Geophysical Research Abstracts,
Vol. 10, EGU2008-A-07336, 2008
SRef-ID: 1607-7962/gra/EGU2008-A-07336
EGU General Assembly 2008

Representing meaning of geographic concepts with semantic elements

M. Kokla

School of Rural and Surveying Engineering, National Technical University of Athens, Greece (mkokla@survey.ntua.gr / Fax: 30+2107722734 / Phone: 30+2107722637)

Semantic interoperability can be achieved only if the meaning of geographic concepts may be properly exchanged, analyzed, and interpreted among the interoperating parties. This necessitates the formalization of the meaning of geographic concepts, in order to avoid semantic loss and support integration of the underlying ontologies. The notion of meaning has been intensively studied by different disciplines (e.g., philosophy, linguistics, psychology, semiotics, and communication theory) along two fundamental directions: (a) the nature and definition of meaning and (b) the methods of semantic description and representation.

The present work proposes the representation and formalization of the meaning of geographic concepts using self-contained, comparable and reusable meaning elements. These *meaning* or *semantic elements* are extracted using a specific methodology from definitions of geographic concepts derived from information sources such as ontologies, categorizations, dictionaries, and lexical databases. Definitions consist in the substitution of the meaning of a complex word for other simpler words. Definitions exhibit a specific lexical and syntactic structure, since they are formed based on syntactic and lexical patterns. The proposed methodology takes advantage of these patterns in order to identify and extract the simple, self-contained semantic elements found in definitions.

The approach is based on the important distinction drawn in semantics between the two modes of meaning: *extension* and *intension*. The *extension* refers to the set of things which are described by a concept. For example, the extension of the concept

"river" refers to all the rivers that exist or existed or will exist in the future, the river Thames, the Mississippi river, etc. The *intension* refers to the set of fundamental, distinctive properties, features, or characteristics, which are possessed by the things described by a concept. These properties provide the concept its individual character, i.e., its identity. In that way, the intension of the concept river is "a large natural stream of water". Definitions of geographic concepts are mainly intensional, i.e., they indicate the essential properties of each concept, which provide its identity and which differentiate the concept from other similar ones. Some definitions may also include an extensional part, by citing prominent instances of the concept.

The semantic elements found in definitions are distinguished into two types: *semantic properties* and *semantic relations*. Semantic properties constitute the inherent features of the concept, which are independent of other concepts, for example, "purpose", "size", "shape", "time", "cover", etc. Semantic relations on the other hand are formed by the relations of the concept with other concepts. Semantic relations include the subsumption ("is-a") and partonomy ("part-of") relations, as well as topological relations (e.g., "surroundness", "adjacency", "overlap", etc.).

The proposed approach evolves in two directions. The first one deals with the identification of semantic elements found in definitions of geographic concepts. Therefore, different geographic information sources were examined to determine the basic semantic elements used in the definitions of geographic concepts. These information sources include data standards (e.g., the Spatial Data Transfer Standard (SDTS) and the Digital Geographic Information Exchange Standard (DIGEST)), lexical databases (e.g., WordNet), ontologies (e.g., CYC) and nomenclatures (e.g., CORINE Land Cover). Geographic concepts are described by general semantic elements, as well as spatially-oriented ones, such as location, topology, direction, proximity, etc.

The second direction involves the extraction of these semantic elements from definitions and their formalization in such a way as to be convenient and practicable for use in a task such as ontology comparison and integration. For that purpose, each definition is decomposed into a set of semantic elements and their corresponding values. For example, the following definition:

"lake: body of water surrounded by land"

includes a subsumption relation ("is-a") which takes the value "body of water" and a "surroundness" relation which takes the value "water". Therefore, the task of comparing concepts' semantics based on natural language definitions is reduced to the task of comparing the sets of semantic elements and values. Concepts from different ontologies which are defined using the same semantic elements and values are considered to be equivalent; whereas concepts from different ontologies which are defined

using either different semantic elements or the same semantic elements but different values, are considered to be different. The proposed methodology is effective when the meaning of geographic concepts is described by definitions, which are based on specific syntactic and lexical rules. A possible extension of the methodology to analyze and extract semantic elements from free text, though not straightforward, may be worthwhile to examine.