Monotone Decreasing Quantifiers in a Scope-Free Logical Form

Jerry R. Hobbs
Artificial Intelligence Center
SRI International

1 Introduction

In Hobbs (1983) (henceforth, ITQ) and Hobbs (1985) (OP) I developed the outlines of an approach to semantic representation in which the logical form of an English sentence is a flat (i.e., scope-free) conjunction of existentially quantified, positive literals, with roughly one literal per morpheme. In this representation scheme the logical form of a sentence is vague with respect to quantifier-scoping decisions, and further information about scoping relations is encoded in the form of further existentially quantified positive literals. In the DIALOGIC system for syntactic analysis, developed in the early and middle 1980s, translations into such a logical form were implemented for a great majority of English syntactic constructions, and this system was used successfully in a number of applications. In Hobbs et al. (1993) (IA) my colleagues and I developed an approach to the interpretation of discourse in which to interpret a text is to find the least-cost abductive proof of the logical forms of the sentences of the text, essentially by back-chaining on mostly Horn-clause axioms in the knowledge base and making assumptions when necessary.

One shortcoming of the proposal advanced in ITQ was in the treatment of monotone decreasing quantifiers, such as "few" and "no" (cf. Barwise and Cooper, 1981). A monotone increasing quantifier, like "most", is "monotone increasing" beacuse when the predicate in the body of the quantified expression is made less restrictive, the truth value is preserved. Thus,

Most men work hard.

entails

(1) Most men work,

By contrast, for monotone decreasing quantifiers, when the predicate in the body of the quantified expression is made less restrictive, the truth value is not necessarily preserved. Quite the opposite. It is preserved when the body is made *more* restrictive.

(2) Few men work.

entails

Few men work hard.

Since "x works hard" entails "x works", a flat, scope-free representation for "few men work hard" runs into problems, because it would seem to allow the incorrect inference "few men work".

In ITQ I suggested very briefly a logical form for such sentences in which the quantifier "few" is translated into a predicate that means "all but a few" and the predication of the body of the quantified expression is negated. Thus, sentence (2) would be interpreted as if it were

All but a few men don't work.

This solves the entailment problem. "x doesn't work" entails "x doesn't work hard." Thus, "Few men work" would be equivalent to "all but a few men don't work", which entails "all but a few men don't work hard," which would be equivalent to "few men work hard." This approach is similar to that of van Eijck (1983).

However, this is not a felicitous solution, since the negation of the main verb makes the compositional semantics of the quantifier nonlocal, in that information from the noun phrase other than its referent is required in the interpretation of the rest of the sentence.

In this paper, I use the insights afforded by IA to propose a different analysis of monotone decreasing quantifiers, one in which the right interpretation arises from a combination of a single rule for interpreting quantifiers, both monotone increasing and monotone decreasing, and the pragmatic process of specializing or strengthening interpretations that is the basis of the abduction approach. Along the way I redo or repair several other features of the ITQ approach that were infelicitous or incorrect in the original, and

I mention in passing the scope-neutral representation of functional dependencies among quantified variables that this approach makes possible. The result is a picture wherein syntactic analysis and semantic translation yields a representation that makes fewer distinctions than we might wish, but is strictly locally compositional, and strengthening to the desired representation is done by pragmatic processes that already have such strengthening as their task.

2 Background

The IA approach may be thought of as dividing the interpretation of a sentence into a "compositional semantic" phase, in which the explicit content of the sentence is represented in a logical notation (referred to here as the logical form), and a "pragmatic" phase that inferentially determines the contextually appropriate specific information that the speaker intended to convey (although, in fact, both phases use the same abductive inferential mechanism and can intermix freely). The compositional semantic phase is strictly local, in the sense that the interpretation of noun phrases does not require information from elsewhere in the sentence, and the only information about a noun phrase that is used in the semantic interpretation of the rest of the sentence is a variable indicating its referent.

Using as the logical form of a sentence a flat conjunction of existentially quantified positive literals becomes possible through an approach to representation called "ontological promiscuity", in which there is extensive reification of such things as eventualities, possible and even impossible individuals, sets, typical elements of sets, and so on. The introduction of eventualities is a key move. In addition to having predications of the form

$$work(J)$$

saying that John works, we also have predications of the form

$$work'(E, J)$$

saying that E is the eventuality of John's working. This eventuality may or may not obtain in the real world. If it does, this is just another one of its properties, expressed by

$$Rexists(E)$$

Existential quantification in this approach is over a Platonic universe of possible (or impossible) individuals, that may or may not exist in the real world.

The relationship between primed and unprimed predicates is captured by the following axiom schema:

(3) $(\forall\, x)[p(x) \equiv (\exists\, e)[p'(e,x) \wedge Rexists(e)]]$

That is, p is true of x if and only if there is an eventuality e that is the eventuality of p being true of x and e exists in the real world. In fact, whenever in this paper the notation $p(x)$ is used, it should be viewed as an abbreviation for the right side of the biconditional.

Those desiring to use model theory to strengthen their intuitions about eventualities can think of the denotation of E in $p'(E, X)$ as the ordered triple of the intension of p, the denotation of X, and an integer serving as an index. The function of the index is to allow multiple events with the same predicates and arguments. There will normally be many events of John's working.

The approach to quantifiers taken in ITQ and the present paper is motivated by two considerations. The first is the desire to treat quantifiers in the same way that every other morpheme in the language is treated, in accordance with a principle that might be stated

All morphemes are created equal.

Every morpheme in English conveys information, and this information can be encoded in the form of a proposition consisting of a predicate applied to one or more arguments. One aim of ITQ and the present paper is to show this is as possible for quantifiers as it is for every other morpheme.

Another consequence of this principle, by the way, is that in the OP approach virtually every morpheme has a corresponding eventuality, even conjunctions. Thus,

John works and George sleeps.

has the logical form

$$(\exists\, e, e_1, e_2)[Rexists(e) \wedge and'(e, e_1, e_2) \wedge work'(e_1, J) \\ \wedge\, sleep'(e_2, G)]$$

That is, the eventuality e that both e_1 and e_2 hold holds where e_1 is John's working and e_2 is George's sleeping. It will be convenient in this paper to use the abbreviation $e_1\&e_2$ to stand for the eventuality e such that $and'(e, e_1, e_2)$.

The second consideration motivating the approach to quantifiers is the need for scope-neutral representation of quantifiers. The sentence

> In most democratic countries most politicians can fool most of the people on almost every issue most of the time.

has 120 readings. Moreover, they are distinct, in that for any two readings one can find a model under which one is true and the other isn't. Yet when people hear this sentence, they have the impression they understand it. They do not compute the 120 possible readings and then choose the best among them. Rather, they use world knowledge to constrain some of the dependencies among quantified expressions and leave other dependencies unresolved. For example, for me, the sets of politicians and the sets of people depend on the country, but I have no view on whether or not the politicians outscope the people. A representation is needed that allows this underspecification of meaning.

In brief, the approach to quantifiers advocated in ITQ consisted of four elements:

1. Sets are individuals. Quantifiers are relations between sets.
2. Sets have typical elements. Ordinary elements inherit the properties of typical elements.
3. Functional dependencies are expressed as relations between typical elements.
4. Disambiguating scope is done by learning functional dependencies.

The first two elements of this approach are discussed in some detail here. The last two are orthogonal to our present purposes and are discussed only briefly below, but they are one of the foci of ITQ.

To begin with, if we accept sets as first-class objects, then a determiner like "most" can be viewed as expressing a relation between sets. The expression $most(s_2, s_1)$ says that set s_2 is a subset of s_1 consisting of more than half the elements of s_1. Then sentence (1) can be represented as follows:

(4) $(\exists\, s)[most(s, \{x \mid man(x)\}) \wedge (\forall\, y)[y \in s \supset work(y)]]$

That is, there is a set s that is most of the set of all men (i.e., it is a subset with more than half the elements), and for every entity y in s, y works.

We can unwind this into a flat notation by introducing two new predicates. The first is *typelt*, and it takes two arguments—a typical element of a set and the set itself. The expression

$$typelt(y, s)$$

says that y is the typical element of the set s. The precise nature of typical elements is discussed in Section 3, but for now they can be viewed as a kind of reified, universally quantified variable. (McCarthy (1977) suggests a similar approach.) I will write about typical elements as though each set had a unique typical element, although this property will not be required (except once) in this paper. The principal property that typical elements should have is that their properties should be inherited by the ordinary elements of the set. A first cut at expressing this property is the following axiom schema:

(5) $$(\forall\, x, s)[typelt(x, s) \\ \supset [p(x) \equiv (\forall\, y)[y \in s \supset p(y)]]]$$

That is, if x is the typical element of set s, then p is true of x if and only if p is true of every ordinary element y of s.

Two obvious problems with this rule are as follows:

1. Because of the Law of the Excluded Middle, it would seem that for any predicate p, either $p(x)$ or $\neg p(x)$ would be true of the typical element x. Then by (5), the elements of s could not differ on any properties. They all would inherit either p or $\neg p$ from x.

2. There is a question as to whether the typical element of a set is itself an element of the set. Both choices seem to lead to difficulties.

The solution to these problems is described briefly in Section 3 and at length in ITQ.

There is another problem that was not dealt with in ITQ. The statement of this rule is not quite right because of the flat notation we are using, and it must be complicated somewhat, as described in Section 4 below.

The fact that sets have typical elements is captured by the axiom

$$(\forall\, s)[set(s) \supset (\exists\, x)typelt(x, s)]$$

The second new predicate, *dset*, is more specific than *typelt* in that it relates not only a set and its typical element, but also its defining condition. It takes three arguments—a set, its typical element, and the defining condition of the set. The expression

$$dset(s, x, e)$$

says that s is a defined set whose typical element is x and whose defining condition is the eventuality e. If e is, for example, the eventuality of x's being a man

$$man'(e, x)$$

then s is a defined set whose typical element is x and whose defining condition is the eventuality e of x's being a man, or the set of men. Thus, the expression

(6) $(\exists\, s, x, e) dset(s, x, e) \;\wedge\; man'(e, x)$

is equivalent to the more conventional expression

(7) $(\exists\, s)\; s = \{x \mid man(x)\}$

The principal property we need for the predicate *dset* is expressed, at a first cut, in the following axiom schema:

(8) $(\forall\, s)[[(\exists\, x, e) dset(s, x, e) \;\wedge\; p'(e, x)]$
$\qquad \equiv\; (\forall\, y)[y \in s \;\equiv\; p(y)]]$

That is, s is the defined set whose typical element is x and whose defining condition is the eventuality e of p being true of x if and only if for all y, y is in the set if and only if p is true of y. Again, this encoding will have to be revised in Sections 3 and 4, but modulo this revision, Axiom Schema (8) implies the equivalence of (6) and (7), since the left side of the outer biconditional in (8) is equivalent to (7).

The relation between the predicates *dset* and *typelt* is expressed in the following axiom:

(9) $(\forall\, s, x, e)[dset(s, x, e) \;\supset\; typelt(x, s)]$

If s is the defined set whose typical element is x and whose defining condition is the eventuality e, then x is the typical element of s. The predicate *dset* is thus a specialization of the predicate *typelt*, a fact that will play an important role in the treatment of monotone decreasing quantifiers.

There should probably not be a rule of the form

$$(\forall\, x, s)[typelt(x, s) \supset (\exists\, e)dset(s, x, e)]$$

since this would entail that every set is definable by some eventuality. This strikes me as an undesirable property. Some linguistically described sets, such as the set referred to by "all men", have natural defining properties. Others, such as the set referred to by "many men", do not, but they are sets nevertheless. In any case, we will not need this property.

The predicates corresponding to quantifiers, such as *most* and few, will be viewed as expressing relations (e.g., comparing cardinalities) between two sets. The principal properties of specific quantifiers can be stated as axioms. For example, one property of "few" and "most" is that they pick out subsets:

(10) $(\forall\, s_1, s_2)[most(s_2, s_1) \supset subset(s_2, s_1)]$
(11) $(\forall\, s_1, s_2)[few(s_2, s_1) \supset subset(s_2, s_1)]$

The monotone increasing and monotone decreasing properties can also be expressed as axioms:

(12) $(\forall\, s_1, s_2)most(s_2, s_1) \wedge subset(s_2, s) \wedge subset(s, s_1)$
$\supset most(s, s_1)$
(13) $(\forall\, s_1, s_2)few(s_2, s_1) \wedge subset(s, s_2) \wedge \neg null(s)$
$\supset few(s, s_1)$

That is, if s_2 is most of s_1 and s_2 is a subset of s which in turn is a subset of s_1, then s is also most of s_1. This is the monotone increasing property. If s_2 constitutes few members of s_1, then so does a non-null subset s of s_2. This is the monotone decreasing property.

Further axioms specify that the arguments of *subset* are both sets.

$$(\forall\, s_1, s_2)subset(s_2, s_1) \supset set(s_1)$$
$$(\forall\, s_1, s_2)subset(s_2, s_1) \supset set(s_2)$$

With this machinery, we can now rewrite logical form (4) as follows:

(14) $(\exists\, s_2, s_1, x, e, x, y)[most(s_2, s_1) \wedge dset(s_1, x, e)$
$\wedge\, man'(e, x) \wedge typelt(y, s_2) \wedge work(y)]$

That is, there is a set s_1 defined by the property e of its typical element x being a man, there is a set s_2 which is most of s_1 and has y as its typical element, and y works. Accepting Axiom Schemas (5) and (8) as written, it is straightforward to show that (14) is equivalent to (4).

Although this property will not be required in this paper, it is easy to see that distinct sets must have distinct typical elements.

It is easy to see how a logical form like (14) could be generated compositionally in a strictly local fashion. The common noun "men" introduces a set, its typical element, and its defining property, generating the conjuncts $dset(s_1, x, e) \wedge man'(e, x)$. The determiner "most" introduces another set and its typical element, along with the conjuncts $most(s_2, s_1) \wedge typelt(y, s_2)$. The latter typical element becomes the logical subject of the predication of the main verb, which generates the conjunct $work(y)$.

The logical form for

Most men like several women.

is

$$(\exists\, s_2, s_1, x, e, x, y, z, s_3)[most(s_2, s_1) \wedge dset(s_1, x, e) \wedge man'(e, x) \wedge typelt(y, s_2) \wedge like(y, z) \wedge several(s_3) \wedge typelt(z, s_3) \wedge woman(z)]$$

That is, there is a set s_1 defined by the property e of its typical element x being a man, there is a set s_2 which is most of s_1 and has y as its typical element, and y likes z, where z is the typical element of a set s_3, z is a woman, and s_3 has several members.

This is a scope-neutral representation. In the course of further processing, we may discover that s_3 is an *actual* set of several women, corresponding to wide scope for "several", or we may discover that s_3 is functionally dependent upon s_2, in which case s_3 is the typical element of a set of sets of women, one for each man in s_2, corresponding to the narrow scope.

This treatment of functional dependencies is elaborated on in ITQ, and is similar to the ordering constraints of Allen (1987) and Poesio (1991).

Section 3, sketchily, and Section 4, more thoroughly, discuss two complications that arise in this approach. The aim of the complications, however, is to bring us back to the original simplicity of notation.

3 The Nature of Typical Elements

There are three ways one might try to view typical elements:

1. The typical element of a set is one of the ordinary elements, but we will never know which one, so that anything we learn about it will be true of all.

2. The typical element is not an element of the set, and only special kinds of predicates are true of typical elements.

3. The typical element is not an element of the set, and ordinary predicates are true of them, except in set-theoretic axioms, which must be formulated carefully.

The first alternative is similar to the stance one takes toward instantiations of universally quantified variables in proofs. In proving $(\forall\, x \in s)p(x)$, one might consider an element a of s and show $p(a)$ while relying only on properties of a that are true for all elements of s. This alternative seems dangerous, however. The set consisting of John and George would have as its typical element either John or George, so by the desired properties (5) and (8), any property one has the other has too. The variable a in the proof is used only in a very limited context and in a very constrained way, whereas we want typical elements to exist in a persistent fashion in the Platonic universe and sometimes in the real world as well.

The second approach was taken in ITQ. The problem that arises when the typical element is assumed to be something other than an element of a set is that if the property p in Axiom Schema (5) is taken to be $\lambda x[x \notin s]$, then we can conclude that none of the members of the set are members of the set. I worked around this difficulty in ITQ by introducing a complex scheme of indexing predicates according to the kinds of arguments they would take. Essentially, for every predicate p, there was a basic level predicate p_0 that applied to ordinary individuals that are not typical elements, and a number of other predicates p_s that applied to the typical element of set s. More precisely, if x is the typical element of s, then $p_s(x)$ was defined to be true if and only if $p(y)$ was true for every y in s, and otherwise p_s was equivalent to p_0.

Axiom (5) can then be stated

$$(\forall\, x, s)[typelt(x, s) \\ \supset [p_s(x) \equiv (\forall\, y)[y \in_0 s \supset p_0(y)]]]$$

This solves the first difficulty with formulation (5). It is true that either $p_s(x)$ or $\neg(p_s)(x)$ holds, but this does not imply that all elements of s have all the same properties. That would hold only if either $p_s(x)$ or $(\neg p)_s(x)$ were true, but this is not what the Law of the Excluded Middle entails. The difference is the same as the difference between having negation outscope universal quantification and having universal quantification outscope negation.

The second difficulty with formulation (5) is solved as well. Suppose x is the typical element of s. We can simply stipulate that $x \notin_0 s$, and since this is a basic level rather than an indexed predicate, no consequences follow for real elements. To determine whether $x \in_s s$ is true, by the indexed version of Axiom (5), we have to ask whether

$$(\forall y)[y \in_0 s \supset y \in_0 s]$$

and this of course is trivially true. So $x \in_s s$ is true.

This solution is inconvenient, however, because it forces us to carry around complex indices in many contexts where they are irrelevant to the content being expressed. For example, the axiom

$$(\forall x)[man(x) \supset person(x)]$$

is true regardless of whether x is an ordinary individual or a typical element of a set. If all the members of a set are men, they are all persons. We would not like to have to specify indices in such axioms, and most axioms are exactly of this nature.

The primary place where the indices must be attended to is in set theoretic axioms. If x is the typical element of s, then $x \notin_0 s$ but $x \in_s s$. Thus, axioms that depend crucially on whether an entity is or is not in a set must be stated in terms of indexed predicates.

This leads to the third alternative, which we will adopt. We can avoid the complexity of indices by considering a bit how they are actually used in discourse processing. One must reintroduce the unindexed predicate p to use in the logical form of sentences, before interpretation, that is, before quantifier scope ambiguities are resolved. The relation between the indexed and unindexed predicates can be expressed, *inter alia*, by the following axiom schemas:

$$(\forall x)[p_0(x) \supset p(x)]$$
$$(\forall s, x)[p_s(x) \supset p(x)]$$

That is, the indexed predicates are specializations or strengthenings of the unindexed predicates, and in the course of discourse interpretation by abduction, one of the things that happens is that, as the existentially quantified variables are resolved to ordinary entities or to typical elements, the predicates that apply to them are specialized to the corresponding indexed predicate.

In this context of use, the indexing of the predicate is uniquely determined by the nature of its arguments. This would hold if constraints such as the following were stipulated:

$$\begin{aligned}&(\forall\, x, s)[p(x) \wedge \mathit{typelt}(x, s)\\ &\qquad \supset [p_s(x) \wedge \neg p_0(x) \wedge (\forall\, s_1)[s_1 \neq s \supset \neg p_{s_1}(x)]]]\end{aligned}$$

That is, if p is true of the typical element x of a set s, then the specialization p_s of p is true of x, and no other indexing of p is true of x.

A more thorough development of this idea depends on a treatment of functional dependencies, and therefore is beyond the scope of this paper.

It is worth noting that the consistency of the formulation I have given of typical elements can be demonstrated by taking as a model one in which the denotation of the typical element of a set is the set itself. In this case, *typelt* is simply identity. However, I wish to admit as well interpretations in which the set and its typical element are distinct, since there are a number of contexts in which this distinction is a useful one to make, including representing the difference between collective and distributive readings.

For the remainder of this paper, only the unindexed predicates are used.

4 Substitution

As noted above, there is a problem with the statement of Axiom Schemas (5) and (8) that arises because what in more conventional logical notations is represented via embedding gets strung out in the OP notation. Consider

John believes men work.

The logical form of this sentence is

(15) $$\begin{aligned}&(\exists\, e_1, m, s, e_2)\mathit{believe}(J, e_1) \wedge \mathit{work}'(e_1, m)\\ &\qquad \wedge\, \mathit{dset}(s, m, e_2) \wedge \mathit{man}'(e_2, m)\end{aligned}$$

That is, John believes the eventuality e_1 to obtain where e_1 is the eventuality of m working, where m is the typical element of a set s whose defining property is the eventuality e_2 of m's being a man.

Suppose John believes George is a man and thus in the set s. We would like to conclude that John believes George works. But this does not follow from Axiom Schemas (5) and (8). The entity m is the typical element of s, John believes m works, and so John should believe that George works. The predication $p(x)$ in Axiom (5) would have to be "John believes m works". If p is restricted to be an atomic predicate, this won't do, because "John believes m works" is not represented by an atomic predicate. Suppose p can be an arbitrary lambda expression. Then given that m is the typical element of s, Axiom (5) implies that any property of m must also hold of G, specifically, for the property

$$\lambda m[believe(J, e_1) \wedge work'(e_1, m)]$$

Thus it would follow from (15) that

$$believe(J, e_1) \wedge work'(e_1, G)$$

But this is the wrong result. The problem is that e_1 is the eventuality of men working, not the distinct eventuality of George's working. If Sam is also a man, then this approach leads to e_1's also being the eventuality of Sam's working.

To get around this difficulty, we can introduce a predicate *Subst* that expresses substitution relations among expressions directly. In a way, it mimics in the flat notation what substitution does in conventional notations, and one may thus suspect it is just a formal trick. However, I think that substitution itself is one particular formalization of an intuitive, commonsense concept—that of "playing the same role". $Subst(a, e_1, b, e_2)$ can be read as saying that a plays the same role in e_1 that b plays in e_2. (*Subst* differs from "playing the same role" in one aspect noted below.) Viewing it in this way, one need feel no compunction about applying the predicate to entities other than reified, universally quantified variables or typical elements. For example, if

$$work'(e_2, G) \wedge work'(e_3, S)$$

then

$$Subst(G, e_2, S, e_3)$$

since George plays the same role in George's working that Sam plays in Sam's working.

Subst turns out to be a useful concept in discourse interpretation wherever the similarity of two entities must be established.

In conventional notations, the first important property of substitution is the following:

$$p(t_1, \ldots, t_n)|_b^a = p(t_1|_b^a, \ldots, t_n|_b^a)$$

That is, the substitution of a predicate applied to a number of terms is the predicate applied to the substitution of the terms.

We can remain maximally noncommittal about the identity conditions among eventualities if we translate this schema into the following four axiom schemas, where p is now restricted to atomic predicates.

(16) $$\begin{aligned}&(\forall\, a, b, e_1, e_2, \ldots, u_i, \ldots)[Subst(a, e_1, b, e_2)\\ &\qquad\qquad \wedge\, p'(e_1, \ldots, u_i, \ldots)\\ &\qquad \supset\ (\exists\ \ldots, v_i, \ldots)[p'(e_2, \ldots, v_i, \ldots)\\ &\qquad\qquad \wedge\ \ldots\ \wedge\ Subst(a, u_i, b, v_i)\ \wedge\ \ldots]]\end{aligned}$$

This says that if a plays the same role in e_1 that b plays in e_2, p is the predicate of e_1, and the arguments of e_1 are u_i, then e_2 also is an eventuality with predicate p and arguments v_i where a plays the same role in each u_i that b plays in the corresponding v_i. This allows us to proceed in substitution from predications to their arguments.

(17) $$\begin{aligned}&(\forall\, a, b, e_1, \ldots, u_i, v_i, \ldots)[\ldots\ \wedge\ Subst(a, u_i, b, v_i)\ \wedge\ \ldots\\ &\qquad\qquad \wedge\, p'(e_1, \ldots, u_i, \ldots)\\ &\qquad \supset\ (\exists\, e_2)[p'(e_2, \ldots, v_i, \ldots)\ \wedge\ Subst(a, e_1, b, e_2)]]\end{aligned}$$

This says that if e_1 is an eventuality with predicate p and arguments u_i, where a plays the same role in each u_i that b plays in a corresponding v_i, then there is an eventuality e_2 whose predicate is p and whose arguments are v_i and a plays the same role in e_1 that b plays in e_2. This allows us to proceed from arguments to predications involving the arguments.

Two more axiom schemas are required because eventualities are not necessarily uniquely determined by their predicates and arguments. $p'(E_1, X)$ and $p'(E_2, X)$ can both be true without E_1 being identical to E_2. Axiom Schemas (16) and (17) guarantee a "substitution" eventuality of the right structure. The next two axiom schemas say that an eventuality is of the right structure if and only if it is a substitution eventuality.

$$
\begin{aligned}
(18)\quad & (\forall\, a, b, e_1, e_2, \ldots, u_i, v_i, \ldots)[\mathit{Subst}(a, e_1, b, e_2) \\
& \qquad\qquad \wedge\, p'(e_1, \ldots, u_i, \ldots) \\
& \quad \supset [p'(e_2, \ldots, v_i, \ldots) \\
& \qquad\qquad \equiv \ldots \wedge \mathit{Subst}(a, u_i, b, v_i) \wedge \ldots]]
\end{aligned}
$$

This says that if a plays the same role in e_1 that b plays in e_2, p is the predicate of e_1, and the arguments of e_1 are u_i, then e_2 also is an eventuality with predicate p and arguments v_i if and only if a plays the same role in each u_i that b plays in the corresponding v_i.

$$
\begin{aligned}
(19)\quad & (\forall\, a, b, e_1, e_2, \ldots, u_i, v_i, \ldots)[\ldots \wedge \mathit{Subst}(a, u_i, b, v_i) \wedge \ldots \\
& \qquad\qquad \wedge\, p'(e_1, \ldots, u_i, \ldots) \\
& \quad \supset [p'(e_2, \ldots, v_i, \ldots) \equiv \mathit{Subst}(a, e_1, b, e_2)]]
\end{aligned}
$$

This says that if e_1 is an eventuality with predicate p and arguments u_i, where a plays the same role in each u_i that b plays in a corresponding v_i, then the eventuality e_2 has predicate p and arguments v_i if and only if a plays the same role in e_1 that b plays in e_2.

The next two axioms enable substitution to bottom out.

$$(20)\quad (\forall\, a, b)\mathit{Subst}(a, a, b, b)$$

That is, a plays the same role in a that b plays in b.

$$(21)\quad (\forall\, a, b, c)\neg \mathit{eventuality}(c) \wedge c \neq a \supset \mathit{Subst}(a, c, b, c)$$

That is, if c is not an eventuality and not equal to a, then a plays the same role in c that b plays in c.

Notice that Axiom (21) allows c to be b. Substituting b for a in b results in b. This is the one asymmetry in the *Subst* predicate, and the reason that *Subst* is really more like substitution than like playing the same role. This asymmetry will allow us to draw from the fact that everyone in a set including John likes John the conclusion that John likes himself. That is, from $\mathit{typelt}(x, s)$, $p(x, y)$, and $y \in s$, we can conclude $p(y, y)$. The one constraint on *Subst* is that the first and fourth arguments cannot be the same. Substitution for the first argument would have eliminated such occurrences.

(22) $(\forall\, a, b, t_1, t_2)[a \neq b \wedge Subst(a, t_1, b, t_2) \supset a \neq t_2]$

That is, substituting b for a will never result in a.

Axiom Schemas (5) and (8) can now be recast as Axioms (23) and (24), respectively.

(23) $$(\forall\, x, s, e)[typelt(x, s) \supset [(\exists\, e_1)[Subst(x, e, x, e_1) \wedge Rexists(e_1)] \equiv (\forall\, y)[y \in s \supset (\exists\, e_2)[Subst(x, e, y, e_2) \wedge Rexists(e_2)]]]]$$

This property is now expressed as an axiom rather than an axiom schema. The explicit specification of the structure $p'(e, x)$ has been eliminated here. Instead, the eventuality e represents that pattern and the predicate *Subst* is used to stipulate that other eventualities exhibit the same pattern. This axiom says that if e is such a pattern and x is the typical element of s, then there is a really existing eventuality e_1 exhibiting that pattern if and only if for every ordinary element of s, there is a corresponding eventuality e_2 exhibiting the same pattern that really exists.

Suppose, in (23), that x is the typical element of s. If e is not an eventuality, then it is either x or something else. If it is x, then $e = e_1 = x$ and $e_2 = y$, so the axiom is valid. If it is something else, then $e = e_1 = e_2$, and the axiom is valid. Suppose e is an eventuality and $p'(e, x)$ holds. Then $p(x)$ is equivalent to $(\exists\, e_1)p'(e_1, x) \wedge Rexists(e_1)$, which is equivalent to $(\exists\, e_1)Subst(x, e, x, e_1) \wedge Rexists(e_1)$. Similarly, $p(y)$ is equivalent to $(\exists\, e_2)Subst(x, e, y, e_2) \wedge Rexists(e_2)$. Thus, Axiom (23) captures the intent of Axiom (5).

Replacing Axiom (8) is Axiom (24):

(24) $$(\forall\, s, x, e)[eventuality(e) \supset [(\exists\, e_1)[dset(s, x, e_1) \wedge Subst(x, e, x, e_1)] \equiv (\forall\, y)[y \in s \equiv (\exists\, e_2)[Subst(x, e, y, e_2) \wedge Rexists(e_2)]]]]$$

That is, if e is an eventuality (representing a pattern expressed in terms of the typical element x of a set s), then there is an eventuality e_1 of the same pattern that is the defining eventuality for s if and only if for every ordinary element y of s there is a corresponding eventuality e_2 of the same pattern that really exists. Here it is necessary to express the constraint that e be an eventuality, because the third argument of *dset* must be an eventuality.

Let us return to (15). If $dset(s, m, e_2)$ and $man'(e_2, m)$ hold and George is a man, then we have

$$\begin{aligned} man(G) &\equiv man'(e_3, G) \wedge Rexists(e_3) && \text{(by 3)} \\ &\equiv Subst(m, e_2, G, e_3) \wedge Rexists(e_3) && \text{(by 19)} \\ &\equiv G \in s && \text{(by 24)} \end{aligned}$$

Suppose $Rexists(e_0)$, $believe'(e_0, J, e_1)$, and $work'(e_1, m)$ all hold. Since $typelt(m, s)$ holds, and letting e and e_1 in (23) both be e_0, there is, by (23), an e_4 such that

$$Subst(m, e_0, G, e_4) \wedge Rexists(e_4)$$

By (16) there is an e_5 such that

$$believe'(e_4, J, e_5) \wedge Subst(m, e_1, G, e_4) \wedge Rexists(e_4)$$

By (19),

$$believe'(e_4, J, e_5) \wedge work'(e_5, G) \wedge Rexists(e_4)$$

By (3),

$$believe(J, e_5) \wedge work'(e_5, G)$$

That is, John believes George works. (I ignore here the problem of what inferences it is legitimate to draw inside belief contexts. Think of this expression as saying that, merely by virtue of the fact that George is a man, John believes George, whoever he may be, works.)

5 Monotone Decreasing Quantifiers

Virtually every utterance describes a situation in a more general fashion than the speaker actually means to convey. If I say, "I went to Tokyo," you are likely to interpret this as saying that I *flew* to Tokyo, even though I did not specify the means of transportation, and I would expect you to interpret it in this way. Indexicality is one example of this phenomenon. If I say "He went to Tokyo," I am saying that a male person went to Tokyo, but my listener will generally use contextual information to arrive at a more specific interpretation. This observation is at the core of the IA framework. To interpret a sentence is to find the "best" proof of its logical form, together with the selectional constraints that predicates impose on their arguments, allowing for coercions to handle metonymy, making assumptions where necessary. In brief, we must find the best set of specific facts and assumptions that imply the generalities conveyed explicitly by the utterance.

The parts of the logical form that we are able to prove constitute the *given* information that provides the referential anchor for the sentence. The assumptions that we must make in order to interpret a sentence constitute the *new* information; this is what the sentence is asserting. Typically, information in the main verb is what is asserted, and information that is grammatically subordinated is given, or presupposed. But this is not necessarily the case. In

> An innocent man was convicted today.

the listener may already know that someone was convicted, and the new, asserted information is that the man was innocent. Similarly, in

> I have a sore throat.

you know I have a throat. The new information is that it is sore. Reinterpreting what is asserted by the sentence will be a key move in dealing with monotone decreasing quantifiers.

The solution to the problem of monotone decreasing quantifiers that I propose consists of three steps.

1. We first generate the logical form of the sentence exactly as we would for other quantifiers. For sentence (2), the logical form would be analogous to (14), namely,

$$(25) \quad (\exists\, s_1, s_2, x, y, e_1, e_2) few(s_2, s_1) \wedge dset(s_1, x, e_1) \\ \wedge\, man'(e_1, x) \wedge typelt(y, s_2) \wedge work'(e_2, y) \\ \wedge\, Rexist(e_2)$$

That is, there is a set s_1 defined by the property e_1 of its typical element x being a man, there is a set s_2 which is few of s_1 and has y as its typical element, and the eventuality e_2 of y's working exists in the real world. Note that all of this is true, as far as it goes; there is a set consisting of few men, and the members of this set work. It just doesn't go far enough, because it does not rule out a much larger set.

2. The next step is to specialize or strengthen the predication $typelt(y, s_2)$ to the more specific $dset(s_2, y, e_2\&e_3)$, by back-chaining on Axiom (9), and

instantiating the defining eventuality to the conjunction of the two eventualities we see in the sentence, or rather, in place of the eventuality e_1 the "substitution" eventuality e_3 such that

$$Subst(x, e_1, y, e_3)$$

That is, we have further specified the set s_2 to be not just some subset of s_1 that has few elements, but the subset defined by the conjunction of conditions e_3 and e_2, where

$$man'(e_3, y) \wedge work'(e_2, y)$$

which, by (24), is the set of men who work.

It is Axiom (9) that places this interpretation in the space of possible interpretations, but nothing so far guarantees that this is the interpretation that will be selected. I would like to suggest one way this could happen, without, however, denying other possible accounts.

To promote this particular strengthening, we can associate as a selectional constraint on the arguments of few the requirement that its first argument be a set with a defining property.

(26) $few(s_2, s_1)$: $(\exists\, y, e) dset(s_2, y, e)$

This requirement then becomes something that has to be proven in addition to the logical form to arrive at an interpretation. It forces us to look for an eventuality e that defines the set s_2. The three most readily available eventualities are those explicit in the sentence itself—e_1 (or rather, e_3), e_2 and the conjunction of the two. e_3 (being a man) is impossible as a defining condition for s_2 since it is the defining condition for s_1, of which s_2 is a proper subset. e_2 (working) is also impossible as a defining condition, since the members of s_2 are men, and more than just men work. That leaves the conjunction $e_2 \& e_3$. The set s_2 is the set of men who work.

3. The proposition $few(s_2, s_1)$ is taken to be the assertion of the sentence, rather than $work(y)$. That is, the sentence would be interpreted as saying

> The men who work are few.

Increasing the plausibility of this part of the analysis is the fact that it is hard to unstress the word "few" when it is functioning as a monotone decreasing quantifier, and high stress is an indication that the information conveyed by the morpheme is new.

To demonstrate that this approach goes through, under this formulation, I need to show that from "Few men work" we can indeed conclude "Few men work hard," (assuming anyone works hard) once these two sentences have been interpreted as in Steps 1-3, and assuming, for the sake of this paper, that we have an axiom

(27) $(\forall x) work\text{-}hard(x) \supset work(x)$

The logical form of "Few men work", generated in Step 1, is given in (25). By Step 2, $typelt(y, s_2)$ is strengthened to $dset(s_2, y, e_3\&e_2)$. Since this sentence is the premise, we assume the strengthened logical form is all true.

The logical form of the sentence "Few men work hard" is

$$\begin{aligned}&(\exists\, s_1, s_4, x, z, e_1, e_4) few(s_4, s_1) \wedge dset(s_1, x, e_1)\\ &\qquad \wedge man'(e_1, x) \wedge typelt(z, s_4) \wedge work\text{-}hard'(e_4, z)\\ &\qquad \wedge Rexist(e_4)\end{aligned}$$

By Step 2, we strengthen $typelt(z, s_4)$ to $dset(s_4, z, e_5\&e_4)$, where

$$Subst(x, e_1, z, e_5).$$

Step 3 tells us that what is asserted is $few(s_4, s_1)$, while the rest is presupposed. Thus, we assume the rest is true, and we must demonstrate $few(s_4, s_1)$.

The conclusion $few(s_4, s_1)$ will follow from Axiom (13) if we can demonstrate $few(s_2, s_1)$ and $subset(s_4, s_2)$. But $few(s_2, s_1)$ is part of the premise assumed above. To demonstrate that $subset(s_4, s_2)$ holds, we need to show that any member v of s_4 is also a member of s_2. We do this in three steps. First we show, using the premises $dset(s_4, z, e_5\&e_4)$, $man'(e_5, z)$ and $work\text{-}hard'(e_4, z)$, together with a rightward use of the inner biconditional in Axiom (24), that any member v of s_4 is a man and works hard. We then use Axiom (27) to conclude that v works. We then use Axiom (24) in a leftward direction to show that v is in the set s_2. This establishes that the monotone decreasing property of the word "few" is preserved in this formulation.

It would be good if the predicate few used in "Few men work" captured the same notion of few-ness that is expressed in "A few men work." I will only sketch a possible account in which this would be the case. Consider

A few men work.

The word "a" expresses a relationship between the entity referred to by the noun phrase and the description it provides; it says roughly that the entity is not uniquely identifiable in context solely on the basis of that description. The logical form of this sentence would be almost the same as (25). But we need first to introduce the eventuality e_0 corresponding to the few-ness relation between s_2 and s_1—$few'(e_0, s_2, s_1)$. Then to express the relation conveyed by the determiner "a" we add the predication $a(y, e_0\&e_3)$, saying that y is not uniquely identifiable in context on the basis of the properties e_0 and e_3. If we were to proceed in Step 2 as before and specialize $typelt(y, s_2)$ to $dset(s_2, y, e_2\&e_3)$, then we would have a contradiction, for the properties e_2 and e_3 would uniquely identify y as the typical element of the set defined by these properties (assuming sets have a unique typical element). The word "a" thus blocks this strengthening of "few", the eventuality e in (26) remains unresolved, and we are left with only the few relation between the sets s_2 and s_1.

It is often argued that one way of drawing the line between compositional semantics (Step 1) and pragmatics (Step 2) is to say that the results of compositional semantics are not defeasible whereas the results of pragmatics are. This would appear to be an argument against the approach suggested here, since the interpretation of "Few men work" as "The men who work are few" does not seem to be defeasible. But another force that strongly constrains likely interpretations is conventionalization. The IA account of discourse comprehension traces out a space of possible interpretations and provides a graded mechanism for choosing among them, given a context. But conventionalization picks out among the possible interpretations a particular interpretation of a given word, phrase, or grammatical structure. It collapses the space of possible interpretations to only the conventional interpretation. It thus eliminates the defeasibility one ordinarily associates with pragmatic processing.

An example of this, unrelated to quantifiers, involves "let's". This is a contraction of "let us". But the sentence "Let us go" could be said by two victims to a kidnapper, whereas the sentence "Let's go" would not be. The general meaning of "Let us go"—

Don't cause us not to go.

is, for the contraction, conventionally specialized to

Don't cause us (inclusive) not to go by not going yourself.

The favored interpretations of "few men" and "a few men" are no doubt conventionalized, even though they can be derived *de novo* according to the accounts given above.

"Only" could be viewed as a determiner, and as such it would be monotone decreasing. Its interpretation would be derived very much as that of "few", but differing in one crucial respect.

"Only" is indeed monotone decreasing, since "Only men work" entails "Only men work hard." But unlike "few" it is not conservative. The conservativity property can be illustrated as follows: The sentence "Few men are men who work" entails "Few men work," and "few" is hence conservative. By contrast, "Only men are men who work" does not entail "Only men work." In fact, the first is tautologically true, and the second is false. "Only" is hence nonconservative (cf. van Benthem, 1983). This means the process of interpreting "only" must differ at some point from the process of interpreting "few". In fact, it differs in Step 2.

The logical form of "Only men work" would parallel (25).

$$\begin{aligned}(\exists\, s_1, s_2, x, y, e_1, e_2) & only(s_2, s_1) \wedge dset(s_1, x, e_1) \\ & \wedge man'(e_1, x) \wedge typelt(y, s_2) \wedge work'(e_2, y) \\ & \wedge Rexist(e_2)\end{aligned}$$

Under this analysis, as before, "only" will be taken to express a relation between a set s_2 and the set s_1 of all men, and the noun phrase "only men" will be taken to refer to the set s_2 in the sense that it is the members of s_2 who work. Thus, Step 1 in the analysis of "only" does not differ from Step 1 in the analysis of "few".

In Step 2, however, the set s_2 is not specialized to the set of men who work. Rather it is specialized to the set of workers. That is, $typelt(y, s_2)$ is strengthened to $dset(s_2, y, e_2)$. The relation that *only* expresses between s_2 and s_1 is then simply the subset relation. The set of workers is a subset of the set of men. That is, only men work.

Step 3 is the same as for "few". The predication $only(s_2, s_1)$ is picked as the assertion of the sentence. That is, "Only men work" is interpreted as though it were "The set of workers is a subset of the set of men," or "All workers are men."

This is a limited account of the interpretation of "only" as a determiner. In fact, a proper account would encompass adverbial uses as well. My real view is that *only* is a predicate of three arguments—an entity or eventuality x, a scale s that has x as its lowest element, and a property that is true of x but not of the other, higher elements of s. In "John only walked", x is

John's walking, s is a scale of actions ordered, say, by energy requirements, and the property is the property of having John as an agent. When used as a determiner, x is the entity or set referred to by the NP, s is the set of subsets containing x and ordered by inclusion, and the property is the main predication of the sentence. In "Only men work", x is a set of men, s is the set of subsets of the relevant entities containing x, and the property is working. The sentence says that the members of x work, but the members of no larger set in s works. This implies that the workers are a subset of all men, the meaning of "only" assumed in the account above.

6 Conclusion

The interpretation of quantifiers is a complex area of semantics, and one's simple, elegant notions of how the information in sentences can be represented run up against difficulties as soon as quantifiers are considered. Everyone who examines quantifiers is obliged to introduce substantial complexities into their logical notations to accommodate them. My approach has been no exception. The appeal to eventualities in Section 2, the indexing of Section 3, and the treatment of substitution of Section 4 are all examples of these complexities. But whereas in most approaches to semantics, the logical notation *remains* complex, the whole aim of my detour into the complexities was to regain the original simplicity and elegance, and I believe this has been achieved. The logical form of a sentence is still an existentially quantified conjunction of atomic predications, roughly one for each morpheme in the sentence. Once such a logical form has been generated for the sentence, only one interpretation process is needed, namely, the abductive process of determining the facts and assumptions that will provide the most economic proof of that logical form.

Acknowledgments

I have profited from discussions about this work with Mark Gawron, David Israel, Bob Moore, Kees van Deemter, Ed Zalta, and especially Massimo Poesio. They of course do not necessarily endorse nor even condone the proposals made here. This material is based on work supported by the National Science Foundation and the Advanced Research Projects Agency under Grant IRI-9314961 (Integrated Techniques for Generation and Interpretation).

References

[1] Allen, James, 1987. *Natural Language Understanding*, Benjamin Cummings, Menlo Park, California.

[2] Barwise, Jon, and Robin Cooper, 1981. "Generalized Quantifiers and Natural Language", *Linguistics and Philosophy*, Vol. 4, No. 2, pp. 159-219.

[3] Hobbs, Jerry R., 1983. "An Improper Treatment of Quantification in Ordinary English", *Proceedings of the 21st Annual Meeting, Association for Computational Linguistics*, pp. 57-63. Cambridge, Massachusetts, June 1983.

[4] Hobbs, Jerry R. 1985. "Ontological Promiscuity." *Proceedings, 23rd Annual Meeting of the Association for Computational Linguistics*, pp. 61-69. Chicago, Illinois, July 1985.

[5] Hobbs, Jerry R., Mark Stickel, Douglas Appelt, and Paul Martin, 1993. "Interpretation as Abduction", *Artificial Intelligence*, Vol. 63, Nos. 1-2, pp. 69-142.

[6] McCarthy, John, 1977. "Epistemological Problems of Artificial Intellegence", *Proceedings, International Joint Conference on Artificial Intelligence*, pp. 1038-1044, Cambridge, Massachusetts, August 1977.

[7] Poesio, Massimo, 1991. "Scope Ambiguity and Inference", Technical Report 389, Computer Science Department, The University of Rochester, Rochester, New York, July 1991.

[8] van Benthem, Johan, 1983. "Determiners and Logic", *Linguistics and Philosophy*, Vol. 6, pp. 447-478.

[9] van Eijck, Jan, 1983. "Discourse Representation Theory and Plurality", in A.G.B. ter Meulen, ed., *Studies in Modeltheoretic Semantics*, Foris Publications, Dordrecht, Netherlands, pp. 85-106.