A Knowledge Representation Format for Virtual IP Marketplaces

Martin Schaaf, Andrea Freßmann, Marco Spinelli, Rainer Maximini & Ralph Bergmann
University of Hildesheim
Data- and Knowledge Management Group
PO Box 101363
31113 Hildesheim, Germany

{schaaf | fressmann | spinelli | r_maximi | bergmann }@dwm.uni-hildesheim.de
www.dwm.uni-hildesheim.de

Abstract: For the design of *Systems on Chip* (SoC) it is essential to reuse previously developed and verified *virtual components* in order to meet nowadays requirements in reliability, correctness, and time-to-market. On the downside, deciding about reusing a third-party component in a design situation can consume substantial time and resources. This is especially true in situations where many potential candidates exist and due to the large amount of functional, non-functional, and quality related aspects of each component. In order to facilitate the search for components in an internet-based market scenario, we have developed a retrieval system that utilizes structural CBR. The approach relies on XML descriptions of each component that constitute the case base. In this paper we present IPCHL, which was originally intended to be a simple schema for such descriptions. As a consequence of the growing demands for structured documentation, IPCHL has become a representation format for virtual component descriptions incorporating many of the CBR knowledge containers.

1 Introduction

The notion of *Electronic Design Automation* (EDA) has been assigned to the current academic and industrial research activities that tackle the development of new methodologies, standardized description techniques, and tools for the design of microelectronic circuits. An important area that depends on efficient EDA is the design of so-called ***S**ystems **o**n **C**hip* (**SoC**). SoCs combine multiple functionalities on one chip and they are the building blocks for a wide range of consumer electronics devices such as mobile phones or DVD players. Nowadays, SoCs are designed by reusing already existing modules, which are called *Virtual Components* or *Intellectual Properties* (IPs). SoC designers do no longer rely only on their own developed components but also on IPs offered by external vendors. Companies like sci-worx are even specialized in providing IPs for functionalities of general interest like audio decoding or error correction. The amount of IPs offered in the World Wide Web grows daily and the task of searching for potential IPs suitable for a given design situation can be very time consuming. This is especially due to the huge amount of functional and non-functional aspects that have to be taken into account. In order to allow virtual marketplaces to emerge, a standardized IP documentation is required, which can be read by humans and processed by computers. Furthermore, there

must be a tight integration of computer support into the development methodology of designers. This is the aim of the project *IP Qualification* (IPQ), an EDA project founded by the German Ministry of Education and Research (BMBF) and initiated with several industrial partners. Within IPQ, we have developed an intelligent retrieval approach that supports designers in their search for IPs. The approach utilizes structural CBR by comparing the designer's situation with the characterization of IPs. The basic retrieval strategy combining functional, non-functional, and quality related aspects of IPs has been described in [Sc02b]. In the following paper we present the IPQ market scenario based on web services that is the environment of the retrieval application. The focus lies on the concepts of the **IP C**haracterization **L**anguage (IPCHL) that was designed for representing IP characterizations and motivated by CBR-specific requirements. Because of the growing demand from the industrial partners for standardization of all IP related information, newer releases of IPCHL incorporate many of the CBR knowledge containers providing the necessary semantic for communication and cooperation between the different services within the virtual marketplace for IP.

2 The Virtual Marketplace Scenario of IPQ

A System on Chip typically combines several functionalities on one integrated circuit, for instance a SoC may contain a java execution engine together with an audio decoding unit. In contrast to traditional integrated circuits, which are individually designed and highly optimized, the focus of SoCs lies on high integration and cost reduction. Such aspects are prevalent for consumer electronics like mobile phones. Nowadays, technology allows SoCs to become very complex and, due to the immense functionality aggregated on one chip, verification is time consuming.

Figure 1 IPs Integrated on a *System on Chip*

Very similar to current advancements in software engineering, a possible solution for SoCs is to reuse already existing and fully verified components so called *Virtual Components* or *Intellectual Properties (IPs)*. IPs are specifications in a hardware description language like VHDL or Verilog and one distinguishes between *Soft IP*, *Firm IP*, and *Hard IP*. The difference between Soft IPs and Hard IPs lies in the degree of flexibility and, therefore, reusability. While a Soft IP is a high-level specification of the intended functionality independent from the technology (e.g. CMOS) of the resulting SoC, a Hard IP is specifically designed for one target technology. Everything between Hard and Soft IP is called Firm IP. Figure 1 shows an example of a SoC that has been produced from a specification incorporating IPs from multiple vendors.

2.1 Architecture of the IPQ Virtual Marketplace

In order to support designers for selecting and integrating IPs into their actual design, the IPQ Virtual Marketplace for IPs offers a variety of web services for purchasing IPs as well as transferring between different IP repositories (see Figure 2). A so-called *IP Basar* currently under development can broker request from the design gate to registered services.

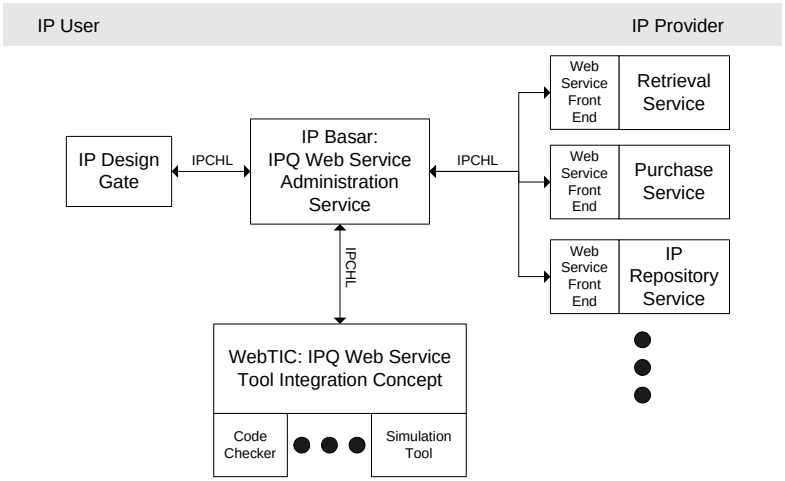

Figure 2: IPQ Web Service Architecture

The *WebTIC* interface connects also web service unaware applications, e.g. code checkers or a simulation tool.
All of these services in Figure 2 utilize XML formats for the IP itself (content) and its description or characterization. The XML application for the IP Characterization is named IPCHL. Because it has been derived from requirements of CBR-based retrieval services it is a structured representation format for the various CBR specific knowledge containers.

3 Representing Knowledge For CBR-based IP Retrieval

When applying the structural CBR approach, knowledge items are described by a *characterization* constructed from a previously developed domain vocabulary. State-of-the-art CBR systems make use of an object-oriented vocabulary representation [Ma94; AP95]. Such representations can be seen as an extension of the attribute-value representation. They make use of the data modeling approach of the object-oriented paradigm including *is-a* and other arbitrary binary relations as well as the inheritance principle. In the domain of IP retrieval, each IP is represented by an XML characterization that acts as a semantic index or semantic markup. The set of all characterizations constitutes the case base and, because that is in fact a set of XML files, it is possible to transfer each characterization across the Internet by utilizing standard web communication facilities like *web services*. That was an important requirement from the market scenario described above.
When using a modern CBR retrieval engine like the Open Retrieval Engine orenge from empolis Knowledge Management GmbH [Em01], cases can be accessed from various data sources, e.g. relational databases, but the domain vocabulary itself is represented in a proprietary format, typically, only of interest for the engine itself. For the aim of IPQ it was a basic requirement to have an explicitly specified vocabulary because it contains much important knowledge about the domain that is of interest for other services as well. For example, a *catalog builder* can use the vocabulary for rendering a human readable document from the IP characterization.
Another requirement from our application scenario was a certain degree of flexibility because, with respect to the current state of EDA research, it is not possible to develop a single problem oriented domain vocabulary sufficient for all IP vendors. The situation very much resembles the problems currently tackled by Ontology-based Knowledge Management approaches where a proper and explicit description of all kinds of knowledge is nearly as important as the problem solution itself [BS03]. Hence, our objective was to develop an explicit XML-based representation of a CBR domain vocabulary tailored for IP retrieval specific needs that is a conceptualization of each particular IP characterization. From a slightly different perspective, such a vocabulary defines the primitives and the structure of each IP characterization. In the following we will present the **IP C**haracterization **L**anguage (IPCHL) that has been developed as a knowledge representation language specific for CBR applications within the EDA domain. We start with the first release of IPCHL that was basically for domain vocabularies. We will then show further enhancements of IPCHL that capture other knowledge types of the CBR knowledge container model as well.

3.1 The Initial Version of IPCHL

As mentioned before, the initial release of IPCHL has been developed with the aim of being an explicit CBR vocabulary representation tailored for IP retrieval and is extensively described in [Sc02a]. In the following, we will give only a brief overview of the language and the rationales behind.
On a very high level, two different attribute types constitute the set of potential properties for the characterization of IPs:

- *Application attributes* that refer to properties important to decide about the applicability of an IP in a given design situation
- *Quality criteria* that characterize the IP and its deliverables according to its quality.

Both types are subject to current standardization efforts of the VSIA (Virtual Socket Alliance). For the application attributes, the document „Virtual Component Attributes (VCA) With Formats for Profiling, Selection and Transfer Standard Version 2 (VCT 2 2.x)“ [VS01] presents a variety of attributes together with proposals for their syntactical representation. For the quality criteria, the decision about several standardization proposals donated by the VSIA members is still pending. A candidate, for example, is the OpenMORE Assessment Program from Synopsis [Sy01]. For the characterization of IPs by elementary properties it is essential to be compliant with the standards of the VSIA. It will ensure a wider acceptance because it is expected that tools like design checkers will commit to these standards, too. However, because of the fact that the standardization is still in progress, it is necessary to rely only on the stable parts and allow enhancements as required. Hence, a major part of IPCHL is an XML Schema definition of both catalogues as depicted in Figure 3.

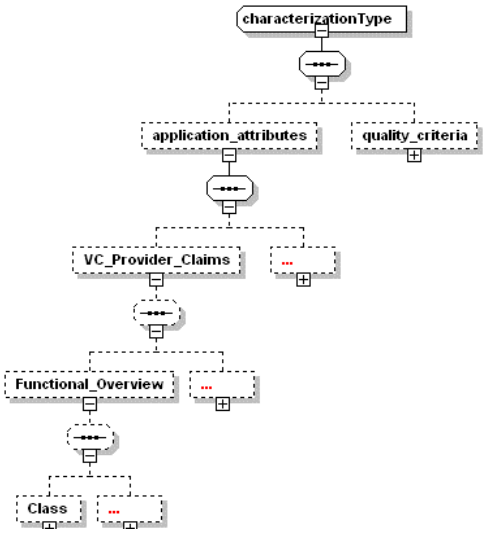

Figure 3: Excerpt of the initial XML Schema of IPCHL

For unique identification, each attribute is referenced by a complex name that is a path of the XML document tree. The complex name is built from the name of the attribute itself and a sequence of categories, which are counterparts of the corresponding sections of the VSIA respective OpenMORE catalogues [Sy01; VS01]. For example, the fully qualified name of the attribute *class* in Figure 3 is *CharacterizationType/Application_Attributes/-VC_Provider_Claims/Functional_Overview/Class*. Note that these categories structure the various attributes but do not classify the particular IPs. They are much like packages in UML that serve the purpose to organize the different elements of the model (e.g. diagrams, classes etc.) but should not be confused with the hierarchical structure of the model itself.

Following the CASUEL approach [Ma94], a CBR domain vocabulary is a hierarchy of classes that is built from a taxonomical property elicited from the domain of discourse. Such a property becomes a classifier. In our application scenario, it is a property from the IP characterization. For example, a potential candidate for this is the functional class taxonomy defined in [VS01] by the VSIA. As with typical object-oriented modeling approaches, the classifier determines the set of remaining properties relevant for a corresponding instance. E.g. an attribute like *sampling frequency* does only apply to IPs of the class *audio converter* respective its subclasses. In fact, an attribute for itself has no meaning without the context given by the class for which they apply. This was the crucial point when developed the domain vocabulary. Because each industrial partner had his own directives for structuring his IP asset, it was simply not possible to fix a particular taxonomy within IPCHL as a standard. Hence, we specified only XML constructs for defining taxonomies and building associations between individual taxonomy nodes of and subsets of the other properties. Consequently, the checking for compliance of particular cases (IP Characterization and IP Content) as instances of the vocabulary could no longer be delegated to standard XML Schema parsers. The combination of the XML Schema specification of IPCHL and a concrete taxonomy makes the CBR domain vocabulary, which is also named an *IPCHL Profile* or, for emphasizing that it is a format for the IP characterization, *IP Characterization Schema*.

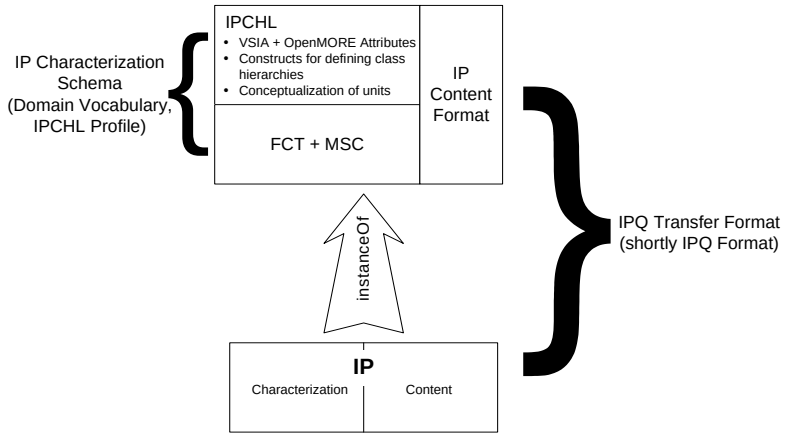

Figure 4: IPCHL Container Overview

Figure 4 provides an overview of the different types of knowledge involved in the IPQ market scenario. It furthermore shows the role of IPCHL within the scope of the *IPQ Transfer Format*, that is the set of all formats, which must be transferable across organizational boundaries. Because the taxonomical properties *functional class taxonomy* (FCT) and *market segment classification* (MSC) are explicitly mentioned in the standardization catalogue of the VSIA [VS01] they are provided as default IPCHL profiles. The IP Provider deploying the IP retrieval application can define other taxonomies as classifiers. In addition, IPCHL provides a conceptualisation for units used within the IP context e.g. *mV* for power consumption or *Hz* for frequencies. The IP is an instance and consists of the

characterization and the content. From the perspective of traditional CBR systems, the IP is the case with the characterization as problem description and the content as solution.

3.2 Extending IPCHL toward a flexible IP Knowledge Representation Language

With the initial release of IPCHL only the domain vocabulary and the IP characterization are explicitly represented. Although the remaining knowledge types of the CBR container model, e.g. similarity measures, provide highly valuable knowledge as well, they were only internally available in the retrieval engine we used for the CBR application (orenge, see [Em01]). Beside a growing demand from the industrial partners to have at least the similarity measure explicitly represented, it also became clear that the set of properties for IP characterization identified by the VSIA [VS01] was not sufficient for specialized application domains e.g. error correction. Hence, it was required that properties can be defined on behalf of the IP Provider.

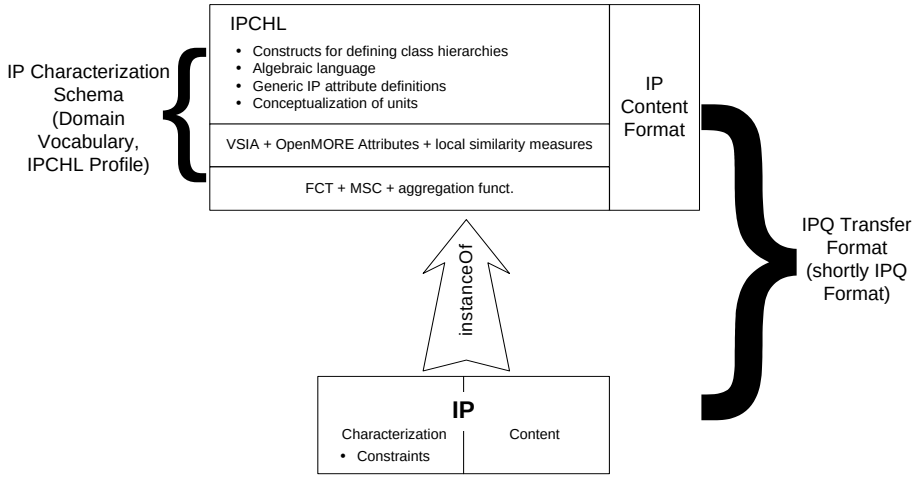

Figure 5: Extended IPCHL V2 Knowledge Containers

As shown in Figure 5, the new release of IPCHL now only contains a set of generic attribute type definitions, constructs for defining class hierarchies, a conceptualisation for units, and an algebraic language specification for defining similarity measure, and constraints used in generalized cases (see below). The VSIA and OpenMORE attributes, which were first-class citizens of the initial IPCHL release, are now refinements of these attribute types provided as a default IPCHL profile that can be enhanced by IP Provider specific profiles. Attribute Types refined in an IPCHL profile must define a local similarity function by making use of the algebraic language specified in IPCHL. Depending on the particular set of attributes, profiles containing the taxonomies that provide the hierarchy of the domain mode. Again, the functional class taxonomy and the market segment classification are provided by default. In addition, each taxonomy node specifies an aggregation function for the attributes associated to the particular concept. At the time of writing, only simple aggregation function utilizing weighted sums are defined. IPCHL knows of the following attribute types:

- ValueType: ValueType is a basic attribute type.
- SingleValueType: Based on ValueType and a description element.
- IntervalType: An attribute of this type implies an interval with boundaries “Maximum” and “Minimum”.
- IntervalWithTypicalValueType: This type is composed of an IntervalType and a SingleValueType that captures a standard value.
- TaxonomyType: When referencing a taxonomy this type is used. It contains one or more elements called “Node” by which taxonomy paths can be selected.
- TaxonomyDependendType: This type is composed of one or more SingleValueTypes and optionally makes possible to define additional taxonomies that are not standardized before.
- PickListType: Here an element of pick list can be declared. The pick list itself is defined in the Meta model.
- CategoryNodeType: This type specializes category types.

Finally, we had to represent so-called *parametrizable IPs* (PIPs), which refer to a kind of IP that has a certain degree of flexibility for improved reusability. Parametrizable IPs can be seen as generalized cases because they do not cover only a point of the case space but a whole subspace of it [MB02]. Following the approach given in [MB02], generalized cases are specified by a set of constraints that span the subspace of a generalized case by defining the functional dependencies between parameters. Although generalized cases are still subject to current research, we had to take precautions for defining constraints on the IP characterization instance level

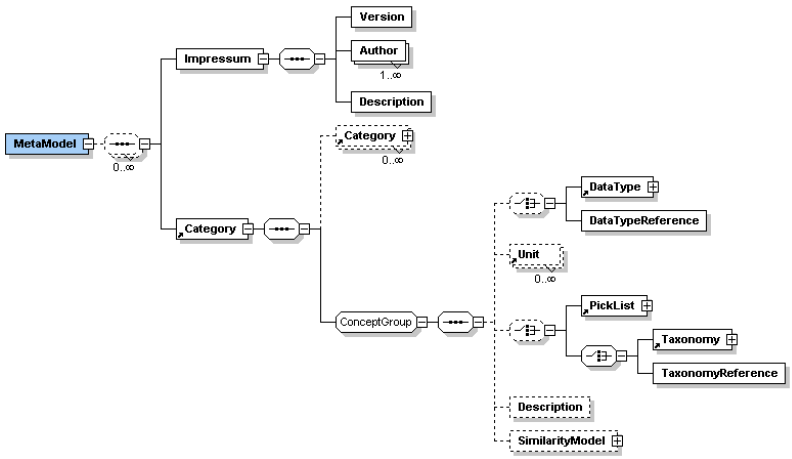

Figure 6: Excerpt of the new XML Schema of IPCHL

Figure 6 gives a brief impression of the XML Schema defined behind IPCHL. As shown, each IPCHL profile already contains information about the author and comes with a version number assigned. This is useful for tracing the history of a profile. Furthermore, it is necessary information for distinguishing different profiles in a distributed setting.

3.3 Future Extensions of IPCHL

As mentioned before, the purpose of IPCHL is not only restricted for knowledge representation, it is furthermore a part of the contract (or protocol) between different web services of the IPQ market scenario affecting other project partners as well. From the beginning, we have been in a strong cooperation with our industrial IPQ project partners and they contributed to the following tasks:

- Elicitation of retrieval relevant attributes and their weights as feedback from daily practice.
- Visualization of the retrieved IPs
- Feedback for attributes and taxonomies

A first result from the cooperation was the necessity for proprietary attribute or taxonomy beside the IP attributes identified by the VSIA und OpenMORE. Both standards are simple catalogues not sufficient for automatic reasoning about IPs. Consequently, they contain only an enumeration of attributes and lack the necessary structure. Typically, IP providers have a large amount of information about IP, which depends on the functionality, not covered by the standard attributes. The demand for proprietary attributes resulted in the new release of IPCHL. Because it no longer relies on already existing standards, knowledge acquisition from the industrial partner became even more important. Therefore, a java-based IPCHL editor has been developed that facilitates the development of IPCHL profiles as well as the editing particular IP Characterizations.

Future extensions of CBR-based IP retrieval should cover some kind of explanations because designer typically want to double-check retrieval results. It has been shown that under certain circumstances retrieval results are totally unexpected when the intended relevance of designers does not match the pre-defined weights of the IP vendor. For that reason, an explanation component will be integrated that can give very useful information if slight changes to the weight model would result in a totally different retrieval sets. A first prototype of an explanation module utilizes a sensitivity analysis interpreting the similarity model. By achieving this, users can interactively play with the retrieval system and test the influences of similarity modifications to the result set. Very related to this, is the general requirement for user specific weight models. The representation of user specific models as well as their integration in IPCHL will be an issue for the next release.

4 Related Approaches

In the following we present two related approaches to SoC development support in general. Both have not been chosen arbitrarily because they represent the complementary directions of current EDA solutions. The first approach comes from the company Design & Reuse [DR03] that hosts an Internet portal with IP retrieval functionality for the IP community. The second approach is from Synchronicity [Sy03] and provides solutions for design collaboration and design management in order to speed the development of SoCs. Although both companies are targeting different markets, they offer support for the integration of IPs into SoC, one of the objectives of the IPQ project.

4.1 Design And Reuse

Design And Reuse (D&R) hosts an Internet portal and offers several services related to IP reuse. D&R claims to be the world's largest directory of silicon IP and SoC design platforms. The D&R approach is a web-based solution with access to registered users. The aim of D&R is to facilitate the search for IPs by distinguishing different catalogues:

- Silicon IP/SoC: Search for silicon IPs
- Verification IP: essential companions to virtual component or IPs in order to verify IPs
- Software IP: IPs ranging from embedded OS to communication stacks and application software.
- IP Search/Find Club: The worldwide place to trigger most strategic IP business deals.

The first three of these different IP catalogues allow a keyword search with or without an additional guidance by functional taxonomy.
Under certain conditions, the search process is facilitated by some optional attributes like IP type (soft, firm, hard, model), technology (ASIC, FPGA), and verified technology. Unfortunately, these attributes are not represented later in the result set that only contains the block name and provider.
The IP Search/Find Club is different. This kind of search facilitates strategic IP business deals. Large system houses and customers are allowed to post their IP demands.
For composing an IP request users have to fill out a form with item, functionality, integration requirements etc. D&R forwards incoming requests to qualified IP providers who can contact the customer directly while being tracked by D&R. Without any pre-selection, this can cause a lot of work for providers processing the requests. Unfortunately, D&R has no information about if the pre-selection is done manually or automatically. Finally, D&R supports IP users in finding experts like providers, appropriate tools or other general information about IPs.
Compared to the CBR-based IP retrieval approach of IPQ, the D&R search mechanisms neither utilize intelligent retrieval techniques nor is it based on structured data. The combination of text-based search combined with the functional taxonomy that acts as a decision-tree leaves much of the work to the searching users. However, as long as the amount of functional similar IPs does not grow too large, the search facilities seem to be sufficient.

4.2 Synchronicity

Synchronicity provides solution for deploying IP design methodologies including multi-site engineering. Hence, the focus is on products enabling team communication, data sharing, and third-party tool integration. Team communication comprises sharing of ideas, bug reports, and engineering change information in distributed environments. Synchronicity's solution for IP searching is IP Gear that incorporates an IP catalog and as a helpdesk application. IP Gear splits into two suites. The Publisher Suite for moving design related information within and between companies. This suite provides a comprehensive infrastructure for minimizing design chain latency across the enterprise. The Catalog is a server that manages IP information by representing web pages. The helpdesk works with

past solution to answers requests. The Consumer Suite provides IP retrieval by connecting to suppliers using the Publisher Suite. Furthermore, this suite possesses the same access to the helpdesk as the Publisher Suite. Consumers are assured to retrieve the latest and correct IP versions that include updates, notifications and incremental releases. Consumers can search for IPs by browsing through a component hierarchy or by using a key word search. Results are presented as links to their web pages. Similar IPs are indicated and their differences are visually highlighted. Although Synchronicity provides solutions for nearly steps of a design flow, IP search facilities are restricted to text-based search or navigation in taxonomies. A structured representation of the IP documentation, which is typically created during IP design, is not considered.

5 Conclusion and Future Work

In this paper we presented IPCHL, a flexible and extensible representation format for IP-specific knowledge. Originated from a CBR application that enables efficient and intelligent support for IP selection in an internet-based market scenario, IPCHL meets the requirements to become a standard committed by a variety of services. This is typically not easy to achieve because the different CBR knowledge containers do not distinguish between ontological and problem-specific knowledge. Beside the domain vocabulary underlying a structural CBR application, other knowledge types like similarity models can be either problem or domain specific. In our approach we reflected the different scopes by distinguishing between local similarity measures for generic types defined by IPCHL itself and local similarity measures defined within an IPCHL profile.

IPCHL may become a language that supports distributed CBR approaches in the near future. As a declarative representation language it can become the starting point of interesting new approaches like CBR applications tightly integrated into the IP design workflow.

Other future work refers to explanation rules that are considered to become the next major enhancement of IPCHL. Explanation rules play an important role in our future knowledge acquisition scenario where important feedback from the users will be used to improve the retrieval results.

References

[AP95] Arcos J., Plaza E. (1995). *Reflection in NOOS: An object-oriented representation language for knowledge modelling.* In IJCAI-95 Workshop on reflection and meta-level architecture and their applications in AI.

[Be96] Bergmann, R., Wilke, W., Vollrath, I. & Wess, S. (1996) Integrating general knowledge with object-oriented case representation and reasoning. In: H.-D. Burkhard & M. Lenz (Hrsg.) *4th German Workshop: Case-Based Reasoning - System Development and Evaluation*, Informatik-Berichte Nr. 55, Humboldt-Universität Berlin, 120-127.

[Be99] Bergmann R., Breen S., Göker M., Manago M., Wess S. (1999). Developing industrial case-based reasoning applications. LNAI 1612, Springer, 1999.

[BS03] Bergmann, R., Schaaf, M. (2003). On the Relations between Structural Case-Based Reasoning and Ontology-based Knowledge Management, accepted for German Workshop on Experience Management, March 2003.

[DR03] Design & Reuse Homepage. http://www.us.design-reuse.com/, last visited 01/27/2003.

[Em01] empolis knowledge management GmbH (2001). The orenge Framework - A Platform for Intelligent Industry Solutions Based on XML and Java. Whitepaper, empolis knowledge management GmbH, Kaiserslautern.

[Ma94] Manago, M., Bergmann, R., Wess, S., Traphöner, R. (1994). *CASUEL: A common case representation language.* ESPRIT Project INRECA. Deliverable D1, University of Kaiserslautern.

[MB02] Mougouie, B., Bergmann, R. (2002). Similarity Assessment for Generalized Cases by Optimization Methods. In: S. Craw & A. Preece (Eds.) European Conference on Case-Based Reasoning (ECCBR'02). Lecture Notes in Artificial Intelligence, Springer, 2002.

[Ri95] Richter, M. M. (1995). *The Knowledge Contained in Similarity Measures.* Invited talk at the First International Conference on CBR (ICCBR-95).

[Sc02a] Schaaf, M., Visarius, M., Bergmann, R., Maximini, R., Spinelli, M., Lessmann, J., Hardt, W., Ihmor, S., Thronicke, W., Franz, J., Tautz, C., Traphöner, R. (2002). IPCHL - A Description Language for Semantic IP Characterization, Forum on Specification & Design Languages, September 2002.

[Sc02b] Schaaf, M., Maximini, R., Bergmann, R., Tautz, C., Traphöner, R. (2002). Supporting Electronic Design Reuse by Integrating Quality-Criteria into CBR-Based IP Selection, Proceedings 6th European Conference on Case Based Reasoning, September 2002.

[Sy01] Synopsis Inc., Mentor Graphics Corporation (2001). OpenMORE Assessment Program for Hard/Soft IP Version 1.0, http://www.openmore.com.

[Sy03] Synchronicity Homepage. http://www.synchronicity.com/, last visited 01/27/2003.

[VS01] VSI AllianceTM, Virtual Component Transfer Development Working Group (2001). Virtual Component Attributes (VCA) With Formats for Profiling, Selection, and Transfer. Standard Version 2.2 (VCT 2 2.2).