Abstract. This paper presents a conceptual graph formalism called simple graph boolean formulas that extends the \textit{SG-family} with boolean connectors. This formalism is used to define categories of objects in a classification service that can be turned into a legal content management system. We define the \textit{SGBF-model} of graph boolean formulas, present two decidable fragments of this formalism (relying on the first order logic BSR and guarded fragments), and describe the functional architecture of a generic classification service that can be used in the legal domain.

1 Introduction

Since the production of knowledge has substantially increased in the last previous years, organisations intend to mutualize knowledge or know-how of their members in systems that process data at a semantic level. This evolution has given rise to knowledge engineering, which aims at eliciting the semantics of data by way of a knowledge representation language controlled by ontologies which are semantic referentials that define the meaning of the language symbols.

Mondeca is a software publisher that is developing ITM (Intelligent Topic Manager), a knowledge management tool based on these principles. The core of this software is a three-level knowledge base (see. Figure 1):

- the highest level is a meta-model of all ITM knowledge bases. It consists of a representation ontology reflexively specifying the semantics of all representations used by ITM. It is common to all customers;
- the intermediate level is composed of models that specify the vocabulary used to describe customer data managed by ITM; it generally comprises a domain ontology describing the conceptual vocabulary used to annotate the content of the customer’s resources, some (representation) ontologies defining primitives related to specific data structures enabling ITM services (primitives for representing the thesaurus, logical organization of documents or collections of objects called categories);
- the lowest level contains the instances: annotations describing the content of information managed by ITM (data, documents), terminological resources (thesaurus) used to index this information, description of the logical organization of documents, hierarchies of categories, etc. This level is managed by the customer.
ITM provides several services as: indexing of documents from the thesaurus, creation of annotations controlled by a domain ontology, semi-automatic building of document annotations, semantic navigation and search features through the annotation base. The rising number of customers in various fields (as tourism, terminological management, military, etc) requesting knowledge quality control, content enrichment or classification features led Mondeca to provide ITM with several reasoning services. As the knowledge representation formalism in ITM is quite similar to labeled graphs, Mondeca chose the conceptual graph formalism (see [7] and [8]) and especially the $SG$-family [3] to define these reasoning services. The $SG$-family provides ITM with a formal semantics in logic, inference rules and constraints, graph-based reasoning operations sound and complete with respect to the logical deduction (see [5] and [10]), and combinatorial-based efficient algorithms for these graph operations. The reader is referred to [4] for a complete description of ITM, its knowledge representation formalism and its relationships with the $SG$-family.

One feature increasingly requested by customers is the ability to classify knowledge in hierarchically organized categories. This is particularly needed in the legal domain and especially by law publishers whose main activity is to publish the last up to date version of various regulations in reference books. Throughout the past few years, law publishers sought to diversify their offer by providing legal expertise services. A part of this diversification strategy relies on three other key actors of the legal domain: (1) the lawmaker whose work consists in maintaining the law up to date, (2) the individual which is expected to respect the law and (3) the compliant-seller whose work consists in selling products that must be compliant with the law. For each of these, law publishers intend to provide law management or expertise services: (1) regulation management services to the lawmaker, (2) triggering services to alert the individual when his personal situation is concerned by a change in the law, and (3) search features for...