Guest Editorial

Ontology Research

Christopher Welty

In this issue, I have collected a fairly broad, although by no means exhaustive, sampling of work in the field of ontology research. To define a field is often quite difficult; it is more a collection of people and ideas than it is a specific technology. To represent our field, I present six articles that cover several of the major thrusts of ontology research from the past decade.

Ontology is a discipline of philosophy whose name dates back to 1613 and whose practice dates back to Aristotle. It is the science of what is, the kinds and structures of objects, properties, events, processes, and relations in every area of reality. Ontology is, put simply, about existence. Like so many things, the term was borrowed by computer science and is rapidly becoming a buzzword in industry, tossed about by salesfolk, like all buzzwords, as if it were something everyone knew about. As it turns out, of course, very few people who use the word actually know what it means, and as a result, the actual meaning has changed, and is changing, over time.

All computer scientists who claim allegiance to this field are constantly peppered with the same question, "What is an ontology?" The answer is often argued back and forth by well-meaning people to clarify confusion, but often, the argument causes more confusion than it eliminates. Like many things, one must actually do ontology to understand what it is. I have, however, found that a little history lesson and some discussion can be informative, though not definitive.

Probably the most common citation in ontology is attributed to Gruber, who, in 1993, offered, "An ontology is a specification of a conceptualization." This definition has led the way in causing more confusion than it has eliminated. Others point to Gruber's article as the start of ontology research in computer science; however, the term was already in widespread use by that time, having been used first by John McCarthy in 1980 and subsequently by Hayes in 1984, Sowa in 1984, and Alexander et al. in 1986. The article by Alexander et al. appears to be the first published departure from the philosophical meaning of *ontology* and the start of a new computer science sense of the word. There is then a steady increase in mentions of *ontology* in the AI literature after 1986.

In fact, what the field of ontology research attempts to capture is a notion that is common to a number of disciplines: software engineering, databases, and AI to name but a few. In each of these areas, developers are faced with the problem of building an artifact that represents some portion of the world in a fashion that can be processed by a machine. Since at least 1976, the database community has recognized the key role this process must play in the design of information systems and named it *conceptual modeling*. In software engineering, the introduction of object-oriented systems also led to this realization some time in the early to mid-1980s, and it was named *domain modeling*. It is far more difficult to pin down when this realization was made in AI; certainly scientists were modeling the world in a logical form from the very first days in the late 1950s; however, these models tended to be examples that were used merely to test systems and theories. It was not, it seems, until the era of expert systems that knowledge engineering came into the light as a specific area of study.

In each of these areas, however, it was not until the mid-1990s that it became commonly understood that information systems, built on sound engineering principles, should be able to interoperate but could not. Each of these fields encountered the same problem and realized a similar solution: The meaning of what has been expressed in some formal system is embedded in operational semantics that cannot be divulged easily by inspection. Asking what a symbol such as *author* means in a library system is like asking what the red button on the dashboard does in one of James Bond's supercars: to find out you have to push it. Moreover, you have to push it again and again in a variety of circumstances until you feel you've gained an understanding.

Computer science ontology is, therefore, about meaning. Even more, it is about making meaning as clear as possible. This is an important and crucial point—it is certainly a departure from philosophy—yet it is still fairly vague. After all, a dictionary is supposed to be about meaning; is a dictionary an ontology?

To address this thorny issue, we must go back to the origin of the word. Computer science ontology does still share something in common with its origin—an ontology is not only about meaning, it is also about existence. An ontology can tell you what kinds of things exist in the domain of some system, how these things can be interrelated, and what they mean. Again, despite the em-

phasis, this discussion is still meant to be informative, not definitive.

Another common misconception concerning ontology is that it has something to do with classes versus instances, or logic or entity-relationship diagrams. These notions are actually related more to implementation details than to ontology itself, rather like claiming object-oriented programming has something to do with JAVA. Finally, it is usually a bad idea to identify particular artifacts as ontologies or not.

In this issue, I have collected a fairly broad, although by no means exhaustive, sampling of work in the field, and it should be clear that no such arbitrary distinction was made. To define a field is often quite difficult; it is more a collection of people and ideas than it is a specific technology.

To represent our field, I present six articles that cover several of the major thrusts of ontology research from the past decade. Unfortunately but necessarily, I have left out a number of outstanding researchers in the field—far more than I was able to include.

The article by my colleagues Aldo Gangemi et al. describes work using analysis techniques from formal ontology to "clean up" the WORDNET top level. Although I have avoided giving a definition of what an ontology is, Nicola Guarino and his group have been working on techniques to characterize what "a good ontology" is. This is the first published large-scale application of these techniques, and I should note the work has been welcomed by Christiane Fellbaum and her colleagues working on WORDNET, which is arguably the most widely used lexical resource and candidate ontology today. Some readers might find the discussion a bit heavy—it is certainly steeped in some potentially difficult terminology—however, the issues presented are very relevant to ontology design, and worth the effort.

The article by Mike Uschold represents the perspective of someone who has been in the ontology field for some time, watching it become popular and trendy and now a part of the next generation web—the semantic web. Uschold considers just what it is supposed to mean—figuring out what things mean is, as discussed earlier, part of what ontology is all about.

The article by Julio Arpírez et al. represents one of the largest subfields of ontology—that of tools targeted specifically at ontology designers. To the aspiring ontologist, this place might well be the best point to start because it is fairly straightforward and discusses a particular effort in ontology tool support at the Polytechnic University of Madrid. As with all the articles here, there are certainly other groups working on similar problems; this group was selected mainly because of the breadth of coverage in their tool.

The article by Lee Obrst et al. represents an interesting commercial venture that was based on a significant ontology component. Although the authors ended up, unfortunately, not being able to reveal as much as I would have liked because of commercial privacy concerns, the article does provide a number of interesting perspectives for the nonacademic side of ontology and serves as evidence of the tremendous commercial interest in "doing it right." Personally I find this to be the major contribution of ontology research—there is a real cost benefit to doing things the right way.

The article by Michael Gruninger and Christopher Menzel offers something that truly skirts the boundary of the ontology question and perhaps by itself motivates the discussion that heads up this introduction. Gruninger and Menzel describe PSL—the PROCESS SPECIFICATION LANGUAGE—a logical formalism designed as a standard to support interchange of business and manufacturing process information. PSL should, in theory, support such pie-in-the-sky business dreams as supply-chain integration and virtual enterprises. The inclusion of this article in this special issue should speak for itself about whether a language can possibly be an ontology. Certainly, the terms are not synonymous, but there is quite a bit of ontology work in PSL.

Finally, the article by Martin Doerr represents the world of actual ontology building from a unique perspective—museums and cultural history. Although they certainly do not have a "killer-app" for ontology, Doerr and his colleagues have come across some of the most interesting and challenging ontology problems I have ever seen in trying to develop standards for information interchange in this sphere. Consider, for example, the problem of representing two pieces of pottery that might, or might not, be part of the same pot. How can one catalog such pieces in such a way that someone might discover it? Add a glass or two of wine, and this is the kind of discussion that can occupy an ontology group all evening.

Early versions of many of these articles appeared in the *Proceedings of the Second International Conference on Formal Ontology in Information Systems (FOIS)* (ACM Press, 2001). The conference was held in October 2001 in Ogunquit, Maine.

Acknowledgments

I would like to take this opportunity to thank all those who showed considerable courage in attending the conference and making it a remarkable success.

Chris Welty is a research staff member at IBM's TJ Watson Research Center and serves on the editorial board of *AI Magazine*. His research interests are in formal ontology and the application of knowledge representation and reasoning to solving real problems. He holds a Ph.D. in computer science from Rensselaer Polytechnic Institute.