Steps Toward Sharable Ontologies for Design Rationale

Jeffrey M. Bradshaw, John H. Boose, David B. Shema
Research & Technology, Boeing Computer Services
P.O. Box 24346, M/S 7L-64, Seattle, WA 98124; (206) 865-3422; jbrad@atc.boeing.com

Douglas Skuce, Timothy C. Lethbridge
Department of Computer Science, University of Ottawa
Ottawa, Ontario, Canada K1N 6N5; (613) 564-4518; doug@csi.uottawa.ca

ABSTRACT

We are interesting in improving collaboration among designers through better knowledge sharing. Building models is not only a way to *capture* design information, but serves more importantly as a means to *communicate* and *come to understand* a design as it evolves. An important prerequisite to effective communication is mutual agreement on important terms and concepts, the *ontology* of the domain. In our approach we are looking at several layers of ontologies: Top level ontologies with very general concepts, ontologies for particular domains, ontologies of concepts relevant to modeling those domains, and ontologies of concepts relevant to partitioning knowledge bases using theories and contexts. We also describe research into mechanisms for exchanging these ontologies.

1. INTRODUCTION: MODELING AS COMMUNICATION

Modeling design rationale is largely a matter of communication. It begins when a group of individuals determine that they have valuable information to share. It evolves as participants collaborate to define models that represent a common understanding of certain aspects of the domain. It succeeds when participants can use these models effectively to promote and enrich communication.

From this perspective, a model is not a 'picture' of the problem, but rather a device for the attainment or formulation of knowledge about it (Kaplan, 1963). Indeed, sometimes the most important outcome of the modeling process may not be the model itself, but rather the insight we gain as we struggle to articulate, structure, critically evaluate, and agree to it (Moore & Agogino, 1987). Thus, the value of such an effort derives not simply from a final 'correct' representation of the problem, but additionally from our success in framing the activity as a self-correcting enterprise that can subject any part of the model to critical scrutiny, including our background assumptions. We must ask ourselves not only, How do we know the model is correct? (every model is an incorrect oversimplification); but also, How useful is the model (and the modeling process) in facilitating our understanding of the domain?

Concern for these issues has become the rallying cry of a number of researchers in conceptual modeling and knowledge acquisition. Recent work in knowledge acquisition has emphasized that the creation of knowledge bases is a constructive modeling process, and not simply a matter of 'expertise transfer' or 'knowledge capture' (Ford & Bradshaw, 1992). Mylopoulos (1991) describes the field of conceptual modeling as follows:

> *"Conceptual modeling* is the activity of *formally* describing some aspects of the physical and social world around us *for purposes of understanding and communication.* Such descriptions, often referred to as *conceptual schemata,* require the adoption of a formal notation, a *conceptual model* in our terminology. Conceptual schemata capture relevant aspects of some world, say an office environment and the activities that take place there, and can serve as points of agreement among members of a group, for example the workers in that office, who need to have a common understanding of that world. Conceptual schemata can also be used to communicate that common view to newcomers, through a variety of graphic and linguistic interfaces. Conceptual modeling has an advantage over natural language or diagrammatic notations in that it is based on a formal notation which allows one to 'capture the semantics of the application'. It also has an advantage over mathematical or other formal notations developed in computer science because unlike them, conceptual modeling supports structuring and inferential facilities that are psychologically grounded. After all, the descriptions that arise from conceptual modeling activities are intended to be used by *humans*, not machines."

Unfortunately, until recently developers of modeling and knowledge representation tools have paid scant attention to the needs of the kind of users who are building a model with the intent of coming to understand some domain. For example, traditional data modeling tools introduce assumptions about the way conceptual schemata will be realized on a physical machine, which constrain and confound domain experts preferring to leave these considerations to others. Furthermore, the need of such users to model the 'real world' drives a requirement for much richer representations than modeling tools typically afford. Knowledge representation tools, on the other hand, almost always introduce the assumption that the resultant knowledge bases will be directly usable for some computational task. Considerations of efficiency and performance inevitably lead to limitations in the richness and flexibility of the knowledge representation. No less important, there are few modeling tools that have been optimized for *usability*. Tools are cumbersome, unnatural, and difficult to use without specialized training and extensive experience. The comprehensibility of textual and graphical notations chosen is rarely considered. As a result, many potential users abandon modeling tools altogether and use simple graphical drawing tools instead.

Elsewhere we have discussed our ideas about how the application of certain general principles of knowledge representation and user-interface design can make our tools more usable and extensible by non-experts (Bradshaw, Ford & Adams-Webber, 1991; Bradshaw, Holm, Kipersztok & Nguyen, 1992; Skuce, 1991b). In this paper, we discuss another aspect of usability for tools intended to promote communication as part of design rationale modeling: support for design and sharing of reusable ontologies.

2. ROLE OF COMMON ONTOLOGIES IN PROMOTING COMMUNICATION

Recently, a number of researchers have emphasized the importance of defining common ontologies to specify the concepts and terms upon which knowledge-based systems

depend (e.g., Gruber, 1991, 1992; Lenat & Guha, 1990; Neches, Fikes, Finin, Gruber, Patil, Senator & Swartout, 1991; Skuce and Monarch, 1990):

> "Consider a planning system based on a theory in which plans are composed of 'steps' which form 'sequences' with specific kinds of 'resource dependencies', and that the search for plans is guided by 'ordering heuristics' and 'optimization criteria.' If one wished to use this planning system, one would need to understand what these words mean, and build a knowledge base in which domain-specific knowledge was formulated in terms that the planning *program* also understands." (Gruber, 1992)

The desire for large-scale knowledge sharing and reuse across different tools and applications provides a strong motivation for common ontologies. The desire for large-scale knowledge sharing and reuse across different tools and applications provides a strong motivation for common ontologies. Careful design of the ontology and the equally careful choice of terminology also benefit developers and users of the system. Our experience confirms that terminological confusion breeds conceptual confusion (and vice versa). Such confusion tends to become more probable the 'higher' one goes in a concept hierarchy. For example, it is much easier to reach consensus on the meaning of the term "pencil" than on terms like "object," or "entity." Hence the most general 'ontological' concepts are the ones most in need of standardization, yet also the most difficult to agree on. Standardization has two aspects:

a) agreeing what the concepts *are*

b) agreeing what to name them, i.e. the terminology

At present, agreement at such high levels is probably unlikely. Our approach is therefore two-pronged:

a) Concentrate on agreement at as high a level as possible within sub-domains

b) Separately, try to develop a set of truly top-level, general ontological concepts and terminology for future agreement

We are currently developing ontologies that contain concepts useful in formally describing our own work in design rationale. These ontologies are being developed within a modeling framework called DDUCKS (Decision and Design Utilities for Comprehensive Knowledge Support; Bradshaw, Boose & Shema, 1992; Bradshaw, Fulton, Holm, Boose, Kipersztok & Nguyen, 1992)[1]. To assist in the definition of ontologies, we have integrated DDUCKS with an enhanced version of CODE4 (Skuce & Lethbridge 1992; Lethbridge, 1991). CODE4 provides a rich, paradigm for the definition of knowledge level concepts. A collection of integrated tools support the important and frequently overlooked aspects of

---

[1] Either the first or second *D* in *DDUCKS* is silent, depending on whether one using the tool in a decision or design context.

conceptual, ontological, and terminological analysis (Skuce 1991b, Lethbridge & Skuce 1992b) . We are developing extensions to the representation to allow the system to make use of additional inferencing and representation facilities similar to those found in Sowa's (1991) conceptual graphs and Gaines' (1991) KRS, which interpret taxonomic and entity-relationship structures in terms of typed formal logics. A first order logic system and a simple natural language system allow various types of syntactic and semantic checks to be performed, if desired. A comprehensive lexicon allows references to concepts to be automatically maintained and quickly accessed. We emphasize the importance of comprehensive lexical support so that terminology can be carefully chosen and subsequently controlled. Concept libraries and default inferencing mechanisms can be augmented by users employing graphical views and an integrated scripting and query language.

3. ELEMENTS OF AN APPROACH TO ONTOLOGY SHARING

We suggest that a successful approach to the development of sharable ontologies for design rationale representation involves the following elements:

1. Very general ontologies

2. Domain ontologies

3. Modeling ontologies

4. Theories and contexts

5. A mechanism for sharing

6. Lexical (terminological) support

The first four of these elements represent increasingly specialized knowledge. We anticipate that research on very general ontologies will eventually result in a consensus, although this may be many years away. More specific ontologies use knowledge from more general ontologies, therefore the closer to consensus the higher-level ontologies can be brought, the more effective will be the sharing of lower level ontologies.

The speed with which consensus can be reached is largely dependent on the fifth element above: a mechanism for sharing. Ontological research is currently shared very little, and where sharing takes place, it is usually either a) in the form of paper reports that take time to distribute and get outdated rapidly, or b) among researchers using some specific piece of not-widely-distributed research software. To facilitate development of ontologies it will be necessary for determine how divers software tools can be facilitated to exchange their knowledge. It will also be necessary to determine *conceptual* formats, i.e. exactly what pieces of information are necessary to convey to accurately communicate ontological concepts among a community of researchers.

3.1 Very General Ontologies

An important area for of work concerns the most general categories that ontologies can have, notions like: thing, object, entity, property, attribute, event, process, state, situation, collection, relation, etc., to name a few favorites. At the moment, everyone uses these concepts and terms, but a) they probably use them in very different ways and b) they cannot tell anyone else what they mean by them. A well known example is the top of the Cyc ontology, which we find difficult to understand from the published descriptions. (Skuce, 1991a).

We believe that such notions can best be clarified by studying linguistic and psychological data. An important AI ontology based on linguistic research is the "Penman" ontology (Bateman, Kasper, Moore & Whitney, 1990). This ontology, derived from Halliday's studies of English, is well-documented, with reasonable if minimal descriptions of each of some 200 categories. Another linguistically-based ontology is Miller's WordNet (Miller 1990). One of us (Skuce) has been working on an empirical approach to very general ontologies for a number of years. By the end of the summer, he hopes to have an initial proposal ready with approximately fifty categories of the kind listed in the previous paragraph.

3.2 Domain Ontologies

In conjunction with a number of research groups interested in knowledge sharing issues, we are participating in an effort to develop specific ontologies for particular domains of interest. For example, in one project we have been developing ontological primitives to represent knowledge of the airplane design and manufacturing enterprise (Bradshaw, Holm, Kipersztok & Nguyen, 1992). In another project, we are interested in representing medical decision making knowledge (Bradshaw, Chapman & Sullivan, 1992). Others have begun projects that will result in ontologies relevant to disciplines such as process planning, software engineering, circuit layout, and device modeling (see Gruber, Tenenbaum & Weber, 1992).

3.3 Modeling Ontologies

While the domain ontologies are designed to be relatively neutral with respect to a particular rationale modeling framework, the modeling ontologies are meant to capture a particular theoretical or practical point-of-view. In other words, the modeling ontologies characterize the *roles* that domain concepts play within a particular modeling framework.

We are performing an analysis of the knowledge acquisition, decision support, and design rationale tools we have developed over the past several years to understand the concepts and assumptions, implicit and explicit, they embody. We have found that the process of developing a formal modeling ontology for our tools has greatly increased our understanding of them. As we increase our understanding of the ontological commitments of our own framework, we plan to develop and evaluate similar modeling ontologies derived from other perspectives. The elements of Lee and Lai's DRL (1991), for example, could be represented quite naturally within a modeling ontology. MacLean et al's (1991) questions, options, and criteria could be formally represented in the same way. The

separation of domain from modeling ontologies will allow us to begin comparing and contrasting different approaches to design rationale for the same application.

3.4 Theories and Contexts

Practical concerns about the engineering of large systems have led to the development of techniques for creating and operating on multiple partitions of the knowledge base (Guha & Lenat, 1990; Gruber, 1992). Each partition, or *theory*, constitutes a set of axioms tailored for a particular domain or purpose. Each theory is associated with a *context* (McCarthy, 1987) that describes, among other things, the set of assumptions made by the theory. We are exploring a layered approach to theory and context management within DDUCKS. Starting with any layer in the system, a user can produce a set of tools, models, ontologies, and representations that imports selected assumptions and axioms into the layer below.

2.5 A Mechanism for Sharing

Gruber's work on Ontolingua (1992) currently provides the most promising mechanism for sharing ontologies between different tools and formalisms. Ontolingua extends the knowledge interchange format (KIF; Genesereth & Fikes, 1991) defined by the DARPA knowledge sharing effort with standard primitives for defining classes and relationships, and organizing knowledge in object-centered hierarchies with inheritance. Ontolingua facilitates the translation of KIF-level sentences to and from forms that can be used by various knowledge representation systems (currently LOOM, Epikit, or Algernon). We are working with Gruber to define an Ontolingua interface for CKB, the CODE4 knowledge base file format (Lethbridge & Skuce, 1992a).

The goal is to use the Ontolingua project as a basis for processing and distributing the ontologies. As good examples become available, the Sharing and Reuse of Knowledge Bases subgroup of the knowledge sharing effort can coordinate evaluations and experiments. In this way, we can pool our results with those of projects with similar goals (e.g., Gruber, Tenenbaum & Weber, 1992).

In order to facilitate sharing of ontologies a number of key research issues must be addressed.

- Knowledge primitives: *Of what primitive elements is the knowledge composed?* We take the approach that all elements are to be considered as concepts[2], and in the CODE4 system we categorize concepts. into types, instances, statements, properties, terms, metaconcepts, etc. Simply saying that the primitives shall be some set of 'predicates' does not answer the difficult question of where we get this set from. This is where linguistic and terminological research becomes especially relevant.

---

[2] i.e., by *concept*, we mean anything that can be referred to and thought about.

- Standardization: *Should there be only one format or more than one?* The types of knowledge in ontologies developed within the design rationale community may have special properties that differ significantly from knowledge exchanged in other communities. Some may consider these differences mandate different interchange standards. We believe that the differences are not great enough, and the benefits of a single standard are of over-riding importance.

- Inference primitives: *What are the basic semantics of the exchanged knowledge?* For example in CKB format some basic rules of inheritance are assumed. If somebody attempted to process a CKB knowledge base but assumed different or no inheritance rules, the results of the processing would likely be invalid (or at the very least, substantial knowledge would be lost). We believe it is necessary to have an absolute minimum number of inference primitives otherwise each system will be forced to reimplement every other system. In order to allow the exchange of knowledge between systems that have differing inference primitives, we believe it is important to be able to exchange 'informal' knowledge. Such knowledge is only manipulated by systems that recognize it. Informal knowledge can also facilitate human understanding by allowing such things as comments.

- Physical format: Obviously it will be necessary to develop a common syntax, but there are a number of meta-issues upon which to be agreed: Should compactness be a priority? What about human readability? Should the structure be hierarchical or flat? Should it look like predicate logic or some extension thereto such as conceptual graphs (Sowa, 1984)? How are cross-references represented: by unifiable symbols (English words), by indexes, or by pointers?

We have chosen our own positions on the above issues, and so have others. Now the issues have been raised we anticipate a lively debate. However, we hope existing formats for sharing knowledge will begin to accelerate the exchange of ontologies.

3. TERMINOLOGICAL SUPPORT

As we have noted, an important aspect of knowledge sharing is lexical, or terminological knowledge. In addition to its other features, CODE4 provides a high level of integrated support for lexical functions such as:

defining a term

comparing meanings of closely related terms

using a term in several senses, or using synonyms

checking that a term is used consistently and correctly

changing a term throughout a knowledge base

relating verbs to associated nouns (for example, "extract", "extractor")

translating a term into another language

critiquing and assisting in choice of terms.

CONCLUSION

We hope that our initial efforts at developing ontologies and sharing mechanisms will spur other researchers to make similar efforts, with the goal of increasing knowledge sharing and reuse within the design rationale community.

REFERENCES

Bateman, J., R. Kasper, et al. (1990) *A General Organization of Knowledge for Natural Language Processing: the Penman Upper Model*, , USC/Information Sciences Institute.

Bradshaw, J.M., Boose, J.H., Holm, P., Kipersztok, O. & Nguyen, T. (1992). DDUCKS: A framework for corporate knowledge representation in an enterprise integration architecture. *Proceedings of the Concurrent Engineering and CALS Conference*, Washington, D.C., June 1-5.

Bradshaw, J.M., Boose, J.H. & Shema, D.B. (1992). A knowledge acquisition approach to design rationale. In J. Carroll and T. Moran (Eds.), *Design Rationale*. Hillsdale, N.J.: L. Erlbaum, in preparation.

Bradshaw, J.M., Chapman, C.R. & Sullivan, K.M. (1992). An application of DDUCKS to bone-marrow transplant patient support. *Working Notes of the AAAI 1992 Artificial Intelligence in Medicine Session of the Spring Symposium,* Stanford, CA, March.

Bradshaw, J.M., Ford, K.M. & Adams-Webber, J. (1991). Knowledge representation for knowledge acquisition: A three-schemata approach. *Proceedings of the Sixth Banff Knowledge Acquisition Workshop,* Banff, Canada, October.

Bradshaw, J.M., Fulton, J.A., Holm, P., Boose, J.H., Kipersztok, O. & Nguyen, T. (1992). DDUCKS: A pluggable architecture for reuse and tailorability of models, tools, ontologies, and representations. Submitted to the AAAI-92 Knowledge Representation Aspects of Knowledge Acquisition Workshop. San Jose, CA, July.

Bradshaw, J.M., Holm, P., Kipersztok, O. & Nguyen, T. (1992). *eQuality:* An application of *Axotl II* to process management. In T. Wetter, K-D Althoff, J. Boose, B. Gaines, M. Linster & F. Schmalhofer (Eds.), *Current Developments in Knowledge Acquisition: EKAW-92.* Berlin/Heidelberg: Springer-Verlag.

Dixon (1991). *A New Approach to English Grammar, on Semantic Principles.* Oxford: Clarendon Press.

Ford, K. & Bradshaw, J.M. (1992) (Eds.), special knowledge acquisition issue of the *International Journal of Intelligent Systems,* in preparation. Also to appear in K. Ford & J.M. Bradshaw (Eds.), *Knowledge Acquisition as a Modeling Activity.* New York: John Wiley, volume in preparation.

Gaines, B.R. (1991). Empirical investigations of knowledge representation servers: Design issues and applications experience with KRS. AAAI Spring Symposium: Implemented

Knowledge Representation and Reasoning Systems, pp. 87-101. Stanford (March). Also in SIGART Bulletin, 2(3), 45-56.\

Genesereth, M. R. & Fikes, R. (1991). Knowledge Interchange Format Version 2.2 Reference Manual. Logic Group Report, Logic-90-4. Stanford, CA: Stanford University Department of Computer Science, March.

Gruber, T.R. (1991). The role of common ontology in achieving sharable, reusable knowledge bases. Stanford Knowledge Systems Laboratory Report No. KSL 91-10, February. To appear in J.A. Allen, R. Fikes, and E. Sandewall (Eds.), *Principles of Knowledge Representation and Reasoning: Proceedings of the Second International Conference.* San Mateo, CA: Morgan Kaufmann.

Gruber, T. (1992). Ontolingua: A mechanism to support portable ontologies. Stanford Knowledge Systems Laboratory Technical Report KSL 91-66, Final revision February 1992. Stanford, CA: Stanford University Department of Computer Science.

Gruber, T.R., Tenenbaum, J.M. & Weber, J.C. (1992). Toward a knowledge medium for collaborative product development. In J.S. Gero (Ed.), *Proceedings of the Second International Conference on Artificial Intelligence in Design,* Pittsburgh, PA, June 22-25, 1992. Kluwer Academic Publishers.

Guha, R.V. & Lenat, D.B. (1990). Cyc: A midterm report. *AI Magazine,* Fall 1990, 33-58.

Kaplan, A. (1963). *The Conduct of Inquiry*. New York: Harper and Row.

Lee, J. & Lai, K-Y. (1991). What's in design rationale? *Human-Computer Interaction,* 6(3-4), ??.

Lenat, D.B. & Guha, R.V. (1990). *Building Large Knowledge-based Systems.* Reading, MA: Addison-Wesley.

Lethbridge, T.C. (1991). Creative knowledge acquisition: An analysis. *Proceedings of the 1991 Banff Knowledge Acquisition for Knowledge-Based Systems Workshop,* Banff, Canada, October.

Lethbridge, T.C. & Skuce, D. (1992a). Informality in knowledge exchange. Submitted to the AAAI-92 Knowledge Representation Aspects of Knowledge Acquisition Workshop. San Jose, CA, July.

Lethbridge, T.C. and D. Skuce (1992b). Integrating Techniques for Conceptual Modeling. *Submitted to: 7th Banff Knowledge Acquisition for Knowledge-Based Systems Workshop*, Banff, October

MacLean, A., Young, R.M., Bellotti, V.M.E. & Moran, T.P. (1991). *Human-Computer Interaction,* 6(3-4), ??.

McCarthy, J. (1987). Generality in artificial intelligence. *Communications of the ACM,* 30(12), 1030-1035.

Miller, G. (1990). WordNet: an on-line lexical database. *International Journal of Lexicography.* **3**(4): whole issue.

Moore, E.A. & Agogino, A.M. (1987). INFORM: An architecture for expert-directed knowledge acquisition. *International Journal of Man-Machine Studies*, 26, 213-230.

Mylopoulos, J. (1991). Conceptual modeling and Telos. Technical Report DKBS-TR-91-3. Department of Computer Science, University of Toronto, Toronto, Ontario, Canada, November 1991.

Neches, R. , Fikes, R., Finin, T., Gruber, T., Patil, R., Senator, T. & Swartout, W.R. (1991). Enabling technology for knowledge sharing. *AI Magazine,* Fall, 36-55.

Skuce, D. and I. Monarch (1990). Ontological Issues in Knowledge Base Design: Some Problems and Suggestions. *Proceedings of the Fifth Knowledge Acquisition for Knowledge Based Systems Workshop*, Banff.

Skuce, D. (1991a). A review of 'Building large knowledge based systems' by D. Lenat and R. Guha. *Artificial Intelligence,* in press.

Skuce, D. (1991b). A wide spectrum knowledge management system. *Knowledge Acquisition Journal,* in press.

Skuce, D. and T. C. Lethbridge (1992). A knowledge representation for interactive knowledge management. *Submitted to: Third International Conference on the Principles of Knowledge Representation. and Reasoning*, Cambridge, Mass., October

Sowa, J. (1984). *Conceptual Structures: Information Processing in Mind and Machine*. Reading, MA, Addison Wesley.

Sowa, J.F. (1991). Toward the expressive power of natural language. In J. Sowa (Ed.), *Principles of Semantic Networks.* San Mateo, CA: Morgan Kaufmann.