# *Book Review*

## *Review of Knowledge Representation: Logical, Philosophical, and Computational Foundations*

### *Nicola Guarino*

Like his previous book, *Conceptual Structures* (Addison-Wesley, 1994), Sowa's new book is a unique blend of philosophy, computer science, linguistics, and mathematics. It is intended to be a "general textbook of knowledge-base analysis and design" (p. xi), but I see it more as a collection of intellectual provocations, stimulating exercises, and cross-disciplinary comparisons for the appreciation of a more mature audience. Its great strength is recognizing the need for an interdisciplinary approach, and the attempt at presenting the logical and philosophical foundations of knowledge representation under a unified view. Its great weakness is a lack of consistent rigor, which is needed in a textbook for newcomers to a subject.

After some historical remarks and a first introductory chapter devoted to logic, Sowa immediately attacks the hard problems involved in choosing ontological categories, which lie at the heart of any knowledge representation project. This chapter is the densest (and most problematic), where the basic distinctions used and refined throughout the book are presented. Then the author overviews the main representational paradigms and delves into the muddy waters of times, events, processes, purposes, contexts, and agents. The subtle issues related to the limits of logic when dealing with vagueness, uncertainty, and ignorance are also addressed in some detail, in the context of what Sowa calls the knowledge soup, "a collection of signs—images, symbols, words, and

concepts with associated feelings" (p. 394); this discussion ends with a synthesis of Peirce's and Saussure's accounts of semiotics, the study of signs. A final chapter overviews the problems and the techniques for knowledge sharing and acquisition, including a discussion about relationships between different representation systems. The book is supplemented by an extensive appendix, which includes a sample ontology (with a preliminary axiomatization of the top-level categories), an extended example, and

---

*Knowledge Representation—Logical, Philosophical, and Computational Foundations,* John F. Sowa, Pacific Grove, California, Brooks/Cole, 2000, 534 pp., ISBN 0534949657.

---

answers to selected exercises.

My feeling is that this book is meant for adults only. I mean that although Sowa's first book was suitable for a large audience (and I myself still recommend it to knowledge engineering newcomers), this one is less systematic and more problematic, reflecting years of passionate inquiry into the deepest foundations of conceptual analysis and knowledge representation. What emerges is a vivid picture of the author's peculiar view of history of logic and philosophy, which is deeply intertwined with the analysis of the big problems of knowledge representation and knowledge engineering. Unfortunately, clarity and formal rigor are sacrificed in this highly interdisci-

plinary attempt, so that a good background in logic, philosophy, and knowledge engineering is necessary to avoid being puzzled by the many confusions, inaccuracies, and formal glitches and to extract the useful lessons.

This tolerant and open-minded attitude is especially necessary when reading the first few chapters of the book, the tough ones. In chapter 1, a simple sentence is chosen to introduce the reader to the problems of knowledge representation, "every trailer truck has 18 wheels" (p. 12). This example is analyzed as saying that if $x$ is a trailer truck, then a particular set $s$ of cardinality 18 exists, whose members are both wheels and parts of the truck. Now, this formalization doesn't exclude the case where the truck has more than 18 wheels, because if it had, say, 20 wheels, then the set with 18 wheels would still exist. Of course, the problem is not serious, but together with similar glitches, it leaves the reader with a feeling of approximation that does not help when trying to digest the more substantial (and interesting) parts. Moreover, this chapter is intended to introduce logic through an historical survey, but instead, it mixes together logic, representation problems (measures and musical entities), and logic-based formalisms (knowledge interchange formats and conceptual graphs) in a way that generates in the reader some sense of confusion.

This sense of confusion increases

throughout chapter 2, the most delicate (and crucial) one. I mainly focus on this chapter in this review. One of the chapter's goals is to present the basic ontological categories proposed by philosophers in the past. Categories are introduced as "what in databases are called domains, in AI are called types, in object-oriented systems are called classes, and in logic are called types or sorts" (p. 51). The famous slogan of Quine, "To be is to be the value of a quantified variable," is presented as "a criterion for distinguishing the ontological categories that are implicit in a knowledge representation" (p. 52). This statement seems a first signal of a confusion that will often arise throughout the book, the one between types and tokens or—if you want—between properties and property instances, or individuals. Quine's criterion seems to me useful for deciding on the entities that a certain theory commits to (that is, about its domain), not for distinguishing among the categories it assumes (that is, distinctions among sorts within the domain), because such categories are not quantified over.

The core of chapter 2 is a very personal overview of the history of ontology, from Heraclitus to Heidegger through Aristotle, Kant, Hegel, Peirce, Husserl, and Whitehead. Sowa shows how Aristotle's system of categories (substance, quality, quantity, relation, and so on) was reorganized by Kant in groups of three, more or less according to the thesis-antithesis-synthesis pattern. These triadic patterns, according to Sowa, were considered by Kant as a fundamental aspect of reality, further explored by Hegel, and ultimately by Peirce. Indeed, Sowa's main thesis is that nearly all the ontological systems proposed by classic philosophers hinge on Peirce's three basic categories: (1) firstness, (2) secondness, and (3) thirdness. Unfortunately, no attempt is made to formally clarify the meaning of these categories, which are used as fundamental primitives throughout the book. The only "definition" that Sowa offers is in terms of Peirce's own words:

First is the conception of being or existing independent of anything else. Second is the conception of

being relative to, the conception of reaction with, something else. Third is the conception of mediation, whereby a first and a second are brought into relation. (p. 60)

In addition to this description, Sowa says that firstness can be "defined by a monadic predicate," secondness by a "dyadic relation," and thirdness by an "irreducible triadic relation" (p. 61). This informal introduction of Peirceian categories is accompanied by an attempt toward a more rigorous characterization in appendix B2, which unfortunately does not answer the reader's doubts. For example, firstness (independence) is characterized in terms of a further primitive, the "has" relation informally introduced on pages 84 to 85, in the following way: "if $x$ is an independent entity, it is not necessary that there exists an entity $y$ such that $x$ has $y$ or $y$ has $x$" (p. 499). Because the interpretation of "$x$ has $y$" includes the case where $y$ is a proper part of $x$ (p. 84), this means that every entity that necessarily has some proper parts (that is, is nonatomic) is not independent; so, this definition excludes most of the entities considered by Sowa as independent (such as those mentioned in exercise 2.2) from being such. The rest of the book is full of examples of application of Peirce's threefold pattern, which unfortunately don't help to dissolve the doubts raised by the previous quotation. For example, I can't understand how emotions, which are described as dependent on external stimuli or internal bodily conditions, can be mapped onto the three Peirceian categories (p. 67), including firstness, which is assumed to be independent.

Sowa's deepest attempt to use Peirce's distinctions as a unifying ontological paradigm is his personal revisitation of Whitehead's categories: actual entity, prehension, and nexus. Although Sowa admits that "Whitehead never mentioned Peirce" (p. 63), he considers these categories as strictly reflecting Peirce's. If successful, this unifying attempt could have been the most intriguing contribution of the whole book; unfortunately, however, the matter addressed is difficult, and the results are terribly confusing (and sometimes just inconsistent). Again,

the problem is that Whitehead's categories are presented in an informal and obscure way, by including a few quotations with no attempt to explain or formalize them:

For Secondness, he [Whitehead] used the term prehension for "concrete fact of relatedness." He explained "that every prehension consists of three factors: (a) the subject which is prehending, namely, the actual entity in which that prehension is a concrete element; (b) the datum which is prehended; (c) the subjective form which is how that subject prehends that datum."

For Thirdness, Whitehead adopted the Latin word nexus [I will use the plural nexuses] which represents an instance of connecting or binding together two or more actual entities: "Actual entities involve each other by reason of their prehensions of each other. There are thus real individual facts of the togetherness of actual entities, which are real, individual, and particular, in the same sense in which actual entities and the prehensions, are real, individual, and particular. Any such particular fact of togetherness among actual entities is called a nexus" (pp. 63–64).

It took me months to grasp and digest these notions. I admit I haven't read Whitehead, but this reading shouldn't be a prerequisite for an introductory book….Eventually, I got to this simplified picture of Whitehead's terms: Suppose I sell my car to John Sowa; my own participation to this event is my prehension of it, which, of course, is different from John's. The selling event is the nexus that includes me and John as participants. However, why should a prehension be considered a secondness if it involves three factors? In addition, why is a "fact of togetherness" "an instance of connecting"? Isn't the prehension between entities the mediating factor within a nexus?

Unfortunately, things become more confusing when Sowa applies Peirce's distinction to the abstract counterpart of Whitehead's basic categories.

According to Sowa, Whitehead "classified the abstractions in the categories of Eternal objects, Propositions, and Subjective forms, which constitute a triad of abstract Firstness, Secondness, and Thirdness" (p. 64). Propositions are therefore the counterpart of prehensions, and subjective forms such as emotions and purposes are the counterpart of nexuses.

This mapping between and Whitehead's and Peirce's categories sounded very strange to me because subjective forms and propositions appeared to be switched in their correspondence to prehensions and nexuses; propositions seem to me as a nexus composed of a predicate and a logical subject, which prehend each other. After some serious puzzlement, I finally decided to resort to *A Key to Whitehead's Process and Reality* by Donald W. Sherburne,[1] which apparently confirms my view.

Indeed, on reading Sherburne's explanations, I realized that Whitehead's notion of nexus might be different from Sowa's because Whitehead sees it more as composed of actual entities that are interrelated through their prehensions of one other rather than as a mediating entity that links the actual entities together. A person, for example, is seen by Whitehead as an example of nexus (a "structured society"), but Sowa considers it an actual entity. Under this view, a nexus should be classified under firstness. Moreover, a prehension should better be seen as a thirdness because it "mediates" between the prehending and the prehended entities. Indeed, intentionality, which Sowa considers thirdness, has been seen in philosophy as an example of prehension (Gier 1976).

In sum, this "synthesis" of Peirce and Whitehead is very problematic. Unfortunately, the consequences of this reconstruction are not limited to the philosophical background chapter because the whole book hinges on six top-level categories, obtained by adding to the physical trichotomy actuality-prehension-nexus its abstract counterpart form-proposition-intention (Whitehead's names have been changed slightly). As discussed earlier, the most serious problem here is that the intuitions behind prehension and nexus seem to be switched with respect to Sowa's interpretation.

These six categories are combined with a further dichotomy, the distinction continuant-occurrent, to form a top-level ontological lattice known as Sowa's diamond (p. 71). This lattice is presented as the final result of a deep work of philosophical unification, although the author wisely remarks that "the philosophers who inspired this approach can no longer comment on how those categories may capture or distort their insights" (p. 68).

Sowa's diamond is a nice symmetric structure, which accounts for all possible combinations of the basic distinctions: independent-relative-mediating, physical-abstract, and continuant-occurrent. Continuants and occurrents are technical names taken from philosophical literature, which can be intended as synonymous with physical objects and events (or maybe processes; terminology is not important here). The problem, however, is that—in the way Sowa introduces them (p. 71, pp. 499–500) and in the usual philosophical understanding—they are both physical entities, in the sense that they are located in space-time; so, no abstract continuants or occurrents exist. However, according to Sowa, a script that describes an occurrent, like a musical score, is an abstract occurrent, and a proposition that describes a continuant is an abstract continuant (p. 74). The feeling I have is that the desire of formal symmetry prevailed over plain common sense.

After the presentation of the top level, a variety of deep ontological issues is addressed in chapter 2, with no attempt at systematicity. This "casual" style continues in the next chapter, which is devoted to knowledge representation. Chapter 3 is more a collection of (often interesting) comments on the history of AI than an introductory exposition of the most fundamental concepts. A serious glitch concerns the way Brachman's knowledge representation levels are relabeled "knowledge levels" (p. 186–187). Indeed, knowledge representation aspects are often intermixed with knowledge analysis aspects in the whole book.

The subsequent chapters, devoted to the modeling issues behind processes, purposes, contexts, and agents, are a bit more technical and less problematic than the preceding ones. The underlying ontological assumptions, which have been introduced in the previous chapters, are not clarified, however. The delicate issue of semiotics (the theory of sign) is addressed in the chapter on the knowledge soup, toward the end of the book. This topic is one of Sowa's favorites. He tries to generalize Peirce's trichotomy introduced previously, presenting a trichotomy of trichotomies (p. 397). Again, this discussion might be stimulating, but at a closer look, it remains as definitely mysterious (and sort of obsessive) to me.

To conclude this review, I recommend this book to those who want to better understand Sowa's view of the world, which is presented in an intriguing and vivid style, and to learn more about using conceptual graphs for hard representational problems. I definitely do not recommend this book as a general textbook on knowledge representation.

## Note

1. Reprinted in 1981 from University of Chicago Press. An excerpt of the most important terms is available online at http://hyattcarter.com/glossary.htm.

## References

Gier, N. 1976. Intentionality and Prehension. *Process Studies* 6(3): 197–213.

**Nicola Guarino** is a senior research scientist at the Institute for Systems Theory and Biomedical Engineering of the Italian National Research Council (LADSEB-CNR), Padova. He graduated in electrical engineering at the University of Padova in 1978. He has been active in the ontology field since 1991. His current research activities regard ontology design and ontology-based metadata standardization.