FINAL PAPER

Proceedings of DETC '09
ASME 2009 Design Engineering Technical Conferences
August 30- September 2, 2009, San Diego, CA, USA

DETC2009-87711

A METHOD FOR MEASURING PART SIMILARITY USING ONTOLOGY AND A MULTI-CRITERIA DECISION MAKING METHOD

Duhwan Mun
Wing Ship Technology Corp.
Daejeon, Korea

Junmyon Cho
Electronics and Telecommunications Research Institute
Daejeon, Korea

Karthik Ramani
Purdue University
West Lafayette, IN, USA

ABSTRACT

When existing parts are re-used for the development of a new product or business-to-business transactions, a method for searching parts from a part database that is similar to the user's requirements is necessary. To this end, it is important to develop a part search method that can measure similarity between parts and the user's input data with generality as well as robustness. In this paper, the authors suggest a method for measuring part similarity using ontology and a multi-criteria decision making method and address the technical details of the approach. The proposed method ensures interoperability with existing engineering information management systems, represents part specifications systematically, and has generality in the procedure for measuring part similarity in specifications. A case study for ejector pins conducted to demonstrate the proposed method is also discussed.

1. MOTIVATION AND PROBLEM DEFINITION

Designers spend about 60% of their time searching for the right information during product design [1] and 80% of their design could be created from an existing design or by modifying an existing design [2]. Electronic parts catalogs or digital parts libraries could be basic means to support this. They are the reference for part selection and supplier selection, describing the parts information which corresponds to an expertise on the criteria to select a part and on the condition of part usage [3,4,5,6].

When existing parts are re-used for the development of a new product or business-to-business transactions, a method for searching parts from a part database that are similar to the user's requirements is necessary. The important aspect in a part search method is measuring the similarity between parts and the user's input data with generality as well as robustness. Besides, domain knowledge about manufacturing process, cost and material selection play an important role in the search process [7]. Hence, this part-related knowledge should be reflected in the part search process.

Electronic catalog systems or part management modules of PLM systems provide part searching capabilities with an exact function that measures whether two values are equal or with a range function that measures whether a value is between the given ranges. However, they do not provide various similarity measuring functions that reflect specific comparison characters of part properties, evaluate the degree of similarity for every part property of a part as a whole, and utilize parts' design and manufacturing knowledge in the part similarity measurement.

Previous studies on measuring part similarity could fall into classification system-based similarity measurement and shape-based similarity measurement.

The classification system-based similarity measurement is an approach wherein similarity between parts is determined by comparing parts' data that are described in accordance with a predefined part classification system. The types of classification system could be categorized into code-based, group technology-based, data dictionary-based. Standardized information modeling resources required for the definition of classification system are provided in both ISO 13584 PLIB and ISO 15926 PLIB [8,9,10]. Recently, several researchers recommend utilizing ontologies as the metadata descriptions of part information sources [11,12]. Because ontologies are explicit and formal specifications of the knowledge, especially implicit or hidden knowledge, of information sources they help us with part of the integration problem by disambiguating information items.

After part data is described according to a part classification system, the degree of similarity between part data and user input data is calculated. Various studies on similarity measurement at the level of data model [13,14] have been conducted with the aims of reusing software [15], integration of part classification systems [16], and improving interoperability between information systems [17]. Contrary to those studies, similarity measurement of part data is conducted at the level of data, in other words, instance of data model. For this, Simpson et al. [18] proposed a new method for similarity measurement of manufacturing process data. In this method, properties of candidate manufacturing processes are compared against user input data and similarity of each property is quantified as value in a range of 0 and 1 by using the Gaussian distribution function. Quantified similarity values are then utilized as input data of ELECTRE IS [19], a multi-criteria decision making method, to find the most outranking candidates with respect to similarity with user input data. The ELECTRE IS has the advantage of considering all the properties of manufacturing processes when measuring similarity. This method has a generality in the procedure of similarity measurement and is a very good basis for research in this area. We build on the same basis of similarity measurement and address formalization of data model for parts' specification data, diversity of similarity measuring functions required for the comparison of part properties, priority of orders of part properties on similarity measurement, and representation and utilization of design and manufacturing knowledge in similarity measurement.

A shape-based similarity measurement is a method of measuring similarity using geometric and topological characteristics including overall shape, features, topology, and a convex decomposition tree [20]. In this method, similarity is measured by comparing feature vectors or relational data structures – global features, manufacturing features, graph, histogram, and so on- of two parts with the preprocessing stage of a conversion into a canonical representation, which is invariant to rotation, translation, and scaling [7]. However, this method is not designed to reflect non-geometric properties. It does not utilize part data represented according to a predefined classification system in engineering information management systems. Although it is necessary to verify the user's input data from the viewpoints of design and manufacturing knowledge when searching similar parts, the shape-based similarity measurement does not address this requirement. One can envision however that such shape based system may be combined with our similarity metric based on user requirements of non-geometric aspects.

In this paper, the authors suggest a method for measuring part similarity using ontology and a multi-criteria decision making method and address the technical details of the approach. The proposed method ensures interoperability with existing engineering information management systems, represents part specifications systematically, and has generality in the procedure for measuring part similarity in specifications. A case study for ejector pins conducted to demonstrate the proposed method is also discussed.

2. PROPOSED SOLUTION AND TECHNIQUES IN USE

In this study, part similarity measurement is defined as "*Selecting the most similar parts by comparing user's input data with part data stored according to a predefined part classification system in the database of an information management system*".

In this part similarity measurement problem, the authors utilize ontology and a multi-criteria decision method in order to address the following requirements:

- Formalization and systematization of part specifications;
- Interoperability with existing information management systems;
- Generality and accuracy of comparison procedure.

<u>Formalization and systematization of part specifications</u>

Part specification data can be classified into general data, domain-specific data, and company-specific data, depending on the level of specialization. Company-specific data has low interoperability along with high details. On the other hand, general data has high interoperability along with low details, opposite to the company-specific data. Considering the aforementioned characteristics, a part specification ontology, a data model for describing part specifications, is defined. The part specification ontology consists of upper ontology, domain ontology, and company ontology which are respectively placed at different information model layers in order to reflect that part specification data have various levels of specialization. The upper ontology is relatively static and one can consider the domain and company ontologies to be dynamic. In addition, data modeling resources for representing design and manufacturing knowledge of parts are defined and then linked to part specification data.

<u>Interoperability with existing information management systems</u>

Part specifications are often described according to a predefined classification system in engineering information management systems. For this, a data dictionary is usually employed. The data dictionary consists of part categories, a classification tree defining the hierarchy of the part categories (parent-child relation or specialization relation), and part properties specifying characteristics of parts belonging to each part category [21]. A part specification ontology is defined with the consideration of a part classification system in the form of the data dictionary, thus ensuring interoperability with existing information management systems.

<u>Generality and accuracy of comparison procedure</u>

In order to measure similarity of parts' specifications with user input data, we need solutions in two areas: a method to quantify the degree of similarity of various part properties against user input data; a method to decide the best similar parts, considering all the part properties as a whole. For this, the similarity measuring procedure proposed by Simpson et al. [18] is adopted and tailored for the purpose of part similarity measurement. First, new types of similarity measuring functions are introduced instead of the Gaussian distribution function. Second, part similarity measurement is conducted sequentially according to priority orders in comparison assigned to part properties.

Part data related to various disciplines including design, manufacturing, and transaction is represented by part properties belonging to each part category of a part classification system. Some part properties of a part stored in a database may be similar to the user's input values, but other part properties may not be. Part similarity measurement is therefore a problem of finding the most appropriate parts by evaluating the degree of similarity for every part property of a part as a whole. For this reason, there are multiple criteria in measuring part similarity, and thus a method that can take these criteria into account should be used when conducting the part similarity measurement. In this study, ELECTRE IS, a multi-criteria decision making method, is used for assessing part similarity.

A part property has different characteristics in comparison with another part property since ways to compare a part property with user input value are closely dependent on type of the part property, data type of the part property value, or design and manufacturing knowledge related to the part property. Considering this aspect, parts are compared with user's input values with the use of similarity measuring functions that reflect specific comparison characters of part properties. These measuring functions quantify the similarity after comparing part properties with user's input values.

Part properties may have different priorities in the order of comparison depending on users. Hence, resources for representing comparison priority are added to the part specification ontology and the similarity measurement is conducted sequentially according to priority orders of part properties from part properties commencing with the highest priority and proceeding to those with the lowest priority.

2.1 Upper ontology for part specifications

The part specification ontology is defined with the use of OWL (Web Ontology Language) [22]. OWL is a Web ontology language and one of the Recommendations published by the World Wide Web Consortium (W3C) for the Semantic Web. OWL is designed for use in applications that need to process information content instead of merely presenting information to users. OWL provides three increasingly expressive sublanguages designed for use by specific purposes: OWL Lite, OWL DL, and OWL Full. OWL is based on Extensible Markup Language (XML) [23], which makes it possible to define the schemas of user-defined tags, and the Resource Description Framework (RDF) [24], which enables flexible representation of data.

An OWL ontology consists of individuals, properties, and classes [32]. Individuals represents objects in eh domain that we are interested in. Individuals are also known as instances of classes. Properties, that are binary relations on individuals, link two individuals together. Classes are interpreted as sets that contain individuals. They are described using formal descriptions that state precisely the requirements of membership of the class. Fig. 1 shows a representation of some classes containing individuals. It should be noted that properties link not two classes but two individuals. For this reason, when we want to define relationships between classes and properties, they are presented by specifying domain and ranges of properties or by adding restrictions to properties or classes.

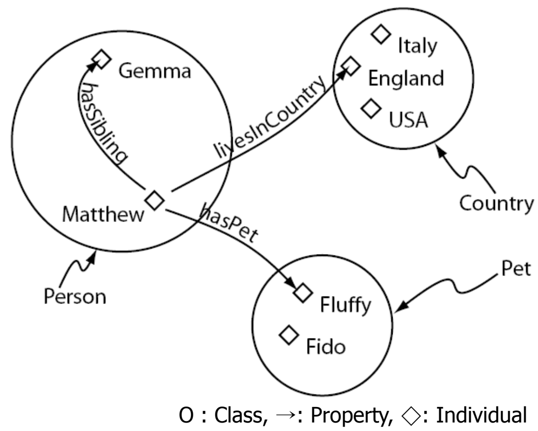

Fig. 1. Components of an OWL ontology [32]

In order to represent part specifications with ontology, a layered ontology definition approach is adopted in this study. After an upper ontology independent of application domains is defined, a domain ontology specific to application domains is constructed. Part knowledge depending on manufacturing companies is defined in the company ontology, referring to the upper ontology and the domain ontology. Among them, constituting the part specification ontology, the upper ontology is addressed in this Section. The domain ontology and the company ontology defined for mold parts in a case study are explained in Section 3.

The upper ontology serves as an information model that provides basic vocabularies and grammar required for the definition of the domain ontology and the company ontology and is necessary to address the following:

- Data dictionary-based classification system: Engineering information management systems such as PDM and ERP usually represent and manage part specifications using a classification system. Hence, the upper ontology should provide information modeling resources for the definition of the part classification system;
- Comparison priority orders of part properties: Comparison orders of part properties vary due to the difference in degree of importance from the various viewpoints such as cost, delivery, and quality. Hence, the upper ontology should provide information modeling resources for specifying the comparison order of each part property;
- Similarity measuring functions for part properties: A variety of similarity measuring functions should be introduced in order to measure the similarity of part properties and the upper ontology should provide resources for identification of the type of similarity measuring function required for each part property;
- Part knowledge: Manufacturing companies have different levels of design methods, production facilities, and worker abilities which could be represented as relations and constraints on part specifications. When part specifications

are described with the data dictionary-based classification system, those relations and constraints are represented as diverse types of relations and constraints between part properties. The upper ontology should provide resources for representing part knowledge of manufacturing companies. Among many types of relations and constraints, exclusive relation, dependent relation, and property precision level are within the scope of this study.

Addressing the aforementioned information requirements, the upper ontology for representation of part specifications is defined as shown in Fig. 2. In the upper ontology diagram, a rectangle with round corners denotes an OWL class and an arrow denotes an OWL property and its domain and range (See the example below).

```
<owl:Class rdf:about="#PartClass">
    <rdfs:subClassOf
    rdf:resource="#ClassAndPropertyElement"/>
</owl:Class>

<owl:ObjectProperty rdf:about="#hasPropertyOf">
    <rdfs:domain rdf:resource="#PartClass"/>
    <rdfs:range rdf:resource="#PartProperty"/>
</owl:ObjectProperty>
```

ClassAndPropertyElement is a superclass of **PartClass** and **PartProperty**. The **PartClass** is a class for representing the type of part and the **PartProperty** is a class for representing a property of the part. An instance of the **PartClass** has several instances of the **PartProperty** with a **hasPropertyOf** OWL property. In the upper ontology, these resources are defined for representing part types, the hierarchy among the part types, and part properties belonging to each part type by referring to the data dictionary information model of ISO 13584 PLIB [25]. **Manufacturer** denotes a company that manufactures parts.

The **PartProperty** has a similarity measuring function (**SimilarityFunctionType** OWL class) and the order of similarity comparison (**itsPriorityOrder** OWL property) as explained in Section 2.1.

The **PartProperty** is classified into **Parameter**, **SubComponent**, and **OptionalManufacturing**. The **PartProperty** is used to define various properties of a part including shape dimensions, functions, business conditions, and manufacturing options. **Parameter** is further categorized into **ShapeDimension**, **Tolerance**, **Transaction**, **Function**, and **Machining**. **SubComponent** is used when a part has sub-components; whether the sub-components are selected or not becomes one of the part properties. **OptionalManufacturing** denotes additional manufacturing operations such as milling and drilling, which are carried out to a part according to the user’s requests. **Parameter** is also used to define properties of the **SubComponent**s and **OptionalManufacturing**.

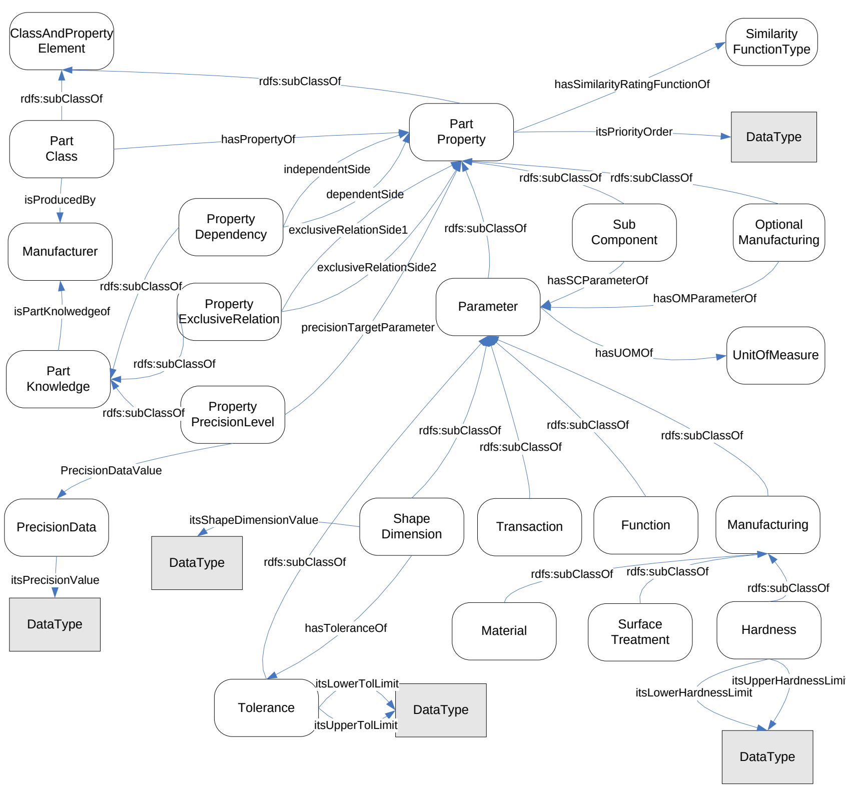

Fig. 2. Upper ontology for the representation of part specifications

ShapeDimension is a class that represents a measurement such as length, width, and height of a part. **Transaction** is used to denote business-related properties such as price and delivery and **Function** is a class for part properties regarding part's capabilities or functions. **Manufacturing** is related to characteristics such as material, surface treatment, and hardness of a part from the viewpoint of manufacturing the part.

PartKnowledge is used in order to specify design and manufacturing knowledge for parts produced by a company and is classified into **PropertyDependency**, **PropertyExclusiveRelation**, and **PropertyPrecisionLevel**. **PropertyDependency** is a class for representing the dependency relation between part properties and the **PropertyExclusiveRelation** denotes the exclusivity between part properties. Considering manufacturing facilities and skills of workers, the allowable precision level of a part property in a company is represented by the **PropertyPrecisionLevel**.

2.2 Similarity measuring functions for part properties

For the comparison between part specifications and the user's input data, the first step is to calculate the degree of similarity of parts property values of a part with the input values. Similarity measuring functions used for this purpose quantify the similarity between a part property value and an input value from the user by a value in a range of 0 to 1.

Similarity measuring functions for part properties have a variety of forms according to their characteristics. The similarity measuring functions adopted in the present study are as follows. Here it is postulated that the user input value is P_i (P_{i_min}, P_{i_max}) and the part property value stored in a database is P_d (P_{d_min}, P_{d_max}).

If part property **P** of part **K** has a single scalar value such as a dimension-related property, the similarity measuring function type used is a uniform distribution given by the following equation.

Uniform distribution function

IF ($P_{d\_min\_in\_database} \leq P_i \leq P_{d\_max\_in\_database}$)

THEN $|P_i\text{-}P_d|/(P_{d\_max\_in\_database} - P_{d\_min\_in\_database})$ ELSE 0

Where If subtype ==“plus range” Then $P_i < P_d$

Else If subtype ==”minus range” Then $P_i > P_d$

In the uniform distribution function, $P_{d\_min\_in\_database}$ is the minimum value of the part property P recorded in the database and $P_{d\_max\_in\_database}$ is the maximum value of the part property P recorded in the database. When a user selects a company as a search criterion, $P_{d\_min\_in\_database}$ and $P_{d\_max\_in\_database}$ become the minimum and maximum values, respectively, of the part property P among parts that the selected company produces and are also stored in the database.

When a part property value is compared with a user input value using the uniform distribution function, the closeness between the two values is usually calculated without considering whether the part property value is greater or smaller than the user input value (this is referred to as a “*plus-minus range”* comparison in this paper). However, design and manufacturing characteristics should be considered in order to precisely measure the similarity of part properties. When a part property P of part K is a dimensional property for shapes formed by machining operations, it is unnecessary or meaningless to conduct similarity measurement for a part property having a value that is smaller than the user input value. In this case, a “plus range” comparison is required, wherein part property values that are equal or greater than the user input value are measured. A “minus range” comparison, opposite to the “plus range” comparison, also exists. Hence, according to the range of comparison on the basis of the input value, there are three subtypes of the uniform distribution function: plus-minus range, plus range, and minus range.

If part property **P** of Part **K** has a string data type or value **a** is the only allowable value for part property **P** due to the manufacturing capability of a company, it is not reasonable to use the uniform distribution function. Instead, the exact function that only measures whether two values are equal is appropriate.

Exact function

IF (P_i==P_d) THEN 1 ELSE 0

If part property **P** of Part **K** is given in the form of a range having minimum and maximum values such as tolerance, similarity is quantified by calculating the overlapped ratio; this is the ratio of the overlapped range between a part property value and a user input value to the range of the user input value.

Overlapped ratio function

overlapped_range/(Pi_max - Pi_min)

2.3 Applying the multi-criteria decision making method

In a part classification system, the number of part properties for each part category typically is more than one. Since part property values usually constitute the data used for measuring similarity of parts with user input data, the number of criteria for deciding the similarity is therefore more than one. When measuring part similarity, there could be a case in which part **A** has higher similarity than part **B** with respect to part property $\mathbf{P_i}$ but part **A** has lower similarity than part **B** with respect to part property $\mathbf{P_j}$. Hence, required here is a method to find the most similar parts with user input data with all the part properties of a part considered as a whole.

The similar part selection problem can be viewed as a multi-criteria decision making problem in which outranking relations among a set of alternatives A={a_1, a_2, …, a_n} are built in terms of the predefined criteria F={g_1, g_2, …, g_m}. In this study, the ELECTRE IS, one of the various ELECTRE methods, is adopted in order to search similar parts. Compared with other decision making methods such as SMART[26] and AHP[27], ELECTRE methods have the advantage of addressing the fuzzy nature in the decision making process by introducing the concept of thresholds. A brief introduction of the ELECTRE IS method is provided in Annex A.

2.4 Procedure for searching similar parts

Priorities of part properties in the comparison order can change since part properties to which users attach importance change depending on the purpose of the part search. The conventional technique of controlling the importance of the part property by applying different weight factors to part properties in multi-criteria decision making methods cannot completely exclude the effects of part properties that have lower priority in the comparison order. The similarity measurement for parts should therefore be carried out sequentially according to comparison priority orders—the **isPriorityOrder** OWL property in the upper ontology for part specifications—of part properties. The procedure for searching similar parts is as follows, where it is postulated that part specifications are presented by using the part specification ontology.

1. A user selects a part type and inputs desired values for part properties. Default priority orders and similarity measuring functions for parts properties are provided by the domain ontology for part specifications. The user can modify the comparison priority order or the similarity measuring function if necessary.
2. With the use of the company ontology for part specifications and rules expressed by Semantic Web Rule Language (SWRL) [28], input values from the user are verified and errors are reported if they exist.
3. Start the similarity measurement with part properties having the highest priority order—part properties of which **isPriorityOrder** value in the domain ontology is one—and determine a set of candidate parts
   A. With the use of similarity measuring functions, the similarity between a part property value and a user input value is quantified as a value in a range of zero to one.
   B. Based on those similarity values for part properties, apply the ELECTRE IS method in order to build outranking relations on parts stored in a database. Select parts that are at least as good as all other parts with respect to similarity with the input values and designate them as a set of candidate parts.
4. Conduct a similarity measurement for part properties with a priority order one lower than the priority order in the previous step and update the set of candidate parts. The similarity measuring process is the same as step 3.
5. Repeat step 4 until part properties with the lowest priority order are reached.
6. Return the set of candidate parts as a result.

3. CASE STUDY – MOLD PARTS (EJECTOR PINS)

Ejector pins, a mold part, are selected as test parts in the case study for demonstrating the proposed part similarity measuring method. Design and manufacturing knowledge regarding ejector pins gathered from paper catalogs or through interviews with technical experts of manufacturing companies in a previous study [29] is utilized for this case study.

3.1 Ejector pins

Ejector pins, as shown in Fig. 3, can be categorized into three types: straight, stepped, and rectangular. Ejector pins have the following part properties:

- Parameter
  - Diameter (P), length (L), head thickness (T), head diameter (H), shoulder diameter (D), shoulder length (N), diameter tolerance, length tolerance, head thickness tolerance, head diameter tolerance, material, surface treatment, hardness
- Sub component
  - None in this category
- Optional manufacturing
  - Head area: Head cutting, tapping, knock pin hole machining, numbering
  - Shaft area: Gas vent machining
  - Tip area: Tip shaping, finishing

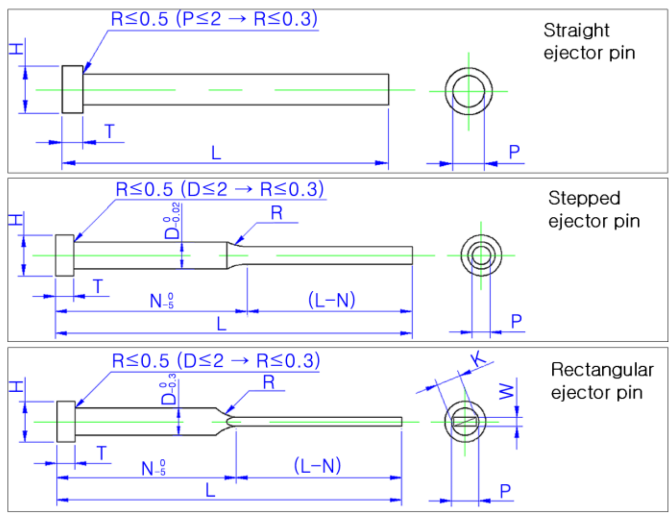

Fig. 3. Types of ejector pins and their part properties

Rectangular ejector pins have width (P) and height (W) instead of diameter (P). Stepped and rectangular ejector pins have additional part properties of shoulder diameter (D) and shoulder length (L).

3.2 Ontology for ejector pin specifications

The ontology for ejector pin specifications consists of the domain ontology and the company ontology.

3.2.1 Domain ontology for ejector pins

The domain ontology specifies a data model (data dictionary-based classification system) for the description of part specifications by referring to the upper ontology. The domain ontology differs for part categories. In this study, the domain ontology for ejector pins has been defined; however, the domain ontology for other disciplines, i.e. mechanical switches, can also be defined by referring to the upper ontology. For the part similarity measurement, this domain ontology at least is necessary. The part properties of ejector pins explained in Section 3.1 are represented in the form of a classification system in the domain ontology for ejector pins.

Part of the domain ontology for ejector pins is diagrammed in Fig. 4. In the domain ontology diagram, an arrow represents an existential restriction (See the example below).

```
<owl:Class rdf:about="#StraightEjectorPin">
    <rdfs:subClassOf rdf:resource="#EjectorPin"/>
    <rdfs:subClassOf>
      <owl:Restriction>
        <owl:onProperty rdf:resource="#hasPropertyOf"/>
        <owl:someValuesFrom
             rdf:resource="#EjectorPinLength"/>
        </owl:Restriction>
 :
 </owl:Class>
```

Since ejector pins selected in this case study is a mold part in which shape dimensions and manufacturing characteristics are important, function of the upper ontology is not used. However, for other parts, many part properties are defined as subtypes of function. For example, power switch has part properties of rated voltage and current which are subtypes of function.

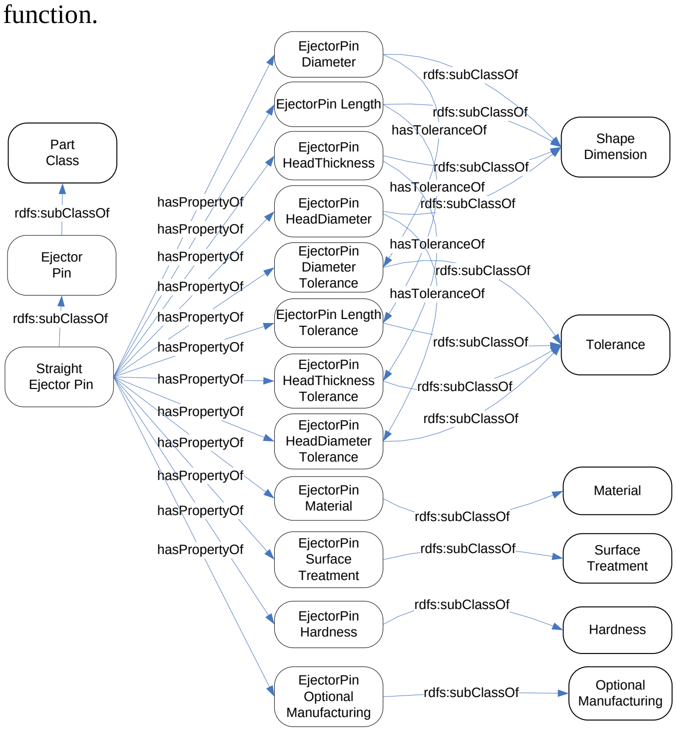

Fig. 4. Domain ontology for part specifications – ejector pin

3.2.2 Company ontology and rules for ejector pins

The company ontology defines design and manufacturing knowledge on specific part types specific for each manufacturing company by referring the upper ontology and the domain ontology relevant to the manufacturing company. The company ontology is usually represented in the form of relations between part properties, and differs for parts suppliers. The company ontology is not mandatory but is useful for checking user's input values or validating integrity of data stored in parts database. In the case study, the company ontology for Company A that produces ejector pins is defined. For the definition, the upper ontology and the domain ontology for ejector pins are referred to. For instance, Company A, which manufactures ejector pins, may have dependency relations between ejector pin properties as shown in Table 1 as part of its knowledge. **Material** and **Surface treatment** part properties of ejector pins affect machining precision of shape dimension rather than shape dimension value itself on manufacturing ejector pins. Therefore **Ejector pin length tolerance**, **Ejector pin head thickness tolerance**, and **Ejector pin head diameter tolerance** are dependent on **Material** and **Surface treatment** whereas **Ejector pin length**, **Ejector pin head thickness**, and **Ejector pin head diameter** are not.

Table 1. Dependency relations for straight ejector pin properties

Independent property		Independent property		
Material	Surface treatment	Ejector pin length tolerance	Ejector pin head thickness tolerance	Ejector pin head diameter tolerance
SACM645	chromium plating	+0.1~+5	-0.05~0	-0.3~0
SKD61	nitriding	+0.1~+5	**-0.02~0**	-0.3~0
SKH51	No	+0.1~+5	-0.02~0	-0.3~0

The company ontology that represents the knowledge on the straight ejector pin, i.e., dependency relations, of Company A is defined with the use of the upper ontology and the domain ontology, as shown in Fig. 5. In the company ontology diagram, a dotted arrow is a hasValue restriction while a solid arrow denotes an existential restriction.

PD_EPLT_CompanyA represents the dependency relation among **Ejector pin length tolerance**, **Material**, and **Surface treatment**. **PD_EPHTT_CompanyA** represents the dependency relation among **Ejector pin head thickness tolerance**, **Material**, and **Surface treatment**. **PD_EPHDT_CompanyA** represents the dependency relation among **Ejector pin head diameter tolerance**, **Material**, and **Surface treatment**.

However, as shown in Fig. 5, the OWL has a limitation in specifying details pertaining to the company's knowledge precisely – by equations or conditions – in the company ontology. Hence, the company ontology only defines existences of exclusive relations, dependency relations, and property precision levels and details regarding these relations are expressed as rules with the use of SWRL.

SWRL is a rule language for semantic web recommended by W3C and combines the OWL DL and OWL Lite sublanguages with the Unary/Binary Datalog RuleML sublanguages. The basic syntax of SWRL has the following form:

Antecedent → *consequent*

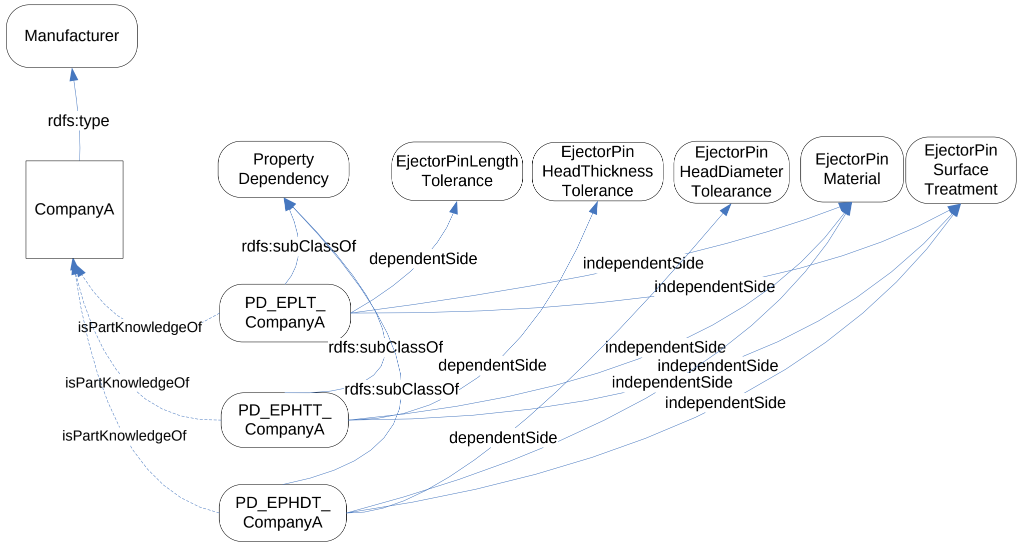

Fig. 5. Company ontology – straight ejector pin produced by Company A

This syntax means that the conditions specified in the consequent must hold whenever the conditions specified in the antecedent are satisfied.

Both the antecedent and consequent consist of more than one atom. Forms of atoms can be OWL class C(x), OWL property P(x,y), sameAs(x,y), differentFrom(x,y), and SWRL built-in functions, where x and y are OWL individuals or OWL data values.

The following is a SWRL rule to retrieve user input data that violates the knowledge listed in List 1 – "If a straight ejector pin manufactured by Company A employs the material SKD61 and has undergone nitriding as a surface treatment, the head thickness tolerance of the straight ejector pin should be -0.02~0.".

List 1. SWRL rule to validate user input data

StraightEjectorPin(?part) ∧
isProducedBy(?part, CompanyA) ∧
hasPropertyOf(?part, ?EPM) ∧
Material(?EPM) ∧
itsMaterial(?EPM, "SKD61") ∧
hasPropertyOf(?part, ?EPST) ∧
SurfaceTreatment(?EPST) ∧
itsSurfaceTreatment(?EPST, "Nitriding") ∧
hasPropertyOf(?part, ?p1) ∧
EjectorPinHeadThicknessTolerance(?p1) ∧
itsLowerTolLimit(?p1, ?l1) ∧
swrlb:notEqual(?l1, -0.02)
→ query:select(?part, ?p1, ?l1, ?l2)

StraightEjectorPin(?part) ∧
isProducedBy(?part, CompanyA) ∧
hasPropertyOf(?part, ?EPM) ∧
Material(?EPM) ∧
itsMaterial(?EPM, "SKD61") ∧
hasPropertyOf(?part, ?EPST) ∧
SurfaceTreatment(?EPST) ∧
itsSurfaceTreatment(?EPST, "Nitriding") ∧
hasPropertyOf(?part, ?p1) ∧
EjectorPinHeadThicknessTolerance(?p1) ∧
itsUpperTolLimit(?p1, ?l2) ∧
swrlb:notEqual(?l2, 0)
→ query:select(?part, ?p1, ?l1, ?l2)

In this study, part knowledge is defined at the company ontology level since mold part data was gathered at the company level. However, if there is standard part knowledge commonly applicable to part type regardless of part manufacturers, part knowledge can be represented at the domain level. This knowledge is usable even if a user does not specify the name of the company.

Part knowledge is used to verify user's input values. Therefore, it is possible to measure similarity of parts against user inputs even if there is no part knowledge available. However, accuracy of similarity measurement results could drop when part data stored in a database has errors.

3.3 Measuring similarity of ejector pins

Experiments for measuring the similarity of straight ejector pins manufactured by Company A were conducted. Four straight ejector pins were selected, and their specifications were then described with the domain ontology. User input data was verified with the use of the company ontology and the SWRL rules. For each part property of the straight ejector pin, similarity with the corresponding user input value was quantified by the similarity measuring function defined for that part property. Finally, ELECTRE IS was applied sequentially according to comparison priority orders of part properties of the straight ejector pin and the most outranking candidates among parts tested are selected.

Similarity measuring functions for part properties of the straight ejector pin and their priority orders in comparison are listed in Table 2. Three types of similarity measuring functions were used in these experiments: the exact, the uniform distribution, and the overlapped ratio functions. The exact function is used for comparing a property that has a string type value, the overlapped ratio function for a property with maximum and minimum values such as tolerance, and the uniform distribution function for a property that has a scalar value. When a part property has a uniform distribution function as its similarity measuring function, the values inside [] in Table 2 indicate the minimum and maximum values of the part property for straight ejector pins manufactured by Company A.

Table 2. Similarity measuring functions for part properties of straight ejector pins and their priority orders in comparison

	Similarity measuring function	Priority order in comparison
Diameter	Uniform Distribution Function: Minus Range [0.3-12]	1
Diameter Tolerance	Overlapped Ratio Function	1
Length	Uniform Distribution Function: Plus Range [50-350]	1
Length Tolerance	Overlapped Ratio Function	1
Head Diameter	Uniform Distribution Function: Plus Range [3-17]	2
Head Diameter Tolerance	Overlapped Ratio Function	2
Head Thickness	Uniform Distribution Function: Plus Range [4-8]	2
Head Thickness Tolerance	Overlapped Ratio Function	2
Hardness	Overlapped Ratio Function	2
Material	Exact Function	2
Surface Treatment	Exact Function	2

Table 3. Similarities of part property of the test parts against user input data

	TP00000001		TP00000002		TP00000003		TP00000004		Modified Input Value	Initial Input Value
	Property Value	Property Similarity	Property Value	Property Similarity	Property Value	Property Similarity	Property Value	Property Similarity		
Diameter	1.5	0.812	3.5	0.983	3.8	0	3.5	0.983	3.7	3.7
Diameter Tolerance	-0.005~0	1	-0.005~0	1	-0.005~0	1	-0.02~-0.01	0	-0.005~0	-0.005~0
Length	250	0.5	150	0.883	60	0	60	0	100	100
Length Tolerance	0.1~5	1	0.1~5	1	0.1~5	1	0.1~5	1	0.1~5	0.1~5
Head Diameter	7	1	9	0.857	7	1	7	1	7	7
Head Diameter Tolerance	-0.3~0	1	-0.3~0	1	-0.3~0	1	-0.3~0	1	**-0.3~0**	-0.4~0
Head Thickness	4	1	6	0.5	6	0.5	4	1	4	4
Head Thickness Tolerance	-0.02~0	1	-0.02~0	1	-0.02~0	1	-0.02~0	1	-0.02~0	-0.02~0
Hardness	58~60	1	56~59	0.5	58~60	1	58~60	1	56~59	56~59
Material	SKH51	1	SKH51	1	SKH51	1	SKH51	1	SKH51	SKH51
Surface Treatment	No	1	No	1	No	1	No	1	No	No

From the part property values of the test parts TP00000001, TP00000002, TP00000003, and TP00000004, similarities of part properties with the user input data were quantified with the use of similarity measuring functions, and the results are listed in Table 3.

The initial input data listed in the right side of Table 3 was verified with the use of the company ontology and the SWRL rules ,and it was found that a head diameter tolerance value "-0.4~0" violates the part knowledge of Company A. Hence, the head diameter tolerance value was changed to "-0.3~0".

When similarities for part properties of the test parts are measured, outranking relations aSb among the test parts were calculated according to the ELECTRE IS method. The parameters -p, q, v, and s- required for calculation of the outranking relations aSb were set as 0.1, 0, 0.3, and 0.5 respectively.

When comparison priority orders of part properties are not considered, outranking relations aSb for the test parts are calculated as shown in Table 4. In one outranking relation aSb, "a" denotes a part in column one in Table 4 and "b" indicates a part in row one in Table 4. If a part "b" has an aSb value as zero for all other parts—in other words, if for a part in row one in Table 4 all aSb values are zero in the column direction—part "b" becomes one of the most outranking parts. As shown in Table 4, TP00000001 and TP00000002 are found as the most outranking parts and returned.

Table 4. Outranking relations aSb for the test parts – without considering comparison priority orders

	TP00000001	TP00000002	TP00000003	TP00000004
TP00000001		**0**	1	1
TP00000002	**0**		0	0
TP00000003	**0**	**0**		0
TP00000004	**0**	**0**	0	

However, when parts properties of Diameter, Diameter tolerance, Length, and Length tolerance, which have the highest priority order, are only considered initially, outranking relations aSb for the test parts are calculated as listed in Table 5. As shown in Table 5, only TP00000002 among the test parts is not outranked by the other parts and thus is returned as the most outranking part. Since TP00000001 is outranked by TP00000002 if only Diameter, Diameter tolerance, Length, and Length tolerance are considered, TP00000001 cannot be selected.

Table 5. Outranking relations aSb for the test parts – with consideration of comparison priority orders

	TP00000001	TP00000002	TP00000003	TP00000004
TP00000001		**0**	1	1
TP00000002	1		1	1
TP00000003	0	**0**		0
TP00000004	0	**0**	0	

As can be seen from the outranking relations built in Table 4 and Table 5, differentiating priority orders in comparison for part properties makes it possible to conduct a comparison of parts while reflecting the user's intention in searching parts.

4. CONCLUSIONS AND FUTURE WORK

In this study, a part similarity measuring method using ontology and a multi-criteria decision making method is proposed. This method has the following technical features:

- Representation of part specifications with the use of a layered ontology consisting of an upper ontology, domain ontology, and company ontology;
- Expression of part knowledge by ontology and SWRL rules and verification of input data from the user through their use;
- Defining various similarity measuring function reflecting characteristic of each part property ;
- Differentiation of priority orders in comparison for part properties.

The proposed method ensures interoperability with existing engineering information management systems, enables structuralized representation of part specifications, and provides a generalized procedure for comparing part specifications.

Future work will be focused in three parts: 1) testing the proposed method in various scenarios beyond the ejector-pin case in order to prove generality of the proposed method, 2) data mining [33,34] for the automatic extraction of part knowledge from existing data sources, and 3) developing an additional layer such as the Engineering Oriented Search [30,31], in order to make users familiar with the proposed similarity measurement technique, where users are intuitively given some suggestions on and educated of the layered ontology and main parameters used in the ELECTRE-IS.

REFERENCES

[1] W. Leizerowicz, J. Lin, M.S. Fox, “Collaborative design using WWW”, Proceedings of the WET-ICE’96, CERC, University of West Virginia 1996.

[2] T.G. Gunn, “The mechanization of design and manufacturing”, Scientific American, Vol. 247, No. 3, 1982.

[3] J.P. Baron, M.J. Shaw, A.D. Bailey, “Web-based e-catalog systems in B2B procurement”, Communications of the ACM, Vol. 43, No. 5, 2000.

[4] E. Satdet, G. Pierra, H. Murayama, Y. Ait-Ameur, “Simplified representation of parts library: Model, practice and implementation”, Proceedings of the 10th Symposium on Product Data Technology Europe, Brussels, 2001.

[5] Z. Cui, J.W. Shepherdson, Y. Li, “An ontology-based approach to eCatalogue management”, BT Technology Journal, Vol. 21, No. 4, 2003.

[6] S. Handschuh, B.F. Schmid, K. Stanoevska-Slabeva, “The concept of a mediating electronic product catalog”, EM-Electronic Markets, Vol. 7, No. 3, 1997.

[7] N. Iyer, S. Jayanti, K. Lou, Y. Kalyanaraman, K. Ramani, "Three-dimensional shape searching: state-of-the-art review and future trends", Computer-Aided Design, Vol. 37, No. 5, 2005.

[8] G. Radack, “Report on The Task Force on Dictionaries, Part Libraries and Reference Data Libraries”, ISO TC184/SC4 N2016, 2005.

[9] G. Pierra, J.C. Potier, E. Sardet, “From digital libraries to electronic catalogues for engineering and manufacturing”, International Journal of Computer Applications in Technology, Vol. 18, No. 1-4, 2003.

[10] D. Leal, “ISO 15926 Life cycle data for process plant: An overview", Oil & Gas Science and Technology - Rev. IFP, Vol. 60, No. 4, 2005.

[11] F. Hakimpour, A. Geppert, “Resolving semantic heterogeneity in schema integration: an ontology based approach”, Proceedings of the International Conference on Formal Ontology in Information Systems, Ogunquit, Maine, USA, ACM Press, 2001.

[12] H. Wache, T. Vogele, U. Visser, H. Stuckenschmidt, G. Schuster, H. Neumann, S. Hubner, “Ontology-based integration of information - A survey of existing approaches”, IJCAI-01 Workshop: Ontologies and Information Sharing, Seattle, WA, 2001.

[13] A. Rodriguez, M. Egenhofer, “Determining Semantic Similarity among Entity Classes from Different Ontologies”, IEEE Transactions on knowledge & data engineering, Vol.15, No.2, 2003.

[14] R. Mili, J. Desharnais, M. Frappier, A. Mili, “Semantic distance between specifications”, Theoretical Computer Science, Vol. 247, 2000.

[15] R. Prieto-Diaz, P. Freeman, “Classifying software for reusability”, IEEE Software, Vol. 4, No. 1, 1987.

[16] J. Cho, S. Han, H. Kim, “Meta-ontology for automated information integration of parts libraries”, Computer-Aided Design, Vol. 38, No. 7, 2006.

[17] A. Sheth, " Changing Focus on Interoperability in Information Systems: From System, Syntax, Structure to Semantics”, in M. Goodchild, M. Egenhofer, R. Fegeas, C. Kottman (eds.), Interoperating Geographic Information Systems, Kluwer Academic Publishers, 1999.

[18] F. Alizon, S.B. Shooter, T.W. Simpson, “Reuse of manufacturing knowledge to facilitate platform-based

product realization", Journal of Computing and Information Science in Engineering, Vol. 6, No. 2, 2006.

[19] B. Roy, "The Outranking Approach and the Foundations of ELECTRE Methods", Theory and Decision, Vol. 31, 1991.

[20] T. Hong, K. Lee, S. Kim, "Similarity comparison of mechanical parts to reuse existing designs", Computer-Aided Design, Vol. 38, No. 9, 2006.

[21] D. Mun, H. Kim, K. Jang, J. Cho, J. Kim, S. Han, "A Method for Automatic Generation of Parametric CAD Models in a Mold Base e-Catalog System", Journal of computing and information science in engineering, Transactions of the ASME, Vol. 6, No. 3, 2006.

[22] W3C OWL Web site, http://www.w3.org/2004/OWL/, 2007.

[23] Extensible Markup Language (XML) 1.0 (Fourth Edition), W3C Recommendation, http://www.w3.org/TR/2006/REC-xml-20060816/, 2006.

[24] Resource Description Framework (RDF): Concepts and Abstract Syntax, W3C Recommendation, http://www.w3.org/TR/rdf-concepts/, 2004.

[25] ISO, Industrial automation systems and integration – Parts Library – Part 42: Methodology for Structuring Parts Families, ISO 13584-42, 1998.

[26] W. Edwards, "How to Use Multiattribute Utility Measurement for Social Decision Making", IEEE Transactions on Systems, Man and Cybernetics, Vol. 7, 1977.

[27] T.L. Saaty, "The Analytic Hierarchy Process", McGraw-Hill, 1980.

[28] A Semantic Web Rule Language Combining OWL and RuleML, W3C Member Submission, http://www.w3.org/Submission/SWRL/, 2004.

[29] D. Mun, K. Jang, S. Han, J. Kim, H. Hwang, "Engineered-to-order Approach for Providing Flexibility in e-Commerce of Mold Parts", Concurrent Engineering: Research and Applications, Vol. 15, No. 4, 2007.

[30] Z. Li, K. Ramani, "Ontology-based Design Information Extraction and Retrieval", Journal of Artificial Intelligence for Engineering Design, Analysis and Manufacturing, Special issue: Computational Linguistics for Design, Maintenance and Manufacturing, Vol. 20, No. 2, 2007.

[31] Z. Li, M.C. Yang, K. Ramani, A methodology for engineering ontology acquisition and validation, Journal of Artificial Intelligence for Engineering Design, Analysis and Manufacturing , Vol. 23, Special Issue 01, 2009.

[32] Matthew Horridge, Holger Knublauch, Alan Rector, Robert Stevens, Chris Wroe, "A Practical Guide to Building OWL Ontology Using The Protégé-OWL Plugin and CO-ODE Tools", The University of Manchester, 2004.

[33] Andrew Kusiak, Mattew Smith, "Data mining in design of products and production systems", Annual Reviews in Control, Vol. 31, No. 1, pp. 147-156, 2007.

[34] Witten, I.H., Frank, E., Data mining: Practical machine learning tools and techniques. New York: Elsevier, 2005.

ANNEX A

ELECTRE IS METHOD

In multi-criteria decision making problems, a set of N alternatives A={a_1, a_2, …, a_n} and a set of M criteria F={g_1, g_2, …, g_m} are given. While the evaluated value of an alternative $a_i \in A$ over a criterion $g_k \in F$ is expressed as $g_k(a_i)$, the evaluated values of an alternative $a_j \in A$ over the multi-criteria F is expressed as $g(a_j) = \{g_1(a_j), g_2(a_j),\ldots, g_m(a_j)\}$.

The ELECRE method is based on the outranking relation (aSb), meaning that "a is at least as good as b with respect to given criteria." The ELECTRE method introduces the concept of thresholds in order to take account hesitations of a decision maker when he or she builds an outranking relation between two alternatives. The thresholds includes the indifference threshold q_k, preference threshold p_k , and veto threshold v_k. The binary relation aI_kb, which is determined by the indifference threshold q_k, means that a is indifferent to b for a given criterion g_k. The binary relation aP_kb, which is related to preference threshold p_k, means that a is strongly preferred to b. Finally, aQ_kb which resides between indifference and preference thresholds, means that a is weakly preferred to b.

aP_kb: $g_k(a)- g_k(a) \geq p_k$
aQ_kb: $q_k \leq g_k(a)- g_k(a) \leq p_k$
aI_kb: $- q_k \leq g_k(a)- g_k(a) \leq q_k$

The outranking relation aS_kb holds in the case of aP_kb, aQ_kb, or aI_kb. On the other hand, if $g_j(b)-g_j(a)>v_j$, the assertion aSb is negated even if aS_kb ($j\neq k$) is built over all other criteria.

The ELECTRE IS starts with the setting of required parameters – the weight factor w_k, indifference threshold q_k, preference threshold p_k , and veto threshold v_k. After the parameters are determined, a concordance index $C(a_i, a_j)$ and discordance index $D(a_i, a_j)$ between two alternatives $a_i, a_j \in A$ are calculated in order to build the outranking relations. Finally, a set of the most outranking alternatives are selected from the outranking relations among alternatives.

The concordance index is defined as follows:

$$c_k(a_i, a_j) = \begin{cases} 1, & a_i S_k a_j \\ \frac{g_k(a_i)-g_k(a_j)+p_k}{p_k-q_k}, & a_j Q_k a_i \\ 0, & a_j P_k a_i \end{cases} \quad (1)$$

$$C(a_i, a_j) = \sum_{k=1}^{m} w_k c_k(a_i, a_j) \quad (2)$$

The discordance index is defined as follows:

$$d_k(a_i, a_j) = \begin{cases} 0, \text{if } g_k(a_i) - g_k(a_j) > q_k \frac{1-C(a_i,a_j)-w_k}{1-s-w_k} - v_k \\ 1, \text{otherwise} \end{cases} \quad (3)$$

$$D = (a_i, a_j) = \begin{cases} 0, d_k(a_i, a_j) = 0, \forall\, i, j, k, i \neq k \\ 1, \text{ otherwise} \end{cases} \quad (4)$$

In the ELECTRE IS method, the outranking relation is built using the following equation (5).

$$a_iSa_j = \begin{cases} 1, & C(a_i, a_j) \geq s \text{ and } D(a_i, a_j) = 0 \\ & 0, \text{otherwise} \end{cases} \qquad (5)$$
$$\text{where } 0.5 \leq s < s^*, s^* = 1 - \min_{K \in F} w_k$$

The minimum value 0.5 of the concordance level s means that a sufficiently high majority of criteria – at least half of them – should be in favor of the assertion a_iSa_j. On the other hand, setting s greater than s^* means that a_iSa_j holds only if $a_iS_ka_j$ is imposed for all criteria $g_k \in F$.

Finally, the most outranking alternative a_j is defined as "an alternative that is not outranked by any other alternatives a_i ($i \neq j$)" and can be expressed as the following equation (6).

$$\sum_{i=1, i \neq j}^{n} a_iSa_j = 0 \qquad (6)$$