year / month / day

2013 / 6 / 28

Department	Electronic and Information Engineering	ID	D099310
Name	Suriati Akmal		

Supervisor	Rafael Batres

Abstract

Title	A Semantic Similarity Method for Products and Processes

(800 words)

Similarity measures are concerned with quantifying of the likeness of the things that are compared. Similarity measures have been practically applied in a wide variety of fields ranging from data mining, case-based reasoning system, image interpretation and pattern recognition. Several researchers have proposed similarity measures that evaluate the likeness between values of numeric properties. However, in many applications some attributes are non-numeric. One solution is to use syntactic similarity measures that calculate the similarity between two words. However, syntactic approaches are limited as they fail to produce good matches when confronted with the meaning associated to the words they compare.

To overcome the above drawbacks semantic similarity measures are been investigated. A semantic similarity measure is a function that quantifies the degree of likeness between two things based on the meaning associated to each thing being compared. This research contributes to the field of semantic similarity measures for products and processes. A novel approach has been proposed in this research, based on Formal Concept Analysis (FCA) and a set of criteria for the characterization of products and processes called Formal Attribute Specification Template (FAST).

This research focuses on countable objects that are represented in terms of their physical aspects and processes in which they are involved. Processes can be intentional or unintentional. In an intentional process, a particular objective is accomplished. Unintentional processes include natural phenomena and undesired processes such as harmful explosions or fires.

The proposed approach is composed of semantic similarity measures that compare classes in a taxonomy obtained with FCA and a template for the specification of formal attributes (FAST).

The semantic similarity measures of the proposed approach compare classes of products or processes. The comparison is based on the assumption that the more common attributes that are shared by two classes the more similar they are. Therefore, a class is 100% similar to another class if both classes have exactly the same attributes. In particular, the attributes are the formal attributes from the FCA. For this purpose, several similarity equations are investigated in this research by using formal attributes as the sets they compare.

Class taxonomies are defined by means of the subclass relation. A class is a subclass of another class if every member of the subclass is also a member of the super class. Formal Concept Analysis (FCA), which is a method based on applied lattice and order theory, is selected as the taxonomy generator.

FAST helps to describe the formal attributes common to all members of a given class that distinguish them from members of another class. The product formal attributes are expressed in terms of its mereological and topological structure and its involvement with one or more processes. The process formal attributes are expressed in terms of: (1) objects that are always changed by the process (a.k.a inputs); (2) objects that are always produced by the process (a.k.a outputs); (3) participating physical objects (including locations, agents, and performer) other than inputs and outputs; (4) sub-activities that compose the process (a.k.a sub-activities).

The proposed approach was evaluated against edge-counting and information-based similarity measures. In order to quantify the efficacy of each similarity measure, the degree of correlation with human judgment was used. The results of the evaluation show that the proposed approach performed better than existing similarity measures.

The proposed approach is illustrated with two case studies. The first case study demonstrates the use of FAST for the construction of an ontology for machining processes. The resulting machining processes ontology was evaluated and compared against a third-party ontology. The degree of correlation with Internet-search engine using the value of the Normalized Google distance evaluated the accuracy of each ontology. The results of evaluation show that the ontology obtained with FAST is slightly better than the existing ontology. It was also found that FAST can provide the design rationale of the ontology.

The second case study focused on the application of the proposed semantic similarities for selecting the service strategy for Product-Service systems (PSS) at the early stage of design. It is often the case that the PSS designer is faced with limited amount of knowledge at the early stage of design. One solution is to use the case-based reasoning (CBR) system to facilitate the service strategy selection in which PSS design problems are solved by using or adapting previously obtained design solutions. Existing CBR-systems use numerical similarity measures to search the relevant solution to the problem to be solved. In this case study, a semantic CBR-system was developed by incorporating product-class-comparison based on the proposed semantic similarities. The results of evaluation show that the proposed approach proved useful when some details of information are not available.