

Intent of Work Statement

Yale, BBN, Kestrel Group

Award number: F30602-00-2-0600

Drew McDermott, PI

Subcontract PIs: Mark Burstein (BBN), Doug Smith (Kestrel)

Technical Goals and Accomplishments

The intent of our work from the beginning has been to create tools for managing mappings among clashing ontologies. We have progressed to the point where we have a formal theory, and have implemented pieces of a software system to carry it out. Our goal is an *ontology-mapping server*, which would accept requests for translation from agents that need to communicate across ontology boundaries, and would respond to those requests with either individual translations, or a *glue agent* that could carry out a series of translations with no further intervention from the server. (The server does not create ontology mappings, but manages mappings created by human experts.) Pieces of this server are already in place, including preliminary translators from RDF to the internal representation used by our system.

The key technical idea behind our approach is that ontology mapping is best thought of in terms of ontology merging. The merge of two ontologies is obtained by taking the union of the axioms defining them, using XML namespaces to avoid name clashes. We then add *bridging axioms* that relate the terms in one ontology to the terms in the other. Inferences can be conducted in this merged ontology in either a demand-driven (backward-chaining) or data-driven (forward-chaining) way. Because users want to see conclusions expressed only in their own ontologies, conclusions might have to be *projected* back into the component ontology.

Our principal technical goals are to

- Show that these inference problems are well defined.
- Overcome the software-engineering issues that will arise (such as management of ontology revisions, mappings among multiple ontologies, handling inconsistencies, and so forth).

The first set of goals have been largely achieved, although we have yet to put the pieces together. The second set will become crucial once we have completed implementing Version 0.5 of the ontology-mapping server this spring.

In addition to the work originally envisaged, we have become active in the DAML-S consortium, which is building a formalization of web-service agents in DAML+OIL. Our main contribution is to show how sophisticated service representations coupled with planning algorithms can make it possible for agents to find novel ways of making use of web services.

Our Role in the DAML Experiment

We believe that our work will fit nicely into the "DAML Experiment" to be carried out this year, or it will when there are more ontologies to be mapped. We will explain how it fits, then talk about the problem of "ontology shortage."

The DAML Experiment Vision talks of "multiple, distributed ontologies." This seems entirely appropriate, given that ontologies will surely grow out of existing databases, web-search directories, intelligence analysts' guidelines, and such. These systems are complex and in some cases rather ad hoc. It will be hard enough to represent their external interfaces and data models in ontologies without requiring them to cohere into a general framework.

The instance data envisioned by the Experiment document will inevitably be quasi-formal database fragments entered by military intelligence analysts, scrounged from web pages, or automatically extracted from databases. Using as a guide this document's vision of how these sources are to be used (by "an agent to roll up and apply a

Intent of Work Statement

Yale, BBN, Kestrel Group

Award number: F30602-00-2-0600

Drew McDermott, PI

Subcontract PIs: Mark Burstein (BBN), Doug Smith (Kestrel)

Technical Goals and Accomplishments

The intent of our work from the beginning has been to create tools for managing mappings among clashing ontologies. We have progressed to the point where we have a formal theory, and have implemented pieces of a software system to carry it out. Our goal is an *ontology-mapping server*, which would accept requests for translation from agents that need to communicate across ontology boundaries, and would respond to those requests with either individual translations, or a *glue agent* that could carry out a series of translations with no further intervention from the server. (The server does not create ontology mappings, but manages mappings created by human experts.) Pieces of this server are already in place, including preliminary translators from RDF to the internal representation used by our system.

The key technical idea behind our approach is that ontology mapping is best thought of in terms of ontology merging. The merge of two ontologies is obtained by taking the union of the axioms defining them, using XML namespaces to avoid name clashes. We then add *bridging axioms* that relate the terms in one ontology to the terms in the other. Inferences can be conducted in this merged ontology in either a demand-driven (backward-chaining) or data-driven (forward-chaining) way. Because users want to see conclusions expressed only in their own ontologies, conclusions might have to be *projected* back into the component ontology.

Our principal technical goals are to

- Show that these inference problems are well defined.
- Overcome the software-engineering issues that will arise (such as management of ontology revisions, mappings among multiple ontologies, handling inconsistencies, and so forth).

The first set of goals have been largely achieved, although we have yet to put the pieces together. The second set will become crucial once we have completed implementing Version 0.5 of the ontology-mapping server this spring.

In addition to the work originally envisaged, we have become active in the DAML-S consortium, which is building a formalization of web-service agents in DAML+OIL. Our main contribution is to show how sophisticated service representations coupled with planning algorithms can make it possible for agents to find novel ways of making use of web services.

Our Role in the DAML Experiment

We believe that our work will fit nicely into the "DAML Experiment" to be carried out this year, or it will when there are more ontologies to be mapped. We will explain how it fits, then talk about the problem of "ontology shortage."

complex algorithm to the values that have been assigned to the enemy and friendly capability”), we foresee a number of potential opportunities to demonstrate our translation framework. Potential new sources of input to these algorithms could be found that have useful *content*, but whose data formats are incompatible with the algorithms. We are currently looking, for example, at the many forms of locational information used by the military, which could be intertranslated to make the data that talks about those locations more widely useful.

The whole purpose of an ontology is to create self-describing datasets. That is, when an agent requires a deduction to be done, it can make use of a dataset if it uses the same ontologies the dataset does, and it can make use of the subclass, cardinality, and axiomatic relationships described in the ontology to make inferences from this dataset and others. The capacity to make inferences from multiple sources that could not be made from a single source is what makes ontologies interesting.

Our work addresses the obvious next question: What if the agent’s ontology and the dataset’s ontology are not the same, as they surely will fail to be in many cases (see above)? We are eager to get involved in case studies of military ontologies that overlap in domain coverage but disagree on conceptualization and syntactic details. The biggest obstacle to proceeding with this plan, or indeed with the whole DAML Experiment, is that there is a **shortage of real world ontologies**. Therefore, we believe that, to make the case that DAML will improve interoperability, we will have to help the users develop their ontologies in the first place. They would tell us what they want to input, and we would infer from such requests what it is they want to represent and what the formal relationships are among those concepts.

We propose teaming with two different military users with overlapping ontologies. (Much more fruitful at this stage than to hook up with another set of researchers.) Another possible source of input into ontology formation is existing database schemas. **If we can get our hands on some of those it would be extremely useful.** (We don’t need the contents of the database, only the schemas, which are hopefully less sensitive.) By examining existing data models from two different military sources we can develop, bottom up, some very realistic examples of interoperable ontologies (we already have seen examples of this at Air Mobility Command, where they move data between systems).

In summary:

1. We will finish implementing the ontology-mapping server. We should have a version up and running by the end of March, 2002.
2. We will begin working as soon as possible to find, or, more likely, develop ontologies that capture various military concepts.
3. We will develop merged ontologies (that is, translation or ‘bridging’ axioms) and put them into the server.

The timeline for the DAML Experiment shows that demonstrations are to occur in May, 2002. That should be no trouble for the first version of our server, running on hand picked examples. We should in principle be a key component in almost any experiment run in June or October (as the timeline suggests), but we have serious doubts that very many realistic ontologies can be made available that quickly.

Forecast

It is difficult to forecast how the military need for “semantic web” markup languages will evolve in the next two years, except to predict that it will evolve rapidly. There is explosive growth of proposals in the area of web services. There are areas of the military that fit this model well, such as logistics. We expect DAML-S to be a key factor in formalizing the way such “military web services” operate. In fields such as Operational Net Assessment, the future is not so clear.

There are two possible ways the area will evolve over the next two years: Either the problem of “ontologizing” quasiformal knowledge will prove to be much harder than we have been assuming, or there will be slow but steady progress in marking up many quasiformal knowledge sources. Our judgment is that the latter outcome will

be boosted by finding a few high-gain data sources in which adding a certain amount of self description suddenly makes the data available to many more military users.

Assuming that happens, the problem of ontology translation is going to become more and more crucial. The alternative is for ontologies to merge together spontaneously so that eventually there is one big ontology that everyone uses. This outcome may or may not occur in the long run, but it certainly will not happen in the next two years. Our observation of progress on common data models at places like AMC indicates that this is a very difficult process, because data models are designed for specific systems and purposes. "One size fits all" is very hard to achieve. We are therefore convinced that translation will continue to play an important role as the number of systems that may be usefully connected grows rapidly. If ontology development enjoys a slow but steady growth, the effort to keep track of their mergers will be able to keep up, and it will be essential that it do keep up.