increases in pervasive mobile computing
and the proliferation of 3D image data comes the need for
advances in automated spatial reasoning. One of the particular
challenges is the need for a practical mapping between
qualitative and quantitative spatial reasoning and human
cognition, the latter being expressed principally through
natural-language terminology. With respect to human understanding, errors about spatial relations typically tend to
be metric rather than topological [1, 2]; however, topology
alone has been found to be insufficient for conveying spatial
knowledge in natural-language communication [3, 4]. The
consensus is that topology matters while metrics refine [5]. To
accommodate natural-language spatial queries, an effective
interface between automated spatial reasoning and natural
language requires an appropriate blend of natural language,
topology, and metrics.

Based on the work that has been done to define metrics for
two lines [4] and a line and a 2D region [3] with topological
relations in order to facilitate a mapping to natural-language
terminology, herein we define metrics appropriate for two 3D
regions and the topological connectivity relations used in
VRCC-3D+ [6, 7, 8]. These metrics extend the notions of what
previous authors [3, 4, 9] have referred to as: splitting (i.e.,
how much is in common between two objects), closeness (i.e.,
how far apart parts are), and approximate alongness (i.e.,
a combination of splitting and closeness). The association
between this collection of metrics, 3D connectivity relations,
and several English-language spatial terms was tested in a
human subject study. The results of that study provide
preliminary insight into how the 3D topological relations and
metrics correlate in distinguishing natural-language terms.

The paper is organized as follows. Section II briefly
discusses the region connection calculus, VRCC-3D+, and the
topological relations pertinent to this study. Section III defines
the metric relations for splitting, closeness, and approximate
alongness, which are similar in concept to those that have been
proposed for a line and a 2D region [3], but are significantly
redefined to be appropriate for objects in 3D space. Section IV
identifies dependencies between the topological relations and
the metrics, as well as intra-relationships within the metrics.
Section V examines associations between the topological
relations, the metrics, and various natural-language terms
based on the results of a human subject experiment. Section
VI outlines directions for future work, followed by a summary
and conclusions in Section VII.

II. TOPOLOGICAL RELATIONS

A. Mathematical Preliminaries

R³ denotes the three-dimensional space endowed with a
distance metric. Here the mathematical notions of subset,
proper subset, equal sets, empty set (∅), union, intersection,
universal complement, and relative complement are the same
as those typically defined in set theory. The notions of
neighborhood, open set, closed set, limit point, boundary,
interior, exterior, and closure of sets are as in point-set
topology. The interior, boundary, and exterior of any region
are disjoint, and their union is the universe.
A set is connected if it cannot be represented as the union of disjoint open sets. For any non-empty bounded set A, we use symbols $A^c$, $A^1$, $A^b$, and $A^e$ to represent the universal complement, interior, boundary, and exterior of a set A, respectively. Two regions A and B are equal if $A^c = B^c$, $A^b = B^b$, and $A^e = B^e$ are true. For our discussion, we assume that every region A is a non-empty, bounded, regular closed, connected set without holes; specifically, $A^b$ is a closed curve in 2D, and a closed surface in 3D.

B. Region Connection Calculus

Much of the foundational research on qualitative spatial reasoning is based on a region connection calculus (RCC) that describes 2D regions (i.e., topological space) by their possible relations to each other. Most notable is the RCC8 model [10] which defines the following eight relations (illustrated in Fig. 1): disconnected (DC), externally connected (EC), partial overlap (PO), equality (EQ), tangential proper part (TPP), non-tangential proper part (NTPP), converse tangential proper part (TPPc), and converse non-tangential proper part (NTPPc). Topological relations in a region connection calculus are typically defined using first-order logic (as in the work of Randell, Cui, and Cohn [10]) or using the 9-Intersection model [11] which looks at whether the intersections between the interiors, exteriors, and boundaries of two regions are empty or non-empty.

Figure 1. RCC8 relations.

Whereas a 2D object is in a plane, a 3D object is in space. The simplest examples of 3D objects are a pyramid, a cube, a cylinder, and a sphere. A concave pyramid is a complex, simply connected 3D object. Since concave objects can be partitioned into convex objects, for all practical purposes, we work with convex objects. For the rest of this discussion, we will base our analysis on convex objects; in particular, spheres are used in our natural-language human study.

VRCC-3D+ [6, 7, 8] is the implementation of a region connection calculus that qualitatively determines the spatial relations between 3D objects, both in terms of connectivity and obscuration. The VRCC-3D+ connectivity relations are named the same as in RCC8; however, the VRCC-3D+ connectivity relations are calculated in 3D rather than 2D. Fifteen obscuration relations also are defined in VRCC-3D+. Considered from a 2D projection, each VRCC-3D+ obscuration relation is a refinement of basic concepts of no obscuration, partial obscuration, and complete obscuration. A composite VRCC-3D+ relation specifies both a connectivity relation and an obscuration relation. Herein our discussion is limited to the VRCC-3D+ connectivity relations, which heretofore will be referred to as topological relations; application of this work to the VRCC-3D+ obscuration relations is beyond the scope of this paper. For a more in-depth discussion of VRCC-3D+, including how it compares to the other RCC models, see [6, 7, 8].

III. METRIC PROPERTIES

Metric relations focus on the quantitative differences in spatial features between the two regions or objects being compared; typically, these relations are expressed as scaled (normalized) volumes, areas, distances, lengths, or size differences. Three metric concepts were introduced in the study that compared a line and a 2D region [3]: (1) splitting, which determines how much is in common between two objects; (2) closeness, which determines how far apart parts of the two objects are; and (3) approximate alongness, which combines splitting and closeness along the boundaries of objects. Metric relations for these concepts were defined in another study [4] to be suitable for comparing two lines. Metrics for additional spatial concepts such as angular direction and overlap have been considered for objects in 2D space [4]. Here we adopt and adapt some of the basic metrics to apply to objects in 3D space. The intersection between the convex volumes can be a volume, a surface, a line segment, or a point. We use normalized metric values, (i.e. dimensionless units) to distinguish between them. For each of our metric relations $M(A, B)$ for objects A and B there is a converse metric relation denoted $M^{-1}(A, B)$, defined by $M^{-1}(A, B) = M(B, A)$.

A. Splitting

Interior volume splitting (IVS) considers the scaled (normalized) part of one object that is split by the other object. Here boundary does not matter as the volume of the boundary is zero.

$$\text{IVS}(A, B) = \frac{\text{volume}(A \cap B)}{\text{volume}(A)}$$

Exterior volume splitting (EVS) describes the proportion of one object’s interior that is split by the other object’s exterior. Again, boundary does not matter.

$$\text{EVS}(A, B) = \frac{\text{volume}(A \cap B^c)}{\text{volume}(A)}$$

Observe that $\text{volume}(A) = \text{volume}(A \cap B) + \text{volume}(A \cap B^c)$, hence $\text{EVS}(A, B) = 1 - \text{IVS}(A, B)$.

We define another splitting metric to specifically examine the proportion of the boundary of one object that is split by the boundary of the other object; we denote this metric BS for boundary splitting. It should be noted that there are two versions of the equations for this metric. If the intersection of the objects’ boundaries is a line, then the metric should be computed as a scaled length; otherwise, the intersection must be a surface area and the metric should be computed as a normalized area.

$$\text{BS}(A, B) = \frac{\text{area}(A^b \cap B^b)}{\text{area}(A^b)} \text{ or } \frac{\text{edgeLength}(A^b \cap B^b)}{\text{edgeLength}(A^b)}$$
B. Closeness

For the 3D object shown in Fig. 2(a), let \( N_\delta(A^b) \) be the exterior 3D \( r \)-neighborhood of \( A^b \) of radius \( r > 0 \) (Fig. 2(b)), and let \( N_\delta(A^b) \) be the interior 3D \( r \)-neighborhood of \( A^b \) of radius \( r > 0 \); see Fig. 2(c). The smaller the value of \( r \), the closer the objects are; thus \( N_\delta(A^b) \cap B \neq \emptyset \) determines the minimum volume common to the exterior neighborhood of \( A^b \) and \( B \). Let \( \delta \) be the least upper bound of \( r \) for which \( N_\delta(A^b) \cap B \neq \emptyset \). Then \( \delta \) is the closest distance between the exterior neighborhood of \( A^b \) from \( B \). Let \( \Delta_e A = N_\delta(A^b) \) be the minimum volume in a pre-specified exterior \( r \)-neighborhood \( (r > \delta) \) of \( A^b \). Similarly let \( \Delta_e A = N_\delta(A^b) \) be the minimum volume in a pre-specified interior \( r \)-neighborhood \( (r > \delta) \) of \( A^b \). The value of \( \delta \) is specified by the application. In general, for numerical calculations, it is approximately one percent of the sum of the radii of two spheres. Intuitively, \( r \) accounts for the thickness of the boundary for the object.

![Figure 2](image)

Figure 2. (a) An object, (b) the exterior neighborhood of the boundary of the object, and (c) the interior neighborhood of the boundary of the object.

Considering the interior neighborhood of an object, we define interior volume closeness (IVC) as follows:

\[
IVC(A,B) = \frac{volume(\Delta_e A \cap B)}{volume(\Delta_e A)}
\]

Similarly, we can consider the exterior neighborhood of an object, and can define a metric for exterior volume closeness (EVC) as follows:

\[
EVC(A,B) = \frac{volume(\Delta_e A \cap B)}{volume(\Delta_e A)}
\]

This metric is a measure of how much of the interior neighborhood of \( A^b \) covers \( B \); this is the extent to which the interior neighborhood of \( A \) is closer to \( B \).

C. Approximate Alongness

Approximate alongness is essentially a combination of splitting and closeness. There are three types of metrics in this case: (1) normalized boundary of an object common to the interior neighborhood of the boundary of the other object; (2) normalized boundary of an object common to the exterior neighborhood of the boundary of the other object; and (3) normalized boundary of an object common to the other entire object. Similar in concept to the inner and outer approximate alongness metrics that were proposed for a line and a 2D region [3], we can define metrics to assess the relative amount of the boundary of one object that is shared with the neighborhood of the other object. Interior boundary (IB) is the proportion of the boundary of one object along the interior neighborhood of the other object. Exterior boundary (EB) is the proportion of the boundary of one object along the exterior neighborhood of the other object. As was the case for boundary splitting (BS), there are two versions of the equations for each of these metrics. If the intersection is a line, then the metric should be computed as the normalized length; otherwise, the intersection must be a surface area and the metric should be computed as the normalized area.

\[
IB(A,B) = \frac{area(A^b \cap \Delta_e B)}{area(A^b)} \text{ or } \frac{edgeLength(A^b \cap \Delta_e B)}{edgeLength(A^b)}
\]

\[
EB(A,B) = \frac{area(A^b \cap \Delta_e B)}{area(A^b)} \text{ or } \frac{edgeLength(A^b \cap \Delta_e B)}{edgeLength(A^b)}
\]

We define another alongness metric, which we shall call boundary alongness (BA), to assess how much of the scaled boundary of one object is shared with the entirety of the other object. Here the edge length of intersection is the length of the arc or line segment of intersection. The edge length of the boundary (i.e., the denominator) corresponds to the edge length of the super arc or line segment that contains the intersection edge (i.e., the numerator).

\[
BA(A,B) = \frac{area(A^b \cap B)}{area(A^b)} \text{ or } \frac{edgeLength(A^b \cap B)}{edgeLength(A^b)}
\]

IV. CONSTRAINTS AND DEPENDENCIES AMONG TOPOLOGICAL RELATIONS AND METRICS

For a topological relation, if the value of a metric varies as the two objects vary then there is a dependency between that topological relation and metric. For example, if objects \( A \) and \( B \) are disconnected (i.e., \( DC(A,B) \)), then \( A \cap B \) always will be empty, and consequently both \( IVS(A,B) \) and \( IVSc(A,B) \) always will be 0. Hence the values of \( IVS(A,B) \) and \( IVSc(A,B) \) are not dependent on \( A \) and \( B \); that is, there is no dependency between the relation \( DC(A,B) \) and metrics \( IVS(A,B) \) and \( IVSc(A,B) \). In contrast, if objects \( A \) and \( B \) partially overlap (i.e., \( PO(A,B) \)), then \( A \cap B \) will be non-empty; the values of both \( IVS(A,B) \) and \( IVSc(A,B) \) will vary depending upon the particular configuration of the partially overlapping objects \( A \) and \( B \), so we say that there is a dependency between the relation \( PO(A,B) \) and the metric \( IVS(A,B) \) (as well as \( IVSc(A,B) \)). Table 1 lists these dependencies where a highlighted box indicates that there is a dependency relation between the metric (listed in the top row) and the topological relation (listed in the left column).

<table>
<thead>
<tr>
<th>Metric</th>
<th>Relation 1</th>
<th>Relation 2</th>
<th>Relation 3</th>
<th>Relation 4</th>
<th>Relation 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>IVS</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
</tr>
<tr>
<td>IVSc</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
</tr>
<tr>
<td>BA</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
</tr>
</tbody>
</table>

When a metric is not dependent upon a topological relation, the value of that metric always will be unchanged; it turns out to always be equal to 0 or always be equal to 1 regardless of the particular objects that are in that topological configuration. However, when a metric is dependent upon a topological relation, we can deduce the most restrictive
Table 1. Dependencies among VRCC-3D+ topological relations and metric relations. Highlighted boxes denote metric relations that are dependent upon a topological relation.

<table>
<thead>
<tr>
<th>DC(A,B)</th>
<th>ECV(A,B)</th>
<th>EVS(A,B)</th>
<th>EVSc(A,B)</th>
<th>BA(A,B)</th>
<th>IB(A,B)</th>
<th>BS(A,B)</th>
<th>RS(A,B)</th>
<th>EVC(A,B)</th>
<th>EVCc(A,B)</th>
<th>IB(A,B)</th>
<th>Esd(A,B)</th>
<th>FB(A,B)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>2</td>
</tr>
</tbody>
</table>

Table 2. Value constraints among VRCC-3D+ topological relations and metric relations.

<table>
<thead>
<tr>
<th>DC(A,B)</th>
<th>ECV(A,B)</th>
<th>EVS(A,B)</th>
<th>EVSc(A,B)</th>
<th>BA(A,B)</th>
<th>IB(A,B)</th>
<th>BS(A,B)</th>
<th>RS(A,B)</th>
<th>EVC(A,B)</th>
<th>EVCc(A,B)</th>
<th>IB(A,B)</th>
<th>Esd(A,B)</th>
<th>FB(A,B)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>2</td>
</tr>
</tbody>
</table>

On examining Table 1 for all sixteen metrics (eight metrics and their converses), it is determined that some of the metrics have the same set of dependencies for the topological relations. For example, in Table 1, IVS(A,B) and EVS(A,B) have identical values in all rows. Two metrics are equivalent if they have the identical dependencies corresponding to all topological relations. A set of equivalent metrics is called an equivalence class of metrics. A set of equivalence classes forms a partition of the set of metrics. Upon complete examination of Table 1 in terms of the dependencies, the set of metrics can be partitioned into eleven equivalence classes: \{IVS(A,B), EVS(A,B), IVC(A,B)\}, \{IVSc(A,B), EVSc(A,B), IVCc(A,B)\}, \{BA(A,B), BAc(A,B)\}, \{BS(A,B), BSc(A,B)\}, \{EVC(A,B), EVCc(A,B)\}, \{IB(A,B), IBc(A,B)\}, \{Esd(A,B), FB(A,B)\}. Table 3 shows the number of topological relations that differ (in terms of dependencies) for each pair of these equivalence classes. This can be used as the basis for creating a conceptual neighborhood graph (CNG) to get a sense of the conceptual closeness of the metrics; see Fig. 3. The CNG is a connected minimal spanning graph. Edges selected to form the CNG are highlighted in Table 3.

The metrics at mutual distance 1 are: (1) metrics that refer to interior volume separation and interior volume closeness, and (2) metrics for exterior volume and boundary (interior or exterior) closeness. Of the remaining metrics, the Snapshot Model [12] is used to include all metrics at distance 2 in the connected minimal spanning graph.

The CNG based on topological-metric dependencies has no direct correlation with the CNG for RCC8 relations, which can be constructed based on the 9-Intersection model; in the latter, two RCC8 relations are conceptual neighbors if one relation can be deformed into the other (i.e., via translation, rotation, or scaling) without encountering any other relation in between. The topological-metric dependencies CNG is useful for easily identifying conceptual similarity between the various metrics; for example, the concept of interior volume splitting (IVS) is more similar to the concept of interior boundary (IB) than it is to exterior boundary closeness (EVC).

A more practical use of the topological-metric dependency CNG is as a means of measuring the semantic distance between two natural-language terms. The similarity between terms \(t_1\) and \(t_2\) could be measured by finding two nodes in the graph that correspond to metrics applicable to both terms, and finding the minimal-cost path between those nodes. This would need to be done for each pair of metrics that apply to the two terms, summing all those metric path costs, and dividing by the number of paths. If the result for terms \(t_1\) and \(t_2\) is smaller than the result for terms \(t_1\) and \(t_3\), then terms \(t_1\) and \(t_2\) are conceptually more similar than terms \(t_1\) and \(t_3\) (i.e., \(t_1\) and \(t_3\) are more likely to be considered synonyms than are \(t_1\) and \(t_2\)).

V. ASSOCIATION BETWEEN THE NATURAL-LANGUAGE TERMS, TOPOLOGICAL RELATIONS, AND THE METRICS

The primary objective of this work is to define a collection of metrics and topological relations that will facilitate a mapping to natural-language terminology for 3D objects, and thereby establish a practical mapping between qualitative and quantitative spatial reasoning and human cognition. To that end, insight can be gleaned from conducting human subject experiments in which human perception of images and words can be associated with the mathematical notions represented by the topological relations and metrics.

A. Related Experiments

Human subject experiments were conducted by Shariff, Egenhofer, and Mark [3] wherein subjects were given several

constraint on the range of values for that metric; namely, we know whether the metric will produce a value > 0 or ≥ 0 regardless of the size and shape of the two objects being compared. For example, if object A is a non-tangential proper part of object B (i.e., NTPP(A,B)), then A can be intersected with the exterior neighborhood of B that is empty, so E(B,A) will always be equal to 0, regardless of the particular objects A and B that are in that spatial configuration. However, B can be intersected with the exterior neighborhood of A that is non-empty, so E(B,A) will always be ≥ 0 depending on the pre-defined radius r. Table 2 lists these constraints (i.e., 0, 1, > 0, and ≥ 0), color-coded as specified by the legend in the table.

On examining Table 1 for all sixteen metrics (eight metrics and their converses), it is determined that some of the metrics have the same set of dependencies for the topological relations. For example, in Table 1, IVS(A,B) and EVS(A,B) have identical values in all rows. Two metrics are equivalent if they have the identical dependencies corresponding to all topological relations. A set of equivalent metrics is called an equivalence class of metrics. A set of equivalence classes forms a partition of the set of metrics. Upon complete examination of Table 1 in terms of the dependencies, the set of metrics can be partitioned into eleven equivalence classes: \{IVS(A,B), EVS(A,B), IVC(A,B)\}, \{IVSc(A,B), EVSc(A,B), IVCc(A,B)\}, \{BA(A,B), BAc(A,B)\}, \{BS(A,B), BSc(A,B)\}, \{EVC(A,B), EVCc(A,B)\}, \{IB(A,B), IBc(A,B)\}, \{Esd(A,B), FB(A,B)\}. Table 3 shows the number of topological relations that differ (in terms of dependencies) for each pair of these equivalence classes. This can be used as the basis for creating a conceptual neighborhood graph (CNG) to get a sense of the conceptual closeness of the metrics; see Fig. 3. The CNG is a connected minimal spanning graph. Edges selected to form the CNG are highlighted in Table 3.

The metrics at mutual distance 1 are: (1) metrics that refer to interior volume separation and interior volume closeness, and (2) metrics for exterior volume and boundary (interior or exterior) closeness. Of the remaining metrics, the Snapshot Model [12] is used to include all metrics at distance 2 in the connected minimal spanning graph.
sentences each containing a natural-language spatial term such as “connects” and asked to draw a picture depicting that specified spatial relation between a park (i.e., a 2D region) and a road (i.e., a line or a curve). The drawings were then analyzed to measure the metrics that had been defined for a 2D region and a line. The subjects also were asked which of fifteen line-region topological relations they most closely associated with the natural-language spatial term in each sentence. Among the conclusions drawn from that study were that: (1) for the majority of the natural-language spatial terms, the topological relation is a more important influence (i.e., distinguishing feature) than any of the metrics, and (2) several natural-language spatial terms fall under the same topological relation, but have different metric values [3].

A similar experiment was conducted by Xu [4] whereby human subjects were given pictures of two lines and sentences containing a natural-language spatial term; the collection of terms used were not exactly the same as those used in the line-region study [3]. The subject was asked to rank his/her agreement as to whether the term described the spatial configuration in the picture. As in the line-region study, the actual values of the metrics for each spatial configuration used in the experiment were measured and included in the dataset. However, it is important to note that the collection of metrics used in the line-region study was considerably different than those examined in the line-region study: the only metrics that were conceptually common to both studies were for three types of splitting. The line-region study also differed from the line-region study in its consideration of topological relations: (1) the collection of topological relations simply consisted of the entries in the 9-Intersection matrix (i.e., the intersection of each object’s interior, exterior, and boundary with that of the other object), and (2) each natural-language spatial term was pre-classified with applicable topological relations; the human subjects were not asked whether or not they thought that a certain topological relation applied to a particular natural-language term. One of the conclusions drawn from this study was that in most cases using topological and metric properties together produces better results for distinguishing natural-language terms than using only topological properties [4].

B. Design and Analysis of Experiment with 3D Objects

B.1. Design of the Experiment

To investigate the association between natural-language spatial terms, topological relations, and metrics for 3D objects, we conducted a human subject experiment with design and analysis aspects similar to those of the aforementioned studies. 119 human subjects (88 male, 31 female) were given a test consisting of 48 questions. Each question contained an image of two 3D spheres, one blue and one green, in the spatial configuration of disconnected (DC), externally connected (EC), partial overlap (PO), non-tangential proper part (NTPP), or tangential proper part (TPP); see Fig. 4. The converse relations TPPc and NTPPc were not tested as they would just be the reverse cases of TPP and NTPP. The relation EQ was not tested because we anticipated that the subjects might be confused if they could not detect both a green sphere and a blue sphere (i.e., because the spheres are “equal”). Each question also included a sentence of the form “The blue sphere term the green sphere,” where term was one of the 26 natural-language spatial terms listed in Table 4; with the exception of “disconnected”, none of the topological relations (e.g., “tangential proper part”) was used as a term. We did use some of the same terms that were used in the line-region [3] and line-line [4] studies. The subject was asked to rank his/her agreement as to whether the term described the spatial configuration in the picture, choosing from seven rankings: ‘strongly agree’, ‘agree’, ‘somewhat agree’, ‘neutral’, ‘somewhat disagree’, ‘agree’, and ‘strongly disagree.’ We subsequently categorized the answers of ‘strongly agree’, ‘agree’, and ‘somewhat agree’ as ‘yes’; all other answers were categorized as ‘no.’ Not every natural-language term and topological relation pair was tested; for example, in the case of two disconnected spheres, we deemed that it would have been highly unlikely that anyone would have said that one sphere “encloses” the other sphere. The frequency of ‘yes’ and ‘no’ responses for the cases that were tested are shown in Table 4.

In most cases, there was a very clear consensus as to whether the majority of the subjects agreed or disagreed that a
particular natural-language term corresponded to a topological relation (as indicated by the blue-highlighted entries in Table 4 showing ≥72% frequency, and to a lesser extent the orange-highlighted entries in Table 4 showing ≥65% frequency). However, there were six cases where the decision was fairly evenly split (i.e., approximately 50% agreeing/disagreeing and approximately 50% disagreeing/agreeing): “is connected to”, “cuts across”, and “intersects” for partial overlap; “enters” and “goes into” for tangential proper part; and “goes into” for non-tangential proper part. These cases (highlighted in green in Table 4) are indicative of a natural-language term that is ambiguous for describing a particular topology.

As was concluded in the line-region study [3], we found that there were some natural-language spatial terms for which more than one topological relation applied (e.g., “is contained within”, “encloses”, “is inside of”, “is outside of”, and “is within”). It was noted in the line-region study that in such cases the metric values associated with a natural-language term could be different [3]. We computed the metric values for each of the five spatial configurations used in our study. Using k-means clustering (SimpleKMeans in the WEKA software, http://www.cs.waikato.ac.nz/ml/weka), we too found that some of the natural-language terms that mapped to more than one topological relation had different values for their metrics.

B.2. Analysis of the Experiment

Analysis of the 3D object experiment dataset was conducted with the objective of investigating the extent of the association between the natural-language terms, the topological relations, and the metrics. Each row of the dataset consisted of a topological relation (DC, EC, PO, TPP, or NTPP), the eleven metric equivalence classes with the dependency values for that particular topological relation, the calculated value of the metric for the depicted spatial configuration, and a natural-language spatial term. For this part of the analysis, we removed all responses (rows) where the subject did not agree that the spatial term described the topological relation that had been depicted between the two 3D objects (spheres in this case). We also performed the testing on smaller datasets, each containing the results for only a single natural-language term. Analysis was done both with and without considering the topological relation as an attribute, for both the smaller datasets and the complete dataset as a whole.

We first used multinomial logistic regression (SimpleLogistic in the WEKA software) to see which of the eleven metric equivalence classes would be the most likely predictors of (and hence the most important for distinguishing) the natural-language terms. Our results showed that the classes \{IVS(A,B), EVS(A,B), IVC(A,B)\}, \{IVSc(A,B), EVSc(A,B), IVCc(A,B)\}, and \{BS(A,B), BSc(B,A)\} were most significant in distinguishing natural-language terms based on both the metric dependency constraints and the metric values. All other metric classes either had zero weight or were dependent on other metric equivalence classes. Using logistic regression alone, the resulting model was able to correctly classify the natural-language terms using these metrics 79% of the time. These three metric classes were equally important; when we removed any one or two of them, the model would use one of the other metrics in this group and achieve the same accuracy. However, if we removed all three, the accuracy dropped noticeably. When we included the topological relations in the dataset, it made no difference in the results under this model. When we tested classification of the natural-language terms in the dataset using the topological relation alone without the metric attributes, the number of correctly classified terms dropped to only 13% using the logistic model.

We then examined the C4.5 [13] decision trees that could be built for each of our datasets using J48 in the WEKA software. Recall that each node of a C4.5 decision tree contains the attribute of the dataset that most effectively splits the instances into one class or another [13]. What we determined in every case was that, in the absence of the topological relation, these same three metric classes (e.g., \{IVS(A,B), EVS(A,B), IVC(A,B)\}, \{IVSc(A,B), EVSc(A,B), IVCc(A,B)\}, and \{BS(A,B), BSc(B,A)\}) always were included when building the decision trees, and that the accuracy of correctly classified natural-language terms ranged from 76% to 98% for datasets containing a single natural-language term. An example for a case where we tested the dataset as a whole and eliminated the topological term as an attribute is shown in Fig. 5. Although the accuracy was low on the total dataset, this example demonstrates the significance of the three metric classes (e.g., \{IVS(A,B), EVS(A,B), IVC(A,B)\}, \{IVSc(A,B), EVSc(A,B), IVCc(A,B)\}, and \{BS(A,B), BSc(B,A)\}) for defining multiple natural-language terms. The notation of the form \(N/E\) next to a node in the tree represents \(N\) as the number of instances that reached that node and \(E\) as the number of instances that differed in the value identified at that node. The first rule in the tree shown in Fig. 5 would be interpreted as “if the average value of the metrics in class \{IVSc(A,B), EVSc(A,B), IVCc(A,B)\} = 0.333333 and the average value of the metrics in class \{BS(A,B), BSc(A,B)\}...
Table 4. Frequency of responses for natural-language spatial terms used in human subject experiment.

<table>
<thead>
<tr>
<th>Term</th>
<th>DC - yes</th>
<th>DC - no</th>
<th>EC - yes</th>
<th>EC - no</th>
<th>PO - yes</th>
<th>PO - no</th>
<th>TPP - yes</th>
<th>TPP - no</th>
<th>NTTP - yes</th>
<th>NTTP - no</th>
</tr>
</thead>
<tbody>
<tr>
<td>is adjacent to</td>
<td>27%</td>
<td>73%</td>
<td>91%</td>
<td>9%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>is connected to</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>76%</td>
<td>24%</td>
<td>50%</td>
<td>50%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>is contained within</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>81%</td>
<td>19%</td>
<td>85%</td>
<td>15%</td>
</tr>
<tr>
<td>crosses</td>
<td></td>
<td></td>
<td>13%</td>
<td></td>
<td></td>
<td></td>
<td>87%</td>
<td>72%</td>
<td>28%</td>
<td></td>
</tr>
<tr>
<td>cuts</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>30%</td>
<td>70%</td>
<td></td>
</tr>
<tr>
<td>cuts across</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>47%</td>
<td>53%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>cuts through</td>
<td></td>
<td></td>
<td>35%</td>
<td></td>
<td></td>
<td></td>
<td>85%</td>
<td>25%</td>
<td>71%</td>
<td></td>
</tr>
<tr>
<td>is disconnected from</td>
<td></td>
<td></td>
<td></td>
<td>91%</td>
<td>9%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>divides</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>20%</td>
<td>80%</td>
<td>17%</td>
<td>83%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>encloses</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>77%</td>
<td>23%</td>
<td>83%</td>
<td>17%</td>
</tr>
<tr>
<td>enters</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>50%</td>
<td>50%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>goes across</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>69%</td>
<td>31%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>goes into</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>29%</td>
<td>71%</td>
<td>50%</td>
<td>50%</td>
<td>52%</td>
<td>48%</td>
</tr>
<tr>
<td>goes through</td>
<td></td>
<td></td>
<td>9%</td>
<td>91%</td>
<td>41%</td>
<td>59%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>is in</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>82%</td>
<td>18%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>is inside of</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>82%</td>
<td>18%</td>
<td>87%</td>
<td>13%</td>
</tr>
<tr>
<td>intersects</td>
<td></td>
<td></td>
<td>24%</td>
<td>70%</td>
<td>40%</td>
<td>52%</td>
<td>40%</td>
<td>66%</td>
<td>38%</td>
<td>62%</td>
</tr>
<tr>
<td>merges with</td>
<td></td>
<td></td>
<td>34%</td>
<td>66%</td>
<td>31%</td>
<td>69%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>is near</td>
<td></td>
<td></td>
<td>42%</td>
<td>58%</td>
<td>94%</td>
<td>6%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>is outside of</td>
<td></td>
<td></td>
<td>87%</td>
<td>13%</td>
<td>75%</td>
<td>25%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>splits</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>18%</td>
<td>82%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>touches</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>80%</td>
<td>12%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>transects</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>20%</td>
<td>80%</td>
<td>21%</td>
<td>79%</td>
</tr>
<tr>
<td>transverses</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>13%</td>
<td>81%</td>
<td>10%</td>
<td>20%</td>
</tr>
<tr>
<td>is within</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>79%</td>
<td>21%</td>
<td>86%</td>
<td>14%</td>
</tr>
</tbody>
</table>

---Classifier model (full training set)---

J48 pruned tree

------------------------

IVSc = 6.333333
  | BS = 0: disconnected (308.0/196.0)
  | BS = 0.03%: 1Near (598.0/484.0)
  | BS = 0.07%: 1Near (308.0/196.0)
  | BS = 1.78: ISNear (0.0)
IVSc = 6.3422: EVS(A,B) = 6.38966: intersects (586.0/502.0)
IVSc = 6.666667: Isinside (672.6/569.0)

Number of leaves: 7
Size of the tree: 9

Time taken to build model: 0.05 seconds

---Stratified cross-validation---

Summary

<table>
<thead>
<tr>
<th>Metric</th>
<th>Correctly Classified</th>
<th>Incorrectly Classified</th>
<th>Overall Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>IVSc</td>
<td>430</td>
<td>2330</td>
<td>17.3297%</td>
</tr>
<tr>
<td>EVS(A,B)</td>
<td>436</td>
<td>2330</td>
<td>17.3297%</td>
</tr>
<tr>
<td>IVSc</td>
<td>430</td>
<td>2330</td>
<td>17.3297%</td>
</tr>
<tr>
<td>IVSc</td>
<td>430</td>
<td>2330</td>
<td>17.3297%</td>
</tr>
</tbody>
</table>

Figure 5. J48 pruned tree for distinguishing natural-language terms based on metric equivalence class values.

= 0 then A is disconnected from B” where “is disconnected from” is a natural language term (not the topological relation DC). Due to attribute naming constraints in WEKA, metric class {IVSc(A,B), EVS(A,B), IVSc(A,B)} was named simply IVSc, and BS was used for metric class {BS(A,B), BSc(A,B)}.

The certainty of this particular rule is not high (35.5%); of 304 data instances that reached this node in the decision tree built using J48, only 196 of those instances actually agree with this decision. In general, the datasets that contained multiple natural-language terms had much lower classification accuracy (than the datasets that contained a single natural-language term) because there were instances where multiple natural-language terms mapped to multiple topological relations, and hence multiple metric equivalence classes.

Our J48 analysis found that every natural-language term used some combination of the aforementioned three metric classes in its decision. If we removed these metrics from the dataset, the classification accuracy drastically dropped. Adding the topological relations to a dataset did not change the accuracy results; however, we did note that if we removed {EVS(A,B)} from a dataset, J48 would replace it in the tree with a topological relation and reduce the tree size, making it slightly more efficient computationally.

What the results of this experiment tell us is that there are three metric equivalence classes that can fairly accurately define the majority of the natural-language terms in this dataset for 3D objects based on metric values (when the dataset contains a single natural-language term); similar results were obtained based simply on metric dependencies with topological relations. The other metrics were either dependent on these three metric equivalence classes, or were of very little significance in the final decision. In fact, two of the metric equivalence
classes, \{EB(A,B)\} and \{EBc(A,B)\}, were never used in any of the decision trees and had zero weight in the logistic regression analysis. We also found that using the topological relation with the metrics made no difference in terms of accuracy, although it did make computation of the decision (i.e., the natural-language term determination) slightly more efficient in a few cases, namely those involving EVS(A,B). Using only the topological relations (no metrics) did not improve the results for distinguishing one natural-language term from another; see Fig. 6 for the J48 decision tree that results from making decisions based only on the five topological relations tested in this experiment (DC, EC, NTPP, PO, and TPP).

These results are in contrast to those reported in the line-line and line-region studies wherein topology was found to play a more important role in distinguishing natural-language terms. At this preliminary point in our research, we cannot attribute this difference simply to the 3D nature of the spatial topology; further experimentation is necessary whereby, for example, measurement of our metrics for more examples of each spatial configuration should be analyzed for possible correlations. At this stage we also cannot conclude that only the metrics in the equivalence classes \{IVS(A,B), EVS(A,B), IVC(A,B)\}, \{IVSc(A,B), EVSc(A,B), IVCc(A,B)\}, and \{BS(A,B), BSc(B,A)\} are valid for 3D objects.

VI. FUTURE WORK

Some of the metrics presented herein are simplifications; they do not consider all relevant aspects of the topology for 3D objects. While the 9-Intersection model is sufficient to determine the topological connectivity, it is not sufficient to determine the qualitative extent of connectivity. For example, PO(A,B) embodies that A∩B, A∩B, and A∩B, are all nonempty, but it does not quantify the precise or qualitative commonality; we need metrics. Since the metrics measure the commonality in addition to the topological connectivity, for some of the metric equations we need to consider additional parameters such as the interior, exterior, boundary, interior neighborhood, and exterior neighborhood for each object. This will be explored in our subsequent future research.

In the future we also will present new metrics for the obscuration relations of VRCC-3D+, and investigate how the connectivity and obscuration relations combined with the metrics affect distinguishing natural-language spatial terms. With that integration of connectivity and obscuration, we expect practical applications of this work will include robotic navigation via voice control as well as natural-language user interfaces for 3D spatial querying.

VII. SUMMARY

Topology alone has been found to be insufficient for conveying spatial knowledge in natural-language communication. Based on previous work that has been done to define metrics for two lines and a line and a 2D region in order to facilitate a mapping to natural-language terminology, herein we defined metrics appropriate for 3D regions. The association between this collection of metrics, 3D connectivity relations, and several English-language spatial terms was tested in a human subject study. We found three metric equivalence classes that could define the natural-language terms in our experiment dataset for 3D objects. In contrast to the results reported for line-line and line-region studies, we found that using the topological relation with the metrics made no difference in terms of the accuracy in defining natural-language terms. This work is too preliminary to attribute this as a phenomenon of 3D topology. However, we believe our work is an interesting starting point in a world that is being saturated with 3D data and thus is in need of automated spatial reasoners with a natural-language interface.

REFERENCES


