Open Research Online

The Open University's repository of research publications and other research outputs

Ontological Foundations for Scholarly Debate Mapping Technology

Conference or Workshop Item

oro.open.ac.uk

Ontological Foundations for Scholarly Debate Mapping Technology

Neil BENN[*], Simon BUCKINGHAM SHUM, John DOMINGUE, Clara MANCINI
Knowledge Media Institute, Centre for Research in Computing, The Open University

Abstract. Mapping scholarly debates is an important genre of what can be called Knowledge Domain Analytics (KDA) technology – i.e. technology which combines both quantitative and qualitative methods of analysing specialist knowledge domains. However, current KDA technology research has emerged from diverse traditions and thus lacks a common conceptual foundation. This paper reports on the design of a KDA ontology that aims to provide this foundation. The paper then describes the argumentation extensions to the ontology for supporting scholarly debate mapping as a special form of KDA and demonstrates its expressive capabilities using a case study debate.

Keywords. Scholarly Debate Mapping, Macro-argument Analysis, Ontologies

Introduction

Research into tools to support both quantitative and qualitative analysis of specialist knowledge domains has been undertaken within the two broadly independent traditions of *Bibliometrics* and *Knowledge Management*. Knowledge Domain Analysis (KDA) tools within the first tradition (e.g. CiteSeer [1] and CiteSpace [2]) follow a citation-based approach of representing knowledge domains, where citation links are used as the basis for identifying structural patterns in the relationships among authors and publications. Tools within the second tradition (e.g. Bibster [3], ESKIMO [4], CS AKTIVE SPACE [5], and ClaiMaker [6]) extend the representational approach to include more features of knowledge domains – e.g. the types of agents or actors in the domain, their affiliations, and their research activities – with the aim of enabling more precise questions to be asked of the domain. This second approach depends on the development of software artefacts called ontologies, which are used to explicitly define schemes for representing knowledge domains.

This paper describes exploratory research into how these two traditions can be bridged in order to exploit both the benefit of ontologies to enable more feature-rich representations, as well as the established techniques of Bibliometrics for identifying structural patterns in the domain. The first section describes the design of a *merged* KDA ontology that integrates the existing ontologies specified in [3]- [6] (§1). Next, the paper describes how the merged ontology can be extended to include both a scheme for representing scholarly debates and inference rules for reasoning about the debate (§2). Thirdly, the extended ontology is applied to the representation and analysis of the

_____

[*] Correspondence to: Neil Benn, Knowledge Media Institute, The Open University, MK7 6AA, UK. Tel: +44 (0) 1908 695837; Fax: +44 (0) 1908 653196; E-mail: n.j.l.benn@open.ac.uk

abortion debate, which demonstrates the benefits of reusing techniques from both traditions (§3). The key lessons from this research are then discussed (§4), before concluding with directions for future research (§5).

1. A merged KDA ontology

One method for merging heterogeneous ontologies requires that the existing ontologies are aligned to a more generic reference ontology that can be used to compare the individual classes in the existing ontologies. Based on the fundamental assumptions that knowledge representation and communication are major activities of knowledge domains, and that knowledge representation and communication constitute semiotic activities, we propose to align the existing ontologies to a reference ontology that describes the interactions between components of any given semiotic activity.

1.1. Invoking a generic theory of semiotics as a reference framework

Semiotics is the study of signs and their use in representation and communication. According to Peirce's theory of semiotics [7], the basic sign-structure in any instance of representation and communication consists of three components: (1) the *sign-vehicle*, (2) the *object* referred to by the sign-vehicle, and (3) the *interpretant*, which is the mental representation that links the sign-vehicle to the object in the mind of some conceiving agent.

Recent research within the ontology engineering field has introduced a reusable *Semiotic Ontology Design Pattern (SemODP)* [8] that specifies, with some variation of Peircean terminology, the interactions between components of any given semiotic activity. The SemODP[1] is shown in Figure 1.
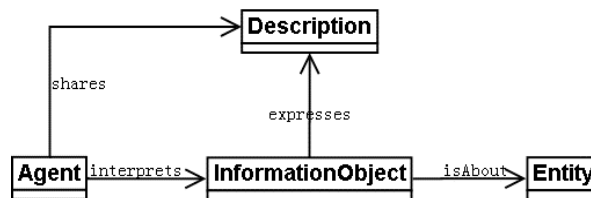

Figure 1. The Semiotic Ontology Design Pattern

In Peircean terminology, the *InformationObject* class of the SemODP represents the 'sign-vehicle'. In the context of knowledge domains, the most typical examples of information objects are publications, which are the main vehicles of knowledge representation and communication. A single publication can be regarded as an Information Object, as can each clause, sentence, table, graph, and figure that is either a verbal or non-verbal expression of knowledge within a publication.

The SemODP classes *Description* and *Entity* respectively correspond to the 'interpretant' and the 'object' in Peircean terminology. The Description class is the abstract, communicable knowledge content that an information object expresses. For

---

[1] The SemODP pattern shown here is based on an updated version of the 'home' ontology from which it is extracted (accessible at http://www.loa-cnr.it/ontologies/cDnS.owl)

example, a single publication is an information object that expresses a thesis (in much the same manner as a novel expresses a particular plot). The Entity class covers any physical or non-physical entity that an Information Object refers to via the 'is about' relation.

Finally, the SemODP specifies the *Agent* class. An agent is required to interpret a given Information Object and in such a case, the agent is said to conceive the Description expressed by that particular Information Object. In knowledge domains, instances of the Agent class include both "agentive physical objects" such as persons and "agentive social objects" such as organisations.

1.2. The core KDA ontology classes

The SemODP is used to merge the existing ontologies through a process of aligning the classes in the existing ontologies to the SemODP classes. The process also reveals the core KDA ontology classes that are based on consensus across the existing ontologies. For example, the consensus classes across the existing ontologies that can play the role of 'Agent' are 'Person' and 'Organisation'. However, it should be noted that core classes are not fixed indefinitely and constantly evolve as applications generate experience and consensus changes about what is central [9]. Table 1 shows the core KDA classes and their relationships to SemODP classes as well as existing ontology classes.

Table 1. The SemODP classes, the existing KDA classes they subsume, and the core KDA classes with their main properties.

SemODP class	Existing ontology classes[2]	Core KDA class	Properties (Type)
cdns:Agent	swrc:Person, eskimo:Person, akt:Person	mkda:Person	name (String)
cdns:Agent	swrc:Organisation, eskimo:Organisation, akt:Organisation	mkda:Organisation	name (String)
cdns:InformationObject	swrc:Publication, eskimo:Publication, akt:Publication	mkda:Publication	has Author (Agent); hasTitle (String); expresses (Description)
cdns:Description	scholonto:Concept	mkda:PropositionalContent	verbalExpression (String)
cdns:Description	scholonto:Concept	mkda:NonPropositionalContent	verbalExpression (String)

[2]'swrc:' is the prefix for the ontology of the Bibster tool [3]; 'eskimo:' is the prefix for the ontology of the ESKIMO tool [4]; 'akt:' is the prefix for the ontology of the CS AKTIVE SPACE tool [5]; 'scholonto:' is the prefix for the ontology of the ClaiMaker tool [6].

2. Extending the ontology to support scholarly debate mapping

The work in this section builds on the debate mapping approach of Robert Horn [10] who has produced a classic series of seven debate maps for analysing the history and current status of debate on whether computers can think[3]. What has emerged from this debate mapping approach is a theory of the structure of debate, which has subsequently been articulated by Yoshimi [11] in what he calls a "logic of debate". Whereas most argumentation research concentrates on the *micro*structure of arguments (e.g. the types of inference schemes for inferring conclusions from premises), the concern of a logic of debate is how arguments themselves are "constituents in *macro*-level dialectical structures" [11]. The basic elements of this logic of debate are implemented as additional classes and relations in the merged KDA ontology.

2.1. Debate representation

Issues
In the proposed logic of debate, *issues* can be characterised as the organisational atoms in structuring scholarly debate[4]. Indeed, according to [13], one of the essential characteristics of argumentation is that there is an issue to be settled and that the argumentative reasoning is being used to contribute to a settling of the issue. An *Issue* class is introduced into the ontology as a specialisation of the core KDA class *NonPropositionalContent*.

Propositions & Arguments
The other basic elements in the proposed logic of debate are claims (propositions) and arguments, where the term 'argument' is used in the abstract sense of a set of propositions, one of which is a conclusion and the rest of which are premises. However, an argument can play the role of premise in another argument, thus allowing the chaining of arguments. *Proposition* and *Argument* classes are introduced into the ontology as specialisations of the core KDA class *PropositionalContent*. The main relations between claims and arguments in the logic of debate are *supports* and *disputes*.

2.2. Debate analysis

Analysing a debate involves reasoning with representations of that debate to detect potentially significant features of the debate. Here we propose to draw on the well-developed network-based analytical techniques employed within the Bibliometrics tradition. However, this reuse is not straightforward since the analytical techniques are typically designed to operate on single-link-type network representations of domains, where the links between nodes are used to signal *positive association* between nodes. For example, network-based analytical techniques are often applied to co-citation networks where a link between two publications is established when they are both cited by a common third publication, and that link signals positive association between the two publications.

This single-link-type assumption presents a challenge because the ontology-based debate representations can be regarded as 'multi-faceted' representations – i.e. there are

---

[3] This debate largely takes place within the knowledge domain of *Artificial Intelligence*
[4] This is analogous to the proposal by [12] that issues are the "organisational atoms" of so-called *information-based information systems* for solving intractable design problems.

a number of node types and a number of link types. Thus, before network-based analytical techniques can be reused for debate analysis, transformation rules need to be defined that can project the 'multi-faceted' debate representation onto a single-link-type representation.

In order for such a projection to work, the multiple relations in the ontology need to be uniformly interpreted from a single perspective. We propose to interpret the relations in the ontology in a 'rhetorical-discourse' context. Indeed, it can be be argued that the analytical techniques of the Bibliometric tradition interpret the publications in a citation-based network as rhetorical viewpoints, which implies that the positive association link between nodes can be interpreted as rhetorical agreement.

Furthermore, the work of Mancini and Buckingham Shum [14] provides the basis for an efficient implementation of the transformation rules. An important feature of their work is the use of a limited set of cognitively grounded parameters (derived from the psycholinguistic work on discourse comprehension [5] by [15]) to define the underlying meaning of discourse relations in the ontology of the ClaiMaker tool. Mancini and Buckingham Shum [14] anticipate that using discourse coherence parameters as the underlying definition language will allow different discourse-relation vocabularies to be used for representing discourse without changing the underlying discourse analysis services provided by their tool. The four bipolar discourse parameters proposed by [15] are: *Additive/Causal*, *Positive/Negative*, *Semantic/Pragmatic*, and *Basic/Non-Basic*.

The 'Additive/Causal' parameter depends, respectively, on whether a weak or strong correlation exists between two discourse units. Note that 'causal' is generally given a broad reading in discourse comprehension research to include causality involved in argumentation (where a conclusion is motivated *because* of a particular line of reasoning), as well as more typical cause-effect relationships between states of affairs.

The 'Positive/Negative' parameter depends, respectively, on whether or not the *expected* connection holds between the two discourse units in question. For example, in the sentence *"Because he had political experience, he was elected president"* the connection between the two units is Positive since the reader would typically expect "being elected president" to follow from "having political experience". However, in the sentence *"He did not have any political experience, yet he was elected president"* the connection between the two units is Negative since the expected consequent of "*not* having any political experience" is "*not* being elected president", but what is actually expressed is a violation of that expectation – i.e. "*yet* he *was* elected president".

The 'Semantic/Pragmatic' parameter depends, respectively, on whether the connection between the two discourse units lies between their factual content or between the speech acts of expressing the two discourse units. At this stage we are primarily focussed on enabling debate analysis, and since debate analysis falls within the realm of speech acts [16], the relations between entities that make up a debate representation will be parameterised as 'Pragmatic' by default.

The 'Basic/Non-Basic' parameter depends, respectively, on whether or not, in the case of Causal connections, the cause precedes the consequent in the presentation of the discourse. The example – *"Because he had political experience, he was elected president"* – is parameterised as Basic, whereas the example – *"He was elected*

---

[5] Discourse comprehension research in general is concerned with the process by which readers are able to construct a *coherent* mental representation of the information conveyed by a particular piece of discourse.

president because he had political experience" – is parameterised as Non-Basic. This parameter is largely about presentation and does not affect the essential nature or meaning of the discourse connection. Thus it can be omitted from the basic parameterisation of relations in the ontology.

These coherence parameters are then used as a grammar for defining relations in the merged KDA ontology, including relations between publications, between persons, and between arguments. The benefit of this approach is that rather than implement a multitude of inference rules for inferring positive association, only a limited set of parameterised inference rules need to be implemented. For example, Figure 2(i), which shows a parameterised rule for inferring a +ADDITIVE connection between some Y and some Z, covers typical positive association inferences such as when two arguments support a common third argument or when two publications cite a common third publication. Figure 2(ii), which also shows a parameterised rule, covers typical positive association inferences such as when two persons author a common publication or when two arguments are expressed by a common publication. Finally, Figure 2(iii) covers a typical 'undercutting' pattern in argument analysis which is a variation of the social network analysis adage that "the enemy of my enemy is my friend".
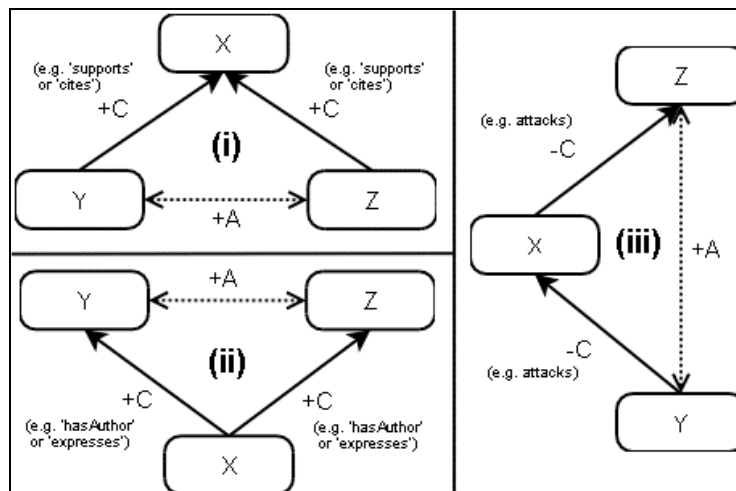

Figure 2. Some of the CCR-parameterised inference rules in the ontology. The dotted line indicates that a +ADDITIVE connection is inferred based on the other connections.

3. The abortion debate case study

This section briefly describes the representation and analysis of the debate about the desired legality/illegality of abortions, as laid out in the *Abortion Debate* entry of the online Wikipedia [17]. The description is a summary of what appears in more detail elsewhere [18].

3.1. Capturing representations of the debate in a knowledge base

Figure 3 (unshaded portion) shows an extract from the Wikipedia entry which expresses some of the debate issues. Figure 3 (shaded portion) then shows how the

first of the questions in the extract (underlined) is captured as an *Issue* instance in the knowledge base[6]. Note that this issue instance is asserted as a sub-issue of the main issue being debate – i.e. *"What should be the legal status of abortions"*.

Some of the most significant and common issues treated in the abortion debate are:

- *The beginning of personhood (sometimes phrased ambiguously as "the beginning of life"): <u>When is the embryo or fetus considered a person?</u>*
- *Universal human rights: Is aborting a zygote, embryo, or fetus a violation of human rights? ...*

...

(def-instance ISS1 Issue
  ((verbalExpression "What should be the legal status of abortions?")))

(def-instance ISS2 Issue
  ((verbalExpression "When is the embryo or fetus considered a person?")
   (subIssueOf ISS1)))

…

Figure 3. (Unshaded portion) An extract from the Wikipedia entry showing some of the debate issues (Shaded portion) 'Issue' instances modelled in the knowledge base.

Next the process turns to representing the claims and arguments in the debate. According to the Wikipedia entry, the argumentation in the debate is generated by two broadly opposing viewpoints – *anti-abortion* and *pro-abortion*. Figure 4 (unshaded portion) shows an extract that expresses three basic 'anti-abortion' claims. Figure 4 (shaded portion) then shows how the first of these three claims is captured as a *Proposition* instance. An *Argument* instance is then coded – BASIC-ANTI-ABORTION-ARGUMENT – which groups the anti-abortion claims together. Similar steps are performed to represent the basic pro-abortion viewpoint. Next, an 'addresses' link is asserted between the BASIC-ANTI-ABORTION-ARGUMENT *Argument* instance and the main debate issue. Finally, a 'disputes' link is asserted between the two *Argument* instances BASIC-ANTI-ABORTION-ARGUMENT and BASIC-PRO-ABORTION-ARGUMENT.

The view that all or almost all abortion should be illegal generally rests on the claims: (1) that <u>the existence and moral right to life of human beings (human organisms) begins at or near conception-fertilisation</u>; (2) that induced abortion is the deliberate and unjust killing of the fetus in violation of its right to life; and (3) that the law should prohibit unjust violations of the right to life.

...

(def-instance P1 Proposition
  ((verbalExpression "The existence and moral right to life of human organisms begins at or near conception-fertilisation")))

…

(def-instance BASIC-ANTI-ABORTION-ARGUMENT Argument
  ((hasPremise P1 P2 P3)
   (hasConclusion P4))

…

(def-relation-instances
  (addresses BASIC-ANTI-ABORTION-ARGUMENT ISS1)
  (disputes BASIC-ANTI-ABORTION-ARGUMENT BASIC-PRO-ABORTION-ARGUMENT))

…

Figure 4. (Unshaded portion) An extract from the Wikipedia entry showing the basic anti-abortion viewpoint in the debate. (Shaded portion) Examples of claims and arguments modelled in the knowledge base.

---

[6] The knowledge base is coded using the OCML [19] knowledge modelling language.

The Wikipedia entry also includes information about publications and authors. Figure 5 (unshaded portion) shows an extract from the Wikipedia entry detailing the reference information for two publications by an author participating in the debate. Figure 5 (shaded portion) shows how the author is modelled as a Person instance and how the first publication is modelled as a Publication instance in the knowledge base. The shaded portion of the figure also shows how the fact that a publication 'expresses' an argument is captured in the knowledge base. In this case, the Thomson (1970) publication expresses an argument labelled as the 'bodily-rights argument', which supports the basic pro-abortion viewpoint in the debate.

- *Thomson, J. "A Defense of Abortion". Philosophy and Public Affairs 1:1 (Autumn 1971): 47-66.*
- *Thomson, J. "Rights and Deaths". Philosophy and Public Affairs 2:2 (Winter 1973): 146-159.*
  …

```
(def-instance JUDITH_THOMSON Person)

(def-instance THOMSON1971DEFENSE Publication
 ((hasAuthor JUDITH_THOMSON)
  (hasTitle "A defense of abortion")
  (hasYear 1971)))

(def-relation-instances
 (expresses THOMSON1971DEFENSE BODILY-RIGHTS-ARGUMENT)
 (supports BODILY-RIGHTS-ARGUMENT BASIC-PRO-ABORTION-ARGUMENT))
```

Figure 5. (Unshaded portion) An extract from the Wikipedia entry showing the reference information for two publications. (ii) (Shaded portion) The corresponding instances modelled in the knowledge base.

3.2. Detecting viewpoint clusters in the debate

The purpose of modelling a debate in a knowledge base is to enable new kinds of analytical services for detecting significant features of the debate. One such feature is the clustering of viewpoints in the debate into cohesive subgroups. A service to detect viewpoint clusters is of significance to a hypothetical end-user because understanding how scholars in a debate fall into different subgroups has already been established as an important part of understanding a debate [10].

The first step in executing this service is to translate the ontology-based representation into a single-link-type network representation so that the clustering technique reused from Bibliometrics can be applied. This involves executing the previously introduced inference rules on the entire ontology-based representation, which results in a network representation with a single +ADDITIVE link type. Next, a clustering algorithm is run over the network representation, which yields a number of clustering arrangements, ranging from 2 clusters to 25 clusters. The algorithm designers [20] propose a 'goodness-of-fit- measure as an objective means of choosing the number of clusters into which the network should be divided. Figure 6 shows the clustering arrangement with five clusters, which the algorithm determines in this case is the arrangement with the 'best fit' for the given network data[7]. However, it should be noted that, as is typical in the Bibliometrics tradition (e.g. [21]), the goal of this type of analysis is typically not to find the definitive clustering arrangement, but rather to detect interesting phenomena that will motivate more focussed investigation on the part of the analyst.

---

[7] The Netdraw network analysis tool (available at http://www.analytictech.com/Netdraw/netdraw.htm) is used here to perform the clustering and visualisation.
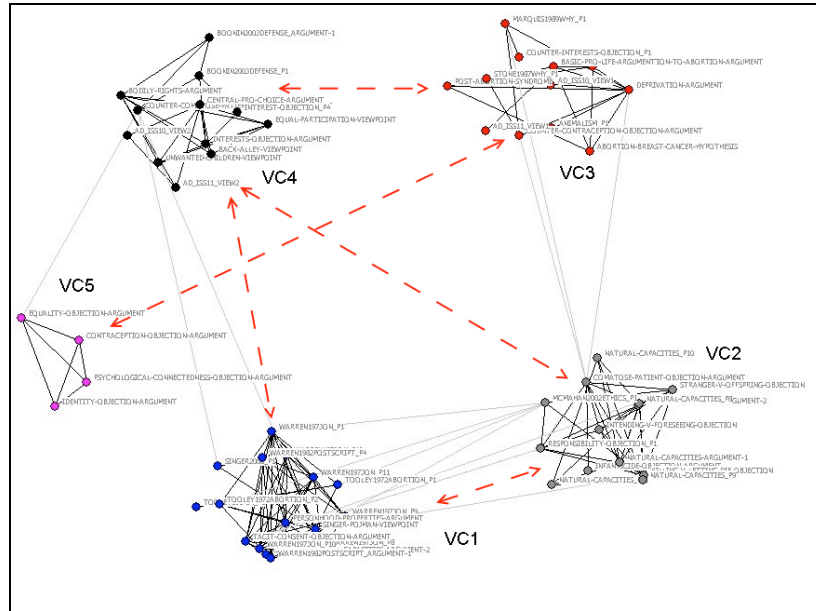
Figure 6. Five viewpoint clusters (V.C. #1 – V.C. #5) identified in the abortion debate network. The dashed red lines between clusters indicate opposition (which is determined by further analysis once the clustering results are translated back into the knowledge base)

Once the viewpoint clusters have been identified the results are translated back into the knowledge base where each of the clusters detected in the network is captured as a *ViewpointCluster* instance. Further analysis is then performed on these instances to determine, for example, the persons who are associated with each cluster and the clusters which are deemed to be opposing each other. Two clusters are regarded as opposing if at least half of the viewpoints in one cluster have a '*disputes*' relation with the viewpoints in the other cluster. Table 2 shows for each cluster, the associated viewpoints, the associated person(s), and the opposing cluster(s).

Table 2. Further details of the viewpoint clusters detected in the abortion debate, including the persons associated with each cluster and the clusters that are opposing each other.

VC#	Associated Viewpoints	Associated Person(s)	Opposing Cluster(s)
VC1	COUNTER-NATURAL-CAPACITIES-ARGUMENT[8], PERSONHOOD-PROPERTIES-ARGUMENT[9]	Bonnie Steinbock, David Boonin, Dean Stretton, Jeff McMahan, Louis Pojman, Mary-Anne Warren, Michael Tooley, Peter Singer	VC2 VC4
VC2	COMATOSE-PATIENT-OBJECTION-ARGUMENT[10],	Don Marquis, Francis Beckwith, Germaine Grisez, Jeff McMahan, John Finnis, Katherine	VC1

---

[8] Summary: "The argument that the fetus itself will develop complex mental qualities fails."

[9] Summary: *"The fetus is not a person because it has at most one of the properties – consciousness – that characterizes a person."*

[10] Summary: *"Personhood criteria are not a justifiable way to determine right to life since patients in reversible comas do not exhibit the criteria for personhood yet they still have a right to life"*

	INFANTICIDE-OBJECTION-ARGUMENT[11]	Rogers, Massimo Reichlin, Patrick Lee, Robert George, Robert Larmer, Stephen Schwarz	VC4
VC3	ABORTION-BREAST-CANCER-HYPOTHESIS[12], BASIC-ANTI-ABORTION-ARGUMENT	Don Marquis, Eric Olson, Jim Stone	VC4 VC5
VC4	BODILY-RIGHTS-ARGUMENT[13], BASIC-PRO-ABORTION-ARGUMENT	David Boonin, Dean Stretton, Jonathan Glover, Judith Thomson, Peter Singer	VC1 VC2 VC3
VC5	CONTRACEPTION-OBJECTION-ARGUMENT[14], IDENTITY-OBJECTION-ARGUMENT[15]	Dean Stretton, Frederick Doepke, Gerald Paske, Jeff McMahan, Lynne Baker, Mary-Anne Warren, Michael Tooley, Peter McInerney, William Hasker	VC3

4. Discussion

The discussion section is organised around a series of questions adapted from the GlobalArgument.net experiment[16]. These questions were used to evaluate various computer-supported argumentation (CSA) approaches to modelling the Iraq Debate. The discussion is concerned with two main points: the added value of the approach, and the limitations of the approach.

4.1. In what ways does this CSA approach add value?

How does this CSA approach guide a reader/analyst through the debate?
The aim of this approach is to provide analytical services that enable the reader to identify interesting features of the debate and gain insights that may not be readily obtained from the raw source material alone. As viewpoint clusters provide a way of abstracting from the complexity of the debate, an approach that enables the detection of viewpoint clusters in a debate is an improvement on what a user would have been able to obtain from looking only at the online Wikipedia entry of the abortion debate. Another interesting feature of the debate that the table reveals is that some persons are associated with more than one cluster. For example, the person of *Jeff McMahan* is associated with two clusters (VC1 and VC2), and furthermore these clusters happen to oppose each other. These are features of the debate that can then be investigated down to the level of the individual arguments.

---

[11] Summary: *"Using personhood criteria would permit not only abortion but infanticide"*
[12] Summary: *"There is a causal relationship between induced abortion and an increased risk of developing breast cancer"*
[13] Summary: *"Abortion is in some circumstances permissible even if the fetus has a right to life because even if the fetus has a right to life, it does not have a right to use the pregnant woman's body."*
[14] Summary: *"It is unsound to argue that abortion is wrong because it deprives the fetus of a valuable future as this entails that contraception, which deprive sperm and ova of a future, is as wrong as murder, something which most people don't believe."*
[15] Summary: *"The fetus does not itself have a future value but has merely the potential to give rise to a different entity, an embodied mind or a person, that would have a future of value"*
[16] http://kmi.open.ac.uk/projects/GlobalArgument.net

To what extent is the modeller's or analyst's expertise critical to achieving the added value?

Capturing the debate in the knowledge base often relied on the modeller's ability to reconstruct argumentation to include parts of arguments not expressed in the original information resource as well as inter-argument relations no expressed in the original information resource. This has an impact on what kinds of connections can be inferred during the reasoning steps, which then has an impact on what features of the debate can be identified.

4.2. What are the limitations of the CSA approach?

What aspects of the debate proved difficult to model?

Using this approach it was difficult to account for the different types of *disputes* relations between arguments. For example, in the case study, one argument often disputed another, not just because of disagreement with the conclusion, but because of the perceived 'unsoundness' of the reasoning used to arrive at the conclusion. Also, because of the focus on macro-argumentation, it was difficult to account for different inference schemes for moving from premises to conclusion in individual arguments.

What missing capabilities have been identified?

One important missing capability is (semi)automatically capturing debate representations from primary literature sources (e.g. experimental articles carried by scholarly journals). The approach relied on a manual process of constructing representations of the debate based on the Wikipedia, which is classified as a tertiary literature source [22]. Tertiary literature – which includes encyclopaedias, handbooks and review articles – consolidates and synthesises the primary literature thus providing an entry point into the particular domain. This was used for the case study to enable the manual coding of the debate, which would have been too vast to code using all the primary literature that was synthesised. However, this has meant that the debate representations rely on the accuracy of the tertiary-level synthesis of the primary literature.

5. Conclusion and Future Work

This paper has described how an ontology of the structure of specialist knowledge domains can be extended to support the representation and analysis of scholarly debates within knowledge domains. The benefit of this extension has then been demonstrated by representing and analysing a case study debate. In particular, a service was demonstrated for detecting 'viewpoint clusters' as significant features of scholarly debate. Other debate modelling case studies are currently being written up.

Future directions for this research need to address the difficult question of how to best support the use of primary literature as the source for capturing debate representations. To cover a significant area of any knowledge domain, the modelling of primary literature would need to be conducted in a (semi)automated distributed fashion using possibly many modellers. This would introduce new challenges of trying to ensure consistent modelling across the different modellers. Finally, to address the current limitations of representing micro-argumentation, future research needs to

investigate how the ontology can incorporate the argument specification of the Argument Interchange Format (AIF) [9].

References

[1]   Lawrence, S., C. Giles, and K. Bollacker, *Digital Libraries and Autonomous Citation Indexing.* IEEE Computer, 1999. **32**(6): p. 67-71.

[2]   Chen, C., *CiteSpace II: Detecting and Visualizing Emerging Trends and Transient Patterns in Scientific Literature.* Journal of the American Society for Information Science and Technology, 2006. **57**(3): p. 359–377.

[3]   Haase, P., et al., *Bibster - a semantics-based bibliographic Peer-to-Peer system.* Journal of Web Semantics, 2004. **2**(1): p. 99-103.

[4]   Kampa, S., *Who are the experts? E-Scholars in the Semantic Web*, in *Department of Electronics and Computer Science*. 2002, University of Southampton.

[5]   Schraefel, M.M.C., et al. *CS AKTive space: representing computer science in the semantic web.* in *13th International Conference on World Wide Web, WWW 2004*. 2004. New York, NY, USA: ACM.

[6]   Uren, V., et al., *Sensemaking tools for understanding research literatures: Design, implementation and user evaluation.* International Journal of Human-Computer Studies, 2006. **64**(5): p. 420-445.

[7]   Atkin, A., *Peirce's Theory of Signs*, in *The Stanford Encyclopedia of Philosophy (Fall 2007 Edition)*, E.N. Zalta, Editor. 2007, URL = <http://plato.stanford.edu/archives/fall2007/entries/peirce-semiotics/>.

[8]   Behrendt, W., et al. *Towards an Ontology-Based Distributed Architecture for Paid Content.* in *The Semantic Web: Research and Applications, 2nd European Semantic Web Conference (ESWC 2005)*. 2005. Heraklion, Crete, Greece: Springer-Verlag.

[9]   Chesñevar, C., et al., *Towards an argument interchange format.* Knowledge Engineering Review, 2006. **21**(4): p. 293-316.

[10]   Horn, R.E., *Mapping great debates: Can computers think? 7 maps and Handbook*. 1998, MacroVU: Bainbridge Island, WA.

[11]   Yoshimi, J., *Mapping the Structure of Debate.* Informal Logic, 2004. **24**(1): p. 1-21.

[12]   Kunz, W. and H.W.J. Rittel, *Issues as Elements of Information Systems*. 1970, Center for Planning and Development Research, University of California, Berkeley.

[13]   Walton, D., *Argumentation Schemes for Presumptive Reasoning*. 1996, Mahwah, New Jersey: Lawrence Erlbaum Associates.

[14]   Mancini, C. and S. Buckingham Shum, *Modelling discourse in contested domains: A semiotic and cognitive framework.* International Journal of Human-Computer Studies, 2006. **64**: p. 1154-1171.

[15]   Sanders, T.J.M., W.P.M. Spooren, and L.G.M. Noordman, *Towards a taxonomy of coherence relations.* Discourse Processes, 1992. **15**: p. 1-35.

[16]   Eemeren, F.H.v., et al., *Reconstructing Argumentative Discourse*. 1993, Tuscaloosa/London: The University of Alabama Press.

[17]   Website, *The Abortion Debate*, URL = <http://en.wikipedia.org/wiki/Abortion_debate> (accessed on 22 September 2006).

[18]   Benn, N., et al., *Designing the Ontological Foundations for Knowledge Domain Analysis Technology: An Interim Report*. 2008, Knowledge Media Institute, The Open University, UK. Technical Report KMI-08-02. Available at: http://kmi.open.ac.uk/publications/pdf/kmi-08-02.pdf.

[19]   Motta, E., *Reusable Components for Knowledge Modelling: Case Studies in Parametric Design Problem Solving.* Frontiers in Artificial Intelligence and Applications, ed. J.L.d.M. Breuker, R.; Ohsuga, S.; Swartout, W. Vol. 53. 1999, Amsterdam: IOS Press.

[20]   Newman, M.E.J., *Fast algorithm for detecting community structure in networks.* Physical Review E, 2004. **69**(066133).

[21]   Andrews, J.E., *An author co-citation analysis of medical informatics.* Journal of the Medical LIbrary Association, 2003. **91**(1): p. 47-56.

[22]   Anderson, J., *The role of subject literature in scholarly communication: An interpretation based on social epistemology.* Journal of Documentation, 2002. **58**(4): p. 463-481.