Attribute Intelligence in Product Ontology

Rouzbeh Meymandpour
M.Sc. Student in Information Technology Engineering and Electronic Commerce, Shiraz University, Shiraz, Iran
dr.rouzbeh@yahoo.com

Mohammad Hadi Sadreddini
Assistant Professor, Department of Computer Science and Engineering, School of Engineering, Shiraz University, Shiraz, Iran
sadredin@shirazu.ac.ir

Abstract

The product information is an essential component of e-commerce collaborations, especially in large-scale enterprise transactions. Clearly defined product information is a necessary foundation for effective collaborative business. Furthermore, the use of ontologies, as a method for formal description of concepts, enables semantically enriched product information, which can enhance the quality and effectiveness of business communications. A product ontology system contains all related information about the products and everything that may describe them. Product attributes, also known as properties, are the most important feature in the product information management, which characterize and describe the product. In this paper, we present methods to describe the products by including more detailed information about product attributes. In order to add more intelligence to the product ontology system, we study different types of attributes and different information that may be related to them, such as data type of the attribute, its unit of measurement, value recommendation, and attribute inheritance. We present the system by operational models, which can be used as a reference model in product ontology design.

1. Introduction

Product information is an essential component in e-commerce. A clearly defined products and services is a necessary foundation for collaborative business processes. Sharing an unambiguously product information by using ontologies enables rich semantics and a high-level of interoperability for e-business systems. With the increasing importance of product information in today's e-commerce, especially in B2B (business-to-business) environments and large-scale business collaboration, it is critical to benefit the power of ontologies in the product information management. Ontologies allow computer understandable product information. It provides a rich description of products and services in addition to intelligence search and using semantic web technologies to add reasoning capability to the management system [1; 2].

Ontologies make sense only if we clearly distinguish things, related things, parts and component of those things, documents describing those things, and similar objects that are held together mainly by being somehow related to a joint topic [3]. Correspondingly, a product ontology system contains all related information about the product, everything that may distinguish it and help the system to describe it, including its features, attributes or properties, its components, related accessories, etc.

Product attributes, also known as properties, are the most important feature in the product information management, which characterize and describe the product. For example, 'Weight', 'Display Size', 'Supported Networks', and 'Internal Memory Capacity' are the features and attributes of the product 'Cell Phone'. Therefore, in order to have a semantically enriched product ontology system, we need to clearly distinguish products, their attributes, values of attributes, and everything related to them.

In this paper, we present methods to describe the products by including more detailed information about product attributes. To add more intelligence to the product ontology system, we study different types of attributes and different information that may be related to them, such as data type of the attribute, its unit of measurement, value recommendation, and attribute inheritance. We present the system by operational models, which can be used as a reference model in product ontology design.

2. Background

2.1. Ontology

Ontology is the philosophical discipline and is the study of what there (possibly) is describing the nature and structure of reality, domain of entities, categories, and relations, and characterizing its properties. However, an ontology is a theoretical or computational artifact which expresses the intended meaning of a vocabulary, in terms of the nature and structure of the entities it refers to [4]. Reference [5] cited that the term ontology was used inconsistently. They mention seven different notions of the term:

1. "Ontology as a philosophical discipline
2. Ontology as a an informal conceptual system
3. Ontology as a formal semantic account
4. Ontology as a specification of a conceptualization
5. Ontology as a representation of a conceptual system via a logical theory
   5.1 Characterized by specific formal properties
   5.2 Characterized only by its specific purposes
6. Ontology as the vocabulary used by a logical theory
7. Ontology as a (meta-level) specification of a logical theory"

The [6]'s definition of the ontology, is the most famous definition. It defines ontology simply as "an explicit specification of a conceptualization." Gruber notes, "A conceptualization is an abstract, simplified view[s] of the world that we wish to represent for some purpose" [6]. Another definition of the ontology is "a logical theory which gives an explicit, partial account of a conceptualization" [5].

Concepts can represent concrete entities (books, cameras, toads, clouds), abstract notions (fictional places, ideas, theories), beliefs, processes, tasks, goals, events, states, or methods—in short, anything that needs to be modelled in the knowledge domain (the universe of discourse). Entities can further be specified to make explicit some chosen characteristics or attributes, such as colour, size, price, manufacturer, location, name, and the like. Reference [7] summarized the uses of ontologies as follows (p. 40):

1. "For communication
2. Between implemented computational systems
3. Between humans
4. Between humans and implemented computational systems
5. For computational inference
6. For internally representing plans and manipulating plans and planning information
7. For analyzing the internal structures, algorithms, inputs and outputs of implemented systems in theoretical and conceptual terms
8. For reuse (and organization) of knowledge
9. For structuring or organizing libraries or repositories of plans and planning and domain information"

Ontologies play an important role in the development of large-scale, computer mediated, and enterprise wide knowledge management projects. It enables computer understandable concepts and allows the automated sharing and reuse of essential explicit knowledge.

3. Modeling Product Attributes

The more detailed product descriptions, the more intelligent product ontology system. Product attributes is the most important part of classification schema, which enables including more information about products. There are two approaches to model the attributes of products. The first method is to define a set of attributes for each product in classification schema. The second approach is to define a product-independent attribute collection and create attribution relationships when needed. To minimize the number of attributes [8] and to prevent the growth of redundant data, we use the second approach and design an independent attribute collection, called [Attributes].

In addition to identification number, name, and description, the attributes class can also contains other important features. A mandatory Boolean field, [att_IsMandatory], is required to show whether it is an optional property or not. It is helpful for reducing the number of essential properties that have to be set for product description [8].

To avoid a huge list of properties and to divide them into the more relevant lists, reference [8] recommends using a list of property groups, [Property_Groups], which includes a set of attributes in each group. An ID, a name, and a description characterize each group. This approach enables better management and prevents the flat structure of large lists of attributes.

4. Identifying Different Types and Aspects of Attributes

To increase the intelligence of the product ontology system, we need to distinguish different kinds of attributes and include specific details about them. Reference [9] identified four types of attributes:

1. Valid for a specific product category
2. Vendor- or model- specific
3. Valid for single product instances

4. Dynamic vs. static

The value of an attribute, which will be assigned to it as a product feature, is one of the most important information that should be stored effectively in the product ontology system. Different types of attributes can be distinguished by two aspects of their values: (1) the value assignment, i.e. how values will be assigned to the attributes, and (2) the nature of the values itself. In this section, we study literal vs. selectable attributes, vendor- specific attributes, and single- vs. multi- value attributes from the first aspect and data type of attribute values and their unit of measurement from the second one.

An attribute may be characterized by more than one value. As an instance, 'Mobile Networks' property of the 'Cell Phone' product is a multi-value attribute. A typical instance of the 'Cell Phone' product, for example 'Nokia N96', supports 'GSM 850, 900, 1800, and 1900' in 2G networks. The Boolean [att_IsMultiValue] field represents this feature in our model.

Another aspect of attributes is how values may be assigned to them. Some values are selectable from a predefined list of values while others are not collectable in a list, i.e. literal attribute values. For example, simply think about difference between 'Weight' of a mobile device and its 'Manufacturer' that produce the cell phone. The Boolean [att_IsSelectable] item expresses it. In addition, we need a list of recommended values for each attribute to permit the selection of values; named [Recommended_Values].

Another issue is that different vendors or retailers may offer a single product with different special features. For instance, 'Vendor 1' may sell the 'Dell 1530' notebook with a 'One year' 'Warranty' at '1250 EUR' while 'Vendor 2' may sell the same with a 'Three years' 'Warranty' at '1400 EUR'. To provide the ontology system with the facility of storing products with vendor-specific features we need special kind of attributes called *Vendor-Specific Attributes*. Therefore, the attribute 'Warranty' is a vendor-specific attribute. To represent these kinds of product attributes, we add a [att_IsVendorSpecific] Boolean field to the [Attributes] class. In the implementation phase, we can add number of tables to hold the values that are assigned to different types of attributes.

To allow even more intelligence about values, attributes, and accordingly products in the system, we add two extra capabilities to the model: *Data Type* and *Unit of Measurement (UOM)*. Data type indicates the type of information, which may be assigned to an attribute. Common data types include *Boolean*, *Integer*, *String*, *Float*, and *Currency*. An attribute-independent list, [Data_Types], contains a list of data types. Identification of data type of attributes enable advance searching and browsing in product ontology system. For instance, searching for 'all mobile phones which their weight is under 100gr' will be possible, sufficiently. It also permits the dynamic and statistical grouping of values based on their common appearance, 'lower than 80gr', 'between 80gr and 120gr', and 'higher than 120gr', for example. These beneficial advantages enhance the quality of the e-catalog system.

Storing units of measurement is another capability to boost the quality of the ontology by gathering more information about the products. Frequently, attributes have a unit of measurement, which implies how the associated attribute is measured and what its scale is. Examples of UOMs are Kilogram (kg), Pound (lb), Inch, Millimeter (mm), Second (s), Megabyte (MB), etc. The range of measurement units is limited and described in standards such as *The International System of Units (SI)* [10] and some of *ISO (International Standards Organization)* standards, including ISO 31, 1000, and 80000 [11]. Hence, we create a separated list for UOMs, titled [UOMs], with N:N (N-to-N, many to many) relationship to [Attributes], and add an ID, [uom_ID], standard name, [uom_Name], a description, [uom_Description], and its standard symbol, [uom_Symbol], fields.

Including the UOMs requires suitable considerations (see [9]). One aspect is that related measures can be converted to each other. For instance, the weight of a 'Dell 1530' 'Notebook' is 2.7 kg or 6 lb. Moreover, we may use different UOMs for one attribute. As an example, the 'Capacity' of 'Memory Card's may be 512 MB, 2 GB, or 16 GB, not 0.5 GB or 16,384 MB. In addition, some UOMs measured differently in different domains or industries, e.g. Pound and Kilogram or Foot and Meter. To resolve these issues in our system, we add two items to the [UOM], a reference to the base standard UOM, [uom_BaseUOMID], and a string field to hold the conversion formula for converting an UOM to its base unit, [uom_ToBaseUOM]. Each base quantity is independent of each other and has a corresponding base unit. For example, the base UOM of length and time is Meter (m) and Second (s), respectively and the conversion formula for changing meter to centimeter is 'm/100'. Reference (Taylor & Thompson, 2008, p. 23), defines seven base quantities and its base units based on the SI standard, including *length (m)*, *mass (Kg)*, *time (s)*, *electric current (A)*, *thermodynamic temperature (K)*, *amount of substance (mol)*, *and luminous intensity (cd)*. The (Lee, Shim, Lee, & Lee, 2006), also adds the *computer data* base quantity with *byte* base unit (Table 1).

Additionally, reference [9] adds *MultipleBound* and *SubmultipleBound* fields to the [UOM] to solve the problem of conversion between compatible UOMs. MultipleBound and SubmultipleBound show the limit

Table 1. Eight base quantities and their base units

Quantity	Name	Base Unit
length	meter	m
mass	kilogram	Kg
time	second	s
electric current	ampere	A
thermodynamic temperature	kelvin	K
amount of substance	Mole	mol
luminous intensity	candela	cd
computer data	Byte	byte

of multiplying and submultiplying the UOM to its higher and lower prefix measure, respectively. For instance, 'GB' (gigabyte) has 1 and 1 for its MultipleBound and SubmultipleBound. Therefore, the 'Capacity' attribute of 'Memory Card' can be represented by 'GB' unit of measurement and simply convert to 'TB' (terabyte) or 'MB' using conversion equations if needed.

Identifying different dimensions of product attributes and its values not only allows better understanding of products but also it improves the intelligence of the ontology system by adding more semantic functionalities and meaningful capabilities. It helps the system to understand the meaning of an attribute. Figure 1, represents the developed system.

5. Attribute Inheritance

To reduce the number of properties in the product ontology system [8] and share common characteristics among different products in a same category, we can assign properties to product classes in the class hierarchy, not only to the products (hierarchy leaves). To enable property inheritance in the system, the class-property relation and the class-property-value relation should be available for all classes. With attribute inheritance, all shared properties are inherited to all lower classes. Additionally, they can be modified on lower levels.

6. Related Work

A group of researchers in the product ontology and related areas are researchers from Korea, which develop one of the best governmental e-procurement services in the world. They develop this system based on product ontology, during 10 years research. In [12; 13; 14], they first present their system and then cite some issues encountered in their product ontology development, such as ontology-based searching system and ontology visualization, construction, and maintenance. In [9], the authors point out some of experienced problems in product ontology building, product attributes, and especially UOMs. They also propose operational

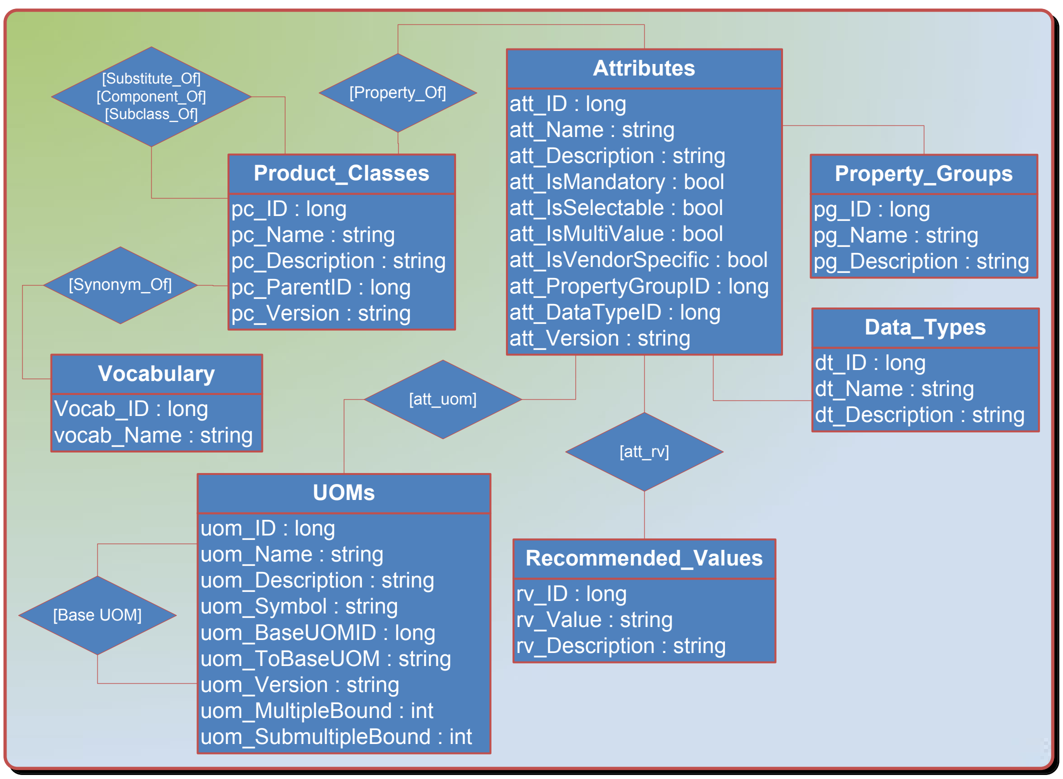

Figure 1. A model for product ontology systems with focus on attributes and attribute values

modeling guidelines for each of them.

Professor J. Leukel in [8] presents the problem of property growth and proposes measures for coping with the problem of extensive and steadily growing property libraries. He recommends solutions for fixing the problem and finally, he develops a corrected model for the property growth problem of product ontology systems. He suggests property groups, attribute inheritance, and optional and mandatory properties to control the product attributes in the product classification schemas. In [15], J. Leukel et al, describe a modeling approach for classification schema. They develop current standards and provide an enhanced framework for product classification. In [16] and [17], J. Leukel and professor Martin Hepp measure the quality of product classification standards by a number of factors, including the quality of their attribute lists. They also present a framework of metrics for the quality and maturity of categorization standards.

7. Conclusions

Including more details about products, by means of attributes and attribute values, increases the intelligence of the ontology system. Storing data types of the attribute value and its unit of measurement, enable us to extract meaningful features from the system. Moreover, indentifying different types of value assignment, such as single-value and multi-value properties and selectable vs. literal attributes, creates a flexible infrastructure that allow advanced searching and browsing techniques in addition to special semantic capabilities. In this paper, we mainly focus on some aspects of attributes that was not previously considered well enough. We present operational models that benefits ontologies to enable the understanding of attributes and improve the intelligence of the system about products.

8. Acknowledgment

We should thank Ms. Kyung-Soon Chang, director of the Public Procurement Service, South Korea, and professor Sang-goo Lee and Mr. Sangwoo Kim, for sending us useful materials about their ontology system. We also thank Dr. Amini Lari and Dr. Ahmadzadeh for their helps. Finally, we thank Dr. Khosravi Farsani for his useful comments.

9. References

[1] D. Fensel, et al., “Product data integration in b2b e-commerce”, *IEEE Intelligent Systems*, 16, 54–59, 2001.

[2] J. Leukel, “Standardization of product ontologies in b2b relationships – on the role of iso 13584”, *10th Americas Conference on Information Systems*, AIS, 2004, pp. 4084–4091.

[3] M. Hepp, *Ontology Management: Semantic Web, Semantic Web Services, and Business Applications*, M. Hepp, P. D. Leenheer, A. d. Moor, & Y. Sure (eds.), Springer, 2008.

[4] N. Guarino, “Ontology and terminology”, Retrieved April 4, 2008, from www.eaft-aet.net/fileadmin/files/VAKKI/nicola_guarino.pdf, 2006.

[5] Guarino, N., Giaretta, P. “Ontologies and Knowledge Bases: Towards a Terminological Clarification”, N. Mars (ed.), *Towards a Very Large Knowledge Bases: Knowledge Building and Knowledge Sharing*, IOS Press, Amsterdam, 1995, pp. 25-32.

[6] T. R. Gruber, “A translation approach to portable ontology specifications”, *Knowledge Acquisition*, 5(2), 1993, pp. 199–220.

[7] M. Gruninger and J. Lee, “Ontology applications and design”, *Communications of the ACM*, 45(2), 2002, pp. 39–41.

[8] J. Leukel, “Controlling Property growth in product classification schemes: a data management approach”, *ICEIS 2006*, LNBIP 3, Springer, 2008, pp. 363–374.

[9] H. Lee, J. Shim, S. Lee, & S.-g. Lee, “Modeling considerations for product ontology”, *The First International Workshop on Ontologizing Industrial Standards*, LNCS 4231, Springer, 2006.

[10] B. N. Taylor & A. Thompson, “NIST special publication 330 - 2008 edition”, Retrieved January 3, 2009, from International System of Units (SI): http://physics.nist.gov/Pubs/SP330/sp330.pdf, 2008.

[11] ISO. Retrieved Novamber 21, 2008, from International Standards Organization-ISO: www.iso.org, 2008.

[12] T. Lee, I.-h. Lee, S. Lee, S.-g. Lee, D. Kim, J. Chun, et al., “Building an operational product ontology system”, *Electronic Commerce Research and Applications*, 5, 2006, pp. 16-28.

[13] T. Lee, J. Shim, H. Lee, & S.-g. Lee, “A pragmatic approach to model and exploit the semantics of product information”, S. Spaccapietra (ed.), *Journal on Data Semantics*, VII (LNCS 4244), 2006, pp. 242-266.

[14] I.-h. Lee, S. Lee, T. Lee, S.-g. Lee, D. Kim, J. Chun, et al., “Practical Issues for Building a Product Ontology System”, *Proceedings of the International Workshop on Data Engineering Issues in E-Commerce (DEEC05)*, IEEE Computer Society, 2005, pp.16-25.

[15] J. Leukel, V. Schmitz, & F.-D. Dorloff, “A modeling approach for product classification system”, *The 13th International Workshop on Database and Expert Systems Applications (DEXA'02)*, 2002.

[16] M. Hepp, “Measuring the quality of descriptive languages for products and services”, F.-D. Dorloff (ed.), *E-Business – Standardisierung und Integration*, 2004.

[17] M. Hepp, J. Leukel, & V. Schmitz, “A quantitative analysis of product categorization standards: content, coverage, and maintenance of ecl@ss, unspsc, eotd, and the rosettanet technical dictionary”, *Knowledge and Information Systems (KAIS)*, 2005.