AXIOMATIC SET THEORIES AND COMMON SENSE

Mehmet Surav and Varol Akman

BILKENT UNIVERSITY

Department of Computer Engineering and Information Science

Technical Report BU-CEIS-94-11

AXIOMATIC SET THEORIES AND COMMON SENSE

Mehmet Surav and Varol Akman
Department of Computer Engineering and Information Science
Faculty of Engineering, Bilkent University
Bilkent, Ankara 06533, Turkey

ABSTRACT

Various axiomatic set theories (ZF, NBG, NF, and KPU) are studied with a critical eye. The basic mathematical and philosophical reasons behind their axioms are given, as well as their review from the commonsense point of view. An introduction to a "commonsense set theory" is attempted at the end.

1. Introduction

In the foundations of mathematics, one of the most popular approaches is set theory. Almost all of the mathematical objects can be constructed out of sets. In this view, mathematics deals only with the properties of sets and all of these properties can be deduced from a suitable list of axioms [14, 21].

In an application of set theory to commonsense knowledge representation and reasoning, two basic mental (cognitive) processes are abstraction and categorization. In fact, a set might be considered as the image (in mind) of a collection of things under some abstraction mechanism. This mechanism should be similar to the one Cantor thought [13, 11, 32, 36, 38] and can use the color, size, use, location, and other things that might be useful in this categorization process. (Here we did not say "property" instead of "thing" because the differentiation of the objects might not depend on the object, viz. location.)

According to Piaget [27], human abstraction mechanism is different from a simple (Aristotelian) mechanism. Piaget calls this "reflexive abstraction." The Aristotelian mechanism is that given some external object, such as a crystal and its shape, substance, and color, the subject simply separates different qualities and retains one of them—the shape, maybe—rejecting the rest. Usually, the Piagetian categorization mechanism depends on the person (his personal cognitive capabilities) and the situation in which this categorization occurs. Technically speaking, if we are to construct a commonsense set theory, we have to consider the situations (thus situation theory and situation semantics (STASS)) [4, 5, 6], and define the axioms of our system to be somewhat compatible with the axioms of STASS. However, the axioms of a commonsense set theory cannot be totally independent of the classical (axiomatic) set theory [2, 26]; the development of the classical theory sheds considerable light into our axiomatic system and must not be summarily rejected.

In this paper, various classical set theories (ZF, NBG, NF, and KPU) are discussed. We will examine each axiom of these theories, both from the mathematical and philosophical points of view. If applicable, we will also study each axiom from a commonsense angle.

2. Zermelo-Fraenkel (ZF)

In any axiomatization, axioms can be used in three ways: (i) to express the basic truths about the universe of objects of interest, (ii) to give the basic building blocks and the principles of construction of a universe of objects, and (iii) to give the rules of the "game to be played" with newly introduced symbols [36].

In general, ZF is *the* basic axiomatization used heavily in mathematics. The origin and the underlying mathematical properties and results of its axioms were extensively discussed in the literature [9, 13, 15, 16, 17, 18, 19, 20, 30, 31, 33, 35, 36, 37]. The axioms are defined in first order logic and only the membership relation ($\in$) is considered to be a basic relation. The axioms of ZF are *Extensionality*, *Null Set*, *Pair Set*, *Union*, *Infinity*, *Power Set*, *Separation (Subset)*, *Replacement*, and *Foundation (Regularity)*. *Choice* is not an original axiom of ZF, but is a must vis-à-vis some advanced topics such as large cardinals, independence proofs, and the continuum hypothesis.

Extensionality.

$$\forall x \forall y \forall z[(z \in x \leftrightarrow z \in y) \rightarrow x = y]$$

Basically, this axiom defines the notion of being a set: a set is a collection of elements, whose identity is completely determined by those members. In some places [13, 28], equivalence of two sets is also defined as a basic notion. In this case, *Extensionality* is used to define the property of this relation [36].

From a mathematical point of view, this axiom is one of the least problematic ones, but from a philosophical vantage point and from the commonsense side, things are not so simple.

First, a questionable thing is $\in$. This relation might be considered with modal operators, i.e., we might think of possible memberships, believed memberships, etc. Consider the barbers of Springfield (home-town of Bart Simpson). There might be three barbers working for money, and one barber who does not work for money (since he has another job) but serves the poor children. If we look at the situation from the commonsense perspective, there are four barbers in the town, but from say, the mayor's point of view, there are only three (official) barbers. Thus, the context is crucial and we must have some knowledge about the situation we are facing.

Another issue is fuzzy membership. There is a probabilistic approach, i.e., a is a member of b with probability $x\%$, which might result in mathematically and philosophically interesting points and discussions, but is not very attractive for a commonsense set theory. Normally, people would think of straightforward membership and non-membership, not of a probabilistic relation.[1]

Consider again Springfield.[2] Springfield Fire Department and Springfield Barber-shop Quartet might have the same staff members. Are these two equivalent? We hope not.

Null Set.

$$\exists x \forall y \neg(y \in x)$$

This axiom guarantees the existence of a set with no members (i.e., empty set $\emptyset$). $\emptyset$ is unique according to *Extensionality*. In ZF, it is the only set whose existence is guaranteed. This is a good feature for the mathematicians since there are no problematic objects (e.g., people, trees, collections, etc.) other than what they have created for their own purposes [13]. Pure ZF does not support urelements[3] but urelements are frequently useful: the universe of commonsense reasoning and knowledge is sometimes the real world, sometimes the thoughts, and sometimes abstract things like $\emptyset$.

In our daily vocabulary, the corresponding word for $\emptyset$ might be "nothing," but does this really agree with the the empty set of ZF? Here, the notion of oppositeness and complementarity comes to mind. In daily life, the opposite of "nothing" matches anything other than nothing. From this point of view, the complement of $\emptyset$ must be any set other than $\emptyset$, but from a mathematical point of view, the complement of empty set is the "universal set."

Pair Set.

$$\forall x \forall y \exists z \forall w(w \in z \leftrightarrow w = x \vee w = y)$$

[1] Here, we are omitting the "not-sure" cases, since being unsure is quite similar to a "don't know." In this case, the conclusion is usually obtained by omitting this knowledge.

[2] Example suggested by M. A. Jorgensen on July 20, 1993 in a discussion in the news group `sci.math`.

[3] Urelements are the objects (or individuals) with no elements. They are sometimes called "atoms."

Basically, this axiom enables us to form $\{a, b\}$ from the sets a and b. In the beginning, taking $a = b = \emptyset$ gives us the set $\{\emptyset\}$. This is the second set we form in ZF. Using this axiom, we can generate finite number of sets with only one or two elements.

This axiom seems to be an innocent one from the mathematical point of view. However, from philosophical and commonsense views, it is somewhat problematic. Consider the example, a = "my coffee cup," and b = "my leather suitcase with two books in it." If we put these two objects into a plastic bag, we form a set. The members of the set is a suitcase and a coffee cup. However, the books might be considered as members of the new set. But with this pairing schema, the books are only members of the leather bag. Similar issues are raised by Zadrozny [40] in his study of the *part-of* relations for multi-media indexing. Zadrozny introduces *part-of* to deal with membership in structured objects (e.g., a first-aid kit). A hereditary membership relation can be a solution in our case (i.e., using a hereditary membership relation we can say that books are the members of plastic bag).

Union.

$$\forall x \exists y \forall z [z \in y \leftrightarrow \exists w (w \in x \wedge z \in w)]$$

This axiom states the existence of $\bigcup a$, and specifies its members:

$$\bigcup \{a_1, a_2, ..., a_n\} = a_1 \cup a_2 \cup ... \cup a_n$$

It assures the existence of a set containing any desired finite number of elements by asserting the existence of the union of sets already defined. Using this axiom, we can generate sets with three or more members; this was so far impossible. However, all the sets generated up to this point (starting from $\emptyset$, using *Pair Set* and *Union*) are finite. This is because only finitely many sets can be proven to exist by finite applications of *Pair Set* and *Union*.

Infinity.

$$\exists x [\emptyset \in x \wedge \forall y (y \in x \rightarrow \exists z (z \in x \wedge \forall w (w \in z \leftrightarrow w \in y \vee w = y)))]$$

This states that there is a set which has $\emptyset$ as an element and which is such that if a is an element of it then so is $\bigcup \{a, \{a\}\}$. In other words, the axiom guarantees the existence of $\{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}, \{\emptyset, \{\emptyset, \{\emptyset\}\}\}, ...\}$. This set is well-ordered with respect to $\in$ and has cardinality ω, the first infinite ordinal.[4]

A set a is called *finite* if there exists a natural number n such that a is equinumerous (has the same cardinality) with $\{0, 1, ..., n-1\}$. A set a is called *denumerable* if it is equinumerous with the set of all natural numbers. A set is called *reflexive* if it is equinumerous with a proper subset of itself. Clearly, every reflexive set is infinite, since the set $\{0, 1, ..., n-1\}$ cannot be reflexive for any natural number n.

Now, how many types of infinity do we have in mind in our daily life? We think that there is only one, i.e., the one denoted by the word "infinity." In daily use, infinity means that the members of the set cannot be counted and that the set is not accessible. But, before drawing the conclusion "there is a unique commonsense infinity," we have to be very careful about the antinomies of Russell and Burali-Forti [13, 36].

From a mathematical point of view, there is no controversy so far (i.e., the first five axioms). In fact, these axioms are necessary for any mathematical axiomatization of sets, even for doing simple arithmetic [10, 11, 14, 21, 32, 33, 36]. With the help of Peano axioms and the constructs up to this point, we can define integers, rationals, reals, complex numbers, and derive their usual arithmetic and analytic properties [15]. Peano axioms are as follows:

[4] The above definition of infinity was not quite straightforward. Before the above definition (or definitions that are equivalent to it), there were serious attempts by Dedekind and Bolzano [13], that led to Russell's antinomy.

1. $0 \in \omega$, taking $0 = \emptyset$;
2. if $n \in \omega$ then $succ(n) \in \omega$, where $succ(n) = n \cup \{n\}$;
3. if $(S \subset \omega) \wedge (0 \in S) \wedge (n \in \omega \rightarrow succ(n) \in S)$ then $S = \omega$;
4. $succ(n) \neq 0$, for all n in ω;
5. if $(n \in \omega) \wedge (m \in \omega) \wedge (succ(n) = succ(m))$ then $n = m$.

These axioms clearly correspond to the basic counting mechanism of humans. The axioms begin with 0 and and continue with the next counting number in each step. This approach to the natural numbers and counting is definitely a powerful tool for a commonsense set theory. The principle of induction say that the set of natural numbers are infinite (since we can always find the following number for each number). The induction mechanism of humans is also similar to this principle, thus the principle is commonsense acceptable. According to the fourth axiom, we cannot find a natural number smaller than 0, in other words natural numbers begin with 0 (so counting starts from a predefined point).

Power Set.

$$\forall x \exists z \forall z [z \in y \leftrightarrow \forall w (w \in z \rightarrow w \in x)]$$

According to this axiom, if a is a set, then there is a set $P(a)$, the power set of a, whose elements are all the subsets of a. The cardinality of the power set is always greater than the original set. (Cantor used this fact in his transfinite numbers [13, 36].) Using this axiom, one can prove that for every infinite set, there is a "larger" infinite set (i.e., its power set). Thus the existence of some "large" cardinals (e.g., $\aleph_1$ (cardinality of real numbers), $\aleph_2, \aleph_3$, etc.) can be proven from $\aleph_0 = \omega$, and size comparisons can be made.

In this axiom, the method of subset formation is not given and is postponed to the next axiom.

Separation (Subset) and *Replacement.*

The axiom schema of separation can be stated as

$$\forall x \exists y \forall z [z \in y \leftrightarrow z \in x \wedge \varphi(z)]$$

where x and y are not free in φ.

This axiom schema really offers a restrictive option over the sets. Until now, all of the axioms were used to create the universe of sets. With this axiom, we have a mechanism to select an appropriate subset for our operations. This selection is done using the well formed formula (wff) $\varphi(z)$. Whatever the condition is imposed by $\varphi(z)$, the axiom produces a subset a_φ of given a set a. (At this point, it essentially refers to the power set of a, i.e., $P(a)$.)

Replacement is a stronger version of *Separation*:

$$\forall x [\forall y [y \in x \rightarrow \exists z (F(y,z) \wedge \forall w (F(y,w) \rightarrow w = z))] \rightarrow \exists v \forall u [u \in v \leftrightarrow \exists t (t \in x \wedge F(t,u))]]$$

From a mathematical point of view, *Separation* does not provide the workspace that mathematicians would like to have. In mathematics, functions, transformations, and mappings from one definite domain to another range are frequently needed. In this respect, the restriction of the range to be a subset of the domain seems unnecessary. For example, we can cite the following: using *Separation*, we cannot prove the existence of the ordinal $\omega + n$ where n is a finite ordinal.

The last thing to note is that both *Replacement* and *Separation* are indeed axiom schemata. This fact is often stated as "ZF is not finitely axiomatizable" and was proved by Mostowski [13].

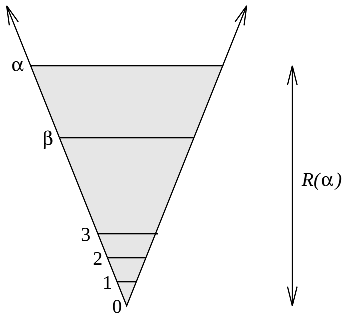

Figure 1: WF universe

Using the above axioms, we can prove the existence of intersection ($a \cap b$), the existence of outer product of a set ($\Pi(t) = a \times b \times ...$ for $t = \{a, b, ...\}$), and the specific case of Cartesian product using $P(P(a \cup b))$ for $a \times b$.

We can also define relations and functions. Then we can talk about the well-ordering of sets and size comparisons over sets. In short, we can develop all simple mathematics (e.g., the sets that we learned in high school and that we usually identify with Venn diagrams). Here, we are using the term set in a loose manner, and large collections and sets usually require the Axiom of Choice to predict such properties (e.g., the real numbers cannot be shown to be well-ordered without reference to *Choice*) [13].

From the commonsense point of view, these axioms constitute the basic set formation processes. In our view, they are rather limited. In a commonsense set theory, we have to allow conditions like $Above(x, y)$ or $On(x)$ which are very common in AI applications. Therefore, φ should take some "external" conditions, but this newly introduced φ must be limited to work only with urelements.

Our discussion will later be continued with the Separation and the Replacement axioms of NBG.

Foundation (Regularity).

$$\forall x[\neg(x = \emptyset) \rightarrow \exists y(y \in x \wedge \forall z(z \in x \rightarrow \neg(z \in y)))]$$

This states that if a is a non-empty set, then there is an element b of a such that there are no sets which belong to both a and b. It simply restricts infinite descending sequences of sets under the membership relation (i.e., $... \in s_2 \in s_1 \in s$ is not allowed) [36].

This axiom answers a fundamental but so far unanswered question: does there exist a set which is a member of itself?[5] Since *Separation* requires a predefined domain, this axiom does not help us whereas Cantor's naive *Comprehension* gives the answer "Yes."

One of the important issues in the theory of sets is the well-founded sets. A set s is *well-founded* if there does not exist an infinite descending sequence $... \in a_{n+1} \in a_n \in a_{n-1} ... \in a_1 \in b$. This axiom guarantees that all of our sets will be well-founded.

One of the most elegant features of ZF is its strong support of the *cumulative hierarchy*. The idea of cumulative hierarchy comes from Russell's theory of types [23, 38]. We start with $\emptyset$ and iterate the power set operation at each stage (Figure 1). Details of this process can be found in [23, 24, 36]. The formal recursive definition of the class of well-founded sets (WF) in ZF is as follows:

- $R(0) = \emptyset$,

[5] This question was originally asked by Mirimanoff in 1917 [1].

- $R(\alpha + 1) = P(R(\alpha))$,
- $R(\alpha) = \bigcup_{\beta<\alpha} R(\beta)$ when α is a limit ordinal,
- $\mathrm{WF} = \bigcup\{R(\alpha) : \alpha \in Ord\}$.

where $R(\alpha)$ is the *rank* function defined for ordinals α and β.

The axiom is designed to reject $X = \{x : x \in x\}$, and thus guarantees the cumulative hierarchy. It can be omitted by the practical mathematician with no overhead. because the notions "set of all sets," "set that contains itself," etc. are not common in daily mathematics.

From a philosophical point of view, this axiom prohibits circularity which is, in some cases, a desirable property. Sometimes, we need to represent self-referring collections, and with this axiom we cannot call them sets, but philosophers do not seem too unhappy with this fact. They use the term *class* for this type of collections. Indeed, classes are like sets except that they can be highly comprehensive; an example of a class is the one that contains all sets.

In the commonsense side, we definitely need to represent circularity. An example for this might be the non-profit organizations [26]. Non-profit organizations are sets of individuals and the set of all non-profit organizations is also a set; all these are expressible in the cumulative hierarchy. Then, we might form a new non-profit organization, which collects (is an "umbrella" organization for) all non-profit organizations including itself. The new organization refers to the set of all non-profit organizations and is thus self-referring. Therefore, in a commonsense set theory, *Foundation* should be dropped. Instead, an *Anti-Foundation Axiom* (AFA) might be helpful. Aczel [1] chooses a graph representation of sets and his AFA states that every graph, whether well-founded (acyclic) or not, pictures a unique set. Removing *Foundation* from ZF and adding the AFA would result in Hyperset theory (also known as $\mathrm{ZFC}^-/\mathrm{AFA}$). What is advantageous with the new theory is that since graphs of arbitrary form are allowed—including the ones containing proper cycles—we can represent self-referring sets. This issue is further discussed in [1, 5, 7, 8, 23, 24, 25].

Choice.

$$\begin{aligned}\forall t[\forall x[x \in t \rightarrow \exists z(z \in x) \wedge \forall y(y \in t \wedge \neg(y = x) \rightarrow \neg\exists z(z \in x \wedge z \in y))]\\ \rightarrow \exists u \forall x(x \in t \rightarrow \exists w \forall v[v = w \leftrightarrow (v \in u \wedge v \in x)])]\end{aligned}$$

The meaning of the axiom is that, if t is a disjointed set which does not contain the $\emptyset$, its outer product Πt is different from $\emptyset$. In other words, among all subsets of $\cup t$ there is at least one whose intersection with each member of t is a singleton.

In 1904, Zermelo formulated the axiom formally and used it in his proof of the well-ordering theorem (i.e., every set can be well ordered). For finite sets, the well-ordering theorem can be proven without reference to *Choice* (AC), but in the infinite case, AC is essential [13]. In 1906, Russell restricted the domain set (t above) to be a disjointed set.

AC is not the part of original ZF, and when it is included in ZF, the new system is called ZFC. In 1922, Fraenkel proved the independence of AC from "ZF + Urelements." In 1938, Mostowski and Lindenbaum extended this proof to pure ZF.

From a mathematical point of view, in order to deal with large sets (e.g., non-denumerable ones), we need this axiom. Typical mathematical applications of AC can be found in [13, 16, 33, 35]; they include operations concerning cardinals (e.g., addition, multiplication, exponentiation) and real analysis, particularly theories of point sets and of real functions, etc.

From a commonsense point of view, AC seems quite complicated. On the other hand, we still have to deal with well-orderings of large collections, cardinalities, and typical mathematical operations that deal with large numbers. The commonsense version of AC must be compatible with our (newly defined) notion of infinity (and of course, new notions of cardinality and well orderings). A problematic case will be fuzzy membership. In this case, even the comparison of two sets will be difficult. Accordingly, AC needs reconsideration after the formation of basic principles of a commonsense set theory.

3. von Neumann-Bernays-Gödel (NBG)

In the 1925 axiomatization of von Neumann, there are three primitive notions: sets, classes and membership. His system can be seen as a conservative extension of ZF to allow comprehensive classes (e.g., the set of all sets, etc.). Later, in 1937 Bernays simplified the system [13]. The system mentioned in this section is his system. In 1940, Gödel used a version of this system in his famous consistency proof and the generalized continuum hypothesis.

The advantage of NBG over ZF is that NBG is finitely axiomatizable, i.e., it does not contain axiom schemata.

Before the explanation of each axiom of NBG differing from ZF, we have to clarify the notion of class. A *class* can be thought of as all the sets x which fulfill a condition $\varphi(x)$. (The condition $\varphi(x)$ is a wff of first order logic).

In formal statements of the system, we use capital letters to denote classes, and lower case letters to denote sets. In the definition of membership, we have two kinds of membership: membership of a set in a set (i.e., $x \in y$), and membership of a set in a class (i.e., $x \in Y$). In fact, we can use different symbols for this relations (η is used by Bernays for class membership). In this system, formulas of the type $X \in y$ and $X \in Y$ are not allowed. Similar to $\in$, the use of "=" is also limited. (It can only be used between sets or between classes.) Quantification over the classes is allowed in this system.

Here are the axioms of NB:

Extensionality for Classes.

$$\forall A \forall B \forall x[(x \in A \leftrightarrow x \in B) \rightarrow A = B]$$

This axiom defines the machinery for comparisons of classes. It states that two classes are identical if and only if their extensions are equivalent and is very similar to *Extensionality.* The same commonsense problems and discussions are valid for this axiom as well.

Predicative Comprehension for Classes.

$$\exists A \forall x(x \in A \leftrightarrow \varphi(x))$$

where $\varphi(x)$ is a condition which does not contain quantifiers over classes.

The intended meaning of the term *class* lies in the following definition: there exists a class which consists exactly of those elements x which fulfill the condition $\varphi(x)$ where $\varphi(x)$ in any pure condition[6] in x.

The reason of the small change in axiomatization is the weakness of the latter condition. With the latter definition, we cannot express some of the properties of our intended notion of class. An example for this is the complement of a class: for every class A, there exists a class B which consists exactly of all elements which are not members of A.

From the mathematical point of view, the above axiom allows us to define and prove the usual properties of the following classes:

- *Null Class.* $\wedge = \{x : x \neq x\}$
- *Universal Class.* $V = \{x : x = x\}$

[6] A pure condition is a condition which mentions only sets.

- *Union of Classes.* $A \cup B = \{x : x \in A \vee x \in B\}$
- *Intersection of Classes.* $A \cap B = \{x : x \in A \wedge x \in B\}$
- *Complement of a Class.* $-A = \{x : x \notin A\}$

After the introduction of the classes into the theory, some of the axioms of ZF require slight revision. *Extensionality*, *Null Set*, *Pair Set*, *Union*, *Infinity*, and *Power Set* do not need any change since they are single axioms dealing with set construction and identity [13].

The remaining axioms of *Separation (Replacement)* and *Foundation* need revision, in order to extend them to be compatible with classes. AC can be added to the system without any difficulty. Here, we have to note that AC will be used only over the domain of sets (i.e., we will accept the same AC of ZFC). If we add this AC to the system, we call the resulting system NBC.

Separation and *Replacement.*

The axiom of separation can be expressed as follows:

$$\forall P \forall a \exists y \forall x (x \in y \leftrightarrow x \in a \wedge x \in P)$$

In the previous ZF version of *Separation*, we had a condition (wff $\varphi(x)$) instead of P. Here, the separation of elements are in fact done during the formation of the class P. After the formation of the class of all elements (not necessarily members of a) that corresponds to φ, we simply intersect this class with the set a. This type of definition for *Separation*, results in a single axiom (shown above). Since the predicative comprehension for classes is finitely axiomatizable, the resulting system does not contain any axiom schema.

A similar attitude towards *Replacement* results in:

$$\forall F (Fun(F) \rightarrow \forall a \exists b \forall y (y \in b \leftrightarrow \exists x (x \in a \wedge x \in Dom(F) \wedge y = F(x))))$$

The predicates $Fun(F)$ and $Dom(F)$ intuitively represent "F is a function" and "domain of F," respectively. These predicates can be expressed (can be formulated as a wff) in the language of NBG. Then, the meaning of the axiom is as follows: if F is a function and a is a set, then there is a set which contains exactly the values $F(x)$ for all members x of a which are in the domain of F.

This axiom is finitely axiomatizable since F can be seen as a class, and Fun and Dom can be defined in the language of NBG.

From the commonsense point of view, these axioms shift the discussion to the axiom of comprehension for classes, since the abstraction mechanism is embedded in this axiom.

Foundation.

$$\forall P (\exists u (u \in P) \rightarrow \exists u (u \in P \wedge \forall x (x \in u \rightarrow x \notin P)))$$

In other words, every class P which has at least one member has a minimal member u, i.e., u is a member of P, but no member x of u is a member of P.

This axiom will not be useful in the commonsense set theory, because of the need for the representation of circularity.

In the beginning of the section, we have said that the "=" would not be defined between a set and a class. The axiomatization is carried out under this constraint. But, our intuitive notion of class requires each set to be equivalent to a class (i.e., the extension of the set and of the class are the same)[7], but the opposite is not true (i.e., for every class there is no equivalent set, e.g., $\{x : x \notin x\}$). Therefore, we can extend "=" to work between a class and

[7] For each set y, we have the class $\{x : x \in y\}$ whose extension is equivalent to extension of y.

set. "$z = A$ (and $A = z$), if and only if z and A have exactly the same members." After these definitions, we see that $\wedge = \emptyset$. Then, A is a *set* if, for some $z = A$. A will be a *proper class* if it is not a set.

From the commonsense point of view, the only novel thing introduced with this system is the notion of class. It has a quite simple idea: "every thing that is not a set is proper class." Although, the system incorporates the power of finite axiomatizability, its workspace is almost the same as the workspace of ZF. The limitations include the representation of urelements, external predicates, etc.

It should be noted that every theorem of ZF is still valid in NBG.

4. New Foundations (NF)

In 1937, Quine tried to streamline Zermelo's approach by giving up the unrestricted use of the axiom of comprehension with the restriction of stratification [12, 13, 28].

His idea of stratification goes back to the Theory of Types of Whitehead and Russell [38]. An assignment of levels to variables is said to *stratify* a formula of the form $\bullet\bullet\bullet \in ---$ if it assigns two consecutive levels to its left-hand and right-hand variables, and such an assignment is said to stratify a formula of the form $\bullet\bullet\bullet = ---$ if it assigns the same levels to its variables. A formula ψ is called stratified if there is an assignment which assigns levels to all the variables which occur in the expanded form of the formula ψ [13].

Quine's system is called New Foundations (for Mathematical Logic). The system contains the following axioms:

Extensionality.

$$\forall x \forall y [\forall z (z \in x \leftrightarrow z \in y) \rightarrow x = y]$$

Comprehension.

$$\exists y \forall x (x \in y \leftrightarrow \varphi(x))$$

where $\varphi(x)$ is a stratified formula in which y does not occur free.

This system is quite powerful in the handling of logical antinomies. Since the decisive formulae of logical antinomies cannot be stratified, antinomic sets cannot be proven directly by *Comprehension*. Here, there is no guarantee against establishing the existence of antinomic classes by an indirect proof. A simple example can be found in [13]. As a set theory, NF has serious difficulties. With the use of NF, one can prove some strange "properties", and cannot prove some important facts. As a relevant example for our discussions on cardinalities, in NF, it is impossible to prove for every finite cardinal number n that there are exactly n cardinal numbers (provided that NF is consistent) [13]. *Choice* is not compatible with NF (i.e., the formulation of AC cannot be stratified). However, the existence of infinite sets is provable in NF.

To continue with the undesirable properties, NF is a strong theory for its approaches to the mathematics, but there is no mental image of set theory which leads to this axiom. Thus, in order to form a playground for mathematics, we need to add a lot of ad hoc axioms. This is not true for ZF (and NBG). This lessens the simplicity of the system, thus its beauty.

Another issue in NF has to do wit relatively small sets. In NF, "relatively small" sets are expected to well-behave (i.e., they can be well-ordered, used in arithmetics, etc.), whereas it is not the case for the sets in general terms [13, 28]. Thus, in the borderline cases, we might need to limit the size of the relatively small sets. This results in the introduction of a new axiom.

In NF, the underlying formal language is not specified. This might result in semantical antinomies.

The consistency of the system cannot be proven by means of the proof mechanism of Gödel [13]. In the literature, we cannot find a proof of the sort "If ZF is consistent, then so is NF" either.

How about the individuals? Quine chosen to introduce individuals as sets x which are equal to their singletons, $\{x\}$. In this way, he did not need to change *Extensionality*. However, if he had chosen to introduce individuals in the usual way, then he would have need to change *Extensionality*, thus the existence of an infinite set would not be granted without an axiom. Thus, NF could be shown to be consistent by a proof similar to that of Gödel.

From the commonsense point of view, this system has various advantages over the other axiomatizations (i.e., ZF and NBG). By adding of simple axioms a workspace can be constructed easily. An example for this is the axiom that states that "relatively small" sets can be well-ordered [28]. Another advantage is that the universal set can be represented in NF, since $\varphi(x) =$ "$x = x$" is a stratified formula [12, 28].

The list of the disadvantages is quite long. First of all, the mathematical weakness of the system (mentioned in the preceding paragraphs) causes problems. Another disadvantage is that it is difficult to have a mental image of the system. It is quite to identify the sets and formulas with the "real world" objects. Since NF has a type theoretical approach, it carries the difficulties regarding circularity and self-reference. Handling of individuals seems quite irrelevant to human cognition, thus to common sense.

The consideration of classes in NF results in *Mathematical Logic* (ML) [12, 13, 28, 29]. The axioms of ML are as follows:

Extensionality.

$$\forall A \forall B[\forall x(x \in A \leftrightarrow x \in B) \rightarrow A = B]$$

Comprehension by a Set.

$$\exists y \forall x(x \in y \leftrightarrow \varphi(x))$$

where $\varphi(x)$ is a stratified formula with set variables only, in which y does not occur free.

Impredicative Comprehension by a Class.

$$\exists Y \forall x(x \in Y \leftrightarrow \phi(x))$$

where $\phi(x)$ is any formula in which Y does not occur free.

The relation between NF and ML is analogous to the relation between ZF and NBG. In both cases, invalid sets are considered as classes and operations are limited to the sets. This approach guarantees the mathematical completeness of the theory (i.e., any collection falls into the range of the theory either as a set or a class), but this ad hoc extension does not increase the power of the theory so much since sets are still the only useful objects in the universe.

5. Admissible Set Theory (KPU)

Admissible sets were first introduced by Kripke and Platek [3]. Their point of view was primarily that of recursion theory. They generalized ordinary recursion theory of integers to ordinals smaller than a fixed, well-behaved *admissible ordinal* [22]. The switch from admissible ordinals to *admissible sets* renders an elegant first order theory. From a technical point of view we switch from admissible ordinal α to set L_α of sets constructible before α.

Barwise [3] weakened KP to KPU by re-admitting urelements. Although the reason for this decision was purely mathematical, the resultant system became stronger from the commonsense point of view.

In the language of set theory with urelements, we have a predicate U which stands for urelements. The set of urelements are denoted by M. The membership relation $\in$ changes to $\in_M$ to work in the range of urelements. However, since the intended meaning did not change, we will use $\in$ both for the original $\in$ and for $\in_M$.

Before going into a discussion of the axioms of KPU, we define the class of Δ_0 formulas. The collection of Δ_0-*formulas* of the language $L(\in, ...)$ is the smallest Y containing the atomic formulas of $L(\in, ...)$ closed under:

(i) if φ is in Y, then so is $\neg\varphi$;

(ii) if φ, ψ are in Y, then so are $(\varphi \wedge \psi)$ and $(\varphi \vee \psi)$;

(iii) if φ is in Y, then so are $\forall u \in v\varphi$ and $\exists u \in v\varphi$, for all variables u and v.

The importance of Δ_0-formulas is that any predicate defined by them is absolute[8], and many predicates occurring in usual mathematics can be defined by Δ_0-formulas (e.g., subset, pairing, ordered pair, etc.). (This fact reduces the number of axioms.) The axioms of the system are as follows:
Extensionality.

$$\forall x(x \in a \leftrightarrow x \in b) \rightarrow a = b$$

In Makkai [22], the extensionality axiom specifically ignores the urelements:

$$(\neg U(a) \wedge \neg U(b)) \rightarrow (\forall x(x \in a \leftrightarrow x \in b) \rightarrow a = b)$$

Foundation.

$$\exists x\varphi(x) \rightarrow \exists x[\varphi(x) \wedge \forall y \in x \neg\varphi(y)]$$

for all formulas $\varphi(x)$ in which y does not occur free.

This axiom, just like *Extensionality*, is concerned with the basic nature of the sets. The function of the axiom in this system does not change, i.e., its primary use is to restrict of the circular sets.
Pair.

$$\exists a(x \in a \wedge y \in a)$$

Union.

$$\exists b \forall y \in a \forall x \in y(x \in b)$$

Δ_0-*Separation.*

$$\exists b \forall a(x \in b \leftrightarrow x \in a \wedge \varphi(x))$$

for all Δ_0 formulas in which b does not occur free.

Pair and *Union* are the same as their counterparts in ZF. The last three axioms are concerned with set formation. Δ_0-*Separation* is also similar to the one in ZF, but here the range φ is limited to Δ_0-formulas.

After these five axioms, we can define the universe of admissible sets over the set of urelements M:

- $V_M(0) = \emptyset$,
- $V_M(\alpha + 1) = P(M \cup V_M(\alpha))$,
- $V_M(\lambda) = \bigcup_{\alpha<\lambda} V_M(\alpha)$, if λ is a limit ordinal.
- $V_M = \bigcup_\alpha V_M(\alpha)$.

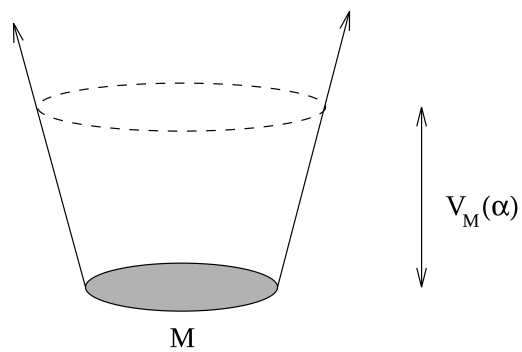

Figure 2: The universe of admissible sets (adapted from [3])

Here α and λ are ordinals.

This universe can be depicted as in Figure 2. It is very similar to that of ZF (except that the urelements) and also supports the cumulative hierarchy.

Δ_0-*Collection.*

$$\forall x \in a\ \exists y\ \varphi(x, y) \rightarrow \exists b\ \forall x \in a\ \exists y \in b\ \varphi(x, y)$$

for all Δ_0 formulas in which b does not occur free.

From the mathematical point of view, Δ_0 collection axiom assures that there are enough stages in the (hierarchical) construction process.

The elegance of the system is due to the fact that Δ_0-formulas are very powerful, and with Δ_0 collection this power is used in the system. The result is *the principle of parsimony* (i.e., simple facts should have simple proofs). One example of this principle is the proof of the existence of the Cartesian product. In ZF, we had to refer to $P(P(a \cup b))$ for this proof, but here the predicate "$< x, y >$ where $x \in a$ and $y \in b$" is a Δ_0-formula and with two applications of collection (one for x and one for y) the proof is completed [3].

An *admissible set* is a transitive set A (in some V_M) that is a model of KPU; more technically if $\mathcal{M}$ is a structure for L, then an *admissible set over* $\mathcal{M}$ is a model $\mathcal{A}_{\mathcal{M}} = \{\mathcal{M}; A, \in, ...\}$, where A is the nonempty set of non-urelements and $\in$ is the restriction of $\in_M$ to $M \cup A$. A *pure admissible set* is an admissible set with no urelements ($M = \emptyset$).

One of the most notable things in the theory from the commonsense point of view is Σ-*Reflection.* We will only state the theorem; a proof can be found in [3].

The Σ *Reflection Principle.*

For all Σ-formulas φ, the following is a theorem of KPU:

$$\varphi \leftrightarrow \exists a (\varphi)^a$$

In particular, every Σ-formula is equivalent to a Σ_1-formula in KPU[9]. Since $(\varphi)^a$ is obtained from φ by relativization of all quantifiers to a, the theorem says that a Σ_1-formula is true if and only if it is true in some set. This type of fact is called reflection principle because it describes a situation in which the truth about the whole universe is locally desirable [39].

Another point to note is that the axioms of infinity, power set, null set, etc. are missing. The system can easily be expanded (without any consistency and independence problems) with such axioms.

[8]Absolute formulas do not shift their meaning on us as we move from a model $\mathcal{A}_{\mathcal{M}}$ to its extension $\mathcal{B}_{\mathcal{N}}$.

[9]A formula of the form $\exists u \varphi(u)$, where φ is Δ_0, is called a Σ_1-formula. The class Σ is the smallest class containing Δ_0 and closed under conjunction, disjunction, bounded quantification, and existential quantification.

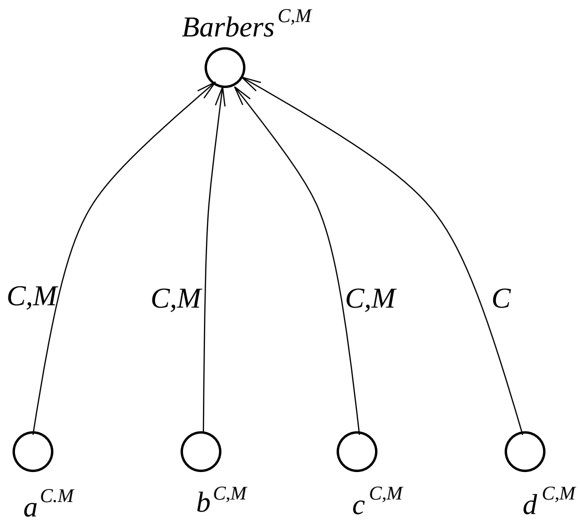

Figure 3: Barbers of Springfield

In our opinion, although, the admissible sets remove some of the drawbacks of ZF, they cannot be a candidate for a commonsense set theory alone. The addition of some new axioms (the axioms in previous paragraph and more) and review of existing axioms (beginning from the extensionality) are necessary.

6. Steps into a Commonsense Set Theory

The commonsense set theory that we propose will be quite similar to the one in [39]. Since that theory deals with cardinalities and well-orderings in an acceptable way, it eliminates further discussion on these subjects.

We have chosen the graph representation of the sets in the explanation of our proposal. In the representation, edges will represent the membership relation. We represent fuzzy membership with labels on the edges. The situation in Springfield can be represented by the graph in Figure 3.

The labels on the edges tell that in which situation the edge (i.e., the membership) is valid. The meaning of the graph is the following:

From the commonsense point of view:

$$Barbers^C = \{a, b, c, d\}$$

From the mayor's point of view:

$$Barbers^M = \{a, b, c\}$$

In Figure 3, we allowed individuals (urelements) as leaf nodes.

When we choose the above labeling schema, what are the basic building blocks of our set theory? These building blocks must include:

1. A set of previously defined situations (contexts) to be used as labels on the edges and the nodes. In the example above, this set is $\{C, M\}$, where C represents the situation from the people's point of view, and M represent the situation from the mayor's point of view.

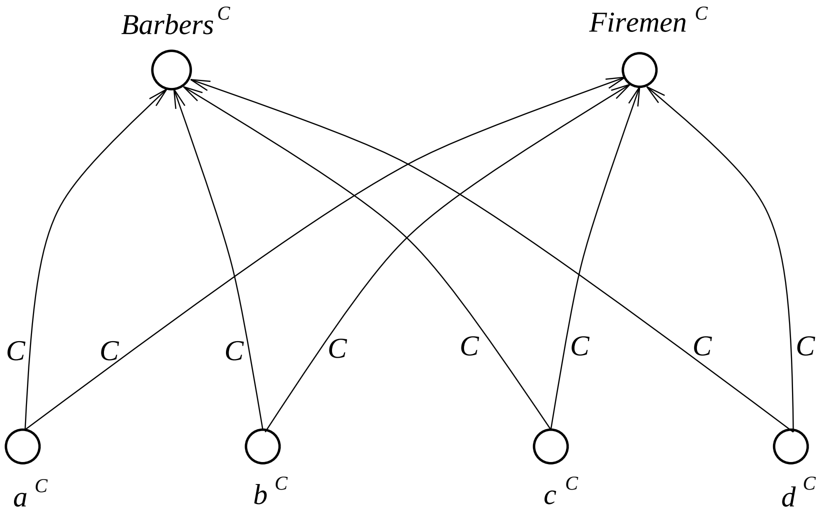

Figure 4: Set representation of Barbers and Firemen of Springfield

2. A set of individuals. ($\{a, b, c, d\}$ is this set.)
3. The edge relation over the nodes. In this relation, the labeling schema must be considered.
4. Set of nodes.
5. Set of edges.

We can define the equivalence of two sets relative to the situation (e.g., $Barbers^C$ and $Barbers^M$ are not compatible). However, since we have used the same node, they have to be the same set (although their extensions are different in each situation).

In the equivalence issue on the commonsense set theory consider the representation of Figure 4. This example represents two different sets (i.e., the barbers of Springfield and the firemen of Springfield). That these two sets are not equal is due to the fact that they are represented by different nodes. But, from a mathematical point of view, they are equivalent since their extensions are equivalent. If we adopt the extensionality axiom for a commonsense set theory, we cannot say that these sets are different. Therefore, the commonsense notion of equivalence must be different than the mathematical notion.

In treating some "useful" sets, we have chosen the the following representations.

Null Set.

The null set can be viewed as a special individual. From the commonsense side this representation does not cause major problems, since the adoption of commonsense definition equivalence. There might be more than one node to represent a set with no member.

Natural Numbers.

Zadrozny [39] introduced a predefined set to find the cardinality of sets. The adoption of this set in the theory seems unproblematic. But, the problem occurs on the mathematical side. We are explicitly defining the cardinals. The cardinal number notion of ZF seems quite different to form an isomorphism between our notions. The natural numbers are represented with the graph of Figure 5. Intuitively, this set corresponds to a common sense counting mechanism. However, when people talk about large quantities, they make a generalization with the word "many." In Figure 5, this counting mechanism is tried to be shown. The

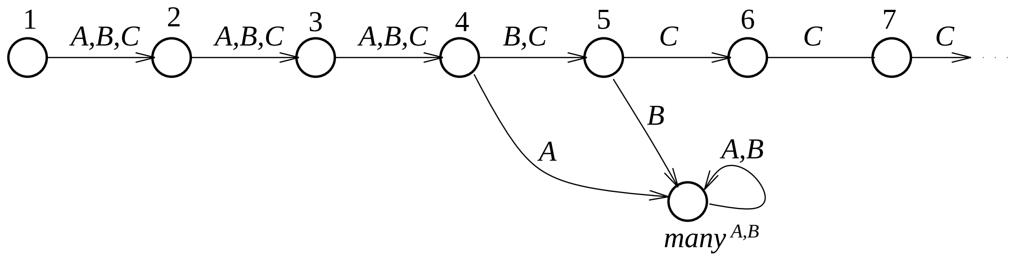

Figure 5: Natural Numbers

labels on the edges and on the nodes correspond to the person who makes this counting. Depending on who this person is (A, B, or C) the notion of "many" changes drastically. For example, A thinks that any quantity greater than 4 can be called "many" whereas for C even 7 is not good enough to qualify as "many." Note also that once A and B reach "many" further operations with this notion results in itself, e.g., many apples plus many oranges gives many fruits.

Infinity.

Infinity is quite an unclear concept in our daily life. In fact, we usually say "a lot" when we want to use the infinity. Infinity, is usually used in a cardinality sense. We usually do not refer to infinities themselves (i.e., we need not compare the cardinalities $\aleph_0$ and $\aleph_1$). In our opinion, the place of the infinity is the end of the chain of natural numbers. It must be represented with self-reference to itself, but indeed this reference will not be used since this node is inaccessible.

In order to have a commonsense set theory, we may begin with the above representations. Extending (and correcting) the above ideas with other commonsense issues, might result in a useful commonsense set theory. But, in our opinion, the theory will grow in two different branches: purely mathematical and pure common sense. If we increase the number of intersection (points of contact) points of these two branches, we will probably get a better commonsense set theory.

7. Conclusion

We tried to point out to the commonsense issues in some classical axiom systems, together with their related mathematical and philosophical background. This review shows us the ad hoc nature of the ZF, i.e., the system is good for the "practicing" mathematician but sometimes it is just too strong from a commonsense point of view. The advantage of the NBG system is that it is finitely axiomatizable. The addition of the notion of class in NBG is sometimes an advantage and sometimes a liability. NF is a rather logical approach to the subject and looks like the part II of Russell's "the theory of types." On the plus sides of the system, we should mention the existence of a "universal set." Finally, we reviewed KPU. Intuitively, KPU seems better than ZF, but since it is primarily recursion theory oriented, its target area is analysis and descriptive set theory.

After above reviews, we offered a skeleton of a commonsense set theory. This part of our work has not been completed yet, but we think that our approach may deliver a "feasible" commonsense set theory. In the current state of the theory, we believe that we can deal with the fuzzy sets of the daily life and address (in some sense) the cardinality and well-ordering issues. We are still studying the formalization of our approach and think that the review of previous axiomatizations will give us useful tools and ideas for this task.

Acknowledgments

The second author is grateful to Wlodek Zadrozny (IBM T. J. Watson Research Center) and Patrick Suppes (Stanford University) for general advice and help with the references. He also acknowledges the partial support of the Scientific and Technical Research Council of Turkey (TÜBİTAK) under grant number TBAG–992.

References

1. P. Aczel. *Non-well-founded Sets*. Number 14 in CSLI Lecture Notes. Center for the Study of Language and Information, Stanford, California, 1988.
2. V. Akman. "Undaunted Sets," *SIGACT News*, 23: 47–48, 1992.
3. J. Barwise. *Admissible Sets and Structures*. Springer-Verlag, New York, 1975.
4. J. Barwise and J. Perry. *Situations and Attitudes*. MIT Press, Cambridge, Massachusetts, 1983.
5. J. Barwise and J. Etchemendy. *The Liar: An Essay on Truth and Circularity*. Oxford University Press, New York, 1987.
6. J. Barwise. "Situated Set Theory," In *The Situation in Logic*, Number 17 in CSLI Lecture Notes. Center for the Study of Language and Information, Stanford, California, 1989.
7. J. Barwise and L. Moss. "Hypersets," *Mathematical Intelligencer* 13: 31–41, 1991.
8. J. Barwise. *Applying AFA: Notes for a Tutorial on Non-well-founded Sets*. Available via anonymous ftp from `phil.indiana.edu:/pub/Barwise-afa.ps`.
9. D. Cantone, A. Ferro, and E. Omodeo. *Computable Set Theory*. Oxford University Press, New York, 1989.
10. D. van Dalen, H. C. Doets, and H. De Swart. *Sets: Naive, Axiomatic, and Applied*. Pergamon Press, New York, 1978.
11. K. J. Devlin. *Sets, Functions, and Logic: An Introduction to Abstract Mathematics*. Chapman and Hall, London, 1992.
12. T. E. Forster. *Set Theory with a Universal Set*. Oxford University Press, New York, 1992.
13. A. A. Fraenkel, Y. Bar-Hillel, and A. Levy. *Foundations of Set Theory*. North-Holland, Amsterdam, 1973.
14. P. C. Gilmore. "Logic, Sets, and Mathematics," *Mathematical Intelligencer* 15: 10–19, 1993.
15. P. R. Halmos. *Naive Set Theory*. Springer-Verlag, New York, 1974.
16. T. Jech. *Set Theory*. Academic Press, San Diego, California, 1978.
17. P. T. Johnstone. *Notes on Logic and Set Theory*. Cambridge University Press, Cambridge, UK, 1987.
18. E. Kamke. *Theory of Sets*. Dover, New York, 1950.
19. K. Kunen. *Set Theory: An Introduction to Independence Proofs*. North-Holland, Amsterdam, 1980.
20. F. Landman. *Structures for Semantics*. Kluwer, Dordrecht, The Netherlands, 1991.
21. S. Mac Lane. *Mathematics: Form and Function*. Springer-Verlag, New York, 1986.
22. M. Makkai. "Admissible Sets and Infinitary Logic," In *Handbook of Mathematical Logic*, J. Barwise (ed.), North-Holland, Amsterdam, 1977.
23. M. Pakkan and V. Akman. *Issues in Commonsense Set Theory*. Technical Report, Department of Computer Engineering and Information Science, Bilkent University, Ankara, 1992.

24. M. Pakkan. *Solving Equations in the Universe of Hypersets.* M.S. Thesis, Department of Computer Engineering and Information Science, Bilkent University, Ankara, 1993.

25. M. Pakkan and V. Akman. "HYPERSOLVER: A Graphical Tool for Commonsense Set Theory," In *Proceedings of the Eighth International Symposium on Computer and Information Systems (ISCIS VIII)*, Istanbul, Turkey, 1993.

26. D. Perlis. "Commonsense Set Theory," In *Meta-Level Architectures and Reflection*, P. Maes and D. Nardi (eds.), Elsevier, Amsterdam, 1988.

27. J. Piaget. *Biology and Knowledge.* University of Chicago Press, Chicago, 1971.

28. W. V. Quine. *Set Theory and Its Logic.* Harvard University Press, Cambridge, Massachusetts, 1969.

29. W. V. Quine. *Mathematical Logic.* Harvard University Press, Cambridge, Massachusetts, 1981.

30. J. Roitman. *An Introduction to Modern Set Theory.* Wiley, New York, 1989.

31. J. R. Shoenfield. "Axioms of Set Theory," In *Handbook of Mathematical Logic*, J. Barwise (ed.), North-Holland, Amsterdam, 1977.

32. R. R. Stoll. *Set Theory and Logic.* Dover, New York, 1979.

33. P. Suppes. *Axiomatic Set Theory.* Dover, New York, 1972.

34. P. Suppes and J. Sheehan. "CAI Course in Axiomatic Set Theory," In *University-Level Computer-Assisted Instruction at Stanford: 1968–1980*, P. Suppes (ed.), Institute for Mathematical Studies in the Social Sciences, Stanford University, Stanford, California, 1981.

35. G. Takeuti and W. M. Zaring. *Introduction to Axiomatic Set Theory.* Springer-Verlag, New York, 1982.

36. M. Tiles. *The Philosophy of Set Theory.* Basil Blackwell, Oxford, UK, 1989.

37. R. L. Vaught. *Set Theory.* Birkhäuser, Boston, 1985.

38. A. N. Whitehead and B. Russell. *Principia Mathematica.* Volume 1. Cambridge University Press, Cambridge, UK, 1927.

39. W. Zadrozny. "Cardinalities and Well-Orderings in a Common-Sense Set Theory," In *Proceedings of the First International Conference on Principles of Knowledge Representation and Reasoning*, R. J. Brachman, H. J. Levesque, and R. Reiter (eds.), Morgan Kaufmann, San Mateo, California, 1989.

40. W. Zadrozny and M. Kim. *Computational Mereology: A Prolegomenon Illuminated by a Study of Part-of Relations for Multi-media Indexing.* Preliminary Report, IBM T. J. Watson Research Center, Yorktown Heights, New York, 1992.