Towards a Uniform Strategy of Plural Quantification? Iterative Cumulativity and the Witness Condition

Eyal Hurvitz*

March 5, 1999

Abstract

This paper relates to the question whether there is a uniform method of collective quantification. Several approaches to plural quantification are presented as a background for Winter's (1998) unified *dfit* lifting strategy. The main proposal is that so-called *cumulative* readings, which are known to be non-iterative, are in fact derivable, under certain circumstances, by an iterative procedure. Cumulativity is shown to be one of the readings that are generated by Winter's *dfit* approach, which contains a *modificational* component called the *witness condition*. However, in order to get these readings, a cumulative operator has to be applied. This operator is needed to overcome an under-generation effect of the witness condition. Further, it is shown that otherwise iteration of the *dfit* strategy generates counter-intuitive readings. Some motivation is given for the cumulativity operator as well as for the underlying constraints.

1 Introduction

This paper is an attempt to answer the following question: is there a uniform strategy of plural quantification? The problem is to extend the notion of quantification over individuals to quantification over collections. The question is asked in the light of several approaches that have been proposed so far. There are two main aspects to this question.

The first aspect has to do with the multiplicity of readings. Apart from the fact that simpler approaches which assume a smaller degree of ambiguity should be preferred,[1] there

*My thanks: to my devoted supervisors, Jaap van der Does and Paul Dekker for their guidance and encouragement, for so many enlightening discussions, and for their comments; to Johan van Benthem, who can make things not only better but also possible; to Yoad Winter for stimulating discussions, comments and encouragement; to the friendly environment I had around me under the cloudy sky of Amsterdam: Carlos Areces (to whom I'm also grateful for his comments), Rosella, Alessandra, Nikos, Paulo (who killed all the zombies on my way) and Andreas; to the ILLC staff and administration who welcomed me warmly: Dick de Jongh, Ingrid van Loon, Marco de Vries and many others.

[1] Moreover, there is an embarrassing methodological question raised in Link (1991) as to where to draw the demarcation line between different readings on the one hand and different types of situations on the other hand.

is also a problem of over-generation of readings, which challenges the empirical adequacy of theories.

The second aspect has to do with the quantificational structure of natural language; and in particular with the question: is it possible to derive *all* the readings we get in natural language by iterative means? This question is crucial as iterative quantification is one of Frege's main contributions, and plays a central role in fragments dealt with in the framework of Montague Grammar.

This paper focuses on the second aspect of the main question, and in particular tries to deal with one kind of readings which have been claimed to be non-iterative, namely *cumulative* readings. However, the two aspects are shown to be closely related.

As far as the first aspect is concerned, I argue that the unifying attempt as implemented in Winter (1998) is problematic when the strategy described there is iterated. Although there is a motivated way to handle this diffuculty, namely an introduction of a cumulative operator, this is not generally applicable since closure under cumulation is not a general property of predicates. On the other hand, uniformity is successfully achieved to some extent with the help of this operator.

As far as the second aspect of the question is concerned, I argue that cumulative readings do not require a new form of quantification. Rather, they are derivable iteratively, under the assumption that the predicates involved are cumulatively closed. However, this restriction, along with further constraints, points to the fact that the availability of cumulative readings is highly restricted to special circumstances.

The paper begins with an introduction to atomic quantification in section 2, and highlights the notions of iterative vs. non-iterative quantification, with cumulative readings naturally chosen as an example for the latter. Keenan's (1995) result about the non-iterativity of these readings is quoted and the required parts that are not given in Keenan are proved.

Section 3 introduces collective quantification. It begins with a method of classification proposed in Winter (1998) and then sketches several approaches to plural quantification proposed by Scha (1981), Van der Does (1992) and Verkuyl and Van der Does (1995). The section ends with a definition of collective cumulation and some comments on the status of the empty set as an argument of predicates, which will become relevant in the sequel.

Section 4 deals with problematic aspects of some of the lifting operators proposed before. It then proceeds to describe Winter's *dfit* lifting strategy and its relation to previous proposals. One component of this strategy, namely the witness condition, is then examined thoroughly and modified in several ways. The last parts of the section discuss a problem associated with the witness condition which gets connected both to some problematic aspect of previous approaches and to cumulativity, which is the subject of the next section.

Section 5 contains the main proposal, i.e. an iterative way of deriving cumulative readings. It begins with an extension of Winter's approach to transitive verbs and deals with the problems associated with this extension. Cumulative closure is proposed and motivation for it is discussed, followed by an extensive discussion of the interaction between the *dfit* strategy and cumulativity closed relations. The proposal is first sketched informally and its scope and generality are discussed at length. Then the formal proofs are given, followed by an extension of Keenan's result to collective quantification and a discussion of how this result is settled with the current findings.

2 Atomic Quantification

This section is structured as follows. It begins with a brief introduction into monadic *atomic* quantification, as originally dealt with in the theory of generalised quantifiers (Barwise and Cooper (1981)). Then it introduces the idea of iteration as a form of polyadic quantification, through the notion of type $\langle 1 \rangle$ functions. It then discusses a case of non-iterative quantification called *cumulative* quantification. A characterization theorem due to Keenan (1995) is presented and Keenan's conclusion that cumulative quantification is not iterative is proved.

The original implementation of the theory of generalized quantifiers (Barwise and Cooper (1981)) was basically designed to deal with sentences of the general pattern Det-N-VP in which the noun N as well as the verb phrase denote sets in a given domain of objects, and Det denotes a relation between sets; or, alternatively, functions from sets to sets of sets: this is the approach we take in most of this paper. For illustration it is sometimes more convenient to use the relational terminology. For example, the determiner *all* denotes the inclusion relation between subsets of the domain. Hence the sentence (1a) gets interpreted as in (1b): the set denoted by 'girls' is included in the set denoted by 'sleep'. In the next examples: "some" denotes a relation between sets such that two sets are related if they are not disjoint; "at most 6" denotes a relation between sets such that their intersection contains at most 6 elements; and finally, "no" denotes a relation between sets such that 2 sets are related if they are disjoint.

(1) a. All girls slept.
b. $G \subseteq sleep$

(2) a. Some girls sleep.
b. $G \cap sleep \neq \emptyset$

(3) a. At most 6 girls sleep.
b. $|G \cap sleep| \leq 6$

(4) a. No girl slept.
b. $G \cap sleep = \emptyset$

The formal definition of a determiner is as follows.

Definition 1 *A global determiner D is a functional which assigns to each domain E a function D_E from subsets of E to sets of subsets of E:*

$$D_E : \mathcal{P}(E) \to \mathcal{P}(\mathcal{P}(E))$$

We will generally constrain the notion of a determiner given in the definition above by assuming several constraints that will characterize a certain class called *logical determiners.* One property which is considered to hold universally of determiners in natural language is *conservativity.*[2] The intuitive idea behind it is that the first argument of a determiner "sets the stage" of quantification. This argument is also called the *restrictor.* The elements of the second argument which are relevant for quantification are those which have the property denoted by the first argument.

Definition 2 *A global determiner is* conservative*, iff for all E and all $A, B, B' \subseteq E$:*

$$A \cap B = A \cap B' \Rightarrow [D_E(A, B) \iff D_E(A, B')]$$

According to an equivalent definition which is very useful in practice:

Definition 2′ *A global determiner is* conservative*, iff for all E and all $A, B \subseteq E$:*

$$D_E(A)(B) \iff D_E(A)(A \cap B)$$

Next comes the principle of *extensionality.* Informally, it says that the meaning of determiners is constant, and does not change from one situation to another.

Definition 3 *The principle of* extensionality *holds for a global determiner D iff for all domains E, E' and all $A, B \in \mathcal{P}(E) \cap \mathcal{P}(E')$:*

$$D_E(A)(B) \iff D_{E'}(A)(B)$$

The principle of extensionality allows us to drop the domain subscript, and use freely D instead of D_E. This is so since extensionality implies that it is not necessary to relativize determiners to a particular domain E: the determiner holds of sets A and B independently of the domain. So the property of extensionality will be assumed from now on.

The third principle which will be assumed is called *permutation invariance.* Intuitively, it means that determiners should not be sensitive to permutations of the domain, since determiners, rather than being about certain individuals, determine quantities. This is captured by requiring that determiners should be indifferent to permutations of their arguments.

Definition 4 *A determiner D is* permutation invariant*, if for all domains E, all $A, B \subseteq E$, and all permutations $\pi : E \longrightarrow E$:*

$$D(A)(B) \iff D(\pi(A))(\pi(B))$$

Definition 5 *The class of* logical determiners *are those determiners which abide by the following three principles:* conservativity, extensionality *and* permutation.

[2]But some counter-examples have been pointed out. The most famous one is "only". Proponents of conservativity as a universal often argue that "only" and other potential counter-examples are not determiners.

As mentioned above, we will consider only *logical* determiners throughout the paper, so the three properties specified in the definition will henceforth be taken for granted.

For the sequel it would be useful to define some other properties that some determiners have. The first one is *upward-monotonicty*, followed by the analogous downward-monotonicity:

Definition 6 *A determiner D is upward-monotone (or monotone-increasing) in the second argument iff for each A, B and C:*

$$D(A)(B) \wedge B \subseteq C \Rightarrow D(A)(C)$$

Definition 7 *A determiner D is downward-monotone (or monotone-decreasing) in the second argument iff for each A, B and C:*

$$D(A)(B) \wedge C \subseteq B \Rightarrow D(A)(C)$$

Another two properties that will be used are *continuity* and *week-monotone-monotonicity.*

Definition 8 *A determiner D is continuous iff for all Y,A,B,C:*

$$D(Y)(A) \wedge D(Y)(B) \wedge A \subseteq C \subseteq B \Rightarrow D(Y)(C)$$

Definition 9 *A determiner D is weekly monotone decreasing iff for all N, Y:*

$$D(N)(Y) \Rightarrow \forall y \in Y : D(N)(\{y\})$$

The last property is *positivity*:

Definition 10 *A determiner D is positive (relative to A) iff for each set B:*

$$D(A)(B) \Rightarrow A \cap B \neq \emptyset$$

We say that D is positive *if it is positive relative to every set A.*

2.1 Iterative quantification

The examples so far illustrated *monadic* quantification only. That is, each quantifier could be considered to be a set of *1-place* predicates. However, in the general case, we would like to deal with *polyadic* sorts of quantification, and to deal with more than one quantifier in a given sentence. Consider for instance:

(5) (Exactly) 3 children wrote (exactly) 4 poems

On what is known as its *iterative* reading, it means that the total number of children who wrote exactly 4 poems is 3. The interpretation is derived by applying a common mechanism of iteration (see Keenan and Westerståhl (1996: 871-873)). The following notation is used, for a binary relation R:

(6) $R_a = \{b : R(a, b)\}$

The iterative procedure for (5) goes as follows. First the denotation of the predicate *write-4-poems* is computed,[3] and then the matrix determiner (in this case the numeral 3) is applied to it. The denotation of *write-4-poems* is expressible in terms of the determiner as follows:

(7) $\{a : 4(poem)(write_a)\}$

This denotation simply contains all those who wrote exactly 4 poems. The next step is an ordinary application of the matrix determiner to this denotation, which is just a denotation of an intransitive verb:

(8) $3(child)(\{a : 4(poem)(write_a)\})$

In general, consider the scheme in (9), where P is a 2-place predicate (e.g. a transitive verb), with R as its denotation. D and D' are the determiners that occur in the subject and object position respectively, and N_1 and N_2 are the restrictors in these positions respectively. The interpretation of (9) is given in (10).

(9) $D - N_1 - P - D - N_2$

(10) $D(N_1)(\{a : D(N_2)(R_a)\})$

2.2 Non-iterative quantification

In the previous subsection iterative polyadic quantification was introduced. Several other forms of polyadic quantification have been proposed in the literature, including cumulative and branching quantification. Here cumulative quantification is presented as the paper focuses on this non-iterative form of quantification.

Consider the following sentence:

(11) 4 people lifted 10 tables

It has been proposed (by Scha (1981)) that (11) has a reading that can be paraphrased as follows: the number of people who lifted tables is 4, and the number of tables lifted by people is 10. This reading is called *cumulative*. The definition is as follows.

(12) $at-cum(D)(D')(N_1)(N_2)(R) := D(N_1)(dom(R \cap N_1 \times N_2)) \wedge D'(N_2)(rg(R \cap N_1 \times N_2))$

The expression $R \cap N_1 \times N_2$ occurs in both conjuncts. It is the restriction of the relation R to the pairs of the form $< M_1, M_2 >$, where $M_1 \in N_1$ and $M_2 \in N_2$. In case of (11), the restriction is to pairs of people and tables. Now the first conjunct looks at the domain of that restriction of R. The result, again in case of (11), is the set of all people who lifted

[3] I ignore scope matters here, and I assume that the order of application is analogous to the linear syntactic order, or rather that the subject scopes over the object (as is widely assumed in generative theories of syntax like G&B, Minimalism, etc.).

tables. The second conjunct looks at the range of the restriction: the set of all tables that were lifted by people.

Unlike iterative quantification, cumulative quantification is known to involve quantifiers which do not scope over each other, and are hence scopeless. Keenan (1995) provides formal ways of characterizing iterative quantification. This allows him to conclude that there are constructions in natural language that involve non-iterative forms of quantification, that is polyadic forms of quantifications that are not reducible to iterative quantification. The next subsection defines this idea of reducibility in a precise manner and presents one of Keenan's characterization theorems.

2.3 Keenan's characterization theorem

In this subsection I present a characterization theorem due to Keenan (1995). The theorem characterizes a subclass of binary quantificational structures that can be reduced to an iteration of unary quantifiers.

First it is useful to define the notion of a type $\langle 1 \rangle$ function. Informally, a type $\langle 1 \rangle$ function stands for a generalized quantifier, standardly of type (ett), i.e. a set of sets. But presenting a generalized quantifier in this way is not general enough, since notationally it is appropriate for the unary case only. In a general perspective, a generalized quantifier can be viewed as a place reducer. For instance, a composition of a generalized quantifier and a 3-place predicate should yield a 2-place predicate, then a further combination with another generalized quantifier should reduces the structure to a 1-place predicate, which is in turned reduced to a "0-place" relation, which is simply a truth value. So for any generalized quantifier a type $\langle 1 \rangle$ function can be assigned, which will serve as a place- reducer. The formal definition of a type $\langle 1 \rangle$ function is as follows.

(13) $F \in$ type $\langle 1 \rangle$, iff:

- the domain of F is $\cup_{n \geq 0} R_{n+1}$, the range is included in $\cup_{n \geq 0} R_n$, and
- for all $n \geq 0$, all $R \in R_{n+1}$,
  $F(R) = \{< a_1, ..., a_n >: F\{b :< a_1, ..., a_n, b >\in R\} = 1\}$.

Next we bring the formal definition of type $\langle 2 \rangle$ functions. These functions stand for quantifiers which are sets of *binary* relations. F is a type $\langle 2 \rangle$ function iff:

(14)
- the domain of F is $\cup_{k \geq 0} R_{k+2}$, and
- the range of F is $\cup_{k \geq 0} R_k$, and
- for all $n \geq 1$, all $k \geq 0$, all $R \in R_{k+2}$:
  $F(R) = \{< a_1, ..., a_k >: F(\{< b_1, b_2 >:< a_1, ..., a_k, b_1, b_2 >\}) \in R\}$

Keenan's theorem characterizes the subclass of type $\langle 2 \rangle$ functions which are reducible to type $\langle 1 \rangle$ functions. Reducibility of type $\langle 2 \rangle$ functions means that they can be represented as a composition of two type $\langle 1 \rangle$ functions, like in the case of iterative quantification. What this characterization says is that the values of a composition of type $\langle 1 \rangle$ functions for all arguments can be determined on the basis of the values of the composition for products.

Proposition 1 *(Keenan 1995) (Reducibility Equivalence) For F,G reducible functions of type $\langle 2 \rangle$, $F = G$ iff for all subsets P, Q of E, $F(P \times Q) = G(P \times Q)$.*

PROOF: See Keenan (1995: 199) □

I use here the notation $it(D_1)(D_2)$ for the iteration of two standard determiners over a non-collective domain. The definition goes as follows:

(15) $it(D)(D')(N_1)(N_2)(R) \iff D(N_1)(\{a : D'(N_2)(R_a)\})$

Now we prove that for relations which are cross products, the iterative reading is equivalent to the cumulative reading. We assume that D and D' are positive.

Proposition 2 *For binary relations R over E which are* products*, and* positive *determiners D and D' the following equivalence holds:*
$at-cum(D)(D')(N_1)(N_2)(R) \iff it(D)(D')(N_1)(N_2)(R)$

PROOF: assume first that the cumulative reading holds of $R = A \times B$:

(16) $D(N_1)(dom(A \times B \cap N_1 \times N_2)) \wedge D'(N_2)(rg(A \times B \cap N_1 \times N_2))$

Note first that $D'(N_2)(B)$ holds. This is so since from the second conjunct of (16), $(A \times B) \cap (N_1 \times N_2)$ which can be rewritten as $(A \cap N_1) \times (B \cap N_2)$ and then taking the range which is $B \cap N2$ (because $A \cap N_1$ is not empty, otherwise the second argument of the D' in the second conjunct of (16) would be empty which is impossible since D' is positive). By conservativity this is what we intended to show.

Consider now the set $C := \{a : D'(N_2)(R_a)\}$. If $a \in C$, then $D'(N_2)(R_a)$ holds, and in particular, from positivity of D', $R_a \neq \emptyset$ (otherwise $N_2 \cap R_a$ would be empty, contrary to the definition of positivity). But since R is a product, R_a can be either $\emptyset$ if $a \notin A$, or it must be equal to B. So we conclude that $a \in A$. Moreover, we conclude that $D'(N_2)(B)$ holds. On the other hand, if $a \in A$, then $R_a = B$, and since $D'(N_2)(B)$ holds we conclude that $a \in C$. So we have shown that $C = A$. By assumption we have $D(N_1)(dom(A \times B \cap N_1 \times N_2))$, but as the last argument can be rewritten as $dom((A \cap N_1) \times (B \cap N_2)) = A \cap N_1$, we have: $D(N_1)(A \cap N_1)$, which by conservativity is equivalent to $D(N_1)(A)$, and in turn to $D(N_1)(C)$, which is exactly the iterative reading.

Conversely, assume the iterative reading:

(17) $D(N_1)(\{a : D'(N_2)(R_a)\}) = D(N_1)(C)$

¿From positivity of D we conclude that C is not empty and hence there must be an $a \in N_1$ such that $D'(N_2)(R_a)$. The latter implies that R_a is not empty (from positivity of D'), whence it must equal B and it must be the case that $a \in A$ (from the same considerations as in the converse proof above). Hence we conclude $D'(N_2)(B)$. On the other hand if $a \in A$ then $R_a = B$ and since we know already that $D'(N_2)(B)$ we conclude that $a \in C$. Again we have $A = C$, and hence we conclude $D(N_1)(A)$. By conservativity we get: $D(N_1)(A \cap N_1)$ and $D'(N_2)(B \cap N_2)$, which in conjunction are equivalent to the cumulative reading. □

Corollary 1 *(Keenan) Cumulative readings are not generally reducible. That is, it is not the case that for all determiners D and D' and NP-denotations N_1 and N_2 there are type $\langle 1 \rangle$ functions such that their composition is equivalent to the type $\langle 2 \rangle$ function which corresponds to the cumulative quantification involving D, D', N_1 and N_2.*

PROOF: We can prove the corollary by finding two positive determiners D and D' and two NP denotations such that $it(D)(D')(N_1)(N_2)$ is not equivalent to $cum(D)(D')(N_1)(N_2)$. Suppose D and D' are positive determiners. By proposition 2, $cum(D)(D')(N_1)(N_2)$ and $it(D)(D')(N_1)(N_2)$ are equivalent on *products*. Hence by proposition 1, if $cum(D)(D')(N_1)(N_2)$ were reducible, then $cum(D)(D')(N_1)(N_2)$ would be equivalent to $it(D)(D')(N_1)(N_2)$, contrary to the assumption.

Such examples are pervasive and not far-fetched. Take for instance D to be the numeral 5, D' to be the numeral 3, N_1 to be the set of children and N_2 to be the set of pianos. Clearly, D and D' are positive. Now it is easy to see that $it(5)(3)(N_1)(N_2)$ is not equivalent to $cum(5)(3)(N_1)(N_2)$.

To see that consider the relation:
$R = \{< c_1, p_1 >, < c_3, p_2 >, < c_4, p_3 >, < c_5, p_3 >, < c_2, p_3 >\}$, as the denotation of $lift$, and suppose $c_1, ..., c_5$ are children and $p_1, ..., p_3$ are pianos. Then $cum(5)(3)(N_1)(N_2)(R)$ holds, but $it(5)(3)(N_1)(N_2)(R)$ does not. Namely, it is true that 5 children lifted pianos and that 3 pianos were lifted by children (the cumulative reading), but it is false that the number of children who lifted 3 pianos is 5; actually it is 0. □

3 Collective Quantification

This section has the following structure. It first presents Winter's (1998) classification of predicates into the categories "atom" and "set". The second category corresponds, for the purposes of this paper, to collectivity. The section proceeds to sketch some proposals to cumulative quantification. Van der Does' definition of collective cumulativity is given, followed by some comments on the status of the empty set as an argument, to become relevant in subsequent sections.

The traditional analysis presented in the previous section does not cover cases of so-called *collective* predicates which take *plural* arguments, such as *gather*, *meet*, *be-similar* or *drink-together-a-whole-glass-of-beer*. It is often suggested to include plural individuals in the universe, and these individuals often get the type $< e, t >$, i.e. the type of sets, in contrast with singular individuals which are simply of type e. Note that it is not sufficient to suggest that some NPs denote plural individuals, which can combine with collective predicates. This is so since quantification can also occur with respect to these kind of NPs and predicates:

(18) All students met on the street.

(19) Exactly 7 girls are similar.

So the main problem is how to relate plural noun phrases with collective predicates.

3.1 Classification of predicates

A pervasive view in the literature classifies predicates into several categories, with the main classification into the categories of *collective* and *distributive predicates*. Under the 'distributive' title one can find predicates like: *sit*, *walk*, *smile*; under the 'collective' title one finds predicates like: *gather*, *be-numerous*, *lift-a-piano-together*, *like-each-other*, *be-a-good-team*, as in the following sentences:

(20) a. The children gathered on the street.
b. The children are numerous.
c. The children lifted a piano together.
d. The children like each other.

The idea behind the 'collective' category seems to be that these predicates have collections as their arguments and that these predicates are involved in predications which are not reducible to predication of individuals only. The following pairs might show a formal contrast between the two categories. In (22) we arguably have equivalence,[4] while in (21) we do not, as it is not clear whether (21b) is intelligible, unlike (21a). Equivalence is also lacking in (23).

(21) a. The children gathered.
b. # Every child gathered.

(22) a. The children smiled.
b. Every child smiled.

(23) a. The children are a good team.
b. # Every child is a good team.

The problem with the test is precisely the judgments of the sentences in (22). As Dowty (1987) concludes, even when a distributive predicate holds of a set denoted by a definite NP, there is some vagueness as to of how many individual members of the set the predicate holds.

Winter (1998) proposes a different distinction between two classes of predicates: *atom* predicates on the one hand, and *set* predicates on the other. The new distinction is based on the following *linguistic* test.

(24) A predicate B is an atom predicate $\Longleftrightarrow$ the sentence "All A are B" is equivalent to the sentence "Every A is B".

[4] (22a) and (22b) are not always equivalent, as while (22b) entails (22a) the inverse is not necessarily true. This weak reading of (22a) which does not imply that *every* child smiled may be termed *vague*; this issue will come up again later.

This definition is based on an empirical linguistic test. It provides linguistic motivation for a distinction between two kinds of predicates. Notice that it is not required that *every* and *all* have different denotations as determiners: in fact they can be considered to be equivalent. However, they clearly differ in their distribution in English.[5] Within the confines of this paper, Winter's dichotomy corresponds to the traditional one: the atom predicates we will deal with are traditionally distributive predicates. In any case, set predicates a la Winter are always traditionally collective predicates.

Consider some examples for illustration. (25a,b) are equivalent, while (26a,b) are not.

(25) a. All children smiled.

b. Every child smiled.

(26) a. All children gathered.

b. # Every child gathered.

So far Winter's distinction overlaps with the traditional one: atom predicates correlate with distributive predicates, and set predicates correlate with collective predicates. However, the point of departure comes in (27), in which *be-a-good-team* is deemed as an atom predicate by Winter's distinction, whereas it is traditionally a *collective* predicate.[6] Note that in this case both expressions are "equally marked". Winter's criterion can be rephrased somewhat more plausibly, perhaps "is equivalent" should be replaced by "does not lead to two unmarked non-equivalent expressions".

(27) a. # All children are a good team.

b. # Every child is a good team.

The first step of the analysis of plural quantification deals with the task of relating predicates with their arguments. For this purpose, the following distinction is made between predicates on the basis of their morphological number:

(28) – Singular predicates, such as *student*, *walk* (which also have an identical plural form), $is - pretty$, unambiguously denote objects of type et

[5] Although "every" and "all" are assumed to have the same denotations as determiners, "every" is never lifted to a higher type which enables it to quantify over collections. The reason is that such a lifting on Winter's approach requires that the restrictor be in the plural; otherwise it cannot be of type (ett). However, it is a syntactic fact of English that "every" can never occur with a plural argument.

[6] Predication such as in "The students are a good team" is claimed by Winter to be based on viewing the set of students as an *impure* atom (see for instance Landman (1989)). On this view, the "extra structure" we associate with "the students" (or the bunch nouns "the team", "the committee") is not different than the structure we may associate with any other atom, such as "a table". Thus, collective effects are possible even with atom predicates. The importance of Winter's distinction, to my mind, is that the focus is put on the way quantification actually works in natural language: whether over atoms or over collections.

– Plural predicates, such as $students$, $friends$, *sleep* (with a plural morphological agreement) are ambiguous between objects of type *ett* and objects of type *et*. They have a basic denotation from which the other denotation derives by means of some operator. The type of the basic denotation depends on the category: set predicates are *ett* objects in their basic denotation; atom predicates are basically *et* objects.

Two operators are defined which map objects between the two types: *pdist* maps atom predicates into set predicates and *sg* has the reversed functionality. *pdist* gives the set of all non-empty subsets of its argument. In is in fact the operator $\mathcal{P}^+$. *sg* gives the set consisting of the members of each singleton in the argument.

(29) – $pdist(A) = \{B \neq \emptyset : B \subseteq A\} = \mathcal{P}^+(A)$

– $sg(\mathcal{A}) = \{x : \{x\} \in \mathcal{A}\}$

3.2 Some proposals for collective quantification

We begin with one approach due to Scha (1981), which provides a semantics of numerals. I am following Van der Does (1992), reprinted as Van der Does (1998), in the terminology and presentation of Scha's ideas. Each numeral can be assigned three different readings: distributive, collective and neutral. The variety of readings is motivated by considerations of the kind described below. Consider

(30) 10 students met in the hall.

This sentence could mean that a bunch of 10 students met in the hall as a group, or it can mean perhaps that more than one meeting took place and hence there was not necessarily a single meeting with 10 students involved. Consider now the following sentence, with the predicate *lift*, which can have plural as well as singular arguments.

(31) 4 men lifted 2 tables.

The sentence seems to have several readings: the set of 4 men, for example, can totally distribute, and thus the sentence could mean that each of the 4 men lifted 2 tables. For instance it can be the case that a total number of 8 tables were lifted, if the tables did not overlap. On the other hand, the set of 4 men may be fully kept integral, which gives rise to a collective reading, on which only 2 tables were lifted, possibly *together* (although it is not clear to what extent this is relevant[7]), by the set of 4 men.

Before going into the details of Scha's analysis, let us fix a domain E of quantification, which represent the universe of objects. Over this domain, three types are most common: small letters will be used for variables of type e, capital letters for variables of type $< e, t >$ and styled capital letters for variables of type $<< e, t >, t >$.

[7]In fact, I suggest that verbs like *lift* do not really receive collective arguments in their second place. On Winter's (1998) analysis, the source of ambiguity is primarily the predicate, whose different arguments may be set or atom- which correspond partially to collective and distributive.

On Scha's analysis the source of ambiguity is the NP, that is the numeral. His analysis of numerals goes as follows, where D,C and N are used for distributive, collective and neutral respectively. The following are the readings he assigns to numerals n (considered as determiners):

(32) (D) $\lambda X \lambda \mathcal{Y}.|\{d \in X : \mathcal{Y}(\{d\})\}| = n$
(C) $\lambda X \lambda \mathcal{Y}.\exists Y \subseteq X[|Y| = n \wedge \mathcal{Y}(Y)]$
(N) $\lambda X \lambda \mathcal{Y}.|\cup \{Y \subseteq X : \mathcal{Y}(Y)\}| = n$

In the plural setting, determiners assume a different type. Van der Does proposes several functions that operate on standard determiners to yield determiners of a higher type. While the type of a standard determiner is $(et)(ett)$, the type of a lifted determiner fitted to the plural setting is $(et)(ettt)$. Since every application of a lifted determiner is reduced to an application of a standard determiner, there will be no need to redefine determiners. As we have already fixed a domain, relativization to E will only be implicit.

The definitions in (32) can be reformulated as in Van der Does (1998: 248) in terms of their numeral denotation. This way of formulating them is much more revealing and is suggestive of what the general treatment of determiners should be, beyond the numeral case. The (D) definition uses the notation AT, which is defined in (33).

(33) $AT := \lambda X \lambda Y.Y \subseteq X \wedge |Y| = 1$

(34) (D) $\lambda X \lambda \mathcal{Y}.n(X)(\cup(\mathcal{Y} \cap AT(X)))$
(C) $\lambda X \lambda \mathcal{Y}.\exists Z \subseteq X[n(X)(Z) \wedge Z \in \mathcal{Y}]$
(N) $\lambda X \lambda \mathcal{Y}.n(X)(\cup(\mathcal{Y} \cap \mathcal{P}(X)))$

For example, consider (35). On its distributive reading, it means that the total number of individual students x such that x met the dean is 8. $AT(student)$ is just the set of all singletons of students. So after applying (34D) to $X = student$ and $\mathcal{Y} = meet - the - dean$, we get (36D). The collective and neutral readings are given in (36C,N) respectively.

(35) 8 students met the dean.

(36) (D) $8(student)(\{x : \{x\} \in meet - the - dean\})$
(C) $\exists Z \subseteq student[8(student)(Z) \wedge Z \in meet - the - dean]$
(N) $8(student)(\cup\{X \subseteq student : X \in meet - the - dean\})$

Consider now the more complex example (31) with a transitive verb. Such sentences have $3^2 = 9$ readings on the present approach, where scope ambiguities are put aside (otherwise there are 2 possibilities for each scope relation, and the number of readings amounts to 18). For instance, the double collective reading is given in (37b). It is derived by means of polyadic iterativity. The iterated part is given in (37a).

(37) a. $\mathcal{B} := \{Y : \exists U \subseteq table(2(table)(U) \wedge lift(Y, U))\}$

b. $\exists Z \subseteq men[4(men)(Z) \wedge Z \in \mathcal{B}] \iff$
$\exists Z \subseteq men[4(men)(Z) \wedge \exists U \subseteq table(2(table)(U) \wedge lift(Z,U))] \iff$
$\exists Z \subseteq men \exists U \subseteq table[|Z| = 4 \wedge |U| = 2 \wedge lift(Z,U)]$

Not all options are wanted. Consider for example the Neutral $\times$ Distributive case:

(38) a. $\{Y : 2(table)(lift_Y)\}$

b. $4(men)(\cup\{X : X \subseteq men \wedge 2(table) \cup (lift_X \cap AT(table))\}) \iff$
$4(men)(\cup\{X : X \subseteq men \wedge |\{d \in table : lift(X, \{d\})\}| = 2\})$

By taking a rich enough cover of the set of 4 men, the number of tables may rise up to $(2^4 - 1) \times 2 = 30$ (each non-empty set of men out of a set of 4 men lifted 2 different tables). This is on the assumption that the empty set of tables is not allowed as either a table or a lifted table. It seems to counter the intuition that at most 8 tables are involved in the situation. Hence there seems to be a problem of over-generation of readings here.

The next step taken in Van der Does (1998) is to generalize Scha's proposal by offering lifting operators, which operate on standard determiners to yield new determiners adjusted for different readings of plural NPs. Here we only deal with the three lifts, out of six, which are selected by Van der Does to give reasonable readings. These lifts correspond to the NP readings above ((32) and (34)).

(39) (D) $\lambda D \lambda X \lambda \mathcal{Y}(D(X)(\cup(\mathcal{Y} \cap AT(X))))$

(C) $\lambda D \lambda X \lambda \mathcal{Y}(\exists Z \subseteq X[D(X)(Z) \wedge Z \in \mathcal{Y}])$

(N) $\lambda D \lambda X \lambda \mathcal{Y}(D(X)(\cup(\mathcal{Y} \cap \mathcal{P}(X))))$

In the next sections Winter's approach (Winter (1998)) will be dealt with extensively. One of its seemingly attractive features is its uniformity. Instead of the standard atomic interpretation, a single mechanism is suggested to adjust determiners for the collective case. It is interesting to point out the close relation between Winter's lift and the ones proposed by Van der Does. Winter explicitly mentions that his approach is based on Van der Does' *neutral* reading, but he attempts to provide formal semantic motivation for the operations involved in this reading.

But this is the right place to mention a previous attempt to unify the three readings under discussion. This proposal is due to Van der Does and Henk Verkuyl (1995). They propose the notion of *covers*, and *partitions*. Let X be a restrictor and $\mathcal{Y}$ be a collective predicate. Then the following specifies the way quantification works with respect to these arguments and a standard determiner D:

(40) $\exists Y \subseteq X[D(X)(Y) \wedge \exists \mathcal{Z} cv Y : \mathcal{Z} = \mathcal{Y} \cap \mathcal{P}(X)]$

The relation $\mathcal{Z} cv Y$ means "$\mathcal{Z}$ covers Y" which in turns simply means that $\cup \mathcal{Z} = Y$. It is easy to see that (40) is equivalent to the neutral reading of X.

These authors consider a strengthening of the cover-reading, which turns out to be crucial in avoiding the wide-scope problem associated with the neutral reading, as exemplified by (31). The new reading is called a *partition* and it is obtained by replacing the *cv* relation by the *partitions* relation. A partition is a special case of a cover, and is defined as a cover with disjoint cells. The examples below will involve *partitions*. The definition of a partition goes formally as follows:

Definition 11 *$\mathcal{Z}$ is a partition of Z iff:*

(i) $\forall X, Y \in \mathcal{Z}[X \neq Y \Rightarrow X \cap Y = \emptyset]$

(ii) $\cup \mathcal{Z} = Z$

The definition of the partition reading is as follows:

(41) $\exists Y \subseteq X[D(X)(Y) \wedge \exists \mathcal{Z} partitions Y : \mathcal{Z} = \mathcal{Y} \cap \mathcal{P}(X)]$

Consider again as an example sentence (35). The predicate $\mathcal{Y}$ in this case would be *meet-the-dean*. The restrictor X is the set of students. At one extreme, we get a partition which verifies the sentence on its distributive reading. The partition of the set of students who met the dean (say: $s_1, ..., s_8$) in this case is as follows.

(42) $\mathcal{Z} = \{\{s_1\}, \{s_2\}, \{s_3\}...\{s_8\}\}$

At the other extreme we get a partition which verifies the sentence on its collective reading, when the partition is simply a singleton:

(43) $\mathcal{Z} = \{\{s_1, s_2, ..., s_8\}\}$

In between these two extremes we get the rest of the readings obtained by the neutral option. For instance, the following is one example of a partition.

(44) $\mathcal{Z} = \{\{s_1, s_2\}, \{s_3\}, \{s_4, s_5, s_6\}, \{s_7, s_8\}\}$

3.3 Collective Cumulation

Van der Does provides a definition of collective cumulative quantification. The idea is natural extension of atomic cumulation as defined in section 2.2. This time R is assumed to be a relation over collections. Consider the following sentence:

(45) 4 children lifted 6 tables

On the collective cumulative reading of (45), the total number of children who were involved in sets of children who lifted tables is 4, and the total number tables which were involved in sets of tables lifted by children is 6. The definition is as follows.

(46) $cum(D)(D')(N_1)(N_2) :=$
$D(N_1)(\cup dom(R \cap \mathcal{P}(N_1) \times \mathcal{P}(N_2))) \wedge D'(N_2)(\cup rg(R \cap \mathcal{P}(N_1) \times \mathcal{P}(N_2)))$

3.4 The empty-set problem

Consider the sentence:

(47) No students gathered on the street

This sentence is problematic for the collective reading. Such a reading would state that there exists a set X of students, such that $No(students)(X)$ holds and X gathered on the street. Hence, X is the empty set, and the collective reading would state that the empty set gathered. So in order for (47) to be true on its collective reading, the empty set has to be in the extension of the set predicate *gather*. We argue that it is just a matter of stipulation to require that the empty set should be a member of such denotations, as a minimal element. However, as far as the collective reading is concerned, such a stipulation will not solve the problem. The reason is that the empty-set is not "fine-grained" to encode information about the restrictor. Thus, "no students gathered" would come out equivalent, under the collective reading, to "no prisoners gathered".

This difficulty is used by Winter as an argument against the modificational strategy (see next section). However, as we shall see, complete refusal to admit the empty set into denotations of predicates gives rise to other difficulties. In particular, relaxing this harsh stance against the empty set may simplify part of Winter's approach. We shall also see in the sequel that this question has substantial implications concerning the cumulative reading.

At this point I would just briefly mention two arguments in favor of admitting the empty set, which will become relevant later. First, it would make a distributive operator more natural, in that such an operator would amount to the power set operation (which gives the ultra-filter of a set), instead of the *positive* power set operation which outputs all the *non-empty* subsets of its argument. Second, if we have good reason to accept a cumulativity principle to the effect that if the sets A and B gathered, then so is their union $A \cup B$, then it would make sense and become handy to have a minimal element in the denotation, namely the empty set.

Thus, I will postulate that the empty set appears in denotations of set predicates like *gather*, however unattractive this postulation seems to be. Notice again that for distributive predicates no such assumption is necessary, since by taking the ultra-filter of a set, the empty set is generated in any case. Moreover, there is also a way by which the empty-set is obtained "for free" also for predicates like *gather*, when the denotation is closed cumulatively.[8] This point will be discussed in section 4.7.[9]

4 Uniform collective quantification?

This section deals with problematic aspects of some of the lifting operators proposed before: first, the collective reading gives rise to what is known as Van Benthem's problem. Second, the neutral reading gives rise to problems when exercising wide scope. Then the section presents Winter's *dfit* lifting strategy, which is introduced as a way to deal with these

[8]However, this will not take care of verbs that have denotations to which cumulative closure is not applied.
[9]Cf. also section 5.4.2.

problems. This strategy is presented as a combination of the collective and neutral readings. The collective component of this lifting strategy is called the witness condition. It is discussed at length in terms of its plausibility. The maximality property of the witness condition is brought up, and finally some critical comments are raised with respect to this condition. In particular, the missing-witness problem is brought up as an under-generation problem, and an operation of cumulative closure is introduced as a solution to this problem.

4.1 The collective reading and Van Benthem's problem

Consider again the collective reading mentioned in the previous section:

(48) $C(D) = \lambda A \lambda \mathcal{B} \exists X \subseteq A(D(A)(X) \wedge \mathcal{B}(X))$

Such an approach to collective quantification is called *modificational*, or *existential-modificational* since it is based on existential quantification over collections. The determiner only applies (along with the restrictor) to the set X bound by this existential quantifier. The collective predicate $\mathcal{B}$ modifies X outside the scope of the determiner. As $\mathcal{B}$ is replaced by the existentially-closed X, this approach is called *modificational.* While it suggests a uniform way of lifting ordinary determiners, it also gives rise to a problem related to non-upward-monotone determiners, known as Van Benthem's problem (Van Benthem (1986)).

Consider the following sentences:

(49) a. Exactly 5 children gathered on the street.
b. Exactly 7 children gathered on the street.

Now the following are their interpretations on the modificational approach:

(50) a. $\exists X \subseteq child(exactly - 5(child)(X) \wedge \mathcal{B}(X))$
b. $\exists X \subseteq child(exactly - 7(child)(X) \wedge \mathcal{B}(X))$

The problem is that (50a) and (50b) are consistent, while sentences (49a) and (49b) are not, when interpreted relative to a single situation, at least on a quantificational reading of $exactly-5$ and $exactly-7$.[10] On the quantificational reading of (49a,b), the point is that the *total* number of children who gathered on the street is exactly 5 and exactly 7 respectively, whence the inconsistency.

The modificational approach seems to work well for monotone-increasing determiners. If D is monotone-increasing, then $C(D)(child)(gather)$ implies that if T is the total number of children who gathered, then $D(child)(T)$ holds as well. But it is not true of arbitrary determiners.

Hence it seems that a modificational approach is not generally available for collective quantification.

[10]Notice that bare numerals exhibit an existential behavior sometimes, and that they are not purely quantificational. Thus "3 men walked" may be interpreted as "There are 3 men who walked", which is equivalent to "At least 3 men walked". However, modified numerals behave differently in some contexts. Consider the contrast (due to Winter (1998)) between "5 students I know well are a good team", which is acceptable, as opposed to "Exactly 5 students I know well are a good team", which is not acceptable.

4.2 The neutral reading and scope problems

In this subsection we discuss a problem associated with neutral readings when exercising wide scope. But we begin with a closely related issue which will also be relevant for the sequel.

Winter (1997) contains a discussion of vagueness which is partly dedicated to accounting for the contrast between sentences such as (51):

(51) a. The men wrote operas

b. The men wrote an opera

Consider the following situation. Two men wrote an opera together, and then another men wrote another opera on his own. While (51a) is verified in this situation, (51b) is not. The sentences get the following interpretation, respectively:

(52) a. $\exists x(opera(x) \wedge wrote(M, x))$

b. $\exists X \subseteq OPERA(wrote(M, X))$

(52b) says that there is a set of operas that the men wrote. It does not go down into details regarding the individuals composing the plural object M. This is the sense in which the predication is "vague" with respect to how men and operas are related exactly.

In this example, there is no question about the scope of the quantifier 'opera' or 'operas', since the NP in the subject position is "referential", i.e. it denotes the plural object 'the men'. So the effect of the quantifier 'opera(s)' is as if it exercised matrix scope. This is different from cases in which a quantified NP occurs in the subject position, as we shall see below.

However, this contrasts occurs also in case of quantified subjects. Hence the vagueness solution proposed here is not sufficient.

(53) a. Exactly 5 people wrote operas

b. Exactly 5 people wrote an opera

(53a) means that the total number of people involved in sets of people who wrote operas is 5. However, (53b) does not have this reading. That is, it is not verified in a situation in which 3 people wrote together one opera, and then other 2 people wrote another opera. Furthermore, (53a) does not mean (54):

(54) There are operas that exactly 5 people wrote

Sentence (54), in which the object has wide scope, means something much weaker than the intended meaning of (53a): that exactly 5 people were involved in writing operas. Moreover, the contrast has nothing to do with an ambiguity between the determiner 'a' (or 'one') and 'at least one', as both (55a) and (55b) are equally strange in the same situation, because both imply that the 5 people worked *together*.

(55) a. There is an opera that these 5 people wrote

b. There is at least one opera that these 5 people wrote

Hence the vagueness proposal is not sufficient to establish the contrast between (53a) and (53b). Below we shall see that the problem with (53b) is related to a more general problem arising when neutral NPs have wide scope.

Van der Does (1998: 268-269) discusses neutral NPs and concludes that they are best behaved when they take *narrow* scope, or when they are scope-free (like he suggests they should be to achieve cumulative readings). To see the problem that arises when they exercise wide scope, consider (56). It has been argued plausible to put a limit to the number of operas involved in situations verifying (56). The most natural bound in this case is $3 \times 5 = 15$. However, on the iterative reading (where *write* is a set predicate on its first argument and an atom predicate on its second argument), it may be the case that the 31 non-empty subsets of a set of 5 men each wrote a distinct set of different operas. In such a case the number of operas could amount to $(2^5 - 1) \times 3 = 93$.

(56) 5 men wrote 3 operas

A similar problem would arise with respect to (51b). What blocks the 'neutral reading' in this sentence is that the singular object exercises wide scope. But when a quantified NP occurs in the subject position, and scopes over the object, like in (53), we again get 'neutral' verifying situations which we are trying to block.

The conclusion is that the neutral reading is problematic when it is applied in transitive sentences in combination with a distributive NP, and it therefore may not serve itself as a general strategy for plural quantification.

4.3 Winter's *dfit* strategy

Winter (1998) defines a single lifting operator which transforms determiners of type $(et)(ett)$ to determiners of type $(ett)(ettt)$. Thus collective quantification is taken care of by one lifting only. The definition is given below.

(57) **(Definition of *dfit*)** $dfit(D) = \lambda\mathcal{A}\lambda\mathcal{B}D(\cup\mathcal{A})(\cup(\mathcal{A}\cap\mathcal{B}))$

For example, consider the sentence "Five students lifted a piano". The truth conditions assigned to this sentence under the current definition are obtained as follows. First, the *dfit* operator yields the appropriate determiner which is then fed with the predicates *student* and $lifted - a - piano$ by students. The expression obtained is: $D(\cup student)(\cup(student \cap lifted - a - piano))$. In other words, the same determiner is applied to two sets: the first is the lowering of the set predicate *student*, whose extension is the set containing all the non-empty sets of students, namely $\mathcal{P}^+(student)$. This is achieved by the *pdist* operator in order to avoid the type mismatch. The desired lowering is obtained by simply taking the union set. In this case, the result is the atom predicate *student*. The second set is the union of the intersection of $lift - a - piano$ with *students*, which comes down to the set of all sets of students which lifted a piano, before applying the union. After applying the union operator, the result is the set of all students who participate in some lifting of a piano *by students*.

Hence the truth conditions thus obtained require that the number of students who took part in an activity of lifting a piano by students is (at least) five, for the sentence to be true. Note that if there were a set consisting of 5 students and one non-student who lifted a piano as a group (of 6 people), the proposition would come out as false, as it should, thanks to the intersection of the partaking sets with the restictor "students". However, consider the following sentence:

(58) All the farmers are similar.

According to the strategy sketched above, the sentence becomes true if every farmer is similar to at least one other farmer, because in such a case for each farmer, the extension of the predicate *similar* contains a set of farmers containing that farmer. This entails that the union of the intersection of *similar* with the set of farmers is precisely the set of farmers, namely *all* the farmers. Clearly, this is not appropriate. All the farmers are similar is equivalent to "the set of (all) farmers is a member of the extension of the set predicate *similar*". I will call this problem simply "the witness problem".

The witness problem is avoided by the modificational approach, which is actually Van der Does' collective reading given in (39C). Consider how this approach manages to derive the right reading of (58). Paraphrasing the modificational reading, we get: "There exists a set such that this set contains all farmers and this set is in the extension of the predicate *similar*. This is indeed an appropriate analysis. However, we have seen that the modificational approach is not available as a general strategy (although it deals successfully with upward-monotone determiners such as "all").

In order to solve the problem with sentences like (58), Winter proposes to strengthen $dfit$ by adding an extra condition, termed "the witness condition", which reads as follows:

(59) $witc(D)(\mathcal{A})(\mathcal{B}) \iff \mathcal{A} \cap \mathcal{B} \neq \emptyset \rightarrow \exists W \in \mathcal{A} \cap \mathcal{B}(D(\cup\mathcal{A})(W))$

The modified $dfit$ is given below.

(60) $dfit_w(D) = \lambda\mathcal{A}\lambda\mathcal{B}[dfit(D)(\mathcal{A})(\mathcal{B}) \wedge witc(D)(\mathcal{A})(\mathcal{B})]$

To illustrate the new definition, consider first the sentence in (58), "All the farmers are similar". The first step is to take, as the first argument of *all* the union of $pdist(farmer)$, which is just the set of all farmers; and the union of the intersection of *similar* with $farmer$. This produces all those farmers who are similar to at least another farmer. Now the witness condition is applied: the only witness set in the case of *all* is the set of all farmers. The set of all farmers is thus required to be in the extension of similar.

Winter's proposal applies successfully also for the case of "Exactly 5 students gather on the street", even though "exactly 5 students" is *not* monotone-increasing. The strategy without the witness condition produces the following reading: the number of students who belong to a set of students that gathered on the street is exactly 5. This guarantees that no more than 5 students gathered. On the other hand, the witness condition requires that there be a witness set that gathered. A witness set is any set of 5 students that gathered, hence there is at least one set of 5 students that gathered. Combined with the first result, this provides us with the desired result.

Notice that the pure *dfit* strategy (that is, without the witness condition) on the one hand and the witness condition on the other hand have complementary roles, in a sense. The pure *dfit* strategy analysis states that the total number of students who took part in a set of students who gathered is 5, which implies that each such set of children has at most 5 members. On the other hand, the analysis provided by the witness condition states that there is at least one such set, which contains 5 members. This gives us uniqueness: besides the set required by the witness condition, there are no other such sets.[11]

The complementary effects between the two components of the *dfit* lifting strategy are illustrated by the fact that Van Benthem's problem does not arise: the pure *dfit* component does not allow a free modificational analysis of sentences.

4.4 *dfit* in relation to the previous proposals

As Winter (1998) himself points out, there is a very close relation between *dfit* and the neutral reading of Van der Does. Further, the witness condition is closely related to Van der Does' *collective* reading.

Consider the special case of the bare *dfit* operator, obtained from (57) by substituting $\mathcal{A}$ for $\mathcal{P}(X)$ and $\mathcal{B}$ for $\mathcal{Y}$. In this form, it is interesting to see that *dfit* is identical to van der Does' neutral reading (39N). The result of the substitutions appear in (61). Although it is not the general case, it is still a very useful instance of *dfit*, covering all the cases in which the restrictor distributes.[12] Since we focus on distributive restrictors, we defer discussion of the general case to section 6. Next consider the witness condition as it is defined in (59). Apply the same substitutions as before, and then consider the simplification of this condition obtained by eliminating the exception for empty sets. This will be the new version of the witness condition adopted here. We shall see later on that it has implications for further issues concerning cumulativity. But notice that this modification is along the same lines of the adopted policy with respect to the empty set, which is not eliminated from extensions of predicates, as explained in section 3.3. The restriction to non-empty sets is introduced in order to avoid a problem that would arise with quantifiers like 'no student' in the sentence 'no students passed the exam'. A witness in this case, i.e. with no exception for empty sets, would be an empty set of students, and this empty set would have to pass the exam. The plural predicate *pass-the-exam* is obtained by the *pdist* operator, which produces all the non-empty set of its arguments. If we instead adopt an alternative operation, which is identical to the power set operation in the unary case, the empty set is generated "for free" and can be considered as a witness.[13]

[11] But cf. the role of maximality with respect to the witness condition, to be discussed in section 4.5; cf also section 5.4.1.

[12] Winter's approach also deals with non-distributive restrictors, such as friends, in the sentence "All friends are loyal". The denotation of *friends*, a set predicate, is lowered from type *ett* to type *et* by the union operator. Thus quantification is over every person who is a friend of someone. However, a difficulty arises in case the predicate (that is, the nuclear scope) is also of type (*ett*). For example: "All friends love each other". This sentence has a reading on which each collection (or pair) of friends is predicated of by the set predicate *love-each-other*.

[13] However, as we shall see in the sequel, the witness condition is sometimes associated with introducing

The new version appears in (62).

(61) $dfit(D) = \lambda X \lambda \mathcal{Y} D(X)(\cup(\mathcal{P}(X) \cap \mathcal{Y}))$

(62) $witc(D)(X)(\mathcal{Y}) = \exists Z \in \mathcal{P}(X) \cap \mathcal{Y}(D(X)(Z))$

Thus the $dfit_w$ approach is the conjunction of Van der Does' readings C and N:

(63) $dfit(D) = C(D) \wedge N(D)$

Although we will focus on the collective and neutral readings, it is in order to comment briefly on the distributive reading. Van der Does provides a special lifting operator to derive these readings. Winter, however, relies on ordinary quantification to get such readings, without using a lift; but on the other hand, he must make use of the operator *sg*. Consider the following example:

(64) Exactly 5 children drank a whole glass of beer.

In some dialects at least, the predicate *drink-a-whole-glass- of-beer* is originally a set predicate, of type (*ett*). By applying *sg* it becomes of type (*et*) (note that the verb occurs in this example in plural form and hence is ambiguous between the two types). Now ordinary quantification applies, without using the *dfit* lifting strategy.

There is one important comment to make about Winter's approach as opposed to the way it is applied here. On Winter's approach, *dfit* does not always apply: ordinary quantification still remains the default mechanism. However, it is possible to assume that *dfit* applies in all cases, by using the *pdist* operator, and applying *dfit*. Winter (1998) proves that equivalence is indeed guaranteed. Formally, the following holds:

(65) For any (logical) determiner D, and for all A and B:
$dfit(D)(pdist(A))(pdist(B)) \iff D(A)(A \cap B)$

It is also possible to extend this equivalence to the full $dfit_w$ approach, as it is very easy to see that $A \cap B$ is a witness (it is in the intersection of $pdist(A)$ and $pdist(B)$, and of course $D(A)(A \cap B)$).[14] In any case, the practice in this paper is a departure from Winter.[15]

4.5 The witness condition and maximality

There is evidence concerning nominal anaphora that points to the need to strengthen the witness condition by requiring the existence of a *maximal* witness sets (for maximality and anaphora see also Van den Berg (1996)). Consider the following sequence:

new references into the context. It is known that quantifiers like "no one" do not have this effect, or in other words they are dynamically opaque (that is, externally static). So this may provide an argument *in favor* of keeping the exception for empty sets.

[14]Later the witness condition will be strengthened, but the result will still hold. In particular, the proposed witness meet the strong and total maximality constraints that will be introduced in the sequel.

[15]Cf. section 5.4.4.

(66) At least 10 people came into the restaurant. They ordered beer.

The pronoun in the second sentence of (66) is bound to refer to the *maximal* set of people who came into the restaurant. That is, if 13 people came into the restaurant and only 11 of them ordered beer, then it is claimed that the sentence "They ordered beer" would not be acceptable. If dynamic interpretation of pronouns has to do with the introduction of witness sets into the context,[16] then the example can be considered as evidence for maximality of witness sets.

There are several alternative definitions of maximality. We will mention two possibilities. *witset* is here the set of potential witness sets, that is:
$witset(D)(\mathcal{A})(\mathcal{B}) := \{W \in \mathcal{A} \cap \mathcal{B} : D(\cup\mathcal{A})(W)\}$

(67) $\exists W \in witset(D)(\mathcal{A})(\mathcal{B})(\forall W' \in \mathcal{A} \cap \mathcal{B}(W \subseteq W' \rightarrow W' = W))$

(68) $\exists W \in witset(D)(\mathcal{A}))(\mathcal{B})(\forall W' \in \mathcal{A}\cap\mathcal{B}(W \subseteq W' \wedge W' \in witset(D)(\mathcal{A})(\mathcal{B}) \rightarrow W' = W))$

We shall refer to (67) as *strong* and to (68) as *weak* maximality. It is easy to see that *strong* maximality indeed entails *weak* maximality.

It remains unclear which approach to maximality should be chosen. Empirical data may shed some light over the issue, but they are barely conclusive themselves. Consider the following sentence:

(69) Between 10 and 30 children drank together a whole glass of beer.

Suppose there are two sets of children that drank together a whole glass of beer: one consisting of 7 children and one consisting of 15 children. The total number of children who took part in a set in the extension of $P = drink - together - a - whole - glass - of - beer$ is 22, which is within the boundaries of the determiner. Contrast this situation with another one in which again there are two sets in the extension of P, but this time they are both within the boundaries of the determiner, say one consisting of 12 children and one containing 10 children. On Winter's maximality proposal, the second situation is a violation of the uniqueness requirement- namely the requirement that there should be a unique maximal witness, while no such violation occurs in the first situation, just because the other set in the extension of P is not a witness, it is just an element in the intersection of the restrictor and the nuclear scope. Note however that while this extra set is in a sense ignored by the witness condition, it is still counted by the bare $dfit$ component. If there were, in addition, 2 other sets of 7 children each in the extension of P, then the total number of sets of children in the extension of P would total at 36, which lies beyond the boundaries set of the determiner 'between 10 and 30'. This would make the sentence false.

The question is whether it is empirically adequate to distinguish these two different ways in which elements in the extension of predicates affect the two components of the $dfit$ strategy. If such a distinction is not appropriate, it would motivate towards preferring (67) over (68) as a notion of maximality concerning the witness set. The approach taken here is

[16] In fact an interesting hypothesis to examine is that the witness condition belongs to the dynamic realm more than it should be considered as contributing to static truth conditions.

along the same lines. Judgments of the above sentence in the light of these various situations seem very delicate.

Another empirical test would be to examine maximality effects of anaphors. For instance, suppose (70) follows (69) directly:

(70) They got drunk.

What sort of maximality effect are shown by *they*? Does it refer to all witness sets, or to all sets of children in the extension of P?

Winter proposes an alternative definition of the witness condition, reinforced with a maximality requirement (as well as uniqueness of the maximal element:)

(71) $mwitc(D)(\mathcal{A})(\mathcal{B}) \iff \mathcal{A} \cap \mathcal{B} \neq \emptyset \to \exists W \in \mathcal{A} \cap \mathcal{B}[max(witset(D)(\mathcal{A})(\mathcal{B})) = \{W\}]$

Instead we adopt a much stronger notion of a maximal-witness condition, we call the *total witness condition*. The total witness condition means that the set $\cup(\mathcal{A} \cap \mathcal{B})$ should be a witness set. It is clear that this notion of maximality entails strong maximality. The total witness condition for D, $\mathcal{A}$ and $\mathcal{B}$ boils down to the following requirement:

(72) $\cup(\mathcal{A} \cap \mathcal{B}) \in (\mathcal{A} \cap \mathcal{B})$

This witness condition, rather than including an existential quantification over a witness, just picks what may be considered intuitively the biggest possible witness.

4.6 The adequacy of the witness condition and the missing-witness problem

4.6.1 "similar" and "different"

Reciprocals like "similar" provide the strongest argument in favor of the witness condition. The intuition about sentences like (73) is very clear. This sentence means that the set of all the rocks is in the extension of the predicate $be - similar$. However, it is not clear that it is adequate to derive the right reading by means of the witness condition.

(73) All the rocks are similar

To see how the adequacy of the witness condition can be challenged, consider the following pair of examples.

(74) a. The children are different

b. All the children are different

First let us see what interpretations are predicted for these two sentences by Winter's theory. (74a) is not quantificational. It simply states that the set of children is in the extension of the predicate *different*, which is a set predicate which bears plural morphology (simply apply the test in (24): # Every child is different $\not\leftrightarrow$ All children are different). Since (74b) is quantificational, the derivation is more complex. First the *dfit* operator applies to lift the determiner *all* to the type $(ett)(ettt)$. This has the effect of applying the ordinary *all* to the union of $pdist(student)$, on the one hand, and to the union of the intersection of $pdist(student)$ with *different*. In other words, (74b) is true iff every child is a member of some set of different children. This is of course too weak, and therefore the witness condition is necessary. The only witness set in this case is obviously the set of *all* children. So this set has to be in the denotation of *different*. Notice that this already entails what had been achieved *before* the witness condition was applied: since the set of all children is in the extension of *different*, it is clear that every child is in a set of different children. Hence, in fact, the entire meaning is contributed here by the witness condition. There is still nothing necessarily wrong with this; it is just the private case of *all*. However, notice that the meaning obtained for (74b) is equivalent to that of (74a). Both sentences come down to: the set of all children is in the extension of *different*. Does this confer with intuition? I believe there is a very strong intuition according to which the two sentences are *not* equivalent. (74b) entails (74a), but not the other way around. For (74a) to be true, it seems sufficient that there is some pair of different children, for instance. This may be an idiosyncratic feature of *different*.

To sum up then, the witness condition contributes the most crucial part to the analysis of "All rocks are similar"; actually it provides the full analysis for this sentence (the contribution of the pure *dfit* strategy being redundant). However, in case of "All children are different", the witness condition is in a sense redundant: since every child is different from at least one child according to the pure *dfit* approach, it seems clear that the entire set of children is different. However, the *dfit* approach fails to predict the existing distinction between (74a) and (74b).

4.6.2 The missing-witness problem

The first potential problem for the analysis is that the witness condition may be too strong. Although it seems essential in (58), it is far more dubious in (75).

(75) 15 students drank a whole glass of beer together.

(75) can arguably be true in cases where there is no witness set consisting of 15 students who drank a glass of beer together, but rather 2 sets of 7 and 8 students respectively. Consider now the more striking case of (76), (77) and (78), where the latter is evaluated in the context (79).

(76) 15 students drank glasses of beer (together).

(77) Between 10 and 30 students drank (together) glasses of beer.

(78) 3 students wrote plays.

(79)
- John, Bill and Mary are students.
- John and Bill wrote a play together.

– Mary wrote a play.

There is another complication which arises in (77). Suppose there are two sets, of 12 students each, that drank together a whole glass of beer. Each set on its own is a witness set in this case, and even the two sets, combined together, still obey the quantifier 'between 10 and 30', since they total at 24. However, the proponents of uniqueness of the witness set argue that this situation does not verify (77). To make this prediction, the witness condition as stated above is insufficient, and some sort of a uniqueness condition is necessary to strengthen it. There is another difficulty here: uniqueness of witness sets is probably too strong. It may be the case that a strong sense of maximality should be the appropriate requirement: that is, that a witness set should include any potential witness set.

Some other evidence that challenges the witness condition comes from the contrast in (80) between (a) and (b):

(80) a. All students gathered on the streets.

b. All the students gathered on the streets.

I think that, quite strikingly, (80a) is more likely to have a neutral reading than (80b). The same effect is evident in (81).

(81) a. All students drank together glasses of beer.

b. All the students drank together glasses of beer.

This may point to the fact that the witness condition has at least something to do with the definiteness of the subject, which is the minimal difference in the above pair. It is, of course, not surprising that definiteness triggers an existence and uniqueness effect, as well as maximality which has been proposed in the context of definites. The question is whether it is adequate to incorporate the witness condition into the semantics of the determiner.

Finally, another point has to do with predicates like *gather*. Winter points out that it is not a very good candidate for testing the predictions of the *dfit* strategy. This is because predicates like *gather* are *vaguely* downward monotone, which is a lexical property concerning its denotation. Ascribing downward-monotonicity to some predicates (e.g. *similar*) seems rather plausible (though the case of *gather* seems more complex, and Winter places him is a class he terms *vaguely* downward-monotone predicates, among verbs like *meet*, *like-each-other*). However, it still causes a problem to the witness condition. Should the following inference be accounted for on the basis of the *denotation* of *gather* alone, under the assumption that the entire academic staff is composed of the professors plus the teaching assistants?

(82)
$$\frac{\begin{array}{l}\text{The professors gathered on the street}\\ \text{The teaching assistants gathered on the street}\end{array}}{\text{The entire academic staff gathered on the street}}$$

This inference is not based on downward monotonicity, but on another principle the validity of which seems to be much more general. Note that this gives a precise sense to what Winter termed *vague monotonicity.* If the validity of inferences like in (82) is not accounted for by the denotation of the predicate, the conclusion would be that the witness condition fails to hold for the conclusion of (82): if the entire academic staff did not gather *together*, as a single unit, and the gathering split into the professors and the teaching assistants, then we would expect to find only the sub-gatherings in the denotation of *gather*, and then the 'complex' gatherings may be a result of a cumulativity principle.

The next subsection will give some more evidence for the use of cumulativity, concerning the missing-witness problem with respect to transitive verbs.

4.7 Vagueness and cumulative closure

Consider the sentence

(83) 1000 children wrote poems

If the sentence is true then 1000 children were involved in pure sets of children who wrote poems. But if these 1000 children did not write anything *together*, as a set of 1000, then the witness condition is apparently not satisfied. In order to satisfy a witness condition, some assumption must be made about the verb *write*, i.e. that its denotation contains the pair $< C, P >$, where C is the set of 1000 children as above and P is the total set of poems that subsets of this big set wrote.

This reminds of the vagueness approach to sentences like "The men wrote operas", to which the following formula is assigned, which underspecifies the precise relations between the individual men and collections of them and the operas they wrote (where M is the set of men):

(84) $\exists X \subseteq opera(wrote(M, X))$

I propose here to save the missing witness problem by introducing an operation of *cumulative closure*, which will operate on binary relations and will be considered as a generalization of the way it operates on unary relation like *gather* which was discussed in the previous section.

Formally I will use the following property of cumulativity for unary predicates R in (85a) and for binary relations in (85b):

(85) a. Whenever $A, B \in R$, then also $A \cup B \in R$.

b. Whenever $< A_i, B_i > \in R$ for $i = 1, 2$, also $< (A_1 \cup A_2), (B_1 \cup B_2) > \in R$.

The cumulative closure of a binary relation R can be defined as follows:

(86) $R^*(X, Y) \iff \exists R' \subseteq R(X = \cup domR' \wedge Y = \cup rgR')$

Now we show how this assumption can help us account for the contrast between (83) and (87):

(87) 1000 children wrote a poem.

And here the neutral reading does not occur. So the assumption that the witness problem with the previous sentence is solved by making an assumption about the nature of the verb *write* is problematic. Instead we can argue that (87) is not verified in a situation which would verify a neutral reading because the witness condition fails: there is no set of 1000 children which wrote *a* poem.

To sum up, the assumption about cumulative closure combined with the witness condition accounts for the contrast between (87) and (83): in (87) a witness is missing in case of a "neutral situation" (that is, a situation that would verify the sentence on its neutral reading, had it existed). The witness is missing "thanks" to the operation of cumulative closure.[17]

5 Collective cumulation as iterative collective quantification

This section begins with an extension of Winter's *dfit* strategy to the transitive case. The first purpose of this extension is to deal with *iterative* readings. Then the section proceeds to examine the interaction of the iterative *dfit* procedure and relations which are closed under the cumulativity operator introduced in the previous section. This operator is shown to solve an over-generation problem associated with the witness condition, in addition to the under-generation problem discussed in the previous section. The main idea is that this interaction results in cumulative readings under certain conditions to be discussed in detail. The main proposal of deriving cumulative readings by iterative means is first sketched informally, followed by a discussion of its scope and generality. Some points of departure from Winter's approach are mentioned and then the formal proofs are given. Finally, Keenan's result is extended to the collective case followed by some comments of how this extended result is settled with the findings of this paper.

5.1 Winter's analysis extended to transitive verbs

In this subsection I illustrate a natural way of extending Winter's *dfit* strategy to deal with transitive sentences. This extension is along the lines of type $\langle 1 \rangle$ functions introduced in section 2. Consider for example the following sentence:

(88) (Exactly) 6 women wrote (exactly) 2 operas.

Under the *iterative* reading, the sentence reads as follows: the total number of women who wrote a total number of exactly 2 operas is exactly 6. How could this reading be derived in the *dfit* approach? First, I assume that the predicate *write* is of the type $set \times atom$. It makes sense to speak of sets of people who wrote things *together*. It may be more appropriate to consider as basic the predicate *write-together*. In any case, I assume this to be the basic

[17] Notice that (87) can be interpreted in two ways: either collectively, on which reading 1000 children wrote a poem together; or distributively, in which case 1000 children each wrote some poem. The neutral option is disallowed. We shall see in the sequel that what "replaces" the illicit neutral reading is the cumulative reading. In case of (87), it is equivalent to the collective reading.

denotation of *write*. Since we are interested here in the distributive reading of (88), we look at $sg^1(write)$ (or $sg^1(write - together)$), and then distribute the complex predicate *write-2-operas* by the *pdist* operator.[18] The following is the expression we get according to the iterative procedure:

(89) $dfit_w(6)(women)(pdist(\{A : dfit_w(2)(operas)(sg(write_A))\}))$

To illustrate this, suppose that the set of women who wrote 3 operas each is as follows:

(90) $\{A : dfit_w(3)(operas)(sg(write_A))\} = \{w_1, w_2, w_3\}$

Then the distributed predicate is simply the power set (or the *positive* power set under another approach that we do not adopt here):

(91) $pdist(\{A : dfit_w(3)(operas)(sg(write_A))\}) =$
$\{\emptyset, \{w_1\}, \{w_2\}, \{w_3\}, \{w_1, w_2\}, \{w_1, w_3\}, \{w_2, w_3\}, \{w_1, w_2, w_3\}\}$

In case of "3 women wrote 3 operas", it is indeed the case that the total number of women who were involved in sets of women who wrote 3 operas is 3. Furthermore, there is a maximal witness $\{w_1, w_2, w_3\}$ which consists of 3 women who wrote 3 operas.

5.2 The scope problem for neutral NPs- revisited

Consider (88) again. Let us keep the predicate *write* in its basic form; that is, with the collective meaning of *write-together*. As shown in section 4, iteration of the neutral reading over a distributive reading results in a reading which we are trying to block. More concretely, in (88) the number of operas could exceed the normal upper bound of $6 \times 2 = 12$, and reach $(2^6 - 1) \times 2 = 126$. Notice that *write* is assumed to be an atom predicate on its *second* argument, hence applying *dfit* should result in a reading parallel to the $N \times D$ reading (because *sg* has been applied, the embedded lifting behaves just like ordinary quantification, which boils down to distributive quantification). Suppose for the moment that cumulative closure is not applied to the denotation of *write*. Then under the assumption that the pure *dfit* component (that is, the "neutral" component is satisfied), sentence (88) is deemed true depending on whether there exists a witness set, which in this case should be a set of 6 women who wrote 3 operas. But whether or not such a set exists does not settle the question of whether the sentence is true or false. What is relevant is whether the *entire* set of operas written by the witness set is 2. That is, the count should also go through all the subsets of the witness, (and in particular through all its occurrences, although this is already taken care of by the *dfit* strategy itself).

For this reason, it seems that the cumulative closure is essential to get a reasonable analysis of the sentence. On the assumption that the operation of cumulative closure has been applied to the denotation of *write*, the witness condition requires that there should be

[18]Notice that alternatively, it is possible to refrain from applying *pdist* and apply ordinary quantification. However, I prefer to use *dfit* here in order to contrast iterativity with cumulativity in the sequel.

a set W of 6 women such that the number of operas written by members of W must total at 2.[19]

Thus, the scope problem for N seems to be circumvented on the assumption that cumulative closure must be applied. Notice what kind of reading actually obtains. The total number of women who were involved in sets of women who wrote exactly 2 operas is 6. The total number of operas written by these women is 2. As it is paraphrased here this reading seems rather unlikely. The next subsection looks more closely at what the cumulative closure produces.

5.3 Cumulativity-closed relations and the double *dfit* approach

Suppose that each w_i where i is some index denotes a woman, and each o_i denotes an opera. Let R be the denotation of *write*. Suppose the restriction of the relation R to $\mathcal{P}(woman) \times \mathcal{P}(opera)$ contains the following pairs:

(92) $< \{w_1, w_2\}, o_1 >, < \{w_3\}, o_2 >, < \{w_4\}, o_3 >$

Now apply cumulative closure. We will take (92) as the basic denotation, and make each second projection into a singleton containing the original object. The following pairs are added by the closure:

(93) a. $< \{w_1, w_2, w_3\}, \{o_1, o_2\} >$
b. $< \{w_1, w_2, w_4\}, \{o_1, o_3\} >$
c. $< \{w_3, w_4\}, \{o_2, o_3\} >$
d. $< \{w_1, w_2, w_3, w_4\}, \{o_1, o_2, o_3\} >$

Notice for example that (94) should be verified if R contains the pairs in (92), and formally it can only be so if (93a) is present.

(94) w_1, w_2 and w_3 wrote 2 operas.

Now examine (93) with respect to the sets of women who wrote 2 operas. There are three pairs where a set of exactly 2 operas appear in the second projection: (93a), (93b) and (93c). From these pairs it follows that all the women who were involved in writing operas by women were also involved in writing a set of 2 operas: w_1, w_2, w_3 and w_4. Is this a just a matter of accident or is there any generalization one can make for such cases?

Consider now the more general pattern, involving arbitrary bare numerals:

(95) m women wrote k operas

[19] Consider the sentence "Exactly 5 children drinked together a glass of beer", on its collective reading. The witness condition does not exclude there being subsets of the witness which also drinked together a glass of beer. Thus the total number of glasses of beer drinked by combinations of members of the witness set may rise up to $2^5 - 1 = 31$. However, if "a glass of beer" is replaced by "2 glasses of beer", for instance, then such a situation should not verify the sentence. I have an explanation for the contrast; perhaps it has to do with the fact that singulars have a more salient existential reading than that of plurals; perhaps it has to do with the idea of "strong functionality" and the "localistic" view to be discussed below.

Suppose also that $k > 0$. Assume that (95) is true for R^*, where R is the denotation of *write*. Then the witness condition must hold, and hence there is some set W consisting of m women, and this set wrote a total number of k operas. Denote this set of k operas by O. Suppose w is a woman who was involved in writing operas by women. Then we know that in R there is a pair of the form $< W', o >$, where o is an opera and W' is a set of women such that $w \in W'$. Therefore the pair $< W', \{o\} >$ occurs in R^*. By assumption $k \geq 1$ and hence O is not empty. We would like to conclude that $w \in W$. In order to do that we want to construct another set O' of k operas which was written by a set of women which includes w.

One possibility is that $o \in O$. In such a case, we could consider the set $W \cup W'$. If we knew that this set wrote k operas, we would be done with this case. What we know is that the set $W \cup W'$ wrote *at least* k operas. It may be the case, as things stand at the moment, that adding the set W' results in adding further operas in addition to o, which is already included, and thus the number of operas would be actually bigger than we want. For instance, $< W', o >$ may not be the only pair in R with W' as its first projection. Another possibility is that R contains pairs with first projections which consist of elements from both W and W'. We shall make the following assumption with respect to R: R should not contain two pairs with non disjoint first projections, unless the second projections of the two pairs are identical. We call this property *strong functionality.* This may appear to be an unintuitive restriction on R. Later we will discuss why such a restriction is sensible.

Under such a constraint on R, the current case is concluded. Let us proceed to the possibility that $o \notin O$. In this case the idea would be to switch some element in O with o, and thus obtain a set of k operas which includes o. The special property assumed about R guarantees also in this case that we can make a similar switching also with respect to the sets of writers, and thus just switch pairs from the basic denotation of R to get O'.

This way, we can prove that W contains *all* the woman who were involved in writing operas by women. Next, it easily follows by the *dfit* strategy that k has to be the total number of operas written by sets of women. Thus we have derived the cumulative reading of (95). The next subsection discusses the generality of this proposal for other kinds of determiners.

5.4 The generality of the proposal

5.4.1 Maximality of the witness

Does the current proposal for deriving cumulative readings work for determiners other than numerals? Below I show that what is proposed here actually relies on some property that bare numerals have, but that determiners in general do not necessarily have. However we shall see that it is the maximality of the witness which takes care of this potential problem. Consider the following example, which in fact counter-exemplifies the current proposal:

Suppose 12 students drank (together) 19 glasses of beer, in the cumulative sense. That is, the total number of students who drank beer are 12 and the total number of glasses of beer drinked by students is 19. Consider the following sentence:

(96) At least 8 students drank 14 glasses of beer.

Suppose further that there is a set of 8 students who drank together a total number of 14 glasses of beer. That is, the cumulative closure of *drink* contains a pair with a set of 8 students as its first projection and a set of 14 glasses of beer as its second projection.

Now according to what is presently proposed, (96) seems to be verified in the given circumstances, which is an undesired outcome. To see why, notice the following. First, the total number of students who took part in a set of students which drank 14 glasses of beer is in fact 12, the total number of students participating in sets of students who drank beer. So the determiner 'at least 8' is appropriate. Second, most importantly, the set of 8 students who drank 14 glasses of beer and whose existence has been stipulated above qualifies as a witness set. This is so since it is a set of at least 8 students which drank a total number of 14 glasses of beers.

In other words, since there is a witness which is "too small", maximality of the second argument is not attained. This defect cannot arise with respect to bare numerals, since a witness of k NPs cannot be too small, otherwise it would contain less than k elements. So in order to generalize the proposal which has been shown appropriate for numerals, the following missing link is required: make sure that the witness set is *maximal*, in the *total* sense indicated in section 4.5. A few alternatives were raised there; but for this purpose the total witness condition seems necessary. Notice that bare numerals always pick the total witness because of cardinality reasons, independently of any maximality requirement.

5.4.2 The empty-set problem and cumulativity

In case of determiners like 'at most 3' occurring in the object position, or more generally *non-positive* determiners, which occur in the object position, what I term 'the empty-set problem' arises: consider a set X of women, some of whom indeed wrote some operas, but suppose one of them did not take part in any set of women who wrote operas. How can it be ensured that the set X is not taken into account? Suppose there are 47 women in total, and that 24 operas were written by women, while the actual number of women taking part in writing operas is only 13. Then if the empty-set problem is not circumvented, the following undesired sentence would be expected to be true:

(97) 47 women wrote at most 13 operas

Notice that there seems to be an equivalent for this, in the given model, which seems much more plausible:

(98) All the women wrote at most 13 operas

In (98) we are interested in the number of operas that were written by women. It is the first conjunct of the cumulative reading, but only the first one, so it is not the cumulative reading which is at stake here. Besides, it is not only (97) which would be considered true in the absence of a solution to the empty-set problem. It would also be the case if 47 is replaced by, say, 41.

For these reasons, we will not allow the empty set to occur in the denotation in pairs other than $< \emptyset, \emptyset >$. It may also be worthwhile to make a distinction between a *basic* denotation of a predicate and a *derived* denotation which is what the cumulative closure produces. Thus we may claim that occurrences of the empty set as an argument is a result of cumulative closure, and that no such occurrences are allowed in the basic denotation. In any case, our assumption is that no pairs other than $< \emptyset, \emptyset >$ contain the empty set as one of their projections.

This however does not seem to circumvent the problem with respect to downward-monotone determiners in the position of the object. Nevertheless, a closer look on the implementation of the $dfit$ strategy in this case would trace back the source of the caveat. The witness condition, as it appears in (59), is not a global condition for sentences, but is rather attached to any application of a $dfit$-tted determiner. Hence, it also applies to the second determiner. This would require that there should also be a witness set of operas in the intersection of $operas$ and $write - 5 - operas_A$ (A here is arbitrary, which indicates that the witness condition for the second argument applies iteratively). A witness for the second argument is required independently of cumulative readings, in cases like the following:

(99) Exactly 2 bartenders sold a glass of beer to exactly 5 children, who gulped it together at once.

To achieve the same effect for 'exactly 5 children' as in (100) (from Winter (1998)), under the assumption that $sold - a - glass - of - beer - to$ is a set predicate on its *second* argument, a witness condition must be introduced for the second argument as well.

(100) Exactly 5 children drinked together a whole glass of beer

As another example, (101) means that the mayor bought one house, and that house was bought for a set of 6 people. It is not true if the mayor bought one house for 4 homeless people and then another house for 2 other homeless people.

(101) The mayor bought one house for exactly 6 homeless people

The witness condition for the second argument should be attached to the iterated component, and is in fact completely parallel with the original one. Going back to the women and operas situation, the iterated part is:

(102) $\{A : 5(\cup operas)(\cup(write_A \cap operas) \wedge \exists W' \in write_A \cap operas(5(\cup operas)(W')))\}$

I propose that the empty-set problem can be solved by the witness condition on the second argument. But for that purpose (and for other purposes as well), I propose to simplify the witness condition Winter suggests by removing the exception for empty sets. This proposal has some advantages, but it is also responsible for some unattractive consequences, as we shall see.

5.4.3 Strong functionality

The requirement that R should be strongly injective seems under-motivated. However, consider the following example (due to Van der Does, personal correspondence):

(103) a. A outnumbers B

b. A outnumbers C

c. $A \cup A$ outnumbers $B \cup C$

d. A outnumbers $B \cup C$

By applying cumulative closure on (103a) and (103b), one gets (103c) which is equivalent to (103d). However, this is not a valid inference. Suppose the set A contains 5 elements, the set B contains 3 elements and the set C contains 4 elements. Suppose further that B and C are disjoint. Then it is true that A outnumbers B and also that B outnumbers C, but $B \cup C$ contain 7 elements, whence B and C together outnumber A, and not vice-versa. However, the following pattern of inference is valid on the assumption that A and A' are disjoint:

(104)
$$\frac{\begin{array}{l} A \text{ outnumbers } B \\ A' \text{ outnumbers } C \end{array}}{A \cup A' \text{ outnumbers } B \cup C}$$

The failure of the inference in (103) results from the fact that it is an instance of (104) where A and A' are not disjoint. What the conclusion should say is that by taking the elements of A twice (as a multi-set), they exceed the elements of $B \cup C$ in number. In other words, A should be treated as an independent object in each occurrence, and it is not identified in terms of the elements of which it consists. We may claim that the occurrence of A as a set in (103) is not extensional.[20]

We now try to apply this line of thought to the case of cumulative readings. Consider the sentence "4 children lifted 6 pianos". I argue that the occurrence of a set A in $lift^*_A$, as it appears in the double $dfit$ iterative procedure, is not extensional. That is, A should not be identified just in terms of the elements it consist of; whence the requirement of strong functionality. On the other hand, when the number of children who were involved in lifting pianos is at stake, we identify identical element in the counting and deal with atoms again.

5.4.4 Clashes with Winter's approach

It is important to point out that the current proposal of how to implement the $dfit$ strategy to derive cumulative readings is not completely in accordance with Winter's $dfit$ approach. For instance, the $dfit$ does not apply to singular determiners. However, we would not like to distinguish cases like the following from the rest, which clearly has a cumulative reading:

(105) Exactly 1 boy met exactly 1 girl.

[20]Note, however, that this should not suggest that strong functionality accounts for all the cases in which cumulative closure fails to apply.

Another deviation from Winter's approach is the application of cumulative closure, which is a polyadic operator. Notice that the operation of cumulative closure is very similar to *pdist*, or simply to the power set operation. By identifying each element of R^* as a subset of R, according to the definition given in (86), we can view R^* as the power set of R.

Winter argues in general against polyadic distributivity (Winter (1997)). The main argument there is based on a sentence like the following:

(106) Jim, Albert and Jerry gave a flower to Ann and Sue.

This sentence is not verified by a situation in which Jim gave a flower to Ann, and Albert and Jerry gave a flower to Sue. However, this example is very similar to (87) which was discussed before. Actually it is due to cumulative closure that (106) is blocked: again because a witness is missing.

5.5 The formal proofs

We introduce a few abbreviations in order to simplify notation:

(107) $dom := \cup(dom(R \cap \mathcal{P}(N_1) \times \mathcal{P}(N_2))$

(108) $rg := \cup(rg(R \cap \mathcal{P}(N_1) \times \mathcal{P}(N_2))$

(109) $\mathcal{B} = \{A : D'(N_2)(\cup(R^*_A \cap \mathcal{P}(N_2)) \wedge witc(D')(\mathcal{P}(N_2))(R^*_A)\}$

The expression $witc(D')(\mathcal{P}(N_2))(R^*_A)$ spells out as follows:

(110) $\exists U \in \mathcal{P}(N_2) \cap R^*_A \wedge D'(N_2)(U) \wedge \forall U' \in \mathcal{P}(N_2) \cap R^*_A(U \subseteq U' \rightarrow U' = U)$

We begin by comparing two strategies to obtain cumulative readings: one which does not constrain the verb meaning and that is based on two monadic quantifications; and one which does constrain the verb meaning by a cumulative closure and which is based on (double) modification and maximization.

The first one is from Van der Does and as given in (12). It is repeated here as (111):

(111) $D(N_1)(dom) \wedge D'(N_2)(rg)$

The second one is the following, where * is the cumulative closure defined in (86).

(112) $\exists X \exists Y[X \subseteq N_1 \wedge Y \subseteq N_2 \wedge R^*(X,Y) \wedge D(N_1)(X) \wedge D'(N_2)(Y) \wedge$
$\forall Z \forall U((Z \subseteq N_1 \wedge U \subseteq N_2 \wedge R^*(Z,U)) \rightarrow (Z \subseteq X \wedge U \subseteq Y))]$

(112) can be reformulated in a more transparent manner, as follows:

(113) $\exists X \in \mathcal{P}(N_1) \cap D(N_1) \exists Y \in \mathcal{P}(N_2) \cap D(N_2)[R^*(X,Y) \wedge R^* \cap \mathcal{P}(N_1) \times \mathcal{P}(N_2) \subseteq \mathcal{P}(X) \times \mathcal{P}(Y)]$

In the literature it has been claimed that the double collective reading amounts to the same thing as the cumulative reading (see Van der Does (1998)). This is of course not the case, in any formal sense. However, under certain circumstances, such a reduction of cumulativity to double collectivity is possible, as the following result shows.

Proposition 3 *For a 2-place predicate R, and for NP denotations N_1 and N_2, van der Does' definition (111) and the double-modificational +maximization (112) are equivalent.*

PROOF: It is useful to notice that $\cup(R^* \cap \mathcal{P}(N_1) \times \mathcal{P}(N_2))$ is identical to the same expression for non-starred R. First assume (113). ¿From $R^* \cap \mathcal{P}(N_1) \times \mathcal{P}(N_2) \subseteq \mathcal{P}(X) \times \mathcal{P}(Y)$ it follows that $dom \times rg \subseteq X \times Y$. For the other inclusion, notice that $R^*(X,Y)$ means by definition that there is $R' \subseteq R$ with $X = \cup dom(R)'$ and $Y = \cup rg(R')$. From this it follows that $X \times Y \subseteq dom \times rg$.

Conversely, assume (111). Take X and Y as $\cup dom(R^* \cap \mathcal{P}(N_1) \times \mathcal{P}(N_2))$ and $\cup rg(R^* \cap \mathcal{P}(N_1) \times \mathcal{P}(N_2))$ respectively (or in the light of the initial remark, simply *dom* and *rg* respectively). Then clearly $X \subseteq N_1$, $Y \subseteq N_2$, $D(N_1)(X)$ and $D'(N_2)(Y)$. That $R^*(X,Y)$ holds is also clear from the definition of R^* in (86) by simply taking $R' = R$ (and notice the remark in the beginning of the proof).

Finally it is obvious that the first projection of $R^* \cap \mathcal{P}(N_1) \times \mathcal{P}(N_2)$ contains only elements of X, and that the second projection only contains elements of Y. Hence, (113) holds. $\square$

Next we prove the main result. The double *dfit* strategy is shown to be equivalent to Van der Does' definition of collective cumulativity, under certain restrictions on R and on the object determiner D'. In addition, equivalence is shown to a third method, which consists in an iteration of the *dfit* strategy over a modificational reading of the object. The modificational reading of the object allows us to implement a localistic view which does not rely on the strong functionality of R. However, in order to insure maximality with respect to the object, a use of strong functionality is indispensable.

We discuss some of the restrictions. We assume that the determiner associated with the object position, D' has several properties. First, we assume that when it is combined with the object restrictor, namely N_2, it is not equivalent to the generalised quantifier "no one" or "nothing", that is $\{\emptyset\}$. Otherwise, for example, the sentence "At most 3 children lifted 0 tables" is verified in a situation where one child lifted one table. The reason is that the empty set is a maximal witness for "At most 3 children lifted 0 tables".

Further, we assume that if D' is not positive, then it is weekly-monotone-decreasing. The last requirement holds for instance for all *continuous* determiners. In fact the crucial requirement is that D' should be continuous at 0, that is 0 should not be a point of discontinuity. More concretely, we want to avoid a case in which $D'(N_2)$ holds of the empty set as well as of a subset of N_2 of some cardinality greater than 1, while it does not hold of a subset of N_2 of cardinality 1.[21]

We would not like to give up on this somewhat complex condition on D' and replace it by requiring that D' be positive, since otherwise it would seem that the proposition does

[21] These problems all arise from the possibility that the empty set could be a witness (for the subject argument). Perhaps another way out of these difficulties is not to let the empty set to be a witness under certain circumstances.

not hold if D' downward monotone, for instance. There are many downward monotone determiners for which the proposition holds, such as "at most 5".

The following is the double *dfit* strategy, for a binary relation R, noun phrases N_1 and N_2, and determiners D and D':

(114) $dfit_w(D)(N_1s)(\{A : dfit_w(D')(N_2s)(R^*_A)\})$

This is equivalent to:

(115) $D(\cup N_1s)(\cup(\{A : D'(\cup N_2s)(\cup(R^*_A \cap N_2))\} \cap N_1s))$

We only deal with the case of distributive restrictors, so we may assume that N_1s and N_2s are just the power sets of N_1 and N_2. We therefore get the following equivalent formula:

(116) $D(N_1)(\cup(\mathcal{B} \cap \mathcal{P}(N_1))) \wedge\ witc(D)(\mathcal{P}(N_1))(\mathcal{B})$

The "hybrid" iteration of *dfit* over a modificational analysis of the object is as follows:

(117) $dfit_w(\mathcal{P}(N_1))(\{A : \exists X \subseteq N_2 : D'(N_2)(X) \wedge R^*(A, X)\})$

It is further assumed that $< \emptyset, \emptyset >$ is the only pair which contains $\emptyset$ as one of its projections. Notice that from the definition of the cumulative closure in (86) it follows that $< \emptyset, \emptyset > \in R^*$ (by just taking $R' = \emptyset$).

The next proposition will show equivalence, under cerain circumstances, between Van der Does' definition, the double-*dfit* approach and the hybrid approach. The proof will be brought in two parts. The first will show equivalence of Van der Does' definition and the double-*dfit* strategy, and the second will show equivalence between Van der Does' definition and the hybrid approach. It is interesting to point out that the "strong functionality" restriction imposed on R play different roles in each part of the proof. The modificational analysis of the object in the case of the hybrid approach incorporates a localistic view, which avoid the need to impose a further restriction on R; however, at another part of the proof, the same restriction becomes necessary. This reflects the double status of sets with respect to extensionality, as will become clearer from the proof itself.

Proposition 4 *Assume R is a binary relation of type set$\times$atom, D and D' are determiners, N_1 and N_2 are NP denotations. We make the following further assumptions about R and D':*

- *R has the following property we call "strong functionality":*

$$\forall A \forall B \forall a \forall b (R(A, a) \wedge R(B, b) \wedge A \cap B \neq \emptyset \rightarrow a = b)$$

- *$D'(N_2) \cap \mathcal{P}(N_2) \neq \{\emptyset\}$, i.e. $D'(N_2)$ is not the quantifier "no N_2".*

- *Either D' is positive, or D' is weakly-monotone-decreasing (or it is possible to require that D' should be continuous instead of the second disjunct).*

Under these assumptions, the following are equivalent: (i) Van der Does' definition of collective cumulative quantification (111); (ii) the double $dfit_w$ iterative strategy (116); (iii) the iteration of $dfit$ over a modificational reading of the object (117).

We first show: $(i) \iff (ii)$.

PROOF: First assumee (ii), the double dfit. It suffices to prove that $dom = \cup(\mathcal{B} \cap \mathcal{P}(N_1))$. For from $\cup(\mathcal{B} \cap \mathcal{P}(N_1)) \in \mathcal{B}$ it follows that $D'(N_2)(\cup(R^*_{dom} \cap \mathcal{P}(N_2)))$, that is $D'(N_2)(rg)$. Combined with $D(N_1)(dom)$, collective cumulation (i) is immediate.

Claim: $dom = \cup(\mathcal{B} \cap \mathcal{P}(N_1))$

Proof of the claim: [$\supseteq$]: We show $\mathcal{B} \cap \mathcal{P}(N_1) \subseteq \mathcal{P}(dom)$. Assume $A \in \mathcal{B} \cap \mathcal{P}(N_1)$. Suppose for a contradiction that $A \not\subseteq dom$. Since $A \subseteq N_1$, and $A \neq \emptyset$, there is some $a \in A$ such that $\forall A'((a \in A' \wedge R(A, b)) \Rightarrow b \notin N_2)$. If $R^*(A, B)$ then for some $R' \subseteq R$, $A = domR'$ and $B = rgR'$. Since $a \in A$, there must be some $b \notin N_2$ and some $A' \subseteq A$ with $R(A', b)$. Hence $B = rgR' \not\subseteq N_2$. Hence: $R^*_A \cap \mathcal{P}(N_2) = \emptyset$. Since $A \in \mathcal{B}$, it must hold that $R^*_A \cap \mathcal{P}(N_2) \in \cup(R^*_A \cap \mathcal{P}(N_2))$, hence it cannot be the case that $R^*_A \cap \mathcal{P}(N_2) = \emptyset$, hence we have arrived at a contradiction.

[$\subseteq$]: Assume $d \in dom$. Then there is some $Z \subseteq N_1$ with $d \in Z$, and some $y \in N_2$ with $R(Z, y)$. We need to find some $W \in \mathcal{B} \cap \mathcal{P}(N_1)$ with $d \in W$. Denote $\cup(\mathcal{B} \cap \mathcal{P}(N_1))$ by A. We have $D(N_1)(A)$ and $A \in \mathcal{B} \cap \mathcal{P}(N_1)$. Suppose for a contradiction that $A = \emptyset$. Then $\cup(R^*_A \cap \mathcal{P}(N_2)) = \emptyset$ (by assumption on R). Since $A \in \mathcal{B}$, $D'(N_2)(\cup(R^*_A \cap \mathcal{P}(N_2)))$, so $D'(N_2)$ is not positive. Hence $D'(N_2)$ is weakly monotone-decreasing. This implies $D'(N_2) \cap \mathcal{P}(N_2) = \{\emptyset\}$, contrary to the second assumption on D'. To see this, note $R^*_Z \cap \mathcal{P}(N_2) = \{y\}$ by strong functionality. By assumption on R, $Z \neq \emptyset$. This implies $Z \notin \mathcal{B}$ (for $\cup(\mathcal{B} \cap \mathcal{P}(N_2)) = \emptyset$). Therefore $\neg D'(N_2)(\cup(R^*_Z \cap \mathcal{P}(N_2)))$ (the negation of the other conjunct $\cup(R^*_Z \cap \mathcal{P}(N_2)) \notin R^*_Z \cap \mathcal{P}(N_2)$ is impossible, because R^* is cumulatively closed). So $\neg D'(N_2)(\{y\})$. By the permutation property of D' it follows that $\forall M \subseteq N_2[M \neq \emptyset \Rightarrow \neg D'(N_2)(M)]$. The reasoning is as follows. Suppose for some non-empty $M \subseteq N_2$, $D'(N_2)(M)$. If $y \in M$, it follows by weak downward-monotonicity that $D'(N_2)(\{y\})$, which is a contradiction. If $y \notin M$ we can choose an element of M and permute it with y. Such a permutation is N_2-invariant. We could then use weak downward monotonicity to get a similar contradiction again. So $D'(N_2) \cap \mathcal{P}(N_2) = \{\emptyset\}$, a contradiction. We conclude that $A \neq \emptyset$.

Since $A \in B$: $D'(N_2)(\cup(R^*_A \cap \mathcal{P}(N_2)))$ and $\cup(R^*_A \cap \mathcal{P}(N_2)) \in R^*_A$. $A \neq \emptyset$, so $\cup(R^*_A \cap \mathcal{P}(N_2)) \neq \emptyset$ (by assumption on R). Choose $p \in \cup(R^*_A \cap \mathcal{P}(N_2))$. Define the following mapping with functionality $E \longrightarrow E$

$$\pi(x) = \begin{cases} y & \text{if } x = p \\ p & \text{if } x = y \\ x & \text{otherwise} \end{cases}$$

Then $\pi(N_2) = N_2$ and $D'(N_2)(\pi(\cup(R^*_A \cap \mathcal{P}(N_2))))$. For each $d \in \cup(R^*_A \cap \mathcal{P}(N_2)) - \{p\}$ choose $W_d \subseteq N_1$ with $R(W_d, d)$. By strong functionality W_d only contributes d. Define $W' := \cup_{d \in \cup(R^*_A \cap \mathcal{P}(N_2)) - \{p\}} W_d \cup Z$. Then: $Z \subseteq W' \subseteq N_1$, and $\pi(\cup(R^*_A \cap \mathcal{P}(N_2))) = \cup(R^*_{W'} \cap \mathcal{P}(N_2)) \in R^*_W$. Since $D'(N_2)(\cup(R^*_{W'} \cap \mathcal{P}(N_2)))$, $W' \in \mathcal{B}$. And from $d \in Z$, also $d \in W'$, and we are done.

Conversely, assume (i), collective cumulation. We show: $dom \subseteq \cup(\mathcal{B} \cap \mathcal{P}(N_1))$. Since $rg = \cup(R^*_{dom} \cap \mathcal{P}(N_2))$, $D'(N_2)(\cup(R^*_{dom} \cap \mathcal{P}(N_2)))$. We also have $\cup(R^*_{dom} \cap \mathcal{P}(N_2)) \in R^*_{dom} \cap \mathcal{P}(N_2)$ because R^* is cumulatively closed. Hence $dom \in \mathcal{B}$. Clearly $dom \subseteq \mathcal{P}(N_1)$, hence $dom \subseteq \cup(\mathcal{B} \cap \mathcal{P}(N_1))$. ¿From the claim above we have $\cup(\mathcal{B} \cap \mathcal{P}(N_1)) \subseteq dom$ (note the claim is independent of the collective cumulation assumption) and so we conclude $dom = \cup(\mathcal{B} \cap \mathcal{P}(N_1))$. $D(N_1)(dom)$ gives us the first conjunct of the $dfit$ iteration. The total witness condition is $\cup(\mathcal{B} \cap \mathcal{P}(N_1)) \in \mathcal{B} \cap \mathcal{P}(N_1)$, and it clearly holds because we have shown that $dom \in \mathcal{B}$. $\square$

We now show $(ii) \iff (iii)$.

PROOF: First assume (ii). It suffices to prove that $dom = \cup(\mathcal{B}_c \cap \mathcal{P}(N_1))$. For from $\cup(\mathcal{B}_c \cap \mathcal{P}(N_1)) \in \mathcal{B}_c$ it follows that $\exists X \subseteq N_2[D'(N_2)(X) \wedge R^*(dom, X)]$. By strong functionality of R we get $(X \subseteq N_2 \wedge R^*(dom, X)) \Rightarrow X = rg$. The reason: from the definition of cumulative closure: there is some $R' \subseteq R$ with $dom = \cup domR'$ and $rg = \cup rgR'$. Let $\mathcal{I}$ be a set of indices s.t. $R' = \{< A_i, B_i >: i \in \mathcal{I}\}$. We claim that $b \in rg \Rightarrow b \in \cup rgR'$: $b \in rg$ implies that for some $A \subseteq dom$, $R(A, b)$. Since $\cup_{i \in \mathcal{I}} A_i = dom$, it follows that for some $i \in \mathcal{I}$, $A_i \cap A \neq \emptyset$. Hence by strong functionality: $b_i = b$. Hence $b \in \cup rgR'$. On the other hand, if $b \notin rg$, then it cannot be the case that $R(A_i, b)$ for any $i \in \mathcal{I}$, since $A_i \subseteq dom$ and $\cup rgR' \subseteq N_2$. Hence $X = rg$ and $D'(N_2)(rg)$. By substituting $\cup(\mathcal{B}_c \cap \mathcal{P}(N_1))$ dom for we get $D(N_1)(dom)$.

Claim: $dom = \cup(\mathcal{B}_c \cap \mathcal{P}(N_1))$: $[\supseteq]$: We show $\mathcal{B}_c \cap \mathcal{P}(N_1) \subseteq \mathcal{P}(dom)$. Assume $A \in \mathcal{B}_c \cap \mathcal{P}(N_1)$. Suppose for a contradiction that $A \not\subseteq dom$. Since $\emptyset \neq A \subseteq N_1$, $R^*(A, B)$ implies that for some $R' \subseteq R$, $A = \cup domR'$ and $B = \cup rgR'$. Since $A \neq \emptyset$, and $R^*(A, X)$ holds for some X (from the fact that $A \in \mathcal{B}_c$), for some $a \in A$, $a \notin dom$, hence for some $A' \subseteq A$ with $a \in A'$ it holds that $R(A', b)$ with $b \notin N_2$. Hence $B = \cup rgR' \not\subseteq N_2$. Hence $\neg\exists X \subseteq N_2[R^*(A, X)]$, which contradicts the assumption that $A \in \mathcal{B}_c$.

$[\subseteq]$: Assume $d \in dom$. Then there is some $Z \subseteq N_1$ with $d \in Z$, and some $y \in N_2$ with $R(Z, y)$. We need to find some $W \in \mathcal{B}_c \cap \mathcal{P}(N_1)$ with $d \in W$. Denote $\cup(\mathcal{B}_c \cap \mathcal{P}(N_1))$ by A. We have $D(N_1)(A)$ and $A \in \mathcal{B}_c \cap \mathcal{P}(N_1)$. Suppose for a contradiction that $A = \emptyset$. Then $R^*(\emptyset, X) \Rightarrow X = \emptyset$, and hence $D'(N_2)(\emptyset)$; whence $D'(N_2)$ is not positive. Hence $D'(N_2)$ is weekly-monotone-increasing. This implies $D'(N_2) \cap \mathcal{P}(N_2) = \{\emptyset\}$, contrary to the second assumption on D'. To see this, note that $R^*(Z, \{y\})$. By assumption on R, $Z \neq \emptyset$. This implies $Z \notin \mathcal{B}_c$, for $A = \cup(\mathcal{B}_c \cap \mathcal{P}(N_1)) = \emptyset$. Therefore $\neg D'(N_2)(\{y\})$. By the permutation property it follows (like in the previous proof) that $\forall M \subseteq N_2[M \neq \emptyset \Rightarrow \neg D'(N_2)(M)]$. We conclude that $A \neq \emptyset$.

Since $A \in \mathcal{B}_c$: $\exists X \subseteq N_2[D'(N_2)(X) \wedge R^*(A, X)]$. Since $A \neq \emptyset$, by assumption on R $X \neq \emptyset$. Choose $p \in X$. Define the following mapping as in the previous proof:

$$\pi(x) = \begin{cases} y & \text{if } x = p \\ p & \text{if } x = y \\ x & \text{otherwise} \end{cases}$$

Then $\pi(N_2) = N_2$ and $D'(N_2)(\pi(X))$. For each $d \in X - \{p\}$ choose $W_d \subseteq N_1$ with $R(W_d, d)$. Then: $R^*((\cup_{d \in X - \{p\}} W_d) \cup Z, \pi(X))$. We conclude that $(\cup_{d \in X - \{p\}} W_d) \cup Z \in \mathcal{B}_c$ and hence $d \in Z \subseteq \cup(\mathcal{B}_c \cap \mathcal{P}(N_1))$.

Conversely, assume (i), collective cumulation. We already know from the claim above that $\cup(\mathcal{B}_c \cap \mathcal{P}(N_1)) \subseteq dom$. We show now: $dom \subseteq \cup(\mathcal{B}_c \cap \mathcal{P}(N_1)) \subseteq dom$. Since $R^*(dom, rg)$ and $D'(N_2)(rg)$ we conclude that $dom \in \mathcal{B}_c$. Hence we have $\cup(\mathcal{B}_c \cap \mathcal{P}(N_1)) = dom$. By substituting dom for $\cup(\mathcal{B}_c \cap \mathcal{P}(N_1))$ we get: $D(N_1)(\cup(\mathcal{B}_c \cap \mathcal{P}(N_1)))$. Moreover, $\cup(\mathcal{B}_c \cap \mathcal{P}(N_1)) = dom \in \mathcal{B}_c \cap \mathcal{P}(N_1)$, hence (iii) holds. □

In the proof of *(ii)* $\iff$ *(iii)*, strong functionality is only used to infer from $X \subseteq N_2$, $D'(N_2)(X)$ and $R^*(dom, X)$ that $X = rg$, in order to show that $D'(N_2)(rg)$ holds. But this can be done alternatively by assuming that D' is upward monotone. In such a case strong functionality of R is not required. Furthermore, maximality of the witness condition is not needed, and the simple witness condition as in (62) is sufficient. The same could be done with respect to the double-*dfit* strategy:

Proposition 5 *Let R be of type set $\times$ atom, and N_1 and N_2 be NP denotations. Suppose further that D' is upward-monotone. Then the iterative, double dfit cumulativity-closed strategy $dfit_w(D)(N_1 s)(\{A : dfit_w(D')(N_2 s)(R^*_A)\})$*
is equivalent to van der Does' definition of cumulative quantification (111).

PROOF: First assume (i). The claim that $\cup(\mathcal{B} \cap \mathcal{P}(N_1)) \subseteq dom$ holds under these circumstances as well and the proof is as in the first equivalence of proposition 4. We only rely on the witness condition as given in (62) (of course we could also assume total maximality). By the witness condition we have: $\exists W \in \mathcal{P}(N_1) \cap \mathcal{B}[D(N_1)(W)]$. Let us fix W as a witness. In particular, $W \in \mathcal{B}$. ¿From the same considerations as in the proof of the first equivalence of proposition 4, $W \in \mathcal{B} \Rightarrow W \subseteq dom$. ¿From the fact that R is cumulatively closed we have: $W \subseteq dom \Rightarrow R^*_W \subseteq R^*_{dom} \Rightarrow \cup(R^*_W \cap \mathcal{P}(N_2)) \subseteq \cup(R^*_{dom} \cap \mathcal{P}(N_2))$. ¿From upward-monotonicity of D': $D'(N_2)(\cup(R^*_W \cap \mathcal{P}(N_2))) \Rightarrow D'(N_2)(\cup(R^*_{dom} \cap \mathcal{P}(N_2)))$. Hence $dom \in \mathcal{B}$, and furthermore $dom \subseteq \cup(\mathcal{B} \cap \mathcal{P}(N_1))$. So we conclude $dom = \cup(\mathcal{B} \cap \mathcal{P}(N_1))$, whence $D(N_1)(dom)$. Since $\cup(R^*_{dom} \cap \mathcal{P}(N_2)) = rg$ we have $D'(N_2)(rg)$.

Conversely, the proof is as in proposition 4. □

5.6 An extension of Keenans's result to plural quantification

Keenan's result as stated in section 2.3 can be straightforwardly extended to the realm of plural quantification. The idea is to restrict ourselves to a certain domain of relations, which can be mapped to the entire set of relations over a simple domain, with no collections. Then we will use the atomic cumulativity case pointed out by Keenan to conclude that the same holds for the plural case.

We want to show that there are no type $\langle 1 \rangle$ functions over collections such that their composition is equivalent to the type $\langle 2 \rangle$ function over collections which gives us the cumulative reading.

Let us fix two NP denotations N_1 and N_2, and also two determiners D and D', that we simply choose to be the numerals 5 and 3 respectively. Now it is clear what type $\langle 2 \rangle$ function is at stake here: $F(R) = cum(D)(D')(N_1)(N_2)(R) = cum(5)(3)(N_1)(N_2)(R)$. By *cum* I mean van der Does' definition of cumulative quantification given in (111). Suppose for the contrary that there exist two unary quantifiers Q_1 and Q_2 of type $(ettt)$, that is sets of predicates over collections, such that their corresponding type $\langle 1 \rangle$ function can be composed to produce the above $F(R)$. Let us extend the notation Q_1 and Q_2 also for their corresponding type $\langle 1 \rangle$ functions, so that the original quantifiers of the ordinary sense are simply the restriction of the type $\langle 1 \rangle$ functions to one-place relations.

We define the following "lowering" of relations and quantifiers over collections. Since we restrict ourselves to the domain of singletons, the mapping is most natural. For a binary relation R over singletons, we define $R'(a,b) \iff R(\{a\},\{b\})$. For unary relations $R_{(1)}$ we define $R'_{(1)} := \cup R_{(1)}$ We define the *lowering* of Q_1, Q_2 as: $Q'_i(A) \iff Q_i(AT(A))$ (or equivalently $Q_i(A) \iff Q'_i(\cup A)$; but only when A consists of singletons, otherwise the lowering is not defined), where AT is as defined in (33), that is a function which produces all the singletons of elements in the set given as the argument. We distinguish between two cases. First assume that $\emptyset \notin Q_2$. It is straightforward to prove that the following holds, for any relation R over SGT, the singletons:

(118) $Q_1(Q_2(R)) \iff Q'_1(Q'_2(R'))$

Proof:

(119) $Q_1(Q_2(R)) \iff Q_1(\{A : Q_2(R_A)\}) \iff Q_1(\{A \in SGT : Q_2(R_A)\})$

The last equivalence holds due to the assumption that $Q_2(\emptyset)$ does not hold and the fact that R is over the singletons. ¿From the definitions of the lowerings of relations and quantifiers it follows that for any singleton A, $Q_2(R_A)$ is equivalent to $Q'_2(R'_A)$. Hence an equivalence chain can continue from (119) to

(120) $\iff Q_1(\{\{a\} : Q'_2(R'_a)\}) \iff Q'_1(\{a : Q'_2(R'_a)\}) \iff Q'_1(Q'_2(R'))$

Next assume that $\emptyset \in Q_2$. Denote by $\mathcal{F}$ the set of all those elements which are not singletons. Define:

(121) $Q^*_1 := \{A \subseteq SGT : A \cup \mathcal{F} \in Q_1\}$

The set $\{A : Q_2(R_A)\}$ can be expressed as $\{A \in SGT : Q_2(R_A)\} \cup \mathcal{F}$. Hence if $Q_1(\{A : Q_2(R_A)\})$ then $Q^*_1(\{A \in SGT : Q_2(R_A)\})$ holds. On the other hand, if $Q^*_1(\{A \in SGT : Q_2(R_A)\})$ then $Q_1(A \subseteq SGT : Q_2(R_A)\} \cup \mathcal{F})$ holds and hence $Q_1(\{A : Q_2(R_A)\})$ (by just expressing the argument of the quantifier differently). So we have:

(122) $Q_1(\{A : Q_2(R_A)\}) \iff Q^*_1(\{A \in SGT : Q_2(R_A)\})$

We proceed as follows: $Q_1(Q_2(R)) \iff Q_1(\{A : Q_2(R_A)\}) \iff Q^*_1(\{A \in SGT : Q_2(R_A)\}) \iff Q^*_1(\{A \in SGT : Q'_2(R'_A)\}) \iff Q^*_1(\{A \in SGT : Q'_2(R'_{A'})\}) \iff Q^{*'}_1(\{A' : Q'_2(R'_{A'})\}) \iff Q^{*'}_1(Q'_2(R'))$.

In any case we have found 2 type $\langle 1 \rangle$ functions the composition of which is equivalent for relations over individuals to the lowering of F. If we show that the lowering of the *cum* function over collection gives us back the atomic cumulative reading, then we have Q'_1 and

Q'_2 or $Q_1^{*'}$ and Q'_2 as counter-examples for Keenan, which is a contradiction. So what is left to show is that the lowering of F defined above is identical to atomic cumulativity, that is: if F' is such that for any R over the singletons, $F(R) \iff F'(R')$, then for any relation R over the simple non-collective domain, $F'(R) \iff at - cum(5)(3)(N_1)(N_2)(R)$. But it is very easy to see that it follows from the following trivial identity, for any R over the singletons:

(123) $R' \cap N_1 \times N_2 = \cup(R \cap \mathcal{P}(N_1) \times \mathcal{P}(N_2))$

This result may be somewhat strengthened to refute the existence of a couple of type $\langle 1 \rangle$ functions whose composition is only partially equivalent to the type $\langle 2 \rangle$ *cum* function, by restricting the domain of relations as long as it is kept sufficiently rich to contain an image of the the class of relations over atoms. So for instance we could constrain the domain to only *set* $\times$ *atom* relations in the format dealt with in the previous subsections. [22]

5.6.1 Getting away with Keenan?

The most perspicuous thing about the iterative method of deriving the cumulative reading discussed in previous sections is that it is valid only for a highly constrained class of relations. Actually the iterative method applies only to relations into which cumulativity is built, that is cumulativity closed relations. So what has been found here does not refute (and is not refuted by) Keenan's findings and their extension to the plural case given in the previous subsection, since the iteration of the type $\langle 1 \rangle$ functions employed there does not compute the iterative reading for an *arbitrary* relation, i.e. for the entire domain of relations. The idea, however, is that an *arbitrary* relation of a particular type should "go through" a closure operation before being fed into the iterative procedure. Here it is important to bear in mind that the cumulative closure is itself a *polyadic* operation, not an iterative one.

The iterative strategy basically finds the maximum of the cumulative closure of R. It should not be surprising that the relational maximality can be found iteratively: to find the maximum of R^*, it is possible to first construct the set of all elements A such that R^*_A is maximal, and then find the maximum among these elements.

We already know from the atomic case that for certain relations the cumulative reading is iterative. For relations over collections, there are other sorts of relations on which cumulativity is an iteration. Relations closed under cumulativity are of that kind. Notice that if D' is monotone-increasing, then R is not constrained by strong functionality. We can take the example $cum(exactly - 3)(at - least - 5)$ as a type $\langle 2 \rangle$ function which is equivalent to an iteration of type $\langle 1 \rangle$ functions on cumulative closures of arbitrary relations.

[22]But then slight changes are required in the definitions of the lowering functions; however these modifications are trivial, such as redefining lowering of binary relations only with respect to one of the arguments: $R'(a, b) \iff R(\{a\}, b)$.

6 Conclusions

Is a uniform strategy of collective quantification at hand? We have examined Winter's lifting strategy, which is a combination of two of Van der Does' liftings: the one which provides the neutral reading of noun phrases, namely the pure *dfit* component and the one which provides the collective reading, namely the witness condition. Each of these liftings considered separately is not generally applicable.

The plausibility of the witness condition has been discussed, but the issue has not been settled. One of the shortcoming of this condition is the missing-witness problem, to which the cumulativity operator has been suggested as a solution. Apart from this, in some cases it is the witness condition which makes the crucial contribution to the truth conditions of the sentence. In cases like "All rocks are similar", the *dfit* analysis appears plausible. However, in a case like "All the children are different", the witness condition provides an analysis which is too weak to distinguish this sentence from the "vague" predication "The children are different".

The witness condition as proposed by Winter has been slightly modified by dropping the exception for empty sets. This move has been carried out as an innocent and purely formal decision, whereas it has been shown to have implications for cumulative readings to be discussed below. The ideas about maximality and the witness condition brought up by Winter have been discusses and incorporated into the main proposal concerning cumulative readings. However, a very strong notion of maximality seemed to be needed, which changed the witness condition quite drastically. Rather than existentially quantifying over a witness set, the witness condition requires that the total cumulation should be a witness. This results in a variant of the original $dfit_w$ approach.

The main problem with the neutral reading discussed here is the non intuitive readings which are generated when a neutral noun phrase exercises wide scope. The main shortcoming of the collective reading is Van Benthem's problem, that is the non intuitive way in which it handles determiners which are not monotone increasing. Each of these liftings seems to balance the negative effects of the other. Nevertheless, the *dfit* strategy as it stands does not provide a fully satisfactory solution to the scope problem associated with neutral readings.

It has been claimed that the only way to deal reasonably with transitive verbs such as *write* on an iterative *dfit* approach is to apply an operation of cumulative closure. In combination with the assumption that this operator has been applied, the *dfit* strategy overcomes the neutral problem as well. So as far as the first aspect of the main question raised in the introduction is concerned, it appears that the *dfit* strategy may indeed be considered as a step towards a uniform strategy of plural quantification. However, cumulative closure is not generally applicable for all predicates. Hence this unifying attempt is not general either. This is also where the second aspect of the main question comes in. It has been shown that the readings obtained after curing this flaw are precisely cumulative readings, which have been claimed to be non-iterative.

However, the assumption that a cumulativity operator has to be applied results in a highly restricted class of relations which are actually fed into the iterative *dfit* procedure that generates cumulative readings. Hence, Keenan's result that cumulativity cannot be generated iteratively does not clash with the present findings.

Under certain circumstances, it is sufficient to assume that a relation R is closed under cumulativity in order to generate the cumulative reading. For example, if D', the determiner associated with the position of the object, is monotone-increasing, no further assumptions are needed about R. However, for a more general class of logical determiners, a further restriction is imposed on R, namely the property which has been termed here *strong functionality*. Some motivation for assuming this highly restrictive property have been given. The main idea is that sets are not always treated extensionally, that is just in terms of the elements they consist of. Hence, an operation of cumulative closure is liable to produce non intuitive results when the requirement of strong functionality is not met, witness the example of *outnumber*.

Yet the proposal has some further limitations. Some of them have to do to the empty-set problem for cumulative readings. It seems that cumulative readings are summations of "positive" events (perhaps a more appropriate term would be *terminative* events, i.e. events with a terminative aspectuality), unlike, for instance, a case in which 5 children lifted 0 tables. Hence, the proposal does not apply when the determiner D' associated with the position of the object produces the generalised quantifier "nothing" when combined with the object restrictor. However, non-positive determiners are not totally excluded. For example, the downward monotone determiner "at most 6" does not give rise to any problems.

On the one hand, it has been shown that under certain motivated assumptions, cumulative readings are iteratively derivable. On the other hand, the nature of these assumption may challenge the idea that it is appropriate for ordinary theories of quantification to deal with these readings to begin with. Apart from this, the question concerning the availability of cumulative reading has been mostly ignored. It has been acknowledged, however, that such readings are not always available. A thorough investigation into the conditions under which such readings obtain would be required in order to settle this issue.

References

[1] J. Barwise and R. Cooper. Generalized quantifiers and natural language. *Linguistics and Philosophy*, 4:159–219, 1981.

[2] J. van Benthem. *Essays in Logical Semantics.* Reidel, Dordrecht, 1986.

[3] Martin H. van den Berg. *the internal structure of discourse.* PhD thesis, Univerisiteit van Amsterdam, 1996.

[4] J. van der Does. Sums and quantifiers. In F. Hamm and E. Hinrichs, editors, *Plurality and Quantification*, pages 239–282. Kluwer Academic Publishers, 1998.

[5] J. van der Does and H. Verkuyl. Quantification and predication. In J. van der Does and J. van Eijck, editors, *Semantic Ambiguity and Underspecification*, pages 27–54. Stanford University, 1995.

[6] D. Dowty. Collective predicates, distributive predicates and *all*. In *proceedings of the Eastern States Conference on Linguistics, ECOL3*, 1987.

[7] E. L. Keenan. Further beyond the frege boundary. In J. van der Does and J. van Eijck, editors, *Plurality and Quantification*, pages 179–201. Stanford University, 1995.

[8] E. L. Keenan and D. Westerståhl. Generalized quantifiers in linguistics and logic. In J. van Benthem and A. ter Meulen, editors, *Handbook of Logic and Language*, pages 837–893. Stanford University, 1997.

[9] F. Landman. Groups I & II. *Linguistics and Philosophy*, 12:559–605, 723–744, 1989.

[10] G. Link. Plural. In D. Wunderlich and A. von Stechow, editors, *Semantics. An International Handbook of Contemporary Research.* de Gruyter, Berlin and New York, 1991.

[11] Lønning. Plurality and collectivity. In J. van Benthem and A. ter Meulen, editors, *Handbook of Logic and Language*, pages 1008–1053. Elsevier Science B.V., 1996.

[12] R. Scha. Distributive, collective and cumulative quantification. In T. Groenendijk, J. Janssen and M. Stokhof, editors, *Formal Methods in the Study of Language*, volume I & II, pages 483–512. Mathematical Centre, Amsterdam, 1981.

[13] Y. Winter. Dependency and distributivity of plural definites. In *De Dag, proceedings of the workshop on definites*, 1997.

[14] Y. Winter. Atom predicates and set predicates: towards a general theory of plural quantification. In *proceedings of SALT8*, 1998.

Contents