38   BUILDING AND CONSTRUCTION EXTENSIBLE MARK-UP LANGUAGE (BCXML): THE C2B / B2C SCENARIO

Frits Tolman[*], Reinout van Rees[*], Michel Böhms[**]
*:Delft University of Technology (NL) **:TNO Building and Construction Research (NL)

Abstract

The EU IST-10303 "eConstruct" project is developing an XML Vocabulary and Grammar for the Building and Construction (BC) industry, together with a number of applications of this vocabulary. The vocabulary - called bcXML (from <u>b</u>uilding and <u>c</u>onstruction <u>XML</u>) - supports eCommerce and eBusiness in BC, both nationally and over the borders of the different European member states. Especially the communication of <u>meaning</u> over the national borders is of crucial importance for the future usage of the Internet as a means to increase the industry's competitiveness, its ability to co-operate, reduce cost of failure, and stimulate the European market of BC products, services and business.

This paper presents the current (March 2001) status of bcXML; what it looks like, why it looks that way, and what the BC industry can expect of it. Special emphasis of the paper is on the Consumer to Business and Business to Consumer (C2B / B2C) scenario.

Keywords: Communication, XML, Supply Chain, Internet, European R&D Project

INTRODUCTION

With the new XML [xml] based Internet, meaningful electronic communication will swiftly become popular in the Building and Construction (BC) industry. The reasons are: (1) the new Internet's speed and security will improve much over the current Internet and (2) it will become possible to eliminate the currently needed human information-transformation role that is hindering the information flows (humans interpret documents, extract computer input, run applications, extract output, create new documents).

With XML (eXtensible Mark-up Language), documents will no longer be inaccessible for computers. An XML file is a text file, containing information on a certain topic (like "*I sell wooden doors 2 metres in height*"), tagged with sort-of nametags that explain what's meant with the text. To say that "*height*" in above example is a property, you would do it like this in an XML file: `<property>height</property>`. Text is the only truly portable data, binary formats being inherently more difficult to use (ever tried importing a Word document using another program?). XML files, containing all the information you need, can be directly inputted into an application and also quite easily formatted for the human eye, either as a web page or in print. So, XML covers both the human and the computer side of information exchange.

The new Internet is seen as an opportunity to improve electronic communication in BC. Therefore, the European (IST 10303) eConstruct project [econstruct] (eCommerce and eBusiness in Building Construction: preparing for the new Internet) works hard to develop, implement, apply, demonstrate and disseminate an XML vocabulary and grammar for BC, called *bcXML*.

The eConstruct project aims to elaborate a number of use case scenarios, one of which focuses on C2B / B2C. Other use case scenarios focus on Computer Aided Selling, integration with design/engineering (through a common object model shared with the IAI-IFC) and on project support. In this paper we will explain bcXML and its role in the Building and Construction (BC) industry from the viewpoint of individual end-user, i.e. focusing on the C2B / B2C scenario.

CONSUMERS AND SUPPLIERS

Computers and Internet play an increasing role in the consumer - supplier relationship in BC. Many suppliers advertise their products in web pages and often provide electronic catalogues with product information. A more recent development is the advent of portals, commercial intermediaries that bring suppliers together and provide improved access to supplier information. Though professionals in BC are often supported by national organisations that may provide lists of web addresses of suppliers and portals, it is still very difficult to find what is needed. For individual consumers (house fathers) it even seems quite an impossible task. Finally, if we look at this problem from an international point of view, we see that it is nearly impossible to use the Internet to find and sell construction products and services over the national borders.

Figure 1 illustrates the current electronic consumer - supplier relation where consumers try to find and access supplier product information through web pages and portals.
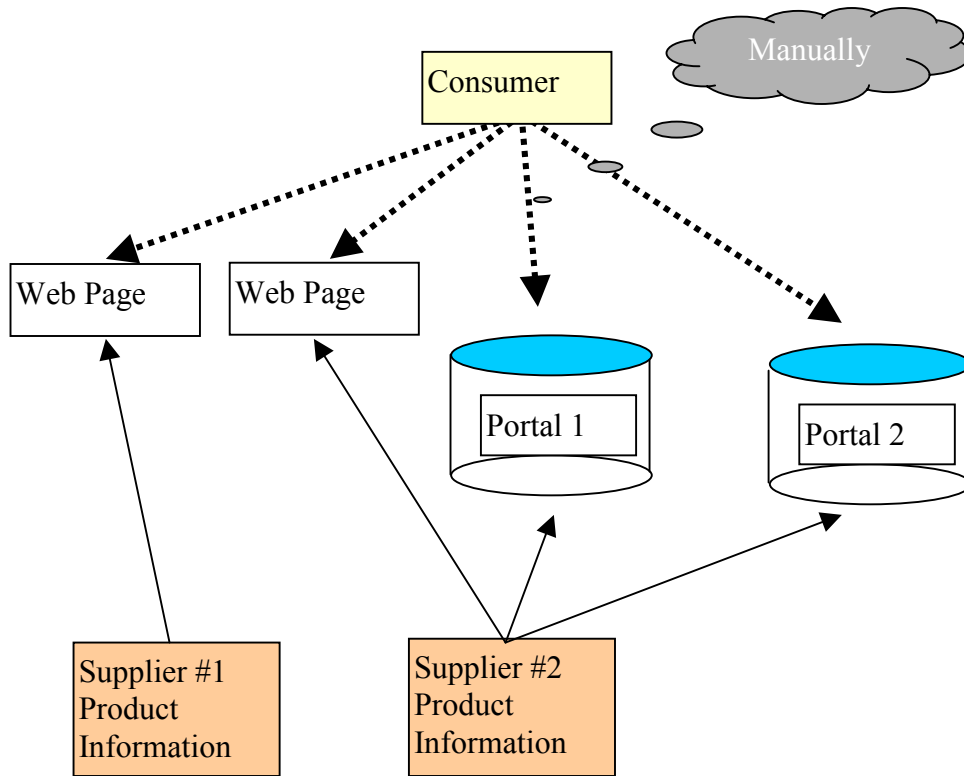

Figure 1. Current situation. Suppliers create web pages from their catalogues and give the same information to Portal organisations that provide (some) added value in PR. Consumers searching for electronic product information have to find suitable suppliers and search manually in web pages to browse electronic catalogues, or search in various Portals doing basically the same.

The main problems with today's (partly) electronic Information System (IS) that is replacing the traditional paper based IS are:

- It is still quite difficult for the consumers to find what they need as no search engines really support this effort
- Suppliers have to maintain their catalogue information in several places in several formats
- Insufficient support is available for finding the right Supplier, especially over the national borders
- International standardisation in this area is not coming off
- Portals normally charge money to the suppliers and often have limited added value
- Portals have different naming conventions and data structures making it difficult for Suppliers to automatically keep their-data-at-the-portal up to date and making it difficult for Consumers to search over different Portals

ECONSTRUCT

In an effort to contribute to the solution of this problem the eConstruct consortium [www.eConstruct.org] is developing bcXML as a neutral language for meaningful communication in BC over the Internet. Figure 2 below shows the role of bcXML as *initially* seen.

Consumer

GUI
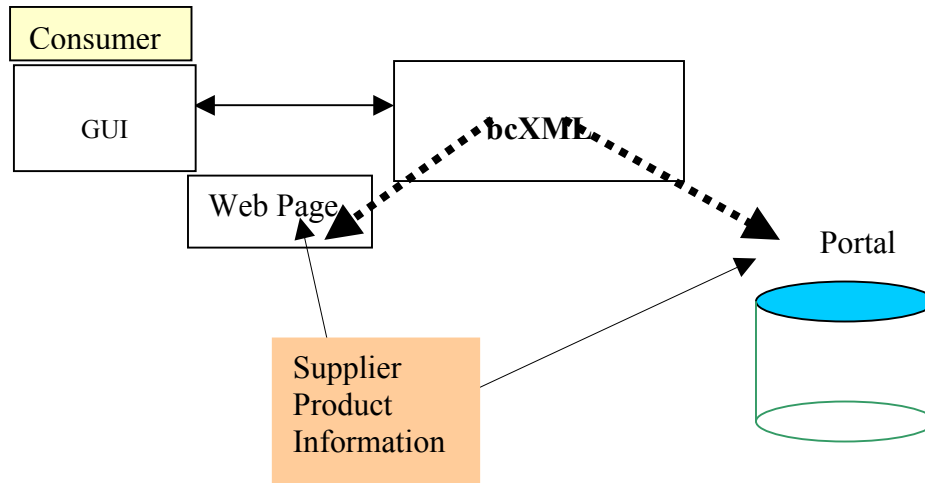
bcXML

Web Page

Portal

Supplier
Product
Information

Figure 2. New situation. Consumers can use any portal to find products of their choice, indirectly searching all suppliers. The portals take care of reading the supplier's bcXML files (available over the Internet) Using bcXML means that this will also work over the national borders.

The suppliers make sure their product information is available in bcXML format. They can do it themselves and put the file somewhere on the Internet in a findable place. They also can use an firm specialised in product information collecting. There are all sorts of possibilities (and business opportunities!) here.

Portals and other organisations can get the data when it is available. This is different from the normal procedure of having the supplier put the data in; and having to pay for it. But at the same time, the costs of maintaining the data fall on the supplier himself (unless he pays you to maintain his data!). And each portal can access every supplier (unless he blocks it one way or another), gaining a massive increase in supplier data. Basically, Portals now can make a living providing *service*, the one with the best service, the one who filters out bad companies, etcetera, can make the most money.

The consumers can use their favourite portal and user interface to search all suppliers, using terms for objects and properties defined by eConstruct, in multiple languages. Expert users can have their own tools directly query the bcXML files of the supplier's. Once the data is universally readable and available, an increasing number of applications can be performed.

How eConstruct initially hoped to achieve all this can be seen in Figure 3. It shows that bcXML basically contains three components (that are relevant for C2B / B2C): bcDictionary, bcTaxonomy[1] and the bcXML meta model that defines the bcXML language syntax.
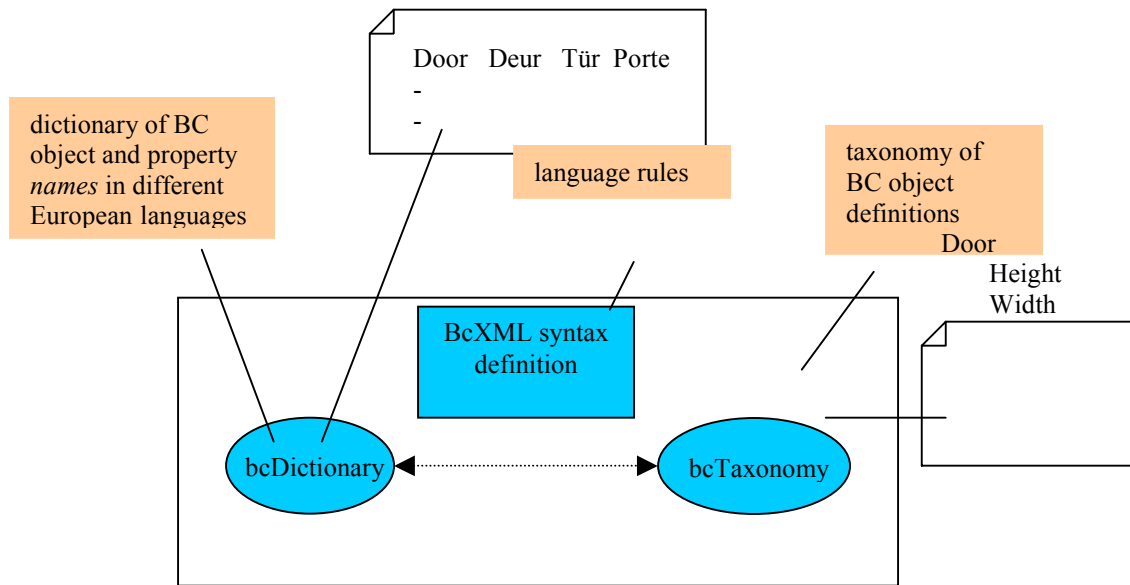

Figure 3. How eConstruct was meant. The bcDictionary contains a set of object and property names in several European languages. The bcTaxonomy defines a subset of properties for the objects in the bcDictionary (i.e. not all the possible properties of an object are defined, only the most important). The bcXML meta model defines a suitable syntax.

The reason for splitting the semantics into a dictionary and a taxonomy were:

- Now two different hierarchies can be used (alphabetic in the dictionary, subtype hierarchy in the taxonomy)
- Division of roles (translation versus definition).
- People might like to use the dictionary with a different taxonomy than the bcTaxonomy (which might be handy in a changing world and might present an opportunity to companies that want to build additional functionality into their taxonomy).

The different components will be explained in more detail below. (Remember, this is our *initial* viewpoint).

---

[1] The word 'taxonomy' is used here in a generic sense for 'a classification-neutral structured collection of objects and their properties/units and such'. There are many other words with overlapping meaning like ontology, lexicon, thesaurus that could have been used instead.

BCDICTIONARY

The bcDictionary is simply an electronic dictionary of names of objects (components, materials, services, etc.) and names of properties, that play a role in eCommerce in BC. The names are translated in several European languages. The bcDictionary names thus become the 'neutral' semantics for meaningful Internet communication. This does not mean that everybody from now on has to use these words to denote the products that they want to sell or buy, but that a mapping from the neutral object name to another prefered object should be provided in case different words are being applied.

It is obvious that eConstruct will not be able to provide a complete bcDictionary in all the European languages. That is simply too much work. What will be done is the development of a limited set of words in a limited set of languages in an open system that can be extended by others with additional words and additional translations.

BCTAXONOMY

More or less the same is true for the bcTaxonomy which describes the objects in relation to their properties in a classification- or view-neutral way.

The role of bcTaxonomy is to provide a basis for agreement while communicating about components, systems, equipment, materials, work methods, services and all that. A Column in bcXML is that object as defined in the bcTaxonomy. It's got the name "Column" (and a few others, as "kolom") and has properties which are needed to describe a column. Not having a central, neutral "something" to identify a "something" makes communicating much more difficult. Again this is not to prescribe to the industry how all these objects should be defined, but as a neutral definition that can be changed or extended. But this then requires some explicit way to describe that in a formal mapping.

RETHINKING THE ARCHITECTURE

After nearly a year it became clear that the architecture as shown in figure 3 was not going to materialise. Basically the problem is that BC and related fields communicate about a very large amount of objects and properties, divided over a great number of sectors and branches. Current estimates about the number of objects are in the 60.000 to 100.000 range, and for the properties in the 1000 range. The number of sub-sectors that play a role are at least of three types. First the sub-sectors within the main sectors Building and Civil Engineering: Office Building, Housing, Tunnelling, Roads, Railroads and such. Second the discipline oriented sectors like HVAC, Durability and Steel that play a role in both main sectors and Electrical and Mechanical that play a role within BC, but also outside BC. And third the related sectors like Offshore, Power and Process Plants, which have considerable overlaps with BC. Moreover these sectors are not precisely the same in different countries, as much depends on local circumstances like temperatures, soil conditions and history. Moreover even inside BC different sub-sectors use different names for the same objects, or the same words for different objects. This realisation, plus the fact that other industries and other projects that face the same problem follow other

roads, made us come back on the decision to develop *the* standard BC taxonomy. Not our turf, too much work, too little funding, manpower, and of course time[2].

The main focus of eConstruct therefore shifts to bridging out to other sectors and sub-sectors that are trying to pursue similar goals and see if bcXML can be used as a mechanism to provide meaningful communication over these worlds. The solution is that the new architecture for bcXML does no longer contain *the* standard bcTaxonomy, but allows everybody to develop and use **bcXML compliant taxonomies**. BcXML compliancy means that the taxonomy returns information in bcXML format and that it implements the interface defined by eConstuct. Figure 4 shows the revised architecture.
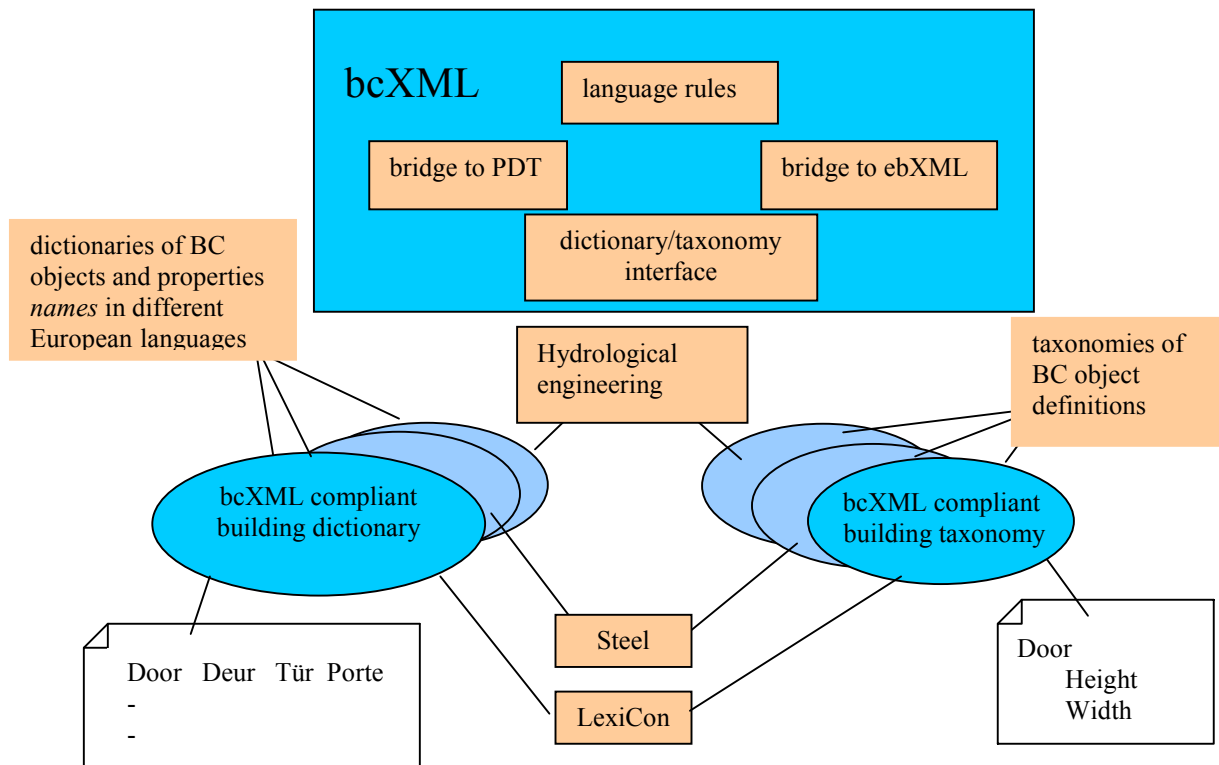

Figure 4 New architecture with external bcXML compliant taxonomies developed for the different sub-sectors of BC, and maintained by the classification and information institutions and companies.

The dictionary/taxonomy interface for the C2B / B2C scenario shown in figure 4 is using the Web Services Description Language (WSDL), an XML schema published as W3C Note [wsdl]. WSDL currently supports communication using SOAP, HTTP and MIME.

---

[2] The Commission routinely brought the original project duration of three years back to two years.

BUILDING TAXONOMY

One partner in the project, the Dutch Specification Institute STABU [stabu], is developing a first bcXML compliant taxonomy, called the LexiCon, partly as a test case and partly with commercial interests. The project also produces a simple example taxonomy that will be available free of charge for general use. The STABU is very active in the ISO/DPAS 12006-3 effort ("a framework for object-oriented information exchange in BC") and they are cooperating with a number of international partners in order to build momentum for... THE Building Taxonomy! It will not happen within eConstruct's timeframe, but once it is more or less completed and becomes available it will probably quickly and painlessly replace the other solutions for general use.

THE META-MODEL

The bcXML meta-model is written in UML (Unified Modelling Language "Class Diagrams"). It is designed to be simple for the end-users. They can use their own language and dictionaries, even for communications over the borders. Simple also, in that all information is basically expressed in *objects* which have *properties*. "*I want a table with a height of 0.70m.*" Advanced users or software applications will be provided with more complex functionality. The model presented is the official model (with a few boxes deleted to fit on the page). This UML model is in its turn is transformed into an "XML Schema" to make it comprehensible for various XML tools. The main schema that describes the bcXML syntax is presented in figure 5.
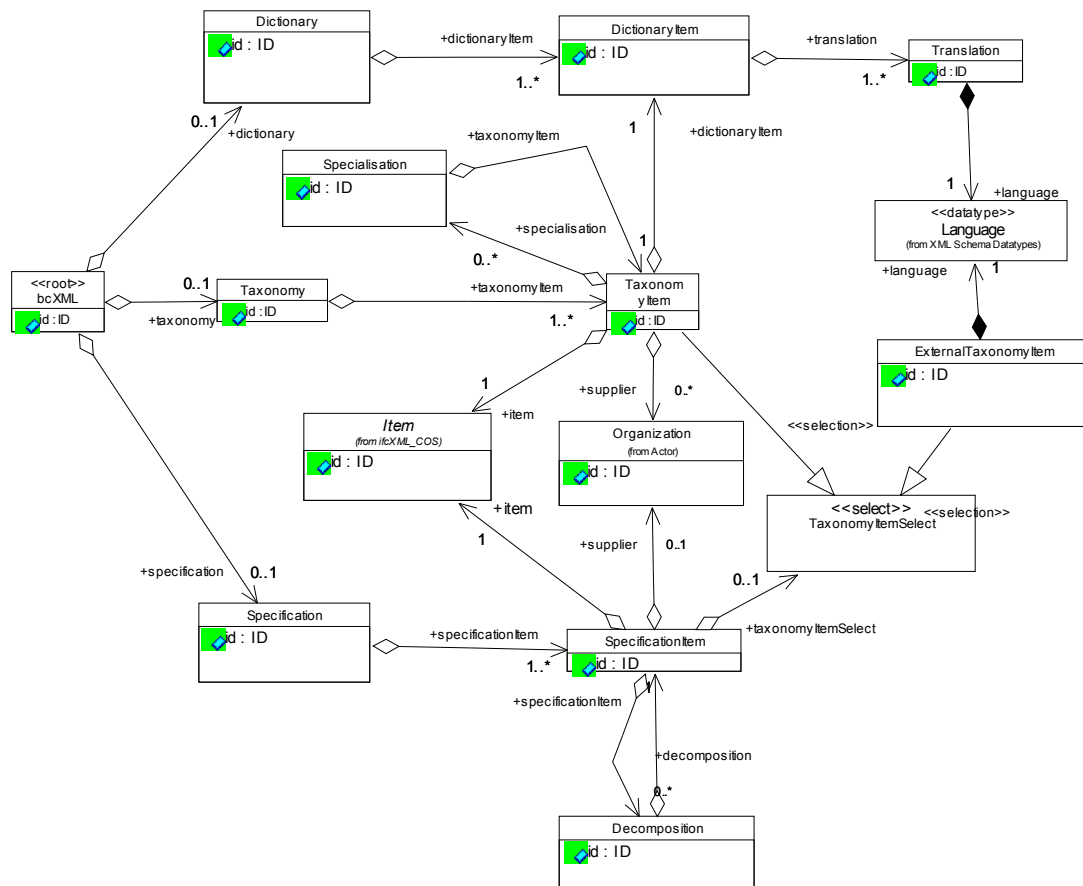
Figure 5. The main UML diagram that defines the bcXML syntax.

eConstruct shares it's core with IFC, both have the jointly developed Common Object Schema (COS). This part allows us to talk about properties, objects, values, etc.
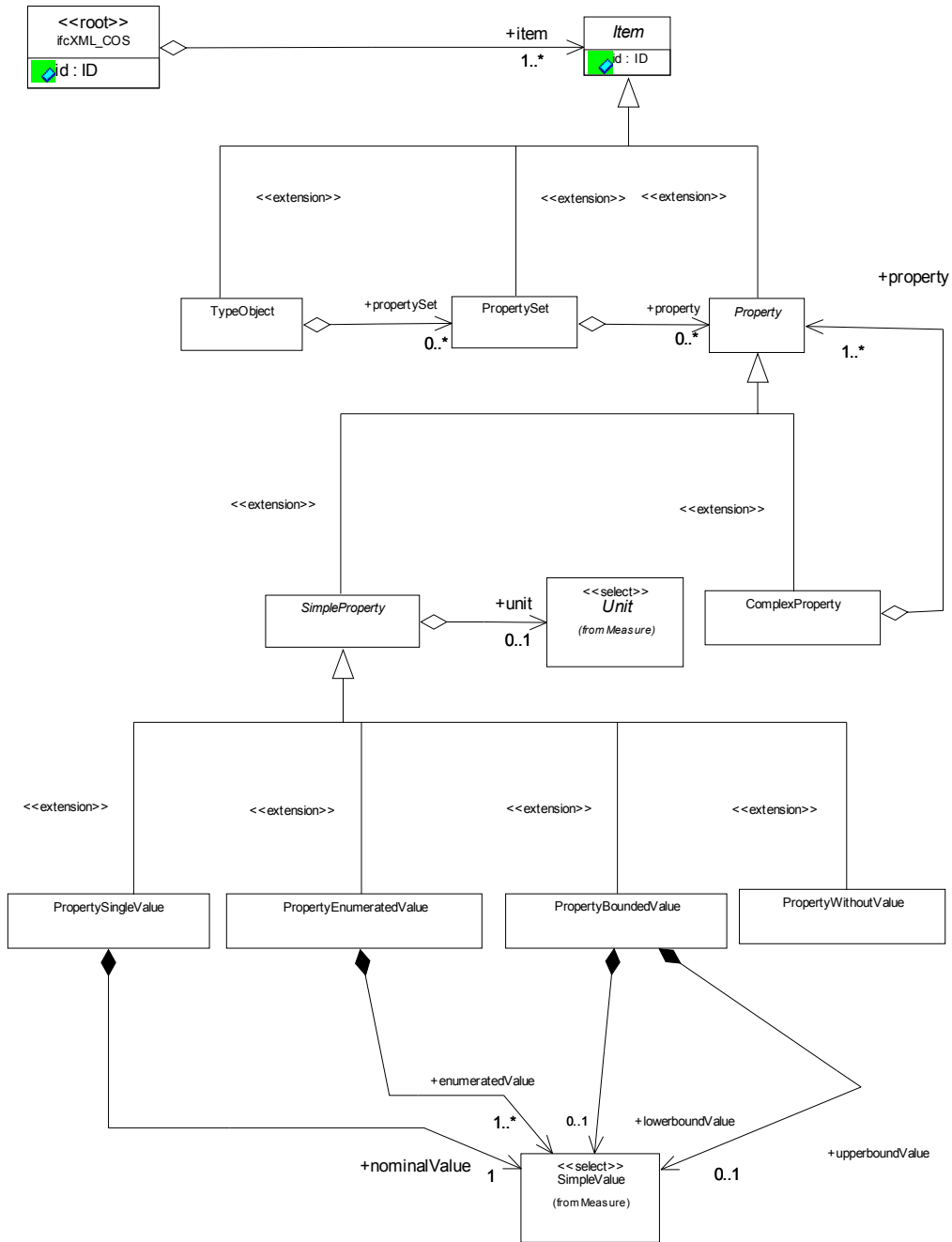

Figure 6. UML diagram of the Common Object Schema (COS), which describes the object, property, value level of bcXML. It is also used in IFC.

It should be noted that there are also other diagrams for other purposes that cannot be explained in this paper. Again, see our website [model].

CONCLUSION

The paper presents and discusses the first working version of bcXML from the Consumer to Business point of view. The XML based communication language bcXML can be used to improve Internet communication in the European Building and Construction industry. In the taxonomy area eConstruct will not produce one giant multi-lingual, multi-sector taxonomy. Instead the idea is that companies and institutions in the classification and information business in various sub-sectors of BC will be able to develop bcXML compliant taxonomies (or provide bcXML interfaces for existing taxonomies) that are better tailored to the needs of different countries, niche markets and consumers in each of the many sub-sectors of BC, providing the means to communicate electronically over the borders and over the sectors.

As to the translation function eConstruct will make a start with an open multi-lingual bcXML compliant taxonomy containing the names of objects and properties (not the links between them). This dictionary will be in the public domain and should be seen as part of the bcXML communication standard.

The eConstruct project will also produce a bcXML compliant building taxonomy, called the LexiCon. The LexiCon is a taxonomy for the Building industry under development by the Dutch STABU. For demonstration purposes eConstruct will also provide a small taxonomy in the public domain. Other taxonomy owners will shortly join in the game, so for not too long bcXML will become the BC communication language for the Internet, at least that is what we hope.

REFERENCES

[econstruct]   http://www.econstruct.org
[xml]   http://www.xml.com
[stabu]   http://www.stabu.nl
[model]   http://www.econstruct.org/6-Public/bcXML-MS-CD_v1.2.1-description.doc
[wsdl]   http://www.w3c.org/TR/2001/NOTE-wsdl-2001-20010315