What do you mean? Arguing for Meaning

Tom M. van Engers[*], Adam Wyner[°]
[*]*Leibniz Center for Law, University of Amsterdam*
[°]*University of Leeds*

Abstract Building ontologies has been proven to be a complex issue in part because a community must commit to the conceptualization that the ontology represents. The community members must align their concepts and co-create. Arguing about a useful conceptualization is therefore an essential part of the process of designing an ontology. Logicians have developed formal argumentation theories, but have not combined formal argumentation with conceptualization. Rather, while conceptualization should play an important role in any argumentation theoretical approach, argumentation theories focus on arguments based on propositional logic and argument structures, which are not sufficient for arguing about domain conceptualization, which requires a more fine-grained logical analysis. In this paper we will explain why conceptualization plays an important role within argumentation and why argumentation support tools, especially if they use Natural Language Processing (NLP), can help in creating domain ontologies.

Keywords: Argumentation, Ontologies, Knowledge Acquisition, Natural Language Processing.

1. Introduction

Building ontologies has proven to be a complex issue in part because a community must commit to the conceptualization that the ontology represents. The community members must align their concepts and co-create. Arguing about a useful conceptualization is therefore an essential part of the process of designing an ontology. The creation of ontologies is usually done in small teams as part of informal knowledge engineering activities where participants discuss the conceptualization. Except where a minority has discretionary power to define the concepts, such a format is not suited for creating shared meaning between members of a larger community. However, in practice, people can cope with the task. For instance, where someone misunderstands, clarifying questions are asked and explanations given. Thus, the shared conceptualisation emerges from discussion; arguing about a useful conceptualization is an intrinsic part of communication. While it is not always easy for human beings to acknowledge and adjust to a different conceptualization, the problems of detecting conceptual differences and creating reconceptualizations are problems which are hard to solve in AI.

While one might expect that logicians working at formal theories on argumentation would have addressed the problems of conceptualization, thus far little attention has been paid to combining formal argumentation with conceptualization. Instead, argumentation theories focus on arguments based on propositional logic, which is not fine-grained enough to argue about domain conceptualization.

Computational linguists have made significant progress in building ontologies from sentences expressed in natural languages. In order to address the hard AI problem of understanding natural language, researchers in this domain usually work with controlled languages (CLS). A good example of this approach is the Attempto Controlled English (ACE, see http://attempto.ifi.uzh.ch/site/description/), which is used by a relatively large number of computational linguists. Sentences expressed in ACE, i.e. in a somewhat restricted subset of the English language, can be parsed into first order logic (FOL) from which the ontology is derived.

One of the reasons to consider the interaction between natural language, argumentation, and conceptualisation is that knowledge engineers must translate from knowledge of a domain, often expressed in natural language, into a representation that is argued about. However, representing each sentence as a proposition hides crucial information that would help to relate statements or the contents of statements, draw inferences, filter redundancy, and identify contradictions.

In this paper we will illustrate why conceptualization plays an important role within argumentation and why argumentation support tools especially if they use Natural Language Processing (NLP) can help in creating domain ontologies.

2. Using CNL for Policy-making Discussions

We work with a scenario in which we want to support stakeholders to participate in policy-making discussions, using forum technology. For this purpose the domain knowledge, i.e. knowledge about the issues being discussed, must be made explicit, formal, and expressed in a language that a machine can process. This machine-readable knowledge representation we call the target form. Translating the knowledge that people have of a domain, which is often implicit, informal, and expressed in natural language, the source form, into the target form is a labour, time, and knowledge intensive task (see also Van Engers 2005), creating a "knowledge acquisition bottleneck" which has limited the adoption and use of powerful AI technologies (see Forsythe and Buchanan 1993).

In Wyner et al. (2010) we propose and outline a framework which extends multi-threaded discussion forums, integrating NLP, ontologies, and argumentation. The proposed framework goes beyond existing debate and argumentation support systems, by making the semantic content of the stakeholders in the policy-making debate formal and explicit. In this paper we will address the formalization rather than the construction of dialectical arguments.

While there are tools which support multi-user ontology development (see WebProtege http://webprotege.stanford.edu/) and there are ontology development tools which use natural language for input (see the AceWiki plug in for Protege http://attempto.ifi.uzh.ch/acewiki/), there is no support for arguing in natural language about an ontology. Rather, current ontology online multi-user systems such as WebProtege rely on the users to converge on an ontology or note the differences. Our proposal motivates the development of systems which not only captures the differences, but represents them as distinct ontologies for reasoning.

Broadly speaking, among the issues that need to be addressed are the following. Even if users enter in well-formed natural language sentences, how can we be assured that they enter in well-formed, meaningful rules for the formulation of arguments? Where we rely on input from public participants, who are not logicians or knowledge engineers with training in building well-formed rules, ill-formed arguments could be entered. This raises a general issue of what prompts can be introduced to make KB construction systematic and meaningful? For instance, at the level of propositions there is nothing incoherent about a rule such as *If P and Q, then R*. However, we see the rule is incoherent where P is *Bill is happy* and Q is *The Great Wall of China is long* and R is Swallows fly south in spring. Indeed, there is nothing preventing users from entering ungrammatical sentences, or sentences that are out of topic of the context of discussion. In the following we develop these issues.

One of the results and in some cases even one of the purposes of argumentation is to clarify issues by finding a shared conceptualization between the participants. Boer (in Boer 2009) citing Schlag (see Schlag 1996) stresses the importance of posing questions in (legal) arguments. He uses the following rhetorical hierarchy guiding those questions:

1. Ontological questions question the truth of terminological axioms and the ontological inferences based upon them.
2. Epistemic questions question the non-terminological inferences made from certain premises to certain thesis.
3. Normative questions address whether something is allowed or disallowed, good or bad etc.
4. Technical questions question the propositions of a case and are about the truth of the facts of a case.

This strength of attacking arguments depends on the rhetorical level, level 4 being the weakest and level 1 being the strongest attack.

In the following section we will explain some conceptualization issues that are relevant to argumentation.

3. Conceptualization issues in arguments

Participants involved in an argumentation process use natural language as the most import means of expressing themselves. In order to understand those expressions, the terms and syntactical information glueing them together has to be transferred into a conceptual model. Where the participants gradually come to understand one another, we have a process of shared conceptualization (Van Engers 2001). The shared conceptual model (ontology) only partly overlaps with the internal mental models of the stakeholders, and making an explicit conceptualization is usually a labour intensive task which requires lots of discussion because the (intended) meaning of concepts depend on the role those concepts play in the cognitive system of the individuals. Shared meaning has to be construed, requiring a 'rewiring' in the minds of these individuals. Mapping terms to a shared conceptualization can result in two typical inferential problems. The first one is class-referential mismatch, and the second is instance-referential mismatch.

An example of a class-referential mismatch is given in the following example where we have the following arguments:

Argument 1 consists of three statements in natural language;

Statement 1. People need a healthy living environment.
Statement 2. Plants are responsible for considerable air pollution.
Statement 3. Therefore plants should be prohibited in living environments.

Argument 2 also consists of three statements in natural language:

Statement 1. People need a healthy living environment.
Statement 2. Plants are responsible for regeneration of air.
Statement 3. Therefore we should have as many plants as possible in living environments.

Obviously the interpretation of these arguments would be quite different depending on what the concept would be that we want the term 'plant' to refer to.

An example of a instance-referential mismatch is the following. Suppose we have the following two arguments:

Argument 1:

Sentence 1: John is rich therefore John is happy.

and a rebuttal

Argument 2:

Sentence 2: John has severe health problems therefore John is not happy.

These arguments can be represented in the following AIF-graph:

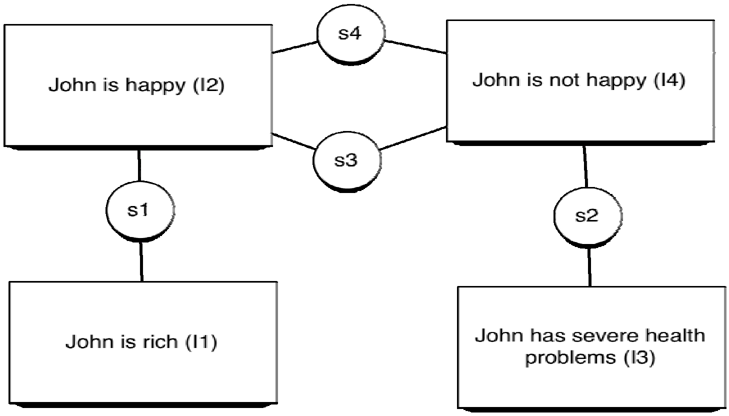

Figure 1. An AIF graph representing two conflicting arguments with a potential instance-referential mismatch. In this AIF-graph we'll find four I-nodes corresponding to

1. John is rich
2. John is happy
3. John has severe health problems
4. John is not happy

Obviously we expect that the John in all of these sentences refers to the same instance (assuming that this is what most readers will infer). But suppose that this is not the case and John in the first two I-nodes is referring to a different instance. In that case the two S-nodes representing the conflict between the second and fourth wouldn't make sense. In order to connect the I-nodes to the conceptualization we could use a mapping function. This mapping function would map the I-nodes 1 and 2 in our example to instance 'John12' and I-nodes 3 and 4 to John'34'. More precisely we would have two situations -- a situation before it was clarified that there are two Johns instead of one and the situation after this was clarified.

In the first AIF-graph the nodes would be functionally mapped to the same instance (John'12'). While in the second AIF-graph the I-nodes 1 and 2 in would be mapped to instance 'John12' and I-nodes 3 and 4 to John'34' and the S-nodes representing the conflict would be 'undercut' with a functional mapping to the 'exclusion' relation between John12 and

John34 in the conceptual model represented by the two sentences in our example.

Another conceptualization mismatch is the caused by the properties that individuals believe to belong to a concept. This problem could be solved to either split the concept in two or more concepts. This can be illustrated by the following example where we reuse the first argument of our previous example,

Argument 1 consists of three statements in natural language:

Statement 1. People need a healthy living environment.

Statement 2. Plants are responsible for considerable air pollution.

Statement 3. Therefore plants should be prohibited in living environments.

Argument 2 also consists of three statements in natural language:

Statement 1. Only some plants cause considerable air pollution.

Statement 2. Plants in living environments can help to reduce t travelling distance to work.

Statement 3. Therefore non-polluting plants should be allowed in living environments.

The second argument introduces a new concept (explicit in Statement 3) that of the non-polluting plant, which will require the splitting of the original concept plant into two concepts, one polluting plants, and another that of non-polluting plants. The reader must have detected the implicit argument in Statement 2 of the second argument that hides the conceptual relationship between travelling to work and air-pollution. Making this relationship explicit would require prompting in order to reveal all deductive steps implicitly made by the individual that made the statement.

The expressivity of AIF-graphs is intentionally limited to represent argument structures and not the content of the constituents of the 'I-nodes'. But this is unfortunately also the case in most other argumentation formalisation formalisms. Understanding the meaning of the arguments however does require a mechanism that allows for connecting the I-nodes to the corresponding conceptualization of the content of these I-nodes.

4. Conclusions and future work

In the IMPACT project we address argumentation in the context of policy modelling, which is a challenge. Firstly the participants in policy-making debates use natural language and understanding natural language is a hard AI problem. Secondly the dialectical form of the argumentation process may shift between different dialogue types (see e.g. Walton 1992).

Persuasion dialogue, information-seeking dialogue, negotiation dialogue, inquiry dialogue and sometimes even eristic dialogue can be mixed in such dialogues. We therefore have to limit the dialogue form and the language used, using a controlled language and a specific dialogue protocol in the forum.

On the argumentation formalization side we have little support yet either. The Dung framework (see also Laera et al. 2006) which we see as a basis for many argumentation theories is not typically useful in the context of policy making. In order to support the users in understanding the arguments, or policies, we need to be able to grasp the meaning of their expressions and give feedback about the consequences of their positions and choices. For this kind of feedback we have to go beyond the fourth level in the rhetorical hierarchy introduced in the section 3, i.e. the technical questions. We claim that in order to really support policy-making we need to be able to also cover the other rhetorical layers, up to understanding the meaning of the propositions, which implies that we have to formalise the participants' expressions using at least in FOL. We intend to further improve the NLP components as well as a component that can prompt participants posing rhetorical questions, as well as critical questions relevant to the argument (a plethora of papers on critical questions in argumentative settings can be found on Doug Walton's website see http://www.dougwalton.ca/papers.htm).

In our approach we hope to bridge between ontology building and argumentation theories which we believe is essential to both fields. As no knowledge will grow without arguments, we hope that our research contributes to more knowledgeable policy-makers and consequently to better policy.

Acknowledgements

The authors wish to thank the European Commission for sponsoring the IMPACT project.

References

Bailin, S. C., Truszkowski, W., 2002, Ontology Negotiation: How Agents Can Really Get to Know Each Other. In Proceedings of the WRAC 2002.

Boer, A., Van Engers, T.M., Van de Ven, S., 2010, Knowledge Acquisition from Sorces of Law in Public Administration, In: Proceedings of the EKAW2010 conference, under review. (2010).

Boer, A., 2009, *Legal Theory, Sources of Law, & the Semantic Web*. Dissertations in AI. IOS Press.

Van Engers, T.M., 2005, Legal engineering: A structural approach to improving legal quality. In Macintosh, A., Ellis, R., Allen, T., eds.: Applications and Innovations in Intelligent Systems XIII, Proceedings of AI-2005, Springer (2005) 3–10

Van Engers, T.M., 2001. *Knowledge Management: The Role of Mental Model in Business Systems Design*. PhD thesis, Vrije Universiteit Amsterdam, 2001. Belastingdienst.

Forsythe, D.E., Buchanan, B.G., 1993, Knowledge acquisition for expert systems: some pitfalls and suggestions. In: Readings in knowledge acquisition and learning: automating the construction and improvement of expert systems. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA (1993) 117–124

Laera, L., Tamma, V., Euzenat, J., Bench-capon, T., Payne, T., 2006, Reaching agreement over ontology alignments, 2006, In *Proceedings of 5th International Semantic Web Conference* (ISWC 2006), Springer Verlag, p 371--384

Rahwan, I., Reed, C., 2009, The Argument Interchange Format, In: I. Rahwan, G. R. Simari (eds.), *Argumentation in Artificial Intelligence,* 383 DOI 10.1007/978-0-387-98197-0 19, c_ Springer Science+Business Media, LLC 2009

Russell, B., 1949, Human Knowledge – Its Scope and Limits, Allen & Unwin, London.

Schlag, P., 1996, Hiding the ball. *New York University Law Review*, 1681, 1996.

Walton, D., 1992, Types of Dialogue, Dialectical Shifts and Fallacies, In Van Eemeren, F., Grootendorst, R., Blair, J.A., Willard, C.A. (eds.), *Argumentation Illuminated,* Amsterdam, SICSAT, 133-147.

Wyner, A., Van Engers, T.M., Bahreini, K., 2010, From policy-making statements to first-order logic. In: Proceedings of eGOVIS 2010. (2010) to be published.

Wyner, A., Krasmir, A., Barzdins, G., Damljanovic, D., Davis, B., Fuch, N., Heofler, S., Jones, K., Kaljurand, K., Kuhn, T., Luts, M., Pool, J., Rosner, M., Schwitter, R., Sowa, J., 2010, On Controlled Natural Languages: Properties and Prospects}, In Fuchs, E. (ed.): Workshop on Controlled Natural Languages, CNL 2009, Springer, LNCS/LNAI 5972.