Collections and Paradox

Remko Scha

Institute for Logic, Language and Computation
scha@uva.nl

1 Preface

This essay is dedicated to Jeroen Groenendijk, Martin Stokhof, and Frank Veltman — but not to each of them separately, but to all of them jointly, as a group, or a set, or a mereological sum. This group was an important part of a community of logicians and philosophers at the University of Amsterdam, that had a stimulating influence on my first ventures into formal semantics in the course of the 1970’s. The discussion below ties in directly with that work.

My topic will be the conceptual viability of the very idea of “collective entities” such as “Frank, Martin and Jeroen” and “the Amsterdam Montagovians.” Most theories of formal semantics assume that such plural noun phrases denote higher-level entities of some sort, that have individuals (such as Frank, Jeroen and Martin) as their members or parts. But this assumption has been challenged. A sweeping argument against ”collections” of any kind was raised in the eighties and nineties by George Boolos, James Higginbotham and Barry Schein [4, 5, 8, 27]: collections engender paradox. The chapters on plurals in some fairly recent handbooks in the Philosophy of Language [28, 15] still discuss this argument at great length, treating it as one of the most important issues in the field. Nonetheless, most researchers in this area seemed not at all disturbed by this challenge, and did not feel the need to deal with it. Fred Landman [11] for instance, in the context of an otherwise detailed discussion of Schein’s proposals about plurals, shrugs it off with a deadpan witticism:

“Schein 1993 invokes an argument for his approach involving Russell’s paradox. The language of Events and Plurality that I will develop is pretty much a standard type theory, so obviously it doesn’t have the means to even formulate Russell’s paradox. And here I go by the semanticists’ First Amendment: The right to solve Russell’s Paradox some other time shall not be restricted.”

Landman has a point. In an empirical enterprise such as formal semantics, it is wise to not be distracted by philosophical worries that feel intuitively irrelevant. But in this case the result has been that the Boolos-Higginbotham-Schein argument, which to some readers does seem convincing, was hardly ever discussed in a critical way. So that it is what I wish to do now.

2 Collective entities.

In logical representations of the meanings of natural language utterances, verbs may often be rendered as n-place predicates on individuals, and noun phrases may be analyzed as quantifiers which compose with these verbs to yield propositions. Simple examples are: “Every boy borrowed a book.” and “John, Paul and Peter walk.”. Natural language locutions may, however, also describe situations whose participants seem to be “collective entities.” Examples are: “Twelve boys gathered in the school yard.” and “John, Paul and Peter carried the piano upstairs.”. In these examples, the subject noun phrases cannot be analyzed as quantifying over
individual boys without yielding incorrect entailments. If we analyze the VP as a predicate, it must be one that does not (or not only) operate on individuals, but rather on collections of individuals.

We may wonder what kind of things these collections are, and how they relate to the individual entities that they consist of. The most obvious idea may be, to treat collections as sets. This is also the oldest idea. Long before modern set theory (let alone logical semantics), Bolzano [3, pp. 2–4] discussed sentences like “Die Sonne, die Erde und der Mond stehen in gegenseitiger Einwirkung aufeinander.” and “Die Rose und der Begriff einer Rose sind ein paar sehr verschiedene Dinge.”, to point out that they make an assertion about a totality (“Inbegriff”) of certain things, i.e., a “whole” consisting of certain parts; he also notes that the identity of such a totality may be conceived as being independent of the arrangement of its parts, and in that case, he calls the totality a set (“Menge”).

The earliest proposals about plurals in the Montague-tradition (by Michael Bennett, Renate Bartsch, and Roland Hauser) have indeed assumed that collective predicates apply to sets [2, 1, 7]. Some variations on this approach were developed in the early eighties. Henk Verkuyl [30] introduced the idea of representing individual entities by singleton sets, so that the distinction between distributive and collective verbs gets blurred and the type system stays simple: formally, all predications involve sets. Roger Schwarzschild [29] made the step to use a set theory in which the singleton set of an individual is formally identical to this individual, as proposed by Quine in Set Theory and its Logic: if \( x \) denotes an individual, \( x = \{ x \} = \{ \{ x \} \} = \ldots \).

All collective entities occurring in the examples above, could be represented by sets that are particularly simple, in that all their members are individuals: “the boys” = “John, Paul and Charles” = \{ J, P, C \}. Thus, in such analyses, the specific possibilities of set theory are in fact not exploited: what is used is a subset-algebra on the domain of individuals. As pointed out by Godehard Link [12], what matters is that the collective entities are elements in a complete atomic join semi-lattice. A subset-algebra implements this, but Leśniewski’s mereology or Leonard & Goodman’s Calculus of Individuals could serve just as well. And Link in fact proposed a mereological framework. Part of his motivation was to be able to account for count-nouns and mass-nouns in one encompassing system: both kinds of nouns are analyzed in mereological terms, and differ merely in that their extensions do or don’t have discrete atoms as their ultimate parts. Fred Landman [10] discusses the connection between set-theoretic semi-lattices and the count-noun component of Link’s mereology, and emphasizes their compatibility. Link [14, Ch. 3, 13, 14] defends the philosophical importance of the distinction, but admits in passing (p. 64): “For practical reasons (for instance, because people are "used to it") we could stick to the power set model as long as we don’t forget it is only a model.”

More complex examples, however, may be brought up to motivate a departure from a purely lattice-theoretical approach. “The sun, the planets and the satellites of the planets hardly influence each other.” has a prominent counterfactual reading where the absence of mutual influence is not asserted about all the heavenly bodies of the solar system, but about three specific parts, two of which are complex and one of which is not. “The rose and the properties of a rose are two very different things” cannot be meaningfully asserted about a mereological or set-theoretical entity which contains on an equal footing all the properties of the rose along with the rose

---

1 The example sentences translate into English as: “The sun, the earth and the moon mutually influence each other." and "The rose and the concept of a rose are two very different things.”.

2 The sketch of these developments in this section is based on joint work with Yoad Winter. But he has not read it and I know he disagrees on many points.

3 I also (reluctantly) adopted this idea in my 1981 paper [24], after Theo Janssen, commenting on a draft version, suggested that its readability would be improved that way. The original version employed a logic with a baroque type system that allows quantification over the extensions of arbitrarily complex “union-types”[23, 25].
itself. "The integers and the rationals are equinumerous." is rendered nonsensical by a lattice-theoretical treatment, because it would be equivalent to "The rationals are equinumerous." Because of such examples, it was argued by many that collective entities have more internal structure [9, 13, 26, 10]. If they are modelled as sets, these sets may have sets as their elements — and so on, recursively. Schwarzschild [29], however, insists that conjunctive NP's always denote sets of atomic individuals. The obvious readings of examples like the above are then to be generated by a pragmatic process which takes into account, among other things, the formulation of the NP. This is an ingenious methodological innovation: to shift syntactic information to the realm of pragmatics, in order to cut down on semantic ambiguity.\(^4\) This approach was endorsed by Godehard Link [14, Ch. 7] and by Yoad Winter [31, pp. 35–39].

Though the proponents of different views on these matters sometimes engage in fierce debates, this quick overview shows that the collective entities employed in different treatments of plurals have much in common. The fundamental problems that have been raised about such collections thus apply to all these approaches to an equal extent. I will now set the scene for the appearance of the Boolos-Higginbotham-Schein challenge.

3 The threat of paradox: “The sets.”

A philosopher with a taste for paradox may wonder what happens if a set-based formal semantic theory of the kind discussed above is applied to a non-trivial fragment of English that includes utterances about sets. Above we saw that most semantic theories analyze nouns as predicates whose positive extensions are sets of individuals. We also saw that many treatments of plurality, building on this, analyze a plural NP such as “the N’s” as denoting the set of individuals that satisfy the predicate denoted by N. In this approach, the noun “set” would be analyzed, for instance, as the predicate \(Set'\), applicable to a universe of discourse that, in the intended interpretation, includes all sets; and the NP “the sets” would be rendered as \(\{x \mid Set'(x)\}\), thus denoting, in the intended interpretation, the set of all sets. (Other relevant lexical items, such as “member”, “subset”, “union”, “intersection”, would be analyzed as functions or relations that involve the elements of this set.)

Now we do seem to have a problem, because the most widely accepted set theories do not acknowledge such a thing as the set of all sets. Most formal set-theories do not aim to be all-encompassing. To avoid paradox, they accept that there is no Universal Set; the totality of all sets is itself not a set. Of course this invites the question: what set-theory should we adopt? Should we perhaps embrace a deviant set-theory which \(\textit{does}\) allow the Universal Set (such as Quine’s \textit{New Foundations}, or, preferably, Jensen’s 1968 version, which adds Urelements)? But the interesting thing about this question is that it cannot be answered as it stands; it is an ill-formed way to ask two different questions. When we say that \(\{x \mid Set'(x)\}\) denotes the set of all sets, and we reflect on what we mean by “set” in the phrase “the set of all sets”, we should realize that it involves the set-theory of the meta-language (that we may choose as we wish for its theoretical properties), as well as the set-theory of the object-language (i.e., the set-theory that is obeyed to be assumed in the discourse that we are analyzing). It is misguided, therefore, to say that the NP “the sets” denotes the set of all sets. It denotes the \(\text{set}_{\text{meta}}\) of all \(\text{set}_{\text{object}}\)’s.

\(^4\)Schwarzschild drew perhaps too many consequences from an observation that is in and of itself worthwhile. “The boys and the girls were fighting” is ambiguous; it may be synonymous with “The boys were fighting with the girls” (it may even be syntactically derived from it), but it can also mean that the children were fighting, without asserting anything further about the nature of the fight. But now imagine that the utterance was preceded by: “The sons and the daughters of Mr. Johnson hate the sons and the daughters of Mrs. Peterson.” Then, the contextual (pragmatically induced) meaning of “The boys and the girls were fighting” may in fact be, that the Johnson’s were fighting with the Peterson’s.”
The notions of \textit{set}_{\text{meta}} and \textit{set}_{\text{object}} are intrinsically distinct. (In fact, the English word “set” is manifold ambiguous, and a semantic analysis should introduce different postulates about its meaning, depending on the distinct authorities or theories that may be implicitly or explicitly invoked when it is used. But \textit{none} of these meanings can involve the mathematical notions assumed by the meta-language.)

Formal semantic theory traditionally assumes a strict separation between the closed-class function-words and morpho-syntactic structures of language on the one hand, and the open-class content words (such as nouns, verbs, and adjectives) on the other. In logical analyses of utterances, the meanings of open-class content words are represented as descriptive terms of the logic — i.e., terms that may get different extensions in different models, reflecting the fact that their meanings may determine different referents, depending on the state of affairs that obtains. These descriptive terms are thus necessarily disjoint with the formal terms of the logic, and with the mathematical terms that are employed to articulate its model theoretic interpretation — because the interpretations of the latter terms are fixed once and for all, not allowed to vary when different models of the logic are considered.

Thus, in particular, the set-theoretical basis of the model-theory and of the representation of plurality (i.e., the properties of \textit{set}_{\text{meta}}) should be chosen once and for all, independently of the analyses of specific lexical items. But if a semantic theorist were so brave as to attempt the analysis of natural language texts about sets, he should be able to introduce and adapt meaning postulates about \textit{set}_{\text{object}} depending on the object-discourse. There is no reason to treat the terminology about sets and other collections as different from other open-class content words. Note, therefore, that the noun “set” is to be analyzed as a descriptive predicate on individuals. In the universe of \textit{set}_{\text{meta}}, these are not sets but Urelements. If we analyze a discourse which involves for \textit{set}_{\text{object}} precisely the meaning postulates that characterize the properties of \textit{set}_{\text{meta}}, an interesting coincidence arises: the \textit{set}_{\text{meta}}-theory would be a copy of the \textit{set}_{\text{object}}-theory. (Cf. Charles Parsons [18, pp. 10-11]: “... there seems to be no intrinsic objection to the postulation of stronger and stronger properties of classes, so that they are conceived as indistinguishable from another layer of sets.”) Note that also in this case, no Universal Set emerges. The totality of all \textit{set}_{\text{object}}’s does not constitute a \textit{set}_{\text{object}}, but a \textit{set}_{\text{meta}} (that contains the \textit{set}_{\text{object}}’s not as sets, but as Urelements).

4 “The sets that are not members of themselves.”

Some readers may think that this is all true, but boring. That the sophist who would try to dismiss a set-theoretical model of plurality is an implausible straw-man. But, as it turns out, there are some brilliant researchers who made important contributions to philosophical logic and formal semantics, who seem to have fallen victim to precisely the fallacy that I just discussed at such length.

George Boolos, James Higginbotham and Barry Schein [4, 5, 8, 27] criticize the set-based analysis of plurals by problematizing its application to the word “set” itself. Their examples are somewhat more complex than the ones we discussed above. Consider, for instance, the phrase “the sets that are not members of themselves”; a set-based treatment of plurals will analyse this as \{x \mid \text{Set}(x) \land \neg \text{MemberOf}(x, x)\}. The most straightforward English paraphrase of this expression would be “the set of all sets that are not members of themselves”. We thus seem to obtain an equivalence between “the sets that are not members of themselves” and “the set of all sets that are not members of themselves”. Accordingly, “The null set is among the sets that are not members of themselves” should have the same truthvalue as “The null set is a member of the set of all sets that are not members of themselves.” Now under any reasonable construal
of the notion of a set, the first of these statements is true — since it is equivalent to “The null set is a set that is not a member of itself,” and thus to “The null set is not a member of itself.” The second statement, however, is bound to come out false or truthvalueless, since no set theory can allow the existence of the set of all sets that are not members of themselves, if it is to avoid the Russell paradox. (If the set of non-self-membered sets exists, “The set of non-self-membered sets is a member of itself” is both true and false.) The set-based treatment of plurals thus seems to entail a contradiction.

The flaw in this reasoning is the same failure to distinguish object-language and meta-language that we discussed in the previous section. If “the sets that are not members of themselves” is analyzed as \( \{ x \mid \text{Set}(x) \land \neg \text{MemberOf}(x, x) \} \), this formula should not be paraphrased as “the set of all sets that are not members of themselves,” but as “the set \( \text{meta of all set object’s that are not member object’s of themselves} \).” The problem thus vanishes completely — not by an ad hoc restriction, but by insisting on an a priori methodological precondition for sound and systematic semantic analysis.

Boolos [4, 5], failing to distinguish object-language and meta-language but wishing to prevent inconsistencies like the one just mentioned, concludes that plural NP’s cannot be analyzed in terms of sets. An analysis in terms of classes, collections or other totalities is excluded by the same token, because such terms allow analogous constructions with analogous inconsistencies. Boolos recommends that locutions involving plural NP’s are analyzed in terms of a second-order logic which allows quantification over predicates, while eschewing terms that denote collections of any kind. In [4] he demonstrates his perspective on second-order logic by translating second-order formulas into a disambiguated fragment of English; this procedure is not completely convincing, however, because this fragment of English crucially employs definite plural pronouns. In [5], a model theory for second-order logic is specified which is more complex than the usual Tarskian formulation, and avoids explicit reference to sets. Michael Resnik [22], however, still discerns a disguised reference to sets in this semantic system.

It should be noted that Boolos was not primarily concerned with natural language, but rather with the relation between logic and set theory. In particular, he was investigating the formulation of the Zermelo-Fraenkel axioms (and their consequences) in second-order logic. In this context, there was no need for him to discuss the intrinsically collective verbs (such as “meet” or “outnumber”) that motivate much of the discussion about plurals in natural language semantics. This is different for Higginbotham and Schein [8, 27]. They acknowledge Boolos’ observations as well as his conclusions; they thus wish to avoid collections in the semantics of plurals, and introduce technical innovations for that purpose.

5 Predicates vs. collections.

In [8], Higginbotham and Schein introduce two new ideas in the literature on English plurals. The first is to analyze plural NP’s not in terms of collections but in terms of predicates; the second is to analyze verbs not in terms predicates but in terms of Davidsonian events. In their proposal, these ideas are combined; but they can be decoupled, and that is what we do for the present discussion. First we discuss the issue of NP semantics.

If, as usual, nouns are analyzed as one-place predicates, collections may be avoided by analyzing plural NP’s in precisely the same way. For instance, a definite plural NP of the form “the Ns” is rendered by the same predicate \( N' \) that translates the noun; an indefinite plural NP of the form “some Ns” quantifies over predicates that entail \( N' \). For instance, assume that Apostle is the predicate that yields True for every Apostle and for nothing else, and that for every numeral there is a second-order predicate that yields True if the positive extension of its
argument predicate has the corresponding cardinality.

“The Apostles are twelve.”

may then be rendered as

Twelve′(Apostle′).

More interestingly, the analysis of

“Some Apostles lift the piano”

would be

∃X[(∀x X(x) ⇒ Apostle′(x)) ∧ LiftThePiano′(X)].

If we employ the subset-sign to indicate the inclusion between the positive extensions of two predicates, this formula can be rendered as:

∃X ⊆ Apostle′ : LiftThePiano′(X).

The move from sets to predicates for the semantics of plural NP’s thus can be carried out if all predicates that intuitively apply to collections are systematically “lifted” to versions that formally apply to predicates. It is not clear what advantage this has.

It may be noted that the replacement of collections by predicates makes no substantial difference for the prevention of paradox, though that is what apparently motivated Higginbotham and Schein. In their system, “the predicates that don’t apply to themselves,” which describes a well-defined category of predicates, would be represented by “the predicate that applies to precisely all predicates that don’t apply to themselves,” which is a standard example of a concept that gives rise to paradox. (Does this predicate apply to itself?) Of course this problem doesn’t arise if object-language and meta-language are distinguished, so that predicate_{object} ≠ predicate_{meta}, but Higginbotham and Schein may not like that solution.

6 Events.

Another innovation by Higginbotham & Schein was to combine the predicate-based treatment of plurals with a radical version of Davidsonian event semantics. The meaning of any sentence is taken to involve quantification over situations (i.e., events and states); the main verb is analyzed as a one-place predicate on situations; and its NP and PP arguments, as well as its tense- and aspect-markings and adverbial modifiers, all specify properties of the situation (such as agent, patient, instrument, location, time, etc.). Thus, [8, p. 170] gives an event-based analysis of

“Some Apostles lift the piano”

as, essentially,

∃X ⊆ Apostle′ : ∃e : [LiftThePiano′(e) ∧ ∀x[Agent(e, x) ⇒ X(x)].

Similarly, [27, pp. 88] renders

“The blue triangles are similar to the red triangles”

6These examples are derived from [8, pp. 169-170], but I simplified the formulas: verbs are rendered here by n-place predicates rather than by events; and I replaced (∀x [x(x) ⇔ P(x)]) by the logically equivalent P.
as:

$$\exists s[BeSimilar(s) \land \forall x[INFL(s,x) \Leftrightarrow (Blue'(x) \land Triangle'(x))]$$

$$\land \forall x[To(s,x) \Leftrightarrow (Red'(x) \land Triangle'(x))]].$$

Note that in this approach, every theta-role corresponds to a two-place predicate, where, for instance, $Agent(A,B)$ is true iff individual $B$ participates as one among possibly several co-agents in situation $A$. This has consequences for the treatment of singular nouns, including proper names, whenever they exhaustively specify the fillers of particular roles in situations. Thus,\(^6\)

“John hits Peter”

is not rendered simply as

$$\exists s[Hit(s) \land Agent(s,J) \land Patient(s,P)],$$

but as:

$$\exists s[Hit(s) \land \forall x[Agent(s,x) \Leftrightarrow x = J] \land \forall x[Patient(s,x) \Leftrightarrow x = P]].$$

Event-based semantics in and of itself has considerable virtues. The predicates expressed by verbs do have an internal structure that an event-based semantics attempts to lay bare, and this may make it easier to account for many semantic phenomena, such as optional verbal arguments, adverbial modifiers, tense and aspect. But, as Landman [11] points out, the optimal version of the event-based approach may be one that allows collective entities to serve as fillers of the event-roles. More often than not, avoiding collections is merely awkward and unwieldy. Schein [27], however, abolishes collections in order to avoid the purported Boolos-paradox, and recommends the event-based approach precisely because it allows this strategy. Oliver & Smiley[16] (2001) and Rayo [19, p. 445] argue that he is mistaken. They buy the Boolos argument, and construct analogous arguments for event-based treatments. Yi [32, p. 186, n. 34] presents a particularly interesting challenge. He shows that a contradictory predicate (F such that $F(e) \Leftrightarrow \neg F(e)$) can be constructed in Schein’s notation. His construction does not require that technical notions from the meta-language are exported into the object-language (as was the case with the Boolos-Higginbotham-Schein argument that we discussed above). Yi’s construction does crucially rely on the fact that in Schein’s logic, one and the same predicate may characterize a set of events (as if it translates a verb) and characterize a set of individuals (as if it translates a noun). This property is of course not a bug but a feature: one of the nice properties of event semantics is that one gets a treatment of nominalizations for free. But Yi’s construction suggests that it has a price; in its untamed form, it creates unwanted possibilities. One is reminded of the Kleene-Rosser paradox for the type-free lambda-calculus.

Schein [28, pp. 760-764] addresses these challenges, but I have not yet been able to assess the implications of this rebuttal. In any case, it seems to focus on Rayo and Oliver & Smiley rather than Yi.

7 Everything.

To return to the main issue of this note: I hope to have established that the Boolos-Higginbotham-Schein argument against collections is deeply flawed, and that Schein’s interesting contributions to the analysis of plurals do not save it. But the problems with the argument

\(^6\)This example is formally identical to some examples in [27, pp. 130].
that I discussed are so basic, that we may wonder what has been going on here. The researchers involved are not sophists. There must be a more constructive, positive correlate of the ideas I criticized. So that is what I want to discuss briefly now.

I already mentioned Boolos’ agenda, but it deserves some further discussion. It seems that there are two philosophical preconceptions at work, which dovetail in a surprising way. One is a very strict idea of reference, where the only things that can be referred to are individuals. This leads to design requirements for logics: variables can only range over individuals, predicates cannot apply to collections. Thus, Boolos’ desire to reinterpret second-order logic in line with these requirements — ultimately spawning a research tradition which is involved with the design of plural logics which satisfy such requirements, but which can nonetheless represent a rich variety of locutions involving plurality [17, 33, 34, 21]. This work must of course to some extent deal with the same issues that are faced in formal semantics, but the perspective is definitely different — the focus is on the properties of the formalism, and natural language is merely a source of inspiration. (In formal semantics, the logic is merely a tool to articulate an empirically motivated theory.)

An interesting consequence of the insistence on singular reference is, that the articulation of the model theory tends to be somewhat informal. A mathematically precise model theory involves reference to domains of discourse, usually characterized as sets or classes, and various subsets of them that constitute ranges of variables and domains of predicates and functions; the Spartan philosopher who wishes to stick to singular reference prefers not to mention such collections. In particular, he prefers not to characterize the domains of discourse as sets, or even better, not to characterize them at all. And this creates an opportunity for another, independently motivated philosophical agenda: the desire to quantify over “absolutely everything” [6, 20].

“Absolutely everything” is an appropriately sublime subject to bring this essay to a close. I will be able to do deal with this forbiddingly large topic within one small paragraph, because this Festschrift-paper is allowed to step outside of the boundaries of strictly academic discourse. I will merely express a feeling. (Articulating it better and more formally is left to future research.) And what I feel is: I do not trust the notion of “absolutely everything”. It ignores that we construct our mental worlds, and that we cannot conceptualize all possible extensions of these constructions. What there is is not given. It is nice if our logics can cast their nets wide, but absolutely everything is too much. Of course, logics with well-defined interpretations will always be caricatures of our actual thought and our actual language — thus, the idea of quantifying over everything may be felt as liberating. But precisely the suggestion that we can get away from the caricatures is badly misleading. Science makes models, and models are caricatures. And that is fine. We do not have to believe in them. They are just food for further thought.

References


