Chapter 90
Bidirectional Spreading Activation Method for Finding Human Diseases Relatedness Using Well-Formed Disease Ontology

Said Fathalla
University of Bonn, Germany & Alexandria University, Egypt

Yaman M. Khalid Kannot
Arab Academy for Science, Egypt

ABSTRACT

The successful application of semantic web in medical informatics and the fast expanding of biomedical knowledge have prompted to the requirement for a standardized representation of knowledge and an efficient algorithm for querying this extensive information. Spreading activation algorithm is suitable to work on incomplete and large datasets. This article presents a method called SAOO (Spreading Activation over Ontology) which identifies the relatedness between two human diseases by applying spreading activation algorithm based on bidirectional search technique over large disease ontology. The proposed methodology is divided into two phases: Semantic matching and Disease relatedness detection. In Semantic Matching, semantically identify diseases in user's query in the ontology. In the Disease Relatedness Detection, URIs of the diseases are passed to the relatedness detector which returns the set of diseases that may connect them. The proposed method improves the non-semantic medical systems by considering semantic domain knowledge to infer diseases relatedness.

DOI: 10.4018/978-1-7998-1204-3.ch090

1. INTRODUCTION

Usually, computers were used to perform tasks with none sort of intelligence or semantics thus there is a necessity to make computers as intelligent as human (Antoniou & Van Harmelen, 2004). The defect of non-Semantic applications is painted by the statement "lack of semantics" particularly when the talk is concerning information retrieval. The semantic web's goal is to extend automation in processing web-based data and to enhance the interoperability of web-based information systems (Payne, Terry & Lassila, 2004). The semantic web gives the meaning of knowledge making software systems can perform intelligent tasks instead of users (Berners-Lee, Hendler & Lassila, 2001) using a semantic knowledgebase called Ontology. Ontologies is a repository in which information are organized and used in semantic-based applications (Shadbolt, Berners-Lee & Hall, 2006). Below a group of definitions that may be used as a basis for mathematical formulation of the term ontology and its components (Fonseca, Davis & Câmara, 2003):

Definition 1: A term is a triple $\tau = [\eta, \delta, A]$, $\tau \in T$, where η is a string of characters containing the name of the term, δ is a string of characters containing its definition and A is a set of attribute domains A1, A2, ..., An, each associated to a value set Vi.

Definition 2: A relation $\varphi: T \rightarrow T$, $\varphi \in \Phi$: is a function from T to T such that for every term $\tau 1 \in T$, there is a term $\tau 1 = \varphi(\tau 1)$, $\tau 2 \in T$.

Definition 3: A semantic relation σ between two concepts is one of the relations among the set of semantic relations $\Sigma = \{$Hypernym, Hyponymy (is-a), Meronomy (part-of), Synonymy$\}$. In linguistics, a hyponym is a word whose semantics is enclosed inside that of another word which called its hyperonym or its hypernym. In other words, a hyponym has the same a type-of relationship with its hypernym. For instance, dog, cat and cow are all hyponyms of Mammals (their hypernym); that, in turn, could be a hyponym of animal (Fromkin, Rodman & Hyams, 2013).

Ontologies are often used in many systems as a semantic knowledgebase (Oberle, 2014). In the field of health informatics, ontology is used for the representing and organizing medical vocabularies. Collins and Loftus (1975) discussed in their research that the Spreading Activation (SA) runs on semantic networks (Anderson et al., 1983) and is used for information retrieval process (Burnset et al., 1981). It runs on a graph structure that comprises of a set of nodes connected by edges. The concepts are nodes which have an activation value and the relations between them are represented by edges between these nodes. An activation value is assigned to each node in the graph and then the algorithm spreads to the nodes with the higher activation value. The algorithm runs in a set of iterations and terminates when a stopping condition is reached.

The main objective of this work is to identify whether two human diseases are related to each other or not. If so, identify the set of diseases that connecting them. Finding a relatedness between diseases is important because the doctor can treat patient not only based on the symptoms they suffer from but treat the real cause of this disease which may be related to another disease that causes these symptoms therefore, the doctor can treat the cause disease not the symptoms.

This paper is organized as follows. Section 1 gives a brief overview of semantic web and ontologies. Section 2 reviews the literature. Section 3 describes the ontology for SAOO. Section 4 introduces the methodology of the proposed work. Section 5 illustrates the methodology by showing a running example. Section 6 describes the prototype used to implement the methodology. Finally, Section 7 presents conclusions and future work.