JON BARWISE[2]

ON BRANCHING QUANTIFIERS IN ENGLISH[1]

One of Hintikka's aims, in the paper Hintikka (1974),[3] was to show that there are simple sentences of English which contain essential uses of branching quantification. If he is correct, it is a discovery with significant implications for linguistics, for the philosophy of natural language, and perhaps even for mathematical logic.

Philosophically, it would influence our views of the ontological commitment inherent in specific natural language constructions, since branching quantification is a way of hiding quantification over various kinds of abstract abstract objects (functions from individuals to individuals, sets of individuals, etc.). Linguistically, the discovery of branching quantification would force us to re-examine, and perhaps re-interpret, Frege's principle of compositionality according to which the meaning of a given expression is determined by the meanings of its constituent phrases. For example, the meaning of a branching quantifier expression of logic like:

$$\begin{matrix} \forall x - \exists y \\ \forall z - \exists w \end{matrix} > A(x, y, z, w)$$

cannot be defined inductively in terms of simpler formulas, by explaining away one quantifier at a time.[4] Rather, the whole block

$$\begin{matrix} \forall x - \exists y \\ \forall z - \exists w \end{matrix} >$$

must be treated at once. This has obvious consequences for any attempt to capture the relation between the syntax and semantics of English sentences in which branching quantification occurs. We will return to the possible implications for mathematical logic at the end of the Appendix.

It is not surprising – given its consequences – that Hintikka's claim has sparked a lively controversy. We refer the reader to Fauconnier (1975), Stenius (1976), the reply to Stenius in Hintikka (1976), as well as the related papers by Gabbay and Moravcsik (1974) and a reply to that in

Journal of Philosophical Logic 8 (1979) 47–80. 0022–3611/79/0081–0047 $03.40

Guenthner and Hoepelman (1974). It seems to us, judging from this literature and from conversations, that many linguists and philosophers are not convinced by Hintikka's arguments in this particular matter. Our principal aim in this paper is to present some additional evidence that Hintikka is correct that branching quantification does occur naturally in English. Furthermore, we will show that branching quantifiers provide a useful tool for the analysis of conjoined noun phrases.

We were led to our examples of branching quantification by trying to understand why Hintikka's own examples seem either confusing or ambiguous. Hintikka was dealing only with English quantifiers that correspond to the logician's $\forall$ and $\exists$. These quantifiers have some unusual properties not shared by most quantifiers, properties which make it difficult to give simple, unambiguous English examples of essential branching quantification using them alone. These properties are discussed in Section 1. We circumvent the difficulties in Section 2 by allowing more general quantifiers, so called 'monotone' quantifiers. More technical matters are discussed in an Appendix.

1. ON THE BRANCHING OF $\forall$ AND $\exists$

To review what a branching quantifier is, recall that Hintikka's starting point is a *first-order* language, first-order in the sense that "only individuals (members of D [the domain of discourse]) are being quantified over", not arbitrary sets of individuals, functions from individuals to individuals or other abstract objects outside the domain D. This language is then strengthened to a language Hintikka calls FPO quantification theory (following a suggestion made in Henkin (1959)) which allows expressions of the form (a) or, more generally (b), below.

(a) $$\begin{matrix} \forall x - \exists y \\ \forall z - \exists w \end{matrix} \Big> A(x, y, z, w)$$

(b) $$\begin{matrix} \forall x_1 - \forall y_1 - \exists z_1 \\ \forall x_2 - \forall y_2 - \exists z_2 \\ \vdots \\ \forall x_n - \forall y_n - \exists z_n \end{matrix} \Bigg> B(x_1, \ldots, z_n).$$

Here the matrix (A or B) is to be some first-order expression. The meaning of (a) is expressed *approximately* by:

> For every x and z there are y and w, where y depends only on x, and w depends only on z, such that $A(x, y, z, w)$.

To eliminate the vague phrases of the form 'y depends only on x', the actual definition of truth for sentences of the form (a) is given by:

(a$'$) There exist functions f and g mapping all of D into D such that for every x and z, $A(x, f(x), z, g(z))$.

Thus, while (a) appears, syntactically, to quantify over individuals, semantically it actually contains hidden existential quantifiers over the set of all functions from D into D. The functions f and g in (a$'$) are called the *Skolem functions* for the formula (a).

This use of quantification over the set of all functions from D into D in (a$'$) is essential in explaining the meaning of FPO sentences. In fact, Walkoe (1970) and Enderton (1970) both show that any second-order expression of the form

> There exist functions $f_1, \ldots, f_m$ such that $A(f_1, \ldots, f_m)$,

where A is first-order, can be written as a sentence of FPO, and there are many known examples of this kind of second-order sentence which are not expressible in first-order logic.

In Hintikka's original article, he attempts to show by examples that "the structure of every FPO sentence is reproducible in English". Here are three sentences he claims to have form (a) and not to be equivalent to first-order sentences.

(1) Some relative of each villager and some relative of each townsman hate each other.

(2) Some book by every author is referred to in some essay by every critic.

(3) Every writer likes a book of his almost as much as every critic dislikes some book he has reviewed.

In our discussion below we choose to concentrate on (1) since it is less confusing than (3) and the arguments for its being of form (a) are better than those for (2).

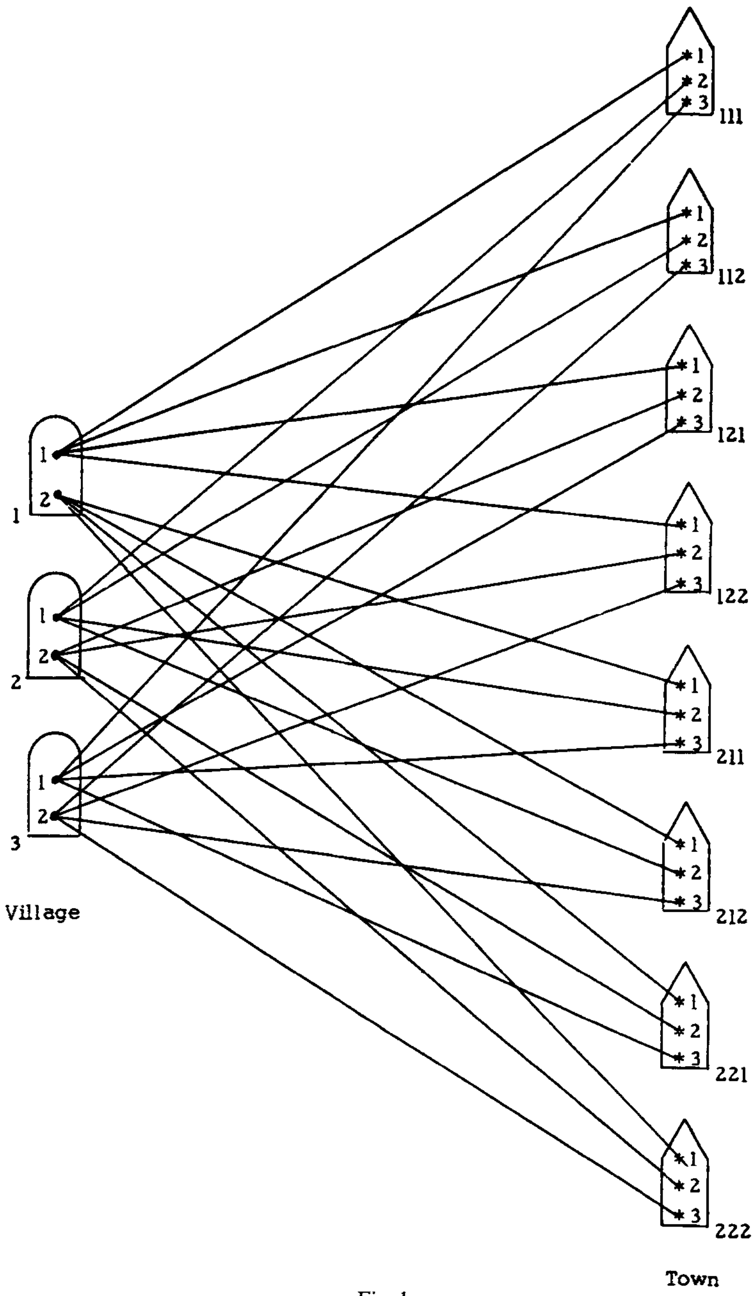

Fig. 1

Before discussing (1), though, we want the reader to make up his own mind as to its meaning, by deciding whether it is true or false in the following model of a village and town (Figure 1). (We will use this model for a proof in the Appendix.)

This model is defined as follows. The village consists of three clans, each with only two members. The town consists of eight clans, each with three members. People in the same clan live together and are related. (In particular, each person is considered related to himself.) People in different clans are not related.

Relations between the town and village are in a terrible state. *Every* villager hates *every* townsman. *Every* townsman hates *every* villager *except one*, the one to whom he is connected by a line in our picture. Thus, of the 144 (= 6 x 24) pairs of villagers and townsmen, 120 hate each other, 24 do not hate each other.

So, question: Is it or is it not the case that some relative of each villager and some relative of each townsman hate each other? In other words, is it or is it not the case that some dot in each hut and some star in each house are not connected by a line?

In our experience, there is almost universal agreement that some dot in each hut and some star in each house are not connected by a line. (In fact, most people agree that at least half the dots in each hut and at least 2/3 the stars in each house are not connected by a line; see example (24) below.) The reader who agrees with this is rejecting Hintikka's claim for a branching reading of (1). The branching reading asserts that we can choose one dot out of each hut, once and for all, and one star out of each house, again once and for all, so that *none* of the three dots and eight stars so chosen are connected by lines. This is impossible. For example, if you choose dots 1, 2, 1 out of huts 1, 2, 3 respectively, then the star chosen out of the house with address 121 will be connected by a line to one of the dots.

Now let us return and review the evidence for a branching reading of (1). Using the obvious notation, the phrase 'some relative of each villager' should have a logical form of

$$\forall x [Vx \supset \exists y (Rxy \; \& \; \ldots)]$$

or, equivalently

(c) $$\forall x \exists y [Vx \supset (Rxy \; \& \; \ldots)].$$

Similarly, the phrase 'some relative of each townsman' should have a logical form

(d) $\forall z \exists w[Tz \supset (Rzw \mathbin{\&} \ldots)]$.

The question is, how should (c) and (d) be put together? Can they be put together in a way that respects the syntactic structure of the sentence (the two 'some of each' phrases)? The sentence certainly cannot be expressed by:

(e) $\forall x \exists y \forall z \exists w[(Vx \mathbin{\&} Tz) \supset (Rxy \mathbin{\&} Rzw \mathbin{\&} Hzw)]$

for that makes the choice of the villager's relative y depend only on the villager x, while the townsman's relative w would be chosen to depend on both the villager x and the townsman z. This can't be correct because, as Hintikka points out, (1) clearly has the same meaning as

(1′) Some relative of each townsman and some relative of each villager hate each other.

Thus, the two quantifier prefixes should be treated as on a par. Neither one can go on first, so we must have the branching reading

(f) $$\left.\begin{matrix}\forall x - \exists y \\ \forall z - \exists w\end{matrix}\right\rangle A(x, y, z, w)$$

where A is the part of (e) in brackets. But this is the reading which is almost universally rejected in the above model. Something has gone wrong – namely namely, the definition of FPO.

There is an expression not in FPO which treats the two prefixes as being on a par, namely:

(g) $$\left.\begin{matrix}\forall x - \exists y \\ \forall z - \exists w\end{matrix}\right\rangle A(x, y, z, w).$$

The lines indicate that y depends on both x and z and that w also depends on both x and z, which is just what the intuition behind the common response to the above model seems to be. This expression is equivalent to various first-order expression, e.g.,

(h) $\forall x \forall z \exists y \exists w A(x, y, z, w)$

and

$$\forall z \forall x \exists w \exists y A(x, y, z, w).$$

These are logically equivalent to (g), but they violate the syntax of (1) and (1′) by shuffling the two 'some of each' phrases together in an *ad hoc* manner. Thus, it seems that the preferred reading of (1) is expressed by (g), which is a form of branching quantification, but an inessential use of branching quantification, since it is equivalent to various first-order expressions.

Our expression (g) is not in Hintikka's language FPO, even though it is equivalent to an FPO sentence, in fact, to a first-order sentence, which brings us to an important point. Henkin, in his paper Henkin (1959), considered such quantifiers. In fact, his proposal was to add arbitrary finite partially ordered strings of quantifiers to the front of first-order formulas. Walkoe (1970) showed that, by certain logically valid quantifier manipulations, Henkin's quantifiers could all be expressed, up to logical equivalence, in the form (b) above. This is Hintikka's justification for using this very special form in his definition of the language FPO.

TABLE I

Branching quantifier expression	*Some equivalent linear expressions*
$\left.\begin{matrix}\forall x\\ \forall y\end{matrix}\right> A(x,y)$	$\forall x \forall y A(x,y)$, $\forall y \forall x A(x,y)$
$\left.\begin{matrix}\exists x\\ \exists y\end{matrix}\right> A(x,y)$	$\exists x \exists y A(x,y)$, $\exists y \exists x A(x,y)$
$\left.\begin{matrix}\forall x\\ \exists y\end{matrix}\right> A(x,y)$	$\exists y \forall x A(x,y))$
$\left.\begin{matrix}\forall x - \exists y\\ \forall z - \exists w\end{matrix}\right> A(x,y,z,w)$	$\forall x \exists y \forall z \exists w A(x,y,z,w)$
$\left.\begin{matrix}\forall x - \exists y\\ \forall z - \exists w\end{matrix}\right> A(x,y,z,w)$	$\forall x \forall z \exists y \exists w A(x,y,z,w)$ and others

A quantifier manipulation may be logically valid without being lingusticially natural. Table I shows a list of some logically valid quantifier manipulations.

It seems that equivalences such as the above often make what would naturally be thought of as a branching quantifier logically equivalent to a first-order formula, and hence what we would call an *inessential* use of branching quantification. To get *essential* uses of branching quantification, we should look at other quantifiers where the manipulations listed above are not valid. We take up this idea in the next section.

Before leaving this section, though, we must take up some unfinished business, by examining Hintikka's remaining argument for giving (1) the reading (f), which involves an essential use of branching quantification.

Stenius,[5] in his attack on Hintikka's article, claims that the formula listed as (h) above represents the intended reading of (1), though without the intermediate step through the logically equivalent (g). To Stenius' claim, Hintikka responded as follows:

> Even apart from Stenius' authoritarian appeal to *his* intuitions as the final arbiter of English semantics, this claim of his is simply wrong. Perhaps the quickest way of seeing that it is (*sic*) to point out that by his token the sentence
>
> (4) Some relative of each townsman and every villager hate each other
>
> must likewise have the logical form
>
> (j) $\forall x \forall z \exists y[(Tx \,\&\, Vy) \supset (Rxy \,\&\, Hyz)]$
>
> whereas it plainly has the force of
>
> (k) $\forall x \exists y \forall z[(Tx \,\&\, Vy) \supset (Rxy \,\&\, Hyz)]$.[6]

If we are to maintain that (g) represents *a* reading of (1), let alone the preferred reading, then we must confront this argument.

We would propose to express (4) by:

$$\left.\begin{array}{l}\forall x - \exists y\\ \qquad \forall z\end{array}\right\rangle [(Tx \,\&\, Vz) \supset (Rxy \,\&\, Hxy)]$$

This is equivalent to (k), whereas (j) is equivalent to

(ℓ) $$\begin{array}{l}\forall x - \exists y\\ \forall z \diagup\end{array} [(Tx \,\&\, Vz) \supset (Rxy \,\&\, Hxz)].$$

So the question is, are we forced into treating (4) by (ℓ) if we treat (1) by (g)? That is, must we give 'every villager' wide scope in (4) if we give both 'each villager' and 'each townsman' wide scope in (1)? The clue lies in the 'each' and 'every'.

Different quantifiers, even *logically* equivalent quantifiers, exhibit different tendencies toward wide scope readings.[7] In particular, 'each' has a greater tendency toward wide scope than 'every', as the following examples illustrate.

(5.1) I have seen a picture of each child.

(5.2) I have seen a picture of every child.

(6.1) Some representative visited each city.

(6.2) Some representative visited every city.

These sentences may be ambiguous, but the first one in each group seems to have $\forall x \exists y(\ldots)$ as its preferred reading, whereas the second has $\exists y \forall x(\ldots)$. This difference between 'each' and 'every' is the reason (4) has the force of (k), where 'each' is given wide scope and 'every' is given narrow scope. Also, note that in reading (1) as (g), we are giving 'each' widest possible scope (over both 'some' s), whereas (f) gives each 'each' scope only over one 'some'. This may in part account for the preferred status of (g) over (f) as a reading of (1).

As further proof that it is this scope question that is the flaw in Hintikka's argument, notice how peculiar his sentence (4) becomes if we change 'every' back to 'each'.

(7) ?Some relative of each townsman and each villager hate each other.

The difference between 'each' and 'every' can also be illustrated by noticing that (1) sounds peculiar if we replace 'each' by 'every'. In fact, there is a reading of (1) that is being blocked by the use of 'each', and by the meaning of 'relative'. Observe what happens if we modify (1) as follows:

(8) Some ancestor of every living villager and some ancestor of every living townsman hated each other.

One of the sensible readings of (8), perhaps the preferred reading, is

$$\begin{matrix} \exists y - \forall x \\ \exists w - \forall z \end{matrix} > [(Vx \,\&\, Tz) \supset (Ayz \,\&\, Awz \,\&\, Hyz)]\,.$$

Of course this is equivalent to various first-order sentences ($\exists y \exists w \forall x \forall z[\ldots]$ and $\exists w \exists y \forall x \forall z[\ldots]$ for example) but again these first-order equivalents are linguistically unnatural since they violate the syntax of the English sentence by shuffling the two 'some of every' phrases together.

For the rest of this paper, FPO will denote the language which allows the more general branching quantifiers discussed above.[8] Up to logical equivalence it is the same as Hintikka's FPO. A simple point that has not been mentioned in the linguistical literature is that the sentences of FPO are not closed under the operations of propositional logic,[9] in particular negation or implication. Notice, for example, that the negation of (a) is expressed in second-order logic by

> For all functions f and g mapping D into D there are individuals x and z so that not $A(x, f(x), z, g(z))$.

Using this and the Craig Interpolation Theorem for first-order logic it is routine to prove:

PROPOSITION 1. If φ is a sentence of FPO and if its negation $\sim\varphi$ is logically equivalent to an FPO sentence, then φ is logically equivalent to some first-order sentence.[10]

This result can be used to give us a reasonable test for whether a given sentence of English is a genuine, unambiguous example of essential branching quantification. To state the test, we define a sentence to be *negation normal* (with respect to subject position) if no quantifier in subject position occurs within the scope of a negation. In a formal language, this definition (without the subject position restriction) has often been used in technical work in model theory. In English it is less precise. We are using it so that the sentences in (9) are not negation normal, whereas those in (10) are negation normal.[11]

(9) It is not the case that everyone owns a car.
It's not true that everyone brought something to the party.
It is not true that someone left early.

(10) Someone doesn't own a car.
Someone didn't bring anything to the party.
No one left early.

There is a clear preference for negation normal sentences in most situations. In fact, the form *It is not true that S* is seldom used unless someone else has just asserted *S*. Followers of Montague may also note that all sentences in the PTQ fragment are negation normal, due to the way his rules of sign operate.

Using the notion of negation normal sentence, we can express our test.

TEST FOR ESSENTIAL BRANCHING QUANTIFICATION: If a sentence **S** of English is an unambiguous example of essential branching quantification, then its negation **NOT-S** should be rather puzzling. For example, there is no way to paraphrase **NOT-S** as a negation normal sentence without in some way using a universal quantifier over abstract objects – functions, sets, 'ways', 'assignments', 'choices', etc. Thus, in particular, **NOT-S** cannot be expressed as a negation normal sentence with branching quantification.

Let us apply this test to Hintikka's villager-townsman example, our (1).

(1) Some relative of each villager and some relative of each townsman hate each other.

Recall from above that Hintikka claims the following as its logical form:

$$\text{(f)}\qquad \begin{matrix} \forall x - \exists y \\ \forall z - \exists w \end{matrix} > [(Vx \,\&\, Tz) \supset (Rxy \,\&\, Rzw \,\&\, Hyw)]\,.$$

Since one can show that this expression is not equivalent to any first-order sentence,[12] it follows that Hintikka is claiming that (1) contains what we are calling an essential use of branching quantification. To test this, we form **NOT-**(1) and ask the reader to paraphrase it with a negation normal sentence.

NOT-(1) It is not true that some relative of each villager and some relative of each townsman hate each other.

It might be less confusing to paraphrase (11), which has the same logical form.

(11) It is not true that some dot in each hut and some star in each house are not connected by a line.

After writing out his own paraphrases, the reader is asked to check to see which (if either) of the following it is equivalent to.

(12.1) There is a villager and a townsman that have no relatives that hate each other.

(12.2) Any way of assigning relatives to each villager and to each townsman will result in some villager and some townsman being being assigned relatives that do not hate each other.

(13.1) Every dot in some hut and every star in some house are connected by a line.

(13.2) Any way of selecting a dot from each hut and a star from each house will result in some selected dot and star being connected by a line.

Again, in our experience, there is almost universal preference for (12.1) and (13.1), whereas (12.2) and (13.2) correspond to the negations of Hintikka's reading (f). We leave it to the reader to apply the test to examples (2) and (3). In our opinion, (2) fails the test but (3) passes. But if (3) is confusing, NOT-(3) is doubly so.

The better a paper on branching quantification is, the more convincing is some example it contains. Our most convincing examples are in the next section, but we do have one example which uses only $\forall$ and $\exists$ which we cannot resist giving.

One of the simplest examples from mathematics of a sentence which cannot be expressed in first-order logic but can be expressed in FPO is:

(14) There is some way to embed the ordering $(L, <)$ into the ordering $(S, \prec)$.

In second-order logic this would be written

(m) $\exists f \forall x \forall y [B(x, y, f(x), f(y))]$

where $B(x, y, z, w)$ is

$$[(Lx \,\&\, Ly \,\&\, x < y) \supset (Sz \,\&\, Sw \,\&\, z \prec w)].$$

Using an FPO sentence, it would be written

$$\text{(n)} \qquad \begin{matrix} \forall x - \exists z \\ \forall y - \exists w \end{matrix} > [B(x, y, z, w) \,\&\, (x = y \supset z = w)].$$

An application of the Löwenheim–Skolem Theorem shows that these equivalent formulas are not expressible by a first-order sentence, or even by an infinite set of such.[13]

There is a special construction in English (and German, and probably other languages) that allows us to express similar things.[14]

(15) The bigger they are, the harder they fall.

(16) The richer the country, the more powerful its ruler.

However, these are definitely not examples of branching quantification. For example, (16) has the logical form

$$\forall x \forall y [(Cx \,\&\, Cy \,\&\, x \text{ richer-than } y) \\ \supset (\text{ruler}(x) \text{ more-powerful-than ruler}(y))].$$

The trouble, from our point of view, is that (16) has a built in Skolem function 'ruler of'. To escape this Skolem function, and to allow for countries with a puppet ruler (say a figurehead Queen), we can try some manoeuver like that in (17).

(17) The richer the country, the more powerful one of its officials.

This is our second candidate for a sentence which is not first-order but which can be analyzed as a FPO sentence. Our first example appeared in the body of the article two paragraphs back.

There is a sentence that is closely related to (17) but which is first-order.

(18) The richer the country, the more powerful is its most powerful official.

Sentence (17) does not assert that it is the most powerful official which accurately reflects the wealth of his country, but rather that there is some way of choosing officials of each country so that the wealth of the country is reflected in the power of the selected official. Thus the two are not equivalent.

Reaction to sentence (17) is mixed. Some people just don't understand it.

Typically they ask something like "Which official?" Thus, they are searching for a Skolem function, as you would expect. Also, the sentence definitely passes the negation test. It is very hard for people to come up with a negation normal statement of NOT-(17), and when they do, it usually begins with "There is no way . . . ". Thus, the puzzling features of (17) are just what you might expect of an essential use of branching quantification.

There is a separate issue surrounding (17). If one grants that it makes sense, then it can be analyzed by either the branching expression (n) or the equivalent second-order expression (m). Which is the more natural and useful analysis? Neither one is very faithful to the syntax of the sentence.

One argument in favor of the branching analysis is that it explains the difficulty in forming the negation of (17). It seems that there are the same sorts of difficulties negating (17) as one would expect of an essential use of branching quantification. We will present an argument in favor of the second-order analysis in the Appendix. The issue will only be settled by a convincing linguistic treatment of sentences like (15), (16) and (17) in a more general setting. But, whichever turns out to be the case, the philosophical and linguistic consequences mentioned in the introduction would still follow.

2. THE BRANCHING OF OTHER QUANTIFIERS

In this section we show that there are many examples of branching quantification when one allows quantifiers like 'many', 'most', and 'quite a few' in addition to 'every' and 'some'. In particular, we will show that the following exemplify branching quantification.

(19) Most relatives of each villager and most relatives of each townsman hate each other.

(20) Few relatives of each villager and few relatives of each townsman hate each other.

(21) Most philosophers and most linguists agree with each other about branching quantification.

(22) Quite a few boys in my class and most girls in your class have all dated each other.[15]

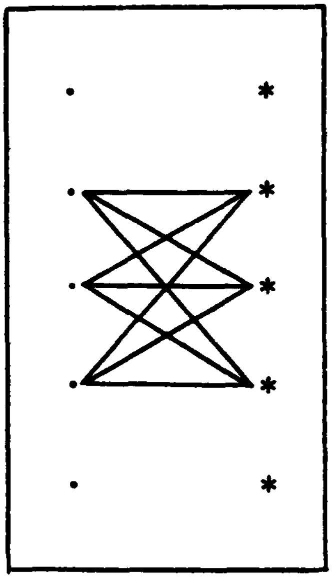

Fig. 2

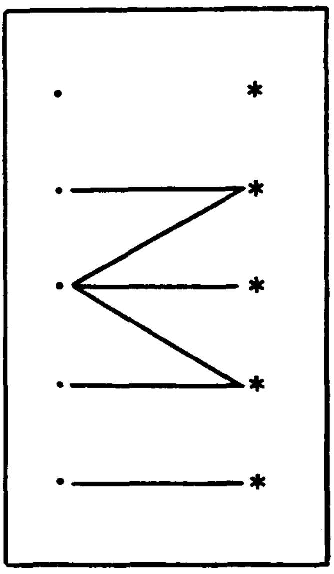

Fig. 3

(23) Most of the dots and most of the stars are all connected by lines.

It is not any supposed vagueness in the quantifiers that accounts for the branching we claim exists, as the following versions of (19) and (23) show.

(24) At least half the relatives of each villager and at least 2/3 the relatives of each townsman hate each other.

(25) More than half the dots and more than half the stars are all connected by lines.

As most people read (25), it is true of Figure 2 but false of Figure 3, since in Figure 3 there is only one dot connected to more than half the stars. As for (24), most would agree that it holds in Figure 1, as we mentioned.

It is beyond the scope of this paper to go into a detailed treatment of how we propose to analyze the model theory of these generalized quantifiers. We are at work on a technical paper on such matters. The reader will get some hints from the discussion that follows together with Section 2 of Barwise (1978). What we attempt here is to explain the basic facts at work that make the above examples of essential branching quantification.

A quantifier Q is *monotone increasing* if for all predicates A, B

$$QxA(x) \ \& \ \forall x[A(x) \supset B(x)]$$

implies

$$QxB(x).$$

Both $\forall$ and $\exists$ are monotone increasing, as are several of the above. On the other hand, Q is *monotone decreasing* if for all A, B

$$QxB(x) \ \& \ \forall x[A(x) \supset B(x)]$$

implies

$$QxA(x).$$

The study of monotone quantifiers has become important in model theory and in recursion theory in recent years. The semantics of monotone quantifiers is handled by assigning to each such Q a collection Q_D of subse of the domain D of discourse. If Q is monotone increasing then Q_D must have the property that $X \in Q_D$ and $X \subseteq Y \subseteq D$ implies $Y \in Q_D$. For example $\exists_D = \{X \subseteq D \mid X \text{ nonempty}\}$ and $\forall_D = \{D\}$. A sentence $QxA(x)$ is said to be true[16] in D if $\{x \mid A(x) \text{ is true in } D\} \in Q_D$.

Note that (as pointed out in Aczel (1975)) if Q is monotone *increasing* then

(p) $$Qx\varphi(x)$$

is logically equivalent to a second-order expression

(q) $$\exists X[Qx(x \in X) \mathbin{\&} \forall x(x \in X \supset \varphi(x))].$$

On the other hand, if Q is monotone decreasing then (p) is equivalent to

(r) $$\exists X[Qx(x \in X) \mathbin{\&} \forall x(\varphi(x) \supset x \in X)].$$

For example,

Most men walk to work

is equivalent to

$\exists X$[Most men $x(x \in X) \mathbin{\&} \forall x(x \in X \supset x$ walks to work)]

whereas

Few men walk to work

is equivalent to

$\exists X$[Few men $x(x \in X) \mathbin{\&} \forall x(x$ walks to work $\supset x \in X)]$.

Note that 'x walks to work' appears on different sides of the $\supset$ in these two examples.

If Q_1 and Q_2 are *both monotone increasing*, then we can make perfectly good sense of

(s) $$\begin{matrix} Q_1 x \\ Q_2 y \end{matrix} \!\!> A(x, y)$$

using (q), since it should mean

(t) $$\begin{aligned} &\exists X \exists Y[Q_1 x(x \in X) \mathbin{\&} Q_2 y(y \in Y) \\ &\quad \mathbin{\&} \forall x \forall y(x \in X \mathbin{\&} y \in Y \supset A(x, y))]. \end{aligned}$$

Notice that

(s) $$\begin{matrix} Q_1 x \\ Q_2 y \end{matrix} \!\!> A(x, y)$$

thus logically implies both of

(s′) $Q_1 x Q_2 y A(x, y)$,

(s″) $Q_2 y Q_1 x A(x, y)$.

We have already seen one case of this with

$$\left.\begin{array}{l}\forall x\\ \exists y\end{array}\right> A(x, y)$$

since it implies $\forall x \exists y A(x, y)$ and is actually equivalent to $\exists y \forall x A(x, y)$. In general, however, there is no reason for (s) to be equivalent to either linear form (s′) or (s″).

Similarly, if Q_1 and Q_2 are *both monotonic decreasing*, then we can define

(s) $$\left.\begin{array}{l}Q_1 x\\ Q_2 y\end{array}\right> A(x, y)$$

by

$$\exists X \exists Y[Q_1 x(x \in X) \,\&\, Q_2 y(y \in Y) \,\&\, \forall x \forall y(A(x, y) \supset x \in X \,\&\, y \in Y)].$$

Let us consider, for example, the sentence (25).

(25) More than half the dots and more than half the stars are all connected by lines.

Using $Q_1 x$ for 'more than half the dots x' and $Q_2 y$ for 'more than half the stars y', we would analyze (25) as

(u) $$\left.\begin{array}{l}Q_1 x\\ Q_2 y\end{array}\right> Cxy$$

where Cxy stands for 'x and y are connected by a line'. Notice that (25) implies both of (26) and (27).

(26) More than half the dots are connected by a line to more than half the stars.

(27) More than half the stars are connected by a line to more than half the dots.

These correspond to

(u′) $Q_1 x Q_2 y Cxy$

(u″) $Q_2 y Q_1 x Cxy$

respectively. Neither of these imply (u). For example, (u′) is true in Figure 4, but (u) is not.

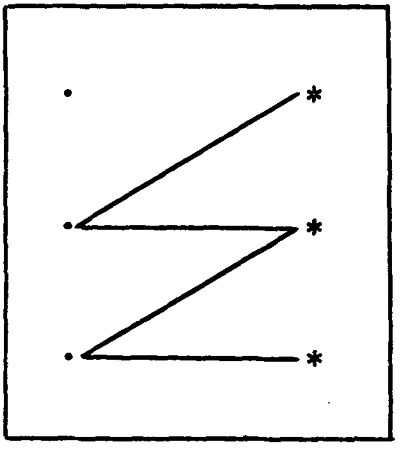

Fig. 4

More complicated examples, similar to that given in the Appendix for Hintikka's (f), can be given to show that (u) cannot be expressed as a first order sentence involving $\forall$, $\exists$, Q_1, Q_2. Of course it can be expressed linearly if we explicitly quantify over *sets* of individuals, as in (t), so we see that (25) contains a hidden quantifier over sets.

We have assigned meanings to

(s) $$\begin{matrix} Q_1 x \\ Q_2 y \end{matrix} > A(x,y)$$

when Q_1 and Q_2 are both monotone increasing or both monotone decreasing, but there is no sensible way to interpret (s) when one is increasing and the other is decreasing. Thus, for example,

(29) ? Few of the boys in my class and most of the girls in your class have all dated each other.

appears grammatical, but it makes no sense (cf. Note 15). It seems hard to know how one would account for this fact without the distinctions we have been making.

More complicated branching quantifiers can be interpreted similarly, for example

$$\begin{matrix} \forall x - Q_1 y \\ \forall z - Q_2 w \end{matrix} > A(x, y, z, w)$$

$$\begin{matrix} \forall x & & Q_1 y \\ & \times & \\ \forall z & & Q_2 z \end{matrix} > A(x, y, z, w)$$

as long as Q_1 and Q_2 are both monotone increasing (or decreasing). Before illustrating these, however, we must pause to take up something we have been putting off.

Our treatment of quantifiers does not count 'most' or 'few' as quantifiers, but rather it is noun phrases like 'most men', 'few dots', 'many relatives of x' that act as quantifiers. For example, if our domain D consists of all men, then

(30) Most congressmen are corrupt

cannot be treated as

$$\text{Most } x\,[\text{Congressman}(x) \supset \text{Corrupt}(x)]$$

for then (30) would be true, simply because most men are not congressmen, even if no congressmen were corrupt. Rather, it must be treated as

$$\text{Most congressmen } x\,[\text{Corrupt } x].$$

Similarly, 'most relatives of x' depends on x and so we need to treat

Most ancestors of each villager are revered by him

as:

For each villager x [Most ancestors y of x (y is revered by x)]

or, more symbolically,

$$\forall x \in V[My \text{ such that } Ayx(Rxy)]$$

or, as we shall write

$$\forall^{V}x\,[M^{Ayx}y(Rxy)]\,.$$

This is actually a better way to treat $\forall$ and $\exists$, too, though we refrained in Section 1 to keep this issue from getting confused with the question of branching quantification any sooner than necessary. For example, it is syntactically more natural to treat

Some ancestor of every villager is revered by him

as

$$\forall^{V}x\,\exists^{Ayx}y\,Rxy$$

than as

$$\forall x\,\exists y\,[V(x) \supset (Ayx \,\&\, Rxy)]\,.$$

We are now in a position to discuss examples like (19), (20) and (24).

(19) Most relatives of each villager and most relatives of each townsman hate each other.

The preferred reading of (19), the one which strengthens the reading (g) of (1), can be symbolized by:

(v)
$$\left.\begin{matrix} \forall^{V}x & M^{Rxy}y \\ & \times & \\ \forall^{T}z & M^{Rzw}w \end{matrix}\right\rangle Hyw$$

which we could linearize a bit by

$$\forall^{V}x\,\forall^{T}z \left\langle \begin{matrix} M^{Rxy}y \\ M^{Rzw}w \end{matrix} \right\rangle Hyw$$

but it cannot be completely linearized. In second-order logic we can express it by making the hidden quantifiers over sets explicit. In a combination of English and second-order logic it would read:

For every villager x and every townsman z there are sets Y containing most of x's relatives and W containing most of z's relatives such that every y in Y and every w in W hate each other.

Those who agree with Hintikka's reading of (1) are forced to find a similar reading for (19), since (19) clearly implies (1). Thus, they are forced to read (19) as:

(w) $$\begin{matrix} \forall^V x - M^{Rxy} y \\ \forall^T z - M^{Rzw} w \end{matrix} \Big> Hyw.$$

To linearize (w) it we must introduce Skolem functions F, G from individuals to sets of individuals:

> There are set valued functions F, G such that, for every villager x, $F(x)$ contains most of x's relative's, and for townsman z, $G(z)$ contains most of z's relatives, and for all y in $F(x)$ and all w in $G(z)$, y and z hate each other.

We favor the reading (v) of (19) over (w), but it doesn't really matter for the general issue, since either reading of (19) is an essential use of branching quantification.

There are sentences similar to our examples which do not require branching quantification for their analysis, but where it still seems useful.

(31) Quite a few boys in my class and most girls in your class all passed the test.

One could treat this as being derived from

(32) Quite a few boys in my class all passed the test and most girls in your class all passed the test.

There is no need for this *ad hoc* move, however, since it can be analyzed simply in terms of branching quantifiers (using Q for 'quite a few boys in my class', M for 'most girls in your class' and Px for 'x passed the test') as

$$\begin{matrix} Qx \\ Mx \end{matrix} \Big> Px.$$

Before concluding, let us point out that there is a version of Proposition 1 that holds for the formulas we are considering here so that the test we gave in section one is still valid. The reader can check for himself that our examples pass the test.

It seems likely that once the misconceptions surrounding branching quantification are cleared up, that they should provide the linguist with a useful tool for discussing questions of quantifier scope, among others. We have tried to show by some examples how they can provide a more faithful translation of certain conjoined noun phrases, even when there are logically equivalent first-order sentences. We have also shown, we feel, that there are certain conjoined noun phrases that really need branching quantification for a logical analysis, though to get any really convincing examples, we had to allow monotone quantifiers other than $\forall$ and $\exists$.

APPENDIX

In this Appendix we want to prove the results stated in the paper itself.

A.1. *First-order approximations of FPO sentences*

We begin with a discussion of the first-order approximations of an FPO sentence for three reasons. In the first place, it shows the wide range of possible first-order sentences that can, under certain conditions, be equivalent to a given FPO sentence, possibilities that aren't hinted at in the papers cited earlier. Secondly, the discussion will be used to *prove* Hintikka's claim that the FPO reading of sentence (1) indeed cannot be expressed by any first-order sentence. Finally, the discussion shows weaknesses in the attack in Stenius (1976) on Hintikka's discussion of game-theoretic semantics.[17]

In the paper Barwise (1976) (which was written for mathematical reasons, in total ignorance of the controversy into which we are now stepping) we showed how to approximate FPO sentences by first-order sentences for certain purposes. To illustrate, consider some FPO sentence φ of the form

$$\text{(a)} \qquad \begin{matrix} \forall x - \exists y \\ \forall z - \exists w \end{matrix} > A(x, y, z, w).$$

We can write down first-order sentences $\varphi_1, \varphi_2, \varphi_3, \ldots$ with the following properties:

(1) For each n, the sentence $(\varphi \supset \varphi_n)$ is logically valid.

(2) For each n, the sentence $(\varphi_{n+1} \supset \varphi_n)$ is logically valid.

(3) Any first-order theory consistent with each of $\varphi_1, \varphi_2, \ldots$ is consistent with the original FPO sentence φ.

From these facts it immediately follows[18] that if φ is logically equivalent to *any* first-order sentence, it must be logically equivalent to one of the φ_n. The first and weakest approximation to φ is the sentence which is most easily confused with φ, the one which came up in the discussion in the paper. Namely,

$$(\varphi_1) \qquad \forall x \forall z \exists y \exists w A(x,y,z,w).$$

The further approximations are more complicated, and it is in interpreting and contrasting these that the game-theoretic interpretation is useful.

$$(\varphi_2) \qquad \forall x \forall z \exists y \exists w \forall x' \forall z' \exists y' \exists w' [A(x,y,z,w) \,\&\, A(x',y',w',z') \\ \&\, (x = x' \supset y = y') \,\&\, (z = z' \supset w = w')]$$

$$(\varphi_n) \qquad \forall x_1 \forall z_1 \exists y_1 \exists w_1 \ldots \\ \forall x_n \forall z_n \exists y_n \exists w_n \left[\&_{i=1}^{n} A(x_i, y_i, z_i, w_i) \,\& \right. \\ \left. \&_{\substack{i=1,\ldots,n \\ j=1,\ldots,n}} [(x_i = x_j \supset y_i = y_j) \,\&\, (z_i = z_j \supset w_i = w_j)]\right].$$

To understand the meaning of these approximations, we view the original φ as a game between two teams named $\forall$ and $\exists$ (Hintikka's 'Nature' and 'Myself'). Each team has two players $\forall_1, \forall_2$ and $\exists_1, \exists_2$ respectively. A *play* of the game consists of $\forall_1$ and $\forall_2$ choosing individuals x and z respectively. $\forall_1$ shows x to $\exists_1$ and $\forall_2$ shows z to $\exists_2$. Then $\exists_1$ and $\exists_2$ choose individuals y and w respectively, *without communicating in any way*. If A holds of (x, y, z, w) then team $\exists$ has won. Otherwise team $\forall$ wins. The sentence φ asserts that team $\exists$ has a fixed winning *strategy* that will allow them to win this game without cheating, no matter what team $\forall$ plays.

In these anthropomorphic terms, which can be made rigorous, the meaning of the first approximation is that team $\exists$ has a winning strategy *if they cheat* by telling each other what $\forall_1$ and $\forall_2$ have played.

Now suppose that Team $\exists$ claims to have a winning strategy. Team $\forall$ challenges them to play, and $\exists$ wins. But Team $\forall$ is suspicious. They think Team $\exists$ cheated by some sort of secret signal. So, Team $\forall$ challenges them to play again, using the same strategy. This means that if $\forall_1$ chooses the same x he chose on the first play, then $\exists_1$ must make the same response and similarly for $\forall_2$ and $\exists_2$. The second approximation asserts that if Team $\exists$ cheats, they can win two successive plays of the game without their cheating being detected by an inconsistency. Similarly for the further approximations.

It is to be noted, however, that just because $\exists$ can't be shown to be cheating in any finite number of moves, it doesn't follow that they aren't cheating. That is, it is possible for all of the approximations to be true in a particular model without φ being true.

A.2. *A proof that Hintikka's formula (f) is not first-order*

Using this game theoretic interpretation of φ and its approximations, we can give a proof of the following.

PROPOSITION 2. Hintikka's FPO sentence

$$\text{(f)}\qquad \begin{matrix} \forall x - \exists y \\ \forall z - \exists w \end{matrix} > [(Vx \,\&\, Tz) \supset (Rxy \,\&\, Rzw \,\&\, Hyw)]$$

is not expressible in first-order logic.

Proof. This proof was worked out in a conversation with Kenneth Kunen. It suffices, by the remarks above, to show that (f) is not equivalent to any of its finite approximations. Given an n we will show how to construct an idealized village V, and town T such that the above is false, but the nth approximation to it is true. The town and village will be very antagonistic toward each other. In particular, every villager will hate every townsman and every townsman will hate all but one villager.

The village will have n clans, say $C_1, \ldots, C_n$ each with two members (they could have more, but it just complicates the argument), say $C_i = \{a_i, b_i\}$. The relation R of being relatives is defined to hold only between members of the same clan.

The town is of course larger than the village. It has 2^n clans, each of size n. Again, the relation R holds between all and only members of the same clan.

There are 2^n ways of choosing one member of each village clan, that is, 2^n functions f with domain $\{1, \ldots, n\}$ such that $f(i) \in C_i$ for all i. We use these functions to label the town clans D_f. The members of D_f we write as $d_1^f, \ldots, d_n^f$. The townsman d_i^f hates all the villagers except for one, namely the villager $f(i)$. Thus the relation H holds between all pairs of villagers and townsmen except for those pairs of the form $\langle f(i), d_i^f \rangle$.

This defines, in a rather informal way, our model in which the FPO sentence is false but its nth approximation is true. Figure 1 given in Section 1 gives a picture of this model for $n = 3$. To see that (f) is false in the model, notice that if it were true we could pick one member of each village clan and one member of each town clan so that all the village–town pairs hate each other. But the selection of one member of each village clan amounts to one of the functions f with $f(i) \in C_i$ for all i. But then no one clan D_f hates each of these choices since d_i^f does not hate $f(i)$.

Rather than go through a formal proof that the nth approximation is true in the nth model, which seems out of place here, we will show the basic strategy in the $n = 3$ case illustrated in Figure 1. You be Team $\forall$, I'll be $\exists$. You get to choose a hut and a house, I choose a dot and star from your chosen hut and house so that no line connects them. But I must play so that in any three plays, whenever you repeat a clan, I repeat by earlier choice. By symmetry, we can assume you always play hut 1, house 111 first. I play 1 in hut 1, 2 in house 121.

	You	Me	You	Me	You	Me
I.	(1, 111)	(1, 2)	(2, 121)	(2, 3)	(3, 211)	(2, 1)
II.	(1, 111)	(1, 2)	(2, 112)	(2, 3)	(3, 121)	(1, 1)

Notice that in II, I can't survive another round of play, for if you chose (1, 121) I am committed to (1, 1) by my earlier play, but these two do not hate each other.

A.3. *A proof that the embedding sentence (m) is not first-order*

PROPOSITION 3. The second-order sentence

$$\text{(m)} \qquad \exists f \forall x \forall y [(Lx \,\&\, Ly \,\&\, x < y) \supset (Sf(x) \,\&\, Sf(y) \,\&\, f(x) \prec f(y))]$$

is not logically equivalent to a first-order sentence, and hence neither is the logically equivalent branching formula (n).

Proof. We could give a proof along the lines of the above proof, using the first-order approximations of (n), but there is a much simpler proof. Suppose (m) were equivalent to some first-order sentence ψ (or even to some first-order theory T, the same proof would work). Consider the structure $\mathscr{D} = \langle D, L, <, S, \prec \rangle$ where $L = \mathbb{R}$, the set of real numbers with its ordering $<$, and $S = \mathbf{Q}$, the set of rationals with its ordering $\prec$ and $D = L \cup S$. Since L is uncountable and S is countable, (m) and hence ψ are false in $\mathscr{D}$. The Lowenheim–Skolem Theorem asserts that we can find a countable submodel $\mathscr{D}_0 \subseteq \mathscr{D}$,

$$\mathscr{D}_0 = \langle M_0, L_0, < \cap (L_0 \times L_0), S, \prec \rangle \quad \text{such that}$$

$\mathscr{D}_0$ and $\mathscr{D}$ satisfy the same first-order sentences. (See Chapter A1 in Barwise (1977) or any textbook in mathematical logic for a proof of the Löwenheim–Skolem Theorem.) But any countable linear ordering can be embedded in the rationals so (m) and hence ψ are true in $\mathscr{D}_0$, a contradiction. □

This seems like a reasonable place to bring up the following example.

(14′) The richer the country, the more powerful are quite a few of its officials.

In terms of orderings $\langle L, < \rangle$ and $\langle S, \prec \rangle$ and a binary relation Oxy for 'y is an official of x' this seems to assert:

(m′) There is a function F from L to subsets of S such that for each x in L, $F(x)$ contains quite a few y such that Oxy, and for every x, z in L if $x < y$ then for all y in $F(x)$ and all w in $F(z)$, $y \prec z$.

However, if we try to express (m′) with a branching quantifier, it turns out to have a different meaning, because it asserts the existence of two functions and there is no trick like we used in (n) to make the two functions equal. This seems like a good argument for treating (14) as a case of hidden second-order quantification, rather than as branching quantification.

A.4. *A proof of Proposition 1 from Section 1*

The proposition follows immediately from the following lemma which is practically a restatement of the Craig Interpolation Theorem for first-order logic. The Interpolation Theorem shows that if $(\varphi_1 \supset \varphi_2)$ is a logically valid sentence of first-order logic, then there is a first-order θ such that θ contains only those non-logical constants common to φ_1 and φ_2 such that $(\varphi_1 \supset \theta)$ and $(\theta \supset \varphi_2)$ are logically valid. (See Chapter A.2 in Barwise (1977) or most textbooks in logic for a proof.)

LEMMA. If φ and ψ are sentences of FPO and if $(\psi \supset (\sim \varphi))$ is logically valid then there is a first-order sentence θ such that $(\psi \supset \theta)$ and $(\varphi \supset \sim \theta)$ are logically valid.

Proof. Using (a′) from Section 1, we can write ψ and $\sim \varphi$ in the second-order forms

$$\exists f_1 \ldots \exists f_n A(f_1 \ldots f_n)$$
$$\forall g_1 \ldots \forall g_k B(g_1 \ldots g_k)$$

respectively, where f_i, g_j are function symbols, $A(f_1 \ldots f_n), B(g_1 \ldots g_k)$ are first-order, and we can use different function symbols and so suppose that none of the f_i are the same symbols as any of the g_j. Thus,

$$[\exists f_1 \ldots \exists f_n A(f_1 \ldots f_n)] \supset [\forall g_1 \ldots \forall g_k B(g_1 \ldots g_k)]$$

is logically valid. We can rewrite this as

$$\forall f_1 \ldots \forall f_n \forall g_1 \ldots \forall g_n [A(f_1 \ldots f_n) \supset B(g_1 \ldots g_k)].$$

So, we have

$$A(f_1 \ldots f_n) \supset B(g_1 \ldots g_k)$$

is logically valid. Now apply the interpolation theorem to get a first-order θ in which none of the f_i or g_j occur such that

$$A(f_1 \ldots f_n) \supset \theta$$
$$\theta \supset B(g_1 \ldots g_k)$$

are logically valid. But then we can universally quantify out the function symbols, pull out the θ and get

$$\exists f_1 \ldots \exists f_n A(f_1 \ldots f_n) \supset \theta$$
$$\theta \supset \forall g_1 \ldots \forall g_k B(g_1 \ldots g_k)$$

as logically valid, as desired. □

A. 5. *Remarks on the principle of compositionality and branching quantification*

We noted in the introduction that the existence of branching quantification in English requires us to reinterpret Frege's principle of compositionality. It is not possible to explain the meaning of an essential use of branching quantification

(a) $$\left.\begin{array}{l} \forall x - \exists y \\ \forall z - \exists w \end{array}\right> A(x, y, z, w)$$

or even

$$\left.\begin{array}{l} Qx \\ Qy \end{array}\right> A(x, y, z, w)$$

for monotone increasing Q, inductively by treating one quantifier at a time in a first-order fashion. Some use of higher-type abstract objects is essential.

This has important consequences for any model-theoretic analysis of the semantics of fragments of English containing branching quantification. For example, Gabbay and Moravcsik (1974) defined a formal fragment of English containing the following sentence.

(33) Every man loves some woman (and) every sheep befriends some girl that belong to the same club.

They assumed that this sentence had a branching reading of the form (a) and attempted to give an inductive definition of a Montague-Style semantics where (33) came out with the desired interpretation.

Güenthner and Hoepelman (1974) pointed out that their semantics did not give (33) a branching interpretation after all, but rather the first-order interpretation where both woman and girl depend on both man and sheep. Our discussion in Section 1 suggests that the Gabbay–Moravcsik semantics does give the correct meaning to (the preferred reading of) (33), but that it is not of form (a) but of form (g) – that is, that it is not an example of essential branching quantification. But if one *does* read (33) as being of form (a), or as having *a reading* of form (a), then one must not try to give the kind of inductive definition of meaning used in Montague – where one quantifier is interpreted at a time in a first-order way.

The mathematical theory of inductive definitions provides a tool for making this claim precise. For an introduction to this theory we refer the reader to Aczel's Chapter C. 7 of Barwise (1977), to Moschovakis (1974), to Chapter VI of Barwise (1975) or to the paper Barwise and Moschovakis (1978).

PROPOSITION 4.

(i) The relation

$$\varphi \text{ is true in } \mathscr{D},$$

where $\varphi \in$ FPO and $\mathscr{D}$ is a first-order structure, is not inductively definable (i.e., it is not an inductive verifiability relation in the sense of Barwise–Moschovakis (1978)).

(ii) The relation

$$\begin{matrix} Qx \\ Qy \end{matrix} \Big> \varphi(x, y) \text{ is true in } (\mathscr{D}, Q_D),$$

where φ is first-order, $\mathscr{D}$ is a first-order structure and Q_D is a monotone increasing quantifier on $\mathscr{D}$, is not inductively definable in $L(Q)$, (i.e., it is not inductive* in the logic $L(Q)$ as defined in Barwise (1978)).

Proof. (i) Otherwise, the Abstract Completeness Theorem (3.3) of Barwise–Moschovakis (1978) would apply to show that the set of valid sentences of FPO is recursively enumerable which, as Hintikka (1974) pointed out, is about as far from the case as is imaginable.

(ii) The same proof would work, except that the set of valid sentences of the given form *is* recursively enumerable – as we will show elsewhere. Thus we must give a little more complicated proof – one that follows the line used in the Lemma in A. 4. Suppose the relation in question were inductive in $L(Q)$. Then, by computations in IV.2 and VI.1 of Barwise (1975), every sentence of the form

$$\begin{matrix} Qx \\ Qy \end{matrix} \Big> \varphi(x, y)$$

would be equivalent to one of the form

$$\forall R\theta(R),$$

where R is a new relation symbol and $\theta(R)$ is in the logic $L(Q)$, as well as to

$$\exists U \exists V[QxUx \,\&\, QyVy \,\&\, \forall x \forall y[(Ux \,\&\, Vy) \supset \varphi(x, y)]].$$

But then, by the interpolation theorem for $L(Q)$, where Q is a monotone quantifier symbol, (due to Sgro and Shelah, presented in essence in Bruce (1978)), the branching expression would be equivalent to a sentence of $L(Q)$. But it is easy to find counterexamples. For example, it follows from a construction in Malitz and Magidor (1977, p. 224) that

$$\begin{matrix} Qx \\ Qy \end{matrix} \Big> [(x<y) \vee (y<x)]$$

cannot be expressed in $L(Q)$. □

We can improve 5. i as follows. The proof is a little more complicated, but it is more basic and shows much more.

PROPOSITION 6. Let $\mathscr{D} = \langle D, N, +, 0, 1, \ldots \rangle$ be *any* first-order structure which contains a copy of the natural numbers N with addition and multiplication. For example, $\mathscr{D}$ might be some universe of set theory. Let $T_{\mathscr{D}}$ be the set of (Godel numbers $\ulcorner \varphi \urcorner$ of) sentences $\varphi \in$ FPO which are true in $\mathscr{D}$. Then $T_{\mathscr{D}}$ is not Π_1^1 definable over $\mathscr{D}$, i.e., not definable by a formula $\Psi(x)$ of the form

$$\forall f_1 \ldots \forall f_n A(x, f_1 \ldots f_n)$$

where A is first-order and $f_1 \ldots f_n$ are function symbols. This is even stronger than that $T_{\mathscr{D}}$ is not inductive since every inductive set is Π_1^1, but, in general, not conversely (if $\mathscr{D}$ is uncountable, e.g.).

Proof. Suppose $T\mathscr{D}$ is Π_1^1. Then there is an FPO formula $\varphi(x)$ which defines $\bar{T}_{\mathscr{D}}$, the complement of $T_{\mathscr{D}}$, by the Enderton–Walkoe result mentioned in Section 1.

Now let sub be the substitution operator so that for any natural number n and any $\theta(x)$, $\text{sub}(\ulcorner \theta(x) \urcorner, \ulcorner n \urcorner) = \ulcorner \theta(\ulcorner n \urcorner) \urcorner$, where $\ulcorner n \urcorner$ is the term denoting n. (See, e.g., Smorynski's Chapter D. 1 of Barwise (1977).) Then let $\psi(x)$ be $\varphi(\text{sub}(x, x))$ and let $m = \ulcorner \psi(x) \urcorner$. Then, by the definition of ψ,

$\psi(\ulcorner m \urcorner)$ is true in $\mathscr{D}$ iff $\varphi(\text{sub}(\ulcorner m \urcorner, \ulcorner m \urcorner))$ is true in $\mathscr{D}$.
iff $\ulcorner \psi(\ulcorner m \urcorner) \urcorner \in \bar{T}\mathscr{D}$ (by choice of φ).
iff $\psi(\ulcorner m \urcorner)$ is false in $\mathscr{D}$ (by definition of $T\mathscr{D}$).

□

There is a way around the difficulty posed by Proposition 6, if one is willing to give up a one-quantifier-at-a-time analysis and use a type-theoretic model theory similar to Montague's, or something equivalent to it. The point is that if $\mathscr{D}$ is a higher-type domain of the kind used in Montague (1974) (PTQ), then there *are* no variables $x, y, \ldots$ of his logic which range over *everything* as in (a), but rather only variables over various subdomains D_a, say x^a, y^a, etc. Thus, to interpret (for example)

$$(\mathrm{a}_1) \qquad \begin{matrix} \forall x^a - \exists y^b \\ \forall z^c - \exists w^d \end{matrix} > A(x^a, y^b, z^c, w^d)$$

we must use

$$(\mathrm{a}_1') \qquad \exists f^{\langle a,b\rangle} \exists g^{\langle c,d\rangle} \forall x^a \forall z^c A(x^a, f^{\langle a,b\rangle}x^a, z^c, g^{\langle c,d\rangle}z^c)$$

where the variable $f^{\langle a,b\rangle}$ ranges over the set $D_{\langle a,b\rangle}$ of *all* functions from D_a D_b and similarly for $g^{\langle c,d\rangle}$. Thus, there is nothing impossible about defining the whole of (a_1) by (a_1'). The problem is to do it in a linguistically natural way, one that respects the basic categories and syntactic structures of English.

A. 6. *Mathematical logic*

The discovery of branching quantifiers in natural language with monotone quantifiers reopens the whole question of the model theory of branching quantifiers – a rather dormant field recently. The subject is much richer in this more general context than in the FPO case, because one doesn't have the quantifier manipulations used by Walkoe (1970) to reduce all branching quantifiers to those of the form (b). It seems that there is a lot of research to be done in this area.

Department of Mathematics
University of Wisconsin, Madison

NOTES

[1] The preparation of this paper was partially supported by NSF Grant MSC76-06541.
[2] I would like to thank Robin Cooper for valuable comments on an earlier draft of this paper.
[3] This paper also appeared as Hintikka (1973).
[4] We will prove a precise version of this claim in Section A. 5 of the Appendix and discuss its relevance to Montague grammar. In particular, we will discover the source of the difficulty in Gabbay and Moravcsik (1974) – the difficulty pointed out in Güenthner and Hoepelman (1974).
[5] Cf. Stenius (1976).
[6] Cf. Hintikka (1976). The numbering given above has been adjusted to fit in with our examples.
[7] See Ioup (1975) for a discussion of this matter. In particular, notice that her Table 1 gives 'each' the very highest tendency toward wide scope, 'every' comes second. Sentence (5.1) comes from this paper.
[8] That is, FPO is exactly what Walkoe (1970) calls H'. See Note 9.

[9] Since Hintikka is attempting to show that every FPO sentence can be mirrored in English, it is important for his argument that FPO is defined to allow just one partially-ordered quantifier in front of a first-order formula (the logic Walkoe (1970) calls H'), not as the much stronger logic (which Walkoe calls H, after Henkin) where (b) is treated as a formation rule to be mixed in with the formation rules from first-order logic.
[10] The proof will be given in the Appendix.
[11] This discussion offers further evidence that NP's should be treated as quantifiers, as they are in Montague Grammar, since (10′) is clearly negation normal, whereas (9′) isn't.

(9′) It is false that John likes worms.

(10′) John doesn't like worms.

But this is beside the point here.
[12] This will be proved in the Appendix.
[13] This will also be proved in the Appendix. It is much easier to prove than that (f) is not first-order.
[14] I have not seen examples like these discussed in the linguistic literature. I include (15) because it is the most famous sentence of this form. Logically it is much simpler than (16) for the underlying sets L and S are the same, and it is the identity function which embeds one ordering *bigger-than* into the other *falls-harder-than*. Thus:

$$\forall x \forall y (x \textit{ bigger-than } y \supset x \textit{ falls-harder-than } y).$$

[15] We are using the phrase 'date each other' in (22) and (29) to apply only to persons of the opposite sex. If we were willing to be more risque we could obviously use a different phrase which more clearly satisfies this assumption.
[16] More formally, satisfaction of formulas in a structure $\mathscr{D} = \langle D, \ldots, Q_D \rangle$, where $Q_D \subseteq$ Power (D), is defined inductively with the clause for the new quantifier being:

$$\mathscr{D} \models Qx\varphi(x, a_1 \ldots a_n) \text{ iff } \{b \in D \mid \mathscr{D} \models \varphi(b, a_1, \ldots, a_n)\} \in Q_D.$$

The model theory of such monotone quantifiers is well understood on a first-order level. We refer the reader to Barwise (1978) for some results and further references. A more complete description will appear in our paper under preparation. See also Keisler (1970) and Chapters A. 1, A. 2, C. 6 and C. 7 in Barwise (1977).
[17] Stenius complains that Hintikka does not apply the game-theoretic semantics. He also asserts that "... game-theoretical rules as these are formulated in Hintikka's paper *cannot* account for branched quantification, since they presuppose a linear order of 'moves' ". These are the parts of Stenius' attack that are relevant here.
[18] For suppose $(\varphi \equiv \psi)$ is logically valid, where ψ is first order. Then, by (3) applied to the theory $\neg\psi$, $\neg\psi$ is not consistent with some φ_n. I.e., $(\varphi_n \supset \psi)$ is logically valid. But so are $(\psi \supset \varphi)$ and $(\varphi \supset \varphi_n)$. Thus $(\varphi_n \equiv \psi)$ is valid.

REFERENCES

Aczel, P., 'Quantifiers, Games and Inductive Definitions', *Proc. of 3rd Scandinavian Logic Symposium*, North Holland, Amsterdam, 1975.

Barwise, J., *Admissible Sets and Structures*, Springer-Verlag, 1975.

Barwise, J., 'Some Applications of Henkin Quantifiers', *Israel J. Math.* 25 (1976), 47–63.

Barwise, J. (ed.), *Handbook of Mathematical Logic*, North Holland, Amsterdam, 1977.

Barwise, J., 'Monotone Quantifiers and Admissible Sets', *Generalized Recursion Theory II*, North Holland, Amsterdam, 1978, pp. 1–38.

Barwise, J. and Moschovakis, Y. N., 'Global Inductive Definability', *J. Symbolic Logic* 43 (1978).

Bruce, K., 'Ideal Models and Some Not So Ideal Problems in the Theory of $L(Q)$, *J. Symbolic Logic* (1978).

Enderton, H. B., 'Finite Partially-ordered Quantifiers', *Z. Math. Logik und Grundlagen der Math.* 16 (1970) 393–397.

Fauconnier, G., 'Do Quantifiers Branch?', *Linguistic Inquiry* VI (1975), 555–578.

Gabbay, D., and Moravcsik, J. M. E., 'Branching Quantifiers, English and Montague Grammar', *Theoretical Linguistics* 1 (1974) 141–157.

Guenthner, F. and Hoepelman, J., 'A Note on the Representation of "Branching Quantifiers" ', *Theoretical Linguistics* (1974), 285–291.

Henkin, L., 'Some Remarks on Infinitely Long Formulas', *Infinitistic Methods*, Warsaw, pp. 167–183.

Hintikka, J., 'Quantifiers vs. Quantification Theory', *Dialectica* 27 (1973), 329–358.

Hintikka, J., 'Quantifiers vs. Quantification Theory', *Lingusitic Inquiry* V (1974), 153–177.

Hintikka, J., 'Partially Ordered Quantifiers vs. Partially Ordered Ideas', *Dialectica* 30 (1976), 89–99.

Ioup, G., 'Some Universals for Quantifier Scope', in J. P. Kimball (ed.), *Syntax and Semantics*, Vol. 4, pp. 37–58, Academic Press, New York, 1975.

Keisler, H. J., 'Logic with the Quantifier "There Exist Uncountably Many" ', *Annals of Math. Logic* I (1970), 1–93.

Magidor, M. and Malitz, J., 'Compact Extensions of $L(Q)$ (Part 1a)', *Annals of Math. Logic* 11 (1977), 217–261.

Montague, R., in R. Thomason (ed.), *Formal Philosophy*, Yale Univ. Press, 1974.

Moschovakis, Y. N., *Elementary Induction in Abstract Structures*, North-Holland, Amsterdam, 1974.

Stenius, E., 'Comments on Jaakko Hintikka's Paper "Quantifiers vs. Quantification Theory" ', *Dialectica* 30 (1976), 67–88.

Walkoe, W., Jr., 'Finite Partially Ordered Quantification', *Journal of Symbolic Logic* 35 (1970), 535–550.