From: AAAI Technical Report SS-97-06. Compilation copyright © 1997, AAAI (www.aaai.org). All rights reserved.

A Constructivist Approach to the Difficulties of Ontological Engineering

Stephen Regoczei
Computer Studies
Trent University
Peterborough, Ontario
K9J 7B8 Canada
stephen@ai.toronto.edu

Ontology recapitulates philology.
W.V. Quine

The map is not the territory.
Alfred Korzybski

... people understand the world by forming mental models.
John Sowa

1 Introduction and Background Information

This brief position statement summarizes the most relevant aspects of our current thinking on ontologies in general, and ontological engineering in particular. In the paper "Creating the Domain of Discourse: Ontology and Inventory" (1986), Regoczei & Plantinga argued for the wisdom of adopting a constructivist approach to knowledge acquisition and the engineering of knowledge-based systems. The domain of discourse about X is a conceptual model of some fragment X of the world (both physical and social) created collaboratively by a knowledge acquisition analyst and an expert informant. All conceptual models are constructed by cogniting agents. They do not merely just "happen to be out there", ready to be "found". The knowledge acquisition analyst does not merely "discover" the conceptual model, but constructs it in collaboration with the expert. The construction of the ontology of the model may be deliberate, in which case we can talk about design and even engineering. More typically, however, the creation of mental models is unconscious. It is the collaborative knowledge acquisition process which makes them explicit and publicly examinable.

In the same paper, we drew a distinction between ontology and inventory. Ontology, in that context, was the set of entities which might exist, or are considered to be admissible to exist. Inventory, on the other hand, was the presence of actual objects as instances of the ontology elements. Thus a statement such as "There are five chairs in the room" has both an ontology component and an inventory component. The ontology component commits us to considering the existence of chairs and a room. The inventory component states that the number of chairs is five, and not four or eight. (Regoczei & Hirst)

In subsequent papers, Regoczei and Hirst elaborated on ontologies, conceptual analysis in knowledge acquisition, meaning-triangle analysis, and the concept cluster attachment (CCA) approach to meaning generation by people, organizations, or machines.

Currently, we maintain that an ontology for X, very much like a conceptual model of X, has a great deal to do with how we think and how we talk about X. Thus ontologies are, at least in part, based on discourse.

Our main goal here is to note that there is a close connection between designed ontologies and

conceptual knowledge. In addition, to help explicate the relationship between an agent's knowledge of the world and the actual world "out there", we draw a distinction between pure constructs and conceptualized referents.

It is very much worth embracing a few important principles about the relationship between theory and practice. One of these is that practice is helped to a great extent by a good theory, and a good theory is good because it is grounded in practice. In ontological engineering, as in any emerging field, there are numerous difficulties -- both theoretical and practical. I shall try to sketch out our approach to coping with at least some of these difficulties.

2 Ontology Construction as Conceptual Analysis

What is an ontology? The answer given by John Sowa (1984) was that an ontology is the end product of a conceptual analysis process.

> Conceptual analysis is the work of ... systems analysts, and database administrators. ... [they] do it when they translate English specifications into a system design.
>
> Every discipline that uses conceptual analysis gives it a different name. In the computer field, the most common names are *systems analysis*, *enterprise analysis*, and *knowledge engineering*. Whatever the name, the ultimate goal is a precise formalizable catalog of concepts, relations, facts, and principles. ... The result of the analysis is an *ontology* for a possible world—a catalog of everything that makes up that world, how it is put together, and how it works.
>
> Sowa 1984

Having defined what an ontology is, we are still left with the task of describing ontologies clearly. Conceptual graphs are Sowa's preferred notation for the recording of ontologies. Is this our only choice? Are there other explicitly concepts-oriented languages and notations that would be equivalent to conceptual graphs? Do we have other options? Recent work indicates that yes we do.

3 Ontologies are Partly Discourse-Based

The Whorfian hypothesis postulates a close connection between the words and concepts of a language user. This is a hypothesis not only about language use but also about ontologies. To the language user, what exists is largely determined by the words available in the used language. While the Whorfian hypothesis is controversial, and its strong form is definitely an exaggeration, it does point to real problems familiar to all designers and implementers. Language helps to a certain extent, but it also does get in the way. What the user of a piece of equipment does or thinks is partly determined by the surrounding, contextualizing discourse, such as user manuals, metaphors, and explanations.

We can indicate the distinction between what people think and say, versus their actions and other happenings in the physical world, on a 2-level diagram in its frame version form.

2-Level Diagram Frame

Level 2: discourse domain

Level 1: action domain

Let us consider some relatively inexperienced people using a computer. What they do is at the action domain level (Level 1). What they think they are doing, and what the manual may say are two—possibly very, very different—components at the discourse level (Level 2). The constructed, and quite possibly tacit, ontology employed by the users severely limits what they can do, or even understand. The mere text of the user manual will not be of much help. Often a new technical language has to be learned, and the

ontology associated with this language has to be acquired, before the manual can be profitably read and put to good use. Even then, the ontology at the discourse level may be a mismatch when compared to the ontology at the action level. Misunderstandings are bound to occur.

If there is an expert at hand, such as a systems engineer, there may be available a "true" and "real" ontology, namely the all-encompassing (relatively speaking) ontology of the expert (this is especially the case if the expert was the original designer). Still the novice user may not be able to benefit. We could think of the ontology of the expert as a kind of "gold standard" against which the ontological attempts of the user may be compared. So we know the "truth". But if a shared ontology is missing, attempts to help may still fail. The grave difficulties inherent in this little horror story are all too familiar to those who work in HCI or conduct usability audits.

If there is theorizing about these domains, as we have just been doing, we may expand the above frame to n levels. We may decide to add a Level 3: abstraction domain 1, and may even recursively create further Level i: metadomain i-3. Thus theorizing about theories can be indicated by an appropriate notation.

4 How Do We Talk About Things That Don't Exist?

It is a mistake to think of ontologies as containing only what we know to exist. Frequently, we want to think and talk about not only what is there, but also what is not there. Or, perhaps something that is impossible for it to be "there". Russell's classical puzzle concerning the baldness of the "present king of France" is the typical nasty example that pops into mind.

But things don't have to be as bad as that. Often the best way to appreciate the difficulties of ontological engineering is to quote specific examples or case studies. Let's take a look at a story about black holes.

> **Elusive Black Holes Almost a Sure Thing**—Astrophysicists tracking down quarry.
>
> TORONTO—Black holes, those much ballyhooed vacuum cleaners of outer space, have moved from the realm of the highly plausible to that of the almost certain, astronomers said yesterday.
>
> "I would not bet everything I own ... but I would be willing to bet my car" ... Narayan ... said at a news conference ...
>
> The concept of black holes originally grew out of theoretical calculations found in Einstein's theory ...
>
> Two separate pieces of evidence presented at the conference yesterday constituted part of the reason the argument over the existence of black holes has effectively come to an end.
>
> Stephen Strauss, *The Globe and Mail*, January 14, 1997, p. A1-A2

Now the ontological engineer has to ask the embarrassing questions, "Do black holes exist?" and "Did black holes exist?" As the quote indicates, black holes did exist as concepts. The evidence now seems to indicate that currently they may also very well exist as referents. Yet, until I saw this news article, it never occurred to me to wonder about the ontological status of black holes.

This is not surprising. It is a well-documented phenomena that we have no difficulty talking cogently about things and abstractions that don't exist (Hirst 1991). We are forced to acknowledge that the ontologies we need to construct have to have elements both for existing and for non-existing entities. Some people would find this paradoxical. Philosophers over the years have worried extensively about this issue, inventing the concept of "intention" in the process.

5 Using the Meaning Triangle

Conceptual models are of interest precisely because they can be connected -somehow—to the referent domain "out there". It is this connection which provides credibility and validity to an ontology. How this connection is to be established, and how others in, let us say, an engineering or problem solving work group can be convinced of an ontology's appropriateness are two of the best known and least tractable difficulties of ontological engineering.

Furthermore, the constructivist stance forces us to acknowledge that conceptual models are formed by agents. They do not drop from the sky, or from Plato's heaven. Once formed, private ontologies of designers have to be turned into publicly-shared ontologies. This sharing and harmonizing of ontologies is effected through a communication process that incorporates both verbal and non-verbal components. This is a complex matter. The difficulty can be alleviated at least by having some analytical techniques at our disposal. The meaning triangle and the meaning tetrahedron are such knowledge tools. (Regoczei & Hirst, Ogden & Richards, Sowa) We can draw meaning triangles, or we can use a text-only frame notation:

Meaning Triangle Frame

A: Agent (Name:.....) Node: subconceptual/embodied/implicit

W: Verbal Node; text/graphic input/output

C: Conceptual Node: the agent's conceptual models

R: Referent Node: physical objects, group ontology for some group G

6 Actual Observers Are Not Omniscient

Many difficulties connected with ontologies are created needlessly. For instance, it is hard for us not to think of reality as everything that exists. We tend to picture this as everything that could be seen by some all-powerful observer who would be capable of seeing it all. An ontology, then, is everything that can be surveyed by this all-seeing, omniscient being.

The remedy for this difficulty is very simple. We merely have to remind ourselves of the vast gap that exists between the ontology of an omniscient observer—however we may picture it to be—and an actual, limited agent's awareness of that ontology, or even of small parts of that ontology. Actual agents are restricted to the confines of a local observation platform. This comment is similar in spirit to the bounded rationality concept of Herbert Simon.

An extension of the meaning triangle, called the meaning tetrahedron, provides the notation to capture these ideas.

Meaning Tetrahedron Frame

A:....

W:....

C:....: the agent's mental models, including various ontologies

R:....

GOT: Omniscient Observer Node: an ontology of an omniscient observer

We should note the singular and plural forms on nodes GOT and C respectively. An actual agent of limited capacity has several different ontologies in mind, each of greater or lesser validity, according to the available information. These ontologies are, in effect, versions of the world available to the limited-capacity agent. The omniscient observer, on the other hand, has knowledge of the one and only one "true" ontology. (Seager, 1995)

7 Pure Constructs

A construct, sometimes also called a logical construct, is a concept or conceptual structure for which there is no corresponding entity in the "real world" (Sloman) Unicorns are often cited as examples of constructs. For our purposes, the average family with 2.3 children is a clearer example. In a very "real" sense, such a family

does not exist. Sometimes it is characterized as a statistical "fiction".
In general, constructs have no instances in the referent world. They may be concepts denoting classes or categories. They may also be abstractions such that there are no concrete entities corresponding to them. As such, pure constructs should be distinguished from conceptualized referents below.

8 Conceptualized Referents

Conceptualized referents are sortally composite entities. (Regoczei & Hirst) One can think of them as an ordered pair, with the first component being a concept, and the second component being a referent. If I take the general concept [cat] or [table] and I apply them to the actual referent object in front of me, and if I decide that the most appropriate combination to form is

([cat] , < <referent object> >)

then we can safely conclude that I am looking at an actual cat, and not a table. The previously mysterious < <referent object> > is now a conceptualized referent. I can turn it into a named conceptualized referent by forming the 3-tuple

("Puss-puss" , [cat] , < <referent object> >)

Such cognitive feats we manage to perform every day. It is precisely because of the ordinariness of the process that the unpacking of the ontological implications presents us with such great difficulties.

9 Constructing Large Ontologies

As a last point, we may say something about "real-world" versus microworld ontologies. "Real-world" ontologies are typically very large, consisting of tens of thousands of elements. Construction of large-scale ontologies requires team work and project management just like large-scale software engineering. We can think of these ontologies as large-scale abstractionware. Can we be sure that microworld ontology construction techniques will scale up without difficulties? This question is similar to the issues faced in chemical engineering trying to go from laboratory techniques to pilot plants and then to feasible industrial production.
In looking at the construction of large-scale ontologies, our best comparison is not computer software, but rather entertainment software. The creation of films, television series, soap operas, Star Trek, Star Wars, *etc.* all require ontology creation for new metaworlds. Other good examples would include the design of large-scale weapons systems such as aircraft carriers, or even large civilian aircraft. The ontology is vast. The problems are nothing new. The "Doomsday Book" of William the Conqueror was a good early example of large-scale ontology construction.
While ontology construction is a widespread, all-pervasive activity, ontology construction for software development purposes is currently a relatively small component of the total infosphere. But as more and more records are kept in digital form, we are confronted with the need to provide for content ontologies as well as for platform ontologies. Developing multimedia titles is a prime example where content and platform ontologies both need to be considered.

10 Conclusion

It is hard to draw great conclusions from a sketchy, brief list of indications as to why ontological engineering is difficult. The very fact that ontological issues are now discussed in an engineering context should give us encouragement. We should continue the research, because the future success of so many artificial systems depend on arriving at greater insights.

Acknowledgements

I would like to acknowledge the helpful assistance of Graeme Hirst, William Seager, and Michael Gruninger in the clarification of many of the ideas in this summary.

References

Carroll, J. M. 1990. *The Nurnberg [sic] Funnel: Designing Minimialist Instruction for Practical Computer Skill.* Technical Communication Series. Cambridge, Mass.: The MIT Press.

Hirst, G. 1991. Existence Assumptions in Knowledge Representation. *Artificial Intelligence* 49(May 1991):199-242. Reprinted as: Brachman, R. J.; Levesque, H. J.; and Reiter, R. eds. 1992. *Knowledge Representation.* Cambridge, Mass.: The MIT Press.

Ogden, C. K., and Richards, I. A. 1923. *The Meaning of Meaning; A Study of the Influence of Language Upon Thought and of the Science of Symbolism.* London: K. Paul, Trench, Trubner & Co., Ltd.; New York: Harcourt, Brace & Company, Inc.

Regoczei, S., and Hirst, G. 1992. Knowledge and Knowledge Acquisition in the Computational Context. In: Hoffman, R. ed. *The psychology of expertise: Cognitive research and empirical AI.* Cognitive Science Series. New York: Springer-Verlag.

Regoczei, S., and Hirst, G. 1990. The Meaning Triangle as a Tool for the Acquisition of Abstract, Conceptual Knowledge. *International Journal of Man-Machine Studies* 33(5):505-20.

Regoczei, S., and Plantinga, A. 1987. Creating the Domain of Discourse: Ontology and Inventory. *International Journal of Man-Machine Studies* 27(3):235-50. Reprinted in: Boose, J. H., and Gaines, B. R. eds. 1988. *Knowledge Acquisition Tools for Expert Systems.* Knowledge-Based Systems Series. London: Academic Press. pp. 293-308.

Seager, W. 1995. Ground Truth and Virtual Reality. *Philosophy of Science* 62(3):459-478.

Sloman, A. 1978. *The Computer Revolution in Philosophy: Philosophy, Science, and Models of Mind.* Hassocks, England: Harvester Press.

Sowa, J. F. 1984. *Conceptual Structures: Information Processing in Mind and Machine.* Reading, Mass.: Addison-Wesley.

Strauss, S. Elusive Black Holes Almost a Sure Thing—Astrophysicists Tracking Down Quarry. *The Globe and Mail.* January 14, 1997, p. A1 A2.