Computers in Industry 65 (2014) 91–107

Contents lists available at ScienceDirect

Computers in Industry

journal homepage: www.elsevier.com/locate/compind

Ontology-based similarity for product information retrieval

Suriati Akmal [a,1], Li-Hsing Shih [b], Rafael Batres [a,*]

[a] Department of Mechanical Engineering, Toyohashi University of Technology, Hibarigaoka 1-1, Tempaku-cho, Toyohashi 441-8580 Japan
[b] Department of Resources Engineering, National Cheng Kung University, No. 1, University Road, Tainan 701, Taiwan, ROC

ARTICLE INFO

Article history:
Received 4 September 2012
Received in revised form 10 June 2013
Accepted 31 July 2013
Available online 21 September 2013

Keywords:
Semantic similarity
Ontology
Product information retrieval
Formal concept analysis

ABSTRACT

Product development of today is becoming increasingly knowledge intensive. Specifically, design teams face considerable challenges in making effective use of increasing amounts of information. In order to support product information retrieval and reuse, one approach is to use case-based reasoning (CBR) in which problems are solved "by using or adapting solutions to old problems." In CBR, a case includes both a representation of the problem and a solution to that problem. Case-based reasoning uses similarity measures to identify cases which are more relevant to the problem to be solved. However, most non-numeric similarity measures are based on syntactic grounds, which often fail to produce good matches when confronted with the meaning associated to the words they compare. To overcome this limitation, ontologies can be used to produce similarity measures that are based on semantics. This paper presents an ontology-based approach that can determine the similarity between two classes using feature-based similarity measures that replace features with attributes. The proposed approach is evaluated against other existing similarities. Finally, the effectiveness of the proposed approach is illustrated with a case study on product–service–system design problems.

1. Introduction

Due to the complexity of products and drastic technological changes, product development is becoming increasingly knowledge intensive. Design is also multi-disciplinary in nature requiring a variety of product life-cycle knowledge [1]. Specifically, design teams face a considerable challenge in making effective use of increasing amounts of information that often accumulate and remain in individual information systems. Also, it is often the case that product designers can reuse past designs rather than designing from scratch [2].

Information retrieval consists of translating and matching a query against a set of information objects. Translation of the query is necessary for converting the user requirements into the language provided by the information retrieval system. The information retrieval system responds to the query using a given algorithm and a similarity measure. Particularly, information retrieval plays an important role in areas such as product family design [3], product embodiment, and detailed design [4]. Shah et al. [5] present a combination framework that consists of software engineering, data engineering and knowledge engineering and design theory.

In order to support product information retrieval and reuse, some authors suggest the use of case-based reasoning (CBR) in which design problems are solved by using or adapting previous design solutions [4,6].

A CBR system is composed of domain knowledge, a case base, and a search mechanism based on a similarity measure. Domain knowledge refers to knowledge about the features of the different objects or entities that a case is about. A case base contains a set of cases, each of which describes a problem and a solution to the problem. The problem is typically defined in terms of specific features of objects. Finally, a similarity measure quantifies the differences that exist between objects [7]. CBR uses similarity measures to identify cases which are more relevant to the problem to be solved.

Most similarity measures evaluate differences between values of numeric properties such as in the numerical difference between two given diameter values. However, many applications also require non-numeric similarities. For example, case-based reasoning systems for the conceptual design of products must be developed to work with a limited knowledge about the product.

Nearly all of non-numeric similarity measures are based on syntactic grounds. For example, the Levenshtein distance [8,9] can be used to calculate the similarity between two words, in terms of the minimum number of operations that are needed to transform one of the words into the other. However, from the point of view of the meaning of the words that are compared, existing syntactic similarity-measures often result in incorrect matches.

* Corresponding author. Tel.: +81 532 44 6716; fax: +81 532 44 6716.
E-mail address: rafaelbatres@hotmail.com (R. Batres).
[1] On leave from the Universiti Teknikal Malaysia Melaka, 76109 Durian Tunggal, Melaka, Malaysia.

0166-3615/$ – see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.compind.2013.07.011

Semantic similarity measures can be used in order to overcome the limitations of syntactic approaches. A semantic similarity is a function that assigns a numeric value to the similarity between two classes of objects based on the meaning associated to each of the objects [10]. For a review of semantic similarity metrics, the reader is referred to the paper of Cross and Hu [11].

Recently, the use of ontologies for evaluating similarity has been reported in the literature [12,13]. Ontologies are formal models that use mathematical logic to disambiguate and define classes of things [14]. Specifically, ontologies describe a shared and common understanding of a domain in terms of classes, possible relations between things, and axioms that constrain the meaning of classes and relations [15]. A class represents a set of things that share the same attributes. A relation is used to represent a relationship among two or more things. Examples of relations are less than, connected to, and part of. 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. Axioms are typically represented as logic constructions that formally define a given class or relation.

Combined with automated reasoning applications, ontologies can be used for several purposes such as knowledge extraction and information retrieval. Unfortunately, ontologies are typically created in an ad-hoc manner, which may influence the accuracy of the similarity calculations.

Formal concept analysis (FCA) is a data processing method that can be used to design ontologies [16,17]. FCA is based on a set of objects and a set of attributes. In this paper, we use FCA along with a theoretical framework for developing product and process ontologies.

Most semantic similarities are defined in terms of the number of edges between the classes that they compare (edge-counting similarity measures). Other semantic similarities are defined in terms of features but use synsets for the comparison between words rather than classes.

The underlying thesis in this paper is that if a class represents a set of things that share the same attributes (such as a class in an ontology), we can state that a class is equivalent to another class if both classes have exactly the same attributes. This implies that the more common attributes that are shared by two classes the more similar they are. In this paper, we show how an ontology-based approach can determine the similarity between two classes using feature-based similarity measures that replace features with attributes.

The paper is organized as follows. Section 2 describes the theoretical framework for product representation used in this paper. Section 3 provides details on the ontology development. Section 4 describes the proposed ontology-based similarity measures. Sections 5 and 6 describe the evaluation of the semantic measures proposed in this paper. In Section 7, the effectiveness of the proposed approach is illustrated with a case study on product-service-system design problems. Section 8 discusses some related work and Section 9 presents conclusions and suggestions.

2. Theoretical framework for product representation

Theoretical frameworks for product representation refer to the world view with which product information models or ontologies can be developed in order to represent a product. In this paper, the theoretical framework for representing a product is based on the ISO 15926 standard which specifies an upper ontology for long-term data integration, access and exchange [18]. It was developed in ISO TC184/SC4-Industrial Data by the EPISTLE consortium (1993–2003) and designed to support the evolution of data through time. The upper ontology was developed as a conceptual data model for the representation of technical information of process plants including oil and gas production facilities but it was designed to be generic enough for any engineering domain [19].

In this theoretical framework, the device is represented in terms of its physical aspects as well as in terms of its relation to some process (*activity* in ISO 15926). These aspects are illustrated in the models of Figs. 1 and 2.

A device is represented as a physical object that is defined in terms of a distribution of matter, energy, or both. The device is also described in terms of its parts. This is possible through a mereological relation that refers to the relationship that a part

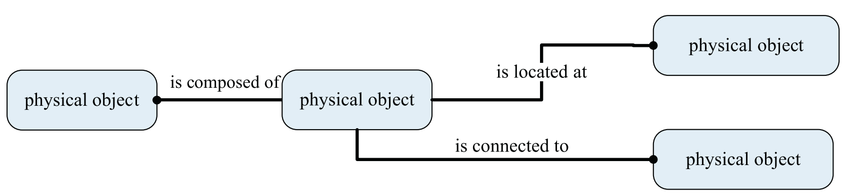

Fig. 1. Composition of device.

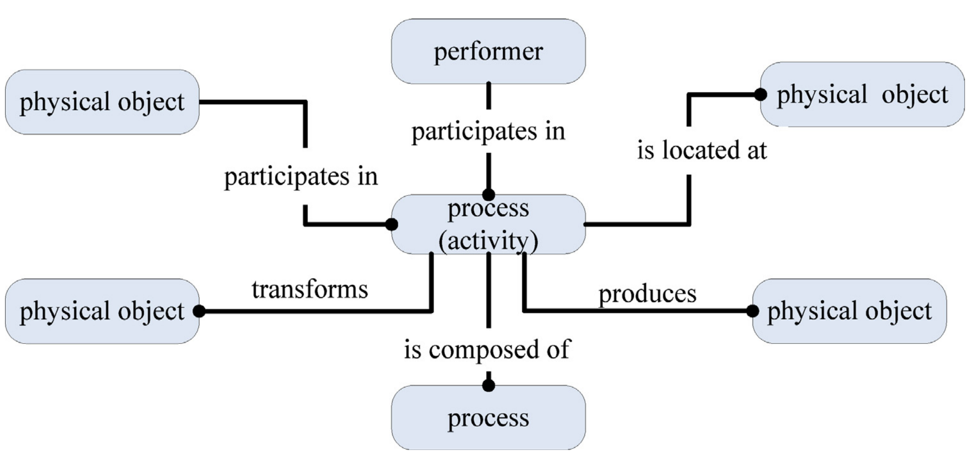

Fig. 2. Relations between device and process.

has in regards to the whole of an object. Mereological relations are reflexive, antisymmetric, and transitive.

Physical objects exist in reference to a specific place. The location relation (*relative location* in ISO 15926) is a kind of mereological relation that is used to locate objects in a particular place.

A stream is another kind of physical object that is applied to material or energy moving along a path, where the path is the basis of identity and may be constrained. For example, the material moving along a pipe is an instance of stream.

The function of a product can be defined as an intended process associated to the device. For example, the function associated to a sofa is represented as the process of seating in which the sofa is involved along with a person that sits on it. Similarly, the function of an electric fan is to generate cool air. In this case, the description of the device includes information about the cooling process. The cooling process is in turn composed of other processes such as conversion of electricity into rotary movement, convection, diffusion and heat transfer. Therefore information about the process or processes associated to the device is an indispensable element to complete the description of the product.

Different objects can participate in a process. Participating physical objects include those objects that are transformed by the process, those objects that are produced by the process, those objects that are not affected by the process (the device itself, other tools, or instruments), as well as agents (such as a person or a control system) that participate or execute the process.

As in the description of the device, a process is also described in terms of its relative location and its mereology.

3. Ontology construction

There are a number of methodologies to develop ontologies including Uschold and King's method [25], Grüninger and Fox's method [26], Noy and McGuiness's method [27], the METHONTOLOGY framework [28], the Cyc methodology, KACTUS, SENSUS, and the On-To-Knowledge Methodology [29].

According to Stevens et al. [30], the general stages for ontology development are: identification of purpose and scope, knowledge acquisition, conceptualization, integration, encoding, documentation, and evaluation. We follow these general steps, but we use FCA (see Appendix A) and the theoretical framework described in Section 2 to guide the knowledge acquisition and conceptualization stages.

Candidate classes that may or may not appear in the final ontology are identified, and the object column of a FCA context table is populated with these classes.

Subsequently, information sources such as scientific papers, technical reports, and Internet resources are consulted to define each class in natural language. When several definitions are found preference is given to those that explicitly describe participating objects, objects transformed by the process (inputs), objects produced by the process (outputs) and/or subactivities. When contradictions among several definitions of a given class occur experts can be consulted to disambiguate.

Formal attributes are identified from the natural-language definitions using the following guideline.

A device (a physical object) is characterized by

1. The process in which the device participates.
2. The composition of the device (its parts).

In turn, a given process or the process in which the device participates is characterized by the following:

- The object that is transformed by the process (the input of the process).
- The object that is produced by the process (the output of the process).
- The tool that is present in any instance of the process.
- The composition of the process.

After adding the formal attributes, the context table is completed and a lattice is generated. Lattices in this paper were generated by means of the Grail algorithm [24], which is implemented in the software Concept Explorer. Finally, the lattice is used to create the ontology. The naming of each class is done based on object or attributes labels from the nodes in the lattice.

Integration is carried out by means of aligning the resulting ontology with an upper ontology that defines domain-independent classes such as physical objects, activities, mereological and topological relations. For example, a class refrigerator is defined as a subclass of *physical_object*, which is a class of the ISO 15926 upper ontology.

4. Ontology-based semantic similarities

In a given ontology, a class is equivalent to another class if both classes have exactly the same attributes[2]. Therefore, the more common attributes that are shared by two classes the more similar they are.

The proposed approach consists of combining feature-based similarities with an ontology obtained with the procedure described in Section 3 and using formal attributes from the FCA as features. The feature-based similarities investigated in this paper are: the Tversky index [21], the Dice's coefficient [22], the Jaccard's coefficient [31], the Overlap coefficient [23], the all-confidence similarity [23], and the Cosine similarity.

For example, Tversky index becomes

$$sim_{Tversky}(C_1, C_2) = \frac{|A_1 \cap A_2|}{|A_1 \cap A_2| + \alpha|A_1 \backslash A_2| + \beta|A_2 \backslash A_1|} \quad (1)$$

where α and β are parameters calculated according to Rodriguez and Egenhoffer [32], A_1 and A_2 are the sets of attributes of classes C_1 and C_2, $|A_1 \cap A_2|$ is the total number of formal attributes shared by C_1 and C_2, $|A_1|$ and $|A_2|$ represent the number of formal attributes of C_1 and C_2.

We also use the similarity equations given by van der Weken et al. [33] but using formal attribute sets instead of fuzzy sets:

$$sim_{vanDerWeken1}(C_1, C_2) = \frac{|A'_1 \cap A'_2|}{\min(|A'_1|, |A'_2|)} \quad (2)$$

$$sim_{vanDerWeken2}(C_1, C_2) = \frac{|A'_1 \cap A'_2|}{\max(|A'_1|, |A'_2|)} \quad (3)$$

where A_1 and A_2 are the sets of attributes of classes C_1 and C_2; and A'_1 and A'_2 are the complements of sets A_1 and A_2.

No restriction exists for one of the classes to be a subclass or a superclass of the other. In other words, the classes to be compared can happen anywhere in the ontology.

In addition, we also investigate a composite similarity obtained by combining semantic similarities:

$$sim_{Composite}(C_1, C_2) = w_1 sim_1 + w_2 sim_2 \quad (4)$$

where w_1 and w_2 are weights and sim_1 and sim_2 represent two different semantic similarity measures.

In order to compare against edge-counting similarity measures, calculations are also carried out using both Wu-Palmer's [20] and Lin's [10] similarity measures.

[2] The attributes of a class also include those attributes inherited from its parent classes.

5. Experimental evaluation of the proposed approach using human judgment

This experiment focuses on the domain of electric home appliances for the evaluation of different semantic similarities. The evaluation is carried out by measuring the degree of correlation between the calculated similarity scores and scores obtained by human judgments. For this purpose, a questionnaire was administered to 30 respondents. The questionnaire asked each respondent to rank the likeness between 'electric kettle' and each of 17 home electric appliances. Respondents then rated the similarity of the pairs on a 1–17 scale, with lower numbers indicating higher similarity.

The comparison was carried out by calculating the correlation coefficient and the sum of squared errors.

The level of inconsistency of each questionnaire was calculated with the following formula.

$$d_i = \sum_j \left| q_{ij} - \mu_{ij} \right| \tag{5}$$

where q_{ij} is the value of the score that participant i submitted for pair j and μ_{ij} is the mean of the scores of all the users except that of user i for pair j. Using this formula, questionnaires with values of d_i above two standard deviations from the mean $\bar{d}_i$ were excluded from the analysis.

The average standard deviations of the scores across respondents were also evaluated to identify inconsistencies. Since one of the questionnaires had a standard deviation lower than average, it was not taken into account. With this last change, the sample size was reduced from 30 to 27.

Finally, individual pair scores with one standard deviation below or above the pair mean $\bar{q}_j$ were eliminated, which accounted for 4% of the total data.

Subsequently, the average scores were normalized using the following transformation:

$$s_j = \frac{\bar{q}_j - q^{\min}}{q^{\max} - q^{\min}} \tag{6}$$

where s_j represents the similarity of pair j, $q^{\max} = 17$ and $q^{\min} = 1$. Values of s_j are shown in the first column of Table 1.

5.1. Ontology development

In this section, we describe the development of an electric home appliance ontology, which is based on the method described in Section 3. The list of potential classes was extracted from product categories in Amazon.com and the attribute information was obtained using the characterization explained in Section 3, using expert consultations and brainstorming. In the development of the ontology, we focused on the process or processes in which the given appliance participates or is involved. Therefore, formal attributes include a reference to the process or a description of the process in terms of the objects that are transformed by the process and the objects that are produced by the process. For example, the formal attribute identification of an electric kettle starts by the analyzing its main process associated to it, which is a process that produces hot water. Heating is a part of that process. In order to produce hot water, the electric kettle consumes electricity that is converted into thermal energy that is used to heat water. Therefore, the formal attributes of an electric kettle become 'heats', 'produces hot water', 'heats water', and 'consumes electricity'.

With formal-attribute information obtained this way, a context table was created (Fig. 3). Subsequently, Concept Explorer [24] was used to generate the concept lattice shown in Fig. 4. After analyzing and correcting the lattice, the final lattice and formal-attribute information were used to develop an ontology using the Protégé ontology editor [34]. Subsequently, the resulting ontology was saved in OWL format [35].

Strictly speaking, formal attribute information must be in the form of axioms as in the following example.

```
Class filtration:
  SubClassOf:
    heating_device
  SubClassOf:
    produces some hot_water
```

However, for simplicity in the similarity calculation, formal attributes were added as OWL properties. For example, the formal attribute for "produces hot water" is declared as follows:

```
Declaration(ObjectProperty(:produces_hot_water))
ObjectPropertyDomain(:produces_hot_water:wa-
ter_heater)
```

Table 1
Comparison between similarity measures.

Electric kettle with:	Human judgment rank	Wu and Palmer	Wu Palmer (with device class)	Lin	Dice	All confidence	Overlap	van der Weken 1	van der Weken 2	Jaccard	Cosine	Tversky
Electric dish washer	0.29	–	0.29	0.01	0.22	0.20	0.25	0.90	0.87	0.13	0.22	0.22
Washing machine	0.32	–	0.29	0.01	0.22	0.20	0.25	0.90	0.87	0.13	0.22	0.22
Electric clothes dryer	0.62	0.40	0.57	0.18	0.50	0.50	0.50	0.93	0.93	0.33	0.50	0.50
Hair dryer	0.76	0.40	0.57	0.18	0.50	0.50	0.50	0.93	0.93	0.33	0.50	0.50
Water heater	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Electric blanket	0.67	0.50	0.67	0.18	0.57	0.50	0.67	0.97	0.94	0.40	0.58	0.58
Toaster	0.73	0.33	0.50	0.18	0.40	0.33	0.50	0.93	0.87	0.25	0.41	0.41
Bread machine	0.52	0.33	0.50	0.18	0.40	0.33	0.50	0.93	0.87	0.25	0.41	0.41
Electric oven	0.80	0.33	0.50	0.18	0.33	0.25	0.50	0.92	0.80	0.20	0.35	0.35
Microwave oven	0.72	0.33	0.50	0.18	0.40	0.33	0.50	0.93	0.87	0.25	0.41	0.41
Vacuum cleaner	0.18	–	0.33	0.01	0.25	0.25	0.25	0.90	0.90	0.14	0.25	0.25
Television set	0.08	–	0.50	0.01	0.33	0.25	0.50	0.97	0.91	0.20	0.35	0.35
Room electric heater	0.81	0.50	0.67	0.20	0.57	0.50	0.67	0.97	0.94	0.40	0.58	0.58
Room air-conditioner	0.49	0.40	0.57	0.18	0.40	0.33	0.50	0.93	0.87	0.25	0.41	0.41
Conventional electric fan	0.25	–	0.33	0.01	0.25	0.25	0.25	0.90	0.90	0.14	0.25	0.25
Refrigerator	0.34	–	0.33	0.01	0.22	0.20	0.25	0.90	0.87	0.13	0.22	0.22
Blender	0.21	–	0.33	0.01	0.25	0.25	0.25	0.90	0.90	0.14	0.25	0.25
Correlation with human judgment	1.00		0.782	0.731	0.777	0.729	0.795	0.608	0.272	0.708	0.781	0.726
Sum of squared errors			0.474	2.713	0.688	0.941	0.484	3.921	3.532	1.576	0.662	0.930

	consumes electricity	generates infrared radiation	generates high energy frequency waves	heats	heats room	heats water	heats body	heats food	removes heat	removes heat from room	removes heat from food	removes water	removes water from hair	removes water from clothes	removes water from dishes	removes dirt	removes dirt from clothes	removes dirt from dishes	removes dirt from surfaces	produces cooked food	bakes	produces baked food	bakes bread	sucks up dirt	produces toasted bread	receives tv-signals	mixes food	chops food	delays bacteria growth	delays mold growth	produces air circulation	produces food	produces hot water
home electric appliances	x																																
room electric heater	x			x	x																												
hair dryer	x			x								x	x																				
electric blanket	x			x			x																										
washing machine	x											x		x		x	x																
electric clothes dryer	x			x								x		x																			
refrigerator	x								x		x																		x	x			
room air-conditioner	x			x	x				x	x																					x		
electric dish washer	x											x			x	x		x															
microwave oven	x		x	x				x												x												x	
toaster	x	x		x				x																	x							x	
electric kettle	x			x		x																											x
television set	x																									x							
conventional electric fan	x								x	x																					x		
blender	x																										x	x				x	
bread machine	x			x				x													x		x									x	
electric oven	x	x		x				x												x	x	x										x	
water heater	x			x		x																											x
vacuum cleaner	x															x			x					x									

Fig. 3. Context table for an ontology of home electric appliances.

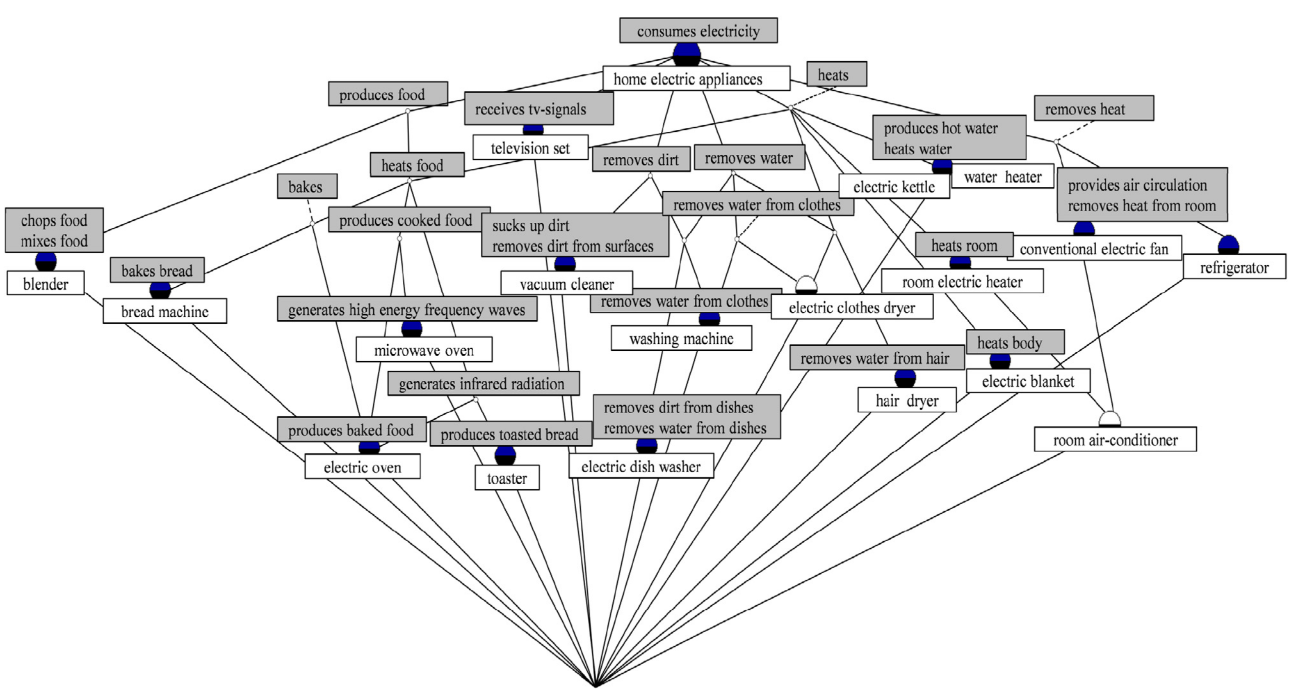

Fig. 4. Concept lattice obtained with the context table of Fig. 3.

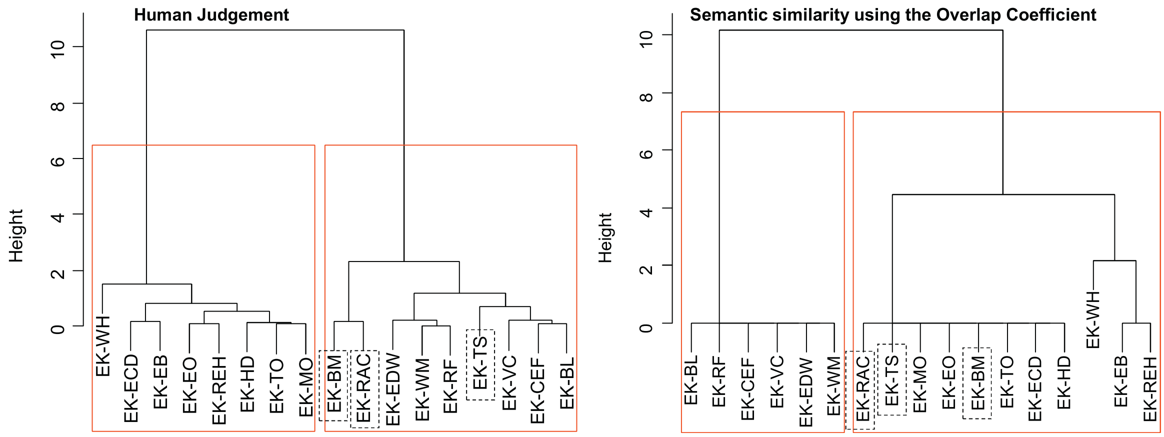

Fig. 5. Results of the cluster analysis.

This resulted in an OWL file with 33 classes, 39 properties, and 5 levels in the class hierarchy.

5.2. Similarity calculation

A program was developed in Java using the ontology library Jena [36]. The program reads the ontology and the names of the two classes to be compared. Firstly, it extracts the formal attribute information of each class in the ontology. Then, the program proceeds to calculate the cardinalities for each set of attributes, the minimum and maximum values, and the number of overlapped attributes. Attributes of a class include those inherited from all its parent classes. Similarity calculations are then carried out using the feature-based similarities as explained in Section 4. Then the Wu-Palmer's and Lin's similarities are calculated by edge counting, using the taxonomy structure of the ontology.

5.3. Experiment results

Table 1 summarizes the calculation results of the investigated similarities rating between 17 class comparisons.

We introduced 'device' as subclass of *physical_object* (defined in ISO 15926) and made 'home electric appliance' a subclass of 'device'. From Table 1, it can be seen that the best performing similarity is the Overlap coefficient ($sim_{Overlap}$) with $R = 0.795$ followed by the Wu-Palmer similarity with $R = 0.782$, the Cosine similarity (sim_{Cosine}) with $R = 0.781$, and Dice's coefficient with (sim_{Dice}) with $R = 0.777$.

After considering every possible combination of feature-based similarities for the composite similarity equation of Eq. (4), the best two combinations were:

$$sim_{Cosine+Jaccard}(C_1, C_2) = 1.887 sim_{cosine} - 0.887 sim_{Jaccard} \quad (7)$$

with a correlation of $R = 0.817$ and

$$sim_{Dice+Jaccard}(C_1, C_2) = 1.966 sim_{Dice} - 0.996 sim_{Jaccard} \quad (8)$$

with a correlation of $R = 0.816$.

The weights w_1 and w_2 were obtained by numeric optimization so as to minimize the residual sum of squares between the composite similarity and s_j of Eq. (6).

5.4. Analysis of the results

To eliminate biases in the analysis of the results, we removed those pairs that produced squared errors greater than two times the standard deviation. The pairs (electric kettle, television set) and (electric kettle, electric oven) produced the biggest squared error. After removing both pairs, the correlation value of the Overlap coefficient increased to $R = 0.947$. Again, sim_{Cosine} ($R = 0.922$) and sim_{Dice} ($R = 0.919$) were second and third in performance, respectively. For the combined similarities, $sim_{Cosine+Jaccard}$ increased to $R = 0.950$ and $sim_{Dice+Jaccard}$ increased to $R = 0.947$.

A hierarchical cluster analysis was conducted in order to compare relatively homogeneous groups of results. The cluster analysis was equally applied to both the human assessment results and the results obtained with the Overlap coefficient. Clustering was carried out using Ward's minimum variance algorithm.

A comparison of the clusters indicates that most of the object pairs that belong to one cluster with the Overlap coefficient also belong to a cluster in the results of human judgment. As shown in Fig. 5, only (electric kettle, television set), (electric kettle, air conditioner), and (electric kettle, bread machine) were grouped into another cluster. This is probably due to missing attributes in the FCA context table. Alternatively, another possible reason is that these three pairs were particularly difficult to judge during the answering of the questionnaire.

6. Experimental evaluation of the proposed approach using Web search

This experiment uses the Web to evaluate the different similarity measures. In order to increase the accuracy of the Web-based evaluation, a more technical domain was selected. For this reason, we chose machining processes as the technical domain. Moreover, the evaluation was carried out with the search engines provided by Google's Scholar and Elsevier's Scirus.

Despite the existence of a vast variety of machining processes, in order to obtain a compact ontology, the scope of this experiment was limited to mechanical material removing processes. In order to develop the ontology, several common textbooks [37–41] and an Internet source [42] were consulted. The potential classes are listed in the first column of Table 3.

During the construction of the ontology, in order to determine the formal attributes for the context table, all the material removing processes were characterized according to the process characterization explained in Section 3.

Fig. 6 shows the context table with the classes of material removing processes. The resulting concept lattice is presented in Fig. 7. The similarity measures were the same as in the experiment of Section 5. The above mentioned Java program was used in all the calculations.

	removes material	transformed object is rotated	transforms a hole object	produces a holed_object	produces a threaded hole	enlarges a portion of an existing hole to a larger diameter	enlarged portion is cylindrical	produces physical object in which the bottom part of the enlarged portion is flat and square	physical object in which the enlarged portion provides seat for a washer	physical object in which the enlarged portion provides a recess for a countersunk flat heat screw or countersunk rivet	produces a physical object in which the bottom part of the enlarged portion is cone-shaped	enlarges the end portion of the hole	removes less than 0.015 in of material	material is removed by means of a cutting tool	material is removed by means of rotating cutting tool	involves in chemical reaction	always uses a drill	uses a boring bar
drilling	x			x										x			x	
boring	x		x	x		x	x							x				x
reaming	x		x	x		x	x						x	x				
tapping	x		x	x	x	x	x							x				
counterboring	x		x	x		x	x	x				x		x				
spot facing	x		x	x		x		x	x			x		x				
coutersinking	x		x	x		x				x	x	x		x				
turning	x	x												x				
miiling	x													x	x			
chemical machining	x															x		

Fig. 6. The context table of material removal process.

This time, the resulting similarity scores were compared against the Web-based similarity denoted by Eq. (9).

$$v(t_i, t_j) = 1 - d(t_i, t_j) \tag{9}$$

where $v(t_i,\ t_j)$ is the Web-similarity of terms t_i, t_j, and $d(t_i,\ t_j)$ represents the distance function proposed by Cilibrasi and Vitanyi [43] also known as the normalized Google distance or NGD. The distance function of Cilibrasi and Vitanyi is described by Eq. (10).

$$d(t_i, t_j) = \frac{\max(\log f(t_i), \log f(t_j)) - \log f(t_i, t_j)}{\log M - \min(\log f(t_i), \log f(t_j))} \tag{10}$$

where $f(t_i)$, $f(t_j)$ and $f(t_i,\ t_j)$ give the number of hits for the terms t_i, t_j and (t_i, t_j), respectively. M corresponds to the number of indexed documents in a given Web search engine. The number of indexed documents of Scirus which is reported on its Web page is 4.6×10^8 as of May 21, 2012. A value of $M = 5.8 \times 10^8$ for Google Scholar was obtained from an earlier estimate [44] and by assuming a growth rate of 2.7% based on the world-wide average annual increase of academic papers [45].

In order to restrict the Web search to the domain of study, keywords in both search engines were formulated with the inclusion of the term "machining" and search was carried out using double quotes.

For example, for the similarity between counterboring and spot facing, search with Scholar for "machining" "counterboring" results in f(counterboring) = 1019 hits; search for "machining" "spot facing" produces f(spot facing) = 620 hits; and search for "machining" "counterboring" "spot facing" results in f(counterboring, spotfacing) = 56 hits. Substituting these values in Eq. (10) gives d(counterboring, spotfacing) = 0.2146. Using Eq. (9) we obtain v(counterboring, spot facing) = 0.7854.

Calculations were carried out for pairwise similarities between all the pairs of processes, resulting in 45 comparisons. Table 2 summarizes the results of the calculations.

The best single similarity measure was the van der Weken similarity ($sim_{vanDerWeken2}$) with correlation coefficients of $R = 0.828$, and $R = 0.916$, for Scirus and Scholar respectively. Then, the Jaccard's coefficient ($sim_{Jaccard}$) came up second.

7. Case study

In this case study we focus on the effectiveness of the ontology-based semantic similarities in the context of product-service system (PSS) design. A PSS is a mix of both products and services that is often associated to better sustainability. In the design of PSS systems, one common design decision is the selection of the type of service that can be integrated with a given product [6].

A CBR system was developed in Java by extending the open source software FreeCBR. As in any traditional CBR system, each case is defined in terms of a problem and a solution. In this case study, the problem is defined in terms of case features that

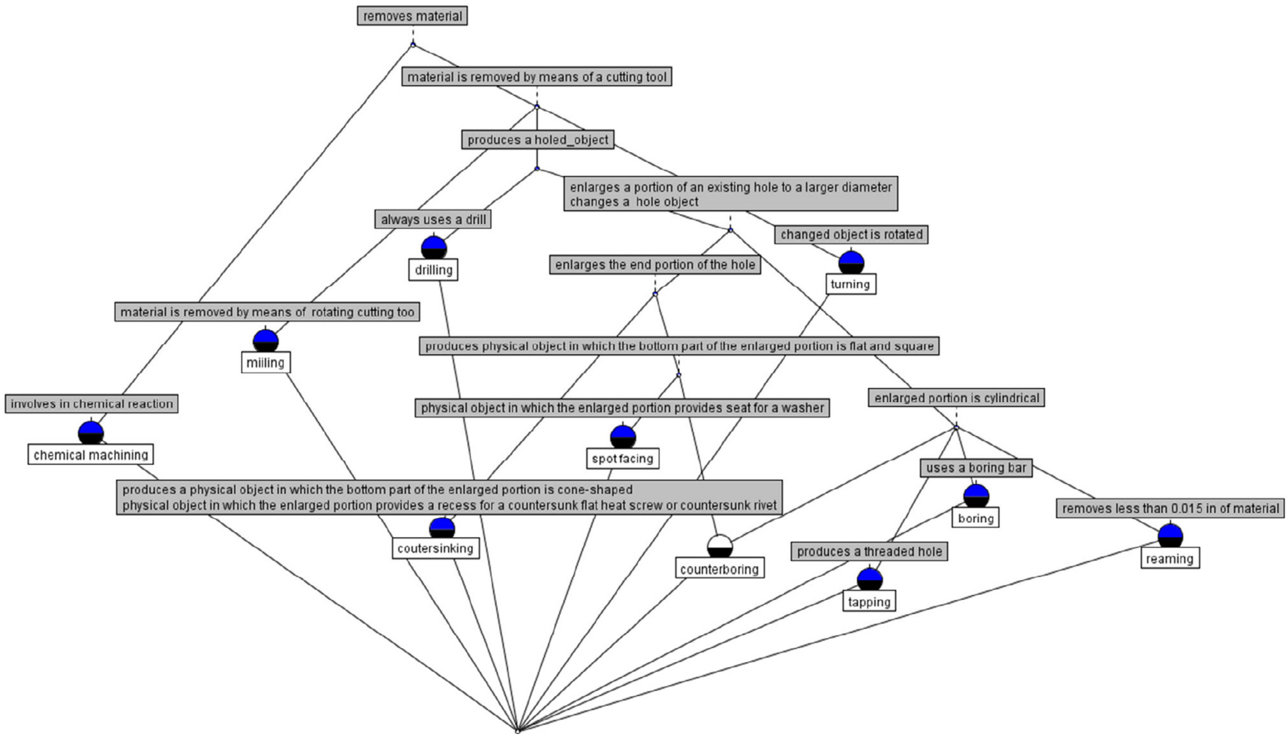

Fig. 7. The concept lattice of material removal process.

represent characteristics of a given product. The case features can be numeric or semantic. For numeric features, the index approach proposed by Lin et al. is used [6]. The semantic feature is specified as the class to which the product belongs, that is defined in a product ontology. The similarity for such semantic feature was calculated using the Overlap coefficient and the formal attributes of each class. A screen dump of the CBR system is shown in Fig. 8.

The case similarity in the CBR system is calculated using the equation proposed by Kolodner and Simpson [46]:

$$S(t,r) = \frac{\sum_{i=1}^{n} w_i \times sim(f_i^t, f_i^r)}{\sum_{i=1}^{n} w_i} \tag{11}$$

where $S(t,r)$ is the global similarity between the target case t and a source case r; w_i is the weight of feature i; f_i^t is the value of feature i of target case t; and f_i^r is the value of feature i of a source case r.

$sim(f_i^t, f_i^r)$ is calculated according to the following criteria, which is based on the Overlap coefficient and a similarity for numerical attributes.

$$m(f_i^t, f_i^r) = \begin{cases} \dfrac{|A_i^t \cap A_i^r|}{\min(|A_i^t|, |A_i^r|)} & \text{if} \quad f_i^t, f_i^r \quad \text{are } \textit{ontology} \text{ classes} \\ 1 - \dfrac{|f_i^t \cap f_i^r|}{f_i^{\max} - f_i^{\min}} & \text{if} \quad f_i^t, f_i^r \quad \text{are numerical features} \end{cases} \tag{12}$$

where A_i^t is the set of formal attributes of the class specified in feature f_i^t; A_i^r is the set of formal attributes of the class specified in feature f_i^r; and $f_i^{\max}$ and $f_i^{\min}$ are the maximum and minimum numeric values of feature f_i, respectively.

7.1. Ontology development

A product ontology was developed based on the procedure described in Section 3 and using the list of products reported in [6]. The resulting concept lattice is shown in Fig. 9. The product ontology extends the upper ontology defined in the ISO 15926 standard [15]. To carry out the alignment, three classes were added as subclasses of *physical object*: substance, mixture, and device based on [47].

During the preparation of the FCA context table, attributes were selected by investigating the process or processes in which the product participates or is involved. Each process was described according to the process characterization explained in Section 3. For example, the objects that are transformed during the operation of a copier are the data input by the user, electricity, and paper and the objects that are produced by the same process are the copied printed paper. Thus, the attributes of the copier become: consumes data, consumes electricity, consumes paper, and produces printed paper.

7.2. Experiment setup

The case base was populated with information about 47 successful product services systems. Each case was described in terms of numerical and semantic features. Based on [6], the following numeric features and weights were used: place of usage of the PSS system ($w_1 = 0.116$), frequency of usage of the PSS system ($w_2 = 0.232$), product fashion cycle ($w_3 = 0.042$), product volume ($w_4 = 0.036$), product weight ($w_5 = 0.034$), product useful life ($w_6 = 0.064$), product price ($w_7 = 0.082$), subsequent expenditure ($w_8 = 0.085$), GDP per capita ($w_9 = 0.119$), population density ($w_{10} = 0.079$), area of territory ($w_{11} = 0.052$), and temperature range of the territory ($w_{12} = 0.059$). The allowable values for each numeric feature and their meaning is also explained in [6]. For example, the index used to describe the place of usage of the PSS system is defined for integer values ranging from 1 to 3, where 1 represents indoor, 3 outdoor and 2 both. Among these features, product fashion cycle, volume, weight, useful life, and price are product features. The objective of this experiment was to evaluate the possibility of using a semantic feature as a replacement of some of the product attributes. The

Table 2
The results for experiments of material removal process.

	Pairs	Scirus	Scholar	Wu Palmer	Lin	Dice	All Confidence	Overlap	van der Weken 1	van der Weken 2	Jaccard	Cosine	Tversky
Counterboring with	Counterboring	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	Milling	0.543	0.546	0.286	0.063	0.533	0.400	0.800	0.900	0.600	0.364	0.566	0.400
	Countersinking	0.862	0.857	0.727	0.443	0.800	0.800	0.800	0.800	0.800	0.667	0.800	0.800
	Drilling	0.607	0.605	0.500	0.140	0.625	0.500	0.833	0.900	0.643	0.455	0.645	0.500
	Spotfacing	0.785	0.797	0.833	0.620	0.900	0.900	0.900	0.900	0.900	0.818	0.900	0.900
	Boring	0.514	0.687	0.800	0.443	0.842	0.800	0.889	0.900	0.818	0.727	0.843	0.800
	Reaming	0.763	0.766	0.800	0.443	0.842	0.800	0.889	0.900	0.818	0.727	0.843	0.800
	Turning	0.548	0.538	0.286	0.063	0.533	0.400	0.800	0.900	0.600	0.364	0.566	0.400
	Tapping	0.695	0.703	0.800	0.203	0.842	0.800	0.889	0.900	0.818	0.727	0.843	0.800
Milling with	Milling	1.000	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	Countersinking	0.581	0.566	0.286	0.063	0.533	0.400	0.800	0.900	0.600	0.364	0.566	0.400
	Drilling	0.836	0.798	0.400	0.063	0.727	0.667	0.800	0.929	0.867	0.571	0.730	0.667
	Spotfacing	0.515	0.526	0.250	0.063	0.533	0.400	0.800	0.900	0.600	0.364	0.566	0.400
	Boring	0.686	0.729	0.286	0.063	0.571	0.444	0.800	0.909	0.667	0.400	0.596	0.444
	Reaming	0.674	0.657	0.286	0.063	0.571	0.444	0.800	0.909	0.667	0.400	0.596	0.444
	Turning	0.883	0.805	0.500	0.063	0.800	0.800	0.800	0.933	0.933	0.667	0.800	0.800
	Tapping	0.706	0.669	0.286	0.063	0.571	0.444	0.800	0.909	0.667	0.400	0.596	0.444
Countersinking with	Countersinking	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	Drilling	0.636	0.628	0.500	0.140	0.625	0.500	0.833	0.900	0.643	0.455	0.646	0.500
	Spotfacing	0.771	0.798	0.727	0.443	0.800	0.800	0.800	0.800	0.800	0.667	0.800	0.800
	Boring	0.509	0.674	0.600	0.203	0.737	0.700	0.778	0.800	0.727	0.583	0.738	0.700
	Reaming	0.789	0.775	0.600	0.203	0.737	0.700	0.778	0.800	0.727	0.583	0.738	0.700
	Turning	0.567	0.542	0.286	0.063	0.533	0.400	0.800	0.900	0.600	0.364	0.566	0.400
	Tapping	0.709	0.712	0.600	0.203	0.737	0.700	0.778	0.800	0.727	0.583	0.738	0.700
Drilling with	Drilling	1.000	0.988	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	Spotfacing	0.566	0.573	0.400	0.140	0.625	0.500	0.8333	0.900	0.643	0.455	0.645	0.500
	Boring	0.697	0.773	0.500	0.140	0.667	0.556	0.833	0.909	0.714	0.500	0.680	0.556
	Reaming	0.729	0.723	0.500	0.140	0.667	0.556	0.833	0.909	0.714	0.500	0.680	0.556
	Turning	0.808	0.767	0.400	0.063	0.727	0.667	0.800	0.929	0.867	0.571	0.730	0.667
	Tapping	0.762	0.736	0.500	0.140	0.667	0.556	0.833	0.909	0.714	0.500	0.680	0.556
Spotfacig with	Spotfacing	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	Boring	0.474	0.546	0.545	0.203	0.737	0.700	0.778	0.800	0.727	0.583	0.738	0.700
	Reaming	0.694	0.635	0.545	0.203	0.737	0.700	0.778	0.800	0.727	0.583	0.738	0.700
	Turning	0.502	0.413	0.286	0.063	0.533	0.400	0.800	0.900	0.600	0.364	0.566	0.400
	Tapping	0.638	0.581	0.545	0.203	0.737	0.700	0.778	0.800	0.727	0.583	0.738	0.700
Boring with	Boring	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	Reaming	0.601	0.794	0.800	0.443	0.889	0.889	0.889	0.909	0.909	0.800	0.889	0.889
	Turning	0.669	0.723	0.250	0.063	0.571	0.444	0.800	0.909	0.667	0.400	0.596	0.444
	Tapping	0.602	0.783	0.800	0.443	0.889	0.889	0.889	0.909	0.909	0.800	0.889	0.889
Reaming with	Reaming	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	Turning	0.661	0.639	0.286	0.063	0.571	0.444	0.800	0.909	0.667	0.400	0.596	0.444
	Tapping	0.805	0.805	0.800	0.443	0.889	0.889	0.889	0.909	0.909	0.800	0.889	0.889
Turning with	Turning	1.000	0.998	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	Tapping	0.707	0.667	0.286	0.063	0.571	0.444	0.800	0.909	0.667	0.400	0.596	0.444
Tapping with	tapping	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	Correlation with Scirus	1	0.983	0.711	0.786	0.764	0.743	0.747	0.628	0.828	0.787	0.770	0.743
	Sum of squared errors (Scirus)			2.409	9.249	0.588	1.04	1.269	2.133	0.499	1.352	0.562	1.040
	2σ (Scirus)			0.125	0.352	0.050	0.051	0.069	0.098	0.051	0.058	0.050	0.051
	Correlation with Scholar	0.983	1	0.835	0.865	0.877	0.858	0.822	0.614	0.916	0.891	0.882	0.858
	Sum of squared errors (Scholar)			2.058	9.259	0.291	0.727	1.019	1.923	0.256	1.124	0.272	0.727
	2σ (Scholar)			0.122	0.317	0.17	0.040	0.061	0.102	0.017	0.059	0.017	0.040

semantic feature consisted of the class of product defined in the product ontology.

Initially, two experiments were carried out. The objective of experiment 1 was to provide a reference for comparing the proposed approach. For this purpose, all the queries in experiment 1 consisted of values for all the numerical features without the semantic feature.

In experiment 2, queries were formulated by replacing two product features (product volume and product weight) by the corresponding class of product from the ontology. The weight for this semantic feature was set to $w_{13} = 0.07$.

The case similarity in both experiments were calculated with Eqs. (11) and (12).

The queries were formulated with the product information from each of the cases stored in the case base. Therefore, 47 problems were defined with the problem data of the 47 cases in the case base, resulting in a total of 94 experiments. The objective was to find the service strategy and then compare the result with the

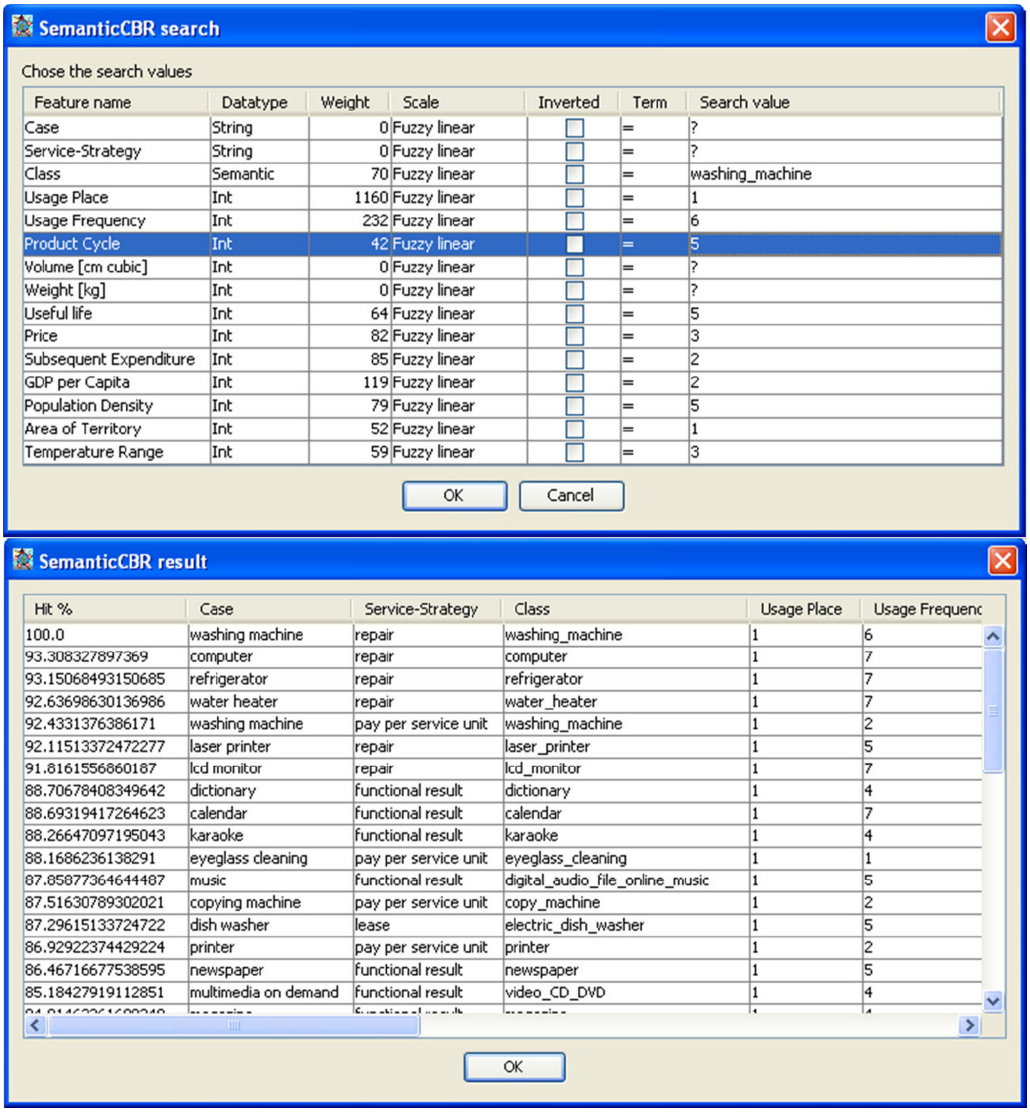

Fig. 8. Screen dump of the user interface of the semantic CBR system.

already known service strategy of the corresponding case. For example, problem 1 describes a certain kind of washing machine that was used in PSS that provided a repair service. In this example, it is thus expected that all if not most of the *n* best matches return repair as the solution.

The execution of each query resulted in a ranked list of matches each of which included product information, the proposed service strategy, and a global similarity value. Then the resulting service strategies were compared against the original service. Table 3 shows the results for both experiments. The best five matches are shown for each problem. From the overall results, it can be observed that there are nine problems (Nos. 1, 5, 10, 14, 18, 26, 28, 43 and 45) in which the results of experiment 1 are identical with those of experiment 2. For example, the best five service strategies in problem 1 were: refrigerator-repair, computer-repair, water heater-repair, laser printer-repair and LCD monitor-repair, all of which are consistent to the repair service corresponding to the solution of problem 1.

Other problems produced slightly different results. For example, in experiment 2, problems 11, 12, 19, 20, 33, 34, 37, 38, 40, 41 and 42 produced the same five best matches found in the results of experiment 1 but with a different ranking. For example, in problem 11 both experiments resulted in treadmill-lease, dryer-lease, LCD TV-lease, refrigerator-lease and dish washer-lease. However, while treadmill-lease has the highest rank in experiment 1, it appears second in experiment 2.

In addition, there were 27 results (such as problems 2–9) that differed in one or two cases. For example, the results for problem 2 include an Internet-based digital calendar which is false positive. On the other hand, some results of experiment 2 were good matches albeit being missing in experiment 1. For example, (sofa-lease and platform bed-lease instead of jewelry-rental and handbag rental) in problem 9 are good matches.

Furthermore, the results of experiment 2 for problems 30 (photocopy-service), 31 (scanning-service) and 32 (laminating-service) are better when compared to the results of experiment 1 in which not only the best 5 matches refer to a service that equals that of the case from which the query was formulated (pay per service unit) but also the product is more compatible with that of the suggested service. For example, experiment 2 for problem 30 resulted in laundry-service, printing-service, eyeglass cleaning-service, scanning-service and fax-service. Among these, printing, scanning and fax can be carried out with a copier machine. These results contrast with those obtained with experiment 1 which included cleaning-service, eyeglass cleaning-service and shoes cleaning service.

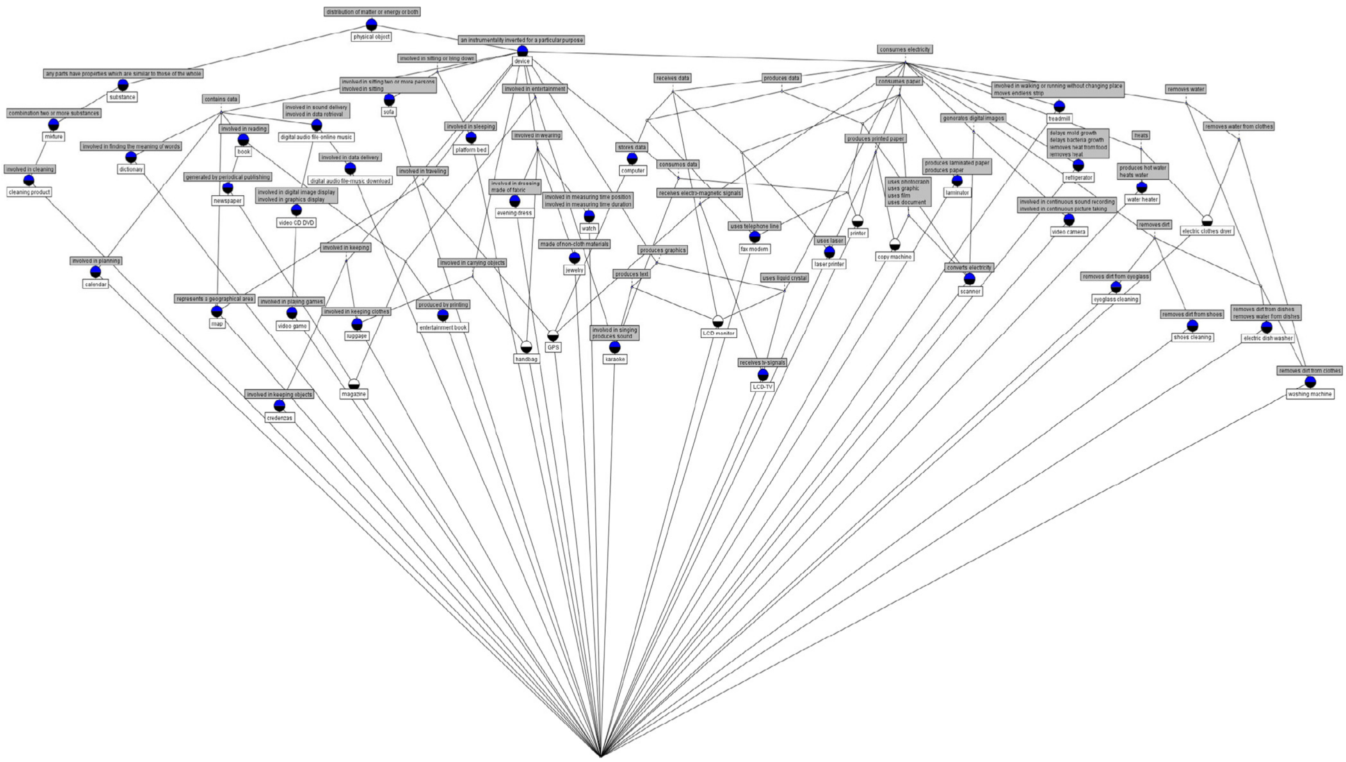

Fig. 9. The concept lattice of products.

Table 3
The results for experiments of product-service system.

Problem	Case (product and known solution)		Best five matches			
	Product	Service	Experiment 1 (numeric features only)		Experiment 2 (class feature replaces weight and volume features)	
			Product-service	Similarity	Product-service	Similarity
1	Washing machine	Repair	Refrigerator repair	94	Refrigerator repair	90
			Computer repair	91.29	Computer repair	89.82
			Water heater repair	89.92	Water heater repair	88.45
			Laser printer repair	89.85	Laser printer repair	87.88
			LCD monitor repair	88.11	LCD monitor repair	87.27
2	Refrigerator	Repair	Water heater repair	95.92	Water heater repair	94.45
			Washing machine repair	94	Computer repair	91.56
			Computer repair	93.02	LCD monitor repair	90.43
			LCD monitor repair	91.26	Washing machine repair	90
			Laser printer repair	83.85	Digital calendar	82.12
3	Computer	Repair	LCD monitor repair	95.394	LCD monitor repair	94.19
			Refrigerator repair	93.02	Refrigerator repair	91.56
			Water heater repair	93.01	Washing machine repair	89.82
			Washing machine repair	91.29	Water heater repair	89.51
			Laser printer repair	85.21	Digital calendar	84.97
4	Laser printer	Repair	Washing machine repair	89.85	Washing machine repair	87.88
			Printing service	88.4	Printing service	87.23
			Computer repair	85.21	Computer repair	82.87
			Refrigerator repair	83.85	Refrigerator repair	81.88
			Water heater repair	83.83	Online karaoke	80.63
5	LCD monitor	Repair	Computer repair	95.39	Computer repair	94.19
			Water heater repair	95.34	Water heater repair	92.98
			Refrigerator repair	91.26	Refrigerator repair	90.43
			Washing machine repair	88.11	Washing machine repair	87.27
			Digital calendar	86.16	Digital calendar	85.59
6	Water heater	Repair	Refrigerator repair	95.92	Refrigerator repair	94.45
			LCD monitor repair	95.34	LCD monitor repair	92.98
			Computer repair	93.01	Computer repair	89.51
			Washing machine repair	89.92	Washing machine repair	88.45
			Laser printer repair	83.83	Digital calendar	84.17
7	Handbag	Repair	Jewelry repair	95	Jewelry repair	94.15
			Watch repair	93.18	Watch repair	92.33
			Audio book	78.20	Treadmill lease	78.39
			LCD TV lease	77.63	LCD TV lease	77.3
			Treadmill lease	77.59	Handbag rental	76.57

Table 3 (*Continued*)

Problem	Case (product and known solution)		Best five matches			
	Product	Service	Experiment 1 (numeric features only)		Experiment 2 (class feature replaces weight and volume features)	
			Product-service	Similarity	Product-service	Similarity
8	Jewelry	Repair	Handbag repair	95	Handbag repair	94.15
			Watch repair	94.08	Watch repair	92.33
			Jewelry rental	80.67	Jewelry rental	80.67
			Handbag rental	79.77	Handbag rental	78.92
			Download audio book	75.00	Treadmill lease	74.29
9	Watch	Repair	Jewelry repair	94.08	Jewelry repair	92.33
			Handbag repair	93.18	Handbag repair	92.33
			Refrigerator lease	75.09	Refrigerator lease	76.99
			Jewelry rental	74.75	Sofa lease	75.86
			Handbag rental	73.85	Platform bed lease	74.25
10	Treadmill	Lease	Washing machine lease	94.85	Washing machine lease	92.95
			LCD TV lease	90.04	LCD TV lease	88.37
			Dryer lease	89.68	Dryer lease	87.78
			Dish washer lease	82.06	Dish washer lease	79.26
			Refrigerator lease	81.12	Refrigerator lease	79.22
11	Washing machine	Lease	Treadmill lease	94.85	Dryer lease	93.67
			Dryer lease	94.83	Treadmill lease	92.95
			LCD TV lease	93.39	LCD TV lease	91.42
			Refrigerator lease	86.27	Dish washer lease	84.31
			Dish washer lease	85.41	Refrigerator lease	82.27
12	LCD TV	Lease	Washing machine lease	93.39	Washing machine lease	91.42
			Treadmill lease	90.04	Treadmill lease	88.37
			Refrigerator lease	88.81	Refrigerator lease	86.84
			Dish washer lease	88.33	Dryer lease	86.76
			Dryer lease	88.22	Dish washer lease	85.46
13	Sofa	Lease	Platform bed lease	91.8	Platform bed lease	90.05
			Credenzas lease	85.29	Credenzas lease	82.69
			Refrigerator lease	85.20	Refrigerator lease	81.89
			Treadmill lease	79.075	Download audio book	77.43
			Dish washer lease	78.02	Treadmill lease	76.68
14	Dryer	Lease	Washing machine lease	94.83	Washing machine lease	93.67
			Treadmill lease	89.68	Treadmill lease	87.78
			LCD TV lease	88.22	LCD TV lease	86.76
			Dish washer lease	82.84	Dish washer lease	81.41
			Refrigerator lease	81.1	Refrigerator lease	77.6
15	Platform bed	Lease	Credenzas lease	93.49	Credenzas lease	90.89
			sofa lease	91.8	Sofa lease	90.05
			Dish washer lease	86.22	Dish washer lease	84.52
			Refrigerator lease	82.89	Music download	84.33
			Music download	80.83	Download audio book	83.03
16	Refrigerator	Lease	LCD TV lease	88.81	LCD TV lease	86.84
			Washing machine lease	86.27	Washing machine lease	82.27
			Sofa lease	85.19	Sofa lease	81.89
			Platform bed lease	82.89	Platform bed lease	80.29
			Credenzas lease	82.45	Treadmill lease	79.22
17	Credenzas	Lease	platform bed lease	93.67	Platform bed lease	90.89
			Sofa lease	85.42	Music download	82.83
			Refrigerator lease	82.42	Sofa lease	82.69
			Dish washer lease	81.5	Download audio book	81.53
			Music download	80.83	Refrigerator lease	78.95
18	Dish washer	Lease	LCD TV lease	88.33	LCD TV lease	85.46
			Platform bed lease	86.22	Platform bed lease	84.52
			Washing machine lease	85.41	Washing machine lease	84.31
			Dryer lease	82.84	Dryer lease	81.41
			Treadmill lease	82.06	Treadmill lease	79.26
19	Luggage box	Rental	GPS rental	94.85	GPS rental	91.55
			Scanning service	87.5	Scanning service	84.2
			Cleaning service	84.95	Cleaning service	80.83
			Video camera rental	83.67	Eyeglass cleaning service	79.95
			Eyeglass Cleaning service	83.25	Video camera rental	79.47
20	Video CD/DVD	Rental	Entertainment book rental	94.71	Multimedia on demand	92.27
			Fax service	92.63	Entertainment book rental	91.21
			Multimedia on demand	92.27	Fax service	90.8
			Online magazine	90.84	Online music	88.4
			Online music	88.4	Online magazine	87.34
21	Evening dress	Rental	Handbag rental	85.38	Handbag rental	85.66
			Jewelry rental	84.48	Jewelry rental	85.66
			Video game rental	73.73	Video game rental	72.47
			Photographer service	70.28	Handbag repair	66.33
			Video camera rental	70.28	Jewelry repair	66.33

Table 3 (*Continued*)

Problem	Case (product and known solution)		Best five matches			
	Product	Service	Experiment 1 (numeric features only)		Experiment 2 (class feature replaces weight and volume features)	
			Product-service	Similarity	Product-service	Similarity
22	Entertainment book	Rental	Video CD/DVD rental	94.71	Video CD/DVD rental	91.21
			Scanning service	90.29	Scanning service	87.66
			Online magazine	88.4	Online magazine	87.23
			Fax service	87.34	Fax service	85.84
			Multimedia on demand	86.98	Eyeglass cleaning service	83.88
23	Video game	Rental	Entertainment book rental	78.18	Video CD/DVD rental	75.73
			Jewelry rental	76.00	Entertainment book rental	74.68
			Video CD/DVD rental	75.73	Handbag rental	72.50
			Handbag rental	75.10	Jewelry rental	72.50
			Audio book	73.77	Evening dress rental	72.47
24	Jewelry	Rental	Handbag rental	99.1	Handbag rental	98.25
			Evening dress rental	84.48	Evening dress rental	85.66
			Jewelry repair	80.67	Jewelry repair	80.67
			Video game rental	76.00	Handbag repair	74.82
			Handbag repair	75.67	Watch repair	73
25	Handbag	Rental	Jewelry rental	99.1	Jewelry rental	98.25
			Evening dress rental	85.38	Evening dress rental	85.66
			Jewelry repair	79.77	Jewelry repair	78.92
			Handbag repair	76.57	Handbag repair	76.57
			Video game rental	75.10	Watch repair	73
26	GPS	Rental	Luggage rental	94.85	Luggage rental	91.55
			Eyeglass cleaning service	88.4	Eyeglass Cleaning service	85.6
			Cleaning service	88.3	Cleaning service	85.08
			Video camera rental	87.02	Video camera rental	84.42
			Laminating service	84.15	Laminating service	80.65
27	DV(video camera)	Rental	Photographer service	95.9	Photographer service	95.9
			Cleaning service	88.72	Eyeglass Cleaning service	85.12
			Eyeglass cleaning service	87.02	Cleaning service	84.6
			GPS rental	87.02	GPS rental	84.42
			Luggage rental	83.67	Laundry service	82.6
28	Fax modem	Pay per service unit	Scanning service	92.95	Scanning service	92.08
			Video CD/DVD rental	92.63	Video CD/DVD rental	90.8
			Online dictionary	89.1	Online dictionary	88.77
			Laundry service	88.72	Laundry service	87.65
			Eyeglass cleaning service	88.7	Eyeglass Cleaning service	87.03
29	Printer	Pay per service unit	Laundry service	93.72	Copying service	92.53
			Laminating service	92.05	Laundry service	92.25
			Copying service	91.67	Laminating service	91.75
			Laser printer repair	88.4	Laser printer repair	87.23
			Cleaning service	87.9	Eyeglass cleaning service	87.03
30	Photostat	Pay per service unit	Laundry service	97.95	Laundry service	93.95
			Printing service	91.67	Printing service	92.53
			Cleaning service	88.07	Eyeglass cleaning service	89.23
			Eyeglass cleaning service	87.97	Scanning service	87.01
			Shoes cleaning service	87.26	Fax service	86.6
31	Scanning	Pay per service unit	Eyeglass cleaning service	95.75	Eyeglass cleaning service	92.95
			Fax service	92.95	Fax service	92.08
			Laminating service	91.5	Laminating service	89.17
			Entertainment book rental	90.29	Entertainment book rental	87.66
			Video CD/DVD rental	87.85	Copying service	87.01
32	Laminating	Pay per service unit	Eyeglass cleaning service	95.75	Eyeglass cleaning service	92.95
			Printing service	92.05	Printing service	91.75
			Scanning service	91.5	Scanning service	89.17
			Laundry service	85.77	Laundry service	86.33
			fax service	84.45	Copying service	85.45
33	Washing machine	Pay per service unit	Copying service	97.95	Copying service	93.95
			Printing service	93.72	Eyeglass cleaning service	92.68
			Cleaning service	90.12	Printing service	92.25
			Eyeglass Cleaning service	90.02	Fax service	87.65
			Fax service	88.72	Cleaning service	86.9
34	Cleaning product	Pay per service unit	Laundry service	90.12	Laundry service	86.9
			Video camera Rental	88.72	Eyeglass cleaning service	85.08
			Eyeglass Cleaning service	88.3	GPS rental	85.08
			GPS rental	88.3	Copying service	84.85
			Copying service	88.07	Video camera rental	84.6
35	Shoes cleaning	Pay per service unit	Copying service	87.26	Eyeglass cleaning service	91.53
			Laundry service	85.21	Copying service	90.36
			Eyeglass Cleaning service	82.96	Laundry service	89.71
			Printing service	78.925	Laminating service	85.88
			Laminating service	78.71	Scanning service	85.88

Table 3 (*Continued*)

Problem	Case (product and known solution)		Best five matches			
	Product	Service	Experiment 1 (numeric features only)		Experiment 2 (class feature replaces weight and volume features)	
			Product-service	Similarity	Product-service	Similarity
36	Eyeglass cleaning	Pay per service unit	Laminating service	95.75	Laminating service	92.95
			Scanning service	95.75	Scanning service	92.95
			Laundry service	90.02	Laundry service	92.68
			Fax service	88.7	Copying service	89.23
			GPS rental	88.4	Printing service	87.03
37	DV(video camera)	Pay per service unit	Video camera rental	95.9	Video camera rental	95.9
			Cleaning service	84.62	Eyeglass cleaning service	81.02
			Eyeglass Cleaning service	82.92	Copying service	80.55
			GPS rental	82.92	Cleaning service	80.5
			Copying service	80.88	GPS rental	80.32
38	Music CD (online music)	Functional result	Online newspaper	97.16	Multimedia on demand	96.13
			Multimedia on demand	96.13	Online newspaper	94.36
			Online magazine	94.71	Online magazine	91.91
			Online dictionary	91.93	Online dictionary	90.18
			Video CD/DVD rental	88.4	Video CD/DVD rental	88.4
39	Magazine	Functional result	Multimedia on demand	98.58	Multimedia on demand	95.08
			Online newspaper	94.71	Online newspaper	94.71
			Online music	94.71	Online dictionary	92.63
			Online dictionary	94.38	Online music	91.91
			Video CD/DVD rental	90.84	Online karaoke	88.78
40	Karaoke	Functional result	Multimedia on demand	91.67	Online magazine	88.78
			Online magazine	90.24	Multimedia on demand	88.7
			Online music	87.8	Online dictionary	86
			Online dictionary	87.47	Online music	85.63
			Online newspaper	84.96	Online newspaper	82.79
41	Music CD (music download)	Functional result	Download audio book	98.7	Download audio book	96.95
			Platform bed lease	80.83	Platform bed lease	84.33
			Credenzas lease	80.23	Credenzas lease	82.83
			Dish washer lease	78.38	Dish washer lease	78.92
			Online music	78.03	Online music	78.03
42	Video CD/DVD(multimedia on demand)	Functional result	Online magazine	98.58	Online music	96.13
			Online music	96.13	Online magazine	95.08
			Online dictionary	95.8	Online dictionary	94.05
			Online newspaper	93.29	Video CD/DVD rental	92.27
			Video CD/DVD rental	92.27	Online newspaper	90.49
43	MAP	Functional result	Online magazine	85.4	Online magazine	83.5
			Online dictionary	83.98	Online dictionary	83.13
			Multimedia on demand	83.98	Multimedia on demand	82.08
			Luggage rental	80.79	Luggage rental	78.63
			Online newspaper	80.11	Online newspaper	78.21
44	Newspaper	Functional result	Online music	97.16	Online magazine	83.5
			Online magazine	94.71	Online dictionary	83.13
			Multimedia on demand	93.29	Multimedia on demand	82.08
			Digital calendar	89.12	Luggage rental	78.63
			Online dictionary	89.09	Online newspaper	78.21
45	Dictionary	Functional result	Multimedia on demand	95.8	Multimedia on demand	94.05
			Online Magazine	94.38	Online magazine	92.63
			Online music	91.93	Online music	90.18
			Fax service	89.1	Fax service	88.77
			Online newspaper	89.09	Online newspaper	87.3
46	Calendar	Functional result	Online newspaper	89.12	Online newspaper	87.37
			Online music	86.27	LCD monitor repair	85.59
			LCD monitor repair	86.16	Computer repair	84.97
			Online dictionary	84.51	Online music	84.52
			Computer repair	84.40	Water heater repair	84.17
47	Book	Functional result	Music download	98.7	Music download	96.95
			Platform bed lease	79.53	Platform bed lease	83.03
			Credenzas lease	78.93	Credenzas lease	81.53
			Handbag repair	78.20	Dish washer lease	78.78
			Dish washer lease	77.08	Sofa lease	77.43

In order to corroborate the influence of the semantic similarity, an additional experiment was conducted (experiment 3). Experiment 3 excluded the numeric product features of volume and weight as well as the semantic feature. For this evaluation, we counted the cases in the best five results that were common to those in experiment 1. In other words, the ideal number of common cases is 5. These results are summarized in Table 4. The presence of the semantic similarity measure in experiment 2 resulted in an average of 4.32 common cases, while its absence in experiment 3 resulted in an average of 3.77. This means that in the absence of data for product volume and weight, the use of the semantic feature shows an improvement of almost 15% compared to not using it. From this, it can be concluded that ontology-based semantic similarities have the ability to emulate (at least to some extent) the numeric product features.

Table 4
Comparison of identical cases.

Problem	Number of best cases that are identical with experiment 1	
	Experiment 2 (using a class feature instead of volume and weight)	Experiment 3 (only numeric features but volume and weight are excluded)
1	5	5
2	4	5
3	4	4
4	4	4
5	5	5
6	4	4
7	4	2
8	4	4
9	3	3
10	5	5
11	5	5
12	5	5
13	4	4
14	5	5
15	4	4
16	4	5
17	4	4
18	5	4
19	5	3
20	5	4
21	3	4
22	4	3
23	4	3
24	4	3
25	4	3
26	5	3
27	4	2
28	5	4
29	4	4
30	3	3
31	4	5
32	4	4
33	5	4
34	5	1
35	4	4
36	3	3
37	5	2
38	5	4
39	4	4
40	5	5
41	5	4
42	5	4
43	5	0
44	3	5
45	5	4
46	4	4
47	4	4
Average	4.32	3.77

8. Related work

Several efforts are reported on the use of ontologies in product design. One interesting example is the work of Patil et al. [48] who describe an ontology that is based on definitions from the NIST's Core Product Model. They describe a methodology for building artifact ontologies which is based on the identification of subclasses of artifact, feature, assembly, and other classes. In the ontology development process, each artifact is characterized in terms of its form, function, and behavior (the implementation of the function).

Annamalai et al. [49] define a general framework for product-service systems that is facilitated by means of an ontology. The terminologies and semantics are based on eight top-level classes that cover product life cycle and the supporting elements such as stakeholder involvement. The ontology is developed using findings from literature review and the opinion of domain experts.

In the area of product customization, Tseng et al. [50] present a CBR system to support conceptual product design. In their work, a numeric similarity measure is combined with part-whole information that has a tree representation. Another similar work is that of Cobb and Agogino [51] who developed a CBR system for designing Micro-Electro-Mechanical Systems (MEMS). They discuss the results of a case-retrieval experiment in which MEMS are described in terms of functional and structural features. These features are numeric, which suggest that case retrieval is carried out by means of a numeric similarity.

In an attempt to generate new product ideas, Wu et al. [52] propose a CBR system in which a product is represented as a numeric vector consisting of 87 elements. Each element represents a product attribute. The product attributes are organized into five dimensions: interface modality, task, physical feature, environment, and users. Some of the attributes in the interface modality resemble the use of the participation relation defined in ISO 15926 such as specifying the parts of the body involved in [the use of] a given product. The task dimension represents the tasks to be performed by the user through the use of the product. Attributes in this dimension are equivalent to specific processes associated to a product. The physical dimension is for attributes such as product sizes. Environment includes attributes such as indoor or outdoor places. Finally, attributes in the user dimension characterize the user in terms of gender, age, etc. Every attribute in the product vector requires a value that represents the relevancy to that attribute.

Lin et al. [6] propose the use of CBR to support the design of product service systems (PSS). Specifically, their CBR selects service strategies for a given product. A case is described in terms of 12 features which are grouped into three categories, namely, user behavior, product, and environmental environment. User behavior is specified in terms of place of usage, and frequency of usage. The product is specified in terms of features describing its fashion cycle, volume, weight, useful life, price, and subsequent expenditure. External environment is defined in terms of GDP per capita, population density, area of territory, and temperature range. Each feature is quantified using integer values. The case similarity is obtained by using a weighted summation of all the feature similarities. The weights are determined by means of the analytic hierarchy process (AHP).

In the area of Web services, Bramantoro et al. [13] propose a similarity measure that quantifies the semantic distance between classes in an ontology of the products that are delivered by the services. Their work is motivated by limitations of other approaches in which only certain superclass–subclass links were taken into account. Their approach was based on the path length, number of downward edges counted between two classes, and the number of common closest ancestors. It is interesting to note that despite the fact that some semantic measures already existed they were apparently unknown in that domain.

9. Conclusions

This paper presented ontology-based semantic similarity measures that determine the degree of likeness between two classes. The main distinguishing aspect of the proposed approach is the use of ontologies obtained with FCA coupled with feature-based similarities. Results of the numeric experiments showed that in all cases, the proposed semantic measures performed better than the similarities of Wu-Palmer and Lin.

In the electric appliance experiment, after removing the least performing pairs (electric kettle, television set) and (electric kettle, electric oven), the correlation saw an increase of approximately

25%. The reason might be that both television set and electric oven were characterized by processes which are unfamiliar to the common user. For example, toaster was characterized as a device that uses infrared radiation. In this case, infrared radiation was considered as a part of heating, which is directly related to toasting bread. Similarly, TV set was defined as a device that receives television signals.

When other devices were characterized in terms of processes and participating objects that were more familiar to the common user, the calculated similarities were close to the human judgments. However, albeit important to the designers, from a user point of view, subprocesses that are not directly perceived by the users (i.e. the mechanism with which a product achieves its given function) are probably not taken into account. This could be a limitation of the questionnaire approach for evaluating the similarities.

A CBR system for product service systems demonstrated the effectiveness of the proposed similarity measures. In the CBR case study, the combination of the ontology and the semantic similarity proved useful when some details such as weight and volume are not available. Therefore, the designer can be relieved by needing less data to define a given design problem, which is particularly important during the conceptual stage of the design.

Nevertheless, in a few instances the proposed approach resulted in mismatches. This could be due to the lack of attributes in the FCA context table. For example, the addition of attributes that emphasize the difference between software and hardware products could reduce the number of false positives for problem 2.

A key element in the proposed approach is the use of FCA. From an information modeling point of view, the use of formal concept analysis is useful but the development of the context table has a large degree of freedom. Specifically, the selection of attributes in the context table of the FCA analysis plays an important role. Therefore, the selection of attributes should be based on an explicit guideline. In this paper, the attributes were selected based on the assumption that every product performs or is-involved-in processes (or activities using the ISO 15926 terminology). Therefore, subclasses of *physical object* are characterized not only by the mereology of the objects, but also by the processes associated to them. A class of process is in turn characterized in terms of the objects that are transformed (inputs), objects that are produced (outputs), other participating objects, and the mereology of the process.

Finally, the results of the correlation between the different semantic similarities and human judgment or Web-based search suggest that multiple similarity measures can be used as a way to validate ontologies. The reason is that the accuracy of the ontology directly influences the correlation values.

Acknowledgments

The authors would like to thank the members of the Industrial Systems Engineering Group of Toyohashi University of Technology for their participation in the human judgment evaluation. We also express our appreciation to Yusuke Tsujioka and Yukihiko Takada for their help in the software implementation. We are grateful to Hongcai Wang for translating the data used in the case study and to Fariza Mohamad for her help in the elaboration of the questionnaire. The authors are also indebted to the reviewer and editor of this journal for their helpful comments and suggestions.

Appendix A. Formal concept analysis

Formal concept analysis (FCA) can be used to design ontologies from a list of potential classes and their respective attributes. FCA is an analysis technique for information processing based on applied lattice and order theory that can be used to generate taxonomies. In FCA, information is organized in terms of a set of formal objects O, a set of formal attributes A, and a set of binary relations $Y \subseteq O \times A$ containing all pairs $\langle o, a \rangle \in Y$ such that the object $o \in O$ has the attribute $a \in A$. For our purposes, the formal objects represent candidate classes for an ontology.

Information about these three sets is typically summarized by a context table such as the one shown in Fig. 3. In a context table, the objects are listed in the first column and the attributes in the first row of the table.

A formal concept is defined as the pair $\langle O_i, A_i \rangle$ such that:

1. $O_i \subseteq O$, $A_i \subseteq A$.
2. Every object in O_i has every attribute in A_i. Conversely, A_i is the set of attributes shared by all the objects in O_i.
3. For every object $p \in O$ that is not in O_i, there is an attribute in A_i that p does not have.
4. For every attribute in A that is not in A_i, there is an object in O_i that does not have that attribute.
5. Formal concepts can be partially ordered into a lattice, such that a concept subsumes another concept. Fig. 4 shows the lattice obtained with the data of Fig. 3. Several lattice-construction algorithms are available some of which have been successfully implemented in several applications.

References

[1] H.A. ElMaraghy, W.A. ElMaraghy (Eds.), Advances in Design, Springer, Berlin, 2006, p. 31.
[2] M.L. Maher, Process models for design synthesis, AI Magazine 11 (4) (1990).
[3] S. Szykman, J. Racz, C. Bochenek, R.D. Sriram, Web-based system for design artifact modeling, Design Studies 21 (2) (2000) 145–165.
[4] S.K. Ong, A.Y. Nee, Q.L. Xu, Design reuse for embodiment and detail design, in: Design Reuse in Product Development Modeling, Singapore, World Scientific Pub. Co Inc., 2009, pp. 237–247.
[5] J.J. Shah, D.K. Jeon, S.D. Urban, P. Bliznakov, M. Rogers, Database infrastructure for supporting engineering design histories, Computer-Aided Design 28 (5) (1996) 347–360.
[6] K. Lin, L. Shih, S. Lu, Y. Lin, Strategy selection for product service systems using case-based reasoning, African Journal of Business Management 4 (6) (2010) 987–994.
[7] M.M. Richter, Similarity, in: Case-Based Reasoning for Signals and Imaging, Springer Verlag, Berlin, Heidelberg, 2007, pp. 25–90.
[8] P. Cunningham, A taxonomy similarity mechanisms for case-based reasoning, UCD-CSI (2008).
[9] A. Maedche, S. Staab, Comparing Ontologies—Similarity Measures and a Comparison Study, Institute AIFB, University of Karlsruhe, Karlsruhe, Germany, 2001.
[10] D. Lin, An information-theoretic definition of similarity, Proc. 15th Int. Conf. on Machine Learning, Madison, Wisconsin, 1998.
[11] V. Cross, X. Hu, Using semantic similarity in ontology alignment, Sixth International Workshop on Ontology Matching, 2011.
[12] V. Olehschuk, A. Pedersen, Ontology based semantic similarity for documents, 14th International Workshop on Database and Expert System Application, Prague, 2003.
[13] A. Bramantoro, S. Krishnaswamy, M. Indrawan, A semantic distance measure for matching web services, LCNS 3807, WISE 2005 Workshops, New York, 2005.
[14] J.S. Madin, Advancing ecological research with ontologies, Trends in Ecology & Evolution 23 (3) (2008) 159–168.
[15] R. Batres, M. West, D. Leal, D. Price, K. Masaki, Y. Shimada, T. Fuchino, Y. Naka, An upper ontology based on ISO 15926, Computers and Chemical Engineering 31 (2007) 519–534.
[16] H. Haav, A semi-automatic method to ontology design by using FCA, Proc. Fourth Int. Conf. on Concept Lattices and Their Applications, (2004), pp. 13–24.
[17] R. Batres, Y. Shimada, T. Fuchino, A semantic approach for incident database development, IChemE Symposium Series No. 155, (2010), pp. 477–484.
[18] ISO-TC-184, ISO-TC-184/SC4, SO 10303, "Industrial automation systems and integration—product data representation and exchange—Part 11: Description methods: The EXPRESS language reference manual, in: ISO TC 184/SC4, 1994.
[19] D. Leal, ISO 15926 "Life cycle data for process plant": an overview, Oil & Gas Science and Technology 60 (4) (2005) 629–637.
[20] Z. Wu, M. Palmer, Verbs semantics and lexical selection, 32nd Annual Meeting of the Associations for Computational Lingustics, New Mexico, 1994.
[21] A. Tversky, Features of similarity, Psychological Review 84 (4) (1977) 327–352.
[22] L.R. Dice, Measures of the amount of ecologic association between species, Ecology 26 (3) (1945) 297–302.
[23] J. Han, M. Kamber, Data Mining: Concepts and Techniques, second ed., Morgan Kaufmann Pub., an imprint of Elsevier, 2006.

[24] S. Yevtushenko, Concept Explorer, Open Source Java Software, 2009[Online]. Available: ⟨http://sourceforge.net/projects/conexp, Release 1.3⟩.
[25] M. Uschold, M. King, Towards a methodology for building ontologies, Workshop on Basic Ontological Issues in Knowledge Sharing, 1995.
[26] M. Grueninger, M.S. Fox, Methodology for the design and evaluation of ontologies, Workshop on Basic Ontological Issues in Knowledge Sharing, 1995.
[27] N.F. Noy, D. McGuinness, Ontology Development 101: A Guide to Creating Your First Ontology, in: Stanford Knowledge Systems Laboratory Technical Report KSL-01-05. Standford, USA, 2001.
[28] M. Fernandez-Lopez, A. Gomez-Perez, N. Juristo, Methontology: From Ontological Art Towards Ontological Engineering, AAAI Technical Report SS-97-06, 1997.
[29] A. Gomez-Perez, Evaluation of taxanomic knowledge in ontologies and knowledge bases, 12th Workshop on Knowledge Acquisition, Modelling, and Management (KAW '99), Alberta, Canada, 1999.
[30] R. Stevens, C.A. Coble, S. Bechhofer, Ontology-based knowledge representation for bioinformatics, Briefings in Bioinformatics 1 (4) (2000) 398–414.
[31] J. Han, M. Kamber, J. Pei, Getting to know your data, in: Data Mining: Concepts and Techniques, Morgan Kaufmann, United States of America, 2013p. 71.
[32] M.A. Rodríguez, M. Egenhofer, Determining semantic similarity among entity classes from different ontologies, IEEE Transactions on Knowledge and Data Engineering 15 (2003) 442–456.
[33] D. van der Weken, M. Nachtegael, E.E. Kerre, Using similarity measures and homegeneity for the comparison of images, Image and Vision Computing (2004) 695–702.
[34] Stanford Center for Biomedical Informatics Research., [Online]. Available:
[35] Web-Ontology working group, OWL web ontology language, second ed., 2004 ⟨http://www.w3.org⟩.
[36] Jena, HP Labs Semantic Web Toolkit. [Online]., 2013 Available: ⟨http://jena.sourceforge.net/⟩.
[37] R.E. Schafrik, Unit manufacturing and assembly process, in: J. Lee (Ed.), Modern Manufacturing, Mechanical Engineering Handbook, CRC Press LLC, Boca Raton, FL, 1999, pp. 13-8–13-34.
[38] S. Kalpakjian, S.R. Schmid, Manufacturing Engineering and Technology, sixth ed., Prentice-Hall, Upper Saddle River, NJ, 2010.
[39] B.S. Nagendra Parashar, R.K. Mittal, Elements of Manufacturing Processes, Prentice-Hall of India Private Limited, New Delhi, India, 2007.
[40] E.P. Degarmo, J.T. Black, R.A. Kohser, Materials and Processes in Manufacturing, 10th ed., Wiley, Hoboken, NJ, 2010.
[41] B. Handley, D.M. Marshall, C. Coon, Principles of Engineering, Delmar, 2012.
[42] M.C. Finishing, Blasting Technical Information [Online], 2012 Available:⟨http://mcfinishing.com/resources/blastingtech.pdf⟩ [Accessed 2012].
[43] R. Cilibrasi, P. Vitanyi, The Google similarity distance, IEEE Transactions on Knowledge and Data Engineering 19 (3) (2007) 370–383.
[44] D. Giustini, How Big is Google Scholar? 500 Million Documents?, 2006[Online]⟨http://blogs.ubc.ca/googlescholar/2006/04/how-big-is-google-scholar-500-million-documents/⟩ (Available) [Accessed 25 May 2012].
[45] National Science Board, Science and Engineering Indicators 2010, National Science Foundation NSB 10-01, Arlington, VA, 2010.
[46] J. Kolodner, R.L. Simpson, Case-based reasoning, in: K.H. Lin, L.H. Shih, S.S. Lu, Y.T. Lin (Eds.), Strategy Selection for Product Service Systems Using Case-based Reasoning. African Journal of Business Management, vol. 4 (6), 2010.
[47] A. Pease, Suggested Upper Merged Ontology (SUMO) [Online], 2010 Available: ⟨http://www.ontologyportal.org/⟩ [Accessed 2010 December 2010].
[48] L. Patil, D. Dutta, R.D. Sriram, Ontology formalization of product semantics for product semantics for product life cycle management, NISTIR 7274 (November) (2005).
[49] G. Annamalai, R. Hussain, M. Cakkol, R. Roy, An Ontology for Product-Service Systems, Cranfield University, Cranfield, Bedfordshire, UK, 2011.
[50] H.-E. Tseng, C.-C. Chang, S.-H. Chang, Applying case-based reasoning for product configuration in mass customization environments, Expert Systems with Applications (2005) 913–925.
[51] C.L. Cobb, A.M. Agogino, Case-based reasoning for the design of micro-electro-mechanical systems, ASME 2006 International Design Engineering Technical Conferences & Computers and information in Engineering Conference, PA, USA., 2006.
[52] M.C. Wu, Y.F. Lo, S.H. Hsu, A case-based reasoning approach to generating new product ideas, International Journal of Advanced Manufacturing Technology 30 (2006) 166–173.

Suriati Akmal is a graduate student in the Department of Mechanical at the Toyohashi University of Technology. She received her BS in Manufacturing Engineering from IIU Malaysia in 2003. In 2006, she graduated with a MSc in Global Production Engineering from TU Berlin and was subsequently appointed to the post of lecturer at Manufacturing Department, University Teknikal Malaysia Melaka (UTeM). Currently, she is on leave from UTeM and working toward her doctoral degree in Industrial System Engineering. Her research interests include semantic similarity methods for products and processes, process planning and product development.

Li-Hsing Shih is a professor of Resources Engineering at National Cheng Kung University, Taiwan. He held BS and MS of Civil Engineering at National Cheng Kung University and MS and PhD on Industrial Engineering from University of Oklahoma. His major research areas include design for recycling, reverse logistics, sustainable innovation management, and sustainable product and service development.

Rafael Batres is an associate professor of Industrial and Systems Engineering in the Mechanical Engineering Department at Toyohashi University of Technology. He received a BS in chemical engineering from the National University of Mexico (UNAM) and both a MEng and PhD in process systems engineering from Tokyo Institute of Technology. His research is currently focused on systematic methods for ontology development, semantic case-based reasoning, operating procedure synthesis, data mining, and multi-agent systems.