Comparing Representations of Geographic Knowledge Expressed as Conceptual Graphs[*]

Athanasios Karalopoulos, Margarita Kokla, and Marinos Kavouras

National Technical University of Athens,
15780, Zografos Campus, Athens, Greece
akaralop@mail.ntua.gr, {mkokla, mkov}@survey.ntua.gr

Abstract. Conceptual Graphs are a very powerful knowledge and meaning representation formalism grounded on deep philosophical, linguistic and object oriented principles [1], [2]. Concerning geographic knowledge representation and matching, the study and analysis of geographic concept definitions plays an important role in deriving systematic knowledge about concepts and comparing geographic categories in order to identify similarities and heterogeneities [4]. Based on the proposed algorithm for the representation of geographic knowledge using conceptual graphs, we also present a method that takes into consideration the special structure of conceptual graphs and produces an output that shows how much similar two geographic concepts are and hence which concept is semantically closer to another. For producing the conceptual graph representation of any geographic concept definition we follow two steps, tagging and parsing, while for measuring the similarity between two geographic ontologies we apply proper modifications to the Dice coefficient that is mainly used for comparing binary structures.

1   Introduction

Conceptual Graphs are a powerful knowledge and meaning representation formalism grounded on deep philosophical, linguistic and object-oriented principles [1], [2]. They provide extensible means of capturing and representing real-world knowledge. Fundamental studies about Conceptual Graphs and some of their applications in the field of Knowledge Representation are found among others in [3].

Concerning geographic knowledge representation, the study and analysis of geographic concept definitions plays an important role in the attempt to derive systematic knowledge about concepts and compare geographic categories in order to identify semantic similarities and heterogeneities [4]. Therefore, the exploitation of effective methods for the representation of geographic definitions forms the basis of the research for analyzing geographic concepts in order to structure their meaning and extract semantic information.

The purpose of the present research is to develop an algorithm for the representation of geographic knowledge using conceptual graphs and then, based on the

---

[*] This work extends the use of conceptual graphs in geographic knowledge representation as first introduced in [18]. It also addresses the issue of comparison.

proposed methodology and the special features and structures of conceptual graphs, to describe a well-defined process for comparing two geographic concept definitions in order to quantitatively measure their semantic similarity. The comparison process takes into consideration the structure of the corresponding conceptual graphs and produces an output that shows how much similar two geographic concepts are and hence which concept is semantically closer to another.

By introducing an algorithm that takes a geographic concept definition as input and produces the corresponding conceptual graph representation, we achieve to break many limitations and obstacles in the extraction of semantic information from definitions of geographic concepts. Furthermore, we provide alternative deterministic means of facilitating semantic interoperability since the similarity between geographic ontologies depends on specific results of the introduced method for comparing geographic ontologies.

2   Related Work

During the last years, research has been done in order to represent and extract information about geographic concepts. Approaches on geographic knowledge representation include methodologies that are based on analyzing geographic concept definitions and finding effective representations. These can be found among others in [5] and [6].

Conceptual Graphs are a diagrammatic and expressive way of knowledge representation that was firstly introduced for the representation of contents of natural language texts. According to the conceptual graph theory [7], a conceptual graph is a network of concept nodes and relation nodes. The concept nodes represent entities, attributes, or events (actions) while the relation nodes identify the kind of relationship between two concept nodes.

Conceptual Graphs are formally defined by an abstract syntax that is independent of any notation, but the formalism can be represented in either graphical or character-based notations. In the graphical notation, concepts are represented by rectangles, conceptual relations by circles and the arcs that link the relations to the concepts are represented by arrows. The linear form is more compact than the graphical and it uses square brackets instead of boxes and parentheses instead of circles.

Research into establishing comparison methods for similarity measurement between two conceptual graphs is included in [8] and [9]. The main goal of the proposed approaches is to determine whether a query graph is completely contained in any given conceptual graph.

On the other hand, in many text-oriented applications, comparison methods for text representations are proposed and implemented. For instance, in [10] different types of coefficients are introduced for similarity measurement of various data structures and representations. Among them, the Jaccard coefficient, the Cosine coefficient and the Dice coefficient are mainly used for comparing binary structures not only because their results are widely accepted, but also because they are very simple.

Our algorithm for similarity measurement is based on the Dice coefficient, which is calculated using the following formula: