Issues in Commonsense Set Theory

Müjdat Pakkan and Varol Akman

Department of Computer Engineering and Information Science Bilkent University, 06533 Bilkent, Ankara

Abstract

This paper discusses commonsense set theory. First a brief review of classical set theory is given. Then the need for a commonsense set theory is questioned and the properties of a possible theory are examined. Finally, previous work in the area is presented.

1. BRIEF REVIEW OF EXISTING SET THEORIES

Set theory is a branch of modern mathematics with a unique place because other branches can be formally defined within it [1, 2]. The theory had started with the work of Cantor on infinite series. In his early conception, he thought of a set as a collection into a whole of definite, distinct objects of our perception or thought [3]. It did not take a very long time for this theory to be revised. Several paradoxes, including Russell's famous paradox of "the set of all objects which have the property of not being members of themselves," were introduced, leading to new axiomatizations of the theory [4].

Among the various axiomatizations, are Russell and Whitehead's *Theory of Types* [5] which brings in a hierarchy of types to forbid circularity and hence avoid paradoxes, and von Neumann, Bernays, and Gödel's NBG which introduces the notion of a class for the same purpose [6]. Strengthening NBG, a new theory, called MK (Morse-Kelley) was obtained [7, 8]. This theory is suitable for mathematicians who are not interested in the subtleties of axiomatic set theory. In 1937, Quine proposed his *New Foundations*, NF, to overcome some unpleasant aspects of the Theory of Types while keeping the main idea same [9]. He introduced the notion of *stratification* for this purpose. NF is a nice theory, avoiding Russell's Paradox, and allowing all mathematics to be defined within it, but has some strange properties. For example, Cantor's Theorem, $a < 2^a$, cannot be proven in NF, because a set used in the proof is not defined since it is not stratified. In general, the most popular axiomatization is ZF, originated by Zermelo [10] and later modified by Fraenkel [11]. The intention was to build up mathematics by starting with the empty set and then construct further sets cumulatively by various operators. This hierarchy works as follows [12].

Initially, the Axiom of Extensionality says that a set is completely determined by its members. The *Null Set Axiom* states the existence of the empty set $\emptyset$. The *Pair Set Axiom* then constructs the sets $\emptyset$, $\{\emptyset\}$ and $\{\emptyset, \{\emptyset\}\}$, and other one and two element sets. It is the *Sum Set Axiom* which constructs sets with any finite number of elements by

Figure 1: ZF universe extending in a cumulative hierarchy

Figure 2: AFA representation of the circular set $\Omega = \{\Omega\}$

defining union of existing sets. Finally, *the Axiom of Infinity* states the existence of at least one infinite set from which others can be formed, viz. $\{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}, \ldots\}$. This hierarchy can be depicted as in Figure 1. There is also the important *Axiom of Foundation* which basically prohibits sets which are members of themselves, thus avoiding the classical paradoxes.

In spite of its popularity in current research, ZF has its drawbacks. It is weak to decide some questions like the *Continuum Hypothesis.* It is also weak in some applications (set-theoretic, linguistic) which make use of self-reference, since circular sets like $x = \{x\}$ are prohibited by the Axiom of Foundation. ZF is also strong in some ways, viz. it violates the *parsimony principle* which states that simple facts should have simple proofs [13]. (This can be seen in the use of the Power Set Axiom in the proof of a simple notion like $a \times b$.) It can also be claimed that it is bad for mathematical practice that all the mathematical objects are to be realized as sets.

New theories have been developed throughout the century besides the huge amount of research in ZF. The *Admissible Set Theory* which originated in the Sixties is one of them. Work on admissible ordinals resulted in a first order set theory called KP (Kripke-Platek). Barwise weakened KP to a new theory KPU by readmitting *urelements* which are indivisible individuals [13]. KPU is an elegant theory which supports the cumulative hierarchy. It overcomes some of the disadvantages of ZF and seems to obey the parsimony principle. But it still cannot handle circular sets.

It is the *Hyperset Theory* which can do that. The origins of this theory are in the late Twenties, but it evolved throughout the century. The theory makes use of graphical representation of sets. In the Eighties, Aczel got interested in the subject and proposed the *Anti-Foundation Axiom (AFA)* [4]. With the existence of AFA, one can easily represent circular sets as in Figure 2. But what is more remarkable with the theory is that it does not require other axioms of ZF to be thrown out. Aczel's theory was successfully used by Barwise and Etchemendy in their thought-provoking work on the *Liar Paradox* [14].

2. COMMONSENSE SET THEORY

Representing all compound entities and the relations between their parts in terms of sets gave rise to the success of set theory in mathematics. This also seems to apply to commonsense reasoning as well. McCarthy has mentioned the possible use of set theory in AI and invited researchers to concentrate on the subject [15]. Unfortunately, not much progress has been made since then. In the sequel, we will study some noteworthy research essentially due to Perlis [16], Zadrozny [17], and Mislove et al. [18].

If we want to design an intelligent machine which will work in the realm of human beings, then we must make sure that it has commonsense knowledge and that it is able to make inferences from that knowledge. Commonsense reasoning involves three main parts: a domain theory, knowledge representation, and inference. We are primarily dealing with knowledge representation issues. The first idea is to represent commonsense notions by sets. To take a classical example, we can consider the commonsense notion of *society* as a relation between a set of people, rules, customs, traditions, etc. Here, we face the problem that commonsense ideas do not have precise definitions as mathematical ideas do. For example, in the definition of *society,* the notions of *tradition* and *custom* are as complex entities as the definition itself, and should best be left to intuition. However, sets may still be useful in conceptualizing such terms. One may, for instance, want to consider the set of societies disjoint from any set of individuals. Moreover, the idea of collecting a set of individuals for further thought is still an important process, e.g., the Comprehension Principle.

A theory proposed for use in commonsense reasoning can be examined in a variety of ways. The first point to examine is the concept of set formation, immediately leading to the question whether to admit urelements or not. It seems intuitive to answer affirmatively because this matches with the naive notion of a set as collecting individuals satisfying a property into a whole.

But then we directly face Russell's Paradox. The problem is due to using a set whose completion is not over yet in the formation of another set, or even in its own formation. Then we are led to the question of when to consider a set of individuals satisfying a property as an individual itself. This brings the notion of *cumulative hierarchy* into picture. Cumulative hierarchy is one common construction of our intuition and can be illustrated by the example in Figure 3 where we start with simple blocks, and make towers out of blocks, and make walls out of towers, and so on.

In the cumulative hierarchy, any set formed at some stage must be consisting of urelements (if any) and the sets formed at some previous stage. At this point the problem of "sets which can be members of themselves" arises, because they are used in their own formation. Circularity is obviously a common means of commonsense knowledge representation. For example, we should allow the unique set of all non-profit organizations to be a member of itself, since it may also be a non-profit organization (which is not an unexpected event). So, a possible theory should allow circular sets. Barwise and Etchemendy's work also demonstrates beyond doubt that circularity is an integral part of our daily discourse and hence should be allowed by a commonsense set theory [14].

One further aspect to be considered is "possible" membership. A commonsense set theory may be helpful in providing representations for dynamic aspects of language by

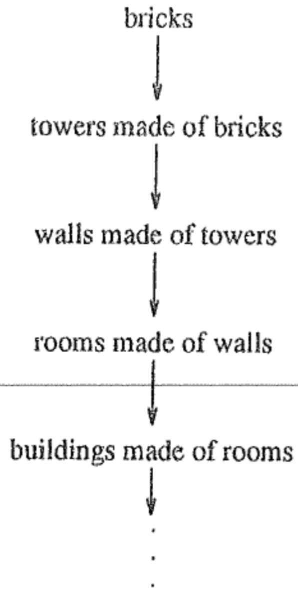

Figure 3: A simple hierarchical construction representing the cumulative hierarchy

making use of *partiality.* We can allow our sets to have possible members as well as real members. Then we can have an operation of *clarification* to determine the real members among possible ones.

Other-set theoretical issues to be considered include cardinality and well-ordering. For example, imagine a box of 34 black and 16 white balls. We know that there are 50 balls in the box, or formally the cardinality of balls in the box is 50. After shaking the box, we would say that the balls are not ordered any more. But this is not true in classical set theory since a set with finite cardinality must have a well-ordering [17]. Counting is another aspect to be considered. While the formal principles of counting are precise enough, we may observe that people also use quantifiers like "many," "more than half." For example, a system which can represent the phrase "A group of people are walking towards me" should not probably answer questions like "Who is the first one?" since there does not exist a well-ordering for the set under consideration.

There is relatively little work done in commonsense set theory. Perlis introduced a series of theories for this purpose. He first proposed CST_0 which is a version of naive set theory with only an axiom of comprehension. Then he extended it to CST_1 using Ackermann's Schema [19] to support cumulative hierarchy but still could not handle circular sets with it. Finally he proposed CST_2 by combining CST_1 with the universal reflection theory of Gilmore and Kripke [20]. While this theory can represent circular sets, its consistency is not proven yet.

Zadrozny, who does not believe in a "super" theory for commonsense reasoning, concentrated on cardinality and well-ordering issues mentioned above and proposed some representation schemes. He introduced a non-standard class of *Nums* for counting purposes and using *Nums*, he gave non-classical interpretations of cardinality and well-ordering. In the context of his interpretations, he proved that there are sets with finite cardinality which do not have a well-ordering and that there are well-orderings elements of which do not form a set (hence solving the Box Problem above).

Mislove, Moles and Oles worked on *protosets*, a generalization of HF, the set of well-founded hereditarily finite sets [18]. Protosets are sets with some packaging which might obscure some elements of the set. They proposed their *Partial Set Theory* based on protosets, with a relativization of Aczel's work. Their theory is consistent with respect to Aczel's, because it is a conservative extension of the latter.

3. CONCLUSION

We conclude that set theory can be useful in commonsense reasoning. The methodology may change: a universal commonsense set theory may be developed, or different set theoretic concepts may be examined and modified. No matter what proposal is followed, we believe that research in this field is promising, considering the current success of set theory in mathematics.

References

[1] P. Suppes. *Axiomatic Set Theory.* Dover, New York, NY, 1972.

[2] P. R. Halmos. *Naive Set Theory.* Springer-Verlag, New York, NY, 1974.

[3] G. Cantor. Fondaments d'une théorie générale des ensembles. *Acta Mathematica*, 2:381–408, 1883.

[4] P. Aczel. *Non-Well-Founded Sets.* Number 14 in CSLI Lecture Notes. Center for the Study of Language and Information, Stanford, CA, 1988.

[5] A. N. Whitehead and B. Russell. *Principia Mathematica (Three Volumes).* Cambridge University Press, Cambridge, UK, 1910–1913.

[6] J. von Neumann. Eine Axiomatisierung der Mengenlehre. *J. für Math.*, 154:219–240, 1925.

[7] J. Kelley. *General Topology.* Van Nostrand, Princeton, NJ, 1955.

[8] A. Morse. *A Theory of Sets.* Academic Press, San Diego, CA, 1965.

[9] W. V. Quine. New Foundations for Mathematical Logic. *Amer. Math. Monthly*, 44:70–80, 1937.

[10] E. Zermelo. Untersuchungen über die Grundlagen der Mengenlehre. *Mathematische Annalen*, 65:261–281, 1908.

[11] A. A. Fraenkel. Zu den Grundlagen der Cantor-Zermeloschen Mengenlehre. *Mathematische Annalen*, 86:230–237, 1922.

[12] M. Tiles. *The Philosophy of Set Theory.* Basil Blackwell, Oxford, UK, 1989.

[13] J. Barwise. *Admissible Sets and Structures.* Springer-Verlag, Berlin Heidelberg, Germany, 1975.

[14] J. Barwise and J. Etchemendy. *The Liar: An Essay on Truth and Circularity.* Oxford University Press, New York, NY, 1987.

[15] J. McCarthy. Acceptance address, 1985. IJCAI Award for Research Excellence, Los Angeles, CA.

[16] D. Perlis. Commonsense Set Theory. In P. Maes and D. Nardi, editors, *Meta-Level Architectures and Reflection*, pages 87–98. Elsevier, Amsterdam, Holland, 1988.

[17] W. Zadrozny. Cardinalities and Well Orderings in a Common-Sense Set Theory. In R. J. Brachman, H. J. Levesque, and R. Reiter, editors, *Proceedings of the First International Conference on Principles of Knowledge Representation and Reasoning*, pages 486–497. Morgan Kaufmann, San Mateo, CA, 1989.

[18] M. W. Mislove, L. S. Moss, and F. J. Oles. Partial Sets. In R. Cooper, K. Mukai, and J. Perry, editors, *Situation Theory and Its Applications*, pages 117–131. Center for the Study of Language and Information, Stanford, CA, 1990.

[19] W. Ackermann. Zur Axiomatik der Mengenlehre. *Mathematische Annalen*, 131:336–345, 1956.

[20] P. Gilmore. The Consistency of Partial Set Theory without Extensionality. In T. Jech, editor, *Axiomatic Set Theory*, pages 147–153. American Mathematical Society, Providence, RI, 1974.