TRENDS IN LOGIC
Studia Logica Library

VOLUME 5

Managing Editor
Ryszard Wójcicki, *Institute of Philosophy and Sociology, Polish Academy of Sciences, Warsaw, Poland*

Editors
Petr Hájek, *Institute of Computer Sciences, Academy of Sciences of the Czech Republic, Prague, Czech Republic*
David Makinson, *Ville d'Avray, France*
Daniele Mundici, *Department of Computer Sciences, University of Milan, Italy*
Krister Segerberg, *Department of Philosophy, Uppsala University, Uppsala, Sweden*
Alasdair Urquhart, *Department of Philosophy, University of Toronto, Canada*

Assistant Editor
Jacek Malinowski, *Box 61, UPT 00-953, Warszawa 37, Poland*

SCOPE OF THE SERIES

Trends in Logic is a bookseries covering essentially the same area as the journal *Studia Logica* – that is, contemporary formal logic and its applications and relations to other disciplines. These include artificial intelligence, informatics, cognitive science, philosophy of science, and the philosophy of language. However, this list is not exhaustive, moreover, the range of applications, comparisons and sources of inspiration is open and evolves over time.

LOGIC, LANGUAGE AND REASONING

Essays in Honour of Dov Gabbay

Edited by

HANS JÜRGEN OHLBACH
King's College, London, United Kingdom

and

UWE REYLE
University of Stuttgart, Stuttgart, Germany

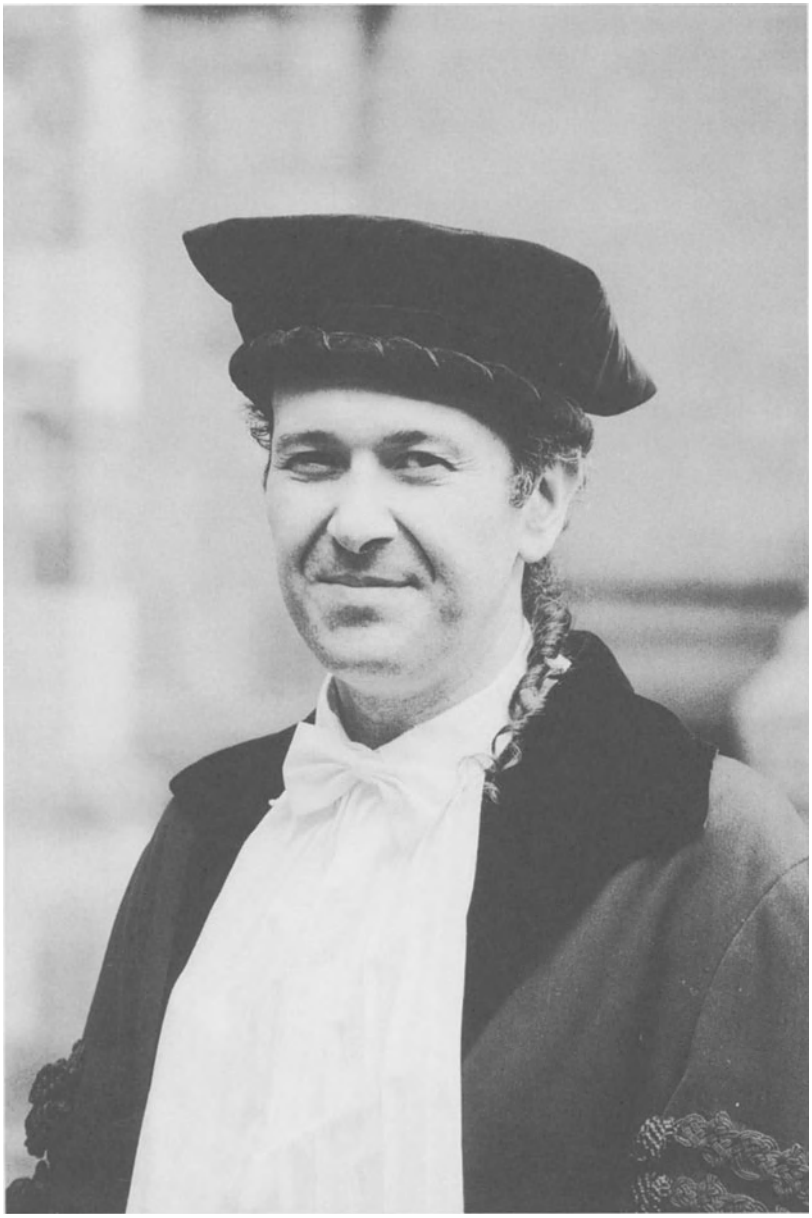
SPRINGER SCIENCE+BUSINESS MEDIA, B.V.

A C.I.P. Catalogue record for this book is available from the Library of Congress.

ISBN 978-94-010-5936-7 ISBN 978-94-011-4574-9 (eBook)
DOI 10.1007/978-94-011-4574-9

Printed on acid-free paper

All Rights Reserved
© 1999 Springer Science+Business Media Dordrecht
Originally published by Kluwer Academic Publishers in 1999
Softcover reprint of the hardcover 1st edition 1999
No part of the material protected by this copyright notice may be reproduced or utilized in any form or by any means, electronic or mechanical, including photocopying, recording or by any information storage and retrieval system, without written permission from the copyright owner.



Photograph by Elsbeth Mastenbroek
Reproduced with kind permission

WHAT A LINGUIST MIGHT WANT FROM A LOGIC OF MOST AND OTHER GENERALIZED QUANTIFIERS

HANS KAMP

When Dov and I received our logical education – Dov is quite a bit younger than I am, still we got our education at more or less the same time – the overall picture of what logic was, seemed comfortably clear. There were four main branches of mathematical logic – model theory, set theory, recursion theory and proof theory. Underlying this clear and simple picture were a number of widely shared assumptions, some of them to the effect that certain basic problems of logic had essentially been solved. Of central importance among these were: the belief that one had, through the work of Peano, Frege, Peirce, Russell, Hilbert, Gentzen and others, a definitive formal analysis of the notion of *logical deduction* (or *logical proof*); the belief that the conceptual problem of defining *logical consequence* and *logical truth*, and of explicating the relationship between those concepts and the concepts of *truth*, *reference* and *satisfaction* on one hand, and their relationship with the concept of a formal deduction on the other, had found a definitive solution in the work of Gödel and Tarski; and, finally, the conviction that with the characterizations of recursive functions proposed by Gödel, Turing and Church, one had uncovered what had to be the right concept of computability. With regard to set theory the situation was perhaps a little different; then as now, one could not help feeling that each of the available systems of set theory (the most popular ones, Z(ermelo–)F(raenkel) and G(ödel–)B(ernays), among them) embodied an element of arbitrariness. Nevertheless, for better or worse even in this domain a certain consensus had established itself which heavily favoured GB and ZF.

True, the picture wasn't really quite as simple as that. At the fringes hovered logical alternatives such as intuitionistic and other constructive logics; the basic concepts of set theory were challenged by the mereological logics; the spectre of undefinedness had produced, in the course of several decades, a still modest, but steadily growing literature on many-valued, probabilistic and partial logics; and the need for new logical tools for philo-

H.J. Ohlbach and U. Reyle (eds.), Logic, Language and Reasoning, 43–72.
© 1999 *Kluwer Academic Publishers.*

sophical analysis was beginning to give rise to a variety of new logical formalisms and to new and better meta-mathematical foundations for the formalisms already in existence. Decisive in this connection was Kripke's work on the semantics of modal and intuitionistic logic, which more than anything gave the impetus to what has developed into the vast and still growing field of modal logic in its comprehensive sense (encompassing such disciplines as tense logic, deontic logic, counterfactual logic, etc.) and which not only contributed to our conceptual understanding of those systems, but also established the foundations for their mathematical investigation.

Still, there was a strong tendency then to see all these alternatives as marginal. The core of logic remained – in the eyes of most, and certainly in the eyes of almost everyone who seemed to count – the four branches mentioned above; and one feature that those four branches shared was a primary, almost exclusive preoccupation with the new Characteristica Universalis, the predicate calculus – in the first place its first-order fragment, but, to a lesser extent, also parts of higher-order logic, or alternative extensions of first-order logic such as the infinitary logics.

If since that time the picture has changed dramatically, Dov Gabbay certainly has been foremost among those to whom that change is due. Already in the days when modal logic was only beginning to develop into the sophisticated field it has become, he made substantial contributions to it, many of which have become so much part of the logician's intellectual equipment that many who have joined the field in the course of the past three decades and who now make up the clear majority of its active representatives, aren't even aware that they owe these tools and insights to him. Yet emphasizing solely the important work that Dov has done – over so many years – on modal and related logics would seriously understate the influence he has had on our general understanding of what logic is and ought to be, an influence which continues to be as strong as it ever was.

It is important in this connection to note in what ways the general conception of logic has changed, and what have been the forces that have led to that change. As it appears to me, the central difference between the views of logic that are held by many today and the earlier one I sketched in the opening paragraphs, is that in the meantime we have arrived at a much more abstract, and, therewith, a more comprehensive, perception of what logic is about – a much more abstract perspective on what can qualify as a formal analysis of reasoning and what counts as a logical formalism (or 'logical language') suitable for the representation and manipulation of information. Pressure towards such a more liberal perspective has come from many different directions – philosophy, computer science, artificial intelligence, linguistics and (inasmuch as that is a discipline in its own right) computational linguistics. Of course, the strongest personal influence on this change has come from those at home in those neighbouring disciplines as well as in the classical branches of symbolic logic itself, and most of all from those, if any, who were conversant in all these fields at the same time. It needs no comment that such individuals are few and far between. Still, their class is not empty; though it may well be that it equals {Gabbay}.

To the needs coming from those neighbouring disciplines – for formalisms providing novel means of expression, new ways of defining the semantic values of expressions, new ways of defining inference or computing inferences – the logical community has not only responded with a forever expanding panopticum of different logical systems; it has also reacted by rethinking its own credo, and tried to come up with abstract, meta-logical analyses of what the central concepts of logic, those which anything deserving the predicate 'logic' should instantiate, might be. And here again, Dov has played a pivotal role, for instance through his work on the question: what is a relation of logical inference? Or, more recently, through his development of the concept of Labelled Deduction.

Labelled Deduction is perhaps the most promising of a number of current proposals of frameworks in which (hopefully) the entire spectrum of logical alternatives which offer themselves to the bewildered observer today can be compared and helpfully classified, at least when this task is seen from a proof-theoretical perspective. Thus it promises to fill the increasingly felt need for a conceptually sound and accessible map through the labyrinthine landscape of contemporary formal logic, in which the potential customer, whether from philosophy, linguistics or computer science, is in danger of getting lost almost as soon as he makes an attempt to enter.

The present paper pursues by no means so lofty a purpose as this. Rather than concerning itself with the labyrinth of logics as a whole, it looks at one little corner of what is itself only a (somewhat larger) corner of that labyrinth. Still, it seems to me that the lesson which can be gleaned from the little exercise we will go through is applicable to all or most of the larger corner as a whole, and even that it throws some light on the larger question that concerns the relationship between logic and one of its domains of application: the semantics of natural language.

As its title makes clear, the paper is about the quantifier *most*. More generally, it tries to address the question what can and should be expected from a logic of generalized quantifiers. The motivation comes from the semantics of natural language and has an eye not only on the correct logical representation of the quantificational devices that natural languages employ, but also on the computability of those representations and their logical properties. I must add emphatically that from the perspective of mathematical logic the paper offers hardly anything that is really new. All the facts discussed in Sections 1 and 2 – they are presented as lore, and rightly so, for most people who are reasonably familiar with the metamathematics of generalized quantifiers have known about these facts for quite a long time, and any competent logician who hasn't actually seen a proof of them should have little trouble concocting one himself – and most of those of Section 3, can be found explicitly or implicitly in the existing literature. See for instance [10], esp. Section 1.7.

To my knowledge, the paper does not relate in any direct way to Dov's own work. My excuse is that it is surely much harder to find a topic which does not directly relate to any of his work than to find one which does. What better way could there be to pay homage to this œuvre than by finding one of the few logical niches which it has left untouched? But then, probably I have failed anyway and all I am going to say, and more, is hidden somewhere in some paper of Dov's that I have missed.

1 Some Established Views on 'most' and Other Generalized Quantifiers

I regard as uncontroversial that nominal quantification in natural languages such as English has the logical form of what have come to be called Generalized Quantifiers: operators which take a pair of formulas as arguments and return a new formula, while binding a variable.[1] In fact, this is as true of the standard quantifiers *every* and *some* as it is of others (such as *many* or *most*); and it is a simple exercise to develop a version of first-order logic, straightforwardly inter-translatable with its standard versions, in which the universal and existential quantifier are formally treated as generalized (i.e. two-place, not one-place) quantifiers.[2] In a way, in the context of this paper such a version, in which even the standard quantifiers are two-place, would make for greater uniformity. But I believe the presentation will be more perspicuous if first-order predicate logic is kept in the form in which most of us are familiar with it. So I will assume, as 'basis logic', a first-order language L_0 with an infinite set of individual variables $x_1, x_2, x_3, \ldots$; infinitely many predicate constants $P_1^n, P_2^n, P_3^n, \ldots$ for each arity n; the connectives $\neg, \wedge, \vee, \rightarrow$ and $\leftrightarrow$, the quantifiers $\forall$ and $\exists$, and the identity $=$. x, y and z are the first three variables x_1, x_2, x_3 and P and Q the first two 1-place predicate constants P_1^1 and P_2^1.

It was one of Frege's insights, which led to the predicate calculus as we now have it, that the universal and existential quantifier *can* be treated as one-place operators. That from the point of view of the grammar of English

[1]Recent work on natural language quantification, especially that of [7] and [8], has shown convincingly that the quantificational possibilities in English and other natural languages go well beyond this – there are expressions that must be analyzed as operators taking more than two formulas as arguments and/or binding more than one variable. Such constructs will play no role in this paper.

[2]To prove the point (if a proof is wanted) see [5], footnote 1.

(or, historically more accurately, German) they rather behave like two-place operators (i.e. as generalized quantifiers) than as the quantificational devices he adopted in his Begriffsschrift, is something of which he was as much aware as anyone. But he noted that for both these quantifiers the contributions made by the two arguments can be contracted into one – by forming material conditionals in the one case and conjunctions in the other; and, for reasons we need not go into here, these are the devices that have remained with us ever since.

It has long been part of the general lore surrounding natural language semantics that *every* and *some* are quite special in this respect. In general such a Boolean reduction of a two-place to a one-place quantifier is not possible. I have called this part of semantic lore since it is a conviction that many take for granted even though it is not established by actual proof. The principal reason for this is that a proof presupposes a well-defined semantics for the quantifier that is to be shown irreducible, and such a semantics is rarely available. A notorious exception – perhaps one should say: *the* notorious exception – is the quantifier *most.*

There is a fairly general consensus that 'Most As are Bs' is true provided the cardinality of the set of As that are Bs exceeds that of the remaining As, or at least that this is so, provided the number of As is finite. Since these two conditions will play a central part in the paper, let us give them a label right away:

(MOST) 'Most As are Bs' is true iff $|A \cap B| > |A \setminus B|$

(MOST^{FIN}) If A is finite, then
'Most As are Bs' is true iff $|A \cap B| > |A \setminus B|$.

This second, weaker assumption suffices to show that *most* is not reducible to a 1-place operator – or, to put it differently, we can show the slightly stronger result that such a reduction isn't possible within the Theory of Finite Models. More precisely we can show Fact 1.

Fact 1 There is no combination of (i) a function F from finite sets U to sets of subsets of U and (ii) a first-order formula $\Phi(P,Q;x)$, built up from the predicate constants P, Q, variables and logical constants, in which at most x occurs free, such that for every finite first-order model $M = \langle U, I\rangle$:

$$|I(P) \cap I(Q)| > |I(P) \setminus I(Q)| \text{ iff } \{u \in U : M \models \Phi(P,Q;x)[u]\} \in F(U).$$

To see that Fact 1 says what it ought to, first observe that a one-place variable binding operator O turns, when it binds, say, the variable x, a formula that has only x free into a sentence. Semantically this means that O maps the satisfaction set of any such argument formula to a truth value. More specifically, if M is any model, the interpretation O_M of O in M must be a function that maps for any such formula ϕ the set of individuals of M which satisfy ϕ in M to one of 0 and 1. Thus O_M must be (the characteristic function of) a set of such satisfaction sets. If we make the additional (highly plausible and generally endorsed) assumption that O_M ought not to depend on the interpretation of any non-logical constant in M and thus that it depends exclusively on the universe U of M, it follows that the meaning of O can be given as a function F from sets U to sets of subsets of U. The interpretation of O in any model M will then be the value $F(U_M)$ which F assigns to the universe of M.

Second, a reduction of *most* of the kind in question will involve a way of combining its argument formulas $A(x)$ and $B(x)$ into a single compound formula $\Phi(A(x), B(x))$ such that the generalized quantifier relation $MOST$ holds between the satisfaction sets of A and B if and only if the satisfaction set of $\Phi(A(x), B(x))$ belongs to the interpretation of the operator O. This will have to be so in particular in cases where the arguments of *most* are the formulas $P(x)$ and $Q(x)$ and in models M in which all non-logical constants other than P and Q are given a trivial interpretation (e.g. every n-place predicate is interpreted as the empty n-place relation.) In such cases $\Phi(A(x), B(x))$ reduces to a formula $\Phi(P, Q; x)$ of the sort mentioned in the statement of Fact 1. Thus Fact 1 entails the irreducibility of *most.*

N.B. the statement made by Fact 1 goes beyond what I promised insofar as the formula $\Phi(P, Q; x)$ may involve (standard first-order) quantification as well as Boolean connectives. In this regard the result is more general than a strict analogue to the reducibility of the generalized quantifiers *every* and *some*, where the combination of the two argument formulas requires only the sentential connectives $\rightarrow$ and $\wedge$, respectively.

The proof of Fact 1 rests on long known facts about monadic first-order logic and would hardly be worth a looking into if it didn't provide some insight into the question what is likely to be needed to obtain similar irreducibility results for other quantifiers than *most.* It is with this purpose in mind that I will take a little time to remind the reader of how the argument might go.[3]

[3]As matters have turned out, no further use of the proof is made in the present paper. However, in more comprehensive joint work with Tim Fernando, in which we investigate other non-standard quantifiers besides *most* and *more*, we intend to exploit this possibility. (See [2].) In retrospect, and thanks to critical remarks by Johan van Benthem, I now feel that this first section should have been written quite differently, and that a much more compact presentation would have served the purpose better.
Another defect of the section is that it does not relate the notions of definability and reducibility for quantifiers sufficiently to those that can be found in the literature on this subject. So to those familiar with this literature the section will appear rather amateurish. And for anyone familiar with the standard techniques for proving results in this domain – such as, in particular, those using Ehrenfeucht games or the notion of partial isomorphism – the old-fashioned, 'syntactic' kind of argumentation I have used will undoubtedly reinforce that impression. This is another reason why the section should have been rewritten. But unfortunately, time prevented me from doing the necessary replace-

Proof of Fact 1. (Sketch) I will state, in a form convenient to the present purpose, the facts about monadic logic which we will need. As said, these facts are standard; they, or something very much like them, is involved in familiar proofs that monadic logic has the finite model property; and they can be established by a well-known quantifier elimination argument. Let $L(P,Q)$ be the language of first-order logic with identity whose only non-logical constants are P and Q. There are sentences of $L(P,Q)$ which express the following properties of models $M = \langle U, I\rangle$ for $L(P,Q)$:

1. For $n \geq 1$ and natural numbers $m(P,Q)$, $m(P,\neg Q)$, $m(\neg P,Q)$ such that $(m(P,Q) + m(P,\neg Q) + m(\neg P,Q)) \leq n$ the proposition that
   (a) $|U| = n$,
   (b) the number of individuals in M satisfying both P and Q is $m(P,Q)$
   (c) the number of individuals satisfying P but not Q is $m(P,\neg Q)$, and
   (d) the number of individuals satisfying Q but not P is $m(\neg P,Q)$.

   (We will refer to these sentences as $\Psi_{n;m(P,Q),m(P,\neg Q),m(\neg P,Q)}$.)
2. For $n \geq 1$ and natural numbers $m(P,Q)$, $m(P,\neg Q)$, $m(\neg P,Q)$, $m(\neg P,\neg Q) < n$, and such that $(m(P,Q) + m(P,\neg Q) + m(\neg P,Q) + m(\neg P,\neg Q)) \geq n$ the proposition that
   (a) $|U| > n$,
   (b) the number of individuals in M satisfying both P and Q is $m(P,Q)$
   (c) the number of individuals satisfying P but not Q is $m(P,\neg Q)$, and
   (d) the number of individuals satisfying Q but not P is $m(\neg P,Q)$.
   (e) the number of individuals satisfying neither P nor Q is $m(\neg P,\neg Q)$.

   (We will refer to these sentences as $\Psi_{>n;m(P,Q),m(P,\neg Q),m(\neg P,Q),m(\neg P,\neg Q)}$.)
3. For $n \geq 1$ and natural numbers $m(P,Q)$, $m(P,\neg Q)$, $m(\neg P,Q) \leq n$ the proposition that
   (a) $|U| > n$,
   (b) the number of individuals that are P and Q, the number of those that are P but not Q and the number of those that are Q but not P are $m(P,Q)$, $m(P,\neg Q)$, $m(\neg P,Q)$, respectively, and
   (c) the number of elements that are neither P nor Q is $> n$;
   this sentence is denoted as $\Psi_{>n;m(P,Q),m(P,\neg Q),m(\neg P,Q)}$; analogously there are sentences $\Psi_{>n;m(P,Q),m(P,\neg Q),m(\neg P,\neg Q)}$; $\Psi_{>n;m(P,Q),m(\neg P,Q),m(\neg P,\neg Q)}$; $\Psi_{>n;m(P,\neg Q),m(\neg P,Q),m(\neg P,\neg Q)}$;

ment job in the way in which it should be done. A humble request to the cognoscenti: Please skip this section!

the first of these says that $|U| > n$, that there are $m(P, Q)$ elements that are P and Q, $m(P, \neg Q)$ that are P but not Q and $m(\neg P, \neg Q)$ that are neither P nor Q, while the number of elements that are Q but not P is $> n$; similarly for the other three.

4. In analogy with the sentences mentioned under 3, there are those which say of two of the four categories that there are $\leq n$ individuals of that category and say exactly how many there are, while of the remaining two categories there are $> n$ – these sentences are denoted as $\Psi_{>n;m(P,Q),m(P,\neg Q)}$, $\Psi_{>n;m(P,Q),m(\neg P,Q)}$ etc. – and there are sentences $\Psi_{>n;m(P,Q)}$; $\Psi_{>n;m(P,\neg Q)}$; $\Psi_{>n;m(\neg P,Q)}$; $\Psi_{>n;m(\neg P,\neg Q)}$, saying of just one category that there is some particular number $m \leq n$ of elements of that category, whereas for each of the other three there are more than n; and finally there is a sentence $\Psi_{>n;>}$ which says that there are more than n elements of each of the four categories.
5. Corresponding to each of the sentences $\Psi_{n;m(P,Q),m(P,\neg Q),m(\neg P,Q)}$ for which $(m(P, Q)+m(P, \neg Q)+m(\neg P, Q)) < n$ there are four $L(P, Q)$ formulae with x as only free variable, to which we will refer as $\Psi_{n;m(P,Q),m(P,\neg Q),m(\neg P,Q)}(P, Q; x)$, $\Psi_{n;m(P,Q),m(P,\neg Q),m(\neg P,Q)}(P, \neg Q; x)$, $\Psi_{n;m(P,Q),m(P,\neg Q),m(\neg P,Q)}(\neg P, Q; x)$ and $\Psi_{n;m(P,Q),m(P,\neg Q),m(\neg P,Q)}$ $(\neg P, \neg Q; x)$. $\Psi_{n;m(P,Q),m(P,\neg Q),m(\neg P,Q)}(P, Q; x)$ is satisfied by $u \in U$ if $|U| = n$, there are $m(P, Q)$ individuals *other than* u which are both P and Q, $m(P, \neg Q)$ individuals *other than* u which are P but not Q, $m(\neg P, Q)$ individuals *other than* u which are Q but not P, while moreover u is a P as well as a Q; $\Psi_{n;m(P,Q),m(P,\neg Q),m(\neg P,Q)}(P, \neg Q; x)$ is satisfied by u if the same conditions obtain except that u is a P but not a Q; and similarly for the remaining two formulas.
6. Similarly there are four formulas with free x for each of the sentences described in 2, 3 and 4. (Thus, to take just one example, there is a formula $\Psi_{>n;>}(P, Q; x)$ which is satisfied by u iff there are more than n individuals, there are more than n individuals different from u which are both P and Q,..., and u itself is both P and Q.)
7. For each formula $\phi(x)$ of $L(P, Q)$ in which only x occurs free there is a number n_ϕ such that $\phi(x)$ is logically equivalent to a disjunction of formulas of the types described under 5 with $n \leq n_\phi$ and those in 6 with $n = n_\phi$.

7 gives us the result we are trying to establish (i.e. Fact 1) fairly straightforwardly. For suppose there was a formula $\Phi(P, Q; x)$ as described in the statement of Fact 1. Then there would be a number n_Φ as described in 7 such that $\Phi(P, Q; x)$ is equivalent to a disjunction D of the indicated kind. Now consider any model $M = \langle U, I\rangle$ such that $|U| = 8 \times n_\Phi$ and in which there are more than $n_\Phi + 1$ individuals that are both P and Q, more than $n_\Phi + 1$ individuals which are P but not Q, etc. It is clear that the set D_M

of those $u \in U$ which satisfy D in M will consist of the union over some subset (possibly empty) of the following four disjoint sets: (i) the set of individuals that are both P and Q in M, (ii) the set of those that are P but not Q, and (iii, iv) similarly for the other two combinations, Q but not P and neither P nor Q. Whether or not the first of these sets is part of D_M depends on whether D contains as one of its disjuncts the formula $\Psi_{>n;>}(P, Q; x)$. For any other possible disjunct of D will fail to be satisfied by a u that is both a P and a Q in M either because what it says about the size of U or else because of what it says about the number of individuals that are P and Q, P but not Q, Q but not P, or neither P nor Q; or, finally, because it requires u to be not a P or not a Q. Similarly, the second set is part of X_1 iff D contains the disjunct $\Psi_{>n;>}(P, \neg Q; x)$ and likewise for the remaining two. This gives us a small, finite number of possibilities for D_M: the empty set $\{\}$, the set of u which are both P and Q, the set of u which are P but not Q, the union of these two sets, i.e. the set of u which are P, etc. with as largest possibility the set U itself. It is tedious, but not hard, to construct for each of those possibilities a pair of models $M_1 = \langle U, I_1 \rangle$ and $M_2 = \langle U, I_2 \rangle$ which satisfy the above conditions for M and which are such that

1. according to our adequacy criteria (MOST^{FIN}) for MOST most Ps are Qs in M_1 but not in M_2, and
2. $D_{M_1} = D_{M_2}$.

We will consider just two cases, that where D_M is the set $I(P) \cap I(Q)$ and that where it is $(I(P) \setminus I(Q)) \cup (I(Q) \setminus I(P))$. In the first case let $I_1(P) \cap I_1(Q) = I_2(P) \cap I_2(Q) = I_1(P)$ be a subset of U of $n_\Phi + 2$ elements and let $I_2(P)$ be U. Then evidently both 1 and 2 are satisfied. For the second case let M_2 be as in the preceding case and let M_1 be like M_2 but with the interpretations of P and Q reversed. Since in the present case D_M is symmetric in P and Q, 2 is satisfied again. Moreover, it should be clear that most Ps are Qs in M_1, as there the Ps are included in the Qs while, as in the first case, only a minority of the Ps are Qs in M_2. ■

The reader will no doubt agree that this proof is every bit as unenchanting as I promised it would be. The point of presenting it nevertheless is, as I said before embarking upon it, that very similar arguments may well be usable to show the irreducibility of other quantifiers, such as, say, *many*, and that this may require comparatively weak assumptions about the semantics of such a quantifier. For instance, it would be enough to assume that (if necessary, only under certain conditions, provided these are compatible with the set of As and the set of Bs being of arbitrarily large finite size) the truth of 'many As are Bs' requires that some suitable proportion of the As are Bs.)

There is a second point to be made, this one not so much about the proof of Fact 1, but rather about what the Fact asserts. What it asserts is *not* that the quantifier MOST is not first-order definable. By the first-order definability of a generalized quantifier we mean the following. First, by a *generalized quantifier relation* we understand a function from sets A to sets of pairs of subsets of A. (Generalized quantifier relations are the kinds of objects that are to serve as meanings of binary generalized quantifiers. The motivation for the definition is the same as the one given above for the meaning of a one-place quantifier as a function from sets to sets of subsets of those sets.) Suppose R is such a relation and that $\Psi_R(P,Q)$ is a sentence of $L(P,Q)$. Then we say that R is *first-order defined by* $\Psi_R(P,Q)$ iff for any model $M = \langle U, I\rangle$ for $L(P,Q)$:

$$\langle I(P), I(Q)\rangle \in R(U) \text{ iff } M \models \Psi_R(P,Q); \tag{1}$$

and R is said to be *first-order definable* iff there exists such a sentence. Similarly, R is said to be *first-order defined by* $\Psi_R(P,Q)$ *in the Theory of Finite Models* iff (1) holds for all finite models.

The point of these notions should be clear: if a generalized quantifier Qu has as its meaning a generalized quantifier relation R which is first-order defined by a formula $\Psi_R(P,Q)$, then any sentence δ containing occurrences of Qu will be equivalent to a sentence η in which Qu does not occur; η is obtained by replacing, going from the inside out, every subformula $Qu_v(\phi,\psi)$ of δ by a formula $\Psi'_R(\phi,\psi)$ which we get by (a) taking an alphabetic variant $\Psi'_R(P,Q)$ of $\Psi_R(P,Q)$ such that the variables of $\Psi'_R(P,Q)$ are disjoint from the free variables of $Quv(\phi,\psi)$ and (b) replacing in $\Psi'_R(P,Q)$ every subformula $P(w)$ by $\phi(w/v)$ and every subformula $Q(w)$ by $\psi(w/v)$.

First order definability is clearly a different concept from the notion of reducibility which was used in Fact 1, and which in general terms can be characterized as follows:

A generalized quantifier R is *reduced to* a one-place quantifier meaning F (i.e. a function from sets U to sets of subsets of U) *by* a formula $\Psi_R(P,Q;x)$ iff for each model $M = \langle U, I\rangle$ for $L(P,Q)$

$$\langle I(P), I(Q)\rangle \in R(U) \text{ iff } \{u \in U : M \models \Psi_R(P,Q;x)[u]\} \in F(U). \tag{2}$$

Again, we say that R *is reduced* to a one-place operator *in the Theory of Finite Models* iff the above condition holds for all finite models for $L(P,Q)$.

It is easy to see that first-order definability entails reducibility to a one-place operator. For suppose that R is first-order definable by $\Psi_R(P,Q)$. Then the formula $x = x \wedge \Psi_R(P,Q)$ will (trivially) reduce R to the one-place operator which maps each set U onto $\{U\}$. Of course, the converse entailment does not hold: there are uncountably many one-place quantifiers F which are not first-order definable, in the sense that there is no sentence

$\Psi_F(P)$ of the language $L(P)$ such that for all $M = \langle U, I\rangle$, $I(P) \in F(U)$ iff $M \models \Psi_F(P)$. For each such quantifier F we can make up any number of 2-place quantifiers reducible to it; consider for instance the generalized quantifier relation R_F defined by the condition that for any set U and subsets A, B of U, $\langle A, B\rangle \in R_F(U)$ iff $A \in F(U)$. This relation is reduced to F by the formula $P(x)$. And it is easy to see that any first-order definition for R_F would yield a first-order definition for F in the sense just given. For suppose that $\Psi_R(P, Q)$ were a first-order definition of R_F. Then the sentence $\Psi'_R(P)$ (= '$\Psi_R(P, \top)$'), which we obtain by replacing in $\Psi_R(P, Q)$ each subformula $Q(v)$ by the corresponding formula $v = v$, would be a first-order definition of F. Thus R_F cannot be first-order definable.

2 Another Piece of 'most'-lore: Non-axiomatizability

The next bit of lore about *most* I must mention is that adding it to first-order logic leads to non-axiomatizability.[4] What is meant is this. Suppose we extend our first-order language L with a generalized quantifier symbol Mo, subject to the syntactic rule that

> if ϕ and ψ are formulas of the new language $L(Mo)$ and v is any variable, then $Mo_v(\phi, \psi)$ is a formula; (3)

and the accompanying semantic principle that for any model $M = \langle U, I\rangle$

$$M \models_a Mo_v(\phi, \psi) \text{ iff } \langle\{u \in U : M \models_{a[u/v]} \phi\}, \{u \in U : M \models_{a[u/v]} \phi\}\rangle \in \text{MOST}(U), \tag{4}$$

where MOST is the binary generalized quantifier we choose to interpret Mo. Together with the familiar clauses of the truth definition for first-order logic (4) provides us with the usual characterizations of logical consequence (as preservation of truth in all models) and of logical truth (as truth in all models). Then, as lore has it, neither the consequence relation nor the set of logical truths of the resulting language $L(Mo)$ is recursively enumerable.

Whether the claim is true depends of course on exactly what the generalized quantifier MOST is taken to be; and here for the first time the distinction between the strong version (MOST) and the weak version (MOST^{FIN}) of our intuitive meaning constraint for the quantifier *most* becomes important. For it is only when we adopt the strong version that the claim holds

[4] Proofs of this fact seem to be ten to the gallon and have been around for (probably) at least two decades. For instance, a slightly different demonstration can be found in [10], leading to a more informative result than will be given here – but one which is for our present aims is not needed in its full strength.

true. This constraint fixes the generalized quantifier relation MOST completely. For now and later reference we repeat the definition:

Definition MOST is the function which maps each set U onto the set of all pairs $\langle V, W\rangle$ such that $V, W \subseteq U$ and $|V \cap W| > |V \setminus W|$.

We state the strongest part of the claim just made, the non-recursive enumerability of the set of logical truths, as Fact 2:

Fact 2 Let $L(Mo)$ be the language defined above, through the clauses (3) and (4). Then the set of all logical truths of $L(Mo)$ is not recursively enumerable.

Proof. Here is a simple proof of this fact. Let L_{Ar} be a sublanguage of L suitable for the formulation of arithmetic (I assume that the operations of successor, plus and times are represented by corresponding predicates) and let T_{Ar} be some finite axiomatization of first-order arithmetic strong enough to yield Gödel's incompleteness theorem and to prove that every model has an initial segment isomorphic to the standard model of arithmetic. Suppose we add the quantifier Mo to L_{Ar}, thus obtaining the language $L_{Ar}(Mo)$ and extend T_{Ar} with a single axiom of the following form (modulo some straightforward definitions)

$$(\forall y)(Mo_x(x \leq y, (\exists z)(x = z + z)) \vee Mo_x(x \leq y + 1, (\exists z)(x = z + z))). \quad (5)$$

Given (4) and our identification of MOST, (5) says that for any number y (finite or transfinite) either the cardinality of the even numbers $\leq y$ exceeds that of the set of the remaining numbers $\leq y$, or else the cardinality of the even numbers $\leq y + 1$ exceeds that of the set of the remaining numbers $\leq y + 1$. It is clear that this condition is satisfied for every finite number y (the first disjunct is true when y is even, the second when y is odd) but that it fails for any transfinite number (for then the sets that are being compared are all denumerably infinite and thus of the same cardinality). Thus the only model of the theory T_{Ar} + (5) (up to isomorphism) is the standard model of arithmetic. But then, if Ψ is the conjunction of the axioms of T_{Ar} + (5), we have that for any sentence ϕ of L_{Ar} ϕ is true in the standard model of arithmetic iff the sentence $\Psi \rightarrow \phi$ is a logical truth of L_{Ar}. ∎

It is important to note that this proof depends crucially on the assumption that the semantics for Mo satisfies the condition (MOST) of the preceding section also for infinite sets A and not only for finite ones. Indeed, we will see in the next section that if we weaken the assumptions of Fact 2 in that we replace (MOST) by (MOST^{FIN}) the assertion it makes is no longer true.

3 An Axiomatizable Logic for 'most'

When reflecting on the implications of Fact 2, we do well to ask once more what and how good is the intuitive justification for conditions such as (MOST) and (MOST^{FIN}). In Section 1 I ventured the observation that there is a firmer consensus concerning (MOST^{FIN}) than there is concerning the more comprehensive condition (MOST). Perhaps this claim is more a reflection of my own preferences than the true description of an actual distribution of opinion. In any case, I have my preferences and this is the place to try and account for them.

It seems to me that when the set A is finite, counting the set of As that are Bs and the set of As that are not Bs and finding there are more things in the first set than there are in the second amounts to a conclusive demonstration that most As are Bs. This is connected with the circumstance that counting a set seems to be *the* criterion for determining its size as long as the set is finite – an intuition that is reflected in the set-theoretic fact that for the finite sets the concepts of cardinal and of ordinal coincide. For infinite sets, in contrast, there is no clear pretheoretic conception of how their size should be assessed, and it seems that precisely for this reason our intuitions about when sentences of the form 'Most As are Bs' are true become uncertain too. The concept of cardinality as a measure of set size was a profound discovery when it was made and since then it has become central to the ways in which we deal with the infinite in mathematics. But cardinality remains a term of art, which has no more than a tenuous connection with the intuitions of the ordinary speakers of natural languages.

As far as those intuitions are concerned, it seems rather that when infinite sets come into play, the concept of 'majority' that one fastens on to form a judgement about the truth or falsity of a *most*-sentence varies with context, and may take factors into account that fall outside the conception of generalized quantifier meaning which has guided us so far. The stock examples

a. Most natural numbers are prime. (6)
b. Most natural numbers are not prime.

remain good illustrations of the point at issue. The tendency to judge the first sentence as false and the second as true – or at any rate, to find it much more plausible that the second should be true and the first one false than the other way round – surely reflects our inclination to think of the *rates* with which we are likely to encounter prime or non-prime numbers when going through the numbers in some special order (e.g. going up the standard ordering) or, alternatively, at random. Indeed, there exists a cluster of number-theoretic theorems which confirm these intuitions: for a vast family of ways to sample the numbers in some order the rate with which one encounters non-primes tends towards 100% while the rate with which one encounters primes tends to 0%.

What morals is the natural language semanticist to draw from these considerations? I do not know of any consensus on this point. But let me put forward my own assessment. First, a realistic semantics should respect speakers' intuitions as much as possible, and this should include cases where speakers' intuitions are unstable or simply missing; in these cases semantic theory should withhold judgement too, or it should try to identify the different conflicting strains of conceptualization that are responsible for the instability. For the case at hand – *most* applied to infinite sets – these recommendations should, I reckon, come to something like this:

(a) Eventually, the different conceptual elements that typically enter into speakers' judgements about sentences such as (6.a) and (6.b) and the ways in which they shape those judgements will have to be identified. This will evidently lead to an analysis of *most* according to which its meaning is something other (and more complicated) than the generalized quantifier relations considered hitherto. As far as I know, this is a research topic on which some work has been done (see the remarks on Colban below), but where there is much to be done still. It is a topic, however, which will not be explored here.

(b) Short of engaging in the kind of investigation advocated under (a), a semantics of *most* should remain agnostic in those cases where speakers' judgements depend on factors which are outside of the conceptual apparatus provided by quantifier meanings in the narrow sense. For a model-theoretic analysis this may have two different implications. First, that of a partial model theory in which sentences need not get a definite truth value in every model. (In particular sentences of the form 'most As are Bs' may fail to be either true or false in models where the number of individuals satisfying A is infinite.) Alternatively, one may adopt a model theory in which every model determines a truth value for all sentences, but where, intuitively speaking, several non-equivalent models may correspond to one and the same possible state of affairs, *viz.* by providing different interpretations for the generalized quantifier. (Intuitively: whenever the judgement about truth or falsity of a *most*-sentence with respect to a given state of affairs depends on such factors, some of the models compatible with that state of affairs may assign the sentence the value true while other such models assign it the value false.)[5]

[5]The difference between these two options – a partial model theory or a non-partial model theory which allows for different models corresponding to a single state of affairs – need not be all that different from each other in the end. This is one of the main lessons

These recommendations seem to me to be in the spirit of a paper by Colban [1], which has played an important part in shaping the thoughts on which the present contribution reports. When compared with the way we have been approaching the problems posed by *most*, Colban's approach could be said to start at the opposite end. Rather than trying to determine of some particular natural language quantifier, such as *most*, exactly what its meaning is and then investigating the logic that is generated by the meaning one has fastened upon, Colban begins by having a look at so-called *weak logic*, the logic for the extension $L(Qu)$ with one new binary quantifier symbol that is generated by the class of all models M for $L(Qu)$ in which the new quantifier is interpreted by any relation between subsets of U_M whatever. (In other words, this is the logic of the concept of a generalized quantifier in its full generality, in which properties that differentiate between such quantifiers are entirely ignored. The idea of weak logic appears to be quite old; one finds it for instance already in one of the mile stones in the history of generalized quantifier theory, Keisler [9], except that Keisler is concerned with a one-place quantifier – 'there are uncountably many' – rather than with the two-place quantifiers considered here and in Colban's work; a discussion of the weak logic of binary quantifiers can also be found in Appendix B of [10].) Once an axiomatization for weak logic is in place, one can then proceed, as Westerstahl and Colban do, to impose conditions on the admissible quantifier meanings and extend the axiomatization of weak logic accordingly. Those interested in the logic of some particular quantifier, semantically given by some particular generalized quantifier relation R, might wish to use this strategy to whittle down the class of permitted quantifier relations step by step until one reaches the singleton class consisting solely of R. But of course, one should be prepared for the contingency that this is too much to hope for: perhaps that no matter how the strategy is applied the resulting class will always contain some relations besides R.

However, in the light of our reflections earlier in this section reducing the class to a singleton set may not be the right goal anyway. In particular, I suggested, the best account of *most* as a generalized quantifier might well be one that admits a variety of quantifier relations, which may yield incompatible predictions about the truth of certain *most*-sentences concerned with infinite sets, while harmonizing in their predictions about sentences speaking of finite sets. Indeed, it is just such an account which I shall present here.

As a basis for our further explanations we need an axiomatization of weak logic for the language $L(Mo)$ (where Mo is, as before, a binary quan-

---

of the so-called supervaluation approach to problems of semantic underspecification. See e.g. [3], [4] or [6].

tifier symbol).[6] As can be shown by a largely standard Henkin argument, addition of the universal closures of all instances of the following schemata to a complete axiomatization of first-order logic (with the rules of Modus Ponens and Universal Generalization) is complete for this logic:

WQL.1 $(\forall v_i)(\phi \leftrightarrow \psi) \rightarrow (Mo_{v_i}(\phi, \chi) \leftrightarrow Mo_{v_i}(\psi, \chi))$

WQL.2 $(\forall v_i)(\phi \leftrightarrow \psi) \rightarrow (Mo_{v_i}(\chi, \phi) \leftrightarrow Mo_{v_i}(\chi, \psi))$

WQL.3 $Mo_{v_i}(\phi, \psi) \rightarrow Mo_{v_j}(\phi', \psi')$,
if $Mo_{v_i}(\phi, \psi)$ and $Mo_{v_j}(\phi', \psi')$ are alphabetic variants.

But where do we go from here? First a decision of convenience. In the remainder of this section I will follow Colban in pursuing an axiomatization not of the quantifier *most*, but instead for the quantifier usually referred to as *more*, which relates its arguments A and B in a way that can be paraphrased as 'there are more As than Bs'. Thus, corresponding to the 'standard semantics' for *most*, which is given by the truth condition

$$M \models_a Most_v(\phi, \psi) \text{ iff } |V \cap W| > |V \setminus W|, \quad \text{where } V = \{u \in U_M : M \models_{a[u/v]} \phi\}, \text{ and } W = \{u \in U_M : M \models_{a[u/v]} \psi\}, \tag{7}$$

we have standard semantics for more given by

$$M \models_a More_v(\phi, \psi) \;\text{ iff }\; |V| > |W|, \;\text{ where }\; V, W \;\text{ as in }\; (7). \tag{8}$$

As shown in [10], on the standard semantics the language with *more* is more expressive than that with *most*. On the one hand, $Most_v(\phi, \psi)$ can evidently be expressed in the language of *more* as $More_v(\phi \wedge \psi, \phi \wedge \neg\psi)$. On the other hand, in the language of *more* we can also express the unary quantifier 'there are infinitely many ϕs' viz. as $(\exists y)(\phi(y/v) \wedge \neg More_v(\phi, \phi \wedge v \neq y))$, where y is a variable not occurring in ϕ. This quantifier cannot be expressed in the language of *most* with its standard semantics. (This is something which will not be shown here, but again, see [10].) This relationship between the two languages with *more* and *most* remains true when the standard semantics is replaced the weaker semantics which I will propose below. For although the above definition of 'there are infinitely many' no longer works in that more liberal semantic setting, the definition of *most* in terms of *more* remains valid; on the other hand there is no hope of defining *more* in terms of *most*, for such a definition, if correct, would be correct a fortiori for the standard semantics; but that is something which we just saw is impossible.

[6] See [10]. Colban presents proof theories in the Gentzen sequence calculus format, which I personally find somewhat more difficult to read and handle than the axiomatic approach we will follow.

So the axiomatizations proposed here leave open the question of an intrinsic axiomatization of *most* for the new semantics (i.e. within the language $L(most)$ rather than $L(more)$).[7] From the linguist's point of view, however, this gap is of little importance. For a satisfactory logic for *more* is as important an item on his wish list as one for *most*, and since the first will automatically give us the second, we may as well concentrate on the first.

From now on we will read the quantifier symbol Mo as short for *more* and we proceed with the question how the weak logic of WQL.1-3. may be extended to one which is a credible reflection of our intuitions about the meaning of *more*.

There are two aspects to this problem. The first concerns the behaviour of *more* on the finite sets. Here, as I have been arguing in relation to *most*, the cardinality principle – there are more As than Bs iff the cardinality of the set of As is greater than that of the set of Bs – seems intuitively right. But then, for the finite sets this principle can be fully axiomatized, albeit by an infinite set of axioms. Note that in view of WQL.1 and WQL.2 it is enough to state, for each $n \geq 0$, that for any pair of sets A, B such that B has at most n members and A has $n+1$ members $more(A, B)$ holds and for any pair A, B such that A has at most n members and B has n members $more(A, B)$ does not hold. The axioms WQL.4^n and WQL.5^n express this:

WQL.4^n $(\forall v_1)...(\forall v_n)(\forall v_{n+1})(\forall w_1)...(\forall w_n)$
$(\bigwedge_{i\neq j} v_i \neq v_j \rightarrow Mo_x(\bigvee_i(x = v_i), \bigvee_i(x = w_i))$.

WQL.5^n $(\forall v_1)...(\forall v_n)(\forall w_1)...(\forall w_n)$
$(\bigwedge_{i\neq j} w_i \neq w_j \rightarrow \neg Mo_x(\bigvee_i(x = v_i), \bigvee_i(x = w_i))$.

(In both WQL.4^n and WQL.5^n the variables $v_1, .., v_n, v_{n+1}, w_1, .., w_n, x$ are all distinct.)

The truth of the axioms WQL.4^n and WQL.5^n in a model M for $L(Mo)$ entails that the interpretation $R(U_M)$ of Mo in M has the property that for *any* two finite subsets A, B of U_M, $\langle A, B\rangle \in R(U_M)$ iff $|A| > |B|$.

The second aspect of the problem concerns the infinite sets A. As we have seen, this appears to be a more difficult matter, conceptually as well as formally. I have already expressed my doubts about the strong logic for $L(more)$ which adopts (8) for infinite as well as finite sets. Still, there surely are some principles which ought to hold also in the case where infinite sets are involved. Arguably the most unequivocal one is that when A is infinite and B finite, then '$more(A, B)$' must be true and '$more(B, A)$' must be false. But there are a number of other plausible candidate principles as well. For instance that if '$more(A, B)$' is true, then '$more(B, A)$' must be

[7]I have not looked at the problem of axiomatizing the logic of *most* in its own terms, i.e. in the language $L(most)$.

false, or that when '$more(A, B)$' and '$more(B, C)$' are both true then so is '$more(A, C)$'; or that when $A \subseteq B$, then '$more(A, B)$' cannot be true. Colban has argued for all these principles as part of what governs our intuitions about the meaning of *more* in the infinite as well as the finite domain. He shows that any set relation satisfying these conditions can be represented as the quasi-ordering induced by a *naive measure*, a function ν on $\wp(U_M)$ with the property that its range is some linear ordering $<$ with a smallest element 0 and a largest element ∞ such that $A \subseteq B$ entails $\neg(\nu(B) < \nu(A))$. With respect to such a naive measure '$more(A, B)$' is interpreted as $(\nu(B) < \nu(A))$. Note that the properties of R that are at issue here are second-order properties, as they involve quantification over all subsets of the given set U_M. For instance, transitivity of R takes the form:

$$(\forall X)(\forall Y)(\forall Z)(XRY \wedge YRZ \rightarrow XRZ) \tag{9}$$

where X, Y and Z are second-order variables. The full force of such a sentence cannot be captured within the language $L(Mo)$ as that language only has individual variables. To express (9) we would have to add second-order variables to $L(Mo)$; then (9) could be expressed as

$$\begin{aligned}&(\forall X)(\forall Y)(\forall Z)(Mo_v(v \in X, v \in Y)\\&\wedge Mo_v(v \in Y, v \in Z) \rightarrow Mo_v(v \in X, v \in Z)).\end{aligned} \tag{10}$$

In the 'first-order' language $L(Mo)$ the force of (12) can only be approximated through the infinite set of sentences which we obtain by dropping the initial second-order quantifiers from (12), replacing the atomic subformulae '$v \in X$', '$v \in Y$', '$v \in Z$' uniformly by formulae ϕ, ψ, χ of $L(Mo)$ (and forming universal closures when the resulting formula is not a sentence). The truth of all these sentences in a model M guarantees that the interpretation R_M of the quantifier satisfies the given property (*viz.* transitivity) with respect to the subset of $\wp(U_M)$ consisting of all the $L(Mo)$-definable sets. But there is no guarantee that the property is satisfied 'absolutely', i.e. with regard to all of $\wp(U_M)$. The problem of transforming a model M in which the property is known to hold only relative to definable subsets into an equivalent model M' in which the property holds absolutely is non-trivial and varies with the property in question. But as Colban has shown, it can be solved for the property under consideration, that of being an asymmetric, transitive relation which respects set inclusion (in the sense that if $A \subseteq B$ then not $more(A, B)$). Moreover, the transformation can be carried out in such a way that the first-order reductions of M and M' (i.e. the models obtained by throwing away the interpretations of Mo) are identical and such that the interpretation $R_{M'}$ of Mo in M' coincides with R_M on the set of definable subsets of M.

This means that if we add to weak logic (i.e. to WQL.1-3) all axioms of the forms:

WQL.6 $Mo_v(\phi, \psi) \rightarrow \neg Mo_v(\psi, \phi)$
WQL.7 $Mo_v(\phi, \psi) \wedge \neg Mo_v(\chi, \psi) \rightarrow Mo_v(\phi, \chi)$
WQL.8 $(\forall v)(\phi \rightarrow \psi) \rightarrow \neg Mo_v(\phi, \psi)$

then we obtain an axiom system that is complete with respect to the class of all models M for $L(Mo)$ in which the interpretation R_M of Mo is a relation that is asymmetric and transitive and respects inclusion on all of $\wp(U_M)$. If we include moreover the axioms WQL.4 and WQL.5, then R_M will coincide with the relation $\{\langle A, B\rangle : |A| > |B|\}$ on the finite subsets of U_M. It should also be clear that transitivity and WQL.4 jointly guarantee that $\langle A, B\rangle \in R_M$ whenever A infinite and B finite.

Is this the axiomatization we want? It comes, I think, pretty close. Still, we can, if we want to, pin the interpretation of *more* for infinite domains down further in various ways and strengthen the logic accordingly. One natural strengthening of the logic, to which my attention was drawn by Johan van Benthem, involves the following principle:

> Suppose that '$more(A, B)$' and '$more(C, D)$' and that A and C are disjoint. Then it should also be the case that '$more(A \cup C, B \cup D)$'. (11)

This principle has a very strong intuitive appeal, and we may well want to add the corresponding schema WQL.9′ to our axiomatization.

WQL.9′ $Mo_v(\phi, \psi) \wedge Mo(\eta, \theta) \wedge (\forall y)(\phi(y) \rightarrow \neg\eta(y)) \rightarrow Mo_v(\phi \vee \eta, \psi \vee \theta)$.

It is not as straightforward, however, to modify the given semantics, based on Colban's notion of a naive measure, in such a way that WQL.9 is verified in a natural way. Intuitively, WQL.9 is an additivity principle, and so one might want it to come out valid in virtue of an operation + of 'addition' on the sizes $\nu(A)$ which the naive measure assigns to subsets A of the universe of any model for $L(Mo)$. + ought to have, in particular, the property that when A and C are disjoint, then $\nu(A \cup C) = \nu(A) + \nu(C)$ (in addition, to the usual properties of commutativity, associativity, and monotonicity w.r.t. the order on the range of ν). At present I do not see how to prove completeness for the axiom system WQL.1-9′ with respect to models in which an operation of addition with these properties is defined on the range of ν; though there may well be some way to do this.

Other possible strengthenings have to do with what happens when a finite set is added to an infinite set. For instance, we can add a schema to the effect that if y does not belong to the extension E_ϕ of ϕ, then there are more elements in $E_\phi \cup \{y\}$ than there are in E_ϕ: and, moreover, that when z is another such element, then neither of the sets $E_\phi \cup \{y\}$ and $E_\phi \cup \{z\}$ has more elements than the other:

WQL.9 $(\forall y)(\neg\phi[y/w] \rightarrow Mo_w(\phi \vee w = y, \phi))$

WLQ.10 $(\forall y)(\forall z)(\neg\phi[y/w] \wedge \neg\phi[z/w] \rightarrow \neg Mo_w(\phi \vee w = y, \phi \vee w = z))$.

(Again, to be precise, WQL.9 and WQL.10 represent the sets of all sentences which are obtained by universally closing any formula of the respective forms displayed; it is assumed that y and z are not among the free variables of ϕ.)

That WQL.9 and WQL.10 can be added consistently to WQL.1-8 will be shown in Appendix A. Of course, the circumstance that these axioms can be added consistently is no compelling reason for taking them on board. In fact, while there seems to be nothing that speaks against adopting WQL.10, WQL.9 is very dubious. If perhaps at first sight it looks like a natural generalization of WQL.4, this impression can hardly stand up to scrutiny. It is not so much that the axiom contradicts the cardinality principle adopted by the standard semantics – it would be odd for me to put this forward as a serious objection against it, after my earlier protests that the standard semantics isn't really what we want. More significant, it seems to me, is that WQL.9 is incompatible with any interpretation of *more* in its application to infinite sets that is based on converging frequency on finite samples. For it is quite clear that the limiting frequencies for two infinite sets which differ by one element only must be the same if they exist at all.

Let us be a little more explicit. Suppose that M is a denumerable model for L and that $\mathcal{S}$ is a nest of finite subsets of U_M the union of which equals U_M (we think of $\mathcal{S}$ as the 'sample sequence'). For arbitrary infinite subsets D of U_M we define the *rate of* D *on* $\mathcal{S}$ to be $lim_{S \in \mathcal{S}, |S| \rightarrow \infty} \frac{|D \cap S|}{|S|}$, in case this limit exists, and to be undefined otherwise. Then, if A is an infinite subset of U_M and $B = A \cup \{b\}$ for some element b from U_M that is not in A and the rate of A on $\mathcal{S}$ exists, then the rate of B on $\mathcal{S}$ exists also and is equal to the rate of A. Thus if we interpret 'there are more As than Bs' as true when the rates of A and B on $\mathcal{S}$ both exist and the former is bigger than the latter, then 'there are more As than Bs' will necessarily be false (if it is defined at all) for the sets A and B in question. So WQL.9 could never be true for a ϕ with an infinite extension.

As I have said already, I cannot see anything amiss with WQL.10. Note that WQL.10 is validated both by the standard semantics and by the converging frequency interpretation just sketched. Indeed, WQL.10 seems a natural candidate for a further strengthening of our theory, even if it is not immediately clear how to give a simple and natural characterization of a class of models with respect to which the logic given by WQL.1-8 + WQL.10 would be complete.

This problem, of finding a natural semantics with respect to which the new theory is complete, brings me back to my earlier plea: to investigate additional concepts in terms of which the meanings of quantifiers like *most* and *more* can be given more life-like analyses than is possible with the purely set-theoretical tools to which generalized quantifier theory has for the most part confined itself in the past. Let me, in this connection, return once more to the frequency interpretation. What I have said about this interpretation so far seems to have the draw-back that, for all we know, the frequency limits in terms of which the truth conditions of $Mo(\phi, \psi)$ are given may fail to be defined, so that models in which Mo is given a frequency interpretation will in general be partial. However, so long as the aim of a model-theoretic semantics is that of defining logical validity, partiality is no serious obstacle. One way to circumvent it is to define ϕ to be a logical consequence of Γ iff for every model in which all sentences in Γ are (defined and) true, so is ϕ.

Someone for whom this analysis of the meaning of *most* and *more* has intuitive plausibility, will want an answer to the following question: For any denumerable model M for L let, as above, a *sample sequence* for M be a chain $\mathcal{S}$ of finite subsets of U_M such that $\bigcup \mathcal{S} = U_M$ and call a *frequency model* for $L(Mo)$ any pair $\langle M, \mathcal{S} \rangle$ such that M is a denumerable model for L and $\mathcal{S}$ is a sample sequence for M. If $\mathcal{M} = \langle M, \mathcal{S} \rangle$ is a frequency model, then $Mo(\phi, \psi)$ is true in $\mathcal{M}$ iff either (i) $\{u \in U_M : \mathcal{M} \models \phi[u]\}$ is finite, and $|\{u \in U_M : \mathcal{M} \models \phi[u]\}| > |\{u \in U_M : \mathcal{M} \models \psi[u]\}|$ or (ii) $\{u \in U_M : \mathcal{M} \models \phi[u]\}$ is infinite, the rates of $\{u \in U_M : \mathcal{M} \models \phi[u]\}$ and $\{u \in U_M : \mathcal{M} \models \psi[u]\}$ on $\mathcal{S}$ are both defined and the former is bigger than the latter. For any sentence of $L(Mo)$ and frequency model $\mathcal{M}$ take $\mathcal{M} \models \phi$ to mean that the truthvalue of ϕ in $\mathcal{M}$ is defined and, moreover, ϕ is true in $\mathcal{M}$. Suppose we define the consequence relation for $L(Mo)$ as in (12).

$$\begin{aligned} \Gamma \models \phi \text{ iff for any frequency model } \mathcal{M} \text{ iff} \\ \text{for all } \psi \in \Gamma, \mathcal{M} \models \psi, \text{ then } \mathcal{M} \models \phi. \end{aligned} \tag{12}$$

Question 1: Is this consequence relation axiomatizable? Question 2: If the answer to Question 1 is *yes*, what is a (nice) axiomatization for this relation?

To repeat, it is questions of this general sort to which I believe quantifier theory should increasingly turn its attention.

4 Conclusion

Let me briefly summarize the principal points and concerns of this paper. I began by rehearsing some well-known facts about the quantifier *most*: its essentially binary character, its undefinability in terms of the classical quantifiers 'for all' and 'there is', and the non-axiomatizability of first-order logic extended with *most* on the standard semantics for it (for all A, B '$most(A, B)$' is true iff $|A \cap B| > |A \setminus B|$). I then argued that the condition $|A \cap B| > |A \setminus B|$ is in agreement with our intuitions about the meaning of 'most As are Bs' only in the case where A is finite. So a more realistic semantics is obtained when we adopt this condition only for the finite case, while treating the infinite case in some other way. Since the restriction of the cardinality condition to the finite case can be axiomatized straightforwardly, axiomatizability is now again within our grasp, although whether we get it, and what an axiomatization will be like, if it can be had at all, will of course depend on what the new semantics will stipulate about the infinite case.

How then should the infinite case be treated? On this score my proposals have been incomplete. I have proposed a number of principles (WQL.6-8) to be adopted universally – for the finite case these are entailed by the axioms reflecting the cardinality condition – as a first approximation and mentioned that completeness can be obtained for the resulting system with respect to a semantics based on Colban's notion of naive measure. But clearly that is not the end of the story. I mentioned one further plausible principle (WQL.10) whose addition presents no difficulties (completeness along essentially the same lines can still be obtained as before), as well as another, (WQL.9′), suggested to me by van Benthem, for which a satisfactory semantics plus completeness is still outstanding.

But will these be enough? What *is* enough? That is, I have tried to argue, a difficult question, which is likely to involve much that goes beyond what can be found within the current model-theoretic toolkit of formal quantifier theory. In particular, the familiar arguments against adopting the cardinality condition for the infinite case suggest that our judgements about *most*-sentences with infinite A and B often involve some notion of rate, or frequency. So, I suggested, to make further progress with the question what logic governs the use of *most* with infinite sets, we should explore a semantics based on such a notion. One option, suggested towards the end of Section 4, would be a semantics which deals with the finite cases by way of cardinality and with the infinite ones in terms of frequency. An implementation of that option will have to make a number of further decisions, possibly with diverging consequences for the resulting logic. So this option alone may yield a spectrum of alternative logics, between which it may be difficult to choose. Moreover, it is possible that whichever way we go, we will have to cope with problems quite unlike those that arise for the comparatively simple model theory which has been used here. (One of the contingencies, I observed, with which a frequency-based semantics must be prepared to deal, is partiality: Some *most*-sentences may come out as lacking truth values in some models.)

In addition, frequency need not be the only conception behind our judgements about *most*-sentences involving infinite sets. Careful thought will have to be devoted to the question whether alternative conceptions might come into such judgements and what these might be like. Pursuing this question may well induce us to look into yet other model theories for *most*.

So, a potentially wide field of possible choices, and corresponding axiomatization problems, opens up to those who accept the need of probing further in these directions.

As far as the present paper is concerned, all this has been no more than a plea. In fact, I have only just begun to look into some of these options. But I am resolved to carry on, and I can only hope that I won't be all alone.

Appendix A

We show that WQL.9 and WQL.10 are consistent with WQL.1-8. As a matter of fact we will prove something slightly stronger that the consistency of WQL.1-10, *viz.* that every consistent set Σ of sentences of L is consistent with all instances of WQL.1-10. It follows from this via the completeness theorem for weak logic (see, e.g. [1], or [9]) that there is an $L(Mo)$ model in which Σ and all instances of WQL.1-10 hold. By the methods of [1] this model can then, if one wants, be turned into an equivalent one in which Mo is interpreted by a naive measure.

Let Σ be any consistent theory of L. Let S be a finite set of instances of WQL.9 and WQL.10. Let M be an at most denumerable model of Σ. We show that M can be turned into an $L(Mo)$ model M' in which Mo is interpreted by a naive measure which verifies all sentences WQL^n.4 and WQL^n.5 as well as the sentences in S.[8] For each of the finitely many ϕ which occur in WQL.9 instances or WQL.10 instances in S let E_ϕ be the set of all $u \in U_M$ that satisfy ϕ in M, and let $Umb(\phi)$ be the set $\{E_\phi\} \cup \{E_\phi \cup \{u\} : u \in U_M \setminus E_\phi\}$. We call $Umb(\phi)$ the *umbrella defined* by ϕ (in M) (thinking of E_ϕ as the *handle* of $Umb(\phi)$ and of the sets $E_\phi \cup \{u\}$ as the *spokes* of $Umb(\phi)$). Umb will be the union of the (finitely many) umbrellas $Umb(\phi)$ with ϕ occurring in S. Evidently a naive measure ν will verify all sentences in S iff it assigns the same value to all spokes of any umbrella $Umb(\phi)$ for ϕ occurring in S and assigns a smaller value to the umbrella's handle. Let $\equiv$ be the relation which holds between two subsets A and B of U_M iff their symmetric difference is finite. It is well-known that this is an equivalence relation. Furthermore, for any two sets A and B such that $A \equiv B$ let *the distance from A to B*, $d(A, B)$, be the integer $|A \backslash B| - -|B \backslash A|$. It is not hard to check that if $A \equiv B$, then $d(B, A) = -d(A, B)$ and that for $A \equiv B \equiv C$, $d(A, C) = d(A, B) + d(B, C)$. It is also clear that if A and B both belong to

[8] In case M is finite, we can directly interpret Mo by the relation which holds betweensubsets A and B of U_M iff $|A| > |B|$. This will then be a naive measure satisfying all the schemata WQL.1 – WQL.10. So we could assume at this point that M is denumerably infinite. As this assumption doesn't seem to simplify the proof, I haven't made it. However, it may help to understand the construction below to think of M as infinite and in particular of the 'umbrellas' $Umb(\phi)$ (which will be defined directly) as (for the most part) infinite.

$Umb(\phi)$ for the same ϕ, then $A \equiv B$ and, moreover, that $d(A,B) = 1$ if A is the handle of $Umb(\phi)$ and B one of its spokes; and $d(A,B) = 0$ if both A and B are spokes of $Umb(\phi)$. Also, if $A \in Umb(\phi)$, $B \in Umb(\psi)$ and $A \equiv B$, then for any other $C \in Umb(\phi)$, $D \in Umb(\psi)$, $C \equiv D$. So $\equiv$ collects the umbrellas $Umb(\phi)$ into equivalence classes. Since any equivalence class contains the members of only a finite number of umbrellas (obviously, as there are only finitely many umbrellas that are being considered altogether), it should be clear from what has been said that for each such class C there is a natural number n_C such that for all $A, B \in C$, $|d(A,B)| < n_C$. Also there will be some member $A_0(C)$ of C (not necessarily uniquely determined) such that $d(A_0(C), B) \geq 0$ for all $B \in C$.

Any two distinct equivalence classes C_1, C_2 consisting of (members of) umbrellas can stand in one of three relations; either (i) there are $A \in C_1$ and $B \in C_2$ such that $B \setminus A$ is infinite and $A \setminus B$ is finite, or (ii) there are $A \in C_1$ and $B \in C_2$ such that $B \setminus A$ is finite and $A \setminus B$ is infinite, or (iii) there are $A \in C_1$ and $B \in C_2$ such that both $B \setminus A$ and $A \setminus B$ are infinite. It is easily seen that in case (i) the same relation, $C \setminus D$ infinite and $D \setminus C$ finite, holds for any other $C \in C_1$ and $D \in C_2$, and similarly for cases (ii) and (iii). So, if we define the following relation $\prec$ between equivalence classes: $C_1 \prec C_2$ iff for some $A \in C_1$ and $B \in C_2$ $B \setminus A$ is infinite and $A \setminus B$ is finite, then (a) this definition does not depend on the choice of A and B, and (b) $\prec$ is a strict partial order on the set of equivalence classes. Since $\prec$ is finite, we can assign to each equivalence class C a degree $deg(C)$ by induction: if C has no predecessors in the sense of $\prec$, then $deg(C) = 1$; otherwise $deg(C) = max\{deg(C') : C' \prec C\} + 1$. Now we define a naive measure ν on the power set of U_M as follows:

(i) $\nu(A) = |A|$, if A is finite;
(ii) $\nu(A) = w \cdot deg(C) + d(A_0(C), A)$, if A is infinite and A belongs to the union Umb of the finitely many umbrellas under consideration;
(iii) $\nu(A) = max\{\nu(B) : B \in Umb \wedge B \subseteq A\}$, if A is infinite but not $A \in Umb$.

It is not difficult to verify that ν is indeed a naive measure (the only condition that needs a little care in checking is that $\nu(A) \leq \nu(B)$ whenever $A \subseteq B$) and that when Mo is interpreted in terms of it, then the sentences in Σ all come out true; that the interpretation also verifies WQL.4 and WQL.5 is obvious and that WQL.6-8 are satisfied follows from the results of [1]. ∎

The consistency of WQL.9 and WQL.10 with any first-order extension of WQL.1-8 is only one of an indefinite number of similar results that one may try to obtain. I have presented the argument in the hope that many such results could be established by similar means, though I do not, at the present time, have a clear conception of how far these methods might carry us.

Appendix B

In Section 2 we noted that $L(\textit{more})$ is strictly more expressive than $L(\textit{most})$. As the proof of this fact in [10] makes clear, the reason for this is that the size comparisons involved in the evaluation of *most*-sentences are always between disjoint sets, whereas *more* permits the comparison of arbitrary sets. It is not clear, however, that this difference – *most* has less expressive power than *more* – remains, when we develop a logic of *most* which covers the full spectrum of uses of the word *most* in a language like English. English has sentences in which *most* requires the comparison of sets that overlap.

For instance, with respect to a situation in which a test was taken by Susan, Fred and Naomi we can say

Susan solved most problems on the test. (13)

to mean that the number of problems that Susan solved was larger than the number of problems solved by either of the others. There is no presupposition that the sets of problems each of them solved are pairwise disjoint – for instance, for all that (13) implies, the set of problems solved by Fred might be a proper subset of the set of problems solved by Susan.[9,10]

The comparison class – here {Susan, Fred, Naomi} – can also be made explicit in the sentence itself, as in

As between Susan, Fred and Naomi, Susan solved most problems on the test. (15)

[9]My attention was drawn to this use of *most* by a remark of Ruth Kempson.

[10]In English it seems that the use of *most* in contexts such as (13) is restricted to comparison classes whose cardinality is at least three; if the comparison is between two cases only, the proper word is not *most* but *more*. It is my impression that in certain other languages this constraint is not as strong as it is in English. For instance, I personally do not feel much resistance (if any) against the use of the Dutch equivalent *de meeste* in comparison between two classes. Thus I can say

Susan en Fred hebben allebei genoeg problemen opglost om voor het examen te slagen. Maar Susan heeft de meeste opgelost, en krijgt dus ook het hoogste cijfer.
(Susan and Fred both solved enough problems to pass the exam. But Susan solved more (literally: 'the most') problems and thus gets the better (literally: 'the highest') mark.) (14)

This issue is of some importance for the present discussion insofar as in languages for which the given constraint (i.e. that the comparison class must consist of at least three elements) does not hold, the question of how *more* could be reduced to *most* can be addressed without the slight complication that the constraint produces.

The presence in (15) of the adjunct *as* between Susan, Fred and Naomi, which makes the comparison class explicit rather than leaving it to be recovered from context, renders (15) unambiguous in a sense in which (13) is not. (13) has besides the reading we have just discussed also one which conforms to the analysis of *most* we have been assuming so far – the reading according to which the number of problems Susan solved was more than half the number of problems on the test altogether. As we will see below, the difference between these two readings is, in a certain sense, a matter of scope.

Before we pursue the semantics of sentences such as (13) further, first a brief remark on how this matter affects the question whether *most* is less expressive than *more*. Speaking somewhat loosely, 'there are more As than Bs' can be expressed by a sentence of the form exemplified by (13), provided we can find

(i) a binary relation R that is expressible as a simple or complex transitive verb,
(ii) a set X of three or more individuals, and
(iii) an individual a in X,

such that

(a) the As are the entities y such that a stands in the relation R to y,
(b) for some b in X with $b \neq a$ the Bs are the entities y such that b stands in the relation R to y, while
(c) for every other element c of X, the set of y such that c stands in the relation R to y forms a subset of the set of Bs.

For we can then paraphrase the statement 'there are more As than Bs' by a sentence of the form

As regards the individuals in X, a (is the one who) Rs most things. (16)

(or something in this vein that obeys the rules of English grammar and doesn't offend the English speaker's sensibilities in other ways).

It is not hard to see what it is about English that enables it to express not only those uses of *most* that can be analyzed correctly by treating *most* as a simple generalized quantifier, but also uses of the sort exemplified by (13). Roughly speaking, an NP the determiner of which is *most* can occur in any of the positions in an English clause that are open to NPs generally. Typical examples of the use of *most* which conforms to its analysis as a generalized quantifier are sentences in which the *most*-NP is the subject and in which the VP acts as a 1-place predicate whose only argument position is that subject. Among these sentences there are in particular those in which the VP consists of the copula *be* followed by a nominal or adjectival predicate – sentences such as 'Most trees in Scandinavia are conifers'. or 'Most Americans are white'. Such sentences fit the schematic paraphrase 'Most *A*s are *B*s' almost to perfection. But other sentences with *most*-NPs as subjects – such as, say, 'Most French businessmen smoke' or 'Most American families own a car'. – can, for the purposes of the present investigation, be considered to be of this form too.

Uses of *most* which display the semantic complication we observed in connection with (13) arise when the *most*-NP occurs as argument to a verb or verb phrase which has other arguments as well, and where, moreover, the *most*-NP can be interpreted as being 'within the scope' of one or more of those other NPs. Typical instances of this are clauses with transitive verbs in which the *most*-NP is the direct object; (13) is a case in point. But it is important to note that these are not the only ones. (17), for example,

Most letters were written by Susan to Fred. (17)

can be used to say that within a certain set of author-recipient pairs (containing three pairs or more) the pair Susan-Fred was involved in the writing and receiving of a larger number of letters than were any of the other pairs.

How should these uses of *most* which we have been ignoring hitherto be formally represented? It takes little reflection to see that what is needed is not some generalized quantifier – in the narrow sense of the term, that of an operator which takes two formulas as arguments, produces a formula as output and binds one variable – other than those which we have explored in the body of the present paper. The *most* that concerns us now diverges from the determiners which we have been looking at so far primarily in that it has a very different 'logical grammar'. Take for instance the occurrence of *most* in (15). Its semantic effect is to establish a certain relation between (i) the comparison class {Susan, Fred, Naomi} given by the *as between* phrase; (ii) the individual Susan given by the subject NP; and (iii) the relation 'u solved problem v' given by the VP. This effect is captured in the following clause:

$$\begin{aligned}&\text{(15) is true iff } (\forall u)(u \in \{Susan, Fred, Naomi\} \wedge \\ &u \neq Susan \rightarrow \text{Susan solved more problems than } u).\end{aligned} \tag{18}$$

If we insist on capturing this semantic relationship while treating *most* as a variable binding operator, the apparent type of this operator is that of one which (a) takes as input one term and two formulas, (corresponding to the subject, the *as between* phrase and the VP, respectively, in (15)); and (b) binds two variables, the first of which represents the relevant member of the comparison class and the subject argument of the VP, while the second represents the object argument of the VP. Thus (15) gets the logical form

$$Most^2_{u,v}(\tau, \rho(u), \chi(u,v)), \tag{19}$$

where t is the term 'Susan', $\rho(u)$ is short for for all '$u \in$ {Susan, Fred, Naomi}' and $\chi(u,v)$ for 'u solved problem v on the test'. The truth conditions of (19) are given in (20)

$$\begin{aligned}&Most^2_{u,v}(\tau, \rho(u), \chi(u,v)) \text{ is true iff}\\&(\forall u)(\rho(u) \wedge u \neq \tau \rightarrow MORE(\{v : \chi(\tau,v)\}, \{v : \chi(u,v)\}))\end{aligned} \tag{20}$$

where MORE is the generalized quantifier (i.e. relation between sets) expressed by *more*; or, alternatively, using the generalized quantifier *Mo* (with interpretation MORE) which we have investigated in Section 3:

$$\begin{aligned}&Most^2_{u,v}(\tau, \rho(u), \chi(u,v)) \text{ is true iff}\\&(\forall u)(\rho(u) \wedge u \neq t \rightarrow Mo_v(\chi(\tau,v), \chi(u,v))\end{aligned} \tag{21}$$

As (21) shows, $Most^2$ is definable in terms of the old Mo. Can we define, conversely, Mo in terms of $Most^2$? Almost. All we need is an antecedent assumption that there are enough things to form at least one proper comparison class; if we stick to the intuitions I mentioned about the use of *most* in sentences like (13) in English, this means that the universe must contain at least three things. So let us assume that there are three distinct objects x, y and z. Consider the formula $Mo_v(\phi(v), \psi(v))$. Let $\chi(u,v)$ be the formula $(u = x \wedge \phi(u)) \vee (u = y \wedge \psi(u)) \vee (u = z \wedge \psi(u))$ and let $\rho(u)$ be the formula $u = x \vee u = y \vee u = z$. Then $Mo_v(\phi(v), \psi(v))$ is clearly equivalent to $Most^2_{u,v}(x, \rho(u), \chi(u,v))$. Thus we have the following conditional definition of Mo in terms of $Most^2$:

$$\begin{aligned}&(\exists x)(\exists y)(\exists z)(x \neq y \wedge x \neq z \wedge y \neq z) \rightarrow (Mo_v(\phi(v), \psi(v)) \leftrightarrow\\&(\exists x)(\exists y)(\exists z)(x \neq y \wedge x \neq z \wedge y \neq z \wedge Most^2_{u,v}(x, \rho(u), \chi(u,v))))).\end{aligned} \tag{22}$$

Since the operator $Most^2$ is definable in terms of Mo, its introduction does not introduce any fundamentally new axiomatization problems. One could still pose the question whether there is a direct, natural and elegant axiomatization for the new $Most^2$. This is a question that I have not explored.

The operator $Most^2$ we have just been discussing arose out of a reflection on the meaning of (15). The need to formalize (15) by means of an operator which binds not one but two variables, one variable for the problem solved and one for the one who solved it, arose from the circumstance that the different sets of solved problems which the sentence asks us to compare depend on who in each case is the solver. By analogy, formalization of a sentence like (17) will require an operator binding three variables, one variable for the letter written, one for the person who wrote it and one for the person to whom it was written. The comparison class is now, as we have seen, a set of pairs; in the setting of variable binding this comes down to a two-free-variable-formula $\rho(u,w)$. And instead of the binary relation expressed by the transitive verb 'solved' in (15) we now have the ternary relation expressed by 'u wrote v to w'; in terms of the operator treatment this amounts to a formula $\xi(u,v,w)$ with free variables u, v and w. These considerations suggest an operator $Most^3$ which binds 3 variables and takes as inputs two formulas (the ρ and ξ just mentioned) as well as two terms – in (17) these are given by the subject and the *to*-PP. Using such an operator, (17) can be represented as.

$$Most^3_{u,v,w}(\tau, \sigma, \rho(u,w), \xi(u,v,w)), \tag{23}$$

where τ is the term 'Susan', σ is the term 'Fred', $\rho(u,w)$ is short for '$\langle u,w\rangle \in C$' with C the relevant class of pairs that acts as comparison class, and $\xi(u,v,w)$ for 'u wrote letter v to w'.

I take it that the meaning of (17) is correctly captured by the following truth clause for $Most^3$:

$$\begin{array}{l} Most^3_{u,v,w}(\tau, \sigma, \rho(u,w), \xi(u,v,w)) \quad \text{iff} \\ (\forall u)(\forall w)((\rho(u,w) \wedge (u \neq t \vee w \neq \sigma)) \rightarrow Mo_v(\xi(\tau,v,\sigma), \xi(u,v,w)). \end{array} \tag{24}$$

Thus $Most^3$ is, just like $Most^2$, definable in terms of Mo.
Of course this is not the end of it. Formalization of a sentence such as

$$\text{Most letters were written by Susan from Ithaca to Fred.} \tag{25}$$

which may report on a comparison of the number of letters which Susan wrote to Fred from Ithaca with the number of letters which Carla wrote to Algie from Corfu, the number of letters that Carla wrote to Fred from Corfu, the number of letters that Susan wrote to Fred from Athens, etc., would require for its formalization an operator binding four variables; and so forth. Operators binding even more variables would be needed to represent sentences in which the sets defined by the *most*-NP depend on four, five,... other arguments (obligatory or optional) to the main verb. Thus, the number of operators needed to formalize arbitrary sentences of this pattern will be finite only if there is an upper bound to the number of optional arguments to any given verb that can be incorporated into a single clause. Those who feel that such an upper bound would, even if it could be argued to exist, testify to an idiosyncrasy of natural language grammar to which the design of logical representation formalisms should be pay no attention, may want to adopt the entire infinite sequence of operators in any case.

From a logical perspective there exists an obvious alternative. Semantically, each of these infinitely many operators is definable in terms of Mo. So the language $L(Mo)$ is all we need in order to capture the truth conditions of any of the sentences that can be represented in the language $L(\{Most^n\}_{n \in N})$. But to what extent is this alternative acceptable linguistically? What the linguist wants is not just a formalism in which the truth conditions of natural language sentences can be stated accurately; he also wants a systematic procedure that gets him, for any one of the sentences of his concern, to a statement of its truth conditions while starting from its syntactic form – a procedure which somehow 'explains' why a sentence of *this* syntactic form has *this* meaning. I find it hard to see, however, how it might be possible to define a systematic transition from syntactic to logical representation for the sentences in question which did not pass via a representation that involves in some form or other the relevant operator $Most^n$.

But this is a matter that will have to be explored in another context than this.

Acknowledgements

Many thanks to Johan van Benthem for a number of important comments and suggestions. Unfortunately it was not possible for me to deal with his criticisms in the way they deserved. All I have been able to do here is to add a few last minute adjustments, but I hope to make better use of his observations in further projected work on the logic of non-standard quantifiers. Many thanks also to Uwe Reyle, whose help in getting this paper into a form suitable for appearance in this volume much exceeded what an author may reasonably expect from an editor.

Universität Stuttgart, Germany.

References

1. Erik A. Colban. Generalized quantifiers in sequent calculus. COSMOS-Report 18, Department of Mathematics, University of Oslo, March 1991.
2. Tim Fernando and Hans Kamp. 'Most', 'more', 'many', and more. SFB-Report, IMS, University of Stuttgart, 1996.
3. Kit Fine. Vagueness, truth and logic. *Synthese*, 30:265–300, 1975.
4. Hans Kamp. Two theories about adjectives. In E. Keenan, editor, *Formal Semantics of Natural Language*. Cambridge University Press, 1975.
5. Hans Kamp. Conditionals in DR theory. In J. Hoepelman, editor, *Representation and Reasoning*. Niemeyer, 1986.
6. Hans Kamp and Barbara Partee. Prototype theory and compositionality. *Cognition*, 57:129–191, 1995.
7. Ed Keenan. Beyond the Frege boundary. *Linguistics and Philosophy*, 15, 1992.
8. Ed Keenan. Natural language, sortal reducibility and generalized quantifiers. *Journal of Symbolic Logic*, 58(1), 1993.
9. H. Jerome Keisler. Logic with the quantifier 'there exists uncountably many'. *Annals of Mathematical Logic*, 1:1–93, 1970.
10. Dag Westerstahl. Quantifiers in formal and natural languages. In D. M. Gabbay and F. Guenther, editors, *Handbook of Philosophical Logic*, volume IV, chapter 1, pages 1–131. Reidel, 1989.