Dynamic Models - A case study for developing curriculum regulation and conformity using Protégé

Mike Hobbs, m.h.w.hobbs@apu.ac.uk
Department of Computer Science, APU, 1 East Road, Cambridge

Abstract

This paper reports on a case study that seeks to use a document driven information system to validate organisational business rules as they change. Protégé is being used to create a supporting ontology for an information system that will help students select appropriate modules to create valid degree programs. A knowledge base will be constructed from documents marked up using RDF compliant XML[10], so that changes made to the documents can be incorporated into the information system. The work is in its initial stages and this paper reports on progress relating to identification of requirements and how they may be supported by an ontology backed knowledge based system.

1. Introduction

In most organisations many different departments, management levels and committees decide decisions effecting operations. Ongoing changes to business rules are typically disseminated via, emails, memos etc. but can usually be traced back to some kind of formal documentation. Keeping track of the current business model in this situation becomes difficult but semantic mark-up of important elements can be utilised by knowledge management and agent based systems to validate changes and automatically disseminate information to the appropriate personnel.

The motivation for the current work is that there is a bewildering choice of tools, systems technologies and methodologies which can be used to provide the kind of functionality we would like to see from a sophisticated integrated information system. We propose that by implementing a system as a case study we can shed light on some of the choices that need to be made and evaluate the practical benefits of a particular approach. A not inconsequential side effect will be a kernel of a system to provide information to students across a range of academic related issues.

In the case study we will develop a 'bare bones' system that will illustrate the needs of a practical system from initial data preparation through to the delivered services for the information users. We are implementing two competing models, a traditional RDBMS approach (the database system), and a document driven knowledge-based approach that uses an ontology to support intelligent software agents (the knowledge based system). It is clear that there are considerable overlaps between the two approaches - one of the key areas of this research is to evaluate the benefits of a knowledge based approach against any extra costs compared with the database approach.

2. Context

2.1. The semantic web

The W3C[1] Web Ontology Working Group [1] released a Working Draft of requirements for the Ontology Web Language (OWL) 1.0. This is part of a larger effort on the part of W3C to improve the semantic content of the web [2] and is central to a number of standards and

[1] W3C is the leading international standards organisation for the internet.

research projects. It has been heralded as the key to unlocking the secrets of the web. Automated tools will be able to use common sets of terms supplied by a domain specific ontology to power services such as more accurate web search, intelligent software agents, and knowledge management. In essence this is going to the road on which agent based systems will use to access, combine and summarise information from a wide variety of sources [3], [4]. Application areas include every area where information from a variety of sources needs to be gathered, assessed and utilised for analysis and decision making. The Working Draft [1] lists the following case studies:

- Corporate Web site management
- Design Documentation
- Intelligent Agents
- Ubiquitous computing

To which can be added applications in Bio-Informatics and Medical Informatics [7].

3. The application domain - student module choice

The application domain needed to exhibit the following characteristics

- Small scale - with clear system functionality.
- Limited semantics - easy to develop a consistent ontology.
- Well structured - to allow automation in the application of rules etc.
- Moderately complex - needs to justify knowledge-based approach.
- Moderate level of change - so that ability to change semantics can be evaluated.
- Well understood - need to be able to design functionality and evaluate outcomes.

3.1. Module choices

An application domain that suits all these requirements is the selection of modules by students to make valid programmes of study. APU operates a modular system with two semesters a year. Each semester students have to be registered for appropriate modules according to the rules laid out for their degree programme. Typically, there are a number of compulsory modules and a number of options with a variety of constraints, some dependant on the degree type, others depend on the needs of individual modules. For example, a particular module may be compulsory for a particular degree and require students to take pre-requisite modules.

3.2. Current problems and limitations

The issue becomes more complex when students change the nature of the degree they are taking or modules themselves are changed or no longer run. Information about the options and choices that are valid for a particular programme are published in a document, by the managing 'field' (a field is loosely equivalent to a department). This can cause difficulties for students with programmes combined from two or more managing fields. Even for single honours students the process of deciding what modules to take and ensuring both a valid and interesting programme of study can lead to problems. Currently the information system tracking student progress has a limited ability to validate module choices and communicate with students. The current problem is twofold - getting information about module constraints to students so that they can make informed choices, and making sure that once they have made their choices they comply with the constraints for their degree programme.

3.3. Solution requirements

An ideal system would need to provide the following functionality and properties:

- Information on module and programme constraints.

- Validation of module choices against programme constraints.
- Verification of choices to students and their tutors.
- Minimal impact on existing administrative processes.

Of all these requirements the last one probably presents the biggest technical problem - it is easy to provide a sophisticated information system if you can implement new ways of working, provide training and appropriate backup. The rules and regulations and basic module information are all currently held and maintained in paper documents. It is likely that these will be replaced or supplemented by web based documents but it is unlikely that a dedicated data management system will be employed to represent the relations between modules in a machine readable format. However, it should be possible to use semantic mark-up[10] in the creation of web pages that will allow the relevant information to be used by a knowledge based system that can provide the necessary functionality. The hope is that enough of the application domain remains static enough to allow minimal changes to the fundamental structure of the KBS and that the day to day changes in information are automatically incorporated.

4. Outline system design

The scope of the current module choice development is considerably smaller than that of an entire system as described in section 3.3 above. The idea is to develop a system to provide a test platform on which to prove the concept of the knowledge-based approach.

4.1. The knowledge based model

Domain	**Component**	**Description/ implementation**
Information source	Documents	module definition forms, field rules, university guidelines etc
Knowledge Acquisition	Mark -up	Semantic mark-up using RDF / OWL
Knowledge Acquisition / Knowledge Engineering	Ontology	Class hierarchy using developed and maintained by Protégé [1]
Knowledge Engineering	Rule base	CLIPS or JESS[6]
Knowledge Engineering / HCI	Interface	Agent based system using JADE [4]

Table 1 Components of the knowledge based system

Documents + Mark-up +Ontology + Rule base + Interface (agents) = Knowledge Based Information system

The core components that make up this model are outlined in Table 1 with a simplistic but descriptive formula that relates them. As the project develops we may alter the implementing technology but the type of component will remain the same. A major main concern for this project will be the integration and creation of applications to interface between them these components. The key is to follow the transformation of data from one technology to another.

Document → Information source with Semantic Mark-up

Initially, this will be simulated by hand coding sample documents with XML to implement the Ontology Web Language. This will give rise to evaluation criteria for, ease of use, expressiveness, flexibility and machine readability of the OWL protocol.

The key integration question will be how will this be automated? There are two possible approaches - incorporate tools / templates into the existing origination process or extract the information as a separate process. Ideally a suitably transparent addition to the existing process would be preferable, but there may still need to be input from a knowledge engineer.

Marked up Documents → Ontology

Information about modules will need to be incorporated into object instances and the structure of the information will be validated against the class hierarchy. The validation of alterations to the knowledge base may also require input from the rule base.

The key evaluation criterion for this step is the degree of change that can be allowed for by the system. The more the Ontology can be self-organising the more automated the process can become. This is likely to require a meta-ontology supported by a meta-rule base, which will be concerned with how concepts represented by the ontology can be altered. The obvious limit to this is related to the classic 'frame problem' but we hope to identify the more common types of change and evaluate the systems ability to cope with respect to the effort required to implement the behaviour.

Ontology Creation

We have chosen Protégé to construct the ontology component for a number of reasons.

- As an effective ontological construction and editing tool
- Used and supported by wide academic community [7].
- Stable tool with continuing development in applications and plug-ins for compatibility with new developments.

As an application generator so that we can develop our own domain specific tools with minimum cost.

Ontology → Rule base

Once the relationships between the various concepts have been systematised in an ontology it should be possible to implement a rule base that utilises this information. Here the key evaluation criteria are to quantify the advantages gained by using an ontology. Many knowledge-based systems are effective at providing useful functionality without the support of an explicit ontology. We expect to find that an approach without an overarching ontology will be less flexible as more of the concepts will have to be implicitly represented in the rules of the knowledge base rather than as an abstract description.

Rule base → Interface

This step is currently being implemented by providing a query service via a web page. Ultimately it is expected that machine-readable information will be able to be used by an agent based system. Graesser [1] provides a good overview of agent systems. Each student will have a software agent that will be the equivalent of a dynamic, interactive student record. The agent will respond to requests from the student and notify the student of changes in information from the central system that may effect the student. A similar system has been proposed by the author for the more complicated area of the Electronic Health Record[4].

4.2. The Database model

RDBMS + interfaces = Information system

This is implemented in MySQL using PHP to generate a simple web based interface.

The data model is deliberately as simple as is necessary to implement the basic functionality of the module choice system. It is expected that this will provide a benchmark to evaluate the effectiveness of the knowledge-based system. Although the system should provide a reasonable level of information it is expected that it will require more 'hands on' work and be less able to change without significant re-design.

5. Conclusion and future work

This paper is halfway between a position paper outlining a proposed project and an initial design document. It identifies the basic requirements and the proposed model and the core areas of concern. It outlines a number of evaluation criteria which are key to this project but will also shed light on wider aspects of the development of knowledge based systems in general and the use of Protégé in providing ontological support.

The next stage for this project will be to determine the detailed conceptual model for the domain and provide a coherent ontological representation. Currently there are overlaps between the functionality provided by different components of the system, we will have to find a way to assign responsibilities and justify the decisions.

It would be hoped that if the current project was successful a wider scale system incorporating more information sources could be implemented to support a wider range of information needs.

6. References

[1] Franklin S. Graesser A. Is it an Agent, or just a Program? A Taxonomy for Autonomous Agents Proceedings of the Third International Workshop on Agent Theories, Architectures, and Languages, Springer-Verlag, 1996.

[2] J. Gennari, M. A. Musen, R. W. Fergerson, W. E. Grosso, M. Crubézy, H. Eriksson, N. F. Noy, S. W. Tu The Evolution of Protégé: An Environment for Knowledge-Based Systems Development. *Technical report Stanford Medical Informatics SMI-2002-0943* 2002.

[3] Hayes-Roth, B, van Gent, R., Reynolds, R., Johnson, M., Wescourt, K. Web guides. *IEEE Intelligent Systems,* 14(2), 23-27, April 1999.

[4] Hobbs, M.H.W., An agent based information system for the NHS Electronic Health Record. *in Information Systems – Research, Teaching and Practice, Proceedings of the 5th UKAIS Conference*, pp 202-210, McGraw Hill 2000.

[5] JADE Home page *http://sharon.cselt.it/projects/jade/*

[6] JESS Home page *http://herzberg.ca.sandia.gov/jess/*

[7] Noy, N.F., Hafner, C.D., The State of the Art in Ontology Design, a Survey and Comparative Review, *AI Magazine*, pp 53- 74, Fall 1997.

[8] Protégé Home Page *http://protege.stanford.edu/*

[9] Semantic Web Activity Statement, W3C Technology and Society Domain 2002 (ref: *http://www.w3.org/2001/sw/Activity*)

[10] WC3 Working Draft Requirements for a Web Ontology Language, Web Ontology Working Group, 07/03/2002 (ref: *http://www.w3.org/2001/sw/WebOnt/)*