

AN IMPROPER TREATMENT OF QUANTIFICATION IN ORDINARY ENGLISH

Jerry R. Hobbs
SRI International
Menlo Park, California

1. The Problem

Consider the sentence

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

In the currently standard ways of representing quantification in logical form, this sentence has 120 different readings, or quantifier scopings. Moreover, they are truly distinct, in the sense that for any two readings, there is a model that satisfies one and not the other. With the standard logical forms produced by the syntactic and semantic translation components of current theoretical frameworks and implemented systems, it would seem that an inferencing component must process each of these 120 readings in turn in order to produce a best reading. Yet it is obvious that people do not entertain all 120 possibilities, and people really do understand the sentence. The problem is not just that inferencing is required for disambiguation. It is that people never do disambiguate completely. A single quantifier scoping is never chosen. (Van Lehn [1978] and Bobrow and Webber [1980] have also made this point.) In the currently standard logical notations, it is not clear how this vagueness can be represented.¹

What is needed is a logical form for such sentences that is neutral with respect to the various scoping possibilities. It should be a notation that can be used easily by an inferencing component. That is, it should be easy to define deductive operations on it, and the logical forms of typical sentences should not be unwieldy. Moreover, when the inferencing component discovers further information about dependencies among sets of entities, it should entail only a minor modification in the logical form, such as conjoining a new proposition, rather than a major restructuring. Finally, since the notion of "scope" is a powerful tool in semantic analysis, there should be a fairly transparent relationship between dependency information in the notation and standard representations of scope.

Three possible approaches are ruled out by these criteria.

1. Representing the sentence as a disjunction of the various readings. This is impossibly unwieldy.

¹ Many people feel that most sentences exhibit too few quantifier scope ambiguities for much effort to be devoted to this problem, but a casual inspection of several sentences from any text should convince almost everyone otherwise.

2. Using as the logical notation a triple consisting of an expression of the propositional content of the sentence, a store of quantifier structures (e.g., as in Cooper [1975], Woods [1978]), and a set of constraints on how the quantifier structures could be unstored. This would adequately capture the vagueness, but it is difficult to imagine defining inference procedures that would work on such an object. Indeed, Cooper did no inferencing; Woods did little and chose a default reading heuristically before doing so.

3. Using a set-theoretic notation like that of (1) below, pushing all the universal quantifiers to the outside and the existential quantifiers to the inside, and replacing the existentially quantified variables by Skolem functions of all the universally quantified variables. Then when inferencing discovers a nondependency, one of the arguments is dropped from one of the Skolem functions. One difficulty with this is that it yields representations that are too general, being satisfied by models that correspond to none of the possible intended interpretations. Moreover, in sentences in which one quantified noun phrase syntactically embeds another (what Woods [1978] calls "functional nesting"), as in

Every representative of a company arrived.

no representation that is neutral between the two is immediately apparent. With wide scope, "a company" is existential, with narrow scope it is universal, and a shift in commitment from one to the other would involve significant restructuring of the logical form.

The approach taken here uses the notion of the "typical element" of a set, to produce a flat logical form of conjoined atomic predications. A treatment has been worked out only for monotone increasing determiners; this is described in Section 2. In Section 3 some ideas about other determiners are discussed. An inferencing component, such as that explored in Hobbs [1976, 1980], capable of resolving coreference, doing coercions, and refining predicates, will be assumed (but not discussed). Thus, translating the quantifier scoping problem into one of those three processes will count as a solution for the purposes of this paper.

This problem has received little attention in linguistics and computational linguistics. Those who have investigated the processes by which a rich knowledge base is used in interpreting texts have largely ignored quantifier ambiguities. Those who have studied quantifiers have generally noted that inferencing is required for

disambiguation, without attempting to provide a notation that would accommodate this inferencing. There are some exceptions. Bobrow and Webber [1980] discuss many of the issues involved, but it is not entirely clear what their proposals are. The work of Webber [1978] and Mellish [1980] are discussed below.

2. Monotone Increasing Determiners

2.1. A Set-Theoretic Notation

Let us represent the pattern of a simple intransitive sentence with a quantifier as "Q Ps R". In "Most men work," Q = "most", P = "man", and R = "work". Q will be referred to as a determiner. A determiner Q is monotone increasing if and only if for any R1 and R2 such that the denotation of R1 is a subset of the denotation of R2, "Q Ps R1" implies "Q Ps R2" (Barwise and Cooper [1981]). For example, letting R1 = "work hard" and R2 = "work", since "most men work hard" implies "most men work," the determiner "most" is monotone increasing. Intuitively, making the verb phrase more general doesn't change the truth value. Other monotone increasing determiners are "every", "some", "many", "several", "any" and "a few". "No" and "few" are not.

Any noun phrase Q Ps with a monotone increasing determiner Q involves two sets, an intensionally defined set denoted by the noun phrase minus the determiner, the set of all Ps, and a nonconstructively specified set denoted by the entire noun phrase. The determiner Q can be viewed as expressing a relation between these two sets. Thus the sentence pattern Q Ps R can be represented as follows:

$$(1) (\exists s)(Q(s, \{x \mid P(x)\}) \& (\forall y)(y \in s \rightarrow R(y)))$$

That is, there is a set s which bears the relation Q to the set of all Ps, and R is true of every element of s. (Barwise and Cooper call s a "witness set".) "Most men work" would be represented

$$(\exists s)(\text{most}(s, \{x \mid \text{man}(x)\}) \& (\forall y)(y \in s \rightarrow \text{work}(y)))$$

For collective predicates such as "meet" and "agree", R would apply to the set rather than to each of its elements.

$$(\exists s) O(s, \{x \mid P(x)\}) \& R(s)$$

Sometimes with singular noun phrases and determiners like "a", "some" and "any" it will be more convenient to treat the determiner as a relation between a set and one of its elements.

$$(\exists y) Q(y, \{x \mid P(x)\}) \& R(y).$$

According to notation (1) there are two aspects to quantification. The first, which concerns a relation between two sets, is discussed in Section 2.2. The second aspect involves a

predication made about the elements of one of the sets. The approach taken here to this aspect of quantification is somewhat more radical, and depends on a view of semantics that might be called "ontological promiscuity". This is described briefly in Section 2.3. Then in Section 2.4 the scope-neutral representation is presented.

2.2. Determiners as Relations between Sets

Expressing determiners as relations between sets allows us to express as axioms in a knowledge base more refined properties of the determiners than can be captured by representing them in terms of the standard quantifiers.

First let us note that, with the proper definitions of "every" and "some",

$$(\forall s1, s2) \text{every}(s1, s2) \leftrightarrow s1 = s2 \\ (\forall x, s2) \text{some}(x, s2) \leftrightarrow x \in s2$$

formula (1) reduces to the standard notation. (This can be seen as explaining why the restriction is implicative in universal quantification and conjunctive in existential quantification.)

A meaning postulate for "most" that is perhaps too mathematical is

$$(\forall s1, s2) \text{most}(s1, s2) \rightarrow |s1| > 1/2 |s2|$$

Next, consider "any". Instead of trying to force an interpretation of "any" as a standard quantifier, let us take it to mean "a random element of".

$$(2) (\forall x, s) \text{any}(x, s) \rightarrow x = \text{random}(s),$$

where "random" is a function that returns a random element of a set. This means that the prototypical use of "any" is in sentences like

Pick any card.

Let me surround this with caveats. This can't be right, if for no other reason than that "any" is surely a more "primitive" notion in language than "random". Nevertheless, mathematics gives us firm intuitions about "random" and (2) may thus shed light on some linguistic facts.

Many of the linguistic facts about "any" can be subsumed under two broad characterizations:

1. It requires a "modal" or "nondefinite" context. For example, "John talks to any woman" must be interpreted dispositionally. If we adopt (2), we can see this as deriving from the nature of randomness. It simply does not make sense to say of an actual entity that it is random.

2. It normally acts as a universal quantifier outside the scope of the most immediate modal embedder. This is usually the most natural interpretation of "random".

Moreover, since "any" extracts a single element, we can make sense out of cases in which "any" fails to act like "every".

- I'll talk to anyone but only to one person.
- * I'll talk to everyone but only to one person.
 - John wants to marry any Swedish woman.
 - * John wants to marry every Swedish woman.

(The second pair is due to Moore [1973].)

This approach does not, however, seem to offer an especially convincing explanation as to why "any" functions in questions as an existential quantifier.

2.3. Ontological Promiscuity

Davidson [1967] proposed a treatment of action sentences in which events are treated as individuals. This facilitated the representation of sentences with adverbials. But virtually every predication that can be made in natural language can be modified adverbially, be specified as to time, function as a cause or effect of something else, constitute a belief, be nominalized, and be referred to pronominally. It is therefore convenient to extend Davidson's approach to all predications, an approach that might be called "ontological promiscuity". One abandons all ontological scruples. A similar approach is used in many AI systems.

We will use what might be called a "nominalization" operator $\hat{}$ for predicates. Corresponding to every n-ary predicate p there will be an n+1-ary predicate \hat{p} whose first argument can be thought of as a condition of p's being true of the subsequent arguments. Thus, if "see(J,B)" means that John sees Bill, "see $\hat{}$ (E,J,B)" will mean that E is John's seeing of Bill. For the purposes of this paper, we can consider that the primed and unprimed predicates are related by the following axiom schema:

$$(3) \quad (\forall x, e) \hat{p}(e, x) \rightarrow p(x) \\ (\forall x)(\exists e) p(x) \rightarrow \hat{p}(e, x)$$

It is beyond the scope of this paper to elaborate on the approach further, but it will be assumed, and taken to extremes, in the remainder of the paper. Let me illustrate the extremes to which it will be taken. Frequently we want to refer to the condition of two predicates p and q holding simultaneously of x. For this we will refer to the entity e such that

$$\text{and}^{\wedge}(e, e_1, e_2) \ \& \ p^{\wedge}(e_1, x) \ \& \ q^{\wedge}(e_2, x)$$

Here e_1 is the condition of p being true of x, e_2 is the condition of q being true of x, and e the condition of the conjunction being true.

2.4. The Scope-Neutral Representation

We will assume that a set has a typical element and that the logical form for a plural noun phrase will include reference to a set and its typical element.² The linguistic intuition

² Woods [1978] mentions something like this approach, but rejects it because difficulties that are worked out here would have to be worked out.

behind this idea is that one can use singular pronouns and definite noun phrases as anaphors for plurals. Definite and indefinite generics can also be understood as referring to the typical element of a set.

In the spirit of ontological promiscuity, we simply assume that typical elements of sets are things that exist, and encode in meaning postulates the necessary relations between a set's typical element and its real elements. This move amounts to reifying the universally quantified variable. The typical element of s will be referred to as $\mathcal{Z}(s)$.

There are two very nearly contradictory properties that typical elements must have. The first is the equivalent of universal instantiation; real elements should inherit the properties of the typical element. The second is that the typical element cannot itself be an element of the set, for that would lead to cardinality problems. The two together would imply the set has no elements.³

We could get around this problem by positing a special set of predicates that apply to typical elements and are systematically related to the predicates that apply to real elements. This idea should be rejected as being ad hoc, if aid did not come to us from an unexpected quarter -- the notion of "grain size".

When utterances predicate, it is normally at some degree of resolution, or "grain". At a fairly coarse grain, we might say that John is at the post office -- "at(J,PO)". At a more refined grain, we have to say that he is at the stamp window -- "at(J,SW)". We normally think of grain in terms of distance, but more generally we can move from entities at one grain to entities at a coarser grain by means of an arbitrary partition. Fine-grained entities in the same equivalence class are indistinguishable at the coarser grain.

Given a set S, consider the partition that collapses all elements of S into one element and leaves everything else unchanged. We can view the typical element of S as the set of real elements seen at this coarser grain -- a grain at which, precisely, the elements of the set are indistinguishable. Formally, we can define an operator σ which takes a set and a predicate as its arguments and produces what will be referred to as an "indexed predicate":

$$\sigma(s, p)(x) = \begin{cases} T, & \text{if } x = \mathcal{Z}(s) \ \& \ (\forall y \in s) p(y), \\ F, & \text{if } x = \mathcal{Z}(s) \ \& \ \neg(\forall y \in s) p(y), \\ p(x) & \text{otherwise.} \end{cases}$$

We will frequently abbreviate this " p_s ". Note that predicate indexing gets us out of the above

³ An alternative approach would be to say that the typical element is in fact one of the real elements of the set, but that we will never know which one, and that furthermore, we will never know about the typical element any property that is not true of all the elements. This approach runs into technical difficulties involving the empty set.

contradiction, for now " $\mathcal{Z}(s) \varepsilon_s s$ " is not only true but tautologous.

We are now in a position to state the properties typical elements should have. The first implements universal instantiation:

$$(4) (\forall s, y) p_s(\mathcal{Z}(s)) \ \& \ y \varepsilon s \rightarrow p(y)$$

$$(5) (\forall s)((\forall x \varepsilon s) p(x)) \rightarrow p_s(\mathcal{Z}(s))$$

That is, the properties of the typical element at the coarser grain are also the properties of the real elements at the finer grain, and the typical element has those properties that all the real elements have.

Note that while we can infer a property from set membership, we cannot infer set membership from a property. That is, the fact that p is true of a typical element of a set s and p is true of an entity y , does not imply that y is an element of s . After all, we will want "three men" to refer to a set, and to be able to infer from y 's being in the set the fact that y is a man. But we do not want to infer from y 's being a man that y is in the set. Nevertheless, we will need a notation for expressing this stronger relation among a set, a typical element, and a defining condition. In particular, we need it for representing "every man". Let us develop the notation from the standard notation for intensionally defined sets,

$$(6) s = \{x \mid p(x)\},$$

by performing a fairly straightforward, though ontologically promiscuous, syntactic translation on it. First, instead of viewing x as a universally quantified variable, let us treat it as the typical element of s . Next, as a way of getting a handle on " $p(x)$ ", we will use the nominalization operator " $\hat{\cdot}$ " to reify it, and refer to the condition e of p (or p_s) being true of the typical element x of s -- " $p_s(e, x)$ ". Expression (6) can then be translated into the following flat predicate-argument form:

$$(7) \text{set}(s, x, e) \ \& \ p_s(e, x)$$

This should be read as saying that s is a set whose typical element is x and which is defined by condition e , which is the condition of p (interpreted at the level of the typical element) being true of x . The two critical properties of the predicate "set" which make (7) equivalent to (6) are the following:

$$(8) (\forall s, x, e, y) \text{set}(s, x, e) \ \& \ p_s(e, x) \ \& \ p(y) \rightarrow y \varepsilon s$$

$$(9) (\forall s, x, e) \text{set}(s, x, e) \rightarrow x = \mathcal{Z}(s)$$

Axiom schema (8) tells us that if an entity y has the defining property p of the set s , then y is an element of s . Axiom (9), along with axiom schemas (4) and (3), tells us that an element of a set has the set's defining property.

With what we have, we can represent the distinction between the distributive and collective readings of a sentence like

(10) The men lifted the piano.

For the collective reading the representation would include "lift(m)" where m is the set of men. For the distributive reading, the representation would have "lift($\mathcal{Z}(m)$)", where $\mathcal{Z}(m)$ is the typical element of the set m . To represent the ambiguity of (10), we could use the device suggested in Hobbs [1982] for prepositional phrase and other ambiguities, and write "lift(x) & ($x=m$ v $x=\mathcal{Z}(m)$)".

This approach involves a more thorough use of typical elements than two previous approaches. Webber [1978] admitted both set and prototype (my typical element) interpretations of phrases like "each man" in order to have antecedents for both "they" and "he", but she maintained a distinction between the two. Essentially, she treated "each man" as ambiguous, whereas the present approach makes both the typical element and the set available for subsequent reference. Mellish [1980] uses typical elements strictly as an intermediate representation that must be resolved into more standard notation by the end of processing. He can do this because he is working in a task domain -- physics problems -- in which sets are not just finite but small, and vagueness as to their composition must be resolved. Webber did not attempt to use typical elements to derive a scope-neutral representation; Mellish did so only in a limited way.

Scope dependencies can now be represented as relations among typical elements. Consider the sentence

(11) Most men love several women,

under the reading in which there is a different set of women for each man. We can define a dependency function f which for each man returns the set of women whom that man loves.

$$f(m) = \{w \mid \text{woman}(w) \ \& \ \text{love}(m, w)\}$$

The relevant parts of the initial logical form, produced by a syntactic and semantic translation component, for sentence (11) will be

$$(12) \text{love}(\mathcal{Z}(m), \mathcal{Z}(w)) \ \& \ \text{most}(m, m1) \ \& \ \text{man1}(\mathcal{Z}(m1)) \ \& \ \text{several}(w) \ \& \ \text{woman1}(\mathcal{Z}(w))$$

where $m1$ is the set of all men, m the set of most of them referred to by the noun phrase "most men", and w the set referred to by the noun phrase "several women", and where "man1 = $\sigma(m1, \text{man})$ " and "woman1 = $\sigma(w, \text{woman})$ ". When the inferencing component discovers there is a different set w for each element of the set m , w can be viewed as referring to the typical element of this set of sets:

$$w = \mathcal{Z}(\{f(x) \mid x \varepsilon m\})$$

To eliminate the set notation, we can extend the definition of the dependency function to the typical element of m as follows:

$$f(\mathcal{T}(m)) = \mathcal{Z}(\{f(x) \mid x \in m\})$$

That is, f maps the typical element of a set into the typical element of the set of images under f of the elements of the set. From here on, we will consider all dependency functions so extended to the typical elements of their domains.

The identity " $w = f(\mathcal{T}(m))$ " now simultaneously encodes the scoping information and involves only existentially quantified variables denoting individuals in an (admittedly ontologically promiscuous) domain. Expressions like (12) are thus the scope-neutral representation, and scoping information is added by conjoining such identities.

Let us now consider several examples in which processes of interpretation result in the acquisition of scoping information. The first will involve interpretation against a small model. The second will make use of world knowledge, while the third illustrates the treatment of embedded quantifiers.

First the simple, and classic, example.

(13) Every man loves some woman.

The initial logical form for this sentence includes the following:

$$\text{love}(\mathcal{Z}(ms), w) \ \& \ \text{man}(\mathcal{Z}(ms)) \ \& \ \text{woman}(w)$$

where " $\text{love} = \sigma(ms, \lambda x[\text{love}(x, w)])$ " and " $\text{man} = \sigma(ms, \text{man})$ ". Figure 1 illustrates two small models of this sentence. M is the set of men $\{A, B\}$, W is the set of women $\{X, Y\}$, and the arrows signify love. Let us assume that the process of interpreting this sentence is just the process of identifying the existentially quantified variables ms and w and possibly coercing the predicates, in a way that makes the sentence true.⁴

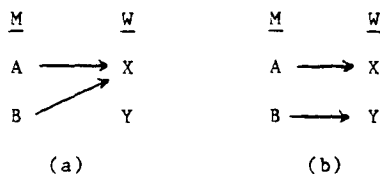


Figure 1. Two models of sentence (13).

In Figure 1(a), " $\text{love}(A, X)$ " and " $\text{love}(B, X)$ " are both true, so we can use axiom schema (5) to derive " $\text{love}(\mathcal{Z}(M), X)$ ". Thus, the identifications " $ms = M$ " and " $w = X$ " result in the sentence being true.

In Figure 1(b), " $\text{love}(A, X)$ " and " $\text{love}(B, Y)$ " are both true, but since these predications differ

⁴ Bobrow and Webber [1980] similarly show scoping information acquired by interpretation against a small model.

in more than one argument, we cannot apply axiom schema (5). First we define a dependency function f , mapping each man into a woman he loves, yielding " $\text{love}(A, f(A))$ " and " $\text{love}(B, f(B))$ ". We can now apply axiom schema (5) to derive " $\text{love}2(\mathcal{Z}(M), f(\mathcal{Z}(M)))$ ", where " $\text{love}2 = \sigma(M, \lambda x[\text{love}(x, f(x))])$ ". Thus, we can make the sentence true by identifying ms with M and w with $f(\mathcal{Z}(M))$, and by coercing "love" to "love2" and "woman" to " $\sigma(W, \text{woman})$ ".

In each case we see that the identification of w is equivalent to solving the scope ambiguity problem.

In our subsequent examples we will ignore the indexing on the predicates, until it must be mentioned in the case of embedded quantifiers.

Next consider an example in which world knowledge leads to disambiguation:

Three women had a baby.

Before inferencing, the scope-neutral representation is

$$\text{had}(\mathcal{Z}(ws), b) \ \& \ |ws|=3 \ \& \ \text{woman}(\mathcal{Z}(ws)) \ \& \ \text{baby}(b)$$

Let us suppose the inferencing component has axioms about the functionality of having a baby -- something like

$$(\forall x, y) \text{had}(x, y) \rightarrow x = \text{mother-of}(y)$$

and that we know about cardinality the fact that for any function g and set s ,

$$|g(s)| \leq |s|$$

Then we know the following:

$$3 = |ws| = |\text{mother-of}(b)| \leq |b|$$

This tells us that b cannot be an individual but must be the typical element of some set. Let f be a dependency function such that

$$w \in ws \ \& \ f(w) = x \rightarrow \text{had}(w, x)$$

that is, a function that maps each woman into some baby she had. Then we can identify b with $f(\mathcal{Z}(ws))$, or equivalently, with $\mathcal{Z}(\{f(w) \mid w \in ws\})$, giving us the correct scope.

Finally, let us return to interpretation with respect to small models to see how embedded quantifiers are represented. Consider

(14) Every representative of a company arrived.

The initial logical form includes

$$\text{arrive}(r) \ \& \ \text{set}(rs, r, ea) \ \& \ \text{and}^{\wedge}(ea, er, eo) \ \& \ \text{rep}^{\wedge}(er, r) \ \& \ \text{of}^{\wedge}(eo, r, c) \ \& \ \text{co}(c)$$

That is, r arrives, where r is the typical element of a set rs defined by the conjunction ea of r 's being a representative and r 's being of c , where c is a company. We will consider the two models in

Figure 2. R is the set of representatives {A,B,(C)}, K is the set of companies {X,Y,(Z,W)}, there is an arrow from the representatives to the companies they represent, and the representatives who arrived are circled.

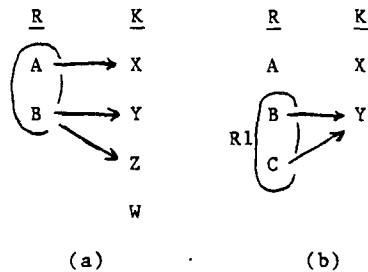


Figure 2. Two models of sentence (14).

In Figure 2(a), "of(A,X)", "of(B,Y)" and "of(B,Z)" are true. Define a dependency function f to map A into X and B into Y. Then "of(A,f(A))" and "of(B,f(B))" are both true, so that "of($\mathcal{Z}(R),f(\mathcal{Z}(R))$)" is also true. Thus we have the following identifications:

$$c = f(\mathcal{Z}(R)) = \mathcal{Z}(\{X,Y\}), rs = R, r = \mathcal{Z}(R)$$

In Figure 2(b) "of(B,Y)" and "of(C,Y)" are both true, so "of($\mathcal{Z}(R1),Y$)" is also. Thus we may let c be Y and rs be R1, giving us the wide reading for "a company".

In the case where no one represents any company and no one arrived, we can let c be anything and rs be the empty set. Since, by the definition of σ , any predicate indexed by the empty set will be true of the typical element of the empty set, "arrive($\mathcal{Z}(\emptyset)$)" will be true, and the sentence will be satisfied.

It is worth pointing out that this approach solves the problem of the classic "donkey sentences". If in sentence (14) we had had the verb phrase "hates it", then "it" would be resolved to c , and thus to whatever c was resolved to.

So far the notation of typical elements and dependency functions has been introduced; it has been shown how scope information can be represented by these means; and an example of inferential processing acquiring that scope information has been given. Now the precise relation of this notation to standard notation must be specified. This can be done by means of an algorithm that takes the inferential notation, together with an indication of which proposition is asserted by the sentence, and produces in the conventional form all of the readings consistent with the known dependency information.

First we must put the sentence into what will be called a "bracketed notation". We associate with each variable v an indication of the corresponding quantifier; this is determined from such pieces of the inferential logical form as those involving the predicates "set" and "most";

in the algorithm below it is referred to as "Quant(v)". The translation of the remainder of the inferential logical form into bracketed notation is best shown by example. For the sentence

A representative of every company saw a sample

the relevant parts of the inferential logical form are

$$\text{see}(r,s) \ \& \ \text{rep}(r) \ \& \ \text{of}(r,c) \ \& \ \text{co}(c) \ \& \ \text{sample}(s)$$

where "see(r,s)" is asserted. This is translated in a straightforward way into

$$(18) \text{see}(\{r \mid \text{rep}(r) \ \& \ \text{of}(r,[c \mid \text{co}(c)])\}, \{s \mid \text{sample}(s)\})$$

This may be read "An r such that r is a representative and r is of a c such that c is a company sees an s such that s is a sample.

The nondeterministic algorithm below generates all the scopings from the bracketed notation. The function TOPBVS returns a list of all the top-level bracketed variables in Form, that is, all the bracketed variables except those within the brackets of some other variable -- in (18) r and s but not c . BRANCH nondeterministically generates a separate process for each element in a list it is given as argument. A four-part notation is used for quantifiers (similar to that of Woods [1978]) -- "(quantifier variable restriction body)".

```
G(Form):
  if [v|R] ← BRANCH(TOPBVS(Form))
  then Form ← (Quant(v) v BRANCH({R,G(R)}) Formv)
  if Form is whole sentence
  then Return G(Form)
  else Return BRANCH({Form,G(Form)})
  else Return Form
```

In this algorithm the first BRANCH corresponds to the choice in ordering the top-level quantifiers. The variable chosen will get the narrowest scope. The second BRANCH corresponds to the decision of whether or not to give an embedded quantifier a wide reading. The choice R corresponds to a wide reading, G(R) to a narrow reading. The third BRANCH corresponds to the decision of how wide a reading to give to an embedded quantifier.

Dependency constraints can be built into this algorithm by restricting the elements of its argument that BRANCH can choose. If the variables x and y are at the same level and y is dependent on x , then the first BRANCH cannot choose x . If y is embedded under x and y is dependent on x , then the second BRANCH must choose G(R). In the third BRANCH, if any top-level bracketed variable in Form is dependent on any variable one level of recursion up, then G(Form) must be chosen.

A fuller explanation of this algorithm and several further examples of the use of this notation are given in a longer version of this paper.

3. Other Determiners

The approach of Section 2 will not work for monotone decreasing determiners, such as "few" and "no". Intuitively, the reason is that the sentences they occur in make statements about entities other than just those in the sets referred to by the noun phrase. Thus,

Few men work.

is more a negative statement about all but a few of the men than a positive statement about few of them. One possible representation would be similar to (1), but with the implication reversed.

$$(\exists s)(Q(s, \{x \mid P(x)\}) \& (\forall y)(P(y) \& R(y) \rightarrow y \in s))$$

This is unappealing, however, among other things, because the predicate P occurs twice, making the relation between sentences and logical forms less direct.

Another approach would take advantage of the above intuition about what monotone decreasing determiners convey.

$$(\exists s)(\bar{Q}(s, \{x \mid P(x)\}) \& (\forall y)(y \in s \rightarrow \neg R(y)))$$

That is, we convert the sentence into a negative assertion about the complement of the noun phrase, reducing this case to the monotone increasing case. For example, "few men work" would be represented as follows:

$$(\exists s)(\overline{\text{few}}(s, \{x \mid \text{man}(x)\}) \& (\forall y)(y \in s \rightarrow \neg \text{work}(y)))^5$$

(This formulation is equivalent to, but not identical with, Barwise and Cooper's [1981] witness set condition for monotone decreasing determiners.)

Some determiners are neither monotone increasing nor monotone decreasing, but Barwise and Cooper conjecture that it is a linguistic universal that all such determiners can be expressed as conjunctions of monotone determiners. For example, "exactly three" means "at least three and at most three". If this is true, then they all yield to the approach presented here. Moreover, because of redundancy, only two new conjuncts would be introduced by this method.

Acknowledgments

I have profited considerably in this research from discussions with Lauri Karttunen, Bob Moore, Fernando Pereira, Stan Rosenschein, and Stu Shieber, none of whom would necessarily agree with what I have written, nor even view it with sympathy. This research was supported by the Defense Advanced Research Projects Agency under Contract No. N00039-82-C-0571, by the National Library of Medicine under Grant No. 1R01 LM03611-

⁵ "few" is pronounced "few bar".

01, and by the National Science Foundation under Grant No. IST-8209346.

REFERENCES

- Barwise, J. and R. Cooper 1981. Generalized quantifiers and natural language. Linguistics and philosophy, Vol. 4, No. 2, 159-219.
- Bobrow, R. and B. Webber 1980. PSI-KLONE: Parsing and semantic interpretation in the BBN natural language understanding system. Proceedings, Third National Conference of Canadian Society for Computational Studies of Intelligence. 131-142. Victoria, British Columbia. May 1980.
- Cooper, R. 1975. Montague's semantic theory and transformational syntax. Ph.D. thesis. University of Massachusetts.
- Davidson, D. 1967. The logical form of action sentences. In N. Rescher (Ed.), The Logic of Decision and Action. 81-95. University of Pittsburgh Press, Pittsburgh, Pennsylvania.
- Hobbs, J. 1976. A computational approach to discourse analysis. Research Report 76-2, Department of Computer Sciences, City College, City University of New York.
- Hobbs, J. 1980. Selective inferencing. Proceedings, Third National Conference of Canadian Society for Computational Studies of Intelligence. 101-114. Victoria, British Columbia. May 1980.
- Hobbs, J. 1982. Representing ambiguity. Proceedings of the First West Coast Conference on Formal Linguistics. 15-28. Stanford, California.
- Mellish, C. 1980. Coping with uncertainty: Noun phrase interpretation and early semantic analysis. Ph.D. thesis. University of Edinburgh.
- Moore, R. 1973. Is there any reason to want lexical decomposition? Unpublished manuscript.
- Van Lehn, K. 1978. Determining the scope of English quantifiers. Massachusetts Institute of Technology Artificial Intelligence Laboratory Technical Report AI-TR-483.
- Webber, B. 1978. A formal approach to discourse anaphora. Technical Report 3761, Bolt Beranek and Newman, Inc. Cambridge, Massachusetts.
- Woods, W. 1977. Semantics and quantification in natural language question answering. Advances in Computers, Vol. 17. 1-87. Academic Press, New York.