ONTOLOGY-BASED COMPUTERIZED REPRESENTATION OF SPECIFICATIONS FOR CONSTRUCTION COST ESTIMATION

Zhiliang Ma, Professor, mazl@tsinghua.edu.cn
Zhenhua Wei, Ph.D. Candidate, weizh04@gmail.com
Zhe Liu, Ph.D. Candidate, zhe.liu.thu@gmail.com
Department of Civil Engineering, Tsinghua University, Beijing, China

ABSTRACT

Construction cost estimation (CCE for short hereafter), which is normally labor-intensive and error-prone, is one of the most important works concerned by multi-participants in the AEC/FM industry during a project's lifecycle. Up to now, estimators have to firstly learn and comprehend specifications for CCE (CCE specifications for short hereafter) before they carry out CCE, thus the proficiency of estimators on CCE specifications greatly affects the efficiency and accuracy of CCE. With the development of information technology, it is expected that the efficiency and accuracy of CCE can be greatly improved by implementing CCE specifications in computer programs in their computerized representation. Due to its logic-based and computer-oriented features, ontology is drawing extensive attention for establishing the computerized representation of CCE specifications. In this study, a typical CCE specification in China is firstly analyzed and relevant models are established. Then, an ontology-based computerized representation of CCE specifications is established based on the models. Finally, a prototype tool for facilitating the establishment, modification and extension of the ontology-based computerized representation for estimators is introduced and the use cases of the tool are illustrated. The findings of this study lay a sound foundation for developing computer programs for CCE not only in China but also in other countries.

Keywords: construction cost estimation, specification, ontology

1. INTRODUCTION

Construction cost estimation (CCE for short hereafter), which is normally labor-intensive and error-prone, is one of the most important works concerned by multi-participants in the AEC/FM industry during a project's lifecycle. In many cases, specifications for CCE (CCE specifications for short hereafter), such as UNIFORMAT and MasterFormat, need to be strictly complied with in order to develop an accurate estimation, especially for bidding (Ma and Wei. 2012). Up to now, construction cost estimators have to firstly learn and comprehend CCE specifications before they carry out CCE, thus the proficiency of estimators on the specifications greatly affects the efficiency and accuracy of CCE.

With the development of information technology, it is expected that the efficiency and accuracy of CCE can be greatly improved by implementing CCE specifications in computer programs (Ma et al. 2013). Considering that specifications are usually written in natural languages, one of the key issues is to translate the human-being-oriented specifications into certain computerized representation (Eastman et al. 2009). Several relevant researches, mainly focusing on design domain, have been carried out and such information technologies as hard coding, parametric or decision tables, and ontology have been used to resolve the issue (Eastman et al. 2009; Pauwels et al. 2011; Wang and Boukamp 2011). For example, in a previous study, the authors have stored CCE specifications in a relational database as their computerized representation (Ma et al. 2013). However, ontology is regarded as a promising technology to represent CCE specifications due to its computer-oriented and logic-based features and thus is drawing extensive attention. For example, Pauwels et al. (2011) deployed ontology to represent the acoustic performance specifications, and the authors (2012) made a preliminary discussion on the ontology-based computerized representation of CCE specifications.

This paper aims to present an ontology-based approach that facilitates the establishment of the computerized representation of CCE specifications efficiently for estimators so that CCE specifications can be better implemented in computer programs for CCE. Firstly, a typical CCE specification in China is analyzed and relevant models are established. Then, an ontology-based computerized representation of CCE specifications is established based on the models. Finally, a prototype tool for facilitating the establishment, modification and extension of the ontology-based computerized representation is introduced and the use cases of the tool are illustrated.

2. ANALYSIS ON CCE BASED ON A TYPICAL CCE SPECIFICATION

Theoretically, CCE specifications specify all the necessary business rules for carrying out CCE. Considering that the bill-of-quantity (BQ for short hereafter) method has become a well-accepted CCE method all over the world, a Chinese national mandatory CCE specification, which is based on the method and called “Code of valuation with bill of quantity of construction works (Chinese standard number GB50500-2008)” (GB50500 for short hereafter), is analyzed and relevant models are established in order to derive its computerized representation.

2.1 Discrimination of BQ items for construction products

According to GB50500, a project is broken down into hundreds of BQ items for CCE, and each BQ item is coded with nine digital numbers and represents a set of construction products (products for short hereafter) which have the same features with specified values. For example, the BQ item “rectangle column of CIP concrete” is coded as “010402001” and represents a set of columns that have the same “material” feature with a “cast-in-place concrete” (CIP concrete for short hereafter) value and “cross shape” feature with a “rectangle” value. In other words, if a product is known to have the type of “column”, material of “CIP concrete” and cross shape of “rectangle”, it can be classified as the BQ item of “rectangle column of CIP concrete”. Accordingly, a discrimination rule can be abstracted for the BQ item, as shown in Figure 1. Based on a deep analysis on all BQ items in GB50500, it is found that at most five categories of features, i.e. product type, material, geometry, functionality and construction, are needed to formulate the discrimination rules for any BQ item, which are called the “discrimination features for BQ items”. Table 1 shows several typical examples of the discrimination features for BQ items and their values.

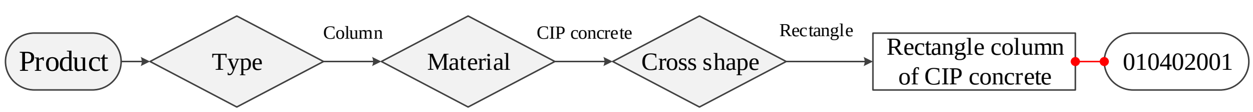


Figure 1: Discrimination rule of BQ item “rectangle column CIP concrete” coded as “010402001”

Table 1: Examples of discrimination features for BQ items and their values.

No.	Category	Name	Level in the discrimination model	Value
1	Product type	Type	Product	Footing/Column/Beam/Slab/Wall/…
2	Material	Material	Column	CIP concrete/Precast concrete
3	Geometry	Cross shape	Column: CIP concrete	Rectangle/Other
4	Geometry	Axis shape	Beam: CIP concrete: Common	Straight/Arc
5	Functionality	Functionality	Beam: CIP concrete	Common/Footing beam/Ring/Lintel
6	Construction	Construction method	Reinforcement bar: Common: Stress	CIP concrete bar/Precast concrete bar

By analyzing all the discrimination rules of BQ items in GB50500, a discrimination model for BQ items was established in the form of tree structure, part of which is shown in Figure 2. In the model, each branch represents a discrimination rule for BQ items; each node represents a discrimination feature for BQ items, whose value differs by its level in the branch as shown in Table 1; and each leaf represents a BQ item. The process for discriminating the BQ items of a product starts from the root node. In order to reach next node, the value of the corresponding discrimination feature is queried, and different values will guide the process to different sub-branches. The process is terminated when it reaches certain leaf, where the corresponding BQ item can be uniquely discriminated for the product. Obviously, the model can be used to formulate the algorithm of computer programs for discriminating the BQ items of the products in a project.

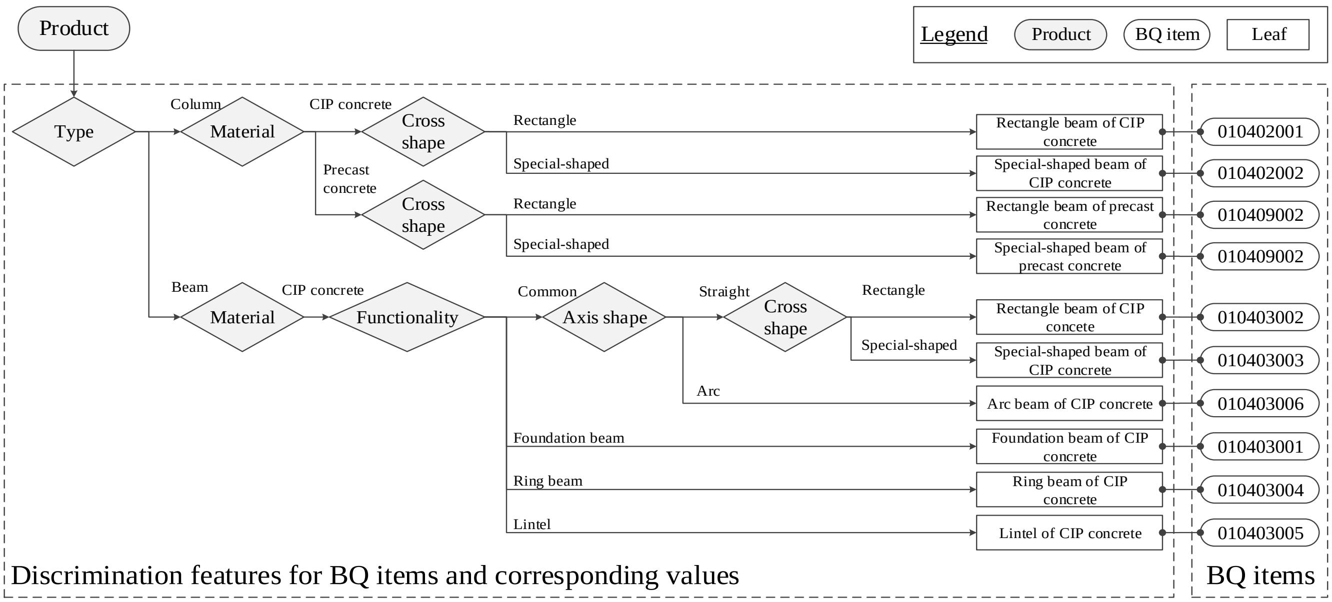
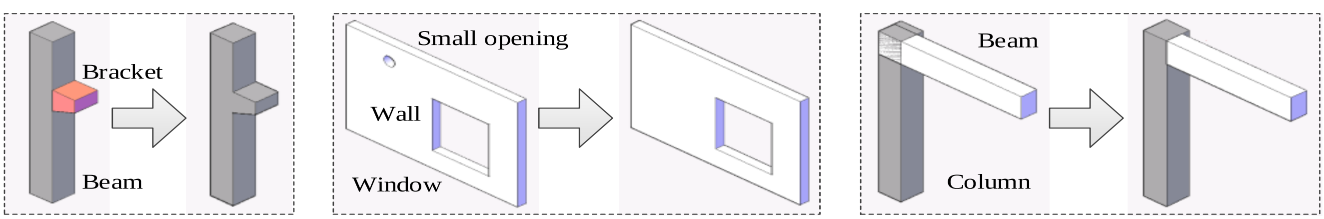


Figure 2: Part of discrimination model for BQ items

2.2 Calculation of Quantities for BQ items

Once the BQ items are discriminated, their quantities, such as the length, area or volume of the corresponding products, should be calculated by strictly complying with the quantity takeoff rules specified in GB50500, which include both unit rules and deduction rules. The unit rules indicate the quantity units of the BQ items, while the deduction rules are mainly related to the products that have the following three kinds of spatial features, i.e. accessory, opening and intersection relationships as described below.



(a) Accessory relationship (b) Opening relationship (c) Intersection relationship

Figure 3. Spatial features for quantity takeoff

a) "Accessory relationship": The quantity of an accessory, such as a bracket and a beam pad, should be added to the quantity of the product that the accessory belongs to, as shown in Figure 3(a).
b) "Opening relationship": If the size of an opening is larger than a prescribed value, its quantity should be deducted from the quantity of the product that the opening belongs to, as shown in Figure 3(b).

c) "Intersection relationship": The quantity of an intersection part (intersection quantity for short hereafter), such as the joint between a column and a beam, should not be double calculated. For the convenience of calculation, a relative priority between two BQ items is defined in this study. The intersection quantity should be deducted from the product whose BQ item is of lower priority, as shown in Figure 3(c).

2.3 Discrimination of quota items for BQ items

In the BQ method, the total cost of a project is computed by summing up the cost of all BQ items, which are calculated by multiplying their quantities by their "comprehensive unit cost" respectively. Obviously, the comprehensive unit cost of a BQ item depends on the works for building the products that belong to the BQ item. For example, the BQ item "rectangle column of CIP concrete" covers the works of producing, transporting, casting, and curing of concrete on site. Therefore, the comprehensive unit cost of each BQ item can be calculated by summarizing the cost of all its related works. It needs to be noted that the cost of the works may vary from countries, regions or firms due to different productivity levels. In China, the unit cost of each work is normally prescribed as a quota item in the national, regional or firm CCE specifications for the quota method (another CCE method adopted in China), for example "National unified basic quota items of construction works" (Chinese standard number GJD-101-95) (GJD-101-95 for short hereafter) which is another Chinese national CCE specification. Thus, the key for calculating the cost of a BQ item is to discriminate the quota items according to its related works. Figure 4 shows an example of the discrimination.

Similar to the discrimination model for BQ items, a discrimination model for quota items was established in the form of tree structure by abstracting discrimination rules and features for quota items in GJD-101-95. Due to the limited space of this paper, the model is not discussed here. It should be noted that, according to their characteristics, the discrimination features for quota items can be further classified into two categories, i.e. the essential features that are independent of the productivity levels, and the construction features that are dependent on the productivity levels, including the construction methods, technologies, machines, etc. of the works.

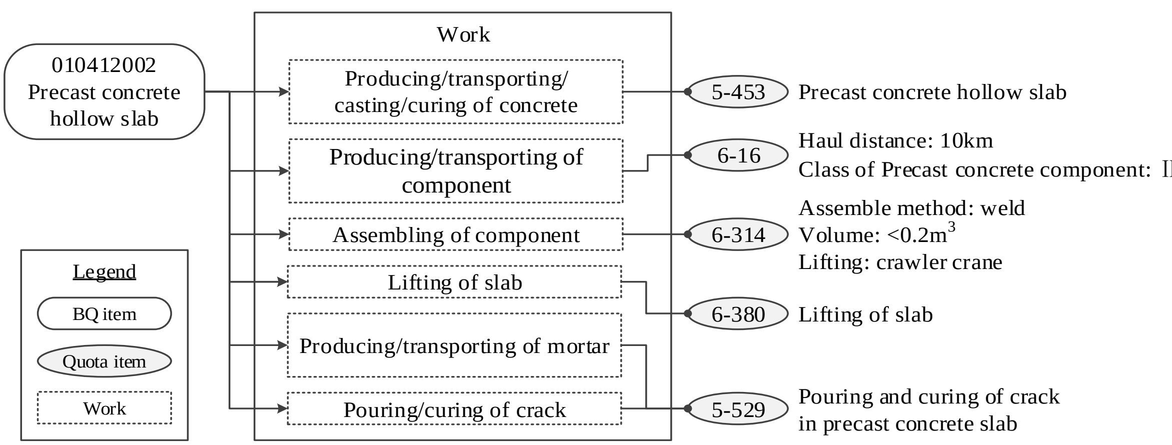


Figure 4: Discrimination of quota items for each work of an example BQ item

2.4 Specification-based information requirement model of CCE

Based on the analysis on CCE specifications stated above, a specification-based information requirement model was established, as shown in Figure 5. It includes (a) four kinds of information subjects (products, BQ items, works, and quota items), (b) three kinds of features (discrimination features for BQ items, spatial features, and discrimination features for quota items), (c) three kinds of corresponding relationships (relationships between products and BQ items, relationships between BQ items and works, and relationships between works and quota items), (d) two kinds of other relationships (Boolean relationships between products and priority between BQ items) and (e) three kinds of rules (discrimination rules for BQ items, quantity takeoff rules, and discrimination rules for quota items). Based on the model, the data of the information subjects, features and relationships can be acquired from the design model and the construction plan of a project, while the rules determine how to process these data and generate new data of quantities, resource consumptions, cost, etc. in carrying out CCE.

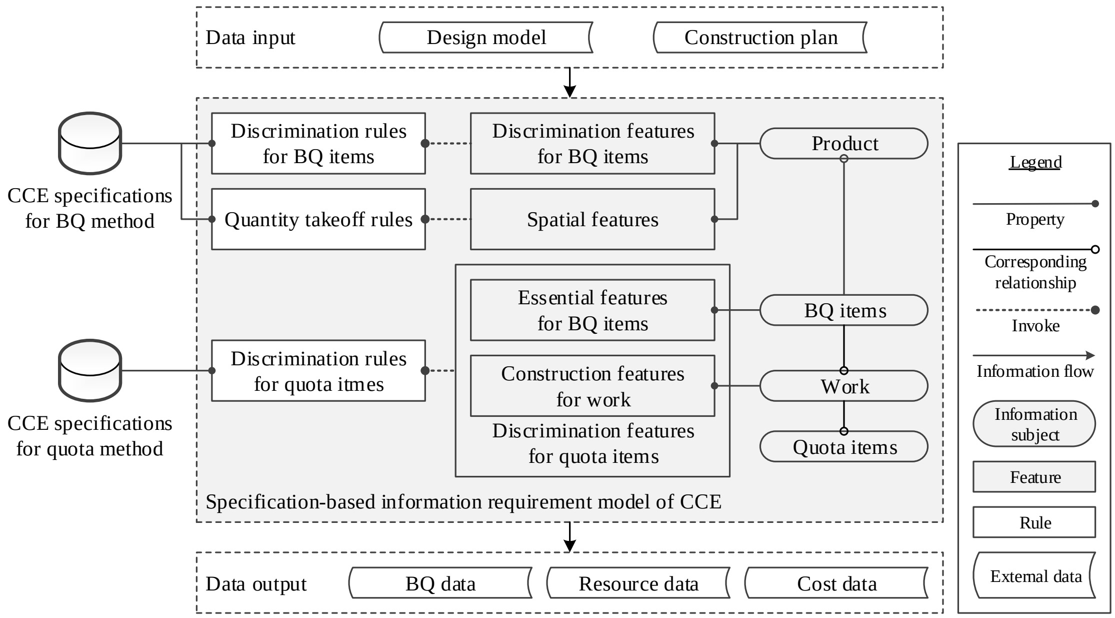


Figure 5: Specification-based information requirement model of CCE

3. ONTOLOGY-BASED COMPUTERIZED REPRESENTATION OF CCE SPECIFICATIONS

According to literatures (Beetz et al 2005; Eastman et al. 2009; Pauwels et al. 2011; Tan et al. 2010), ontology has its advantages in representing the human-being-oriented specifications in computer programs in comparison to such technologies as hard coding and parametric or decision tables. In CCE domain, it is expected that ontology can be used to represent information subjects, features and their values, relationships and rules in the specification-based information requirement model in a complete and extensible approach. As a result, CCE specifications can be implemented in computer programs in a computerized representation, and when CCE specifications changes, estimators can establish, maintain or update the corresponding representation efficiently without recompiling or redeveloping computer programs. Thus, ontology is adopted as the technology for deriving the computerized representation of CCE specifications in this study. Based on the information requirement model, an ontology-based model was established as the computerized representation of CCE specifications, which includes two parts, i.e., a knowledge model and a rule library.

3.1 Knowledge model for CCE

The knowledge model is the computerized representation of information subjects, features and their values, and relationships in the information requirement model. Since OWL (Ontology Web Language), recommended by W3C (World Wide Web Consortium), is more powerful in ontology modeling than other languages such as XML (Extensible Markup Language), RDF (Resource Description Framework) etc. and supports logical reasoning, it is adopted in this study. In OWL, concepts are represented as OWL classes while relationships between concepts are represented as OWL properties. It should be noted that information subjects and values of features are concepts while features represent special relationships between them. For example, if the material of a column is CIP concrete, this material feature is regarded as a relationship between the concepts of “column” and “CIP concrete”.

In this study, OWL classes and properties are predefined based on the information requirement model as shown in Table 2, and they constitute the framework of the OWL-based knowledge model for CCE, as shown in Figure 6. Based on the framework, the data of a project can be stored in an OWL file by defining individuals of

OWL classes and related with each other by OWL properties. For example, products in a project can be represented as individuals of OwlProduct class or its subclasses.

Table 2: Predefined OWL classes and properties in the knowledge model for CCE

No.	Concepts and relationships in the information requirement model		Predefined OWL classes and properties in the knowledge model
1	Information subject	Product	OwlProduct class and its subclasses
		BQ item	OwlBQItem class and its subclasses
		Work	OwlWork class and its subclasses
		Quota item	OwlQuotaItem class and its subclasses
2	Value of feature	Value of discrimination feature for BQ items	OwlFeatureValue_BQJudge class and its subclasses
		Value of essential feature for BQ items	OwlFeatureValue_BQItemFeature class and its subclasses
		Value of construction feature for work	OwlFeatureValue_ConstructFeature class and its subclasses
3	Feature	Discrimination feature for BQ items	RelatingBQJudgeFeature property and its sub-properties
		Essential features for BQ items	RelatingBQItemFeature property and its sub-properties
		Construction feature for works	RelatingConstructFeature property and its sub-properties
		Spatial feature	SpatialRelationship property and sub-properties
4	Corresponding relationship	Relationship between products and BQ items	ContainProducts and BelongBQItem property
		Relationship between BQ items and works	ContainWorks and BelongBQItem property
		Relationship between works and quota items	RelatedWork and RelatingQuotaItem property
5	Other relationship	Boolean relationship between products	BooleanOperation property and sub-properties
		Priority between BQ items	RelativePriority property and its sub-properties

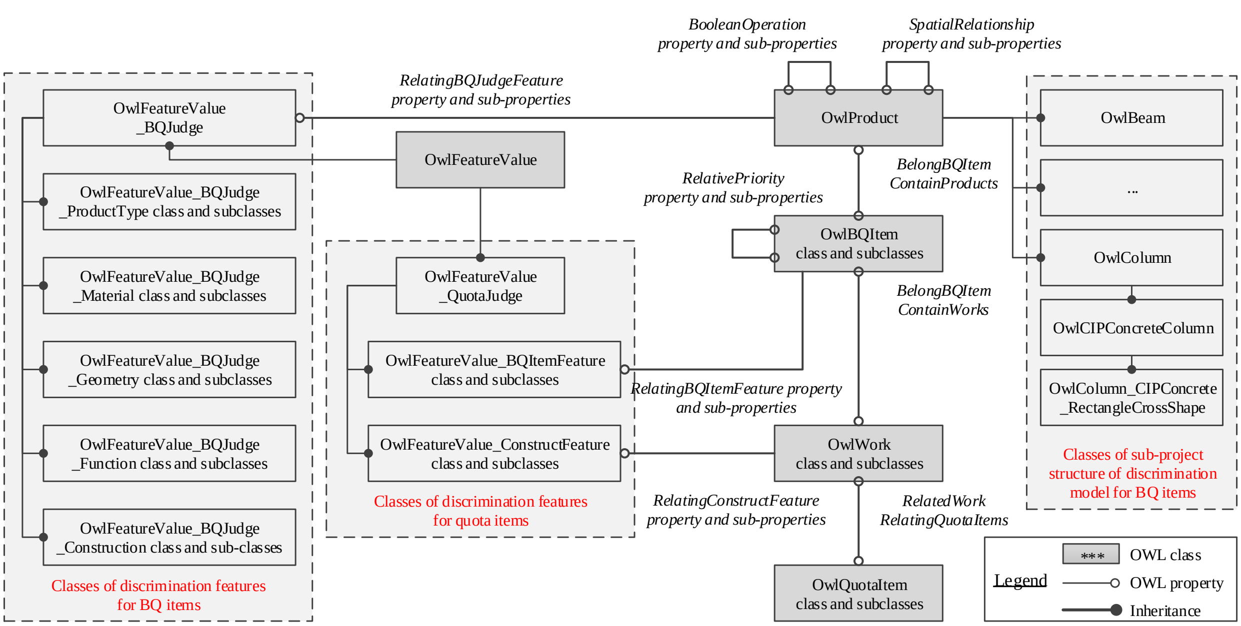


Figure 6: Framework of the OWL-based knowledge model for CCE

3.2 SWRL-based rule library for CCE

The rule library for CCE was established by defining logical reasoning rules corresponding to the three kinds of rules in the information requirement model. It should be noted that all these rules are Horn-like rules, in which a serial of conditions can draw only one conclusion by logical reasoning. Since Horn-like rules cannot be directly represented by OWL, SWRL (Semantic Web Rule Language), recommended by W3C based on OWL and RuleML (Rule Makeup Language), is adopted to represent the rules. The logical reasoning rules represented in SWRL (SWRL rules for short hereafter) are defined by invoking relevant OWL classes and properties in the OWL-based knowledge model.

In this study, all the 70 discrimination rules for BQ items of "sub-project of concrete" specified in GB50500 are translated into SWRL rules. Take the discrimination rule for BQ item "010402001" for instance, as shown in Figure 7, by invoking relevant OWL classes and properties in the OWL-based knowledge model, the corresponding SWRL rule is created as below (the meaning of each part is explained with a comment at the end of the line).

OwlProduct (?x) // A product is defined as x
∧ OwlBQItem_010402001 (?y) // A BQ item 010402001 is defined as y
∧ Column (?z) // Value of discrimination feature: column (defined as z)
∧ CIPConcrete (?w) // Value of discrimination feature: CIP concrete (defined as w)
∧ Rectangle (?v) // Value of discrimination feature: rectangle (defined as v)
∧ OwlProperty_ProductType (?x, ?z) // Discrimination feature: x is of type z
∧ OwlProperty_MaterialofColumn (?x, ?w) // Discrimination feature: x is made of material w
∧ OwlProperty_CrossShapeofCIPConcreteColumn (?x, ?v) // Discrimination feature: x is of cross shape v
→ BelongBQItem (?x, ?y) // Relationship: product x belongs to BQ item y

In SWRL rules, "∧" is an operator for the logical AND between conditions, and "→" is an operator for drawing the conclusion. For the quantity takeoff rules and discrimination rules for quota items, the corresponding SWRL rules can be defined similarly. In this way, estimators can create all the SWRL rules to establish a SWRL-based rule library, which can be used for the logic reasoning tasks when carrying out CCE, i.e., the discrimination of BQ items for products, the calculation of quantities for BQ items, and the discrimination of quota items for BQ items. In detail, based on the data of a project stored in an OWL file stated above, all the individuals and their relationships can be automatically queried and invoked by certain SWRL rule as the required conditions to draw a conclusion as new data with the support of certain logical reasoning machines such as Protégé, Jena, etc. It is expected that the SWRL-based rule library along with the OWL-based knowledge model can be used for better implementing CCE specifications in computer programs for CCE.

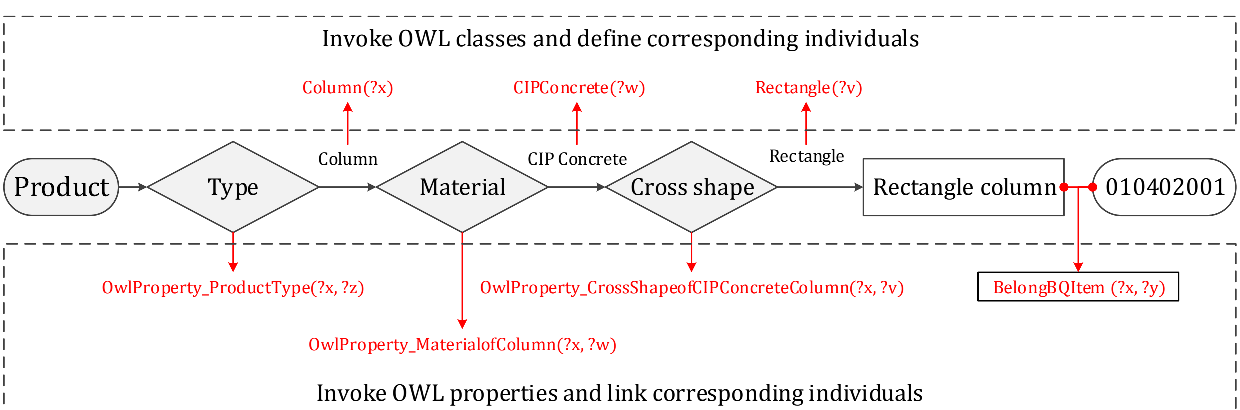


Figure 7: Discrimination of BQ item "010402001" and process for defining the corresponding SWRL rule

4. PROTOTYPE TOOL FOR COMPUTERIZED REPRESENTATION OF CCE SPECIFICATIONS

During the process of CCE, estimators are required to specify and extend the CCE specifications in the existing CCE software according to their understandings of CCE for some special cases, such as specially-shaped columns, new materials and construction methods, which takes time and efforts. Thus, the modifiability and extensibility are required of the computerized representation of CCE specifications. Additionally, it should be leant easily and used efficiently. Although the existing ontology modeling tools, such as Protégé (Matthrew 2011), provide a rich set of functions and user-friendly interface, a specific CCE-orient tool is necessary to facilitate the establishment, modification and extension of the ontology-based computerized representation.

4.1 Process for establishing computerized representation of CCE specifications

Considering that the SWRL-based rule library is established by invoking relevant OWL classes and properties in the OWL-based knowledge model, the establishment of the OWL-based knowledge model is the key for the computerized representation of CCE specifications. A step-by-step process is formulated for estimators to establish the OWL-based knowledge model according to the three relevant models or rules described in section 2, which includes five major steps as shown in Figure 8.

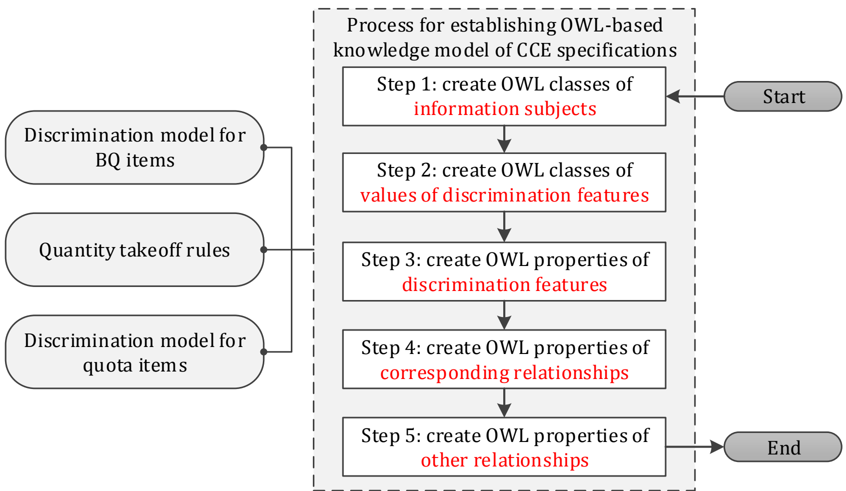


Figure 8: Process for establishing OWL-based knowledge model of CCE specifications

It should be noted that, due to the complexity of CCE specifications, it still takes time and labor to establish their computerized representation. However, since the process is closely related to CCE knowledge and rules which are familiar to estimators, it is a much easier and less labor-intensive approach for estimators to establish the ontology-based computerized representation than that by using hard coding, parametric or decision tables.

4.2 Development of the prototype tool

Based on the process for establishing the OWL-based knowledge model and defining SWRL-based rule library, a prototype tool is under development for estimators to establish, modify and extend the ontology-based computerized representation of CCE specifications efficiently. In the tool, all the OWL classes and properties shown in Table 2 are predefined and a CCE-specification-oriented interface is designed for estimators to reduce the learning cost and raise their working efficiency.

4.2.1 Major functions

Three categories of major functions are designed for the prototype tool as follows.

a) Basic management functions, including creating, editing, searching, browsing, and deleting the information subjects, features and their values, relationships and rules.
b) Functions for importing and exporting the ontology-based computerized representation of the objective CCE specification as an OWL file, which acts as the schema for the data of a project.
c) Functions for graphic display and operation, which can help estimators to understand the computerized representation and logical reasoning process well in a user-friendly way rather than learn, understand and operate OWL classes and properties or even hard coding.

4.2.2 Development environment

To facilitate the integration of the prototype tool with existing CCE software developed by the authors (Ma et al. 2013), the prototype tool is mainly developed in C++ programing language. In addition, JNI (Java native interface) is employed to invoke the Protégé Java APIs to implement the basic management functions as well as the functions for importing and exporting OWL files.

4.3 Use cases of the prototype tool

The user interface of the tool is mainly divided into four parts, i.e., (a) a menu bar containing file operations, import and export of OWL files, graphic tools of ontology modeling operations, etc. on the top, (b) a tree structure browser of CCE specifications on the left, (c) a BQ item property editor on the right, and (d) a graphic display view for the rules in the middle, as shown in Figure 9.

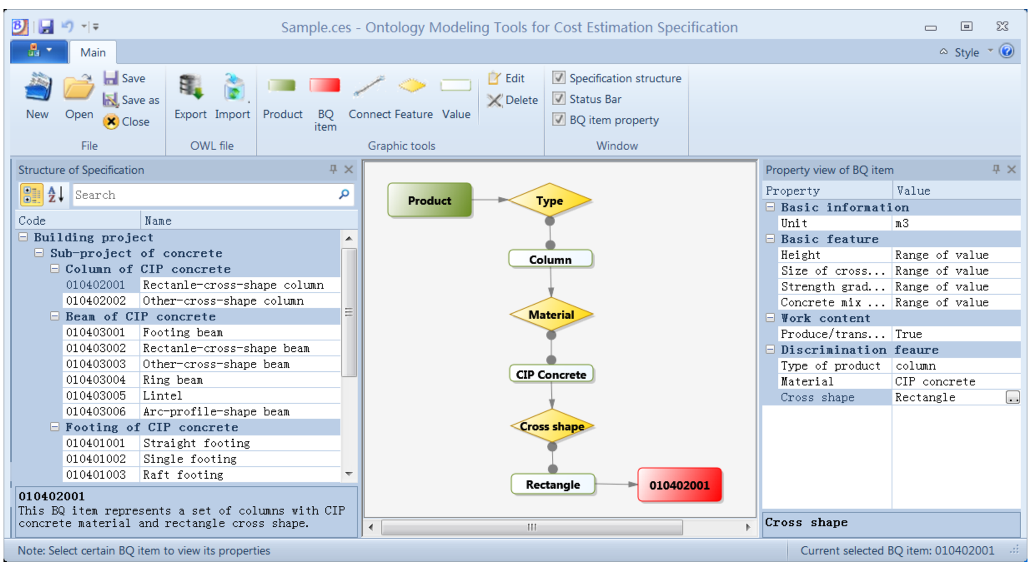


Figure 9. User interface of the prototype tool

Take the use case of the creation of a discrimination rule for BQ item for example. Estimators (a) firstly select a BQ item from the tree structure browser, (b) then drag corresponding "Product", "BQ item", "Feature", or "Value" from the graphic tools in the menu bar into the graphic display view with assigning a specific name for them respectively, and link them together with "connect", (c) the discrimination features and their value in the BQ item property editor keep synchronization with the graphic display view, and estimators can also edit them in the editor, (d) meanwhile, as the background processing, the tool will automatically generate corresponding OWL classes and properties as well as SWRL rules in accordance with the estimators' operation, (e) and the creation of the discrimination rule is finished. Obviously, graphic operations instead of operations on OWL classes and properties will improve the efficiency of the establishment of the computerized representation of CCE specifications, thus will meet the requirement of estimators stated above. The operations for the modification and extension of the computerized representation of CCE specifications are very similar.

5. CONCLUSION

In order to develop an ontology-based approach that facilitates the establishment of the computerized representation of CCE specifications efficiently, a typical CCE specification in China was firstly analyzed and relevant models were established. Then based on the models, the ontology-based computerized representation of CCE specifications was established, including an OWL-based knowledge model and a SWRL-based rule library. Finally, a prototype tool for facilitating the establishment, modification and extension of the ontology-based computerized representation was introduced and the use cases of the tool were illustrated. The proposed approach lays a sound foundation for developing computer programs for CCE not only in China but also in other countries.

The prototype tool is currently under development. In addition, a research on the integration of the tool, generic logical reasoning machine and quantity takeoff module is being carried out in order to implement a semi-automatic or even automatic CCE based on the BIM models of projects.

ACKNOWLEDGEMENTS

This research is supported by "*National Technological Support Program for the 11th-Five-Year Plan of China*" (No. 2007BAF23B02), "*National Natural Science Foundation of China*" (No. 51278279) and "*Tsinghua University Research Fund*" (No. 2011THZ03).

REFERENCES

Beetz, J., van Leeuwen, J. P., & de Vries, B. (2005). "An ontology web language notation of the industry foundation classes". Proceedings of the 22nd CIB W78 Conference on Information Technology in Construction.

Eastman, C., Lee, J. M., Jeong, Y. S., & Lee, J. K. (2009). "Automatic rule-based checking of building designs". Automation in Construction, 18, 1011-1033.

Ma, Z. L., & Wei, Z. H. (2012). "Framework of automatic construction cost estimation based on BIM and ontology technology". Proceedings of the CIB W78 2012: 29th International Conference, Beirut, Lebanon.

Ma, Z. L., Wei, Z. H., & Zhang, X. D. (2013). "Semi-automatic and specification-compliant cost estimation for tendering of building projects based on IFC data of design model". Automation in Construction, 30, 126-135.

Ministry of Housing and Urban-Rural Development of People's Republic of China (1995). National unified basic quota items of construction works, GJD-101-95.

Ministry of Housing and Urban-Rural Development of People's Republic of China (2008). Code of valuation with bill quantity of construction works, GB 50500-2008.

Pauwels, P., Van Deursen, D., Verstraeten, R., De Roo, J., De Meyer, R., Van de Walle, R., & Van Campenhout, J. (2011). "A semantic rule checking environment for building performance checking". Automation in Construction, 20, 506-518.

Matthew, H. (2011). "A practical guide to building OWL ontologies using Protégé 4 and CO-ODE tools". The University of Manchester.

Tan, X., Hammad, A., & Fazio, P. (2010). "Automated code compliance checking for building envelope design". Journal of Computing in Civil Engineering, 24, 203-211.

Wang, H. H., & Boukamp, F. (2011). "Ontology-based representation and reasoning framework for supporting job hazard analysis". Journal of Computing in Civil Engineering, 25, 442-456.

Wei, Z. H., & Ma, Z. L. (2012). "Discussion on the ontology-based computerized representation of specifications for construction cost estimation". Proceeding of the 16th National Conference on Information Technology in Civil Engineering and Architecture, Guangzhou, China.

World Wide Web Consortium. "OWL web ontology language guide". http://www.w3.org/TR/2004/REC-owl-guide-20040210/ (6/3/2013).

World Wide Web Consortium. "SWRL: A semantic web rule language combining OWL and RuleML". http://www.w3.org/Submission/SWRL (6/3/2013).