Ontology-Based Representation and Reasoning in Building Construction Cost Estimation in China

Xin Liu, Zhongfu Li and Shaohua Jiang *

Department of Construction Management, Dalian University of Technology, Dalian 116024, China; xin_liu84@outlook.com (X.L.); lizhongfu@dlut.edu.cn (Z.L.)
* Correspondence: shjiang@dlut.edu.cn; Tel.: +86-0411-8470-7482

Academic Editors: Tamer E. El-Diraby and Jinyue Zhang
Received: 17 February 2016; Accepted: 18 July 2016; Published: 3 August 2016

Abstract: Cost estimation is one of the most critical tasks for building construction project management. The existing building construction cost estimation methods of many countries, including China, require information from several sources, including material, labor, and equipment, and tend to be manual, time-consuming, and error-prone. To solve these problems, a building construction cost estimation model based on ontology representation and reasoning is established, which includes three major components, i.e., concept model ontology, work item ontology, and construction condition ontology. Using this model, the cost estimation information is modeled into OWL axioms and SWRL rules that leverage the semantically rich ontology representation to reason about cost estimation. Based on OWL axioms and SWRL rules, the cost estimation information can be translated into a set of concept models, work items, and construction conditions associated with the specific construction conditions. The proposed method is demonstrated in Protégé 3.4.8 through case studies based on the Measurement Specifications of Building Construction and Decoration Engineering taken from GB 50500-2013 (the Chinese national mandatory specifications). Finally, this research discusses the limitations of the proposed method and future research directions. The proposed method can help a building construction cost estimator extract information more easily and quickly.

Keywords: ontology; cost estimation; semantic relationships; reasoning; specification for cost estimation

1. Introduction

Cost estimation is one of the critical tasks for participants in the AEC/FM (Architecture, Engineering, Construction, and Facilities Management) industry. However, in practice, estimators still have to extract useful information from printed design drawings or CAD drawings manually or rebuild three-dimensional models of cost estimation manually [1,2]. The process of manual work, which is time-consuming, lacking in efficiency, and prone to error, still depends on the estimator’s knowledge and experience, and the design results cannot be used directly due to the subjectivity that results from the cost estimator’s knowledge and experience.

BIM (Building Information Modeling) technology provides a potential solution to generate takeoffs, counts, and measurements directly from a unique and continuously updated model [1,2]. BIM supports the entire project lifecycle and has the capability to integrate costing efforts throughout all project phases [2]. Compared to traditional methods, all of the cost estimation information can be shared directly with the support of open data exchange standards for BIM, such as IFC (Information Foundation Classes) [3], which can avoid heavy human workload, subjectivity, and manual errors. At present, several BIM-based cost estimation software applications [4,5] have been developed to improve the efficiency and accuracy of estimation. However, these software applications rarely support the IFC standard, which is the most comprehensive data exchange standard for BIM. Moreover, there
are differences between the specifications of different regions for cost estimation, and these applications can hardly be used in actual projects in China.

Several studies have revealed that cost estimation work has mainly focused on the schematic design or tendering phase. Ma et al. [1] developed a program of cost estimation for tendering based on BIM, and Staub-French et al. [6] presented an ontology to represent estimators’ rationale regarding the features relating building product models to construction cost. Lee et al. [7] gained detailed information through the use of cost estimation based on BIM and ontology, such as the specific material and construction type. However, these studies are mainly concentrated on the design or tendering stage, and do not consider the project lifecycle. Although the cost estimation based on the BIM model is more efficient and accurate than the traditional methods, the intervention of the cost estimator’s subjective decisions and semantic ambiguity is inevitable in determining costs [1,7].

In dealing with the current cost estimation issues in the building construction industry, it is necessary to use a proactive approach rather than a reactive one. Thus, to solve the above problems, this research introduces the notion of ontology into cost estimation, and the built ontology model can be used in various phases of the project lifecycle. In this way, cost estimation ontology supports the extension or modification of knowledge and rules from the perspective of pure cost estimation rather than software implementation, thereby enhancing the extendibility of BIM-based cost estimation software for different specifications.

This research is organized as follows. The research background is outlined in Section 1. Section 2 provides a brief review of related studies. The cost estimation ontology is proposed in Section 3. Section 4, which is based on a cost estimation ontology, discusses the implementation of building construction cost estimation, and in Section 5, the proposed approach is illustrated with cost estimation examples, and a process of cost estimation is discussed. Finally, conclusions are drawn in Section 6.

2. Research Background

The research background is related to two issues, the first of which is BIM-oriented construction cost estimation and its problems. The second issue is ontology modeling and representation.

2.1. BIM-Oriented Construction Cost Estimation

BIM is one of the most promising developments in the AEC/FM industry [8]. With the development of BIM technology, a common information repository for all project participants is offered [2]. IFC is an industry-developed product data model for the whole lifecycle of buildings; supported by buildingSMART, it is a widely utilized standard for representing BIM [9]. Some widely known cost estimation software applications are Vico Estimator [10], Innovaya Visual Estimating [11], and Success Estimator [12]. Estimators still have to set conditions to comply with a specification manually, and quantity takeoff still depends on the imported design model. Due to the working complexity and compatibility, cost estimation is time-consuming and inaccurate.

In the AEC/FM industry, there have been numerous studies regarding the application of BIM-oriented cost estimation. Faraj et al. [13] developed an IFC Web-based collaborative construction computer environment named WISPER, which built an IFC-based object-oriented database to help users realize the network integration and sharing of the design, budget, schedule, and other information in construction projects. Staub-French et al. [14] suggested an IFC-based cost estimation system that is capable of directly using the results of IFC files and then automatically applying corresponding prices to accomplish the cost estimation according to the component geometries and properties. Fu et al. [15] developed a system for life-cycle cost assessment that can automatically extract cost-estimating data from the design results of IFC files and then transfer the data to a preexisting component of the life-cycle cost assessment. Yabuki et al. [16] applied the IFC standard to the cost estimation of earthwork and accomplished the cost estimation function according to the 4D model. Ma et al. [17,18] established an information model that can be applied to the development of a construction cost estimation software and established a framework for BIM-based Construction Cost
Estimating (CCE) software that utilized the Chinese national mandatory specification (GB 50500-2008), which laid a solid foundation for the development of next-generation CCE software. Simultaneously, they also proposed an IFC-based discrimination model for BOQ (Bill of Quantity) items and the corresponding rules and created a semantic database by analyzing related Chinese specifications; they also formulated a mechanism to intelligently generate BOQ from IFC data [19].

In summary, there are quite a few studies regarding cost estimation. Some mainly concentrate on cost estimation software applications based on BIM, and others propose methods of information modeling for BIM cost estimation. However, limited attention has been paid to the information integration of multiple data sources and its implementation in construction cost estimation.

2.2. Ontology Modeling and Representation

Berners-Lee et al. [20] presented an extension to the World Wide Web (WWW) named the Semantic Web, which is able to handle Web data without human intervention. Compared with the existing network, the Semantic Web can make data retrieval more accurate and improve knowledge sharing and reuse. Ontology is a key technology of the Semantic Web, originates from the philosophy domain, and has been widely adopted in research efforts in the artificial intelligence and computer science domain over the past two decades [21,22]. The widely accepted ontology definition is “an explicit and formal specification of a conceptualization,” as given by Gruber [21].

Several studies have been conducted related to the application of ontology to the AEC/FM industry. Zhu et al. [23] demonstrated the usability of IFC [3] for constructing a metadata model for RFI (Request for Information) that improved the retrieval of RFI-related information. El-Diraby and Osman [24] developed a domain ontology for construction concepts in urban infrastructure products, and their ontology provided a conceptualization for knowledge in civil infrastructure. Tserng et al. [25] presented an ontology-based risk-management framework of construction projects throughout the project lifecycle. Hsien et al. [26] proposed a method to extract concepts, instances, and relationships from the engineering domain handbook and quickly built a base domain ontology. Wang and Boukamp [27] built a domain ontology for JHA (Job Hazard Analysis) that used the existing JHA documents from a construction company and checked the applicability of the safe approaches under specific conditions. In the infrastructure product ontology (IPD-Onto), entities are used to represent the core concepts within the domain of the utility infrastructure, and the five main entities used are Project, Process, Actor, Resource, and Product [24]. Wang and Boukamp [28] presented a framework aiming to improve access to a company’s JHA knowledge by using ontologies for structuring knowledge regarding activities, job steps, and hazards. In the construction safety management domain, Zhang et al. [29] put forward an ontology consisting of three main domain ontology models: the Construction Product Model, Construction Process Model, and Construction Safety Model. Abanda et al. [30–32] proposed an ontology-based labor cost estimation model and New Rules of Measurement (NRM)-based ontology modeling by using ontology engineering methodology [33].

In summary, there have been many studies regarding ontology modeling and representation. Some studies have proposed ontologies for product modeling, and others have proposed ontologies for key concepts in the construction domain. However, there have been very limited studies regarding automating reasoning in construction cost estimation by means of specific conditions, related concepts, and the smallest unit of work. This paper proposes a construction cost estimation ontology model by combining concept models, construction conditions, and work items.

3. The Building Cost Estimation Ontology Model

Although approaches for developing ontologies have been discussed in many papers, every approach has its own purpose, depending on the research field. A building ontology approach that is suitable for one research field may not suit another field. Therefore, this study presents an ontology representation model for building construction cost that does not fit other fields. The architecture
of ontology can be viewed as a set of concepts related to cost estimation knowledge, and these concepts and the relationships between them should be organized using the characteristics of building construction cost estimation. The combination of theory and practice in ontology building can build relationships and axioms better. During the taxonomy-building stage, some existing taxonomies related to the AEC industry are considered and partially integrated into the proposed cost estimation ontology, where the concepts fit its range [34]. The fundamental structure of ontology [35] can be seen as:

\[ O = \{C, I, R, A\} \]  

(1)

where “O” represents the ontology for the building construction domain, “C” represents the set of classes (sometimes called concepts) in the ontology, “I” represents the set of instances, and is used to represent elements or individuals in the ontology, “R” represents the set of relationships and usually uses binary relations to represent associations between concepts of the ontology, and “A” represents the set of axioms, which are used to describe concepts, individuals, and relationships by various constraint conditions in the ontology.

3.1. Ontology Representation Model

The representation model of cost estimation in this research includes three sub-models: the concept model ontology, the work item ontology, and the construction condition ontology. The model aims to model cost estimation concepts in the ontology, and the data of the representation model are computer-readable. A building construction project is a collection of work items, and the concept model can be further sub-classified by using the work items and construction conditions that correspond to particular projects. The proposed approach is based on three models of the recognized base-core-actual modeling pattern. To make better use of the cost estimation ontology, mapping is required among the ontologies of the concept models, work items and construction conditions. Figure 1 summarizes the basic interpretation steps of the proposed ontology representation model.

Figure 1. Steps for ontology representation model interpretation.

3.1.1. Concept Model Ontology

The “Code of Valuation with Bill Quantity of Construction Works (Chinese standard number GB 50500-2013)” (GB 50500, hereafter) [36] is a Chinese national specification for cost estimation, and provides the unified rules for classifying and determining the bill quantity items by means of the bill quantity method. It defines a classification system similar to MasterFormat [37] for classifying construction products into cost items and provides many quantity takeoff rules for calculating the quantity takeoff for cost items [1]. MasterFormat, a publication of CSI (Construction Specifications Institute) and CSC (Construction Specifications Canada), is a master list of numbers and titles that
are classified by work results. Thus, this research takes the GB 50500 specification as the main reference specification.

The concept model ontology leverages ontological modeling to model cost estimation concepts and the relationships among these concepts, which are extracted from building specifications, construction documents, and BIM data. By analyzing the relationships between concepts, reasoning rules are defined. In the reasoning procedure, the concepts are selected to represent a list of potential conditions that must be considered in the cost estimation. These concepts are hierarchical, and the semantic relationships among them are organized to benefit the reasoning by identifying the reasoning rules.

According to the characteristics of building construction cost estimation, this research proposes the following two steps to build the concept model ontology.

Step 1: Concept taxonomical structure and hierarchy.

The first step is to define the related concept taxonomical structure and hierarchy for building construction cost estimation. Based on previous studies, combined with the characteristics of cost estimation in China, four primary classes, i.e., project classifications, building products, building features, and cost items, are defined as the cost estimation concept model ontology. Through the defined properties of the primary classes, the concepts of each class can be further categorized. Here are the steps to define these primary classes in this research:

1. Project Classification

GB 50500, published in 2013, divides construction engineering into subsection projects, subentry projects based on measurability and project valuation. GB 50500 defines hundreds of cost items for classifying all product concepts of the AEC and MEP (Mechanical/Electrical/Plumbing) industry in a building project. This specification meets the requirements of cost estimation in China and is taken as the main resource for extracting project classification concepts. Thus, GB 50500 is well-suited for project classification to represent its semantic properties, including monomial projects, unit projects, subsection projects, and subentry projects, and the relationship is shown in Figure 2. GB 50500 has a better classification hierarchy and clearer categorical attributes. UNIFORMAT II [38], a classification standard of ASTM (American Society for Testing and Materials) Standard E 1557-05 (the standard was revised in 2005 and represented as E1557-05), is a format for classifying building elements and related site work. Elements usually perform a given function, regardless of the design specification, construction method, or material used. UNIFORMAT II is a significant advance over the original UNIFORMAT classification in that it has added elements and expanded the descriptions of many existing elements [38]. Hence, UNIFORMAT II is chosen as the hierarchical reference for project classifications of the concept model ontology.

Figure 2. Example of the concept representation of project classifications.
2. Building Product

The classification of OmniClass Construction Classification System (known as OmniClass or OCCS, US) [39] is used as the secondary class to further represent concepts in the concept model ontology. OmniClass provides a detailed classification system for the construction industry, and it incorporates other exist systems that are currently in use as the basis of many classification schemes. For specific information, each of the classification schemes in OmniClass can be classified independently. Thus, OmniClass is well suited for concepts to represent their division-specific properties, and building products of the primary class are further divided into different levels in the concept model ontology by using 23-Products of OmniClass.

3. Building Feature

The description content of building features is combined with the actual construction project in the proposed concept model ontology such that building features can better meet the needs of determining unit prices. In GB 50500, building features include geometry features, material features, and construction features. Geometry feature represents the shapes of building products, such as the shape of a rectangular beam section; material feature represents the materials used to form the building products; and construction feature represents construction process, construction type, strength grade, etc.

4. Cost Item

Cost item is composed of project coding and the names of cost items. However, due to the differences of culture and economy among different countries, classification systems of these countries are not the same; there may even be multiple classification systems of information in one country, and the range, method, structure, and objective of classification in the systems may differ significantly. To better realize building construction cost information integration in China, this research takes GB 50500 as the classification and coding rule. Take a wooden door as an example to describe the cost items shown in Figure 3. Each cost item is expressed by nine numbers that belong to four levels. The first level represents the category of the project, e.g., a construction project (01). The second level represents the category of the work in the project, e.g., a doors and windows project (08). The third and fourth levels are the work subcategory and subentry project, respectively. Three more optional numbers are added to further divide projects in GB 50500.

![Figure 3. Example of the classification and coding system of GB 50500.](image)

Step 2: Defining the semantic relationship.

The relationships among these classes are used to describe the associations that occur between concepts in the ontology. When defining association relationships to connect concepts, the association relationships should be given semantically rich names to facilitate the understanding of how connected concepts are related [40]. Relationships can be categorized as hierarchical and non-hierarchical [41]. Joonhee [42] presented RNA (Relationship Navigation Analysis) and its generic relationship taxonomy, described their use for system analysis, and considered the relationships that exist between classes to be internal and external. Two main types of semantic relationships are used in this research: hyponymy
(superclass-subclass) and association. The relationships of concepts enrich concepts' definitions, so they can be used to better describe related concepts.

In Step 1, the concepts are represented, and the hyponymy relationship between concepts is introduced in the description of four primary classes. Thus, the association relationship, which includes the synonymy relationship (equivalent($x$, $y$)), antonymy relationship (disjoint($x$, $y$)), and meronymy relationship (whole–part($x$, $y$)), is introduced here, and the following describes the association relationship:

A synonymy relationship is used to describe the similarity between concepts. The following relationships represent types and categories of a synonymy relationship: is_the_same_as, is_like, and is similar_to. For example, “混凝土” is the same as “砼” in Chinese, and “concrete” is the same as “con’s” in English because “con’s” is the abbreviation of concrete.

An antonymy relationship is used to represent disjoining between concepts. For example, “pile foundation” is seen as opposite to “earthwork.”

A meronymy relationship (whole–part) is used to describe a whole–part relationship between concepts, that is, one concept is a part of the other. There are two methods for partitioning complex concepts: decomposition into parts (is a part of, is segment of, etc.) and decomposition into subclasses (include, consist of, contain, etc.). For example, “concrete engineering” includes “formwork engineering,” “reinforcement engineering,” and “concrete engineering.”

3.1.2. Work Item Ontology

A work item is considered to be the smallest unit of work for cost estimation purposes [43], including all building elements in the project. In this research, project classification is based on GB 50500, and cost item indicates project coding, as described in Section 3.1.1. Taking a cast-in-place concrete beam as an example work item, the building material in a work item is composed of, e.g., cement, sand, stone, and water. The project classification may comprise a cast-in-place concrete rectangular beam (the cost item id is 010502001) or a cast-in-place concrete constructional column (the cost item id is 010502002), that is, a work item is the smallest unit of work and is indivisible in building construction cost estimation. A construction process is composed of one or more sub-processes, and sub-process is a collection of work items that produce building elements. To estimate the cost of work item, the proposed ontology needs to know the corresponding item and the quantity of the necessary resources. Therefore, the building of work item ontology can improve the utilization of the cost estimation ontology.

The work item ontology is developed from the reuse of existing ontologies (like FreeCLassOWL ontology). FreeClassOWL ontology [44] (FC, hereafter), which was developed by the European building and construction materials market to describe construction materials and services, allows the fine-grained search of products, suppliers, and warehouses for any building-related sourcing needs. FC has over 88 million triples of real business data, 81 manufacturers/brands, 19 resellers, 56,360 product types, and 1,783,798 offerings to describe construction materials.

In this research, entities are used to represent core concepts within the domain of cost estimation, and the four base entities in the work item ontology are actor, knowledge, resource, and schedule. A framework of the work item ontology is shown in Figure 4. Actor contains organization and personnel. Organization is made of government and company, and personnel consists of designer, constructor, operational staff, and maintainers, among others. Every process or event in the construction that involves humans produces and updates of knowledge can be seen as work plans (e.g., schedule plan), work changes (e.g., change list), designs, bills of materials, and similar items [45]. Resource includes labor, equipment, material, and specifications in the design and construction process. The 41-Materials of OmniClass is selected as the classification standard for the building material. Schedule provides the scheduling arrangement of construction, including the actual schedule and the planned schedule. The above information is required for the work item ontology, and the resulting work item ontology aims to integrate the information concerning the construction documents,
developers, contractors, and facility management companies. The ontology can be built by experts’ analysis and information extraction. Therefore, it is a time-consuming job and a long-term process.

![Figure 4. Framework of the work item ontology.]

The relationships of the work item ontology use an existing relationship ontology, the GoodRelations [46] (hereafter GR) ontology, to describe the relationships of primary classes, subclasses, and individuals. The GR ontology is a standardized vocabulary for products, prices, stores, and company data that can be embedded into existing static and dynamic web pages and processed by other computers. In this research, GR is used to present a RDF (Resource Description Framework) description of the construction material and as a work item ontology extension for the terminology dictionary. The relationships in the work item ontology use parts of the GR ontology as reference, as shown in Figure 5.

![Figure 5. Part of the GR ontology.]

3.1.3. Construction Condition Ontology

Mapping is required between the ontologies of work items and construction conditions to realize an information description of a construction condition’s subcategory and subsection by means of the work item. To automate the mapping reasoning, this research proposes a construction condition ontology that consists of the specific classes that are used to select work items. Estimators select a...
work item that is suitable for the construction condition using a reasoning rule based on the expert’s experience (e.g., the construction condition of earth excavation in foundation pit engineering involves the work items of excavator, construction sequence, earthwork transportation). Construction condition can be extracted from BIM, the specifications of cost estimation, or historical data. In the process of cost estimation, all specialized data concerning work items may come from suppliers and/or BIM; for example, cost-related information about a specific material can be found in the resources of the work item ontology or in the documents provided by the supplier. The construction condition ontology, extracted from IFC and documents, is used to recognize BIM data and knowledge items and is converted into RDF(S)/OWL format. Knowledge item is a unit of knowledge; a knowledge item is the physical or symbolic manifestation of such knowledge. For example, a “schedule plan document” is the manifestation of the “schedule plan”. Knowledge item includes documents, CAD drawings, roles, records, reports, knowledge classifications, and communities of common interest in construction cost estimation [45,47].

3.1.4. Heterogeneous Information Integration and Semantic Disambiguation

Heterogeneous Information Integration

The information required for the cost estimation ontology has different sources, so effective heterogeneous information integration is an important issue. The goal of the Semantic Web is to provide tools that make it easier to integrate information across all formats, representations and schemas. RDF provides a flexible and extensible data model that eases the task of combining data sets into a common data model [48]. Because the machine cannot interpret the meaning of the XML (Extensible Markup Language) and IFCXML data exactly, the data need to be converted into machine-readable form by using XSLT (Extensible Style-sheet Language Transformations).

XSLT [49] can solve the problem of integrating heterogeneous data sources. XSLT is a language for transforming XML documents into other XML documents. A special advantage of this approach is that, although XSLT is written independent of any programming language, it can be executed by a program written in almost any up-to-date programming language [50]. Hence, source data can be converted into RDF data via the XML to RDF transformation engine of XSLT. RDF data can be expressed as a triple structure, that is, resource, attribute, and attribute value. The approach is capable of representing meaning of information both by concept and by ontology relationship. The data source transformation process is, however, outside the scope of this research.

Semantic Disambiguation

In practice, differences in language, personal habits, cultures, and customs have resulted in different description forms of concepts in cost estimation. For the same concept, different classification methods may have different expressions, such as abbreviations and synonyms. Particularly in China, semantic problems are more prominent due to the multiplicity of national cultures. Consider “挖土方” and “土方开挖” or “砼” and “混凝土”; the two pairs of concept have the same meanings, of earth excavation and concrete, respectively, in Chinese, so it is difficult for the estimator to address the problem of semantic disambiguation. To solve this problem, the ISO committee has proposed IFD (International Framework for Dictionaries) [51] to allow all of the information in IFC format to be tagged with a Globally Unique ID (GUID). However, the framework lacks cost estimation information. Thus, considering the viewpoint of engineers and estimators, this research defines a terminology dictionary based on GB 50500 and the CSI classification. Figure 6 shows the structure of the terminology dictionary. In the terminology dictionary, standard terminology and common terminology are matched by a terminology relationship. The various expressions of several records of the terminology dictionary are shown in Table 1.
work item, and work condition, can be built. The cost estimation ontology is essentially composed of the hierarchy of taxonomy. As mentioned above, relate to each other through the object properties. The example of object properties relationships are extracted by experts and are treated as explicit knowledge. Through the extracted information, related work item ontology uses OmniClass as the smallest unit of cost management, the collection of cost estimation information based on GB 50500, the work item ontology uses OmniClass and the existing ontologies as the reference standards. As the smallest unit of cost management, the proposed work item can be applied to each phase from the design process to operational stage. Thus, by using these three models in combination, the proposed cost estimation can be used in specific construction condition for certain circumstances.

The ontology model integration architecture based on the relationships among the classes of the cost estimation ontology is treated as implicit knowledge. The classes and instances and their relationships are extracted by experts and are treated as explicit knowledge. Through the extracted knowledge of building construction cost estimation, many significant classes, including concept models, work items, and construction conditions, and their subclasses are obtained in this research. The relationships among these classes are built to construct the cost estimation ontology. The classes, as mentioned above, relate to each other through the object properties. The example of object properties of the cost estimation ontology is shown in Table 2, where the domain and range of the object property relationship “controls” are restricted to two classes: “Actor” and “Resource.” Then, as shown in Figure 7, the general integration architecture of the cost estimation ontology, which includes concept model, work item, and work condition, can be built. The cost estimation ontology is essentially composed of the taxonomy of concepts and the taxonomy of relationships. These concepts and their relationships are extracted from specifications, standards, classifications, and IFC entities and are used to form the hierarchy of taxonomy.

![Figure 6. Structure of the terminology dictionary.](image)

<table>
<thead>
<tr>
<th>No. of Cost Item</th>
<th>Description 1 in Chinese</th>
<th>Description 2 in Chinese</th>
<th>Standard Description in English</th>
<th>Description 1 in English</th>
<th>Description 2 in English</th>
</tr>
</thead>
<tbody>
<tr>
<td>0105</td>
<td>混凝土</td>
<td>混凝土</td>
<td>concrete</td>
<td>conc</td>
<td>con’c</td>
</tr>
<tr>
<td>010502001</td>
<td>矩形柱</td>
<td>长方形柱</td>
<td>rectangular column</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>010809002</td>
<td>铝塑窗台板</td>
<td>铝塑板窗台</td>
<td>aluminum-plastic sill</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

3.1.5. Integration Architecture of the Building Construction Cost Estimation Model

By following the content, framework, and problems of the cost estimation ontology, this research shows an architecture of the integration of the above ontologies, i.e., the concept model ontology, work item ontology, and construction condition ontology. Among them, the concept model ontology is the collection of cost estimation information based on GB 50500, the work item ontology uses OmniClass and the existing ontologies as the reference standards. As the smallest unit of cost management, the proposed work item can be applied to each phase from the design process to operational stage. The construction condition ontology provides specific information extensibility in different states and situations of construction process. Thus, by using these three models in combination, the proposed cost estimation can be used in specific construction condition for certain circumstances.

3.1.5. Integration Architecture of the Building Construction Cost Estimation Model

By following the content, framework, and problems of the cost estimation ontology, this research shows an architecture of the integration of the above ontologies, i.e., the concept model ontology, work item ontology, and construction condition ontology. Among them, the concept model ontology is the collection of cost estimation information based on GB 50500, the work item ontology uses OmniClass and the existing ontologies as the reference standards. As the smallest unit of cost management, the proposed work item can be applied to each phase from the design process to operational stage. The construction condition ontology provides specific information extensibility in different states and situations of construction process. Thus, by using these three models in combination, the proposed cost estimation can be used in specific construction condition for certain circumstances.

3.1.5. Integration Architecture of the Building Construction Cost Estimation Model

By following the content, framework, and problems of the cost estimation ontology, this research shows an architecture of the integration of the above ontologies, i.e., the concept model ontology, work item ontology, and construction condition ontology. Among them, the concept model ontology is the collection of cost estimation information based on GB 50500, the work item ontology uses OmniClass and the existing ontologies as the reference standards. As the smallest unit of cost management, the proposed work item can be applied to each phase from the design process to operational stage. The construction condition ontology provides specific information extensibility in different states and situations of construction process. Thus, by using these three models in combination, the proposed cost estimation can be used in specific construction condition for certain circumstances.

The ontology model integration architecture based on the relationships among the classes of the cost estimation ontology is treated as implicit knowledge. The classes and instances and their relationships are extracted by experts and are treated as explicit knowledge. Through the extracted knowledge of building construction cost estimation, many significant classes, including concept models, work items, and construction conditions, and their subclasses are obtained in this research. The relationships among these classes are built to construct the cost estimation ontology. The classes, as mentioned above, relate to each other through the object properties. The example of object properties of the cost estimation ontology is shown in Table 2, where the domain and range of the object property relationship “controls” are restricted to two classes: “Actor” and “Resource.” Then, as shown in Figure 7, the general integration architecture of the cost estimation ontology, which includes concept model, work item, and work condition, can be built. The cost estimation ontology is essentially composed of the taxonomy of concepts and the taxonomy of relationships. These concepts and their relationships are extracted from specifications, standards, classifications, and IFC entities and are used to form the hierarchy of taxonomy.
Future Internet 2016, 8, 39

Table 2. Examples of object properties of the cost estimation ontology.

<table>
<thead>
<tr>
<th>Object Property</th>
<th>Domain</th>
<th>Range</th>
<th>Note</th>
</tr>
</thead>
<tbody>
<tr>
<td>isPartOf</td>
<td>IFC</td>
<td>BIM</td>
<td>The relationship between BIM and IFC.</td>
</tr>
<tr>
<td>controls</td>
<td>Actor</td>
<td>Resource</td>
<td>The relationship between actors and resources that it controls.</td>
</tr>
<tr>
<td>hasMaterialFeature</td>
<td>Building_Product</td>
<td>Material_Feature</td>
<td>The relationship between a building product and material feature.</td>
</tr>
<tr>
<td>hasOpening</td>
<td>Building_Product</td>
<td>Building_Product</td>
<td>The relationship between building products, and its space relationship between building product and building product.</td>
</tr>
<tr>
<td>decompose</td>
<td>Construction_Condition</td>
<td>Concept_Model</td>
<td>The relationship between a construction condition and the concept that it involves.</td>
</tr>
<tr>
<td>isSimilarTo</td>
<td>Building_Product</td>
<td>IfcBuildingProduct</td>
<td>The relationship between a building product and IFC entities IfcBuildingProduct</td>
</tr>
<tr>
<td>hasQuantity</td>
<td>Work_Item</td>
<td>Quality_Specification</td>
<td>The relationship between work items and quality specification.</td>
</tr>
<tr>
<td>responsibleFor</td>
<td>Actor</td>
<td>Knowledge</td>
<td>The relationship between an actor and what it is responsible for.</td>
</tr>
<tr>
<td>hasSectionShape</td>
<td>Building_Product</td>
<td>Geometry_Feature</td>
<td>The relationship between a building product and a geometry feature.</td>
</tr>
<tr>
<td>hasCostItem</td>
<td>Building_Product</td>
<td>Cost_Item</td>
<td>The relationship between a building product and a cost item.</td>
</tr>
</tbody>
</table>

3.2. Reasoning Mechanism

As an ontology representation language, OWL (Web Ontology Language) is a Semantic Web language that was designed to represent rich and complex knowledge about things, groups of things, and relationships between things. However, OWL has limits in the amount of expressivity that it can offer users. One solution to this issue is to keep the OWL recommendations unchanged and introduce the inclusion of rules to expand the expressive capability [48]. In addition, OWL can only represent property restrictions of a single individual, but the rules of the building construction cost estimation ontology always contain many individuals, which is beyond the representation capability of OWL. Hence, the reasoning rules should be added to the ontology reasoning. The reasoning mechanism developed in this research aims to reason about the cost estimation ontology by means of the domain rules expressed by SWRL [52]. SWRL (Semantic Rule Language), which is based on OWL, allows users to define rules to reason about OWL individuals and new knowledge [53].
According to the cost estimation ontology integration architecture, the classes and their relationships of the cost estimation ontology are related by SWRL. Axiom, the constraint condition of the class, is coded by SWRL and can then be reasoned with by the Jess inference engine. Axioms are used to assert conditions on classes to restrict the individuals that it contains or what properties its individuals can have. Through the definition of encoding regulation in GB 50500, the classes and properties of the cost estimation ontology are related by virtue of SWRL. The SWRL rule is shown in Rule 1.

Rule 1: Building_Product(?x) ∧ Material_Feature(?y) ∧ Building_Material(?z) ∧ havingMaterialFeature(?x, ?y) ∧ consistsOf(?y, ?z) ∧ includedMaterial(?x, ?z) → includesBuildingMaterial(?x, ?z)

In SWRL rules, “∧” is an operator for the logical AND between conditions, “→” is an operator for drawing the conclusion. Thus, the above rule means that if object property “havingMaterialFeature” of the building product “?x” is “?y”, and object property “consistsOf” of “?y” is material “?z”, then object property “includesBuildingMaterial” of the building product “?x” is material “?z” can be reasoned out.

3.3. Framework of the Building Construction Cost Estimation Based on Ontology

The overall framework for the automated reasoning of cost estimation is described in Figure 8. First, construction conditions are extracted from BIM or knowledge items, and BIM data are exported into IfcXML and converted into computer readable RDF data by the IFC-to-RDF converter service, such as the converter service provided by UGent MultiMediaLab [54]. Further, documents, CAD drawings, and photos are considered knowledge items and are converted into RDF data. By means of GB 50500, taxonomies of CSI, and existing ontologies, this research builds ontologies, i.e., the concept model ontology, work item ontology, and construction condition ontology. Second, reasoning rules are created by means of a reasoning process that is based on three ontologies and RDF data in the reasoning layer. Therefore, the knowledge base has all of the factual knowledge about the cost estimation, including the building feature, building product, and building material, and all of the inferred knowledge.

![Figure 8. Overall framework for the automated reasoning of cost estimation.](image_url)

4. Implementation of the Building Construction Cost Estimation Ontology

4.1. Software Environment

The construction of the cost estimation ontology is implemented by the ontology editor. This research uses Protégé, a free and open-source platform that enables the development of
domain models and knowledge-based applications with ontologies [55], to build the ontology [55,56]. The Protégé 3.4.8 (Protégé, hereafter), proposed by Stanford University, provides the environment for creating, editing, saving, and using ontologies. It is convenient to build the class and subclass to specify the class, its hierarchy, and semantic relationships.

The implementation of reasoning rules of this research has been realized by SWRL rules, which are a plug-in tag in Protégé 3.4.8. All related semantic rules are defined by the SWRL rules plugin. However, SWRL is a language that is independent of any inference engine and, hence, cannot be used directly without inference tools. OWL and SWRL need to be translated into a rule language that can be inferred. In this research, the reasoning process is conducted through the Jess rule engine. Jess (Java Expert Shell System) is used in Protégé as an inference engine, and it consists of a fact base and a rule base [57]. Jess’s reasoning is based on a list of known facts and a set of rules that are used to match these facts in the knowledge base [50]. Jess is currently the most mature SWRL engine, has been implemented in many knowledge-based systems, and is capable of dealing with reasoning rules that are written by SWRL. Jess can be embedded into Java applications and provides a flexible mutual runtime communication between Jess rules and Java. The reasoning is carried out in the Jess inference engine by matching facts with the rules in the rule base, as the extensions to Protégé, SWRLJessTab, and JessTab are successful implementations of Jess through SWRLJessBridge.

Based on the defined classes and properties, Protégé automatically generates a graphical user interface (GUI), as shown in Figure 9.

![Graphical user interface of Protégé](image)

**Figure 9.** Graphical user interface of Protégé.

In this research, the Jess rule engine converts the combination of OWL and SWRL into Jess facts. Because a Jess interpreter can only interpret Jess code, the RDF and OWL formats need to be transformed into Jess code. The proposed XSLT-based transformers, i.e., OWL2Jess and SWRL2Jess, can satisfy the integration of heterogeneous information among OWL, SWRL, and Jess. The flowchart of cost estimation inference is shown in Figure 10, the source language is shown at the top of each arrow, and the target language is shown at the bottom of each arrow. First, after the cost estimation knowledge is extracted, the hierarchy of cost concepts is defined according to, e.g., the Chinese national mandatory specification, the classification standard from CSI. The cost estimation model, including...
the concept model ontology, work item ontology, and construction condition ontology, is determined, and the related concepts and their relationships are built. Second, the above-listed ontologies are all expressed by OWL/RDF, and the reasoning rules among different hierarchies are encoded in SWRL. Third, OWL and SWRL-based knowledge is transformed into facts and rules that are expressed in the rule language of Jess. Thus, the knowledge about cost estimation can be deduced automatically by Jess.

Figure 10. Flowchart of cost estimation reasoning.

4.2. The Transformation System Framework of Cost Estimation

SWRL is a descriptive language that is independent of other rule languages and is encoded by XML grammar, which can be embedded in Protégé. The Jess rule engine is used in the integrated framework, and SWRLJessTab and JessTab have been successfully implemented as Jess extension transformers in Protégé. Transformer plugins can be used to transform the fact (OWL) and rule (SWRL rules) knowledge into the fact base and rule base encoded by Jess, and the transformed cost estimation model is implemented by the Jess inference engine. The transformation between them is summarized in Table 3.

Table 3. The transformation between OWL and Jess.

<table>
<thead>
<tr>
<th>OWL/SWRL</th>
<th>Jess</th>
</tr>
</thead>
<tbody>
<tr>
<td>Class, property, individual</td>
<td>Facts</td>
</tr>
<tr>
<td>RDF, OWL, SWRL semantics</td>
<td>Rules</td>
</tr>
</tbody>
</table>

The transformed detail is described as follows:

The classes of ontology knowledge base that are used to describe classes and subclasses. For instance, the cost estimation class and its subclasses are built and transformed into the following Jess templates: “owl: Thing” is the top class, and others are subclasses that are directly and indirectly derived from it:

(deftemplate owl:Thing(slot name))(deftemplate Cost_Estimation extends owl:Thing)(deftemplate Concept_Model extends Cost_Estimation)(deftemplate Building_Product extends Concept_Model)(deftemplate Beam extends Building_Product)

The instances of classes in the cost estimation model can be transformed into the Jess fact through transformers. For instance, as the instances of the class “beam”, “Rectangular_Beam” fact can be transformed into the Jess fact:

(assert(owl:thing(name Rectangular_Beam))) (assert(Cost_Estimation(name Rectangular_Beam ))) (assert(Concept_Model (name Rectangular_Beam ))) (assert(Building_Product (name Rectangular_Beam ))) (assert(Beam(name Rectangular_Beam )))
Relationships between individuals or between an individual and a data value represent an object property or data type property and need to be transformed into the fact-based return of the Jess knowledge base. For instance, the instance "Rectangular_Beam" has the object property "hasMaterialFeature" with the instance "Cast-in-Place_Concrete" of the class of "MaterialFeature", and it needs to be transformed into a Jess fact:

(assert(havingMaterialFeature Rectangular_Beam_of_Cast-in-Place_Concrete)

For the data type property, the transforming method between OWL and Jess is summarized in Table 4.

Table 4. The data type property transformation between OWL and Jess.

<table>
<thead>
<tr>
<th>Data Type Property in OWL</th>
<th>Jess Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>xsd: string</td>
<td>RU.STRING</td>
</tr>
<tr>
<td>xsd: boolean</td>
<td>The atoms “TRUE” and “FALSE”</td>
</tr>
<tr>
<td>xsd: decimal, xsd: double, xsd: int, xsd: long, etc.</td>
<td>RU.INTEGER, RU.FLOAT, RU LONG</td>
</tr>
<tr>
<td>ARRAY_XX_TYPE class, etc.</td>
<td>A Jess multifield</td>
</tr>
</tbody>
</table>

By using the SWRL2Jess transformer, the rules defined by SWRL are transformed into Jess rules in the rule base. For instance, the above-defined rule is transformed into a rule in the Jess rule base, as shown below.


5. Experiment of Cost Estimation Ontology Reasoning

5.1. Execution of Interation Reasoning

This research conducts experiments to demonstrate the proposed cost estimation ontology construction process and its application in an actual situation. Protégé is adopted as the experiment platform. The cost estimation ontology is built on the basis of Measurement Specification of Building Construction and Decoration Engineering in GB 50500, and it is necessary to examine the implementation of the cost estimation ontology.

In the experimental case, the beam of a building product is studied. The beam information from the GB 50500, construction document and construction company, is selected to build an ontology to verify the feasibility of the proposed cost estimation ontology. As shown in Figure 11, the new cost estimation knowledge of the beam can be entered using the ontology editor software Protégé.

When the material feature and geometry feature of the beam are set up in the cost estimation ontology, the building product represented by the cost item can be reasoned by the defined reasoning rule. For example, the cost item instance “ID_010503002” represents "Rectangular_Beam" by the property "hasBuildingProduct". Once the concept model ontology is triggered, Rule 2 is executed. The property "hasSectionShape" of a beam's instance "Rectangular_Beam" indicates the section shape of the beam. Then, the ontology model and SWRL are translated into the Jess knowledge base via OWL/SWRL transformers, and the reasoning is executed. Finally, through reasoning by the Jess inference engine and the Jess transformer, the result is obtained, as shown in Figure 12. The content in shaded parts is the related classes and instances in the reasoning process. The reasoning result shows “ID_010503002”, which represents "Rectangular_Beam_of_Cast-in-Place_Concrete".

Rule 2: Beam(?x) \ Material_Feature(?y) \ hasMaterialFeature(?x, Cast-in-Place_Concrete) \ hasSectionShape(?x, rectangular) -> representsBuildingProduct (id_010503002, ?x)
Rule 3: Beam(?x) ∧ Material Feature(?y) ∧ Building Material(?z) ∧ hasMaterialFeature(Rectangular_Beam, Cast_In_Place_Concrete) ∧ consistsOf(Cast_In_Place_Concrete, ?z) → includesBuildingMaterial(Rectangular_Beam, ?z)

In the cost estimation process, the building product can be a combination of multiple building materials, which is usually not known by professionals. The input building product and its material feature are matched to the list of established classes and instances to determine its corresponding building materials. The aforementioned process involves two ontology models: the concept model ontology and the work item ontology. Once the work item ontology and concept model ontology are triggered, Rule 3 is triggered to compare these two ontologies so the result can be determined (the reasoning result is shown in Figure 13). The property “hasMaterialFeature” of beam instance “Rectangular_Beam”, which indicates a rectangular beam of concrete, has the material feature “Cast-in-Place Concrete”; cast in place concrete includes materials, i.e., “Cement”, “Sand”, “Stone”, and “Water”; and then the reasoning rule is triggered. Next, the ontology model and SWRL are translated into the knowledge base. The Jess inference engine is run, and the result is returned to the ontology, as shown in Figure 13 (Reasoning result), which shows that the rectangular beam has the materials “Cement”, “Sand”, “Stone”, “Water”. The rectangular beam and the property are in the concept model ontology, and cement, sand, stone, and water are in the work item ontology, so the reasoning is realized by combining the concept model and work item ontologies.
To combine an ontology and BIM, the property “isSimilarTo” to can be used to map BIM element properties to corresponding cost estimation properties. Once the cost estimation is triggered, Rule 8 is executed. The property “isSimilarTo” of the beam, which indicates a beam of the building product, is similar to “IfcBeam” (1 in Figure 14.), but “IfcBeam” is part of Architecture in the BIM model (2 in Figure 14). Thus, the reasoning result (3 in Figure 14.) shows that the beam belongs to “Architecture” in the BIM model, and the reasoning rule can be modeled by the following SWRL Rule 4.

\[
\text{Rule 4: Building \_ Product}(?x) \land \text{IFC}(?y) \land \text{BIM \_ Model}(?z) \land \text{isSimilarTo}(?x, ?y) \land \text{isPartOf}(?y, ?z) \rightarrow \text{belongTo}(?x, ?z)
\]

Figure 13. Reasoning result of the beam-related building material.

Figure 14. Reasoning result of the cost estimation ontology.
5.2. A Case Application

The case application is to illustrate the cost estimation ontology on a real bill quantity of building construction. The proposed ontology is applied in the construction cost estimation analysis of the concrete structure engineering, which consists of 22 cost items.

Take ring beam, for example: the bill quantity item "Ring Beam of Cast-in-Place Concrete", coded as "010503004001", represents a set of beams that have the same "Material Feature" and "Geometry Feature" with a "Cast-in-Place Concrete" and "Ring" feature. After the decomposition, if building product is known to have the type of "Beam", material of "Cast-in-Place Concrete", and the section shape in geometry feature of "Ring", it can be classified into the bill quantity item of "Ring Beam of Cast-in-Place Concrete". In the cost item, the three optional numbers of specific ring beam changes with the specific building construction. The "Ring Beam of Cast-in-Place Concrete" has strength grade with a "C20", then it is coded in "010503004001". Accordingly, a reasoning process can be abstracted for this bill quantity item, as shown in Figure 15. The reasoning rule is as follows:

\[
\text{Rule 5: Building Product}(x) \land \text{Material Feature}(y) \land \text{Geometry Feature}(z) \land \text{Strength Grade}(a) \land \\
\text{hasMaterialFeature}(	ext{Concrete Beam, Cast-in-Place Concrete}) \land \text{hasSectionShape}(	ext{Concrete Beam, Ring}) \land \\
\text{hasStrengthGrade}(	ext{Concrete Beam, C20}) \land \text{Building Component}(b) \rightarrow \\
\text{hasSubCostItem}(	ext{Ring Beam C20 of Cast-in-Place Concrete, ID_010503004001})
\]

Figure 15. Reasoning process of bill quantity item “Ring beam C20 of cast-in-place concrete” codes as “010503004001”.

The cost of work items is a relatively important indicator in the process of cost estimation, and it includes the price of resource, material, labor, construction equipment, enterprise management fee, and profit. Estimators start to price each work item by using the construction company’s historical records, which include cost information from previously completed projects. Therefore, many companies have established unit cost databases to support the query, calculation, and update of unit cost. Unit cost changes continuously, but the current estimating applications do not have the built-in capability to update their unit cost databases. Therefore, estimators have to obtain the latest construction costs information and update the unit cost databases manually, which is a time-consuming process. Providing direct access to the latest construction costs would eliminate the need for manual updating and has the potential to improve estimating efficiency [58]. Estimators can extract the quantities of work items from the cost estimation ontology, and the unit cost can be assigned to a work item by the unit cost database.

The reasoning results can be applied in the process of cost estimation, as shown in Figure 16. The reasoning process is reflected in ① and ② Estimators need to extract the unit cost from the unit cost database ③ and find the quantity of work item from drawing or BIM ④. Then, unit cost is combined to work item quantity ⑤ to form the bill quantity.
6. Conclusions

The total lifecycle cost management of a project is of the greatest importance, and the automation of extraction and reasoning information for cost estimation is necessary and remains a challenge.

In this research, we propose our research on the formalization and organization of cost estimation by means of ontology and this research is mainly focused on: hierarchy of building construction cost estimation concept; definition of OWL axiom and SWRL rules; and integration of heterogeneous information. This work is constructed in the context of the Chinese national mandatory specification GB 50500 and includes two aspects: (1) the ontology representation model including concept model ontology, work item ontology, and construction condition ontology; (2) the development of reasoning mechanism that automates the reasoning if a building construction cost estimation satisfies a set of these reasoning rules. Cost estimation examples are given to prove the feasibility and usability of the developed ontology model. The proposed method eliminates the need for manually updating a cost database.

The related works of this research are summarized as follows:

1. The hierarchical classification and reasoning rules are defined on the basis of GB 50500 and the terminology standard. This research solves the problem of localized standards by combining GB 50500 and the internationally accepted taxonomy standard.
2. The cost estimation ontology has the advantages of both consistency and clarity in its structure and concepts of expression. OWL can perfectly represent the knowledge of building construction cost estimation, and the ontology is convenient for sharing and reusing building construction cost estimation knowledge.
3. Reasoning rules can be described by SWRL in the building construction cost estimation ontology. Through the combination of SWRL and Jess inference engine, the reasoning results of cost estimation can be reasoned automatically.
4. This research demonstrates that the proposed ontologies and semantic reasoning rules can be utilized to facilitate cost estimation of concept and taxonomy structures in the proposed ontology.
5. Due to the powerful expansibility of the cost estimation ontology model, the cost information and rules in this proposed method can be continuously updated and expanded in practice, and automatic reasoning and integration of information can be realized.

In addition, developing a mature building construction cost estimation ontology is a cumbersome task, and the theory and technology of the existing cost estimation model based on ontology are imperfect. Hence, a comprehensive comparison and analysis with the existing related ontologies and some related methods is outside the scope of this research.

However, there are also some limitations of the proposed method that need to be studied in the future. First, there are still some limitations regarding ontology representation and reasoning in the proposed concept model ontology, work item ontology, and construction condition ontology. Due to the limitation of considering all possible construction conditions, which differ depending on the construction phase, the ontology model established in this research is simple; therefore, the
classes and their instances do not fully reflect the characteristics and their effects in actual cost estimation. Second, the study on building construction cost estimation ontology is still in its infant stage; thus, the number of classes and the relationships between classes are not sufficient to reflect the systematic and complex nature of cost estimation, and the cost estimation ontology needs to be improved continuously by collecting related information in the future. Finally, this research has not collected sufficient information to evaluate the proposed method, and there are not abundant data used for valid proof. The future research will adopt theoretical and empirical, qualitative and quantitative methods to evaluate the performance of the proposed method.

Acknowledgments: This research was financially supported by National Natural Science Foundation of China under grant number 51178084.

Author Contributions: Xin Liu contributed to the design of the proposed approach, and conducted analysis, discussed the results and implications, wrote and commented on the paper at all stages. Shaohua Jiang contributed to refinement of the proposed approach and the advancement of the paper. Zhongfu Li supervised the research activities.

Conflicts of Interest: The authors declare no conflict of interest.

References


