Chapter 5

Distributivity

5.1 A Challenge to the Union Approach

As noted in chapter 3, sentences of the following type:

(84) The authors and the athletes are outnumbered by the men.

have been used to argue for a more complex theory of plural reference (cf. Landman 1989a; Scha and Stallard 1988). (84) is claimed to have a distributive reading that (86) lacks:

(86) The women are outnumbered by the men.

even in a context in which the authors and the athletes are just the women. It is argued that this difference must be captured in part by distinguishing the possible denotations assigned to the subjects of (84) and (86). This constitutes a challenge to the union theory, which, in the context assumed here, assigns these noun phrases the same denotation.

This challenge to the union theory relies on the view that the denotation of a distributively read verb phrase differs from that of its non-distributive counterpart (see Lasersohn 1995 for extensive discussion of that view and alternatives). Even for those who accept that view, as we shall here, the source of that difference in denotation remains open and that turns out to be a crucial issue for the debate between the union and the sets theory. It is usually assumed that distributivity is a purely semantic matter: a plural predicate has one meaning on its distributive reading and a different meaning on its non-distributive reading, and these meanings differ in such a way that in some situations the two readings lead to different extensions. However, there is another possibility. It could be that a plural predicate has a single meaning, but that that meaning is context-dependent, and will lead to different `readings’ in different contexts. On the purely semantic view, it makes sense to trace the differences between (84) and (86) to differences in the referents of their subjects. However, if this latter,
context-dependent alternative is correct, then a difference in the context-change potentials of the subjects in (84) and (86) may account for the fact that they do not share (in the sense of section 3.2) the distributively read verb phrase common to (84) and (86).

In this section I will, through successive attempts, arrive at an account of distributivity that has a context-dependent element to it. The data in (84) and (86) when analyzed on that account no longer pose a threat to the union theory.

My presentation here will depart slightly from the practice of chapter 1. The account will be cast, at least initially, in a framework in which English is translated into a semantically interpreted language. The purpose of this departure is to remain somewhat closer to existing accounts of distributivity upon which mine is based.

5.2 A Quantificational Account

5.2.1 Cumulativity

Our story begins with the oft cited connection between distributivity and cumulativity. For example, while (133) is gotten from (132) by cumulativity:

(132) John moved the car and Mary moved the car.
(133) John and Mary moved the car.

(132) follows from a distributive construal of (133). Our first guess then, is that by accounting for cumulativity we thereby account for distributivity. This leads to what Lasersohn (1988) calls a closure-condition account. By this we mean that all predicates of the language have a simple translation as well as a translation which is interpreted as the closure under union of the simple translation. Letting $\alpha$ represent a metavariable over predicates of English we have:

(134) $\alpha$ translates as: $\alpha'$ and as $^*(\alpha')$

$\| ^*(\alpha') \| = \text{the closure under union of } \| \alpha' \|$

We have employed here the "$*$" operator familiar from the work of G. Link, though with a slightly different semantics in terms of set union rather than lattice theoretical sum (see Appendix on closure under union of a set of individuals). The cumulative inference of (132)-(133) is now mapped as follows:

Distributivity

(135)

Assumptions just the star-of-\(\alpha\) condition

one way to distri-\(2\) predict the closure condition

(133)
(132)
(136)

Unfortu-

The previous

false one

false one
Distributivity

\[(135) \quad (\text{moved-the-car}'(\text{J}) \land \text{moved-the-car}'(\text{M})) \rightarrow ^*\text{(moved-the-car')}(\text{J} + \text{M})\]

Assuming that \(\text{J} + \text{M}\) is interpreted as the union of John and Mary (this is just the set of John and Mary, see the Appendix), the interpretation of the star-operator in (134) guarantees the inference in (135).

Turning now to distributivity, this setup allows us two translations, one with and one without a star, corresponding respectively to the distributive and the non-distributive or collective construals of a given predicate. Presumably, the inference from (133) on its distributive construal to (132) would be mapped as in (136):

\[(133) \quad \text{John and Mary moved the car.}\]
\[(132) \quad \text{John moved the car and Mary moved the car.}\]
\[(136) \quad ^*\text{(moved-the-car')}(\text{J} + \text{M}) \rightarrow [\text{moved-the-car}'(\text{J}) \land \text{moved-the-car}'(\text{M})]\]

Unfortunately, (136) is not guaranteed by (134) the way (135) was. To see this, consider a situation in which John and Mary moved the car together but neither John nor Mary moved the car individually. Since (133) is true, the set of John and Mary must be in the extension of either \(^*(\text{moved-the-car}')\) or \(\text{moved-the-car}'\). Since it is false that John moved the car, John is not in the extension of either \(^*(\text{moved-the-car}')\) or \(\text{moved-the-car}'\) and likewise for Mary. This means that the set of John and Mary can’t have gotten into the extension of \(^*(\text{moved-the-car}')\) by closure under union, hence it must be in the extension of \(\text{moved-the-car}'\) and then by definition it must also be in \(^*(\text{moved-the-car}')\). This means that in the situation described, the antecedent of (136) is true and the consequent is false.\(^{15}\) In other words, the arrow of (136) is not justified by adding (134) to our system.

The upshot of this result is that in a grammar where distributivity

\(^{15}\)One might think that the star-operator should be redefined as follows:

\[\| ^*\alpha' \| = \| \alpha' \| \text{ excluding the elements of } \| \alpha' \| \text{ itself.}\]

The problem is that this would incorrectly make:

i. John and Mary moved the car.

false on its distributive reading if they moved the car together, even if they also moved it individually.
is accounted for with a rule like (134), there is no translation of (133) from which (132) follows:

(133) John and Mary moved the car.
(132) John moved the car and Mary moved the car.

The prevailing view seems to be that (132) should follow from (133) on some reading (cf. Gillon 1987 and Lasersohn 1988:§2.1). In fact, it is only under this view that the argument presented at the outset against the union theory holds weight, since that argument relied on an entailment of this sort: The authors and the athletes are outnumbered by the men entails The authors are outnumbered by the men and the athletes are outnumbered by the men. While I am not entirely convinced that the prevailing view is correct, I will adopt it here and with it abandon our first attempt at an analysis of distributivity within the union theory.

Before moving on to the next attempt, let me note one more possible flaw in the current system, which was mentioned chapter 1. Theories which include something like (134) seem to overgenerate. They predict that cumulativity is independent of the predicates involved. However, I am uneasy with the following entailments:

(137) The boys look alike and the girls look alike → The boys and the girls look alike.
(138) The students left as a group and the teachers left as a group → The students and the teachers left as a group.

Lønning (1989:125) makes a similar point concerning this example:

(139) a. The black children played with each other and the white children played with each other.
b. The (black and the white) children played with each other.

5.2.2 The D-operator

The next attempt starts with the observation that distributivity can be overtly marked with the floated adverb each:

(140) John and Mary each moved the car.

(140) unambiguously entails that John moved the car and Mary did too. On the basis of this observation, one posits an adverbial D-operator in the translation language with the following semantics, where x,y are variables
Distributivity

over elements in the domain of discourse:

\[(141) \quad x \in \| D(\alpha) \| \iff \forall y[ (\text{singularity}(y) \land y \in x) \rightarrow y \in \| \alpha \| ] \]

An operator of this type is found in Link’s work as well as in Lønning (1987), Roberts (1987) and elsewhere.

The ambiguity of (133) is now captured by allowing the predicate to be translated either as in (142a) or as in (142b).

(133) John and Mary moved the car.
(142) a. moved-the-car'
b. D(moved-the-car')

The distributive entailment from (133) to (132):

(132) John moved the car and Mary moved the car.

is mapped as in (143):

\[(143) \quad D(\text{moved-the-car'})(J + M) \rightarrow [\text{moved-the-car'}(J) \land \text{moved-the-car'}(M)] \]

This entailment is guaranteed by the semantics given in (141). Here I’ve assumed that the simple translation moved-the-car’ denotes a set containing singularities and pluralities. Each member of the set is responsible for a moving of the car.

The introduction of a quantifier into the logical form of distributive predicates is further justified by evidence of scope interaction between it and other quantifiers. One example of this concerns the interaction between the D-operator and indefinite noun phrases as analyzed in Roberts (1987). To see this effect, consider first the simple example in (144) in a context where there is more than one boy:

(144) Every boy killed a dog.

This example has a plausible reading in which the existential has narrow scope with respect to the universal and an implausible reading involving multiple killings of a single dog. As is well-known, the singular indefinite can serve as the antecedent for a singular pronoun later in the discourse only if it has wide scope. Thus we get only the implausible reading when (144) is embedded in a discourse having such a pronoun, as in (145):

(145) Every boy killed a dog. It turned out to have nine lives.
Now consider the example in (146):

(146) John and Mary killed a dog.

This example has two distributive readings. On one reading, (146) is true if John killed a dog and Mary killed a dog. On the other, implausible, distributive reading, (146) is true if there is a dog and John killed it and Mary killed it. The presence of the two distributive readings is explained by taking the indefinite to have narrow or wide scope with respect to the D-operator. Once again, the implausible wide scope reading is the only one possible in a discourse where the indefinite serves as the antecedent for a subsequent singular pronoun:

(147) John and Mary killed a dog. It was buried in the parking lot.

((147) also has a collective reading, involving a single collaborative murder). Another example of scope interaction involving the D-operator was pointed out to me by Angelika Kratzer. In this case, the interaction is with the modal predicate likely:

(148) John and Mary are likely to win the lottery.

(148) has the following two distributive readings:

a. there is a good chance that John will win the lottery and that Mary will win the lottery.

b. John and Mary each have a good chance of winning the lottery.

This difference is explained by taking the D-operator to have scope under the modal in the (a) reading (attached to the lower verb) and over the modal in the (b) reading.

Reviewing so far, we have now helped ourselves to an account that fulfills the basic requirement of guaranteeing the distributive entailments and which analyzes the distributive-collective distinction as one of ambiguity. Furthermore, essential use of a quantifier in the translation of distributive predicates is independently confirmed by its participation in scope interactions.

One might wonder at this point whether or why this approach is an improvement on the approach taken for example by Bennett (1974:193,229) in which definite noun phrases were optionally translated with a universal quantifier, essentially giving a definite plural the meaning of a universal.

5.2.3 Interchangeability of D-operators

D-operators of a second order hold true if and only if they hold true of all entities that are members of the same distributive collective. For example, the sentence "John and Mary each have a good chance of winning the lottery" is equivalent to "John and Mary have a good chance of winning the lottery".

(149) John and Mary are likely to win the lottery.

we can assume a membership relation between the two members of a plurality, like:

(150) The two are likely to win the lottery.

when the quantifier is understood as essentially holding of no hitting collective. Fiengo and Langendoerfer (1975) are likely to win the lottery in an analogy:

\[ A \text{ has a sentiment for } B \]

\[ \{ A, B \} \text{ has a sentiment for } C \]
of a universal noun phrase. The present approach has an advantage. As noted in section 1.3, conjoined verb phrases need not be understood both collectively or both distributively even when they combine with a single subject. This is unexpected on an account in which the ambiguity is located in the noun phrase.

5.2.3 Intermediate Readings: Context Sensitivity

With all that is positive about the D-operator account of distributivity, one significant problem remains. By employing the D-operator, we envision two kinds of situations in which a verb phrase will hold true of a plurality: either the property expressed by that predicate holds of each of the singularities that are parts of the plurality, this is the distributive case, or the property holds of the plurality itself, this is the collective case. Research on plural constructions has uncovered a third case, however. In their work on reciprocals, Fiengo and Lasnik (1973) observed that whereas for simple cases like (149),

(149) John and Bill were hitting each other.

we can say that a reciprocal VP of the form V-each other is true of a plurality if, and only if, the relation expressed by V holds between any two members of the plurality reciprocally, such is not the case with an example like:

(150) The men were hitting each other.

when there are more than two men. If the men are divided into groups, and there is reciprocal hitting between any two members of each group, Fiengo and Lasnik say that (150) might be considered true even if there is no hitting between members of different groups. In other words, reciprocity holds within subpluralities of the plurality denoted by the men. Langendoen (1978) extended this idea of distributing down to sub-pluralities in an analysis of non-reciprocal sentences.16 Within this tradition,

---

16 A reviewer pointed out that Katz (1977:127) expresses a similar sentiment:

"the units of attribution can be individuals, pairs, triplets, and so on, up to the entire membership of the set DES(t;)[roughly, the denotation of the relevant argument of the attributed predicate RSS]. The frequently discussed notions
Higginbotham (1981:100) adopts the following interpretive principle:

\[(151) \quad [S \text{ NP}_{\text{plural}} \text{ VP}] \text{ is true iff there is a partition } C \text{ of the plurality } P \text{ denoted by NP such that VP is true for every element in } C.\]

A partition is a kind of cover, where:

\[(152) \quad C \text{ is a cover of } P \text{ if and only if:} \]
\[1. C \text{ is a set of subsets of } P \]
\[2. \text{ Every member of } P \text{ belongs to some set in } C.\]
\[3. \emptyset \text{ is not in } C.\]

\(C\) is a partition of \(P\) if, and only if, \(C\) covers \(P\) and no two members of \(C\) overlap. In fact, there is some question whether (151) shouldn’t make reference to covers of all types rather than just to partitions. Gillon (1987:212) provides the following example in support of this change:

\[(153) \quad \text{The men wrote musicals.}\]

Suppose the men denotes Rodgers, Hammerstein, and Hart. (153) is true, on at least one reading, when they are the denotation of the subject noun phrase. However, there is no partition of the set containing those three men in which wrote musicals is true of each member. Rather, the sentence is true because Rodgers and Hammerstein collaborated to write musicals and Rodgers and Hart also collaborated to write musicals.

On the strength of this example, we make the recommended change in (151):

\[(154) \quad [S \text{ NP}_{\text{plural}} \text{ VP}] \text{ is true iff there is a cover } C \text{ of the plurality } P \text{ denoted by NP such that VP is true for every element in } C.\]

If the claim in (154) is correct, then surely something is lacking in our analysis in terms of a collective reading and a D-operator based distributive

of the distributive and collective features of quantifiers represent two extremes of this range of possible units."

Incidentally, Katz rejects the notion that there is a genuine ambiguity here, a question to which we return below.
reading, where the distribution is to singularities only. But is the claim correct? Lasersohn (1989) doesn't think so. He asks us to consider a situation in which a department pays each of its three TAs (teaching assistants) $7,000. In such a case (155) is true on a distributive reading, (156) is true on a collective reading and (157) is false:

(155) The TAs were paid exactly $7,000 last year.
(156) The TAs were paid exactly $21,000 last year.
(157) The TAs were paid exactly $14,000 last year.

However, according to (154), (157) should be true as well since the VP in that sentence is true of any set of two TAs and hence will be true of each member of a cover of the TAs containing two two-membered sets. Lasersohn proposes instead that we remain with the simple two-way collective-distributive distinction. He reanalyzes Gillon's (153) by adopting a meaning postulate that guarantees the following:

(158) || write || (w,y) \land || write || (x,z) \rightarrow || write || (w \cup x, y \cup z)

Since the union of the set of Rodgers and Hammerstein and the set of Rodgers and Hart is just the set denoted by the men in (153), that sentence is guaranteed to be true on the construal Gillon is after.

Lasersohn's use of a meaning postulate to handle an "intermediate" distributive reading is in the spirit of Scha (1984) and Scha & Stallard (1988), to be discussed below. While I believe that the statement about the meaning of write made in (158) is likely correct, I think it is misleading to capture this information with a meaning postulate and, more importantly, it is incorrect to account for distributivity in strictly semantic terms. Both of these points deserve elaboration.

Accounting for the difference between (153) and (157) in terms of a meaning postulate amounts to claiming that the presence of the intermediate reading in (153) is a direct result of the presence of write as the main verb and that (157) lacks the intermediate reading because the verb write has been avoided. Observe, however, that (153) can be continued with:

(159) They were paid exactly $2,000 per musical/for their musicals.

If, in fact, the musicals went for $2,000 apiece (159) is true on the same intermediate reading, that is the one corresponding to the same cover as in (153). The main verb of (159) is similar to the one that was used to deny the presence in general of intermediate readings. This indicates that the
presence of the intermediate reading in (153) is not intimately connected with choice of main verb.

Now, regardless of how we capture facts about the extensions of plural predicates like that in (158), we will not achieve a complete analysis of distributivity. This is because there is a pragmatic element to distributivity which is nicely illustrated in Gillon’s reply to Lasersohn. Gillon guessed that the adverb *exactly* was responsible for the possible lack of the relevant reading in Lasersohn’s (157) and suggested we consider his (160) as well:

(160) The T.A.’s were paid their $14,000 last year.

in a context which he describes as follows:

A chemistry department has two teaching assistants for each of its courses, one for the recitation section and one for the lab section. The department has more than two teaching assistants and it has set aside $14,000 for each course with teaching assistants. The total amount of money disbursed for them, then, is greater than $14,000. At the same time, since the workload for teaching a course’s section can vary from one section to another, the department permits each team of assistants for a course to decide for itself how to divide the $14,000 the team is to receive. Suppose that it turns out, as it very well could under such circumstances, that no teaching assistant is paid exactly $14,000 . Yet it seems to me that either of the sentences in (157) or (160) could be truly affirmed, though neither sentence, by hypothesis, is true in virtue of either a collective or a distributive reading.

What Gillon has done here is to change the context in which Lasersohn’s example is uttered. In other words, whether or not a certain intermediate reading is available seems to have to do with the context not with the semantics of particular lexical items. The phenomenon we are looking at is pragmatic, not semantic. The claim I am making can be cast in terms of a revision of the generalization made in (154) above repeated here:

(154) [$s \text{ NP}_{\text{plural}} \text{ VP}] \text{ is true iff there is a cover } C \text{ of the plurality } P \text{ denoted by NP such that VP is true for every element in } C.$

(154) makes reference to covers of a plurality. Now, in different contexts,
Distributivity

different covers may be salient, that is, a given plurality may have parts that are relevant in one conversation but not in another. So (154) should be modified to:

(161) \[ \text{[NP\textsubscript{plural} VP]} \text{ is true in some context } Q \text{ iff there is a cover } C \text{ of the plurality } P \text{ denoted by NP which is salient in } Q \text{ and VP is true for every element in } C. \]

The examples we've seen so far in which an intermediate reading was claimed to have arisen have all involved transitive verbs with bare plural or amount denoting object noun phrases. Before turning to an analysis which captures the generalization in (161), I would like to provide an example of a different form in which an intermediate reading arises.

Imagine a situation in which two merchants are attempting to price some vegetables. The vegetables are sitting before the merchant, piled up in several baskets. To determine their price, the vegetables need to be weighed. Unfortunately, our merchants do not have an appropriate scale. Their grey retail scale is very fine and is meant to weigh only a few vegetables at a time. Their black wholesale scale is coarse, meant to weigh small truckloads. Realizing this, one of the merchants truthfully says:

(162) The vegetables are too heavy for the grey scale and too light for the black scale.

In order to save space in our explanation, let us reword his utterance:

(163) a. The vegetables are too heavy for the grey scale.
    b. The vegetables are too light for the black scale.

(163a) is false on its distributive reading, the one corresponding to a translation employing the D-operator of (141). It is true on its collective reading but that is not what the merchant intended. (163b) is false on the collective reading, the one corresponding to a translation without the D-operator. It is true on its distributive reading, but again that is not what the merchant intended to say. The physical arrangement of the vegetables in baskets suggests a cover of the vegetables with cells of the cover corresponding to baskets of vegetables. (162) is true and informative on the intended intermediate reading because the verb phrase is true of every member of that cover.

Reviewing now, our discussion began with the acceptance of those theories which model collective-distributive distinction with the help of the D-operator interpreted as in (141):
(141) \( x \in \| D(\alpha) \| \) iff \( \forall y [ \text{singularity}(y) \land y \in x \rightarrow y \in \| \alpha \| ] \)

This approach has two advantages. It delivers a clear distributive and collective reading and it accounts for certain ambiguities in terms of a scope interaction between the D-operator and modals and between the D-operator and indefinite noun phrases. The trouble with this approach is that it fails to make enough distinctions. In certain contexts, sentences with definite plural noun phrase arguments are found to have intermediate readings that are not predicted on this approach. We need a new account that will allow for more readings. This account must do justice to the pragmatic aspect of these readings, referred to in the generalization in (161).

5.2.4 A Generalization of the D-operator

Our current defective proposal involves the following semantic rule for the D-operator:

(141) \( x \in \| D(\alpha) \| \) iff \( \forall y [ \text{singularity}(y) \land y \in x \rightarrow y \in \| \alpha \| ] \)

We are happy with the universal quantifier attached to \( y \) and would like to retain it. The problem lies in the restriction to singularities. In the intermediate readings there is universal quantification, that is distribution, but not necessarily down to singularities. What would happen if we simply dropped this restriction:

(164) \( x \in \| D(\alpha) \| \) iff \( \forall y [ y \in x \rightarrow y \in \| \alpha \| ] \)

This doesn't do very much given the kind of model we are assuming (that of the union theory). Being on the left side of the membership sign, \( y \) is effectively restricted to singularities since our domain has in it only singularities and sets of singularities. So we need to change the membership sign to subset:

(152) \( x \in \| D(\alpha) \| \) iff \( \forall y [ y \subseteq x \rightarrow y \subseteq \| \alpha \| ] \)

We don't lose any values for \( y \) in this process, since Quine's Innovation (see the Appendix) guarantees that:

(166) \( \forall y \forall x [ \text{singularity}(y) \land y \in x \rightarrow y \subseteq x ] \)

The problem with (165) however is that it now requires too much for a distributive reading. To see this consider a situation in which the sentence:

Distributed bottles

(167)
Distributivity

(167) The bottles are light enough to carry.

is true but only on its distributive reading and that even two or three bottles would be too heavy to carry. If we go to map the true distributive reading with the operator defined in (165), we end up requiring that every set of bottles be in the extension of be light enough to carry. But then the translation of (167) with the D-operator is false, when in fact (167) is true on its distributive reading.

Getting rid of the singularity restriction was fine but we need some new restriction to replace it. The claim in (154) above, suggests the following:

(168) \( x \in \| D(\alpha) \| \) if and only if

\[
\text{There is a cover } C \text{ of } x: \forall y [ y \in C \rightarrow y \in \| \alpha \| ]
\]

Reconsider (167) in a situation in which each bottle by itself is light enough to carry though two or more bottles together would be too heavy to carry. The set of all the bottles is a cover of itself. This follows from the definition of cover:

(152) \( C \) covers \( A \) if:

1. \( C \) is a set of subsets of \( A \)
2. Every member of \( A \) belongs to some set in \( C \).
3. \( \emptyset \) is not in \( C \)

along with our adoption of Quine’s Innovation, according to which each bottle is a subset of the set of all the bottles. The predicate of (167) is true of every member of this cover. Translating (167) with a D-operator interpreted as in (168) yields a formula that is true in this situation. In addition, notice that if we left the D out we would get the collective reading and hence a formula that is false.

(168) represents progress. It retains the quantificational analysis of distributivity, and though it still allows for situations in which there is distributivity to singularities it is flexible enough to allow for intermediate distributive situations. Nonetheless it is flawed in two ways. First, we no longer have a true distributive (to singularities) reading, much as in the case of our original cumulativity-based analysis ((134), page 58). That is, distributive entailments no longer hold in our system. For example, if the bottles refer to three bottles named A, B and C, the entailment mapped as:

(169) \( D(\text{are-heavy}')(A + B + C) \rightarrow \text{are-heavy}'(A) \land \text{are-heavy}'(B) \land \text{are-heavy}'(C) \).
is not guaranteed by (168). This is because there may be another cover, say the one in which all the bottles occupy one cell, such that each member of that cover is in the extension of be heavy. The other problem with (168) is that it makes no reference to context. But we learned in the previous section that the availability of certain readings was dependent on context; not all covers are equal.

The source of both of these problems is the existential quantifier in the phrase "there is a cover C" in (168). The semantics should make reference to a specific cover, the choice of which is a matter for the pragmatics. This can be done by leaving the variable C free. In this case, the actual truth conditions which a sentence receives on a particular occasion of utterance are determined not by its translation alone, but by the translation interpreted with respect to a certain value assignment to its free variables, which is determined by pragmatic factors. There is some leeway in how we change (168). Here is one possibility:

\[
(170) \quad x \in || D(\text{Cov})(\alpha) || \text{ if and only if } \\
|| \text{Cov} || \text{ is a cover of } x \land \forall y[(y \in || \text{Cov} || \rightarrow y \in || \alpha ||)]
\]

Cov is a free variable over sets of sets. The value of Cov is determined by the linguistic and non-linguistic context. For example, in our vegetable example (163), the non-linguistic context provided a partition of the vegetables corresponding to their physical arrangement. This partition would have been assigned to Cov in the evaluation of (163). A slightly different way to amend (168) is as follows:

\[
(171) \quad x \in || D(\text{Cov})(\alpha) || \iff \forall y[(y \in || \text{Cov} || \land y \subseteq x) \rightarrow y \in || \alpha ||]
\]

In this version Cov is variable over covers of the whole domain of quantification. In future discussion we will assume this alternative, briefly returning to the choice between the two at the end of section 5.3.

In all versions of the semantics of the D-operator there is implicit restriction to the domain of quantification, as with all natural language quantifiers. The change then from (165) repeated here:

\[
(165) \quad x \in || D(\alpha) || \iff \forall y[\ y \subseteq x \rightarrow y \in || \alpha ||]
\]

to (171) is simply that we have quantification restricted to contextually specified covers over the domain rather than to the domain itself.

I would like to end this discussion by showing how we have regained the distributive (to singularities) reading of (167). Before doing

\[
(172) \quad \text{earlier}
\]

which

\[
(167)
\]

To simplify things, let us take the bottle example again. Let Cov be the set in the context of bottle 1, and let \( \beta \) be the assignment to the bottle variable. Then distributivity involves the term (170) to measure ambiguity in the context. It involves the "reading" or syntactic form here (replacing the collection operator) if necessary.

5.3 Inclusion
that I want to make a notational modification. From now on, let us leave the one place D-operator familiar from the literature with exclusive rights to that name and rename our operator "Part" :

\[(172) \quad x \in \parallel \text{Part}(\text{Cov})(\alpha) \parallel \text{ if and only if } \forall y[(y \in \parallel \text{Cov} \parallel \land y \subseteq x) \rightarrow y \in \parallel \alpha \parallel] \]

Turning now to the distributive (to singularities) reading, recall our earlier example:

\[(167) \quad \text{The bottles are light enough to carry.}\]

which on its non-collective readings would get translated as:

\[(173) \quad (\text{Part}(\text{Cov})(\text{are-light-enough-to-carry''})) \quad \text{(the-bottles')}\]

To simplify, let's assume that the domain of discourse doesn't include non-bottle entities. In that case, the reading we are after is the one in which Cov is assigned a set containing each of the bottles. This cover is salient in the discourse (it has been mentioned as the subject of the sentence) so its assignment to Cov is plausible. On this reading, (167) entails that each of the bottles is light enough to carry. This is what we have come to call the distributivity entailment.

This example points up a possible misunderstanding in the use of the term "reading". Throughout this discussion, I use the term "reading" to mean particular interpretation, including the choice of a meaning for ambiguous lexical items as well as the factoring in of specific aspects of context that affect interpretation. In this sense, there is a reading of (167) involving distribution to singularities. Some would limit the terms "reading" and "ambiguity" to differences in meaning deriving from lexical or syntactic ambiguity. On that view, according to the grammar envisioned here (to be modified below), (167) might be said to have two readings: one collective (translated without the Part operator) and one distributive, not necessarily to singularities.

5.3 Incorporating the Account into the Grammar

In this section, I would like to take a closer look at how the grammar needs to change in light of the analysis sketched in (172) below from the previous section:

\[(172) \quad x \in \parallel \text{Part}(\text{Cov})(\alpha) \parallel \text{ if and only if} \]
\forall y [ ( y \in \parallel \text{Cov} \parallel \land y \subseteq x ) \rightarrow y \in \parallel \alpha \parallel ]

To begin with, we need a new translation rule, something like the following:17

(174) Distributive VP rule:
If α is a plural VP with translation α', then Part(Cov)(α') is also a
translation for α.

Next, a semantic rule is needed to interpret these translations. We take e-
type expressions to denote elements of D^*, and type <e,t> expressions to
denote subsets of D^*. The cover variable is of type <e,t> and the Part
operator is interpreted as follows:

(175) Let α and β be variables whose values are object language
expressions of type <e,t> and let u,v be variables whose values are
entities in D^*. For all α,β,u:

\[ u \in \parallel \text{Part}(β)(α) \parallel \text{ if and only if } \]
\[ \forall v [(v \in \parallel β \parallel \land v \subseteq u) \rightarrow v \in \parallel α \parallel ] \]

Finally, a word about the pragmatics. In the introduction I said that
interpretation would be with respect to a model and that I would write
simply \parallel \cdot \parallel instead of \parallel \cdot \parallel^M, omitting the superscripted M. Since we
now have free variables in the translations, we need a mechanism by which they
get interpreted. For concreteness, let’s assume the pre-DRT view of
things, where interpretation is carried out with respect to an assignment
function and where the particular function chosen is somehow
pragmatically determined (Cooper 1979, for example, discusses this method
and attributes it to Montague). This means that from now on, \parallel \cdot \parallel is an
abbreviation for \parallel \cdot \parallel^M, where the superscripted M and g have been
omitted.

Returning to the translation rule (174), we see that it allows for
plural VPs to have two different translations. This is the source of the
distributive-collective ambiguity mentioned at the end of the previous
section. It also makes it apparent that the way that the range of
distributive readings is handled, namely in pragmatic terms, differs from the
way the distributive-collective ambiguity is handled. This conflicts with the

17 This rule will produce translations containing multiple Part-operators. I
suspect that this is harmless and will ignore this possibility.
intuition discussed earlier according to which the distributive (to singularity) and the collective readings are just extremes on a scale which includes distribution to subpluralities of various ‘sizes’. What I would like to explore then, is a way to modify our grammar to reflect this intuition.

Recall, above, a distributive reading of:

(167) The bottles are light enough to carry.

was analyzed with the following translation:

(173) (Part(Cov)(are-light-enough-to-carry’)) (the-bottles’)

by taking the assignment to Cov to be a set containing each of the bottles. It is reasonable to assume that by collecting the bottles together under one noun phrase, the speaker makes salient in the discourse another cover of the domain: one in which the bottles occupy a single cell. Assigning this cover to the variable Cov leads to the collective reading. In other words, we now have two sources for the collective reading, one with and one without the Part operator. Since we don’t in fact need the translation without the Part operator for anything else, we can simplify the translation rule (174) as follows:

(176) Plural VP rule:
If $\alpha$ is a singular VP with translation $\alpha'$, then the corresponding plural VP is translated as Part(Cov)(\alpha').

On this way of doing things, the Part-operator simply reflects the plural marking on the verb and the collective reading is now just one among many that the semantics and context could potentially yield.\(^{18}\) Support for this move, comes from the use of the phrase in a sense, which occurs when a speaker attempts to bring the so-called distributive-collective ambiguity into focus. For example, if John and Mary each made $1,000, one might at first reject (177), but then upon reflection utter (178):

(177) John and Mary made $1,000.
(178) Well, in a sense they did and in a sense they didn’t.

\(^{18}\) It might be harmless to incorporate the Part-operator in singular VPs as well. This would require a different kind of translation mechanism.
Lewis (1970:229) and Kamp (1975:150) have analyzed this phrase in its use with vague adjectives as in *John is clever, in a sense*. The role of *in a sense* according to these authors is to select a context or set of contexts with respect to which the adjective's vagueness is resolved. We might do something similar here. Assume that the predicate of (177) is interpreted with a Part(Cov) operator. In a "collective-context," Cov is assigned a cover in which John and Mary occupy the same cell. In a "distributive context," it is assigned a cover in which John and Mary occupy different cells. In an ambiguous context either assignment is possible. The role of *in a sense* in (178) is to restrict the interpretation now to a collective context, now to a distributive context. On this account then, we say that in the ambiguous context the plural noun phrase makes salient two different covers of the domain, one collective and the other distributive.

While I think this essentially pragmatic view of the distributive-collective ambiguity is correct, there is an apparent problem with the way we've set things up. Earlier (page 63), we argued for the D-operator analysis in part because it could handle the fact that conjoined verb phrases need not be understood both collectively or both distributively even though their conjunction combines with a single subject. On this current view, this type of data becomes a problem. To see this, consider the following example:

(179) These cars were put together in Malaysia and sent to different countries in Europe.

Here I am interested in a construal of the sentence in which the first verb phrase conjunct is understood distributively while the second is understood non-distributively. In other words, the cars were not attached to one another, but rather each car was assembled in Malaysia and each car didn't go to different countries, rather the shipment of cars was dispersed in Europe. Now consider the kind of translation the conjoined VP would receive (assume $||A \land B||$ is interpreted as $||A|| \cap ||B||$):

(180) Part(Cov)(p.t.i.Malaysia') $\land$ Part(Cov)(s.t.d.c.i.Europe')

For simplicity, let's assume the domain includes only the cars in question. Following the recently adopted view of the distributive-collective ambiguity, the cover variable in (180) is either assigned a set containing each of the cars, in which case we get the distributive reading of both conjuncts or it is assigned a set containing the set of all the cars in which case we get the collective reading of both conjuncts. Unfortunately, neither of these possibilities corresponds to the desired reading of the sentence.

Thinking from a different view, a distributive reading of the conversational space on all parts, while a collective reading counts only the same as a distributive reading of the same space. This takes us to the following view of the phrase (179):

(181)

Now, a couple more thoughts: if we think about the cars as the arguments and the language as the arguments, then they are from a different space. Then the D-operator has scope in the language for the operator in the argument space. Later, the whole sentence is interpreted as a sort of existential quantifier, i.e., the existential variable that is interpreted with p.t.i. Malaysia' although it is a proper quantifier and not an existential quantifier.

\[\text{p.t.i.}\]
I think the source of this problem is that we are not yet used to thinking of the distributive-collective ambiguity as pragmatic. On that view, a distributive-to-singularities reading arises because, in some sense, the conversants are thinking of the plurality in question in terms of its singular parts, while the collective reading arises when the conversants think of the plurality in question as a whole. But nothing prevents us from thinking of the same plurality in two ways. The problem in (180) is not the pragmatic view of collectivity, but rather that that representation doesn't allow for this third possibility. In fact, the culprit here too is our translation rule, which forces us into assuming one way of thinking of the given plurality per conversation. The following modification should rectify this:

(181) Plural VP rule:
If \( \alpha \) is a singular VP with translation \( \alpha' \), then for any index \( i \),
\[
\text{Part}(\text{Cov}_i)(\alpha') \text{ is a translation for the corresponding plural VP.}^{19}
\]

Now, a possible translation for (179) would be:

(182) Part(\text{Cov}_i)(\text{p.t.i.Malaysia}') \text{ } \land \text{ } \text{Part}(\text{Cov}_j)(\text{s.t.d.c.i.Europe})

And, on the intended reading, \( \text{Cov}_i \) is assigned the set containing each of the cars, while \( \text{Cov}_j \) is assigned the set containing the set of all the cars.

Although the rule in (181) grew out of a consideration of how language users think about pluralities, one could have arrived at this result from a different route beginning with an argument presented earlier for the D-operator (section 5.2.2). It had been noticed that plural predicates display scope interactions, behaving as if they contained a quantifier. The D-operator was a spelling out of that quantifier, as is our Part operator. Later, after having introduced the cover variable, I remarked that quantification in natural language always or almost always involves some sort of discourse restriction of the domain of quantification. The cover variable is a way of formalizing that restriction for the quantifier associated with plural predicates. Continuing this line of reasoning, we note that although domain restriction may be a restriction of the entire domain of quantification, it is something that is done on a per-quantifier basis, as pointed out in Westerstahl (1985). Here too, we should allow for each quantifier associated with plural predication to have its own domain

---

19One might want here a set of rules, one for each index, as Montague (1974) does with Quantifying-In. In that case, each rule provides a unique translation.
restriction. This is what rule (181) does.

I want to end this section with one more piece of evidence in favor of the grammar set up here and in particular the view it takes on the distributive-collective ambiguity. Consider the following incident. Apparently, in the last five years, an unsavory Mr. Slime has made several purchases from a computer store: 4 computers and 1 cartonful of diskettes. These purchases were made over the course of a few years and each time, Mr. Slime paid an initial amount in counterfeit currency and the remainder he paid for with a valid credit card. The following remark is entered in the police report:

(183) The computers were paid for in two installments and the diskettes were too.

First note, that the intention here is a distributive reading of the first verb phrase and a collective reading of the second, elided, VP. An analysis in which distributive and collective readings correspond to different underlying forms would have to explain how the second VP was elided when in fact it was non-identical to the first. What should we say about this example? We will assume the following formula represents the meanings of the conjuncts of (183) with c denoting the computer-plurality and d, the diskette plurality:

(184) \[ c \in \text{Part}(\text{Cov}_0(p.f.i.t.\text{installments}')) \]
\[ d \in \text{Part}(\text{Cov}_0(p.f.i.t.\text{installments}')) \]

Notice, here we cannot explain the different readings of the VP in terms of assignments to different cover variables, since here we are assuming identity of the VPs.\(^{20}\) In fact, we don’t need to assume different cover variables.

\(^{20}\)It is interesting to compare the reasoning here with that of Klein (1980:15-16). Klein was looking at the contextual evaluation parameters that determine the comparison class for adjectives like \textit{tall} and those that determine the domain of quantification for quantificational noun phrases. Klein observed that an elided VP and its antecedent can differ with respect to the settings of these parameters for material inside the VPs. In contrast, the referent of an indexical pronoun in an elided VP must be identical to the referent of the corresponding pronoun in the antecedent VP. Klein concluded from this that the contextual parameters should not be incorporated like pronouns into the representation. We are not driven to a similar conclusion in the present case, as the reader will shortly see.
variables in this case, since the subjects of the two VPs are not identical. Rather we need to consider how the writer of (183) was thinking about things. As far as the facts of the case were concerned, the purchased items divide up into purchase-parts: one for each computer and one for the set of diskettes. When this cover of the domain (or some extension of it to cover things other than the stolen merchandise) is assigned to the variable Cov, the intended reading of (183) results.

It should be pointed out that the analysis just provided rides on the assumption that Cov gets assigned a cover of the whole domain and not just a cover of the plurality that the VP is predicated of, as suggested earlier (section 5.2.4, (170-171)). Lasersohn (1995) and an anonymous reviewer have pointed out that the decision to assign covers of the whole domain allows for cover choices that would effectively eliminate dependence of the truth of a sentence like John and Mary left on whether or not John left. This would happen, for example, if the cover puts John in only one cell, with someone other than Mary. While such a cell is in the cover, it is not a subset of the set of John and Mary, hence the truth value assigned to the sentence will incorrectly not depend on John's having left or not, given the rule in (172) above. The following variation on the semantics of the Part operator would take care of this.

Alternative Semantics for the Part operator:
(a) For any Y, a set of sets of individuals, and any y, a set of individuals, Y/x is the largest subset of Y that covers x, if there is one, otherwise it is undefined.
(b) \( x \in \llbracket \text{Part}(Cov)(\alpha) \rrbracket \) if and only if \( \forall y (y \in \llbracket Cov \rrbracket /x) \rightarrow y \in \llbracket \alpha \rrbracket \)

I refrain from adopting this formulation because I believe that pathological values for domain of quantification variables should be ruled out pragmatically and not semantically. This point is elaborated in Schwarzschild (1994:228-233).

5.4 Excursus on Plural Quantification: Partitions

In the previous section we replaced the D-operator which was interpreted with a quantifier (implicitly) restricted simply to the domain of quantification with the Part-operator which involved quantification restricted to a cover of the domain of quantification. Let us call this kind of restricted quantification, partitioned quantification. I use the verb "partition" loosely referring to something that results in a cover of any sort,
even one that is not technically a partition. Our discussion gave the impression that partitioned quantification was something unique to the Part-operator. The point of this excursus is to indicate that partitioned quantification is pervasive. To show this, I will briefly review a number of examples whose interpretation is sensible only in case quantification of this sort is assumed. I will not endeavor to analyze these examples in any serious way.

Our first example is of a type common in statistical reports:

(185) One out of every three handguns in America is made by Smith and Wesson.

(185) is not falsified by the fact that three handguns can be found in America all of which are not made by Smith and Wesson. This is because (185) is considered true if there is a partitioning of handguns into threes such that each triplet contains a Smith and Wesson. Interestingly, replacing every with each or any changes the meaning in a way that seems to be related to partitioning:

(186) One out of any/each three handguns in America is made by Smith and Wesson.

If (186) means anything at all, it is that for any partitioning of handguns into threes, each triplet contains a Smith and Wesson. This is a roundabout way of saying that with the exception of at most two, all American handguns are Smith and Wessons. The difference noted here between every and any is especially clear in a situation in which the context provides an obvious partitioning of the domain. Observing a suburban neighborhood in which houses are built in blocks of three, each block in a different style, one may say:

(187) I observed that every three houses {formed a block / were built in the same style}.

but it would be false to say:

(188) I observed that any three houses {formed a block / were built in the same style}.

If the difference between any and every has been correctly analyzed, then I do believe there is a lesson here for our move from the D to the Part operator. The D-operator was modeled on each and for many speakers,

(189)

is true. Collectively, the "cumulatively" is less than 10,000, so we equally assume

(190)

If no other

(191) but

(192) Lasersof all the families must be described, if they are to represent the money we do not pay. counter
Distributivity

floated *each* quantifies over singularities only. For this reason it seemed natural that a covert distributivity operator should also quantify over singularities. However, the comparison between *every* and *any* shows that quantifiers can differ with respect to the partitioning of the domain of quantification. Floated *each* requires a partition into singularities while the Part operator does not.

Our next example comes from the discussion in Lasersohn (1988:Ch.IV) of quantifiers and group-level properties. He distinguishes between three different types of situation in which the sentence:

(189)  John and Mary made $10,000.

is true. Either John and Mary each made $10,000 or in the "pure collective" case they made $10,000 in a joint enterprise or in the "cumulative" case, the combination of their individual incomes amounts to $10,000. Lasersohn points out however that these three situations are not equally relevant for the negative quantifiers *no* and *only*. To see this, assume that in fact (189) is true and that in addition it is true that:

(190)  Bill made $10,000.

If no other individuals besides those mentioned so far made any money, is (191) below true?

(191)  Only Bill made $10,000.

Lasersohn contends that the answer depends on which of the situations described above made (189) true. If John and Mary each made $10,000 or if they made $10,000 in a joint enterprise then (191) is false. But if $10,000 represents the combination of their individual incomes (and no other money was made by them) then (191) is true. In short, negative quantifiers do not pay attention to the cumulative case.

Lasersohn (1988:190) discusses the following putative counterexample to this last claim:

Consider the budget of a small city. The payroll for the police department totals $1,000,000, the payroll for the fire department also totals $1,000,000, and the payroll for the sanitation department totals $500,000. In this situation, sentence (192) seems false:

(192)  Only the police officers get paid $1,000,000.
The sentence is false because the firefighters also get paid $1,000,000. Since it is the combined income of the firefighters that is in question, this appears to be a case where only excludes a group from having a property even if it has that property only by virtue of a totaling operation on the properties of the group’s members.

Compare another example, however. Suppose now that the fire department payroll is only $500,000. In this case, (192) is true – despite the fact that the combined income of the firefighters and the sanitation workers is $1,000,000. What is different about this case? It seems clear that the reason why (192) was false in the original situation but true in this one is that in the original situation there were two distinct payrolls which (each) totaled $1,000,000, while in the new situation there is only one. To calculate the truth value of the sentence, one compares the lump sums allocated to single entries in the overall budget.

Elaborating on this explanation we might say that the context naturally partitions the workers into three sets of workers and only is sensitive to this partition. In the first case there is another member of the partition besides the police officers that gets paid one million dollars, while in the second case there isn’t. Reference to a partition is important here. We cannot simply say that the domain includes only three entities that get paid. The following could be felicitously uttered in a conversation including (192) without affecting the facts outlined above:

(193) Every worker receives his paycheck on the same day so the city estimates it must have at least $50,000 in the bank at all times.

In other words the domain must include individual workers as well.

Another case of partitioned quantification occurs in the use of the word majority. Consider first the indefinite term a majority. According to (194),

(194) If a majority votes for this proposal, we are doomed.

---

\(^{21}\)It could be that the partitioning observed here comes in not from the meaning of only but rather from a Part operator on the verb phrase.

Distributions over a group can contain any group that a sentence that a sentence truly assigns an event.

(195)

Who or what contains the domain? Further, are the events contained there? That must be an issue of assumption.

pluralities... counting events.

following... then isn’t.

an operator... expression.

5.5 Two important... this section.

\(^{22}\)Cases... that set containing a dog counted as a female. The reason is that a dog is an animal, not a female, so it is fine for a lawyer...
any group containing most of the voters could spell disaster. Imagine now that a secret ballot is taken, 40 out of 50 voters are in favor and it is now truly announced that,

(195) The majority voted for the proposal.

Who or what does the majority refer to? There is no unique group containing most of the voters, unless we have partitioned the domain. Further, not just any partition will do. A group containing all those who were against plus 20 of those who voted in favor constitutes a majority. That majority in fact did not (all) vote for the proposal. Rather, what is assumed here is a partition into voting blocks.

Partitioning as a prerequisite for quantification is not limited to pluralities. Various people have noted that partitions are assumed in the counting of kinds (Carlson 1977:346ff)\textsuperscript{22}, facts (Kratzer 1989:608ff) and events.

The omnipresence of partitioned quantification suggests the following open question. If partitioning is a prerequisite for quantification then isn’t a generalization missed by including this as a part of the Part operator? Shouldn’t it simply fall out of a general mechanism in the grammar for interpreting quantifiers?

5.5 Two Place Predicates and Distributivity

Various people have discussed the phenomenon of distributivity as it relates to two or more arguments of a verb at a time. The purpose of this section is to consider some examples, particularly those of Scha (1984)

\textsuperscript{22}Carlson requires that in order to count members of a set of kinds, that set must form a partition, not just a cover. That is, the set of kinds counted must be such that no two kinds have the same realization. His reason is that we cannot say three kinds of dogs are in this room if just one dog is in the room even if that dog instantiates three kinds, e.g. collies, females, loving dogs. Carlson’s requirement is probably too strong. I think it is fine to say:

(i) Three kinds of professionals will attend this conference— doctors, lawyers and college professors.

despite the fact that there may be an individual who is both a doctor and a lawyer.
and Scha and Stallard (1988), which have been used to identify special distributional readings associated with two place predicates and to extend the ideas on distributivity presented above to these cases.

Scha considers the following sentence in connection with the figure in (197):

(196) The sides of R1 run parallel to the sides of R2.

(197)

\[
\begin{array}{c}
\text{R1.} \\
\text{R2.}
\end{array}
\]

He claims that (196) has the reading given in (198):\(^{23}\)

(198) \(\forall x \in \text{SR1} : \exists y \in \text{SR2} : \text{PAR}[x,y] \land \forall x \in \text{SR2} : \exists y \in \text{SR1} : \text{PAR}[x,y] \]

\(\text{[SR1 = the set of sides of R1]}
\text{[SR2 = the set of sides of R2]}

Scha analyzes (196) with the following formula:

(199) \(\text{PAR}[\text{SR1, SR2}]\)

Following the method of Bartsch (1973), the reading in (198) is derived by adding the following meaning postulate to the grammar:

(200) \(\text{PAR}[u,v] \leftrightarrow \forall x \in u : \exists y \in v : \text{PAR}[x,y] \land \forall y \in v : \exists x \in u : \text{PAR}[x,y]\)

\(^{23}\)Scha also includes as part of the reading that sets SR1 and SR2 are non-empty. We leave this out here and in the meaning postulate to be given in (200). Also, Scha maps individuals as singleton sets which are always distinguished from their members. I have adapted his formulae to the approach taken here, as discussed in the Appendix. In this adaptation, \(x,y\) are variables over singularities, while \(u,v\) are variables over pluralities and singularities.
Distributivity

This analysis is similar to Lasersohn’s approach to Gillon’s examples discussed above. As in that case, I would like again to show that the meaning postulate approach cannot account for the effects of varying context. To that end consider the following sentence in connection with either of the diagrams in (202) or (203):

(201) The double lines run parallel to the single lines.

(202)

(203)

I find the sentence false in both situations or maybe difficult to judge. But the meaning postulate in (200) would have it otherwise. Letting SR1 now be the set of the double lines and SR2 the set of the single lines, the formula in (198) is true. This means that if the meaning postulate is correct, (201) should be true. My feeling is that again factors that arise in the interpretation of a sentence in a specific context (197) are mistakenly identified as part of the lexical meaning of the predicate in question. Pursuing the program begun above, I would like to find a way to have the relevant contextual information enter into the interpretation without incorporating it once and for all in the meaning of lexical items.

In order to develop a plan for using necessary contextual information in the interpretation, we first need to decide what the relevant information is. Let’s look again at the diagram in (197) repeated below, to determine why it is that we agree the sides of the first rectangle run parallel to those of the second.
It appears that we compare the horizontal sides of the rectangles and the vertical sides of the rectangles independently or perhaps we compare the top horizontals, the bottom horizontals and left and right verticals. This diagram differs from those in (202) and (203) in that it provides us with an intuitive partition of the lines in question. Whereas in previous discussion we relied on the partitioning of sets of elements, here we need to partition sets of pairs. Building on our earlier analysis, we need an operator like Part and a variable like Cov that work on pairs. To do that, we need to extend the notion "cover" to pairs:

\[(204) \quad T \text{ is a paired-cover of } \langle A, B \rangle \text{ iff:}
\]
\[
\text{there is a cover of } A, C(A), \text{ and there is a cover of } B, C(B), \text{ such that:}
\]
\[
i. \quad T \text{ is a subset of } C(A) \times C(B).
\]
\[
ii. \quad \forall x \in C(A) \exists y \in C(B): \langle x, y \rangle \in T
\]
\[
iii. \quad \forall y \in C(B) \exists x \in C(A): \langle x, y \rangle \in T
\]

If \( T \) is a paired-cover of \( \langle A, A \rangle \)
then \( T \) is a paired-cover of \( A \)

Now we introduce a two-place version of the Part operator, \( \text{PPart} \) (short for paired Part). We let \( X, Y \) be variables whose values are pairs of elements in the domain and we use the symbol \( \leq_2 \) to combine two pair denoting terms where \( \langle a, b \rangle \leq_2 \langle c, d \rangle \text{ iff } a \leq c \land b \leq d \). We call this relation pair-subset. We assume now that some contexts provide a value for the variable \( \text{PCov} \) and that that value is a paired cover of the domain. Our semantics for \( \text{PPart} \) is given as:

\[(205) \quad X \in \| \text{PPart}(\text{PCov})(\alpha) \| \text{ if and only if}
\]
\[
\forall Y[ (Y \in \| \text{PCov} \| \land Y \leq_2 X) \rightarrow Y \in \| \alpha \| ]
\]

Let us see how this works in our test case. The paired cover assigned to \( \text{PCov} \) in the context of Scha's diagram consists of pairs whose members are...
both horizontal or both vertical and it will not contain mixed pairs. Let (SR1,SR2) denote that pair of sets whose first element is the set of sides of rectangle 1, and whose second element is the set of sides of rectangle 2. Take an arbitrary element Y of PCov. If its first element is a subset of SR1 and its second element is a subset of SR2, then \( Y \subseteq_2 (SR1,SR2) \). Furthermore, \( Y \in || \text{run-parallel'} || \) because of the way PCov was set up. That is, in (197), any two lines that are both horizontal or both vertical are parallel. So the following holds:

\[
\forall Y[(Y \in || PCov || \land Y \subseteq_2 || (SR1,SR2) ||) \rightarrow Y \in || \text{run-parallel'} ||]
\]

Assuming \text{run parallel} is translated PPart(PCov)(run-parallel') and the usual rules for assigning meaning to transitive clauses are employed, the semantics given in (205) guarantees that the sides of R1 run parallel to the sides of R2 is true for Scha's diagram, (197). The required translation entails a rule of the following sort:

(206) Plural TVP rule:
If \( \alpha' \) is a singular transitive verb phrase with translation \( \alpha' \), then for any index \( i \), PPart(PCov)(\( \alpha' \)) is a translation for the corresponding plural transitive verb phrase.

In order to further justify the PPart operator, I would like to present some more examples whose interpretation seems crucially to rely on a pair-partitioning of the domain. Before doing that we might ask why (201) is not true of (202):

(201) The double lines run parallel to the single lines.
(202) \[\begin{array}{c}
\hline
\end{array}\]

Our answer will be somewhat tentative, for reasons to be explained in the next section. It may be the case that mention of the double and single lines introduces a paired partition into the discourse consisting of all pairs of a double line and a single line and that this is the only value assignable to PCov. Now there are pairs of non-parallel lines in this cover that are pair-
subsets of the pair whose first element is the double lines and whose second element is the single lines. It would follow then that (201) with run parallel translated as PPart(PCov)(run-parallel') is false.

We move now to other instances of the paradigm examined so far. The first is based on an example that I have discussed elsewhere (Schwarzschild 1990). Imagine you arrive on the first day for a literature class on the relation between fiction and non-fiction. While introducing the course requirements, the lecturer directs your attention to the chart below and says that:

(207) The fiction books in the chart complement the non-fiction books.

<table>
<thead>
<tr>
<th>Fiction</th>
<th>Non-fiction</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alice in Wonderland</td>
<td>Aspects; Language (Bloomfield)</td>
</tr>
<tr>
<td>Fantastic Voyage</td>
<td>Gray’s Anatomy</td>
</tr>
<tr>
<td>David Copperfield,</td>
<td>Das Kapital, The Wealth of Nations</td>
</tr>
<tr>
<td>Hard Times</td>
<td></td>
</tr>
<tr>
<td>Oedipus Rex,</td>
<td>Freud’s</td>
</tr>
<tr>
<td>Agamemnon</td>
<td>Intro. to Psychology</td>
</tr>
<tr>
<td>Richard III</td>
<td>Machiavelli’s The Prince</td>
</tr>
</tbody>
</table>

You are required to read three fiction books and their non-fiction complements. It hardly needs saying that the truth of (207) depends on a comparison of adjacent pairs of books. Note a number of things. The domain is partitioned into pairs of adjacent entries. Some of the pairs have non-singleton sets as members. (207) is true if for every pair in the pair-partition, if it is a pair-subset of (|| the fiction books ||, || the non-fiction books ||) then the first element complements the second element. Crucially, non-adjacent pairs are irrelevant here.

In the two examples we have seen so far the value assigned by the context to PCov was determined by non-linguistic, graphic information. For balance I mention a few other kinds of cases. Scha and Stallard (1988), whose work was intended for use in accessing information about the capabilities and readiness conditions of the ships in the Pacific Fleet of the US Navy, discuss the following example:

They found that the ship’s men are more aggressive when they are in the universal than in the cigar factory. The men in the universal enter faster and stay longer in the smoking room. They search more frequently for the existence of men in the cigar factory. When they do find someone, they either move away or wait longer before approaching. They are more likely to enter the cigar factory than the universal. The smoking room is of course a more rewarding reading room and the men have a much better understanding of their smoking room behavior. The men who are present in the smoking room know the smoking room is their smoking room and that it is larger than the cigar factory. The men in the cigar factory are not sure who is present in the cigar factory between the usual and non-usual times.

Deterministically, the smoking room is a less rewarding reading room.

(209)

Here the men are more aggressive when they are in the universal.

(210)

In this example, the smoking room is the smoking room, the cigar factory is the smoking room which is less rewarding. If the smoking room is the universal, then the smoking room is not the universal. The smoking room is a more rewarding reading room than the cigar factory.

which is an introduction to the smoking room. The smoking room is a more rewarding reading room than the cigar factory. The smoking room is the smoking room. PPart
Distributivity

(208) The frigates are faster than the carriers.

They speak of two translations for this sentence. The first, called a universal-universal translation, leads to a reading where every frigate is faster than every carrier. The other translation, which I will call universal-existential, requires that every frigate is faster than some carrier and that for every carrier there is a frigate faster than it. It is not very hard, however, to envision a situation in which neither of these represents the correct reading of the sentence. We just need to think partitionally. Imagine for example, that (208) is uttered in a context in which it is clear that these ships are sent out in teams to different areas of the globe with each team consisting of frigates and carriers. It may be that one area calls for very fast action while another will tolerate a sluggish response. If that were the case; I would judge (208) true just in case the frigates in a given area were faster than the carriers of that area, regardless of what speed relations obtained between ships of different areas. In this situation the universal-universal reading is too strong and the other reading is too weak. A semantics that incorporates the notion of a contextually determined partition accounts for these facts without having to drum up new translations.

Consider finally an example where the value for PCov is determined linguistically within the sentence containing PCov:

(209) Even though the couples in our study were not married, the men did display aggressive behavior towards the women.

Here the concessive clause raises the salience of a paired-cover in which men and women are paired into couples. (209) seems to be about aggressive behavior within pairs. The point is strengthened if all is added:

(210) Even though the couples in our study were not married, the men all displayed aggressive behavior towards the women.

In this case every man would have had to display aggressive behavior towards his female pair-mate, but crucially he would not have had to be aggressive to non-pair-mate women for (210) to be true.

I would like to end this section by briefly mentioning directions in which the present account could be further pursued. The two operators introduced so far are covert. The floated quantifier each could now be thought of as an overt counterpart of Part with a particular assignment to Cov, namely one where the cells are singularities. The adverb respectively might similarly correspond to the PPart operator. It has the effect of a PPart operator with a particular assignment to PCov. Other phrases that
seem to restrict the values of Cov and PCov are together, especially in its sentence initial position as well as phrases such as one by one and in groups of three. In each of these cases, the lexical item puts restrictions on the value that can be assigned to a free variable. This is similar to analyses of temporal adverbials where they are said to restrict the possible values assigned to a contextual variable over times.

The Part operator is for one-place predicates and the PPart operator is for two places. It is natural then to consider 3 and more place predicates. The following alternative formulation of the semantics of two-place distributivity in terms of functions should suffice to show how things could be extended for more places.

\[(211)\] Alternative Semantic Rule for PPart:

Let $\beta$ be a variable of type $\langle \langle e,t \rangle,t \rangle$, $\alpha'$ an expression of type $\langle e,t \rangle,t$ and $a,b,u,v$, variables over elements of the domain, $D'$, then:

\[
\| \text{PPart}(\beta)(\alpha') \|^{M_{\mathcal{E}}}(a) = 1 \text{ iff } \\
\forall u \forall v \forall (\forall f(\beta)(u)(v) \land u \subseteq a \land v \subseteq b) \rightarrow \|\alpha'\|^{M_{\mathcal{E}}}(u)(v))
\]

5.6 On Collective Readings

The approach that I have outlined here in terms of contextual paired-coverings has been offered as an alternative to Scha's analysis based on meaning postulates. I think that the comparison between the approaches is worth pursuing particularly in light of the elaborations and modifications that are presented in Scha and Stallard's article. My original motivation for employing the PPart operator was simply that the meaning postulate in (200) fails in certain situations. Roberts (1987:133-4) lodges another objection to Scha's program. She points out that certain verbs are ambiguous, having distributive and collective readings and so if the source of distributivity is a meaning postulate one is forced to claim that meaning postulates can be optional. This, Roberts claims, is incoherent and she goes on to propose an analysis in terms of a distributivity operator. But Scha and Stallard have found a solution to this problem. Essentially, they give multiple translations to English expressions and these translations are associated with the various readings.²⁴ For example, the English verb eat

²⁴Scha and Stallard's system actually involves double translation. English is translated into a language which itself gets translated.

\[\]

These phenomena need to be handled in the sentence initial position. In this section, we can identify how things could be extended for more places.

\[(212)\]

\[(213)\]

The grammar postulates only PPart, not the Part. This means that meaning postulate (213) holds of the partake of the gesture, not of the partake. In the present context, we can assume that the context.

The approach that I have outlined here in terms of contextual paired-coverings has been offered as an alternative to Scha's analysis based on meaning postulates. I think that the comparison between the approaches is worth pursuing particularly in light of the elaborations and modifications that are presented in Scha and Stallard's article. My original motivation for employing the PPart operator was simply that the meaning postulate

\[\]

\[(1988,19)\]

and Stallard's system actually involves double translation. English is translated into a language which itself gets translated.

\[\]
has at least these two translations:

\[
\begin{align*}
\lambda u,v : \forall y & \in v : \exists x \in u : \text{EAT}'[x,y] \\
\lambda u,v : [\forall y & \in v : \exists x \in u : \text{EAT}'[x,y]] \land \\
[\forall y & \in u : \exists x \in v : \text{EAT}'[y,x]]
\end{align*}
\]

These particular translations are motivated by Scha and Stallard's belief that the sentence *the children ate the pizzas* has different quantificational readings which differ in whether or not each child has to have eaten a pizza or not. Scha and Stallard allow that different readings are imposed by context, though they do not elaborate on this claim.

Given this new way of incorporating various readings in the grammar, there may now be an answer to my charge that the meaning postulate in (200) leads to incorrect predictions in some situations, since that meaning postulate is now attached only to one particular translation of the phrase *run parallel*. That translation is salient, so the story would go, in the context of Scha's original diagram but not with the figures I presented as counterexamples (202) and (203).

To me this approach is misguided in the role that it attributes to the context. In the examples we have seen so far the context is not providing information about which quantificational formula is appropriate, but rather about specific groupings. It tells us which elements are to be compared, which elements are to be checked and how, in order to verify the sentence. It does not determine a quantity of elements to check nor how many of this type must bear the relation in question for every one of that type.

There is, I think, another more fundamental flaw implicit in Scha and Stallard's agenda. The driving force of that analysis seems to be to translate away plural predications into quantification over and predication of singularities, whenever the predicate in question is applicable to singularities. It is doubtful whether this goal is ultimately attainable. Lønning (1987:124)'s discussion of the sentence *the boys ate the cakes* illustrates the difficulties one encounters in pursuing this goal. He points out that this sentence is true in a situation in which two boys jointly partake of each of two cakes. Notice first that the Scha and Stallard translations for *eat* given in (212) and (213) will not work here. In fact, without moving to quantification over sub-singularities (parts of cakes), it seems well nigh impossible to formalize this as a reading in any meaningful way. Matters get even worse once you consider what Lasersohn (1988,1990) terms "team-credit extensions." These are examples in which a team gets credit for the actions of one of its members. For example, if John and Mary are a couple we may report that John and Mary made
$20,000 last year, even if in fact, only one of them actually worked. Expanding on the discussion in Carlson (1977:102ff), Roberts (1987:147) notes that

(214) The Marines invaded Grenada.

is true, in one sense, even though not all members of the U.S. Marine Corps went to Grenada.\(^{25}\) The problem these examples pose for the agenda I have attributed to Scha and Stallard is as follows. Team credit extensions have a non-logical aspect to them. They cannot be analyzed simply by providing a translation for the verb phrase that has an existential quantifier in it, (e.g. \(\lambda u: \exists y \in u: \text{invaded-Grenada}[y]\)). Thus for example, it cannot be said that Mozart and Einstein won the Nobel Prize even though one of them did, because Einstein and Mozart were not a team in any sense.

The conclusion that many have reached based on the foregoing examples, and with which I concur, is that even predicates applicable to individuals can have a simple collective reading. On this reading, we should not, indeed can not, specify in the grammar how many of the singularities that make up a plurality must satisfy the predicate in order for that plurality to satisfy it. In the system I sketched here, this reading arises when the cover does not partition the plurality in question into more than one part. In this case, the Part and PPart operators don’t do any real work. And they shouldn’t, for they are not present in the grammar in order to specify quantificational refinements of the collective reading; this we have just said is fruitless.

Some of the examples mentioned above might be useful in demonstrating the distinction I am making here. Recall the sentence the frigates are faster than the carriers uttered in a context where the ships are sent out in crews to different areas. If all of the carriers in some high-speed area are faster than all of the frigates in some low-speed area, the sentence isn’t false. The construal we are after which employs the PPart operator tells us not to compare ships across areas. On the other hand, exactly how many frigates have to be faster than how many carriers in a given area is left unspecified. There may be exceptional quick carriers. Next we have

\(^{25}\) Similarly, Gillon (1984) notes that if the soldiers in F-Troop are chasing a band of Indians and the soldier in front sees them we can say the soldiers of F-Troop spotted the Indians. However, if a member of F-Troop sees an Indian while on vacation, we do not say the soldiers of F-Troop spotted the Indians.
the book chart example from (207) in which the PPart operator tells us which particular entries are complementary. It does not tell us how many are complementary. Notice, that if the list was twice as long as it is, we still would not expect complementarity within the fiction or non-fiction column. And again, although the context determines which entities are related to which, there are singularities about which it says nothing. Consider one particular pair of adjacent books, say Oedipus Rex and Agamemnon on the one hand and Freud’s Introduction to Psychology on the other. The sentence in (207) asserts that the former two complement the book by Freud. No relation is claimed to hold here between individual books.

It is important to stress that the distinction I am making here is not simply one of ‘quantity’ versus ‘quality’. The important distinction is that between pragmatics and semantics. Consider again the point made in connection with the Granada example, (214). Here the question is how many and perhaps even which individuals in a group have to possess a property in order for us to say that the group possesses the property. The answer here will depend on the makeup of the group as well as the kind of property ascribed. These are semantic facts, or extralinguistic facts, that I, as a speaker, do not control in uttering (214). Compare this to the point made with the chart example (207). Here the question is what exactly is being said, which proposition is expressed. If the speaker has in mind the context illustrated by the chart, he is simply not claiming anything about non-adjacent book pairs. Think of what happens if the chart is rearranged and the context is changed. In this case, the speaker says something different about these very same books. The facts of the world will not have changed, the speaker is talking about the same books and the same relation of complementarity, yet the proposition expressed will be different.

I would like to end this subsection with an example discussed in the literature. Dowty (1987) expresses the view that it would be appropriate to use (215) to describe a news conference at which only a small number of the reporters present asked questions:

(215) At the end of the press conference, the reporters asked the President questions.

This is so despite the fact that some (or even many) reporters may not have asked questions. Compare this to:

(216) At the end of the press conference, the reporters remained silent.

Here one feels that all of the reporters must have remained silent for the
sentence to be true. In both cases, a group is said to possess a property, but this entails different quantities of individual group members possessing the property in the two cases. Whatever the explanation, it will have to do with facts about groups of reporters and about the properties being ascribed.

Now consider the following variation on (215):

(217) At the end of the press conference, the reporters from NBC, CBS and ABC asked the President questions.

Here again some but not all of the reporters would have to have asked questions for the sentence to be true. However, there is a difference. Even if the reporters mentioned in (215) are also from the major networks, indeed the same exact reporters, (215) doesn’t seem to require that every network got a question in. On the other hand, (217) in at least some contexts, implies that questions came from one or more members of each of NBC, CBS, and ABC. Mention of these networks introduces a partition of the reporters in the discourse and, on the pertinent reading, this partition is assigned as the value of the variable that provides the domain of quantification for a distributivity operator.

5.7 Plurals in Discourse: The Pragmatics of Distributivity

On the view espoused here, the truth conditions for sentences with plural arguments are often determined in part by the assignment to a free variable over coverings or paired coverings of the domain. We have said that the source of this assignment is pragmatic. Can we say more?

The question of what makes a partitioning of the domain salient in the discourse bears some resemblance to the question of what makes the antecedent for a pronoun salient. In many instances there are linguistic clues, some of which will be discussed below, but arriving at a complete answer surely involves other branches of cognitive science. Such is the case for domain partitioning as well. How we divide up our visual space for example is relevant here and yet that is a question which is properly a matter yet to be settled by experts on vision. Recall above we tried to explain why, in contrast to the diagram Scha used, the double lines run parallel to the single lines might be judged false for the diagram in (202):

(218)

I have recently learned from Manfred Krifka that differences like these are discussed in Kang (1994).
Distributivity

A tentative answer was given to this question. A complete explanation of this difference demands an account of why there is no salient paired-cover here like there was in Scha’s example. Such an account lies beyond linguistic theory.

So, while non-linguistic sources for partitioning are important to demonstrate that there is a pragmatic element to distributivity, we shouldn’t expect to say much within linguistic theory about why a particular partition gets chosen in these cases. On the other hand, just as with pronominal anaphora, there are contexts in which the source of a domain partition or cover is linguistic and presumably these should be covered by some part of linguistic theory. What I therefore want to do now is to discuss aspects of anaphora resolution that appear to shed some light in the area of domain partitioning. Before turning to those parallels, I would like to remark on the status of the discussion to follow. I will be comparing the assignment of a value for a cover variable to the choice of antecedent for a pronoun. Pronominal anaphora is a much-studied case in which the semantics underdetermines meaning leaving conversants to resolve things further. But there are many other instances of this, including for example the choice of comparison class for some adjectives and the choice of domain of quantification for quantifiers. Recent research in this last area, especially as it pertains to adverbs of quantification, has shown how complex and difficult things can get. Although I suspect that these last mentioned cases are more closely related to domain partitioning, I will be focussing here on a comparison with pronominal anaphora resolution because of its relative transparency.

In general, to make the referent of a pronoun salient by linguistic means, one has to mention it explicitly. It is not enough to have mentioned a group containing the individual, as the infelicity of the following example (out of the blue) shows:

(218) The boys think that he will win.

Intuitively, a hearer would have no way of determining which member of the group referred to by the subject of (218) is the intended referent of the
pronoun. Apparently this is enough to disqualify members of that group as choices for the pronoun's reference. This is reasonable enough. Somewhat surprising though, are contrasts like the ones in (219) and (220) from Partee (1989:footnote 13) and Carlson (1984:320) respectively, showing just how strong the effect is:

(219)  
a. One of the ten balls is missing from the bag. It's under the couch.

b. Nine of the ten balls are in the bag. It's under the couch.

(220)  
a. I did not catch all of the words. They were spoken too indistinctly.

b. I missed some of the words. They were spoken too indistinctly.

In (219a), as opposed to (219b), the remaining missing ball is not explicitly mentioned and hence is an unlikely candidate for reference by the pronoun it. In (220b), they is likely to refer to just the words missed but this is not possible in (220a) where that group of words is not explicitly mentioned.

The same kind of effect, I would claim, is found when one looks at potential choices for the value of a cover variable. An intermediate distributive reading is pretty much unavailable for an utterance in which the particular intermediate covering is not explicitly mentioned (or salient in the non-linguistic discourse). Consider the following:

(221)  
The children earned seven dollars.

Even though the children could be partitioned into two groups, say one male and one female, it is difficult, if not impossible to interpret (221) as involving distribution to these two groups. If I have that cover in mind, I must explicitly mention it as in:

(222)  
The boys and the girls earned seven dollars.

I have assumed here that to mention a particular covering is to name the cells comprising it. For the time being I will stick with that assumption, though other possibilities will be considered later on. Comparing (221) and (222), as in the pronoun case, intuitively, there are many intermediate coverings of the children, since the hearer of (221) has no way of knowing which the speaker might have in mind, none is available to serve as the value of the cover variable.

Now, although the mention of a particular entity is necessary to
make it salient, we know that this is not always sufficient. The following is a simple example of this aspect of anaphora:

(223) [The boys and the girls] entered the room (separately). They were wearing hats and they were wearing skirts.

The subject of the first sentence contains three noun phrases, referring, on all accounts, to three entities: the boys, the girls and the children. Nevertheless, the plural pronouns that follow cannot be used, in the absence of other contextual clues, to refer to the boys or to the girls. Exactly why this should be the case, I don't know. Comparison of this example with the following example in which the pronoun can refer to the boys:

(224) The boys told the girls that Mary took their hats.

suggests that one way or another, the correct story for (223) will have to take into account the syntactic relationships among the three noun phrases (cf. Smaby (1979) for an attempt at such a theory). Whatever one says about (223), it shows that simply mentioning something doesn't necessarily make it available for anaphoric reference with a pronoun.

So far, we have seen some examples where entities discussed are not likely referents for a following pronoun. Perhaps the more common situation is for a pronoun to have more than one possible antecedent. In this case, a host of factors come in to play in determining which is most likely to be chosen. These factors include appeal to extra linguistic knowledge as well as properties of the text or conversation itself, such as the relative proximity of the antecedent to the anaphor and general notions of textual coherence. In this last category, are cases in which a referent is chosen to reconcile implied or asserted contrast between utterances. This effect is seen in the following pair:

(225) a. Bill is coming for dinner. John is coming too and will bring his book along.
   b. Bill thinks Sam will arrive at 8:00, but John thinks he will arrive at 9:00.

In a., one tends to choose John as the referent for the possessive pronoun, for he is last mentioned and is the subject of the predicate containing the pronoun. In b., John is not the most likely referent of the pronoun, but rather Sam is. In b., John's thought is being contrasted with Bill's. The two thoughts are analogous and hence contrastable if one takes the
pronoun to refer to Sam here.

In light of the foregoing observations, we now turn to some parallel considerations affecting the relative availability of various distributive interpretations. First, consider the following pair, based on the examples at the very beginning of our discussion of distributivity, about a group of women athletes and authors:

(226) a. The authors and the athletes outnumbered the politicians.
b. The authors and the athletes entered the room through different doors. We realized at once that they outnumbered the politicians.

(226a) is interpretable as meaning that the authors outnumbered the politicians as did the athletes. The cells of a cover of the women are explicitly mentioned, and this cover is then assigned as the value for the cover variable. This interpretation is not readily available in (226b). Paralleling what we said above, it appears that while mentioning a cover is necessary, it is not sufficient. In b. the noun phrase in which the cover is mentioned is further away from the verb phrase outnumbered the politicians than is the noun phrase they. Presumably, this works against the assignment of that cover to the variable in the verb phrase.

The following example, based on one from Barry Schein (p.), is a particularly surprising case in which a mentioned cover is nevertheless unlikely to produce the relevant reading:

(227) The vegetables, which are the beets and the carrots, weigh 5 lbs.

Even though the partition into beets and carrots is mentioned, Schein would find the intermediate distributive interpretation impossible. Here, the fact that the covering is mentioned parenthetically and perhaps ‘outside’ the clause containing the cover variable (cf. McCawley 1982) seems reduce its saliency.27

If I understood him correctly, the source of the problem in Schein’s view is that explicit mention of a covering is insufficient and what is needed are individuating events. Schein’s actual example was:

(i) *The integers which are odd numbers and even numbers are (all)
equenumerous.

which he claims cannot mean:

---

27If I understood him correctly, the source of the problem in Schein’s view is that explicit mention of a covering is insufficient and what is needed are individuating events. Schein’s actual example was:

(i) *The integers which are odd numbers and even numbers are (all)
equenumerous.

which he claims cannot mean:
Distributivity

Now just as anaphora resolution is amenable to extra linguistic reasoning and textual clues, so too is the choice of cover values. Compare, the example in (227), with the syntactically similar one below:

(228) The visiting players, who were Italians and Brazilians, outnumbered their opponents.

Here too the relevant covering is mentioned parenthetically, however here, the sentence does seem amenable to the intermediate distributive interpretation according to which the Italian players outnumbered their opponents as did the Brazilian players. Knowledge of what players and opponents do enters in in this case: the cells of the cover correspond directly to games played.

Besides extra linguistic factors, textual clues also have a role to play in the interpretation of plural discourse. Above, concerning (226), we observed that even when a particular intermediate covering is mentioned it may remain unavailable as a value for a cover variable if another cover is mentioned in an intervening noun phrase. However, this effect can be overturned in an effort to understand implied or asserted contrast among utterances in the discourse. The following examples are all cases where this seems to happen (here and there I have used capitalization to indicate contrastive stress):

(229) A: The beets, the potatoes and the carrots (all) cost less than the meat.
B: No, the vegetables (all) cost less than the BEEF, but not less than the chicken.

B’s reply is readily interpreted with distribution to the kinds of vegetables mentioned in A’s comment. This interpretation would be unlikely out of context. Similarly, in the following:

---

(i) The odd integers and the even integers are equinumerous

These examples do not involve (simple) distributivity, (i) doesn’t say that the odd integers are equinumerous and the even integers are equinumerous) hence I used a different example in the text. However, (i-ii) do involve reciprocity and are therefore relevant to the discussion in the next chapter.
(230) The administration thinks that the physics instructors AND the math instructors cover five courses. In fact, THOSE instructors cover only TWO courses. Only the English teachers cover five courses.

In the next example, a contrastive statement receives a paired-covering interpretation which would otherwise, in the absence of contrast, be unlikely:

(231) We expected the male pigs and the male goats to be more numerous than the female pigs and the female goats. In fact, the males were LESS numerous than the females.

On the relevant interpretation of the second statement in (231), the sentence is false, if, for example, there are more male goats than female goats, even if the total number of male animals is less than the total number of female animals. The paired-cover introduced in the first utterance, is carried over into the second. Although these effects may be related to the process of quantifier domain restriction encountered in the interpretation of a noun phrase, they cannot be reduced to that process. The particular interpretation arrived at here could not be achieved by narrowing or widening the set of males considered or the set of females considered.

This completes our comparison of pronominal anaphora and distributivity effects. The purpose here was not to develop an explicit theory of discourse that assigns a saliency ranking to potential pronoun or cover-variable interpretations. Rather, the goal was simply to further demonstrate that distributivity indeed behaves like a pragmatic phenomenon. As such, the presence or absence of a particular distributive interpretation is to some extent dependent on the kinds of things that more familiar cases of pragmatic phenomenon depend on.

5.8 Conclusion: The Union and Sets Theories Reconsidered

In the preceding pages, I have developed an analysis of distributivity within the context of the union theory. It is now time to return to our main theme, the choice between the sets and the union theories. Recall, distributivity was raised as an issue for us, with the following pair of examples:

(232) The authors and the athletes outnumbered the men.

(233) The women outnumbered the men.
Distributivity

It was observed that even in a situation where the women just are the authors and the athletes, for (232) but not for (233), there is a distributive interpretation according to which the authors as well as the athletes outnumbered the men. This difference is explained on the sets theory by the fact that the subject of (232) but not of (233), denotes, or may denote, a set of two entities. We need only assume that a sentence is true on a distributive reading of the subject, just in case the verb phrase truthfully applies to each of the entities making up the referent of the subject and this assumption is needed anyway for simpler cases of distributivity. Since on the union theory the subjects in the pair above are coreferent in the situation described, the union theory appeared initially to be counterexemplified. This of course depended on a purely semantic analysis for distributivity, something that we now claim is incorrect. There is a pragmatic explanation available to the union theorist for the difference in the pair above. According to this explanation, as we have recently said, the relevant distributive interpretation requires the assignment to a variable in the VP of a cover that is salient in (232) but not in (233).

At this point, a proponent of the sets theory might argue as follows. Leaving aside the question of whether there is some pragmatic element to distributivity, at least in cases like (232-233), the sets theory is more desirable, since it makes very clear predictions, commensurate with the clarity of the data. The union theorist has to some extent avoided responsibility for the data here, by passing the problem off to some other part of the grammar or out of the grammar altogether.

To begin with the argumentation itself is a bit shaky. Sure, if we leave aside the evidence for a pragmatic analysis, then the union theorist appears to be avoiding responsibility. However, if we take that evidence into consideration, then the proponent of the sets theory needs to explain why the explanation here is semantic. But even if we grant the sets theorist a limited view of the data, things are not as smooth as might be suspected for that analysis. Trouble arises upon reconsideration of some of the examples provided in the last section, this time in terms of the sets theory. Recall, the modified version of (232-233), discussed above:

(226) a. The authors and the athletes outnumbered the politicians.
b. The authors and the athletes entered the room through different doors. We realized at once that they outnumbered the politicians.
(226a) and the last clause in (226b) differ in just the way (232) and (233) do.\textsuperscript{28} But notice, first, that the most natural interpretation of the first clause of (226b) is one where the authors came in through a different door than the one used by the athletes. On the sets theory, that interpretation arises because the subject denotes a set of two sets, one of athletes, the other of authors. Since the pronoun they in the second clause is anaphoric to the subject of the first clause, it will also denote this set of sets. But this is exactly the denotation that gives rise to the intermediate distributive reading in (233), yet the intermediate reading is missing here. In other words, like with (232-233), (226a) and the last clause of (226b) differ in interpretation, but unlike with (232-233) their subjects are coreferent. The only possible way out here is to assume that the pronoun cannot refer to a set of sets, but just to a set of individuals, in effect, making the pronoun here coreferent with subject of (233), the women. Besides being ad hoc, this cap on the space of pronoun denotations won't work, as the following shows:

(234) The authors and the athletes arrived simultaneously but they left at different times.

According to the sets theory, the pronoun in this case would again have to denote a set of sets.

The examples in (226) show that the sets approach presents little advantage over the union approach for explaining examples where the intermediate distributive reading is lacking. No less troublesome are cases discussed earlier where these readings are available, even though the relevant noun phrases lacked any conjunction. The chart example, repeated here:

(207) The fiction books in the chart complement the non-fiction books.

involved an interpretation where there was distribution over parts of the fiction books and parts of the non-fiction books. Aside from the technical

\textsuperscript{28}I rely here on the judgment of a reviewer of the manuscript for this book, who cited an example like this one as a problem for the pragmatic analysis of distributivity. The claim there was that intermediate readings do not arise if the most recently mentioned NP does not describe the intermediate partition. Note, if the cases differ and the intermediate reading is in fact available in (226b), then the discussion of this example in the previous section would require revision as well, and the reviewers initial point would not go through.
problem of extending the sets approach to the two-place case, if this interpretation was to be achieved purely in the semantics, then the noun phrases here would have to denote higher than first order sets, but nothing in the syntax justifies this. A similar point holds for (230), from our discussion of contrast in the previous section:

(230) The administration thinks that the physics instructors AND the math instructors cover five courses. In fact, THOSE instructors cover only TWO courses. Only the English teachers cover five courses.

(230) is interpretable as entailing that the physics instructors cover two courses as do their colleagues in math. A sets analysis for this would require the semantics to assign to the simple noun phrase those instructors a set containing two sets, one with the math instructors and one with the physics instructors. Such a semantics would undermine the claims made for the differences between (232) and (233), for it would allow the women to corefer with the athletes and the authors.

Initially, distributivity was presented as a semantic phenomenon with respect to which the sets approach appeared to have an advantage. What we have lately seen is that a semantic account within the sets approach fairs poorly overall. On the other hand, a case has been made for viewing distributivity as a pragmatic or semantico-pragmatic phenomenon. In particular, the work that is done by the richer ontology of the sets approach in fact should not be handled semantically at all. This leaves us with the simpler union approach, and a pragmatic theory of distributivity, which, along with other pragmatic phenomena, requires further analysis both in and outside of linguistics.