Proceedings of the Twenty-Third AAAI Conference on Artificial Intelligence (2008)

The Benefits of an Ontological Patient Model in Clinical Decision-Support

Mark Austin and **Matthew Kelly** and **Sir Michael Brady**
Wolfson Medical Vision Laboratory, Robotics Research Group
Department of Engineering Science, University of Oxford
Parks Road, Oxford OX1 3PJ, UK
Tel: +44 1865 283179
mark.austin@worc.ox.ac.uk

Abstract

In this paper we discuss an application integrating an ontological data model with an argumentation-based decision-support system, showing how the combination of leading technologies OWL, SPARQL and Jena can be used to achieve this. In the context of improving a decision support tool that is currently being trialled live in a clinical environment, we describe quantitatively how the incorporation of an ontology leads to an improvement over the existing software, highlighting the benefits of incorporating an ontology in medical applications. Data and clinical feedback is being collected from a live trial at the John Radcliffe hospital in Oxford, where we are able to test the original decision support tool, but also the ontology driven version, and thus will be able to demonstrate that any quantitative improvements in the efficacy of the software are a product of the ontological data model alone.

Introduction

The data structure used in most medical decision-support languages is some form of flat data storage, either storing data as key-value pairs, or in a database with a fixed number of columns (Robert & Raymond 2006). We are developing a decision-support-system *MDTSuite* to support the multidisciplinary team meetings (MDMs) that manage patients with colorectal cancer, the data structure of which currently follows this flat format. The MDMs are forums for the different clinical specialists (surgeons, radiologists, pathologists, oncologists and specialist nurses) to combine knowledge from their respective fields to form a group decision on treatment options for each patient in their care. *MDTSuite* supports this decision making process through a software suite which includes efficient patient data presentation software as well as an argumentation-based decision-support-application. In the development and live testing of this software, we have discovered several limitations which, we contend, derive directly from the use of such a flat data structure, some of which are likely to make such decision-support systems impractical for medical use.

For example, consider an example patient with two tumours as described in Figure 1. To describe this patient using a flat data structure, each piece of data for each tumour would be represented by a separate datatype, thus two sets of tumour stage, two sets of histological grade etc. would need to be present. Now, imagine the problem if a third tumour was to be found in the patient. Should the data structure then be modified to contain three such sets of data? Such cases illustrate clear limitations of using such a database design. If we were to create the fields required to describe the patient in Figure 1, they would then have to appear for every patient in the system, so for the majority of patients with only one tumour, a lot of irrelevant fields would be present. Also, notice that some of the data for rectal cancers is not required for colon cancers, for example the CRM (Circumferential Resection Margin - this is the minimal clearance between the boundary of the mesorectum and the affected tissue), but in a flat data structure, this field would necessarily be present for colon cancers too despite its irrelevance for such cancers.

Colorectal cancer patient:
Tumour in upper rectum:
Anterior, mobile, T1N0M0, Adenocarcinoma, Moderately differentiated, CRM 7mm
Tumour in caecum:
T3N0M0, Adenocarcinoma, Well differentiated

Figure 1: *Example of a patient which causes a problem for a flat data structure*

Such extra fields would result in an increase in input time for the system since a user would have to wade through more fields in order to define a patient. Ontologies (Noy & McGuinness 2001) offer a potential solution to all of these problems as well as providing an excellent data structure platform for the integration of multi-disciplinary clinical data including multimedia data. Perhaps for these reasons, they are the subject of a great deal of research in medical informatics, for example in projects such as HL7 RIM (Smith & Ceusters 2006; Schadow *et al.* 2006), SNOMED (SNOMED 2006) and the NCI Thesaurus (NCI 2006). Much of this research has focused on using ontologies to create a common language with which medical systems and centres can communicate (Smith & Ceusters 2003; Ceusters, Smith, & De Moor 2006; Ceusters & Smith 2003). However, an ontology can also be used to create a richer, more complete representation of the domain over which a computer can reason more effectively in a practical application, providing benefits beyond information integration. This more complete ontology-based representation enables us to describe the patient in Figure 1 fully, as well as handling even more complicated cases, whilst at the same time maintaining an efficient, smaller set of inputs for less complex patients. The relationships between data inherent in an ontological data structure enable an increase in the relevance of input fields such that only the required fields are displayed for each patient. These inherent relationships also facilitate a consistency check of the data capable of reducing the number of human input errors in the system.

Ontology-Driven MDTSuite

While we have focused our trial of the software on colorectal cancer, *MDTSuite* has been written in such a way that the knowledge base is completely separated from the application, so the software can be converted for use in any group decision making environment simply by writing a new rulebase. We have authored an OWL (McGuinness & Harmelen 2004) colorectal cancer ontology to replace the flat-file patient data model used originally and refactored the software to handle this model using the Jena (Carroll *et al.* 2003) java API. A data-driven, AJAX (Garrett 2006) powered patient data input form is generated automatically from the ontology and patient summaries are generated automatically from the populated patient instances. The logic statements associated with the arguments as well as the consistency arguments are written in the W3C recommended SPARQL (Perez, Arenas, & Gutierrez 2006) query language, and this rulebase can be created and maintained within the software itself. This architecture provides a very user-friendly and powerful platform for creating efficient argumentation-based decision support over extremely complex domains.

We have collected patient data for 100 colorectal patients that have been discussed at the MDT meetings at the John Radcliffe in Oxford over several months. The patient data was then represented in both the original flat data format and the new ontological model. As detailed below, we used this information to measure the improvement resulting from use of the ontology. We have seen that data model in the ontology-based *MDTSuite* can represent all of the patients presented at the MDMs, while the original software was unable to cope with at least one patient a week. The dynamic data input form leads to an average reduction in form length of 31% (from 70 to 48 input fields, the improvement was statistically significant with $p < .0001$), while the relevance is increased by 9% (again with statistical significance of improvement with $p < .0001$). Clinical time is the scarcest resource. By reducing the number of input fields displayed, as well as ensuring that those displayed are relevant to the patient, we are able to reduce the workload for the clinician loading patient data into the system as input time is surely correlated with the length of the input form. Input time could be reduced by 31% in line with the number of input fields displayed assuming that each field takes the same time to consider at the input stage. We can also demonstrate a reduction in input errors made possible by the ontology consistency checking. By generating random errors and attempting to detect them using the consistency arguments we can calculate what proportion of errors are picked up. By considering which errors have an actual effect on the recommendations given by the decision support tool we can also find out what proportion of these critical errors are detected. These critical errors were seen to be reduced by 59% when using the ontology consistency checking. In demonstrating these improvements, and the suitability of ontologies for information integration, we have made a strong case for the use of an ontological data model to elevate software from research to a useable system, robust enough for the unpredictable world of medicine.

A future goal for *MDTSuite* is to incorporate reasoning over imaging data in a more integrated manner. We hope to be able to link what is visible to a machine in an image to the underlying disease state of the patient. Ontologies provide an excellent platform for information integration of this sort thus presenting us with a further motivation for pursuing them.

References

Carroll, J.; Dickinson, I.; Dollin, C.; Reynolds, D.; Seaborne, A.; and Wilkinson, K. 2003. Jena: Implementing the semantic web recommendations.

Ceusters, W., and Smith, B. 2003. Ontology and medical terminology: why descriptions logics are not enough. *Proceedings of the conference Towards an Electronic Patient Record. Electronic Publication* 10–14.

Ceusters, W.; Smith, B.; and De Moor, G. 2006. Ontology-based integration of medical coding systems and electronic patient records. *http://ontology.buffalo.edu/medo/Coding-AndEHCR.pdf*.

Garrett, J., ed. 2006. *Ajax: A New Approach to Web Applications*. http://www.adaptivepath.com/ideas/essays-/archives/000385.php.

McGuinness, D., and Harmelen, F. v. 2004. Owl web ontology language overview. *W3C Recommendation*.

NCI. 2006. Nci thesaurus. *http://nciterms.nci.nih.gov*.

Noy, N. F., and McGuinness, D. L. 2001. *Ontology Development 101*.

Perez, J.; Arenas, M.; and Gutierrez, C. 2006. Semantics and complexity of sparql.

Robert, M. J., and Raymond, J. K. 2006. Electronic medical records and medical research databases can they be synonymous? *Business Briefing: US Cardiology*.

Schadow, G.; Walker, C.; Mead, N.; and Mead, D. 2006. The hl7 reference information model under scrutiny. *Studies in Health Technology and Informatics Vol 124*.

Smith, B., and Ceusters, W. 2003. Towards industrial-strength philosophy: How analytical ontology can help medical informatics. *Interdisciplinary Science Reviews*.

Smith, B., and Ceusters, W. 2006. Hl7 rim: An incoherent standard. *Studies in Health Technology and Informatics Vol 124 133138*.

SNOMED. 2006. Snomed clinical terms (snomed ct) core content for the july 2006 release. *http://www.snomed.org-/snomedct/documents/july_2006_release.pdf*.