Qualitative Linguistic Terms and Geographic Concepts

Quantifiers in Definitions

Eleni Tomai and Marinos Kavouras

Cartography Laboratory, School of Rural, and Surveying Engineering,
National Technical University of Athens
etomai@mail.ntua.gr, mkav@survey.ntua.gr

Short Title. Quantifiers in geographic definitions

Keywords. Qualitative linguistic terms, semantic relations, geospatial ontologies, quantifier theory

Corresponding Author.

Eleni Tomai

Cartography Laboratory, School of Rural, and Surveying Engineering,
National Technical University of Athens, Zografos Campus
9, Iroon Polytechniou 15780
Athens, Greece
Abstract. Definitions of categories in existent geospatial ontologies are an invaluable source of information because they provide us with essential knowledge about concepts and their properties. A closer examination reveals that definitions also contain supplementary linguistic items, which are mainly qualitative expressions, such as quantifiers. This inclusion of modifiers in definitions affects the way values are assigned to the categories’ properties (semantic properties and relations). This paper introduces a methodology for a) representing the essence of qualitative information to clarify the identity relations among categories, and b) assessing their semantic similarity in order to disambiguate the taxonomic structure of existent geospatial ontologies.
1 Introduction

To what extent structures and transport networks must cover a land, so that it qualifies as a member of category “Continuous Urban Fabric”? For how long must an area be covered by water to qualify as a member of category “Swamp”? According to CORINE LC (2000), a land cover nomenclature, in “Continuous Urban Fabric” most of the land is covered by structures and the transport network. According to DIGEST (2000), a “Swamp” is a low lying saturated area covered with water all or most of the year. These definitions do not provide a straightforward answer to the problem of category membership as depicted from the questions hereinabove, not to mention that the use of vague linguistic terms, such as quantifiers, complicates ontological analysis.

In this paper, we introduce a methodology for modeling Qualitative Linguistic Terms (QLTs), which can be found in definitions of categories in existing geospatial ontologies. According to Guarino (2002) an ontology is a shared vocabulary and a specification of its intended meaning. Ontologies may differ in complexity and formality and can range from lightweight (a simple glossary of terms) to heavyweight (axiomatized, formal) ones. Fonseca and Martin (2004), advocate that the design of information system ontologies must be aided by hermeneutic analysis while Smith (2003) argues that the construction of robust ontologies has to involve conceptualizations “transparent to reality”, that is, Science should not exclude common-sense. In our case, the ontologies that we analyze are taxonomies and terminological ones where categories are described by an agreed shared lexicon. Richer ontologies than those examined herein are not common in the GIS community, however when new
geospatial ontologies are to be generated several tenets should be followed as discussed by Tomai and Kavouras (2004).

Prior work by Kokla and Kavouras (2002), and Tomai and Kavouras (2003) has focused on the determination of the semantic properties and relations from such definitions. Semantic properties and relations can be considered as the combination of: a) properties of the categories and b) values for these properties. The problem with QLTs is that they affect property values in a non-straightforward manner. The present analysis underlines the need to set the values of properties, in a clear-cut way, therefore, to explicitly define the semantics (semantic relations and properties) of geospatial concepts.

Hence, the objective of the research is to explore ways of assigning numeric values to qualitative linguistic terms, which can be found in definitions of geospatial concepts, and to further use these values when determining similarity among these concepts. The intention is to model these qualitative linguistic terms, in a way to formally express the semantics of the geospatial concepts’ definitions. The current research consists in the clarification of the identity relationships of geospatial categories and the disambiguation in the taxonomic structure of existing ontologies. The straightforward establishment of properties for geospatial categories facilitates any integration procedure. In addition, the modeling of modifiers accounts better for the semantic similarity of the categories since it reveals subtle dissimilarities among them.

More specifically, the next section of this paper deals with theoretical matters of qualitative linguistic expressions found in definitions. Firstly, we defend the use of definitions in the process of educing semantic relations, and we identify the types of the included qualitative linguistic terms. The focus of the paper is to analyze and model the
category of modifiers called Generalized Quantifiers, referred to as quantifiers hereinafter.

We examine four existent geospatial ontologies/ categorizations: the European's Environmental Agency’s CORINE LC (2000), DIGEST (2000), the American National Standard SDTS (1997), and GMap (2000), all of which contain definitions of geographic categories. The third section of the paper details the characteristics of the examined categorizations and provides information on the included quantifiers.

In section four, we present a methodology for turning qualitative expressions into quantitative information, which can be farther, applied to establish similarity among categories. In this analysis, we employ the classic linguistic theory of Generalized Quantifiers and methods from the field of computational linguistics, which can represent QLTs schematically, and are used to assess QLT similarity. Section 5 gives a practical account of the results of the methodology, while the last section summarizes the method presented herein, and sets the ground for further research.

2 Definitions and Qualitative Linguistic Terms

A survey of existing definitions of categories in geospatial ontologies shows that they contain significant lexical knowledge. From this information, we can determine semantic relations and properties for the geographic categories. However, an extensive examination of these definitions shows that they also include qualitative linguistic terms and expressions, which work as degree modifiers on the semantics of geographic concepts.
12.1 Defending the Use of Definitions

At this point, it is important to defend the value of definitions in geospatial ontologies. Linguists have not always shown full appreciation for definitions, and the fact remains, that lexicography’s definitions have many deficiencies. In our case, nevertheless, definitions contribute to ontological research since they usually are the only source of information about the concepts described by the geospatial ontologies/categorizations. What is more, definitions contain organized knowledge about the terms they describe, knowledge that can be extracted, analyzed, and formalized. According to the Cambridge International Dictionary of English (2001): ‘A definition is a statement that explains the meaning of a word or phrase’.

Therefore, a definition is considered as highly structured text, rich in the following types of knowledge, as defined by the participants of SIGLEX workshop, in Barriere (1997):

- **Lexical Knowledge** (rich semantic structure, associative patterns between words’ feature structure);
- **World Knowledge** (ontology, organization of concepts);
- **Semantic Knowledge** (word/ sentence/ text meanings, predicates).

In the case of analyzing concepts of a particular domain such as the geospatial domain, definitions are rich also in:

- **Domain-Specific Knowledge** (of a particular domain);
- **Pragmatic Knowledge** (information coming from context).

Many attempts, such as these by Barriere (1997), and van der Wende (1995), have focused on extracting semantic relations from larger lexical databases, such as dictionaries, under the axiom that the language used in dictionary definitions presents
syntactic patterns (syntactic structure of a definition) to express these relations. Most of
that research is conducted following the genus/differentia type of definition. This kind
of approach results in building a taxonomy, the is-a hierarchy, based on the genus.
Another approach can be found in WordNet 2.0 (2003). Significant relations, for the
structure of this lexical database, are, as discussed by Miller (1990): synonymy,
antonymy, hyponymy, hypernymy, meronymy, and holonymy.

2.2 Qualitative Linguistic Terms

Previous work by Kavouras et al. (2003), Kokla and Kavouras (2002), Tomai and
Kavouras (2003), has introduced the determination of semantic relations for geospatial
categories from their definitions using techniques of Natural Language Understanding.
While, in NLU, much research has been conducted in the direction of nouns as in van
der Wende (1995), and verbs as in Fellbaum (1990), their modifiers have not yet been
explored to such an extent.

Modifiers can be:

- Nouns functioning as adjectives, called nominals e.g., water transportation. For
  reasons of simplicity, we use the term “nominals” for expressions consisting of a
  head noun preceded by one or more noun specifiers.
- Adjectives that modify a noun assigning an attribute to it e.g., brackish water. These
  are also nominals but one or more adjective specifiers precede the head noun.
- Adverbs that modify verbs, adjectives, or other adverbs e.g., frequently includes,
  very large, quite often.

The first two kinds of modifiers can be distinguished from the third and are classified as
qualifiers, while the third is mostly referred to as adverbial qualifier.
Definitions for geospatial categories include but to a lesser extent another category of words that are categorized as quantifiers. Quantifiers, in logic, correspond to “for all”=∀ and “exists”= ∃. In Natural Language however, a small category of words correspond to how quantifiers act in logic. Examples of this category are: all, each, every, some, most, many, and several.

The existence of such words, both qualifiers and quantifiers, in definitions of geospatial categories, necessitates their modeling when elucidating semantic relations, since the assignment of their numeric values is not straightforward. Therefore, modifiers and quantifiers do not make explicit the properties involved. A skeptic would state, at this point, that QLTs present a deficiency of definitions for they do not add to meaning; on the contrary, they are futile verbal forms. An immediate reply to such a position is that QLTs are used widely in definitions with the purpose to differentiate concepts. Furthermore, conducting a definitions’ analysis using the “bag-of-words” approach\(^1\) and therefore, disregard certain aspects of qualitative information, imposes restrictions on conveying meaning.

This paper focuses on modeling Generalized Quantifiers in definitions of geographic concepts, and establishes similarity among these linguistic terms.

### 3 Geographic Categorizations Analyzed

As mentioned above, we analyze four (4) ontologies/ categorizations of geospatial information. These are: CORINE Land Cover (2000), DIGEST (2000), SDTS (1997),

\(^1\) The “bag of words” approach is a non-syntactic based approach of information retrieval, since word order and hierarchical relations are ignored. Each item is considered independently, regardless of syntactic relation, e.g. house representative vs. representative house are considered related in such an approach (Klavans et al.1997).
and GMap (2000). All these categorizations include definitions of geospatial concepts. The majority of these definitions are simple, complete sentences.

WordNet (version 2.0) could have been also included in the analysis. The reason for not doing so is that we wanted to keep this analysis to a limited size of entries. As it will be shown in the next section, the ontologies used, have a limited number of categories, which are to a certain extent manageable. Another problematic aspect of using a lexical database, such as WordNet, is that like all “dictionaries,” it favors polysemy. On the other hand, the selected ontologies describe each category univocally, that is, they only deal with one sense of the given term, which is an advantage in our case. Furthermore, the abovementioned categorizations provide information of a given domain, the geographic one, while WordNet is not domain-dependent.

3.3.1 Characteristics of the Examined Ontologies

This section presents, in brief, the characteristics of the four ontologies.

- CORINE Land Cover is a categorization schema that was established to provide consistent information about land cover for the member states of the EU. It has a three-level hierarchical structure and includes 44 category terms at the lowest level.

- DIGEST is an international exchange standard that has been designed to enable the transfer of Digital Geographic Information (DGI) between geographic information systems. The FACC (Feature and Attribute Coding Catalogue) contains features, attributes, and attribute values. The features of the FACC are organized in 10 categories, which are further divided in subcategories.

- The Spatial Data Transfer Standard (SDTS) gives definitions of spatial features for data transfer. These are described by the concepts: Entity Type, Entity Instance,
Attribute, and Attribute Value; and by the terms Standard and Included. The standard describes 200 entity types, and has no hierarchical structure.

- GMap established in 2000, contains specifications of spatial data. The Global Map Data Dictionary has specifications for vector and raster data.

4 3.2 Qualitative Linguistic Terms in Definitions - The case of Quantifiers

Quantifiers present a small linguistic category in Natural Language. The category consists of words that indicate quantity, which range from all to no. A quantifier included in a geographic definition is shown by the following example:

1. Swamp: A low lying saturated area covered with water all or most of the year, where accumulating dead vegetation does not rapidly decay … (DIGEST).

In the definition, the pattern cover with defines the semantic relation cover, which has water as value. In the meanwhile, the expression: all or most of the year, applies to semantic relation time. The use of most, however, implicates the assignment of a value to semantic relation time; the denoting property of being covered with water does not hold for "all times".

Consider also the following examples:

2. Alluvium: All unconsolidated fragmental material laid down by a stream (SDTS).
3. Airspace: Designated airspace within which some or all aircraft may be subjected to air traffic control (DIGEST).
Another group of quantifiers, that is widely used in the definitions of the four geospatial ontologies analyzed here, is shown in the following examples from GMap (International Geosphere Biosphere Program Land Cover Classification):

4. Closed Shrublands: lands with woody vegetation less than two meters tall with shrub canopy cover $>60\%$. The shrub foliage can be either evergreen or deciduous.

5. Open Shrublands: lands with woody vegetation less than two meters tall with shrub canopy cover $\textit{between 10-60\%}$. The shrub foliage can be either evergreen or deciduous.

In this case, we encounter a two-level problem.

- If the information of the percentage of the $\textit{canopy cover}$ between the two categories is overlooked, we end up with identical definitions for both of them. We want to avoid doing so; these two categories are evidently very similar, but still the establishers of the categorization have reasons to categorize them separately. In addition, from this analysis, it is obvious that the semantic property, which has the most significant role in defining these categories, is the percentage of the $\textit{canopy cover}$. We are therefore, obliged to consider the difference of the values of that property between the two categories.

- The information we have for the $\textit{canopy cover}$ is quantitative but we have to decide on a formalized way of encoding this kind of information when assessing similarity between the two categories. Similarity between categories can be established by comparing their descriptions as e.g., in Kavouras et al. (2003), and Rodriguez et al.
The challenge is how are we to include into the similarity measure, terms like more than, less than, between, equal to etc.

In section 4.2, we present a method to model quantifiers in a way that will address the matter of assessing their similarity.

5 3.3 The Semantics of Semantic Relations when a Modifier is involved

To make the need for the QLTs modeling more explicit, we present here the semantics of such expressions. The use of a QLT in conjunction with a semantic relation, in a sentence, as shown in the previous examples (1-5), denotes that the value of that property is overstated/understated (verbally) but is also “boosted”/ “deducted” (semantically). Therefore, the problem stated by this research is not merely a linguistic issue (that of wordiness) but a matter of how concepts are construed and therefore represented by natural language itself.

4 Modeling Quantifiers

For the modeling of quantifiers, involved in definitions, we base our study on the linguistic theory of Generalized Quantifiers by van der Does (1996), van der Does and van Eijck (1996), and van der Does and de Hoop (2001), and work conducted in the field of computational linguistics by Clark (2004a and 2004b) for the quantifiers automation.

The theory of Generalized Quantifiers adopts a relational view of quantifiers. This perspective, according to van der Does and van Eijck (1996), can be used as a translation procedure from Natural Language to logical representation. In other words, the theory seeks to define a Quantifiers Logics for the analysis of Natural Language.
Section 3.2 focused on an analysis of the major implications of the quantifiers in definitions, on the question of establishing similarity. It was decided to adopt the tree of numbers as in van der Does (1996) to represent quantifiers because certain properties of quantifiers can be perceived as regularities of tree patterns.

Apart from schematically representing quantifiers, the tree of numbers can be additionally used to represent quantifiers as strings of binary values\(^2\). We decided to further draw on these strings to determine similarity among quantifiers. More specifically, in section 4.1.2 we compare different quantifiers to *all* to assign them a numeric value.

### 6 4.1 Schematic Representation of Quantifiers; The Tree of Numbers

The tree of numbers represents quantifiers schematically. To use such a tree we must set first the truth conditions for quantifiers. A quantifier $Q$ has two arguments $X$ and $Y$ (e.g., *most birds fly. birds is arg. $X$ and fly is arg. $Y$*). Its truth depends on sets $\{X-Y\}$ and $\{X \cap Y\}$ for universe $E$. Table 1 shows truth sets of some quantifiers.

In the case of *some* and *most*, further explanations should be given for their truth sets.

- According to Wierzbicka (1996), in the case of *some*, it is better to keep our distances from *not all, which must be regarded*, as similar to *most*. In addition, she points out that *some* and the existential *there is* ($\exists$) should not be reduced to one notion from a semantic point of view. Therefore, in our case, we take the truth set of *some* to be $\{0 < |X \cap Y| \neq E\}$ and not $\{0 < |X \cap Y|\}$ which can be equal to $E$.

- *Most* is a higher-order quantifier; to model it using the tree of numbers, we introduce a formalization of its truth-value sets using first-order quantifiers instead.

---

\(^2\) The method for changing the tree of numbers into strings is called semantic automata and was introduced by van Benthem in 1983.
In the case of most, we use the disjunction of at most and at least as it is described in Clark (2004a) Alternatively, for most, we could also use more than half as suggested in van der Does and van Eijck (1996). Both approaches result in the same tree of numbers.

Once the restriction sets for the given quantifiers are defined, we can visualize quantifiers as a tree of numbers as follows. Figure 1 shows the general format of such a tree adapted to represent quantifiers. The “coordinates” of every node of the tree is the pair (i, j). If (i, j) is a pair number in the tree for quantifier Q, then i+j is the cardinality of universe E, i the cardinality of |X - Y|, j the cardinality of |X ∩ Y|. If a pair (|X - Y|, |X ∩ Y|) is an element of a quantifier, this is indicated by a “+” as in van der Does and de Hoop (2001).

Thus, a ‘+’ in position 1,2 of Figure 2, for statement: most birds fly, means that two birds fly from a total (universe E) of three birds; founded on that we can represent and visualize quantifiers. Figure 3 shows the corresponding trees for all, most, and some.

The trees include patterns, which, once identified, can provide the threshold of ending (closing) the trees to a finite sequence.

4.1.1 Modifying Trees to Strings and Finding their Sizes

The second goal is to explore ways of assessing similarity for the modeled quantifiers.

We, therefore, consider the automation of quantifiers’ trees that uses strings of binary values as in Clark (2004b).

Quoting from Partee and Borschev (2001, p. 7): “The quantifier’s (Q) automation

- Uses as an input a sequence of members of a set X, marked as to whether they are members of Y or not. Representation as a sequence of 0 and 1, 0 representing a member of |X - Y|, 1 a member of |X ∩ Y|, and
• **Gives as output, yes or no, representing whether or not $Q_{\exists}XY$ is true for the sequence thus far observed**.

In the tree of numbers automata, a ‘+’ in position $i, j$ of the tree, is represented by a string, which contains $i$ 0s and $j$ 1s. However, the problem is how to define the sum of numbers of 1s and 0s in the string, that is the size of that string.

According to van Benthem in Clark (2004a), in first order quantifiers, there will be a line in the tree of numbers, below which, the behavior of the quantifiers is the same as in the lines above. This is referred as the “Fraissé threshold” (ibid.) and gives an account of the geometry of quantifiers. The most important result of the threshold is that the tree of numbers can provide a finite pattern that reveals the behavior of the quantifiers.

Therefore, a closer examination of the resulting trees will give us an answer to the size of the string; for every represented quantifier, there is a pattern of pluses (+) and minuses (−) in the tree, that is repeated after some node.

With that in mind, the accepted string for:

• **All** is any string that contains only 1s, e.g., (1,1,1,1);

• **Most** is any string that contains more 1s than 0s, e.g., (1,1,1,0), (1,1,1,1);

• **Some** is any string that contains at least one 0, e.g., (1,0,0,0), (1,1,0,0), (1,1,1,0).

As it can be concluded from the accepted stings of **all**, **most**, and **some**, reveal the relationships between these quantifiers in terms of semantics. **Most** and **some** overlap because the former represents a quantity of *more than half*, rather than an unspecified one, as the latter does, while **all**, since it represents totality, is a refined case of **most**.
4.1.2 Similarity among Quantifiers

The previous section explored the capability of the tree of numbers in revealing the geometry of quantifiers and in ascertaining strings of binary values to account for their truth-value sets. Nevertheless, the modeling of quantifiers is not fully achieved. A method is, therefore, needed which assigns numerical values to them.

It has been decided to establish similarity between quantifiers in order to turn qualitative information into quantitative by comparing strings of the same size for each quantifier. It is important to bear in mind that in quantifier automation, the total number of 0s and 1s is what counts and not their order in the string.

We chose to compare every quantifier, to quantifier *all* because this is represented merely as a string of 1s. In addition, the term *all* may not be present in the definition of a concept, but still it is implied, as a category fully possesses a property. We can also compare percentage quantifiers to *all*, taking the latter equal to 100%. Moreover, the universality of quantifier *all* as a semantic prime, discussed by Wierzbicka (1996) makes it a safe choice for comparing every quantifier to it, for establishing similarity.

In this way, we determine the similarity $S$ between *all* and *most*; for strings of three\(^3\) we have:

- **All**: $(1,1,1)$ and
- **Most**: $(1,1,0)$ or $(1,1,1)$ so $\tilde{S}_{\text{all,most}} = 0.833$

The value of similarity $\tilde{S}$ is estimated by comparing only strings of a size that corresponds to the sum of $i+j$ in which the tree pattern has been firstly identified. For the case of *most*, the pattern is recognized in the fourth row of the tree where cardinality

\(^3\) We compare here strings of three because the pattern of *most* is first identified when $i+j$ is equal to three, that is, in the fourth row of the tree of numbers (Figure 3).
of universe $E = i+j$ is three, that is why we accept $S_{all\_most} = 0.833$. The values for similarity $S$ between all and the quantifiers included in the analysis are found in Table 2. Horn L.R., in Klein (1998) shows how the relations between quantifiers can be represented through two different scales, a positive, and a negative one. Horn’s schema of opposition is shown in Figure 4. The positive scale ranges from 0 to 1 while the negative from 0 to –1. While this approach provides an expressive way of representing the relations between quantifiers, it is not suitable for our ultimate goal, which is the degree of possessing a property in terms of positive numeric values (section 5). The right side of the schema with the negative scale, proved quite thorny for, what it means that a geographic concept possesses a certain property of a negative value?

The advantage of the, herein described, methodology of modeling quantifiers using the tree of numbers it that it consists in revealing the semantics of quantifiers. The tree of numbers approach is a method for proving several properties of quantifiers, on the linguistic side, such as reflexivity, symmetry, and monotonicity. While in terms of logical issues, this approach renders natural language quantifiers as relations between sets.

Another approach on modeling quantifiers giving them a numeric value is by using fuzzy sets. Diaz-Hermida et al. (2003) propose to translate the classic linguistic theory of Generalized Quantifiers into the fuzzy case, through a fuzzification mechanism. While Dvorak and Novak (2002) propose an application of fuzzy logic in broader sense, to extract linguistic knowledge from databases (when QLTs are included in queries) using if-then rules. On the same basis Ribeiro and Moreira (1999) present a fuzzy querying model (for databases) capable of handling various types of questions in a
natural language form (including quantifiers). These approaches however, do not provide an objective explanation of where the numbers come from. On the other hand, they do not take into consideration the relational properties of quantifiers as discussed hereinabove.

5 The Gain of Modeling Quantifiers

At this point, someone could ask:

- What is the gain in modeling quantifiers? and
- What has been accomplished by having turned qualitative verbal information into quantitative?

As it has been demonstrated throughout the paper, existing geospatial ontologies contain in their definitions terms that prevent the explicit assignment of semantic relations to categories. Our analysis has revealed the extensive use of qualitative linguistic terms as in the case of quantifiers, in such ontologies and the need to model this kind of information.

The gain is that we have introduced a consistent way to assign values to these terms so that when they are included in a definition, we know how much they deflect from bivalence. So far, the semantic properties of geospatial categories have been represented by strings of binary values as by Kavouras et al. (2003). At this point, we are able to use in these strings, values other than 1s and 0s. Therefore, we can represent all the information in definitions, without excluding any kind of linguistic data.

The representation of definitions after having modeled the included quantifiers is done therefore, as follows: take for instance example 1, Swamp. A low lying saturated area

---

4 A detailed discussion of this kind of quantifiers’ properties is out of the scope of the paper. The reader can
covered with water all or most of the year, ... (DIGEST) in paragraph 3.2. Table 3 gives an account of the category’s semantic relations and values containing also the quantitative information coming from the modeling of the quantifiers in the definition. This kind of representation is needed when encoding values of semantic relations as strings as in Kavouras et al. (2003). Thus, the representation of the semantic relations and their values of swamp definition in DIGEST, gives two strings: (1,1,1,1,1) and (1,1,1,1,0.833). This can further be used to compare definitions across repositories of geographic information such as standards and geographic ontologies. The assessment of values other than 0s or 1s for the semantic relations and properties of a geographic category, gives a better account of the semantics of the category.

The use of quantifiers in definitions, serves the purpose of differentiating between similar categories. Consequently, their modeling is indispensable when we want to reveal and code semantic heterogeneities across categories even subtle ones. A systematic approach to modeling qualitative linguistic terms, which are met in definitions, can aid the grasping of the semantics of geographic concepts, which is a key issue in geographic information science nowadays.

6 Further Work

This paper explores ways of representing the essence qualitative linguistic terms, which are found in the definitions of categories in geospatial ontologies. QLTs present a problem when trying to determine the semantic relations of geospatial categories. This is so, because, QLTs designate the degree of possessing a property. Therefore, their modeling is necessitated in ontology disambiguation.

nevertheless, for more details, refer to Partee and Borschev (2001)
Natural Language quantifiers can be modeled in terms of computational linguistics, and under the context of Generalized Quantifiers Theory, using the tree of numbers and then can be represented by strings of binary values. It becomes therefore vital to pursue further analysis of this kind of terms, for a lot of information about space and its phenomena involve such qualitative linguistic data. Supplementary research must be conducted to formalize the way these modifiers affect the semantic relations of geospatial concepts.

Future work includes a unified treatment of lexical items in definitions so that the establishment of semantic properties, and their values, for geospatial categories becomes explicit.

Acknowledgements

This work has been partially supported by the Heraclitus Research Programme 2.2.3.b of the Hellenic Ministry of National Education. The authors are also indebted to the anonymous reviewers for their very constructive comments.

References


<table>
<thead>
<tr>
<th>Quantifier Q</th>
<th>Definition from WordNet</th>
<th>Truth Set(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>All</strong> subscripts (_{XY}^{XY})</td>
<td>Indicates the whole number, or amount of, or everyone of a class, used as an adverb means completely, wholly, totally, entirely</td>
<td>({X \cap Y} = {E})</td>
</tr>
<tr>
<td></td>
<td>Antonym of no</td>
<td></td>
</tr>
<tr>
<td><strong>No</strong> subscripts (_{XY}^{XY})</td>
<td>Indicates a complete, or almost complete lack, or zero quantity of</td>
<td>({X \cap Y} = \emptyset)</td>
</tr>
<tr>
<td></td>
<td>Antonym of all</td>
<td></td>
</tr>
<tr>
<td><strong>Some</strong> subscripts (_{XY}^{XY})</td>
<td>Indicates an unspecified number (relatively many) or quantity of (relatively much)</td>
<td>(\emptyset &lt;</td>
</tr>
<tr>
<td><strong>Most</strong> subscripts (_{XY}^{XY})</td>
<td>Indicates the greatest in number (count nouns), in amount, extent, degree (mass nouns)</td>
<td>(</td>
</tr>
<tr>
<td></td>
<td>Superlative of many</td>
<td></td>
</tr>
<tr>
<td>Between (n) and (m)% of the XY</td>
<td>({(n/100)E \leq</td>
<td>X \cap Y</td>
</tr>
<tr>
<td>More than (n)%</td>
<td>(</td>
<td>X \cap Y</td>
</tr>
<tr>
<td>Less than or equal to (n)%</td>
<td>(</td>
<td>X \cap Y</td>
</tr>
</tbody>
</table>

**Table 1.** Examples of quantifiers and their truth sets for universe \(E\)
<table>
<thead>
<tr>
<th>Q</th>
<th>$S_{all,Q}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>No</td>
<td>0.000</td>
</tr>
<tr>
<td>Some</td>
<td>0.500</td>
</tr>
<tr>
<td>Between 10-60%</td>
<td>0.350</td>
</tr>
<tr>
<td>Between 10-30%</td>
<td>0.200</td>
</tr>
<tr>
<td>Most</td>
<td>0.833</td>
</tr>
<tr>
<td>More than 60%</td>
<td>0.850</td>
</tr>
<tr>
<td>Less than 10%</td>
<td>0.045</td>
</tr>
</tbody>
</table>

*Table 2.* The values of similarity $S$ between *all* and other quantifiers met in geographic concept definitions.
<table>
<thead>
<tr>
<th>Category</th>
<th>Semantic Relations/Properties</th>
<th>Hypernym</th>
<th>Location</th>
<th>State</th>
<th>Material/cover</th>
<th>Time (with reference to material/cover)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Swamp</td>
<td></td>
<td>Area</td>
<td>Low-lying</td>
<td>Saturated</td>
<td>Water</td>
<td>All year or most of the year</td>
</tr>
<tr>
<td></td>
<td>Encoded Values (Between 0 and 1)</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1 ∨ 0.833</td>
</tr>
</tbody>
</table>

**Table 3.** Representing a category’s semantic relations including modeled qualitative terms
Fig.1 The format of a tree of numbers
Fig. 2 How an element of a quantifier is indicated
Fig. 3 Trees for: all, most, some and their repeated patterns (the lines indicate that in the specific direction the patterns of minuses/pluses are repeated no matter how far down the rows we go)
Fig. 4 Relations among quantifiers as presented by Horn’s square of opposition, in Klein (1998)
Fig. 1 The format of a tree of numbers

Fig. 2 How an element of a quantifier is indicated

Fig. 3 Trees for *all, most, some* and their repeated patterns (the lines indicate that in the specific direction the patterns of minuses/ pluses are repeated no matter how far down the rows we go)

Fig. 4 Relations among quantifiers as presented by Horn’s square of opposition, in Klein (1998)