Plural Pronominal Anaphora in Context:

Dynamic Aspects of Quantification

Published by
LOT
Trans 10
3512 JK Utrecht
the Netherlands

Phone: +31 30 253 6006
Fax: +31 30 253 6000
e-mail: lot@let.uu.nl
http://wwwlot.let.uu.nl/

Cover illustration: *by Alison Unsworth*

ISBN 90-76864-45-4
NUR 632

Plural Pronominal Anaphora in Context:
Dynamic Aspects of Quantification

Meervoudige Pronominale Anafora in Contekst:
Dynamische Aspecten van Kwantificatie
(met een samenvatting in het Nederlands)

Proefschrift

ter verkrijging van de graad van doctor
aan de Universiteit Utrecht
op gezag van de Rector Magnificus, Prof. dr. W. H. Gispen,
ingevolge het besluit van het College voor Promoties
in het openbaar te verdedigen
op vrijdag 21 november 2003
des middags te 14.30 uur

door

Rick Willem Frans Nouwen
geboren op 12 september 1974 te Heerlen

Promotores: Prof. dr. H.E. de Swart
Prof. dr. D. J. N. van Eijck

Contents

Acknowledgments

First of all, some words of gratitude are in order to both my supervisors, Henriette de Swart and Jan van Eijck. I thank Henriette for stimulating me and patiently putting up with my stubborness. Moreover, I'm grateful that she always gave me the freedom to explore new (sub)topics and approaches, even though these excursions often lead to dead-ends. I thank Jan for his enthusiasm, encouragement and for a fruitful collaboration.

Some people read parts of the manuscript, or even complete non-final versions of the dissertation and gave useful comments. I'm greatly indebted to Kees Vermeulen, Alexander Kaiser and Paul Dekker. The book improved considerably on the basis of their comments. Thanks also to Sharon Unsworth for proof-reading my English.

Then, there is a long list of people who deserve a word of thanks because of one or more of the following things: their teachings, their willingness to discuss whatever linguistic or non-linguistic topic, their occasional or ongoing constructive criticism, their collegiality and/or their friendship or because of whatever way they enhanced the general quality of my life and/or work. In alphabetical order: Albert Visser, Alexis Dimitriadis, Anna Feldman ('We and Anna cleaned the house'), Anna Młynarczyk (shopping in Leiden), Anne-Marie Mineur, Balder ten Cate, Benjamin Spector, Christophe Costa Florêncio (owner of the smallest cinema in Utrecht), Crit Cremers, Dimitra Papangeli, Elisabeth Villalta, Emiel Krahmer, Esther Kraak (personal advisor, movies and Thai food), Evi Vlachou, Frank Drijkoningen, Helen de Hoop, Heleen Hoekstra, Henk Verkuyl, Henk Zeevat, Jaap van der Does, Jack Hoeksema, Jan Don, Jenny Doetjes, Jeroen Groenendijk, Jocelyn Ballantyne, Joost Zwarts, the Baron Karl von Drais von Sauerbronn, Kata Balogh, Kees Vermeulen, Kriszta Szendrői, Linda Eilers, Maaike Schoorlemmer, Maarten Janssen, Marie Safarova, Marijana Marelj, Nicholas Asher, Oele Koornwinder, Olga Borik, Øystein Nilsen (small scale riots), Patrick Brandt (general positivist), Paz González (excellent roommate), Paul Dekker, Raffaella Bernardi, Richard Moot, Robert van Rooy, Sharon Unsworth (heel veel lol), Silke Hamann, Ton van der Wouden, Willemijn Vermaat and Yoad Winter.

Special thanks to Alison Unsworth for enthusiastically designing the cover illustration especially for me.

For the excellent discussion and atmosphere, I thank the organisers and participants of the dip colloquia in Amsterdam, the two sin-days (in Leiden and Nijmegen) and the semantics colloquia in Nijmegen. Thanks to the participants of several reading and discussion groups involving fellow PhD students at the OTS. These get-togethers kept the institute alive. They provided an ideal breeding ground for wild ideas and deep discussions. I would also like to thank the organisers and participants of four ESSLLI summerschools (1999-2002), especially those involved in the student sessions.

Tenslotte zijn er vier dierbare personen die ik hier in deze simpele tekst niet genoeg kan danken voor hun geduld en hun onvoorwaardelijke steun en liefde: Sharon, mijn zus Mariëlle en mijn ouders, mam en pap.

Chapter 1

Introduction

Pronominal reference is central to one of the classic puzzles of formal semantics, namely what does an utterance containing a pronoun, an *open sentence*, mean. In answering this question, *meaning* should clearly be conceived of as something context-dependent. Take a simple example like "he VP". How such a sentence is interpreted is determined by the (con)text it occurs in. In a discussion about X, the utterance expresses that X has the property described by the VP. Whatever precise semantics we decide on, it will have to imply that the pronoun refers to an individual that is salient in the current discourse.

At first sight, plural pronouns are no different from singular pronouns; they merely appear to refer to plural objects salient in discourse rather than salient atomic entities. Although the semantics of a plural pronoun differs only trivially from that of a singular pronoun, we will see that by not limiting our attention to singular anaphora, more *kinds* of anaphora emerge. For instance, whereas quantified singular NPs such as 'every student' can only antecede singular pronouns in their scope, quantified plural noun phrases like 'all students' or 'most students' can stand in an anaphoric relation even at the extra-sentential level.

(1.1) Every student thinks he wrote an excellent paper.

(1.2) Every student wrote a paper. #He thought it was excellent.

(1.3) Most students think they wrote an excellent paper.

(1.4) Most students wrote a paper. They thought it was excellent.

Quantification and its role in setting up antecedents in discourse will play an important part in this thesis, since quantificational noun phrases challenge two of the basic assumptions that I will make here, namely (i) that pronouns are to be interpreted *uniformly* as variable-like entities with hardly any meaning themselves and (ii) that the variation in pronominal reference is *not* due to ambiguity of the pronoun, but rather it is a result of variation in context. These assumptions are well-established in formal semantics and are central to the tradition of *dynamic semantics*. Quantificational noun phrases, with their potential for numerous anaphoric effects, form a well-known complication. The basic problem is this: how can all these effects be due to the (same) context? The challenge then is to devise a sufficiently sophisticated semantics of quantification and a sufficiently fine-grained notion of context and show that the interaction between the two results in the reported patterns of anaphora.

In this first chapter, I will start with a systematic overview of the issues that play a role in this thesis, starting with the very basics of the relationship between pronouns and context.

1.1 Interpreting pronouns in context

It is easy to accept that the meaning of an expression containing a pronoun is dependent on the context such an expression occurs in. It is notoriously difficult, however, to indicate what this relation is like exactly. This is especially so since pronouns seem to be able to stand in several kinds of relations to their antecedent.

For the better part of the second half of the 20th century, there has been a lively debate on how to analyse different occurrences of anaphoric pronouns. An initial contributor to this debate was Peter Geach. In his 1962 book, he stated that anaphoric pronouns in natural language correspond to the bound variables of predicate logic. Together with a second hypothesis due to Geach, namely that indefinite noun phrases are existential quantifications, this makes it possible to analyse (1.7) which contains a pronoun which is said to be 'referentially used' as (1.8). This analysis is comparable to the more straightforwardly bound variable interpretation of (1.5) as (1.6).

(1.5) Every man knows he is mortal.

(1.6) $(\forall x\text{: man}(x))$ (x knows x is mortal)

(1.7) A man walked through the park. He whistled.

(1.8) $(\exists x\text{: man}(x))$ (x walked through the park & x whistled)

Evans (1977) criticised such a view on many points. First of all, he questioned how the pronoun 'he' in (1.7) can be a bound variable when it is outside the scope of its operator. Second, given Geach's strategy, it seems difficult to explain the difference between the alleged bound interpretation of 'he' in (1.7) and the impossibility of such a reading in (1.9).

(1.9) Every student wrote a paper. # He also read a book.

These objections focus on the fact that if pronouns are variable-like, then in (1.7), the pronoun ‘he’ should not be a *bound*, but rather a *free* variable. Just like –intuitively– the pronoun in (1.9) is free since it is out of the scope of the universal quantifier. The example in (1.9) cannot mean that every student wrote a paper and read a book. It appears then that the problem with (1.8) is a matter of compositionality. In order to derive the truth-conditions of the first sentence one will have to analyse it as a closed quantificational sentence. Subsequent anaphoric pronouns should not be able to change this meaning.

One might think, however, that the limit to the scope of a quantifier is merely an unfortunate characteristic of predicate logic and that the first sentence in (1.7) should be meaningful without closing off the scope of the existential quantification. This is the point made by dynamic semantics. Dynamic theories of meaning assume that the meaning of a full expression should not be identified with its truth-conditions, but rather with its potential to change the context.

The notion of *context* used here and in the rest of this thesis is defined as a medium supplying values for variables or other labels of subjects introduced in discourse. Anaphoric reference is possible to those values that are specified somewhere in the context. [1]

As long as this context provides a value for x, an occurrence of x will be dynamically bound whether or not it is in the classical scope of an operator. So, while in predicate logic (1.10) is ‘open’ due to the free last occurrence of x, in dynamic logic x is *dynamically bound* by ‘$\exists$’. In fact, in a dynamic set-up, (1.8) and (1.10) mean the same thing: they express the same potential to change the context.

(1.10) ($\exists$x: man(x)) (x walked through the park) & x whistled

Still, in order to account for the contrast between (1.7) and (1.9), dynamic theories of meaning will have to assume that not every noun phrase has the same kind of context-change potential. According to the second dogma of Geach we mentioned, an indefinite NP corresponds to an existential quantifier, as in (1.10). In the dynamic logic, then, the existential quantifier is dynamic (i.e. it contributes change to the context), while the universal quantifier is static: it does not export a value for its ranging variable. So while after processing ($\exists$ x: φ)(ψ) the context contains information about the value for x, after processing ($\forall$x: φ)(ψ) this is not the case. That is:

(1.11) ($\exists$x: φ)(ψ) & τ = ($\exists$x: φ)(ψ & τ)

[1] ‘Context’ as used here should not be confused with some other frequent appearances of the word, such as the situational configuration in which an utterance is specified, the set of (labels for) discourse subjects introduced or the set of possible worlds one is willing to entertain (see Visser 1998 for discussion). Context as used here represents a set of labelled values which stand for the singular or plural individuals that play a role in the (representation of) the discourse.

(1.12) $(\forall x: \varphi)(\psi) \& \tau \quad \neq \quad (\forall x: \varphi)(\psi \& \tau)$

According to the view held by dynamic semantics, interpretation always occurs *in context* and context is always created *by interpretation*. The second sentence in (1.9) is infelicitous since it is interpreted in a context which does not supply a potential value for the pronoun 'he'. That is, the universal quantification in the first sentence does not change the context in such a way that it contributes a suitable value for 'he'. In contrast, the first sentence in (1.7) does contribute such a value.

While accounting for the issues sketched above there is a third objection raised by Evans against Geach that is more problematic for dynamic semantics. Some cases of pronoun use are neither prototypically bound uses (such as (1.5)), nor are they prototypically referential (such as (1.7)). Consider (1.13):

(1.13) Few senators admire Kennedy; and they are very junior.

Notice that if we were to treat the plural pronoun as a bound variable, the sentence would read that few senators at the same time admire Kennedy and are very junior. This, however, could be true in a situation in which most senators admire Kennedy, but very few of these admirers are junior. But this is not the meaning of the sentence. What the plural pronoun does appear to refer to is *the set of all senators that admire Kennedy.*

Unfortunately, there is no straightforward way of deriving such a reading in a dynamic approach. Obviously, "few senators" cannot be treated as an existential quantification in the same way indefinites were. For a start, "few senators admire Kennedy" does not even guarantee there are *any* such senators. Moreover, were "few senators admire Kennedy" to change the context by adding a set of few Kennedy-admiring senators to it, this would not create the correct anaphoric effect. The pronoun is clearly exhaustive in reference: not just *some* small set of senators that admire Kennedy are supposed to be junior, but rather *the* small set of senators that admire Kennedy have that property.

Evans calls all pronouns that are neither bound nor 'go proxy' for their antecedent 'e-type pronouns.'[2] In his view, both 'he' in (1.7) and 'they' in (1.13) are e-type. He proposes to treat them as terms that receive their reference through the reconstruction of a description from the antecedent sentence. Thus, in (1.13), 'they' is interpreted as 'the senators who admire Kennedy,' just like in (1.7), 'he' could be interpreted as 'the man that walked through the park.' This approach has been called the *e-type strategy*.

[2]It is important to make a distinction between the class of e-type pronouns as defined here and Evans' treatment of these pronouns. In the literature, an e-type pronoun is often used as being synonymous with a pronoun that has been analysed as corresponding to some reconstructed definite description. In what follows, I prefer to use the term to refer pre-theoretically to a class of pronouns.

In Evans' analysis, pronouns are either bound or referential (i.e. e-type). Bound variable pronouns, in contrast to e-type pronouns, have to occur in construction with their antecedent (i.e. they have to stand in some syntactic relation with their antecedent, such as, for instance, the c-command relation.) This means that intersentential pronominal anaphora is always e-type.

Again, some complications arise. It turns out that while the dynamic approach failed for discourse pronominals in a context containing a quantificational antecedent like "few senators," the e-type approach runs into problems with simple indefinites.

If pronouns refer like definite descriptions, then they impose uniqueness requirements on their referent. This is especially clear in examples like (1.14) (from Heim (1982), who credits the stoic Chrysippos). An e-type strategy is in danger of analysing this example as saying that if a man is in Athens, the unique man in Athens is not in Rhodes.

(1.14) If a man is in Athens, he is not in Rhodes.

Although quite some work has been devoted to solving this problem and discussing its relation to the semantics of conditionals, quite simple examples, like (1.15), show that the problem is actually very basic.

(1.15) Three men came in. They ordered beer.

The e-type strategy will have to explain why in (1.15) the pronoun does not refer to the set of all men that came in, but rather to just the three mentioned in the linguistic context. The element of maximality that played an important role in (1.13) is absent from (1.15). Strikingly, a dynamic approach has few problems with (1.15) if we assume that the first sentence expresses the existence of a set of three men that came in and changes the context by adding this set to it.

What I have sketched here is the core of a complicated debate which, following Geach's work in the sixties and Evans' work in the seventies, had its high-point in the eighties with the emergence of dynamic approaches to meaning (Kamp 1981; Heim 1982) and subsequent attempts at reinterpretations of the e-type strategy (e.g. Neale 1988; Heim 1990).

This thesis has its roots in the dynamic semantic tradition, but it tries to analyse exactly those cases which do not naturally follow from the original setup of that tradition. The key aspects of the problematic examples like (1.13) are quantification and plurality and their anaphoric potential. This is what I will focus on in this book. A closer look at the relation between the semantics of quantificational expressions and the antecedents they present will make clear which are the important issues.

1.2 Pronouns and sets

Let us assume that any simple declarative English sentence of the form [[Det N′] VP] is, in principle, representable as a tripartite structure $D(A)(B)$ consisting of a determiner D, a restrictor A and a (nuclear) scope B. The truth-conditions of such a structure depend on the set of individuals that make up A, the set corresponding to B and the relation between sets expressed by D.

A well-known (near) universal of natural language determiners is that they are conservative (Barwise and Cooper 1981). That is, $D(A)(B)$ expresses the same proposition as $D(A)(A \cap B)$. The scope-denotation relative to the restrictor-denotation, i.e. $(A \cap B)$, turns out to be the set of entities which seems to form an antecedent in Evans' example (1.13). That is, if we represent the semantics of this example as in (1.16), then the reference of the e-type pronoun 'they' in (1.13) is the intersection of λx.senator(x) and λx.admire(x,Kennedy). We call this set *the reference set* or, sometimes, simply refset.

(1.16) few(λx.senator(x))(λx.admire(x,Kennedy))

(1.17) very-junior((λx.senator(x)) $\cap$ (λx.admire(x,Kennedy)))

Following Evans' observations, we could radically hypothesize that all e-type pronouns take the reference set of a quantificational structure as their antecedent. Interestingly, there are two reasons why this does not seem to hold: (i) not all quantificational structures license subsequent reference to the (full) reference set, and (ii) sets associated with a quantificational structure other than the refset seem to be able to antecede a discourse pronoun as well.

The first problem relates to the uniqueness and maximality issues we briefly discussed above. For (1.14), the proposal implies that the pronoun 'he' picks up the set of men in Athens, but our world-knowledge tells us that other than the singularity of the pronoun suggests, the set of men in Athens is not a singleton set. For (1.15), according to our faulty hypothesis, the pronoun refers to the set of men coming in. The antecedent sentence, however, merely describes a set of three men and says nothing about this total set of men entering.

In sum, there are two ways the reference set can be involved in anaphora. In other words, e-type pronouns come in two varieties: (i) those that refer to the reference set, and (ii) those that refer to a subset of the reference set.

The second mismatch between e-type pronouns and reference sets is illustrated in (1.18) and (1.19), where in both examples, the plural pronoun refers to a set distinct from the refset.

(1.18) Few senators admire Kennedy. Most of them prefer Carter.

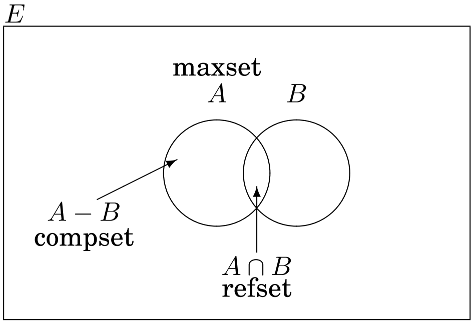


Figure 1.1: Sets associated with a structure $D(A)(B)$

(1.19) Few senators admire Kennedy. They admire Carter instead.

The continuation in (1.18) is interpreted as stating that the majority of senators admire Carter. This means the plural pronoun 'them' has to refer to the set of senators, the denotation of the restrictor. Similarly, the continuation in (1.19) seems to express that the senators that do not admire Kennedy admire Carter instead.[3]

Let me, at this point, introduce some more terminology. Given a quantificational structure $D(A)(B)$, we call the set A the *maximal set* and the set-theoretical difference between restrictor and scope $A - B$ the *complement set*. Sometimes I will use abbreviations as indicated in figure 1.1.

We have seen that reference to the reference set (1.13), the maximal set (1.18) and the complement set (1.19) is possible. Note, however, that these last two types of pronominal reference are not generally available.

(1.20) Two senators who admire Kennedy bullied Carter. They are very junior.

(1.21) Most senators admire Carter. # They admire Kennedy instead.

The example in (1.20) can only be understood to mean that the two bullying senators were junior, not that the senators who admire Kennedy are in general junior. Similarly, in (1.21), the pronoun cannot be used to refer to the senators who do not admire Carter.

We see then that not just one but several sets associated with a quantificational structure allow for subsequent pronominal reference to them, but that not all of these sets are proper antecedents in all contexts. In

[3]If the reader is of the opinion that (1.19) is not felicitous or does not agree with the paraphrase given here, then she is referred to chapter 3, where the exact status of such examples is examined in much detail.

order to understand these restrictions we have to analyse what the relation between a set associated with a quantificational structure and the constraints on its use for anaphora is.

I will show in chapter 3 that whereas reference to the maximal set and the complement set seems to be constrained by the properties of the antecedent determiner, refset reference turns out to be free: any quantified structure provides the context with an antecedent corresponding to the refset or a subset of the refset. Reference sets are completely ordinary salient antecedents. In contrast, the cases in which pronominal reference to the complement set is allowed turn out to be rather special. While the infelicity of (1.21) seems to be generally accepted, the felicity of examples like (1.19) is not undisputed. Pronominal reference to the complement set is a *marked* type of anaphora. Chapter 3 will show that pronominal complement set reference is only possible under strict conditions. The non-emptiness of the complement set should be, first of all, be inferable. Moreover, there should be no set with higher salience available which also leads to a consistent interpretation, when chosen as an antecedent for the pronoun. The status of the maximal set is different from that of the complement set. When possible, pronominal reference to the maximal set is not marked. I will argue that the availability of the maximal set as an antecedent is related to whether or not a determiner presupposes the existence of its domain of quantification.

I have made two points concerning reference to sets here. First, standard examples of e-type pronouns come in two varieties: those that refer to the reference set and those that refer to a subset of the reference set. Second, other occurrences of e-type pronouns may refer to the maximal set or to the complement set.

I assume that theories of e-type pronouns aim at providing a model of the referential possibilities of pronouns. Since reference to the complement set is not a common referential possibility and is constrained by more factors than the semantics of its potential antecedent alone, it is outside the scope of such theories. What is left then is maximal set reference and the two types of reference to the reference set. The first of these, however, seems independent of the mechanisms, which according to the theory put forward here, are responsible for the creation of referential possibilities. The issue of reference to the maximal set is more a question of how a theory of presupposition links up to such a mechanism. The reference set and its subsets thus seem to be the antecedents which pose the real challenge for theories of e-type interpretation. Let us therefore return to the discussion of the dynamic strategy, with its straightforward account of non-maximal reference and the e-type strategy with its straightforward account of maximal reference to the refset, by focusing on a theory which accommodates both kinds of reference.

1.3 Discourse Representation Theory

The discourse representation theory of Kamp and Reyle (1993) offers a reconciliation of Geach and Evans. In many senses, discourse representation theory (or DRT) is set in the dynamic semantics tradition. DRT proposes to treat both indefinites and pronouns as existentially closed free variables (Kamp 1981). Quantificational noun phrases like “few senators” and “every man”, however, are static, but they do allow the creation of an antecedent from the material in the quantificational structure they form. That is,“few senators admire Kennedy” does two things: first, it tests whether the world is such that few senators have the property of admiring Kennedy. Second, it changes the context by making the set of x’s such that x is a senator and x admires Kennedy (i.e. the reference set) available for future anaphoric reference. This antecedent forming operation is called abstraction, and it should be clear that it owes a lot to Evans’ original idea. In (1.22), this process is exemplified in pseudo-formal terms.

(1.22) (few x: senators(x))(x admires Kennedy) &
Y = λ y. y is senator & y admires Kennedy

Notice that the truth-conditions of the sentence (the top line of (1.22)) results in a formula in which the quantifier ‘(few x: senators(x))’ has limited scope. The reference set antecedent, however, is set up without a quantifier, thus leaving the variable Y free. In DRT, free variables are interpreted existentially. In this way, all occurrences of a free variable co-refer. With respect to (1.22), this means that future occurrences of Y are interpreted as the senators that admire Kennedy.

The explanatory power of the abstraction procedure is impressive. For instance, it explains why (1.9) is infelicitous. The only antecedent that can be formed is a set (the set of students that wrote a paper) and such an antecedent cannot be picked up by a singular pronoun. In contrast, as predicted, the plural form is fine.

(1.23) Every student wrote a paper. They also read a book.

(1.24) ($\forall$x: student(x))(($\exists$y: paper(y))(wrote(x,y))) &
X=λz.[student(z) & ($\exists$y: paper(y))(wrote(z,y))] &
X read a book

This example illustrates why plural pronouns deserve special attention in a theory of discourse anaphors. The discourse representation theory of Kamp and Reyle (1993) is no doubt still the most complete theory of plural anaphoric reference. However, one might wonder why if quantificational structures generally allow for subsequent anaphora, they still have an in principle static representation and why they need an additional referential component, whereas for indefinites, truth-conditions and antecedent formation go hand in hand. This triggers the question whether we can rephrase the analysis of e-type pronouns in discourse representation theory in terms of a generalised dynamic treatment of quantification.

This question is of interest not only because we are uncomfortable with Kamp and Reyle's analysis for reasons of simplicity, elegance or uniformity. This thesis aims to devise a theory of (plural) pronominal reference in discourse which fits into the Fregean and Montagovian tradition of compositional, bottom-up meaning derivation, which is in sharp contrast to the top-down architecture of discourse representation theory.

In order to better understand what I mean by this, let us focus some more on discourse representation theory. In chapter 2, we go into the details of DRT's analysis of plural pronouns, but let us for now concentrate on what *kind* of theory DRT is.

DRT enables the semantic analysis of a discourse by translating the discourse into a *discourse representation structure* (a DRS) and, subsequently, interpreting these intermediate representations. The interpretation process starts with a syntactic structure in an otherwise empty DRS. Using so-called *construction rules* the DRS is expanded and transformed to form an interpretable representation. The construction rules use the syntactic representations (trees) as well as the present semantic content of the DRS itself as triggering configurations for adding semantic material to the DRS. That is, the interpretation process is fully mediated by representations.

This is in contrast to the more traditional approach of Montague Grammar (Montague 1974), where an explicit distinction is made between the construction of an interpretation and the logical content. The derivational part is given by the simply typed lambda calculus (Church 1941). Sentence constituents are represented by lambda-terms and combinations of such terms ultimately result in first order expressions through beta reduction. In this way, Montague captured the link between grammatical structure and logical structure. Each lexical entry specifies exactly what its semantic contribution is and in what way this contribution fits into the grammatical structure.

The architecture of Montague grammar follows the principle of compositionality (attributed to Frege (1884)): what matters to the interpretation of an expression is the interpretation of its parts and the way these parts are combined. In Montague's strong interpretation of this principle, both the syntax and the semantics of a disambiguated language map to algebraic structures. Moreover, each syntactic structure maps to exactly one semantic structure. DRT is not strictly compositional, since the semantic contribution of parts of expressions cannot be abstracted from the construction rules. Although there is a mapping from syntax to DRSs and one from DRSs to truth-values, there is no structural correspondence between the syntactic mapping and the semantic one.[4]

[4]Formally, what one would want to derive is a *homomorphism* between the algebra of syntactic structures and the algebra of semantic structures.

Compositionality, however, does not prohibit the use of representations. A compositional theory might, for instance, use some intermediate representational level in order for it to be more comprehensible. However, such representations should, in principle, be dispensable. In DRT, they are not.

1.4 Rational reconstructions

The empirical success of DRT gave rise to the search for theories with potentially the same empirical coverage that can model compositional meaning derivation *without* making use of representations. The rationale is that if meaning corresponds to context-change potential, there should in principle be no need for intervening representations. As Groenendijk and Stokhof (1996) put it: "what undergoes change in the dynamic process of interpretation are semantic objects, not representations." A theory of discourse is non-representational when, in principle, it can do without any level of representation. That DRT is not such a theory is intuitively clear, but it follows most convincingly from the abstraction procedure used for plural reference. The introduction of an antecedent by a quantificational sentence is not due to its interpretation (quantifiers are static in DRT), but due to reusing representational material in an abstraction.

As a compositional alternative to DRT, Groenendijk and Stokhof (1991) introduced dynamic predicate logic (DPL). DPL has the syntax of first order predicate logic, but a semantics which fully respects the slogan 'meaning is context-change potential.' Subsequently, Groenendijk and Stokhof proposed dynamic Montague grammar (DMG, Groenendijk and Stokhof 1990), which incorporated the insights from DPL in the traditional Fregean and Montagovian view on meaning derivation.

In DPL, predicate logical expressions are interpreted as relations between assignment functions: functions mapping variables to individuals. For instance, 'A man came in' is represented as '$\exists x(\text{MAN}(x) \wedge \text{CAME\_IN}(x))$'. This formula expresses a relation between two assignment functions f and g, where g differs at most from f by assigning to x some individual which is a man that came in. Subsequent expressions can now take g as an *input* assignment. For instance 'he sighed' can simply be translated as '$\text{SIGHED}(x)$' which expresses an identity relation (returning the input assignment) if and only if $g(x)$, the man who came in, sighed.

The study of so-called *rational reconstructions* of DRT does not end there, however. In the formulation of the dynamics of a semantics, there are many design choices which lead to different empirical and formal results. Most of the work in the dynamic predicate logical tradition, however, focuses on ordinary cases of dynamic binding and does not have a lot to say about DRT's treatment of quantificational noun phrases. Often, this is because only singulars are studied. Van den Berg 1996, however, presents a plural dynamic predicate logic. This work will play an important role in this thesis.

Van den Berg focuses on the fact that plural antecedents in discourse are often the result of an NP embedded in a quantificational structure. For instance, in (1.25), it is the singular noun phrase 'a paper' which antecedes the plural pronoun in the second sentence, which is to be interpreted as the set of all papers written by a student.

(1.25) Every student wrote a paper. They weren't very good.

This example presents yet another referential possibility triggered by quantification. Rather than the refset, in (1.25) a set closely related to the refset is referred to by the plural pronoun, namely the set of values for the indefinite that were encountered during quantification over students.

It turns out, however, that the relation between the set of papers and the students that wrote them remains accessible in discourse. In (1.26), the singular pronoun 'it' is dependent on the ranging value of the distributor 'each'. That is, the second sentence in (1.26) can only mean that every student submitted the paper he or she wrote to L&P.

(1.26) Every student wrote a paper. They each submitted it to L&P.

The plural pronoun in this example takes the reference set as its antecedent. Quantifying over this set, using the floating quantifier 'each' in this case, accesses the atomic individuals that make up the refset. The VP is said to be true of all these atoms. Somehow, it appears, by accessing atoms in the refset, the related (atomic) papers become accessible as well.

In fact, the quantor 'each' is not necessary to make the use of a singular pronoun to refer to individual papers felicitous. In (1.27), the second sentence, when interpreted distributively, yields the same effect as in (1.26).

(1.27) Every student wrote a paper. They submitted it to L&P.

Examples like these lead van den Berg towards a particular design choice for the plural version of DPL. Instead of having expressions denote relations between assignments of pluralities, in van den Berg's work, predicate logical formulae express relations between *pluralities of assignments*. So, for (1.26) the first sentence collects assignments g such that g assigns a student to (say) x and a paper written by this student to (say) y. While the collection of assignments gives us information about the groups involved in the sentence, the individual assignments carry specific information about the atoms. The details of this analysis will be discussed in chapter 4.

Van den Berg's insight concerning the architecture of context is not the only design choice which interests us, however. In chapter 5, we will focus on the relation between anaphora and variables. If a pronoun is like a variable, then all the potential reference-resolutions of a pronoun should be retrievable by choosing a different variable for the pronoun. All potential antecedents should thus have their own label. This can become problematic with respect to the dependent anaphora of, for instance, (1.26). Both ‘it’ in (1.26) and ‘they’ in (1.25) have the same antecedent, namely the noun phrase ‘a paper’. The two pronouns, however, refer quite differently and, as we will see, this difference cannot be explained by the difference in number feature. Potentially, then, the indefinite ‘a paper’ should be able to contribute two variables to the context.

I will study this problem in the light of a related, more technical problem with predicate logical formalisms. The dynamics of an existential quantifier forces us to manage variables carefully. If we translate ‘a man’ with $\exists x(\text{MAN}(x))$, we assume that in the discourse preceding this noun phrase no other noun phrase quantified over the variable x, since were such a quantification to exist, the introduction of ‘a man’ would cause all previous information we have gathered about x to be lost.

In order to avoid such variable clashes, I will present a variable free reformulation of van den Berg’s work. This will be based on the system of *incremental dynamics* (van Eijck 2001). This way we will not have to worry about what variable we choose for any NP. In fact, what I will show, moreover, is that the move to this variable free system enables a solution to the problem of having to have enough antecedents to cover all referential options of pronouns.

All this will be necessary before we can finally come to a dynamic semantics of quantification which has results comparable to those of DRT. The idea is that the tri-partite structure of quantification is interpreted as a comparison between two extensions of the input context. The restrictor specifies the possible (atomic) values that matter to the quantification. The nuclear scope collects those atoms which satisfy its predicational requirements and finally the determiner is a test of whether the set of possible values and the set of actual values stand in the appropriate quantifier relation. If this test succeeds, the set collected by the scope is automatically accessible in the subsequent discourse.

Working out the details of such a theory is far from easy, however. For instance, the extended context due to the restrictor clause is merely a resource for the interpretation of the nuclear scope. That is, at no point is the maximal set available as an antecedent. Only when the restrictor’s content is presupposed (and thus present in the informative component of context) will reference to the maximal set be possible.

The resulting formalism is a typed incremental variable free formalism which is far from compact. With the growing complexity of the phenomena (e.g. dependency structures), the complexity of the semantics increases as well. I take this to be an important aspect of this thesis as well. Given our goal to formulate a non-representational bottom-up alternative to DRT’s treatment of plural anaphora, we ask how much structure in context is needed once representations are banned.

1.5 Plan of the thesis

This thesis, then, focuses on how the compositional interpretation of an expression *in some context* generates a context allowing the subsequent accessibility of all referential possibilities for pronouns. The emphasis will be on *modelling* complex anaphoric phenomena, in particular those involving plural pronouns. A central topic is the set of restrictions plural pronouns impose on their antecedents.[5]

The structure of the thesis is as follows:

In *chapter 2*, I provide the necessary background information concerning discourse representation theory and dynamic predicate logic. I focus in particular on the simple formal basics of dynamic interpretation that gives us the vocabulary we need to develop the extended models below. Moreover, I give a detailed introduction in DRT's treatment of plural anaphora.

In *chapter 3*, I turn to the issue of reference to sets and discuss the constraints on pronominal anaphora with antecedents other than the reference set. Most attention is paid to to the phenomenon of pronominal reference to the complement set, given that this type of reference cannot be accounted for naturally within theories of quantification and anaphora (especially not in DRT). The informal model of plural antecedents which is proposed forms the basis for the formal proposals to follow and this will justifies the fact that, in our formal model, the complement set is never generated as an antecedent.

In *chapter 4*, I introduce the work of van den Berg and compare it to two other proposals. My main concern is to establish that van den Berg's motivations for his particular notion of context are well-founded and are compatible with the types of structure the other proposals assume context to have. At the same time, it allows us to critically examine some other aspects of existing non-representational theories of plural reference, most notably the dynamics of distributivity.

[5]Since plurality plays such an important role in this thesis, a few words on important studies in the linguist semantics of plural expressions are in order. In the early Montague semantic tradition, at least two extensions to plurality are well-known, namely Hausser 1974 and Bennet 1974. Link 1983 proposes an algebraic account of plural issues like distributivity and mass terms. *The* issue discussed in this thesis is anaphora. Two other branches of research in plurality deserve to be mentioned, although the results of such studies are largely glossed over here and are only indirectly relevant for this dissertation.

First of all, there is the question whether or not plural objects are internally structured. Key papers are: Landman 1989, Schwarzschild 1990 , Schwarzschild 1991 and Krifka 1991. A second important research topic with respect to plurality is the analysis of the distributivity/collectivity distinction, the analysis of other readings and the question of how many readings a sentence with multiple quantification has. Scha 1981 is an important source of issues and proposals of how to deal with distributivity, collectivity and cumulativity. In the late eighties and early nineties of the previous century, this evolved into a lively debate. Key papers are: Link 1987, Gillon 1987, Lønning 1991, Lasersohn 1995, van der Does 1992, Verkuyl and van der Does 1996.

With van den Berg's work at the background, I turn, in *chapter 5*, to the question of how anaphora is represented in such a theory. After arguing that the choice of which label to give to antecedents should be dependent on the context, I reformulate van den Berg's formalism within van Eijck's framework of incremental dynamics. This enables us to define a sufficiently rich dynamic interpretation of distributivity. Subsequently, I present a derivational counterpart of this semantics and discuss how pronominal interpretation is to be integrated in this approach.

Finally, in *chapter 6*, the dynamic semantics of quantification is discussed. Here, the refined dynamic semantics of distributivity of chapter 5 will be generalised and this will lead to a dynamic interpretation of distributive quantification. This will show that the additional mechanisms used by DRT to account for plural pronominal discourse anaphora can be discarded in a fully dynamic framework.

Chapter 2

Background: E-type Interpretation

In this chapter I set out to present a detailed introduction into the basics of interpreting pronouns in context.

We focus on the formal alternatives that are available for the analysis of e-type pronouns in general. As we saw in the introduction there are basically two options: either we account for e-type pronouns by giving them a semantics of their own (e.g. by considering them as reconstructed descriptions) or we dismiss the distinction between e-type and non-e-type pronouns and account for e-type effects using a sophisticated semantics sensitive to the contexts e-type pronouns occur in. Traditionally, the former strategy is called an e-type strategy, while the latter approach is associated with dynamic semantics.

This thesis is not meant to contribute to the discussion about which of these approaches leads to better results. Rather, I try to present a sufficiently strong analysis of quantification and plurality which allows us to view plural pronominal anaphora as objects which take their antecedent from a context which is built up dynamically. Nevertheless, in order to get more grip on the subject, we will start by briefly considering some variants of e-type strategies in section (2.1). Then, in section (2.2), the basics of discourse representation theory and dynamic predicate logic are discussed extensively. This will provide many of the notions that are essential for an understanding of the rest of the thesis. Finally, in section (2.3), I focus on DRT's treatment of plurality.

2.1 E-type strategies

E-type strategies have in common that they try to model the observation that e-type pronouns can be paraphrased by a definite description constructed from material in the antecedent sentence.[1] They differ in how such a reconstruction comes about, e.g. whether a pragmatic, semantic or syntactic operation is involved.[2] Cooper (1979) proposes a pragmatic strategy in which e-type pronouns are analysed as 'the $R(x)$', where the relation symbol R is resolved by pragmatics. Heim (1990) discusses alternative forms of pragmatic e-type strategies and points out some problems. One important defect is that pragmatic strategies ignore the role of a linguistic antecedent, witness the contrast between (2.1) and (2.2).

(2.1) Every man who has a wife sits next to her. = Heim 1990 (ex. 57)

(2.2) *Every married man sits next to her. = Heim 1990 (ex. 58)

For an e-type strategy to work, it seems that the reconstruction should be sensitive to the syntactic environment of the antecedent. Heim proposes such a syntactic account. The link between e-type pronouns and their antecedent is regulated by a single transformation rule (see Heim 1990, p. 170).

(2.3)

X	S	Y	NP_i	Z	$\Rightarrow$	$1 \quad 2 \quad 3 \quad 4+2 \quad 5$
1	2	3	4	5		

where: 4 is a pronoun
2 is of the form $[_S\ NP_i\ S]$

The rule says that a sentence S occurring in some context containing some noun phrase indexed i enables a subsequent co-indexed pronoun to be interpreted using the syntactic material in S. The symbols 'X', 'Y' and 'Z' express the contextuality of this rule. The sentence 'S' and the e-type pronoun 'NP_i' are not required to stand in any particular syntactic relation, they appear in their own context. The newly formed 'augmented pronoun', i.e. '4+2', is interpreted as a definite description.[3]

[1]There has been much discussion on how serious one should take this paraphrase and whether e-type pronouns *go proxy* for definite descriptions (as most of the 'modern' proposals claim) or whether the reference of the pronoun is fixed by a description. Evans himself rejected the proxy view. But see Neale 1988 (p. 158–169) for a defence in favour of the proxy view.

[2]Adopting some e-type strategy to e-type pronoun interpretation does not mean that one accepts Evans' distinction between bound variable and e-type pronouns. One could choose for a strong theory proposing that *all* pronouns are interpreted via the e-type strategy.

[3]Interestingly, Irene Heim herself observes a problematic issue involving plurality (Heim 1990, p. 172–173). When e-type pronouns are accounted for by a syntactic reconstruction process, then plural pronouns as in (i) would be analysed as 'the papers that x turned in' instead of 'the papers the students turned in.'

(i) Every student turned in a paper. They were all identical. = Heim 1990 (ex. 79)

(2.4) $[\![\text{it } [\text{Det}_x\ \alpha]\ \beta]\!]^{g}\ =$ the unique d such that $[\![\alpha]\!]^{g[x/d]} = [\![\beta]\!]^{g[x/d]} = 1$

Alternatively, Neale (1990) introduces a semantic rule to come to an interpretation of the e-type pronoun as a definite description:

> "If x is a pronoun that is anaphoric on, but not c-commanded by '[Dx:Fx]' that occurs in an antecedent clause '[Dx:Fx](Gx)', then x is interpreted as the most "impoverished"definite description directly recoverable from the antecedent clause that denotes everything that is both F and G." Neale 1990 (p. 182).

The uniqueness condition in (2.4) is also implicit in Neale's rule and eventually leads to predictions which are too strong (just like Cooper's pragmatically saturated definite descriptions would). For instance, in (2.5), we are forced to interpret 'he' as the unique man in Athens.

(2.5) If a man is in Athens, he is not in Rhodes.

Most of the modern work on e-type strategies is dedicated to resolving the uniqueness issue (see e.g. Lappin 1989; Kadmon 1990; Neale 1990; Chierchia 1992; van der Does 1993). I would like to leave the discussion at this point, however. The (interesting) finesses of modern e-type strategy proposals are beyond the scope of this thesis. What is important from our point of view is the idea that at least some pronouns are descriptive in nature. Interestingly, when adopting a strong version of the e-type strategy, one wherein *all* pronouns correspond to descriptions, the differences with dynamic semantics start to blur, since such a strong theory will have to explain how in every context the right descriptions are recoverable. The remaining difference seems to be only due to the fact that dynamic semantics naturally restricts itself to talk about predicate logic and variable binding. In van der Does 1993, however, a system is presented in which dynamic semantics and the e-type strategy seem to converge. Moreover, as I already suggested in the introduction and as we will see in more detail in the current chapter, the discourse representation theory of Kamp and Reyle 1993 also uses elements from the e-type strategy by modelling maximal discourse anaphora using a reconstruction process.

2.2 Dynamic semantics for singular anaphora

Since this thesis focuses on the tradition of *dynamic semantics*, I will discuss the formal and conceptual aspects of this approach extensively. What may be called 'dynamic' about dynamic semantics is that it involves a notion of interpretation which contributes some kind of *change*. In dynamic semantic theories, the meaning of an expression is said to be its 'context-change potential.' Apart from a notion of context change, dynamic semantic theories have in common that pronouns are systematically analysed as ordinary bound variables. The distinction between 'types' of pronouns, such a e-type vs. bound, is a result of how these variables are evaluated. The notion of context standardly involves assignment functions and consequently different contexts result in different variable evaluations. Roughly, bound usages of pronouns are explained as variables which are evaluated with respect to iterated contexts; the iteration being instantiated by some operator (quantifier) in the scope of which the pronoun occurs. Referential usages of pronouns are variables evaluated in a context which provides a determined value for this variable. E-type pronouns are now to be explained as variable evaluations which, due to contextual circumstances, do not naturally fall under the previous two classes. The main point, however, is that from a dynamic semantic point of view, there is no useful distinction between *kinds* of pronouns, since all pronouns correspond to variables in context.

For instance, the case of inter-sentential anaphora in (2.6) is explained by considering the variable corresponding to *he* in a context which underdetermines the value for this variable.

(2.6) A man came in. He sighed.

The conditional in (2.7) is analysed as a universal quantifier over contexts satisfying the antecedent clause. (See Lewis 1975; Kratzer 1986.)

(2.7) If a farmer owns a donkey, he beats it.

(2.8) All cases in which x is a farmer, y a donkey owned by x, are cases in which x beats y.

The pronouns in the consequent clause can simply be analysed as bound variables, since they are evaluated in individual contexts in which x is a farmer owning donkey y.

Notice that this has immediate advantages over an e-type strategy. First, with respect to Heim's argument that a linguistic antecedent is a necessary ingredient for pronominal reference, notice that a dynamic approach has a semantic explanation for this need. In Heim's examples (2.1) and (2.2), repeated here as (2.9) and (2.11) respectively, the set of married men coincides with the set of men who have a wife, but the way these constituents change the context differs in an important way. In the paraphrase (2.12), there is no variable available for the pronoun 'her'.

(2.9) Every man who has a wife sits next to her.

(2.10) All cases in which x is a man such that there exists a y being x's wife is such that x sits next to y.

(2.11) *Every married men sits next to her.

(2.12) All cases in which x is a man such that x is married are such that x sits next to -?-.

Second, issues of uniqueness do not play a role in dynamic semantics. The example in (2.5) is repeated here with a dynamic paraphrase. Notice that whatever value is chosen for 'a man' will be the value for 'he', but that nothing is said about how many such values exist.

(2.13) If a man is in Athens, he is not in Rhodes.
All cases in which x is a man and x is in Athens, are cases in which x is not in Rhodes.

In the introduction to dynamic approaches to meaning below, I focus on two frameworks: Kamp (1981)'s discourse representation theory and Groenendijk and Stokhof (1991)'s dynamic predicate logic. This creates an oversimplified image of dynamic semantics. In what follows, many contributions and discussions are not discussed, such as, for instance, the seminal papers Heim 1982, Rooth 1987, Barwise 1987, Chierchia 1992 and Groenendijk and Stokhof 1990.

2.2.1 Discourse Representation Theory

Although of paramount importance to the emergence of dynamic semantics, discourse representation theory (henceforth, DRT) should not really be seen as a part of that tradition. DRT fits in the Geachean tradition of viewing pronouns as variables, but 'meaning' in DRT certainly does not correspond to context-change potential. As we will see, some elements of dynamics do exist, though.

The key feature of the discourse representation theory presented in Kamp 1981 is that all 'types' of pronouns, as well as indefinite NPs, can be accounted for by translating them as predicated plain variables. The specific effects responsible for the seemingly different types of pronouns are caused by interactions with operators introduced by other syntactic material. Likewise, the quantificational force of an indefinite is established by the context it appears in. The difference between an indefinite and a pronoun is due to nothing more than the fact that the former corresponds to a 'new' variable, while the latter corresponds to an 'old' one.

In DRT, given a syntactic form, the representation of the meaning of that form is generated by applying construction rules. For instance, the rule for pronouns says that a variable should be selected and used for predication (conform the rest of the sentence). By recursively applying such rules, sentences result in discourse representation structures or DRSs, which "can be regarded as the mental representations which speakers form in response to the verbal inputs they receive" (Kamp 1981, p. 282). Subsequent syntactic forms are incorporated in the existing representation and transformed in meaning representations by again applying construction rules. This shows an element of dynamics: syntactic forms are interpreted inside representations formed by previously processed material and they also set the stage for forms that follow.

DRSs are formal objects containing a set of variables or '(discourse) referents'[4] called the *universe* and a set of conditions. These structures can be interpreted. Again, this notion of interpretation is not dynamic.[5] Consider a model $M = \langle D_e, I \rangle$ and a set of referent symbols $\textsc{Var}$. A DRS is a pair $\langle V, C \rangle$, where $V \subseteq \textsc{Var}$ and C is a set of conditions. Another way of looking at these structures is by viewing them as a list of conditions paired with (some of) the free variables that occur in them. In this light, a DRS like the one in (2.14) is nothing more than the predicate logical formulae $P(x)$. Its interpretation (in (2.15)), however, tells us that it is *true* as soon as we find a value for x which satisfies P in the model.

(2.14)
$$\begin{array}{|c|}\hline \text{x} \\ \hline \text{P(x)} \\ \hline \end{array}$$

(2.15) $\exists f : \{x\} \rightarrow D_e$ such that $f(x) \in I(\mathrm{P})$

A DRS is true in DRT if and only if there exists a function assigning values to the free variables such that the conditions are satisfied in the model. As a result, free variables are standardly existentially quantified. Indefinite NPs, thus, need not be translated as existentially quantified expressions, but simply as variables. This allows them to adopt the quantificational force of whatever context they appear in. For instance, conditionals introduce embedded DRSs with a more complex interpretation.

(2.16)
$$\begin{array}{|c|}\hline \begin{array}{|c|}\hline \Upsilon \\ \hline \Gamma \\ \hline \end{array} \Rightarrow \begin{array}{|c|}\hline \Upsilon' \\ \hline \Gamma' \\ \hline \end{array} \\ \hline \end{array}$$

(2.17) All functions f that verify $\langle \Upsilon, \Gamma \rangle$ in M can be extended to a function which verifies $\langle \Upsilon', \Gamma' \rangle$ in M.

A function g extends another function f if the domain of g includes that of f and with respect to this joint domain, they agree on the values they provide. This is the source of many effects involving pronouns and indefinites. The embedded boxes introduced by a conditional are interpreted in a *local*

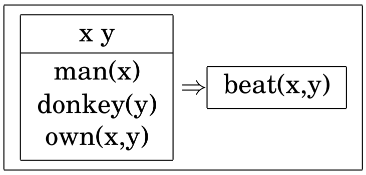
[4] In this introduction, we do not distinguish the notion of variable from the notion of discourse referent as this is not relevant for the purpose of introducing DRT. However, the reader should keep in mind that discourse referents are linguistically motivated objects, while variables can play many roles in natural language semantics, including some for purely derivational purposes. Groenendijk and Stokhof (1990), page 3, for example, point out that in the lambda abstract $\lambda x.\text{farmer}(x) \wedge \exists y(\text{donkey}(y) \wedge \text{own}(x, y))$, y, but not x can be considered a discourse referent (i.e. with the properties they have in DRT).

[5] The interpretation of conditionals (see below), however, *is* dynamic, since the consequent clause is interpreted with respect to an assignment function changed by the antecedent clause.

context. The consequent clause is iteratively interpreted with respect to whatever the function f is changed into by the antecedent clause. Consider for instance the DRS in (2.19) which represents the classic donkey example and its interpretation in (2.20).

(2.18) If a man owns a donkey, he beats it.

(2.19)

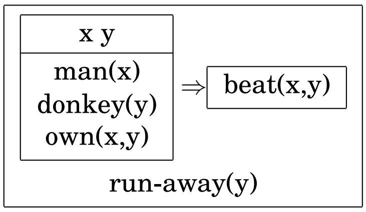


(2.20) All functions $f : \{x,y\} \rightarrow D_e$ such that $f(x)$ is a man and $f(y)$ is a donkey owned by $f(x)$ in model M extend to a function $f' : \{x,y\} \rightarrow D_e$ such that $f'(x)$ beats $f'(y)$.

(2.21) All functions $f : \{x,y\} \rightarrow D_e$ such that $f(x)$ is a man and $f(y)$ is a donkey owned by $f(x)$ in model M are such that $f(x)$ beats $f(y)$.

The simplification in (2.21) of (2.20) is due to the fact that, if two functions have the same domain, one of them can only extend the other if they are, in fact, the same function. The pronouns as well as the indefinites in (2.18) are simply represented as variables. The e-type effect is explained by the semantics of the conditional rather than by a reconstruction procedure.

A final important basic ingredient of DRT is its notion of accessibility. As stated above, the only difference between a pronoun and an indefinite is that the latter corresponds to a fresh variable while the former corresponds to a referent which has already been introduced. The construction rule for pronouns thus states that a pronoun should be replaced by an *accessible* variable. The notion of accessibility is not stipulated, but is semantically based: it follows from the truth-conditions for a DRS and is therefore very similar to the syntactic requirements of variable binding in predicate logic. For instance, from the interpretation in (2.17), it follows that referents introduced in the antecedent DRS of a conditional are accessible in the consequent DRS. This is so because the functions used to test verification of the right DRS are functions verifying the left DRS and, therefore, these functions will have as their domain the universe of the left DRS. In contrast, the referents in a conditional are *not* accessible from outside the two embedded DRSs. That is, the referent 'y' in the predication 'run-away(y)' in (2.22) is *free*. It cannot originate from a pronoun, since the referent 'y' introduced in the embedded DRS is inaccessible from within the main DRS. This is supported by the infelicity of (2.23).

(2.22) 

(2.23) If a man owns a donkey, he beats it. ??It runs away.

As I mentioned above in the introduction chapter, DRT was one of the causes of the rise of a programme of dynamic semantics in which interpretation is context-change potential. The specific representational architecture of DRT, however, begged the question whether a more direct expression of this potential was not possible.

2.2.2 Dynamic predicate logic

In Groenendijk and Stokhof 1991, a dynamic predicate logic is presented which fulfilled the goal of creating "a compositional, non-representational theory of discourse semantics." The dynamics of this system follows directly from the fact that predicate logical formulae are not interpreted in terms of truth-conditions, but in terms of how they change the context, or, to be precise, how they change assignment functions. Every formula denotes a set of pairs of assignment functions, i.e., it denotes a relation between assignments. Conceptually, these pairs should be seen as the inputs and corresponding outputs of the predicate logical 'instructions.' For instance, $\langle f, g\rangle$ is in the interpretation of Γ if g is a possible output for interpreting Γ with respect to f.

Crucial is the interpretation of conjunction as relation composition. That is, in interpreting a formula $p \wedge q$ the output of interpreting p is used as an input for the interpretation of q. Let $[\![]\!]$ map formulae to relations between assignment functions. We write $f\,[\![\varphi]\!]\,g$ to express that $\langle f, g\rangle$ is one of the pairs of assignment functions in the dynamic intepretation of φ.

(2.24) $f\,[\![p \wedge q]\!]\,g \;:\Leftrightarrow\; \exists h : f\,[\![p]\!]\,h \;\&\; h\,[\![q]\!]\,g$

Dynamic conjunction is thus *internally* dynamic: the change brought about by p is passed on to the interpretation of q. Moreover, it is *externally* dynamic: the change brought about by processing p and subsequently q can be recovered from the output assignment. The prime source of such 'changes' is the dynamic existential quantifier. It uses the notion of *random assignment*.

(2.25) $f\,[\![\exists x(\varphi)]\!]\,g \;:\Leftrightarrow\; \exists h :$ f and h differ at most in the value they assign to x and $h\,[\![\varphi]\!]\,g$

The existential quantifier replaces whatever value the input assignment assigns to x with some value that satisfies its scope φ. Combining the definition of the existential quantifier with relation composition as the interpretation for conjunction, we are able to derive the so-called donkey equivalence in (2.26), which says that existentially introduced values have unlimited scope to the right.

(2.26) $f\,[\![\exists x(\varphi) \wedge \psi]\!]\,g \;\Leftrightarrow\; f\,[\![\exists x(\varphi \wedge \psi)]\!]\,g$

In fact, this points out that it is unnecessary to define the existential quantifier with a scopal formula. We could alternatively define just the bare quantifier:

(2.27) $f\,[\![\exists' x]\!]\,g \;:\Leftrightarrow\;$ f and g differ at most in the value they assign to x

Existential quantification changes the context in an important way. Other atomic formulae have no such influence on the context and simply test whether the context is to their satisfaction. Such expressions are accordingly called *tests*. A predication $P(x)$, for instance, returns an incoming assignment whenever this assigns a value to x which is in the extension of P.

(2.28) $f\,[\![P(x_1,\ldots,x_n)]\!]\,g \;:\Leftrightarrow\; f = g \;\&\; \langle f(x_1),\ldots,f(x_n)\rangle \in I(P)$

Dynamic implication is also a test, although a much more complex one. Similar to DRT, conditionals are analysed as introducing universal quantification over contexts. In DPL's terms, this means that given an input assignment f, we first interpret the antecedent clause and for each output assignment this interpretation returns, we check whether we can use that as an input for interpreting the consequent clause successfully (i.e., returning some output assignment function). If this test succeeds, we return the input assignment again.

(2.29) $f\,[\![p \Rightarrow q]\!]\,g \;:\Leftrightarrow\; f = g \;\&\; \forall h : f\,[\![p]\!]\,h \rightarrow \exists k : h\,[\![q]\!]\,k$

All this gives us the tools to analyse a number of ordinarily problematic pronouns as simple (bound) variables. For instance, the semantics of intersentential anaphora in (2.30), represented in (2.31), follows directly from the donkey equivalence in (2.26). The classic donkey conditional in (2.32) can be analysed parallel to the proposal in DRT, except that the indefinite results in an existential quantification as opposed to a predication over a free variable. The effect from (2.33) is the same.

(2.30) A man came in. He sighed.

(2.31)
$$\begin{aligned}
&[\![\exists x(\text{MAN}(x) \wedge \text{CAME_IN}(x)) \wedge \text{SIGH}(x)]\!] =\\
&[\![\exists x(\text{MAN}(x) \wedge \text{CAME_IN}(x) \wedge \text{SIGH}(x))]\!] =\\
&[\![\exists' x \wedge \text{MAN}(x) \wedge \text{CAME_IN}(x) \wedge \text{SIGH}(x)]\!]
\end{aligned}$$

(2.32) If a man owns a donkey, he beats it.

(2.33) $[\![(\exists' x \wedge \text{MAN}(x) \wedge \exists' y \wedge \text{DONKEY}(y) \wedge \text{OWN}(x,y)) \Rightarrow (\text{BEAT}(x,y))]\!]$

The *accessibility* of antecedents follows from the semantics. Consider for instance (2.33). In (2.23), we saw that a conditional like (2.32) does not allow subsequent anaphoric reference to indefinites introduced in one of its clauses. The formula in (2.33) is interpreted as a test. Given an assignment function f, it returns f if and only if all possible output functions for the left-hand side of the dynamic implicator '$\Rightarrow$' yield an output when used as an input function for the interpretation of the right-hand side of the implicator. But when these conditions are met the input assignment (f) is returned and consequently all random assignments considered during the test are lost.

Negation is a test as well. It blocks access from outside its scope to variables introduced within its scope. The negation operation checks whether there exists an assignment function which can serve as a proper output for its scope given some input assignment. If so, the test fails. If not, the input assignment is returned.

(2.34) $f\,[\![\neg(\varphi)]\!]\,g :\Leftrightarrow\ f = g\ \&\ \neg\exists h : f\,[\![\varphi]\!]\,h$

Changes brought about to f by φ can thus never surface outside the scope of the negation. Therefore, a variable x occurring after $\neg(\exists x)$ can never access the value assigned to x inside the negation. This explains why negated indefinites are not proper antecedents.[6]

Something similar occurs with universal quantification.

(2.35) $f\,[\![\forall x(\varphi)]\!]\,g\ :\Leftrightarrow\ f = g\ \&\ \forall k : f\,[\![\exists' x]\!]\,k \Rightarrow \exists j : k\,[\![\varphi]\!]\,j$

This says that any extension of f assigning some value to x can successfully interpret φ. The output assignments j of these successful interpretations, however, are not passed on.

This way, the semantics of DPL mirrors the anaphora facts of natural language. While existential quantification is externally dynamic, universal quantification is externally static. Like the semantics of negation, the semantics of universal quantification does not pass on any values introduced in its scope. Both conjunction and predication are also externally dynamic, the latter trivially so.

2.2.3 Information increase

The relations expressed by DPL formulae carry information about the content of the formulae. For instance, the set of assignments g paired with some assignment f according to such a relation, inform us about the possible values for the variables. They inform us about which values are still

[6]It has often been pointed out that a definition as in (2.34) constrains the accessibility of negated antecedents too strongly. See, e.g. Krahmer and Muskens 1995 and van Rooy 1997for discussion.

open to discussion. Accordingly, the predicate logical expressions in DPL can be seen as *actions* on information states. Predication, being a test, has the capability of reducing the number of possible values for variables. For instance, after introducing 'a man' in the discourse using the action '$\exists x(\text{MAN}(x))$', the possible values we wish to entertain for 'x' are all the objects satisfying the property of being a man. If, somewhere further on in the discourse, an utterance instantiates the action '$\text{OLD}(x)$', the possible values for 'x' are reduced to objects which are old men. This way, predication reduces the number of options that are open and thus increases the information stored in context.

Given a set of assignment functions F and a dynamic predicate logical expression φ, we are able to express the update potential of this expression. That is, given a set of possible value assignments that is consistent with the foregoing discourse, we are able to express how a form like φ influences this set.

Definition 2.1 Update potential

$$F[\varphi] \;:=\; \lambda g.\exists f \in F:\; f\,[\![\varphi]\!]\,g$$

□

Ideally, only three possibilities are open for formulae of a dynamic formalism: (i) φ accepts the possibilities in F, i.e. $F[\varphi] = F$, (ii) φ adds information, i.e. $F[\varphi] \subset F$ or (iii) φ is inconsistent with respect to F, i.e. $F[\varphi] = \emptyset$. In such a formalism, no action can cause a loss of information: for any formula φ it holds that $\forall F:\; F[\varphi] \subseteq F$. For DPL, however, this does not hold.[7] For example, say that some set of assignments F contains only functions assigning some entity d to x. Clearly, however, $F[\exists' x] \nsubseteq F$, for the existential quantifier *resets* the variable x.

Intuitively, this is related to the fact that apart from reduction of the set of possible assignment functions, there seems to be a second form of information increase associated with quite a different notion of information. During the processing of a discourse, not only do we constantly reduce the number of options open as a value for some variable, we moreover keep track of which variables are under discussion. By introducing new topics we expand the information we have about the discourse.

In DRT, this type of information is represented by the universe of a DRS. Each NP that is encountered introduces a fresh referent in the universe. As the discourse unfolds, more and more referents will be introduced, increasing the potential for anaphoric reference. The contexts of

[7]This is in contrast to the update semantics presented in Veltman 1991, which does have this property. See Vermeulen 1994 (also 1993) for an elaborate discussion of this so-called eliminativity property and an elegant variation on DPL's variable management which guarantees eliminativity. Other proposals can be found in Fernando 1992; Dekker 1993; Dekker 1994 and van Eijck 2001.

DPL, however, are assignment functions that assign values to all the variables in their domain and as such do not have a way of discerning between *active* and *inactive* variables. Moreover, the notion of context of DPL is not able to represent something like *an initial state* wherein an item has yet to be introduced in the discourse.[8]

An intuitively attractive way of enriching DPL with information concerning the introduced discourse topic is by using partial (instead of total) assignment functions. This way, the active domain of a function expresses which variables are under discussion, just like the universe of a DRS supplies the domain of the verifying assignment function in DRT. The initial state, in such a set-up, will be the function that is undefined for every variable.

As we will see below, particularly in chapters 4 and 5, we will need to model both kinds of information discussed here. There, we return to the issue of defining DPL with partial functions.

2.3 DRT and plural e-type pronouns

The fourth chapter of Kamp and Reyle 1993 is doubtlessly the first analysis of anaphoric aspects of plurality which is detailed enough to make serious empirical claims about virtually all of the numerous phenomena involving plural reference. Although the framework is DRT, the analysis of plural e-type pronouns presented in this work very much resembles an *e-type* strategy.

Kamp and Reyle distinguish between two types of NPs, namely quantificational NPs and non-quantificational NPs. The latter group is treated like ordinary indefinites. That is, they introduce a (possibly plural) referent in the local universe which is accessible in the local DRS as well as in embedded DRSs. For instance, (2.36) is represented as (2.37).

(2.36) Two students wrote an article.

(2.37)

X y
|X|=2 student*(X) article(y) wrote(X,y)

Uppercase letters are used for referents which correspond to plural indi-

[8]This is not entirely true. For DPL one could take as an initial state the set of all possible assignment functions. In such a state, each variable is associated with every possible value in the domain of entities. The state, therefore, represents a tabula rasa, since all options are still open. However, this is at the cost of representing all non-introduced topics as topics about which nothing is known.

viduals, lower case letters have an atomicity required implicit in them[9]. The conditions in the DRS say that the referent 'X' should refer to a plurality containing two atoms which occurs in the plural closure of the set of students. Furthermore, 'y' corresponds to an article written by these students. Subsequent plural pronouns are able to pick up these two referents.

(2.38) Two students wrote an article. They sent it to L&P.

(2.39)

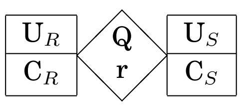
X y
|X|=2
student*(X)
article(y)
wrote(X,y)
senttoL&P(X,y)

There is no maximality involved in the case of anaphora in (2.38). There might have been other students writing a paper, but they need not have sent a paper to L&P to verify (2.38). This is indeed allowed by the DRS in (2.39), since *any* function f assigning a set of two students to X that wrote some article $f(y)$ and sent it to L&P verifies this DRS.

The other class of NPs, quantificational NPs, does not introduce referents. They introduce a type of representation called *duplex conditions*. These representations follow the common wisdom that (quantificational) determiners create a tripartite structure. Duplex conditions take care of the truth-conditions of the structures following the techniques of generalised quantifier theory(Barwise and Cooper 1981) . At the same time, they account for possible anaphoric links between the arguments of the determiner. Let us briefly consider how this is done.

Ignoring some details, duplex conditions consist of two DRSs, one corresponding to the restrictor argument of the determiner and the other to the nuclear scope. These boxes are connected by a symbol expressing the quantificational relation corresponding to the determiner as well as the referent(s) quantified over.

(2.40)



The referents introduced by the restrictor ([U_R|C_R]) are accessible in the scope ([U_S|C_S]). The interpretation is not straightforward. I will ignore several important issues here and aim at communicating the general intuition behind the interpretation.[10] Say that Q corresponds to some relation

[9]Actually, the system of plural and singular referents presented in Kamp and Reyle 1993 is much more refined, but we will ignore these subtle (though important) details.

[10]See especially subsection 4.3.2.1 of Kamp and Reyle 1993 for discussion of some details.

between sets $I(Q)$, then a model and an embedding function f verify (2.40) if the following sets A and B are in the $I(\mathbf{Q})$ relation:[11]

(2.41) A: The set of individuals d, such that there is an extension g of $f \cup \{\langle r, d\rangle\}$ which verifies C_R.

(2.42) B: The set of individuals d, such that there is an extension g of $f \cup \{\langle r, d\rangle\}$ which verifies C_R, for which in turn there exists an extension h which verifies C_S.

The accessibility facts follow from the interpretation. Referents in the scope are interpreted relative to an extension of an embedding function which verifies the restrictor. Referents in the restrictor are thus accessible to the scope. Outside the duplex conditions no referents introduced in it are accessible. This way, DRT needs to use an independent procedure to account for non-bound pronouns with quantified antecedents.

2.3.1 Abstraction

In DRT, the so-called abstraction procedure makes it possible to reconstruct a description using representation material of the antecedent. Plural e-type pronouns can be analysed as variables equated with such a description. This is made possible by a summation operator 'Σ', which creates the plural individual maximally satisfying the description. For instance, the abstraction in (2.43) corresponds to the maximal plurality of farmers owning a donkey.

(2.43) Σx.

x y
farmer(x) donkey(y) own(x,y)

The operator can be used to abstract information from a duplex condition. That is, following a duplex condition as in (2.44), we want the individual in (2.43) to be a potential antecedent for subsequent (e-type) plural anaphors. It stands for the farmers that own a donkey.

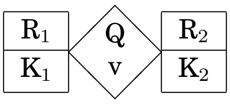
(2.44)

x	⟨Most x⟩	y
farmer(x)		donkey(y) own(x,y)

A construction rule regulated when abstraction can be applied. It is a special kind of construction rule since the "application conditions relate to the DR-theoretical structure of the DRS, not to the syntactic form of particular reducible DRS-conditions" (Kamp and Reyle 1993, p. 344). That is, the

[11] To be slightly more precise, the following conditions hold for such duplex conditions: $r \in \mathrm{U}_R$ and $\mathrm{DOM}(f) \cap (\mathrm{U}_R \cup \mathrm{U}_S) = \emptyset$.

occurrence of a duplex condition in a DRS is enough to trigger application of the procedure. In fact, Kamp and Reyle suggest that Abstraction could be seen as a kind of "inference principle on DRSs" Kamp and Reyle 1993, p. 344). According to the construction rule, given a duplex condition as in (2.45), we may perform the operations specified in (2.46).

(2.45) 

(2.46) Form the union $K_0=K_1\cup K_2$ of the two component DRSs of this condition. Choose a discourse referent **w** from $R_1\cup R_2$. Introduce into the universe of the DRS in which the duplex condition occurs a new discourse referent **Y** and add to its set of conditions:

$$\mathbf{Y} = \Sigma\mathbf{w}.K_0$$

According to this rule, we can extend any DRS containing a duplex condition to one containing an abstraction over the restrictor and scope of the quantification that introduced the duplex condition.

Notice that abstraction makes use of the representational nature of DRT. It is triggered not by a linguistic quantificational structure, but rather by the semantic representation of such a structure. This is necessary since the abstraction procedure comes on top of the ordinary interpretation of quantification, which needs to represent the constituents of the structure piece by piece. Only after the whole quantification is represented can the formation of antecedents start. The duplex condition specifies exactly which referents can be abstracted over, namely those in $R_1\cup R_2$ with respect to $K_1\cup K_2$.

The representational nature of abstraction becomes especially clear if we try to incorporate abstraction in DPL. Of course, it is perfectly possible to find a way in which a formula '$X=\Sigma x.\varphi$' is meaningful in an extended version of DPL. It would have to express the identity relation $\langle f, f\rangle$ if and only if $f(X)$ equals the set of d's such that if we interpret φ with respect to $f \cup \{\langle x, d\rangle\}$ we successfully find some output function.

The question, however, is where φ comes from. The conditions it contains will, of course, not be retrievable from the input function. They will have to be copies of conditions which occurred in a formula representing an antecedent quantificational structure. For instance the example in (2.47) could be successfully modelled in our neo-DPL as (2.48), but this would be completely counter to the compositional ideology of DPL, since the anaphoric reference Y is retrieved not by context change, but through a copy instruction on formulae.[12]

[12]Moreover, the fact that the abstraction procedure is in essence a *copy* instruction illustrates once more the strong resemblance with the e-type strategy tradition.

(2.47) Every student wrote a paper. They worked very hard.

(2.48) $\forall x[[\text{STUDENT}(x)] \Rightarrow [\exists y[\text{PAPER}(y) \wedge \text{WROTE}(x, y)]]]\wedge$
$\exists X[X = \Sigma x.[\text{STUDENT}(x) \wedge \exists y[\text{PAPER}(y) \wedge \text{WROTE}(x, y)]]] \wedge$
$\text{WORKED\_VERY\_HARD}(X)$

For now, it suffices to conclude that there is an interesting tension between the dynamic semantic ideal and DRT's treatment of maximal plural anaphora.

2.3.2 Domain information

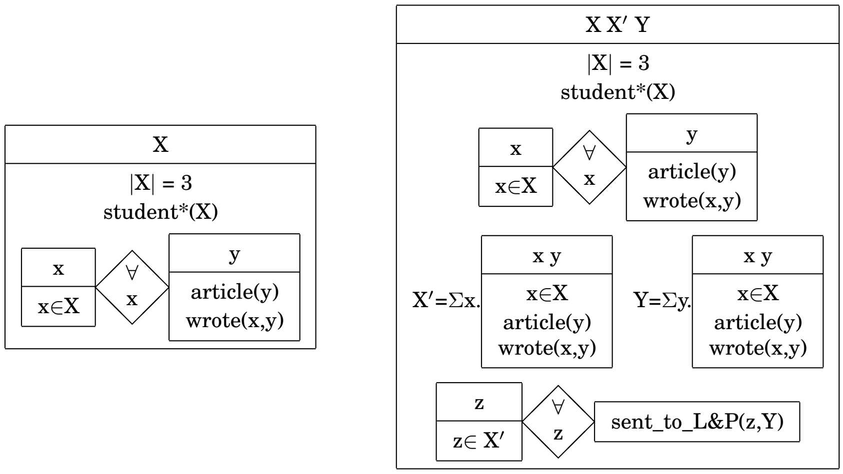
Turning now to more complex kinds of anaphora that involve pluralities, namely dependent interpretations of pronouns, it becomes clear that the abstraction procedure is not enough. Consider (2.49).

(2.49) Three students wrote an article.
They sent it to L&P. Krifka 1996a (ex. 1)

The first sentence of this example can have a distributive reading, wherein three students each wrote a (different) article. Given that reading, the second sentence is to be interpreted as meaning that the students each sent *their own article* to Linguistics and Philosophy. The puzzle this example presents is how the singular pronoun accesses the individual papers when the only procedure for antecedent formation involves the summation of values for the referents in the first sentence. That is, given the representation of the first sentence in (2.49) in figure 2.1a, the second sentence will result in a quantification over the abstracted referent 'X′' (the students that wrote an article),[13], as in figure 2.1b. However, the pronoun 'it' can only access the abstracted referent 'Y' (the papers written by the three students) which leads to the undesirable interpretation that the students did not only send their own article but also those written by the other two students. Kamp and Reyle conclude from a similar example that "[w]hen a set is introduced via Abstraction over some duplex condition δ, then the information contained in the constituent DRSs of δ is available as information concerning the members of that set. This means that when we distribute over such a set, the DRS occurring on the right-hand side of the Abstraction equation may be "copied" into the left-hand DRS of the duplex condition which the distribution introduces" (Kamp and Reyle 1993, p. 379).

In order to derive the intended interpretation of (2.49), we have to execute the "copy" instruction proposed by Kamp and Reyle. Next to the DRS in (2.1b), which we can discard as an interpretation for (2.49) due to the mismatch between the plurality of the group of papers ('Y') and the singular number feature of 'it', we come to a second DRS, one in which the

[13] Or, alternatively, it can access the referent 'X', which, in this case, describes the same group of students.



a. DRS for three students wrote an article

b. potential DRS for (2.49) (ignoring number agreement)

Figure 2.1: Abstraction applied to example (2.49)

individual papers are accessible. This DRS is given in figure 2.2. Quantifying over 'X′' allowed the copying of the descriptive material (i.e. the set of conditions) that was abstracted from the 'antecedent' duplex condition.

Krifka criticises this last move as being unmotivated: "[T]he anaphoric phenomenon of box copying is treated in a quite different and strikingly informal way. This stands in sharp contrast to the narrowly defined and well-motivated constraints for the accessibility of discourse entities that represent standard anaphora" Krifka 1996a (p. 561).

I agree with Krifka's criticism and wish to add that DRT's treatment of dependent pronouns is completely dependent on the use of representations. The copy-instruction under discussion here accesses individuals depending on the members of some group antecedent by accessing the representation responsible for describing the group. From our point of view, we prefer to have these individuals accessible in the form of semantic objects (as in the range of assignment functions). As we saw above with our attempt to integrate an abstraction procedure in a dynamic predicate logic, an extension of a semantic analysis using DPL with a treatment of dependent pronouns in the style of Kamp and Reyle would not comply with DPL's compositional roots. The accessibility of individuals that are (indirectly) involved in group formation when quantifying over such a group would not be due to context change potential, but due to a copy-procedure on formulae.

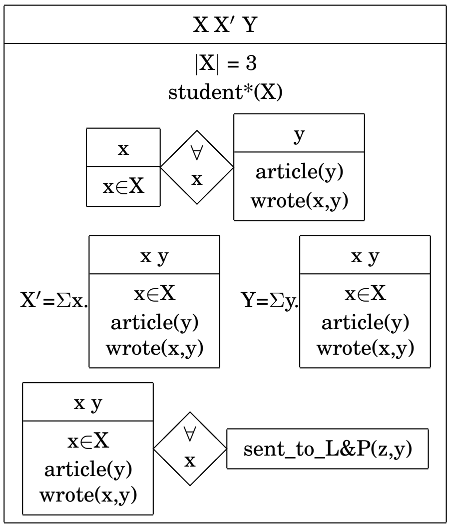


Figure 2.2: Quantification over an inferred domain

2.4 Discussion

The abstraction procedure makes it clear that after processing a quantificational sentence, not only a duplex condition taking care of the internal dynamics as well as of the truth-conditions of the quantification is inserted in the DRSs, but also an abstraction over the material in the conditions of the duplex condition. Under normal circumstances (i.e. when the quantificational structure is not embedded in a negative context), this means that the maximal individual satisfying restrictor and scope is accessible in the subsequent discourse. One might wonder, however, why this individual is introduced in such an indirect way. It is difficult to decide whether the strategy employed by DRT is an e-type strategy or whether it belongs to the more dynamic tradition. It is dynamic in the sense that plural e-type pronouns are still simple variables (equated with an abstracted individual). However, the anaphoric effect these cases of anaphora display are not due to the linguistic context they appear in, but rather due to a stipulated inference principle on a specific representational form. Nevertheless, it is hard to criticise the abstraction procedures on these grounds. The predictions it makes seem to be correct. We could possibly only accuse Kamp and Reyle of a lack of elegance. Still, there are some cases which show weaknesses in the abstraction procedure.

In this section, I discuss the merits of DRT, as well as some questions about the nature of its anaphoric mechanisms.

2.4.1 Merits of DRT's analysis

The empirical coverage of the abstraction procedure (and the other inference principle of box-copying) goes beyond merely accounting for Evans' remarks on e-type pronouns and maximality. First of all, quantificational noun phrases are *distributive*, in contrast to non-quantificational noun phrases.[14]

(2.50) Exactly four students wrote a paper. dist/*coll

(2.51) Four students wrote a paper. dist/coll

This is predicted by duplex conditions, which counts the possible *atomic* extensions to an embedding function. In other words, only assignment functions that are extended with an assignment to an atomic individual are considered. The same goes for the abstraction procedure. In 'X=Σ x.K', no groups that possibly satisfy K are considered, due to the atomicity restriction implicit in the small letter 'x'.

A second success of the DRT treatment of quantificational noun phrases is the maximality of the abstraction procedure. By simply collecting the successful values, what is recovered is not just any set of values which would have verified the duplex condition, but a maximalised one. The distinction above between distributive quantificational noun phrases and the potentially collective referential ones is relevant with respect to maximality as well. Indefinites and bare numerals, which do not introduce duplex conditions and therefore do not depend on the abstraction procedure for the introduction of antecedents, license non-maximal discourse anaphora. We can assure ourself of this fact by applying a test taken from Szabolcsi 1997.

(2.52) A few students wrote a paper.
Perhaps there were other students who did the same.

(2.53) Two students wrote a paper.
Perhaps there were other students who did the same.

[14] Although the intuitions for examples as (2.50) are clear, it remains a simplification to call quantificational noun phrases *distributive*. For instance, both (i) and (ii) below, can clearly be said to be collective.

(i) Exactly four students wrote this paper.
(ii) Exactly forty students gathered in the square yesterday.

There are many complicating subtleties. For instance, (iii) shows that not all quantificational noun phrases behave in the same way. (That is, the example cannot mean that a group of students collaborated in writing the paper and that this group forms a majority.)

(iii) ??Most students wrote this paper.

Since –apart from the complications involving collective predication– the group of quantificational noun phrases behaves the same in many respects, I will ignore this problem and assume that QNPs are, in essence at least, distributive.

(2.54) More than two students wrote a paper.
Perhaps there were other students who did the same.

(2.55) Most students wrote a paper.
Perhaps there were other students who did the same.

Whenever anaphora is maximal, there is nothing 'others' can refer to. Thus, (2.53) shows that anaphoric relations with 'two students' involve two students no matter whether there were actually more than two students that wrote a paper. With quantificational determiners, like 'more than two' and 'most', pronouns will have to pick up the abstracted referent, which is exhaustive.

A third success of DRT's treatment of quantification has to do with accessibility. Since the abstraction procedure can only be applied once its trigger, a duplex condition, is formed, it is predicted that the type of maximal anaphora for which the abstraction forms an antecedent will always occur after the quantificational structure is completed. That is, quantificational noun phrases do not display referential effects within the sentence level. This is supported by the data in (2.56) and (2.57).

(2.56) Two lawyers (each) hired a secretary they interviewed.

(2.57) Most lawyers hired a secretary they interviewed.

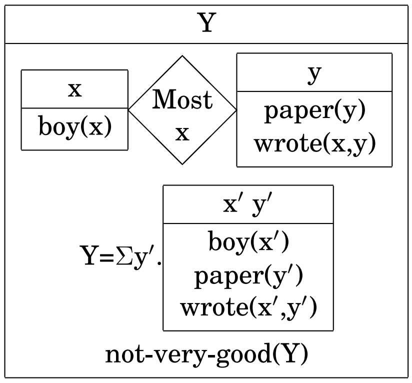
The example in (2.56) is ambiguous between a reading in which each of the two lawyers hired a secretary he or she interviewed and one in which they each hired a secretary that was interviewed by them both. The ambiguity involves the possibility of the plural pronoun 'they' to be construed as a collective subject for 'interview'. In other words, in the first reading, the secretary is interviewed by the lawyer that hired him and in the second, he is interviewed by the two lawyers. In contrast, (2.57) lacks this latter reading, since no plural antecedent has been introduced yet by the subject. For instance, (2.57) cannot mean that a majority of lawyers each hired a (different) secretary they collectively interviewed.[15]

A fourth virtue of DRT is the fact that abstraction is free to abstract over any referent in the universe of either sub-DRS of the duplex condition. This way, indefinite noun phrases introduced somewhere in the quantificational structure trigger exhaustive reference in discourse. For instance, in (2.58) a plural pronoun accesses the set of papers written by the students. The DRT account is in (2.59).

(2.58) Most students wrote a paper. They weren't very good.

[15]Note that, strikingly, the paraphrase given here *does* have this reading, since I replaced the QNP 'most lawyers' with 'a majority of lawyers.'

(2.59)



In sum, we have discussed four successes of the duplex condition and abstraction approach to plural e-type pronouns. We now turn to two of the less clearly advantageous sides of the proposal.

2.4.2 Some questions

2.4.2.1 Entailment, negativity and emptiness

Let us turn to a particular aspect of entailment patterns between examples with downward-entailing quantifiers. These quantifiers license anaphora just like upward-entailing quantifiers do, even though they do not assure us of the existence of anything satisfying both restrictor and scope.

(2.60) Few students went to the party.
$\not\Rightarrow$
Some student went to the party.

With the abstraction procedure, DRT seems to acknowledge that the possibility for anaphora is not influenced by the choice of the determiner. However, by automatically constructing the antecedent as soon as a duplex condition is inserted in a DRS, it is in danger of ignoring the lack of entailment in (2.60). The reason is that a condition '$X = \Sigma x.$K' is interpreted as the condition that there should exist a function which assigns to X a value all the atoms satisfy the condition K. Following downward entailing quantifiers, this could be the empty set. The functions we are considering should therefore include functions assigning the empty set to a referent for if we were to exclude such sets, we would not do justice to (2.60).

What are the effects of allowing empty sets as values for referents?[16] Allowing for the empty set is not straightforwardly harmless. For instance,

[16] Actually, DRT does not talk about sets at all, but about the elements of a free complete atomic upper semilattice with a zero element. Here, we confuse the elements in such structures with sets, which does not have any serious consequences for our purposes given the isomorphism between the structures used by DRT and the set of subsets of the domain of atomic entities together with the subset relation.

a DRS like the one in (2.61) becomes a tautology, since we may assume the empty set not to have any lexical property P.

(2.61)
$$\begin{array}{|c|}\hline \text{X} \\ \hline \neg\ \boxed{\text{P(X)}} \\ \hline \end{array}$$

Nevertheless, it is questionable whether this causes problems for the interpretation of natural language expressions. In general, it seems to me that the restriction of a quantification will already exclude the empty set as a value. That is, there is no natural language paraphrase for (2.61), since there is nothing in the DRS that can play the role of a restrictor. For instance, 'something does not have property P' or 'nothing has property P' both quantify over 'things' and since the empty set is not a thing it does not verify these sentences. Moreover, quantifier domains are contextually restricted and we may assume these restrictions to exclude empty sets.

There might be more danger in an example like: 'Only women do not like Kylie', which would be analysed as 'for all x such that x does not like Kylie, it holds that x is a woman' and thus seems to be immediately falsified by the empty set since it neither likes Kylie, nor is a woman. But here, too, there is a restriction, namely the set of alternatives (e.g. men, women, martians) brought about by focus on 'women', which again excludes the empty set. The empty set, then, seems to be an element in the ontology which never surfaces in linguistic meanings.

There is another assumption which is necessary and relevant to empty sets. Downward entailing quantifiers are to be represented by duplex conditions and not simply by the introduction (and existential closure) of a referent. This is because if we were to represent them as ordinary existential constructions, the predications involved would exclude the empty set as a possible value for the referent involved, thus strengthening the truth-conditions. Here is an example. Say we represent 'Less than three players are holding a card' as the following DRS:

(2.62)
$$\begin{array}{|c|}\hline \text{X} \\ \hline \text{player*(X)} \\ |\text{X}|<3 \\ \text{card(y)} \\ \text{hold(X,y)} \\ \hline \end{array}$$

Since the empty set is not a player, this DRS is only verified by two players or a single player holding a card. The DRS is falsified by a situation in which no players is holding a card. As we will see in chapter 6, the semantics I propose makes an assumption similar to the one described here: downward entailing NP are to be interpreted as quantificational structures, not as predicational structures over some introduced (potentially empty) group of individual.

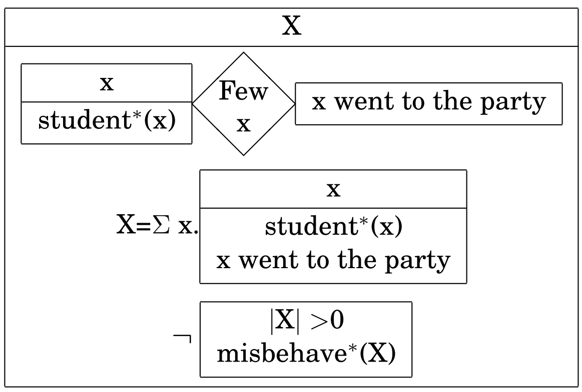
Let us focus some more on the burden of the pronoun. The pattern in (2.63) suggests that with pronominal reference there comes some non-emptiness condition on its referent. What the contrast between (2.60) and (2.63) shows is that this condition is due to the pronoun rather than due to the duplex condition or the subsequent abstractions. Pronouns, then, trigger an extra condition on abstracted referents, namely that they are non-empty.

(2.63) Few students went to the party. They had a good time.
$\Rightarrow$ Some students went to the party.

Thus, in order to do justice to both (2.63) and (2.60), two assumptions have to be made. The first assumption is that abstraction is able to abstract the empty set, that is, that empty sets are allowed in the range of assignment functions. Second, it is necessary to have the pronoun trigger a non-emptiness condition on its reference. Against this latter assumption, however, a more philosophical complication comes to mind. The non-emptiness condition is semantic in nature, that is, it can only be a predication over a referent. This means that there is no distinction between a false predication over an abstracted referent and an empty referent picked up by a pronoun. From the point of view of accounting for (2.63), this is a good thing, since what we see there is obviously a *semantic* fact. However, (2.64) and its supposed representation (2.65) show that the condition triggered by a pronoun is not a simple predicational condition.

(2.64) Few students went to the party. And it is not the case that they misbehaved.

(2.65)



The representation in (2.65) is wrongly verified if no students went to the party. Clearly, the condition that comes with a pronoun is presuppositional. This is confirmed by an example like (2.66), where the condition seems to bind into the antecedent clause, causing the paraphrase in (2.67).

(2.66) If less than ten students pass the test, the teacher will take them out for dinner.

(2.67) If less than ten but more than one student passes the test, the teacher will take them out for dinner.

Similar issues play a role if we consider another alternative understanding of DRT's abstraction procedure, namely one where it is seen as a run-of-the-mill e-type strategy by making the pronoun the crucial factor in the triggering configuration, instead of the quantificational sentence. This means that a plural pronoun introduces a variable equated with a description formed from material from an antecedent duplex condition. This description is taken to correspond to a non-empty set (i.e. the empty set is not allowed in the range of assignment functions). Here, it is the site of insertion of the abstraction equation which becomes crucial. Like the non-emptiness condition in (2.65), the abstraction equation should end up in a higher DRS.

It is clear that with respect to downward entailing quantification, many more details concerning the application of the abstraction procedure have to be clarified.

2.4.2.2 Accessibility of other sets

As we saw above, given a quantificational sentence $Dx(A)(B)$, the abstraction procedure correctly predicts that the reference set, $(\lambda x.A) \cap (\lambda x.B)$, is not the only set that is accessible after processing a quantificational sentence. Sets depending on the reference set, such as $(\lambda y.A) \cap (\lambda y.B)$, are also accessible using abstraction. DRT also makes the strong prediction that the reference set is the *only* set related to the ranging variable (here, x) of the antecedent quantifier that is accessible. That is, Kamp and Reyle predict the unavailability of the maximal set, $\lambda x.A$, and the complement set, $(\lambda x.A) - (\lambda x.B)$. Recall from chapter one, however, that cases of pronominal reference to the maximal set and to the complement set are both reported.

There does exist a procedure in DRT which comes close to an analysis of maxset reference, namely *kind introduction*. Kamp and Reyle's motivation for this procedure is closely related to the phenomenon of maximal set reference. Consider the following example (Kamp and Reyle 1993, p. 391):

(2.68) Few women from this village came to the feminist rally.
No wonder.
They don't like political rallies very much.

The plural pronoun here does not refer to the few women from the village that came to the feminist rally. Instead, it seems to generalise over *all* women from the village (or maybe even over all women in the world). Kamp and Reyle choose to treat this phenomenon as a general option for the plural pronominal reference and introduce the procedure of *kind introduction*, which given some "noun establishes a discourse referent for the genus within the universe of the main DRS" Kamp and Reyle 1993 (p.392).

Kamp and Reyle correctly remark that a genus is not simply a set. Rather than a case of reference to the maximal set, they would argue that the anaphoric phenomenon in (2.68) is independent of reference to the maximal set. However, most examples do not involve generic reference at all. Consider, (2.69).

(2.69) Most marbles in this bag are red. But exactly three of them are black.

The preferred interpretation for the second sentence of (2.69) is that three of the marbles in the bag are black. The genus 'marbles in the bag,' whatever this may mean, does not seem to be involved. Should the treatment of genera in DRT be worked out in more detail, however, so that is allows for examples like (2.69), the question becomes how kind introduction is to be restricted. As we will see in more detail in chapter 3, the maximal set is *not* generally accessible in the case of weak quantifier. Since kind introduction is a general principle working on representations, there does not seem to be a way of restricting it to not operate on structures that are due to those kind of noun phrases.

Turning now to reference to the complement set, Kamp and Reyle explicitly mention that such type of pronominal reference is predicted *not* to exist. They explicitly argue that set-subtraction is not an operation that could be involved in antecedent formation. Again, however, we have seen that cases of pronominal reference to the complement set seem to exist. If Kamp and Reyle are right about the set-subtraction not being one of the tools involved in anaphora, then we should at least be able to explain the existence of such cases (away). In the next chapter, we will turn to an evaluation of the kinds of anaphora Kamp and Reyle do not consider.

2.4.3 Conclusion

Abstraction can account for many of the facts. Yet, as the case of downward entailing quantifiers and reference to sets like the maximal set and the complement set show, a several details of what exactly the relation between pronominal reference and abstraction is remain unclear. Moreover, as Krifka (1996a)'s criticism shows, the procedures that control abstraction are in many ways unmotivated. The operation is ad hoc, since it does not treat plural e-type anaphora as a phenomenon which is derivable from accessibility facts in a straightforward way. This in contrast to singular e-type anaphora. In general, since accessibility is not a purely semantic notion, but dependent on representations, it follows that DRT's analysis of plurality is not compatible with a non-representational compositional enterprise in a straightforward way.

Criticism of the abstraction procedure, however, will always have to deal with the massive empirical coverage of chapter 4 of Kamp and Reyle

1993. DRT's success, however, is only due to the fact that the overgenerating tool of set-abstraction can only be operated by a small set of independently motivated rules.

Chapter 3

Pronominal Reference to Sets

Parts of this chapter appeared in the *Journal of Semantics* 20.1 (Nouwen 2003),

3.1 Discourse and sets

Recall from the introduction that discourse pronouns have the ability to refer to not just one but several sets associated with a quantificational structure. Apart from the reference set, that is, the set of MPs attending the meeting in (3.1), reference to the complement set (the MPs not attending the meeting) and the maximal set (the MPs) also present themselves as potential antecedents for future plural pronouns.

(3.1) Few MPs attended the meeting. $\text{Few}(A, B)$

a. They decided not to discuss anything important. $A \cap B$

b. They stayed home instead. $A - B$

c. But they all attended the drinks afterwards. A

It seems that plural pronoun reference is very flexible. However, as we will see in this chapter, pronominal reference set reference, exemplified in (3.1a), is the only kind of pronominal anaphora with a quantificational antecedent which is robust. For instance, it is difficult to interpret the pronoun in (3.2) as referring to the maximal set. That is, it does not have the reading that five students were at the beach (and not in the garden).[1]

[1]Nevertheless, quite some people disagree with this intuition. As we will see, pronominal maximal set reference is licenced whenever the domain of quantification is presupposed to

Similarly, the pronoun in (3.3) cannot be linked to the complement set. In fact, the only reading available for (3.3) is a nonsensical one, in which the MPs that attended the meeting 'went to the beach instead' as well.[2]

(3.2) There are four students in the garden.
Five of them are at the beach. *A

(3.3) Most MPs attended the meeting.
They went to the beach instead. *(A-B)

The reference set *is* available for these examples.

(3.4) There are four students in the garden.
One of them was picking flowers. $A \cap B$

(3.5) Most MPs attended the meeting.
They made a lot of important decisions. $A \cap B$

The question central to this chapter is the following: does a sentence of the form NP VP generally introduce discourse referents for the reference set, the maximal set and the complement set? At the end of the chapter, we will be able to conclude that a sentence *never* introduces a referent for the complement set, *sometimes* introduces a referent for the maximal set and *always* introduces a referent for the reference set.[3]

Here, I use the DRT terminology of referent-introduction as a convenient way of phrasing the question of whether a certain set is contextually salient. With this use of terminology come the following two assumptions: (i) only accessible referents can be salient antecedents and (ii) pronouns typically refer to salient antecedents. This means that when we want to refer anaphorically to something *not* available via a discourse referent, we will, in principle, have to do so by some linguistic form which bears more semantic content than the pronoun. But as we will see, the existing cases of pronominal reference to the complement set are examples of cases in which a pronoun is used to refer to a set not available by an accessible

exist. This means that quantifiers that do not trigger a presupposition of their domain will not always licence maxset reference. However, in cases like (3.2), it is easy to assume the existence of some (salient) set of students anyway. See section 3.2 for some examples of a different nature.

[2]Here and in the rest of the book, I use '#' to indicate that an expression is pragmatically anomalous. In the case of (3.3), the marking is due to the fact that a resolution for the plural pronoun which would render the continuation consistent does not seem to be available. If the pronoun has a set of MPs that did not go to the meeting as a potential antecedent, then the continuation is not in contradiction with the first sentence. The '#' judgement, therefore, indicates that this antecedent is, in this case, not available.

[3]An obvious exception to the general availability of the refset as an antecedent is caused by the determiner 'no'. However, this follows from the fact that we have assumed that pronouns carry some sort of non-emptiness presupposition (recall the discussion in section 2.4.2), since the reference set for a quantifier 'No N′′' is always the empty set. See section 2.4.2 for how this is realised in the final proposal of this thesis.

referent. I will come to this conclusion because the availability of the complement set for pronominal reference is not only constrained by properties of the antecedent sentence, but also constrained by properties of the utterance the pronoun occurs in. Accessibility of the complement set, I will conclude, is therefore not due to context-change, but due to a last resort strategy on the part of the hearer. In contrast, the availability of the maximal set relies solely on the properties of its antecedent and, therefore, if these relevant properties are met, a referent for the maximal set will be introduced.

In order to compare the constraints on referents to sets in discourse, we will need a way of modelling the restrictions on the different possibilities. Hendriks and de Hoop (2001) developed a framework called *optimality theoretic semantics* which will enable us to do just that. In their theory, different options for interpretation are resolved by resolving the conflict these options present to the hearer. Among many other phenomena, they discuss how domain interpretation can be formalised in such a way and, what is relevant to the present discussion, they briefly apply their theory to pronominal reference as well.

The interpretation of quantificational domains resembles the flexibility of pronominal reference in discourse. Quantification is context-dependent. A statement like the one in (3.6) does not necessarily commit its speaker to believe that everybody in the whole wide world was smiling.

(3.6) Everybody smiled.

The domain in (3.6) is not interpreted as the set of all human beings, but rather as some salient set in the discourse. This dependency may be on many different sets set up by a quantificational antecedent. This becomes especially clear when syntax does not overtly specify a domain. The examples in (3.7) shows a variation similar to those in (3.1).[4]

(3.7) Few MPs attended the meeting,

a. and several walked out. several attending MPs

[4]Hendriks and de Hoop (2001) even report a case where the domain is chosen to be the set denoted by the nuclear scope of the previous quantificational sentence.

(i) Er zijn in Nederland maar weinig vrouwen hoogleraar.
There are in the Netherlands only few women professor.

'There are only a few female full professors in the Netherlands'

(ii) De meesten zijn nog altijd mannen met baarden of brillen.
Most are still always men with beards or glasses.

'Most of them are still men with beards or glasses.'

Hendriks and de Hoop 2001, p. 24, ex. 44

The interpretation here would normally prefer to have 'de meesten' (most) range over female full professors or simply women in general, but these are blocked by the predication of being 'men with beards'. Instead, the set denoted by the scope of the first sentence, namely the set of full professors, is chosen as the domain of 'most'.

b. but all had an excuse. all non-attending MPs

c. most stayed at home. most MPs

The starting point for optimality theoretic semantics is a radical hypothesis about interpretation. De Hoop and Hendriks' so-called *free interpretation hypothesis* states that "all possible elements of a certain semantic type may serve as the elements between which a relation can be established by a relational type in a complex expression" Hendriks and de Hoop 2001 (p. 13). For quantificational domains, this means that a determiner may, in principle, choose any set as its domain. This in principle infinity of choices is then evaluated by a number of soft constraints, which differ in their importance. The interpretation of a quantificational structure is now that choice which violates the least important constraints. Thus, it represents an optimality theoretic approach to semantics.

Optimality theory (OT, Prince and Smolensky 1997) is really a theory of grammar. Given an input, a number of candidate forms is generated, the so-called candidate set. A hierarchy of violable constraints dictates which candidate is optimal, and hence grammatical. In OT semantics, the input is a syntactic string, which generates a set of possible interpretations. An interpretation I is optimal for form F if all its competitors show more serious violations than those of I. A violation of a constraint C is more serious than a violation of all or some of the constraints ranked below C. Optimality is graphically represented using tableau notation, (cf. figure 3.1, below) in which highly ranked constraints are left of lower constraints and violations are marked with '*'. The optimal candidate is marked with the symbol '☞'.

Returning to the subject of domain selection, it is clear what the merits of such a theory are. Since, as we have seen, domain restriction is extremely flexible, all possibilities should in principle have the option of resulting in the actual interpretation. Which set is chosen as a domain depends on the interaction of a number of soft constraints.

One of the constraints adopted by de Hoop and Hendriks describes the preference for choosing a *reduced* domain of a previous quantification as the current domain of a quantification. To be precise, *forward directionality*, as the constraint is called, states that the reference set of a previous quantificational sentence is to be favoured as the domain of the current quantificational sentence.

(3.8) Forward Directionality: The topic range induced by the domain of quantification of a determiner is reduced to the topic range induced by the intersection of the two argument sets of this determiner.

This constraint is motivated by the fact that forward directional interpretation seems to block other choices. For instance, in (3.9), 'three' is readily

(3.9)	AvoidContradiction	ForwardDirectionality
three students		*
☞three attending students		

(3.10)	AvoidContradiction	ForwardDirectionality
☞three students		*
three attending students	*	

Figure 3.1: Avoid Contradiction and Forward Directionality

interpreted as three students attending the meeting, rather than just any three students.

(3.9) Ten students attended the meeting. Three spoke.

Forward directionality, however, cannot be a hard constraint. The example in (3.10) is interpretable even though 'three' clearly cannot be interpreted as ranging over attending students.

(3.10) Ten students attended the meeting. Three didn't.

What happens in (3.10), according to the theory of de Hoop and Hendriks, is that a non-forward directional interpretation is preferred over a contradictive one. Therefore, the constraint 'Avoid contradiction' is postulated and ranked above forward directionality. This predicts the interpretation of both (3.9) and (3.10), as shown in figure 3.1.

Assuming that anaphoricity of pronouns is comparable to the anaphoric behaviour of bare determiners, forward directionality is a likely candidate for the modelling of a preference of pronouns for reference set reference. As we saw above, however, other choices for interpreting pronouns sometimes surface. We also saw that this possibility is not always an option. In fact, with respect to complement set reference, (3.1b) and (3.3), repeated here as (3.11) and (3.12), show a contrast in the availability of the complement set.

(3.11) Few MPs attended the meeting.
They stayed home instead. complement set

(3.12) Most MPs attended the meeting.
#They went to the beach instead. *complement set

The intuitive difference between (3.11) and (3.12) is the monotonicity of the antecedent quantifier.[5] Since (3.12), contains the monotone increasing

[5]A quantifier $D(A)$ is monotone increasing (or upward entailing) if and only if $\forall B \subseteq B' : D(A)(B) \rightarrow D(A)(B')$. A quantifier $D(A)$ is monotone decreasing (or downward entailing) if and only if $\forall B \subseteq B' : D(A)(B') \rightarrow D(A)(B)$

(3.11)	EMPTINESS	FORWARDDIRECTIONALITY
reference set	*	
☞complement set		*

(3.12)	EMPTINESS	FORWARDDIRECTIONALITY
☞reference set		
complement set	*	*

Figure 3.2: Emptiness and Forward Directionality

'most MPs' it does not guarantee the existence of any MPs not attending the meeting. It seems then that reference to the complement set is only allowed once we know it exists. This moves de Hoop and Hendriks toward proposing the following constraint:

(3.13) Emptiness: As the antecedent of an anaphoric expression, do not choose a set which is or may be empty.

Emptiness is ranked higher than forward directionality. This means that whenever the refset could be empty, we disregard forward directionality and selecting to satisfy the emptiness condition by choosing an alternative reference (e.g. the complement set). This yields a reasonable account of the distribution of complement anaphora. Monotone increasing quantifiers have potentially empty complement sets, so complement set reference would violate both emptiness and forward directionality. The corresponding tableaux are in figure 3.2.

Generalising on this approach we may expand the candidate set with the maximal set.[6] The ranking of 'emptiness' above 'forward directionality' now predicts that maximal set reference is possible whenever it is guaranteed to be non-empty and reference set reference is excluded. Intuitively, this seems to be on the right track. As I will discuss below, the class of determiners not licencing pronominal maximal set reference does indeed allow its domain to be empty. However, what the relation is with forward directionality remains to be found out.

The analysis of de Hoop and Hendriks sets the stage for a more detailed investigation into the availability of sets in discourse. What the data clearly points out is the principled flexibility of both pronominal reference and domain selection. This is captured by de Hoop and Hendriks' free generation of interpretations. Various preferences constrain the choices between these interpretations. In the remainder of this chapter, I will focus in much more detail on pronominal reference to sets and show that

[6]The candidate set should really be something which is generated by the free interpretation hypothesis. The presentation here is a schematic account of the theory. It leaves implicit many constraints which filter out the candidates that never surface.

there is more to say than de Hoop and Hendriks' initial successful analysis of.

3.2 Reference to the maximal set

The example in (3.2) showed that maximal set reference is not a general option. Similar to what we saw with the reference set and the complement set, not all determiners guarantee the non-emptiness of their domain of quantification.

(3.14) No unicorns have ever visited this forest.
$\not\Rightarrow$ there are unicorns

(3.15) It is unlikely that most unicorns visited this forest.
$\Rightarrow$ there are unicorns

(3.16) It is unlikely that two unicorns visited this forest.
$\not\Rightarrow$ there are unicorns

This is related to the intersectivity property. Formally, D is intersective iff $D(A)(B) \Leftrightarrow D(A \cap B)(B)$ (see Barwise and Cooper 1981; Keenan 1987). Intersective determiners do not give any information about their domain. For instance, while in (3.16) the reported unlikelihood is challenged by two unicorns visiting the forest, in (3.15) the majority of *some given set of unicorns* should pay the forest a visit. That is, in order to falsify (3.16) it suffices to find two visiting unicorns. The total set of unicorns is not at issue.

The relevant distinction here is that between weak and strong determiners (Milsark 1974). Weak determiners are intersective (Barwise and Cooper 1981) and due to this property fail to present any information about the domain. The classic test for weakness is occurrence in an existential 'there' sentence.

(3.17) There are four/a few/many/few/no/*most/*all/*the cats.

Quite some determiners that pass this test for weakness, however, can also have strong, i.e. non-intersective readings, although not in existential there sentences. Clearly, (3.18) has more than one reading. Its weak reading tells us that the students that came to the party are many in number. But it also has a strong reading in which the number of students that came to the party is high when compared to the total number of students.[7]

(3.18) Many students came to the party.

[7] Westerståhl (1984) gives two more ways to interpret 'many', both of which violate the conservativity property.

It is possible to isolate strong readings from weak ones. For instance, Milsark (1974) notes that only strong noun phrases can be the subject of individual-level predicates (that is, predicates which express inherent properties of individuals).

(3.19) Some students are intelligent.

(3.20) Some students came to my office.

The example in (3.19) cannot be uttered at the start of a conversation about some group of students who happen to be intelligent, while (3.20) could be uttered to introduce some group of students to the hearer. The only possible reading for (3.19) is that while *some* of the students are intelligent, others are not. Stress plays an important role here. The stress pattern within a noun phrase is another way of keeping weak and strong readings apart. While in (3.19), ‘some students’ has to be pronounced with stress on ‘some’ in order to get the felicitous strong reading, the preferred pronunciation in (3.20) is with stress on ‘students,’ yielding the weak reading. In fact, if we stress the determiner in (3.20), we get the partitive reading that some of the students came to my office, implying that other students did not.

(3.21) SOME students came to my OFFICE.

Using these characteristics of weak and strong determiners, we can show that only (occurrences of) strong determiners introduce a maximal set antecedent into the discourse. One clue comes from the definite description ‘the others.’ It needs to be able pick up two sets: a domain and a subset of that domain to which it forms a contrast. Following weak noun phrases, there is no domain which may form an antecedent for ‘the others’. Consequently, use of ‘the others’ is infelicitous following such noun phrases.[8]

(3.22) Some STUDENTS came to my office. # The others didn’t bother.

(3.23) SOME students came to my office. The others didn’t bother.

I take this to show that strong determiners presuppose their domain of quantification. That is, their use signals the existence of a set satisfying the restrictor clause in the context. (See also, Moltmann 1996)

Consequently, maximal set reference is only felicitous if the antecedent noun phrase is strong, since only those cases are cases in which the domain of quantification is accessible. Strong determiners presuppose that the context contains a referent corresponding to the domain of quantification. Weak determiners do not tell us anything about such a referent. The property of intersectivity allows them to simply count the individuals satisfying the property expressed by both arguments.

As we will see, the case of pronominal reference to the complement set is not quite as straightforward.

---

[8]Of course, there does exist an interpretation for (3.22) where ‘others’ refers to non-students.

3.3 Reference to the complement set

The phenomenon where an NP refers to the complement set of an antecedent quantificational structure is called *complement anaphora*. During the end of the eighties and the early nineties of the previous century, this type of reference was extensively studied by psycholinguists. In the nineties, formal semanticists began to get interested in the subject, but this interest resulted in just a few papers. The reason for this seemed to be that, in general, semanticists reacted sceptically to the experimental data. They questioned whether examples like (3.35) really involved reference to the complement set, or rather a more generic-style reference to the set of MPs.

Thus, two questions became central to the study of complement anaphora: (i) do pronominal complement anaphora involve reference to the complement set? and (ii) what constraints explain the distribution of complement anaphora? I will address both questions. First, however, we will consider the psycholinguistic data which sowed the seeds of interest in the semantics of complement anaphora.

3.3.1 Moxey and Sanford's experiments

Studies of complement anaphora have their origin in the psycholinguistic research of Linda Moxey and Anthony Sanford. They found that when subjects were asked to give intuitions about denotations of related determiners like *a few*, *few*, *very few*, *only a few*, *not many*, etc., no reliable differences were found. The goal of a new set of experiments thus became to find out what the functional difference between these determiners is. See Sanford and Moxey 1993 for an overview, but also Moxey and Sanford 1987 and Sanford, Moxey, and Paterson 1994.

In a sentence completion task, subjects were confronted with a single quantified statement and asked to make up a sensible continuation beginning with the plural pronoun 'they'.

(3.24) Q of the MPs attended the meeting. They...

Subjects were also asked to indicate what the plural pronoun referred to in their continuation. Here, they had a choice between five categories: *MPs in general*, *all MPs*, *MPs who went to the meeting*, *MPs who did not go to the meeting* and *none of the above*. Independent judges checked all the utterances and reference indications. In 98% of these cases, the judges agreed with the judgments of the subjects.

In an alternative experiment, intra-sentential complement anaphora were tested, using a structure like (3.25).

(3.25) Q of the MPs attended the meeting, because they...

The results showed that following a structure which had one of the determiners *hardly any, not many, very few* and *few* substituted for Q, subjects preferred to use the plural pronoun to refer to the complement set of the quantification over MPs, that is, the set of MPs who did not go to the meeting. With other determiners (like *many* and *a few*), the tendency was to refer to the reference set (that is, the MPs who did go to the meeting). A special case was *only a few*, which only showed compset reference in the task involving a structure like (3.25).

In a different study, complement set reference was studied with proportional numerical expressions. The continuation method and the use of judges was as in the experiment above (cf. Sanford and Moxey 1993, p. 77). Here, it was found that compset continuations were favoured following determiners *less than n%*. The other determiners (*n%, only n%* and *more than n%*) showed hardly any continuations containing complement set reference.

'Preference' for a continuation containing reference to the complement set should be interpreted rather weakly as roughly indicating that more than half of the subjects used a complement anaphor. That is, refset continuations *did* occur following *few*, *not many*, etc. In fact, Moxey and Sanford 1987 conclude that though "*few* and *not many* licence compset anaphoric mappings [t]hey do not however, require them" (Moxey and Sanford 1987, p. 203). In contrast, they remark that it appears that determiners like *a few* can never be followed by pronominal reference to the complement set.

Sanford and Moxey stress that the results show that determiners do not just identify different proportions, but moreover place focus on different parts of the domain. In fact, there is evidence for yet another role played by determiner choice, namely that of steering the thematic content of a continuation. This is supported by experiments testing for a correlation between determiners and rhetorical choices made. Independent judges were asked to indicate to which types the produced continuations belonged. The possible continuation categories for a sentence like 'Q MPs attended the meeting' were: *reason why there*, *reason why not there*, *consequence of the number of MPs being there* and *other*. It was found that determiners that focus on their complement set displayed the 'reason why not there' type as the most prevalent type of continuation. These determiners thus seem to motivate the subject to indicate "the reason why the predicate is not true of the refset" (see Sanford and Moxey 1993, p. 66).

In sum, taking *few* as the prototypical case of complement set focus, for Moxey and Sanford, the utterance of an expression "Few X do Y" consists of three *actions*: "(i) identify a small percentage of Xs of which "does Y" is true; (2) put into focus the set of which Y is false, and (3) set the system to expect a reason why the division is *as small* as it is" (Moxey and Sanford 1987, p. 203).

3.3.2 Semantic data

The experimental data show a clear preference pattern associated with different types of determiners. From a semantic point of view, it is clear that determiners differ greatly with respect to their potential to licence complement anaphora. Moxey and Sanford's experiments showed that following a determiner like *many*, there was no preference for using a complement anaphor. It is now widely accepted that following such 'positive' quantifiers, reference to the complement set is disallowed, as is illustrated in (3.26).

(3.26) Many/Most/A few MPs went to the meeting.
They were too busy.

The analysis by de Hoop and Hendriks, discussed above, makes a first attempt at explaining the pattern. Their emptiness constraint supposes that the data in (3.26) is explained by the fact that the quantifiers in (3.26) are monotone increasing. A case of complement anaphora will therefore be an attempt at reference to a potentially empty set. Given such an analysis, pronominal complement anaphora are licenced by a non-increasing antecedent. Unfortunately, not all decreasing quantifiers licence complement anaphora. The data show a difference between (modified) cardinal and proportional (decreasing) quantifiers. Cardinal quantifiers (like 'thirty N') and modified cardinal quantifiers (like 'less than thirty N') are intersective. As we saw in the discussion of the maximal set, in a quantificational sentence $D(A)(B)$, if D is an intersective determiner only the A's that are B need to be considered. On the other hand, if D is a proportional determiner the A's which are not B also need to be taken into account. Intersective decreasing quantifiers disallow pronominal reference to the complement set, while proportional decreasing quantifiers allow it.

(3.27) Less than thirty MPs attended the meeting.
They were too busy.

(3.28) Less than thirty percent/Less than half/Few of the MPs attended the meeting.
They were too busy.

The example in (3.27) improves once the decreasing modified cardinal is in a partitive construction.

(3.29) Less than thirty of the fifty MPs attended the meeting.
They were too busy.

These observations are in sharp contrast with data concerning reference set reference. No matter what the antecedent determiner is, reference to the refset is always allowed.

(3.30) Most MPs attended the meeting.
They discussed a lot.

(3.31) Few of the/Less than thirty MPs attended the meeting.
Nevertheless, they managed to discuss a lot.

In contrast to Moxey and Stanford's findings, which deal with focusing and preference patterns, the data given here are strict facts about interpretability. While Moxey and Sanford focus on why decreasing determiners emphasise their complements set, I wish to focus on why increasing determiners completely disallow subsequent pronominal complement set reference, while *all* determiners allow for reference set reference. The data show that De Hoop and Hendriks' elegant explanation in terms of a conflict of constraints needs to be amended in order to explain this asymmetry.

Like de Hoop and Hendriks, I have assumed up to now that the cases of complement anaphora and cases of anaphoric reference to the complement set are the same. However, as mentioned before, it is not immediately clear that what is involved in complement anaphora *really* is reference to the complement set. Quite a few researchers consider this to be a different kind of phenomenon. Before turning to what explains the strict pattern of complement anaphora licencing, it is important to first concentrate on the question of what a complement anaphor actually is.

3.3.3 Complement anaphora and reference to compset

Corblin (1996b) argues that accepting the complement set as a potential antecedent for plural reference is inconsistent with an important generalisation made by Kamp and Reyle. Based on Partee's example (3.32), Kamp and Reyle conclude that "apparently, subtracting one set from another is not a permissible operation for the formation of pronominal antecedents" (Kamp and Reyle 1993, p. 307).

(3.32) Eight of the ten balls are in the bag.
They are under the sofa.

Corblin goes on to argue that complement anaphora is really a case of reference to the maximal set. Following Corblin, we will call this *pseudo reference* to the complement set.

Reference to the restrictor following a quantified structure is in principle always possible. Kamp and Reyle's (3.33) is the standard example:

(3.33) Few women from this village came to the feminist rally. No wonder. They don't like political rallies very much.

Here, the plural pronoun in the final sentence refers to the restrictor of the subject quantifier, the *women from this village*.[9] Corblin argues that it is

[9] There is also a reading of (3.33) where the plural pronoun refers to women in general. Corblin allows two maximal set references: the restrictor and the bare noun inside the restrictor. Like Corblin, I will ignore the latter type of reference.

this type of reference that gets confused with complement set reference.
Corblin gives the following example to strengthen his proposal.

(3.34) Peu d'électeurs français ont voté pour le candidat du
Few of-voters French have voted for the candidate of-the
parti communiste. Ils ont voté pour le candidat de la
party communist. They have voted for the candidate of the
droite à 40% environ.
right with 40% average.

"Few French voters voted for the candidate of the communist party. Approximately 40 percent of the voters voted for the right-wing candidate."

He argues that this example cannot possibly be an example of (genuine) complement set reference, since the adverbial *à 40% environ* would render the sentence with an explicit reference to this set false.

The example above doesn't show that reference to the complement set does not exist, it merely illustrates the fact that maximal set reference is a potential alternative. I will discuss two proposals for a pseudo-reference analysis of complement anaphora more concretely. One is Corblin's own analysis, the other an alternative suggested by Bart Geurts.

3.3.3.1 Two pseudo reference analyses

The general idea in Corblin 1996a and Corblin 1996b is that complement set reference is really a case of reference to the maximal set under an implicit restrictive modifier.

Corblin identifies the complement set licencing determiners as those which demonstrate that the reference set is smaller than expected. Thus, by referring to the maximal set, we know that its majority will consist of elements not satisfying the nuclear scope. In a continuation, we can therefore confuse a restricted reference to the maximal set with reference to the complement set.

(3.35) Few of the MPs went to the meeting.
They stayed home instead.

In (3.35), the first sentence declares that the proportion of MPs that went to the meeting is small. In the second sentence we can therefore generalise over the MPs and explain their (general) absence. Technically, Corblin adopts a DRT approach. Since the antecedent sentence expresses the fact that the refset was smaller than some norm, a partial abstraction operation is triggered creating the reference to the maxset. Restrictive modification then weakens this maxset reference, causing the resemblance to complement set reference. The application of restrictive modification is natural since any sentence, Corblin argues, carries either explicitly or implicitly some quantificational modification. Sometimes, this implicit modifier has to be universal (in a context where strictness is important, for instance), but often the modifier expresses some kind of generalisation.

An alternative to Corblin's analysis was proposed by Bart Geurts in his book review of Moxey and Sanford's book (Geurts 1997). The proposal is based on the concept of *collective reference*.[10] This phenomenon is quite common with plural definite descriptions.

(3.36) The soldiers withstood the attack. (from Link 1991)

Reference in (3.36) can be sloppy in the sense that the sentence is considered suitable to describe even situations where not *all* the soldiers were able to cope with the attack.

According to Geurts, a complement anaphor is an instance of this kind of sloppy reference. The plural pronoun in compset continuations refers to the maximal set, which is collectively held responsible for the negation of the nuclear scope in the antecedent sentence.

3.3.3.2 Arguments against pseudo reference

Several authors have criticised the line of reasoning explained in the previous section, most notably Moxey and Sanford themselves, but Kibble 1997a (p. 260-264), also gave a number of arguments against such accounts. In this section, I will evaluate the arguments found in the literature. In the end, I will come to the conclusion that complement anaphora really *do* involve complement set reference. Still, many of the arguments against pseudo-reference accounts that I discuss below remain rather weak. I will explain why this is so, but I will also explain why one of these arguments is strong enough to seriously consider complement anaphora as reference to the complement set.

Personal judgments

An obvious problem with reducing complement anaphora to a confusion of maxset with compset is that it is not clear why subjects would not be aware of such a confusion. Remember that in Moxey and Sanford's experiments, subjects were asked to give their personal judgment on what they had referred to. These were almost always compatible with judgments of independent judges. Subjects were even given the explicit option of judging their continuation to be a generalisation over the maximal set.

[10]It must be noted that Geurts' proposal is simply an option he considers. In his book-review he still seems to think of reference to the complement set as a realistic possibility. Nevertheless, I will refer to the collective reference variant of pseudo-reference as "Geurts"s proposal".

Moxey and Sanford report that some critics have weakened this argument by suggesting that people in general do not know what they refer to and that their comment on the reference of *they* is simply a naive judgment on language use. Moxey and Sanford themselves find it "hard to see why there is a misunderstanding over the intended referent" (Sanford and Moxey 1993, p. 64), but I tend to agree with the critics. How can naive language users even understand what is meant by the question: where does the word *they* refer to? It is clear that even if complement set reference did not exist and complement anaphora were to reduce to a weakened form of maximal set reference, the objects that matter to the continuation were still just those in the complement set. It could well be that subjects answer the question in that light.

The use of *instead*

A different argument against a pseudo reference account of complement anaphora is that the frequent use of *instead* seems to indicate that a reference is made to a complementary situation. We are not able to use *instead* when no contrast is made. One of the continuations from Moxey and Sanford's experiment is given in (3.37).

(3.37) Hardly any of the MPs went to the meeting.
They were out at the pub or with their secretaries instead.

If this were a case of maximal set reference, then there would be no contrast. Genuine compset reference would make the use of *instead* felicitous since it implicitly denies the fact that more than a handful of MPs attended the meeting.

This is my interpretation of the standard argument against pseudo-reference based on the use of *instead*. It seem, however, that this line of critique cannot be maintained. In fact, we can simulate Corblin's approach to a continuation including *instead* as follows.

(3.38) Few of the students went to the party.
Most of them went to the beach instead.

The plural pronoun 'them' refers to the maxset and there is no explicit negative property being contrasted, yet, the use of *instead* is unproblematic. It seems to me that the semantic aspects of *instead* are not well known and that rejecting pseudo reference on the basis of an occurrence of *instead* is rather preliminary.

Notice that the use of *instead* in (3.39) is ambiguous.

(3.39) Tom went to the cinema, and Bill stayed at home instead.

(3.39) is ambiguous between Bill staying home instead of John and Bill staying home instead of going to the cinema. This means that *instead* does not necessarily seek to contrast full sentences. We could thus easily think of the use of *instead* as contrasting a VP and we are certainly not obliged to think of *instead* as operating on full sentence meanings, including for instance, implicit modifiers. The example in (3.37) could thus very well be an instance of maximal set reference.

Explicit definites

Another common criticism of an approach to complement anaphora involving some sort of implicit modification of maxset-reference is that the restricted reading is absent when the plural pronoun is replaced by a definite description directly referring to the maximal set.

(3.40) Few of the students went to the party.
They were too busy.

(3.41) Few of the students went to the party.
The students were too busy.

In (3.40) the plural pronoun corresponds to the full set of students. Since we know that few of them went to the party we can apply a restrictor to the maximal set and thus identify this weakened maximal set with the complement set. The puzzle is why in (3.41) the implicit operator suddenly cannot do its work.[11]

I am not convinced that this puzzle forms an argument against Corblin's approach. In fact, the oddness of (3.41) is part of a much larger puzzle. Consider, for instance:

(3.42) All students came to the party.
They had a good time.

(3.43) All students came to the party.
The students had a good time.

(3.44) All students came to the party.
The students who came had a good time.

In (3.42), the plural pronoun refers to the maximal set, which is equal to the reference set. However, if we explicitly refer to either maximal set or reference set by a definite (as in (3.43) and (3.44) respectively), then the continuation is out. The puzzle of (3.41) obviously has little to do with the implicit restrictive modifier, which might or might not be present. More insight is needed into the difference between the use of a pronoun and the use of a definite description when used anaphorically in these contexts.

[11]Moreover, in (3.99), below, and example is given which shows that an epithet definite description *is* allowed to refer to the complement set.

Notice also that the argument cannot be turned *against* complement set reference. Definite descriptions are possible as genuine complement anaphora, witness the complement subsectional[12] anaphor in (3.46).

(3.45) Most of the senators attended the meeting.
The democrats sat on the opposite side of the republicans.

(3.46) Few of the senators attended the meeting.
The republicans had a good excuse, the democrats didn't.

(3.46) shows that definite descriptions (here, as so-called partial match anaphors) *do* have the potential to refer to (parts of) the compset. Consequently, there is neither an argument against nor in favour of pseudo reference. Definite descriptions can have complement set reference, but that doesn't tell us anything about whether or not this is real reference to compset or some kind of pseudo-reference.

The strength of the modifier

Corblin's analysis assumes that complement set licencing determiners report on the smallness of the reference set. Moxey and Sanford's studies tell us that there are a few determiners that do licence complement anaphora but do not offer a smallness judgment.

(3.47) Not quite all of the MPs attended the meeting.
They stayed at home instead.

Note that it does not do to dismiss this on account of the vagueness of a determiner like *not quite all*. "True" downward monotone proportional determiners like *less than 90%* also licence complement anaphora.

(3.48) Less than 90% of the MPs attended the meeting.
They stayed at home instead.

Following Corblin's analysis, the second sentences in (3.47) and (3.48) have the plural pronoun referring to the maximal set. By the presence of a restrictive modifier the sentences comment on a generalisation of this set. The problem is, however, that *They stayed at home* reports on a minority group of MPs.

Notice that Geurts seems immune to this line of critique (as he points out himself). Even minorities can be responsible for collective reference. In (3.49), there is an example where the actual subset of individuals satisfying the predication could, in fact, be a minority.

(3.49) The local residents organised a barbecue.

[12] Subsectional anaphora are anaphora that refer to subsets of their antecedent. An example is an anaphoric relation between the definite description *John's pets* and the subsectional anaphor *the dogs*. See Krahmer and van Deemter 1998 and references therein.

Imagine, for instance, a situation where there are a hundred residents and an organising committee consisting of ten of them. In such a situation (3.49) could be truthfully uttered.

3.3.3.3 Plural reference

So far, we have seen that there exist a whole range of arguments against pseudo-reference, but that most of them are far from convincing and can very easily be dismissed. Moreover, the most powerful argument, the strength of the modifier involved, seems to hold only for one type of pseudo reference, namely one wherein an implicit operator which is responsible for the generalisation over maxset is assumed. Other analyses, where the effect is seen as a side-effect of plural reference, cannot be subjected to this criticism. In my view, this allows us to reject the analysis based on some sort of implicit modification of maximal set reference. In this section, I will investigate the potential of the other pseudo reference account, which is based on inherent properties of plurality.

Geurts refers to the type of reference he suspects to be involved in complement anaphora as *collective reference*. As already mentioned, in this type of reference, subsets of the predicate's argument can be responsible for the complete satisfaction of the predication. I should start by acknowledging the existence of such reference and by making it clear that such reference undoubtedly occurs in complement anaphoric continuations. What is at stake here is whether complement anaphora are necessarily instances of collective reference.

Plural reference, distributivity and *all*

Dowty (1987) illustrates in a clear way that non-total definite reference displayed by plurals is not limited to predicates which are clearly collective. He shows that apparently plural reference in general is not committed to strict truth-conditions, witness the example in (3.50).

(3.50) At the end of the press conference, the reporters asked the president questions. Dowty 1987 (ex. 23)

While one would want to classify *to be asking a question* as a distributive predicate, this example does not turn out false if, in fact, not every one of the reporters asked the president a question. Dowty argues that sentences like (3.50) are not true modulo some handwaving, but that they are in fact literally true due to the semantics of the definite determiner. In the lexicon, *distributive sub-entailments* are specified and these give the conditions for individuals to satisfy the verbal predicate. The definite determiner does not force its complement to fully satisfy the distributive sub-entailment.

The effect which might be responsible for complement anaphoric reference is thus not limited to collectivity. There is another property of the type of reference under discussion which is relevant for our evaluation. This is the observation that any kind of sloppy plural reference is strengthened by the quantifier "all".

Consider again an example of sloppy reference, given below in (3.51). Plural definite reference allows a subgroup of the total group of soldiers to be responsible for a collective hunger (notwithstanding the fact that the involved VP is distributive). This means that in (3.51) some soldier might actually be fine. Such a possibility typically disappears once the floating quantifier "all" occurs in the sentence. In (3.52), there are no exceptions allowed.

(3.51) The soldiers were hungry.

(3.52) The soldiers were all hungry.

In Dowty's terms, "all" forces all the sub-individuals to fulfil the predicate's sub-entailments. With distributive predicates this means that each individual should be an independent member of the interpretation of the predicate. With collective predicates this means that each individual has to have some (distributive) property.

How is this relevant for the current discussion on complement anaphora? If complement anaphora involves plural reference to the maximal set, then the presence of a floating quantifier "all" undoes the weakening to pseudo complement set reference. That is, the floating quantifier turns the interpretation into a strict maximal set reference. In case complement anaphora involve reference to the complement set, then floating quantifiers have little effect, since they would totalise reference to the complete complement set.

It so happens that there is an oddity with complement anaphoric continuations. Moxey and Sanford report that in quite a few continuations judged to involve reference to compset, subjects used the floating quantifier "all". One example is given in (3.53).

(3.53) Few of the fans went to the football match. Sanford and Moxey 1993
They all watched the game on television instead. (p. 74)

One might suggest that what the floating quantifier does here is in fact exaggerate the inferior size of the reference set further. By the bold statement that every one of the fans saw the game on television, the reference set is claimed to be empty. Frankly, there is no way to tell and here we stumble on a difficult point in our evaluation. It may well be that the choice between real compset reference and pseudo reference cannot be made. Since the pseudo reference analysis makes use of the vagueness of language, it is not possible to construct counterarguments based on systematic properties of natural language interpretation possible.

Similar considerations arise when we return to the argument against pseudo-reference based on situations where considerable weakening of maxset reference is needed. An analysis in terms of the potential sloppiness of plural reference was thought to be immune to such an argument since even minorities can take responsibility for a property of the mass they belong to. This, however, is only partly true. Consider again (3.47), repeated below.

(3.47) Not quite all of the MPs attended the meeting.
They stayed at home instead.

In contrast to (3.49), it is hard to imagine how in the second sentence in (3.47) a minority could take responsibility for the satisfaction of the VP. How can a handful of MPs be responsible for a collective absentness? The situation in (3.47) is quite different from that in (3.50), where truth-conditions are almost indifferent to how many of the reporters actually asked a question. But would we say that "the MPs stayed at home" is true when in fact only a minority really did? Geurts is right in saying that collective reference can sometimes allow any subset of an argument to satisfy the predicate. But the degree to which this is possible depends on the predicate involved. We do not say that "the soldiers were hungry", when only a few actually were. In the same way, the second sentence of (3.47) will not be true if only a small part of the pronominal reference consists of individuals that stayed at home.

Pseudo reference, entailment and implicature

Once again, one could counter the critique of the previous section by claiming that complement anaphora are simply stylistic exaggerations. This begs the question of how analyses of pseudo reference explain the set of compset licencing determiners. An intuitive answer is the following: compset licencing determiners do not allow the reference set to be equal to the restrictor – i.e. following such a determiner, sloppy reference could never be so sloppy as to refer to an empty set and is therefore allowed.[13] Notice that such a proposal would be based on a rather vague relationship between the logical properties of the determiner (emptiness potential) and the speaker's/hearer's intention of producing/interpreting a pronoun with a weakened reference. The speaker is apparently allowed to use the maxset as a sloppy compset since she knows about some fundamental formal properties of the preceding quantification. This particular way of referring is thus a pragmatic choice of the speaker.

[13] Basing the criterion for compset reference on the *largeness* of the mistake made by sloppy reference makes this account actually a near notational variant to a *real* compset analysis based on some notion of emptiness. Compare with section 3.3.6.

Nevertheless, notice that a language user will not base her choice of reference/ resolution on non-logical facts. Consider (3.54).

(3.54) More than 10% of the students went to the party.
They went to the beach instead.

The determiner "more than 10%" is not a complement set reference licensor. This is logically quite clear, because this determiner is right monotone increasing and therefore allows the quantification to be true even if all students went to the party; this would make pseudo-reference too sloppy, so complement anaphora are disallowed. In addition to the monotonicity entailments, however, (3.54) also evokes implicatures. One of these implicatures is that not all students went to the party. In fact, the first sentence provides the implicature that the real proportion of students that went to the party is not far from ten percent. A potential problem for pseudo-reference is now that if the relation between determiner-type and action of the language user is one of pragmatic considerations, why then doesn't the language user react on the implicatures evoked by the structures.

The distribution of complement set reference seems to be very much governed by logical properties of determiners. This means that a theory of complement anaphora should focus on these formal properties and find the precise causal link between those properties and subsequent anaphoric potential. In my view, pseudo-reference analyses could therefore only be valuable if they adopt some formal mechanism describing what pseudo-reference is all about. Such theories of plural reference do exist (cf.Link 1983; Brisson 1997),[14] but they typically also adopt a strict semantics for floating "all". Such analyses will then additionally have to explain the (flexible) use of "all", as well as account for the normal unacceptability of extreme sloppiness in complement anaphora cases like (3.47).

3.3.3.4 Reference to the complement set

We saw that analysing complement anaphora in terms of pseudo-reference to the maximal set begs the question of how a formal relation can be realised between the set of right decreasing determiners and the possibility of pseudo-reference, especially since these analyses make use of the vagueness of plural reference, while complement anaphora seem to be so closely related to very formal properties of language. In what follows, I therefore assume that complement anaphora involve reference to the complement set. But accepting complement set reference as an unproblematic alternative might be too hasty. Corblin argues that accepting complement set reference would oppose Kamp and Reyle's generalisation. However, it needs to be stressed that acknowledging the existence of complement set

[14] Although Dowty gives a programmatic overview of how to deal with plural reference, he does not give a formal account, he himself admits. He notes that there might be several compositional difficulties involved. See Dowty 1987, p. 111.

reference does not mean that a subtraction operator is freely applicable. In fact in the coming sections, I will argue that complement set reference is marked and is subject to some strict conditions.

3.3.4 Dynamic Quantification

If we accept complement anaphora as anaphora that refer to the complement set, the main question becomes how this type of reference comes about and why it is constrained to such specific contexts. Related to this question is the question of how e-type pronouns refer in the first place. Why are maxset and refset generally available sets for anaphoric reference and why is reference to the extension of the scope, for instance, impossible?

A possible analysis of the anaphoric potential of quantificational sentences is to treat this type of sentence as dynamic and subsequent anaphora as cases of *dynamic binding*. As we saw in chapter 2, a formalism like DPL displays the so-called donkey equivalence: $\exists x(\varphi)\wedge\psi \;=\; \exists x(\varphi\wedge\psi)$. Such a formalism can be exploited to derive a dynamic approach to generalised quantification. Conservativity tells us that only two sets are needed to derive the truth-conditions of a sentence $D(A)(B)$, viz. A and $A \cap B$. These sets are also two likely candidates for future anaphoric reference. Thus, if we existentially introduce these sets and check whether they relate in the way D desires, then these two sets will be accessible for subsequent anaphoric reference. This observation is the central driving force behind some dynamic quantifier accounts of e-type anaphora (e.g. van den Berg 1996b, but especially chapter 4 for a detailed consideration of such approaches). Moreover, Rodger Kibble (1997b, 1997a) claims that such an analysis can even account for the distribution of complement anaphora. He bases his claims on van den Berg's dynamic quantifier logic. In that formalism a quantificational sentence $D_x(A)(B)$ is represented as follows:

(3.55) $\epsilon_{x'}(\max_{x'}(A[x/x']) \wedge \epsilon_x(\max_x(x \leq x' \wedge B) \wedge Q_D(x')(x)))$

Postponing the interpretation details of such a representation to chapter 4, the formula in (3.55) is read as follows. The ϵ operator is a dynamic existential quantifier ranging over plural individuals and the maximality operator 'max' ensures that only maximal collections are exported. The term $x \leq x'$ expresses that the value for x is a part (or subset or sub-individual, depending on ones favourite ontology for plurality) of the value of x'.

The representation in (3.55) introduces a referent x' for the maximal value satisfying the restrictor (i.e. maxset) and a referent x for the maximal subset of x' that satisfies the scope (i.e. refset). The final condition on these two referents is membership of the quantificational relation Q_D associated with the determiner D.

Any new occurrences of x' and x will now be dynamically bound and thus the maximal set and the reference set are available for anaphoric reference.

A well-know problem with representations like the one in (3.55) is that they are truth-conditionally flawed whenever $Q_D(x')$ is monotone decreasing and B involves collective predication. For example, the sentence in (3.56) is true according to (3.55) if there exists a maximal set of students gathering in the square yesterday, which are "few" relative to the maximal set of students.

(3.56) Few students gathered in the square yesterday.

But consider a situation in which there was more than one *gathering* yesterday in the square. One of them could involve only a handful of students (thus satisfying (3.55)) while another could be a massive meeting of students. This second maximum should not be overlooked by the representation of the quantificational sentence. It is not sufficient that there exists *some* local maximal set satisfying scope and restrictor and counting as *few*. In fact, *all* such maximal sets should satisfy the quantificational relation.

A well-known solution to this problem is to represent right monotone decreasing quantificational structures as the negation of their increasing counterpart. That is, (3.56) is treated as (3.57).

(3.57) It is not the case that 'many' students gathered in the square yesterday.

Van den Berg proposes that the same metamorphosis for decreasing quantifiers could also apply in the dynamic case (with a dynamic notion of negation) and thus account for the whole range of e-type anaphora. Kibble notices, based on work by Zwarts (1996), that there are basically two ways of giving a truth-conditionally sound representation for examples like (3.57). These two alternatives are due to the fact that a quantifier has two complements. One gives us a representation based on (3.57) and the other is based on (3.58), with different logical realisations of the complement of *few* (here twice loosely represented as *many*).

(3.58) 'Many' students did not gather in the square.

Truth-conditionally, there is no difference between (3.57) and (3.58), but if we translate the two into dynamic quantificational representations, then, dynamically, they will be different. The referent for the conservative scope (x in (3.55)) will correspond to the set of all the students who were in a gathering in the square yesterday, in the case of (3.57), but in (3.58) this referent will correspond to the complement set, the set of students that did not participate in gathering in the square yesterday.

It appears then that the need for an alternative representation for decreasing quantifiers leads to an ambiguity between either $A \cap B$ or $A \cap -B$ for the reference set referent. Moreover, Kibble argues that when the quantificational relation involved is cardinal, the value for $A \cap -B$ cannot be established since no information concerning the (relative) cardinality of A is known. This accounts for the distribution of complement anaphora.

3.3.5 The interpretation perspective

We have seen that dynamic quantifiers create an independent need for an ambiguity between compset introducing and refset introducing representations of decreasing proportional quantifiers. In this section, I want to point out that however striking Kibble's findings are, it is unwise to treat compset and refset on a par. That is, the complement set and the reference set have quite a different status as antecedents. I will argue that in interpretation there is a natural preference for refset resolution over resolution to the complement set.

We have already seen that Moxey and Sanford found a thematic effect in compset continuations, namely that they usually specify the reason why a relatively large proportion of the domain did *not* satisfy the scopal predication. From an interpretation perspective it is interesting to see what happens if we use complement anaphora in continuations other than the thematic preference found by Moxey and Sanford. Consider the following examples.

(3.59) Few of the students went to the party. I know who they are.

(3.60) Few of the 20th century presidents of the USA were elected for two consecutive terms. My history teacher made me learn their names by heart.

(3.61) Few of these balls are blue.
Can you point them out for me?

In all these examples the predications in the first and second sentence are neutral with respect to one another. That is, there is no relation between the predication in the antecedent sentence and the predication over the plural pronoun, which would lead the interpreter to resolve the pronoun to either refset or compset. In (3.61) for example, the speaker could ask the hearer to point at any set of balls. Resolving the plural pronoun, however, shows a clear preference for refset reference. It is unlikely that anyone would point at the non-blue balls, if asked the question in (3.61). It appears then that the default interpretation for plural pronouns is the intersection of restrictor and scope no matter the formal properties of the determiner.

The point made here is that refset is (if possible) the preferred resolution of a plural pronoun. The complement set has no such status. During resolution it is overruled by the reference set in neutral situations. It seems to me that compset interpretation is the result of a last resort strategy. Notice that a side effect of the type of continuation that was most common in the cases of complement anaphora found by Moxey and Sanford, *reason-why-not* continuations, is that resolving the plural pronoun to refer to the refset results in a contradiction. The only non-contradictory resolution is, of course, the compset. Only in these situations will compset interpretation occur.

In Moxey and Sanford 1987, we find an example showing that *reason-why-not* continuations are not obligatory. In (3.62), there is a clear case of complement set reference, but the fact that the members of compset send their apologies does not really indicate why so few MPs were at the meeting.

(3.62) Few MPs attended the meeting.
They sent apologies for being absent.

Still, once again we see that resolving the plural pronoun to refset reference would result in a contradiction.

Notice that Kibble's account given in the previous section is not able to cope with the general preference for refset reference. On Kibble's story, compset *is* a reference set: the distinction between the two sets is no longer made.

More support for the view that the interpretation of anaphora shows a preference for refset comes from explicit reference to the complement set. Notice the following contrast.

(3.63) Few of the students went to the party.
The others stayed at home instead.

(3.64) Few of the students went to the party.
The others had a good time. non-party goers/$^{??}$party goers

In (3.63), we see that we can replace the complement anaphor with an explicit reference to the complement set: *the others*. This definite description takes the complement of the refset relative to some domain of quantification (the maxset).

If we now take Kibble's analysis serious, there is really no difference between compset and refset other than that they are the reference set of different but logically equivalent representations of a quantificational structure. In other words, there is no apparent reason why *the others* should not accept the compset as an antecedent reference set. This, however, gives us the odd (3.64). Even though the choice of the predicate supports complement set reference, (3.64) is still odd.

Given these observations, the complement anaphora paradigm becomes a bit more complex than the association with monotone-decreasing proportional quantifiers. First of all, we see a general preference for reference to refset. Second, we see that this preference can be overruled by contradictive predication. Finally, we see that the possible cancellation of the preference only occurs with downward proportional antecedents.

3.3.6 Forward directionality and emptiness again

The difference between the complement set and the reference set discussed above is reminiscent of the soft constraint *forward directionality* from de

Hoop and Hendriks' account of quantificational domains. A choice between complement set reference and reference set reference will always turn out in favour of the latter option. As we saw, de Hoop and Hendriks' analysis is too strong, since it predicts that refset reference and compset reference are in complementary distribution. This was because their soft constraint *emptiness* blocks reference to any potentially empty set, including reference sets of decreasing quantifiers. Now, we find a second complication, namely that in the cases where both types of reference are possible, refset reference actually blocks complement anaphora.

It is useful to restate the complement anaphora paradigm in terms of emptiness potential. Below, (A) corresponds to monotone increasing contexts, (B) to intersective downward monotone ones and (C) represents the cases with true compset licensors (other monotone decreasing determiners).

(A) compset possibly empty, refset non-empty: *compset/refset

(B) compset and refset both possibly empty: * compset/refset

(C) refset possibly empty, compset non-empty: refset>compset

Although forward directionality expresses a preference for reference set reference, this preference can never surface in case (C). Even if the reference set is potentially empty, reference to it is still felicitous. Let us therefore, for the sake of the OT analysis, stipulate a variant of the emptiness constraint:

(3.65) Emptiness': As the antecedent of an expression do not choose a set which is potentially empty, except when this set is the reference set of a sentence.

Ranking the alternative emptiness' constraint over forward directionality now accounts for the fact that, in the downward proportional cases, reference to both complement set and reference set is (in principle, at least) possible, but that there is a preference for the latter. The fact that this constraint has a conditional character, suggests that it is not a soft constraint at all, but part of the generator, the set of hard constraints that dictates which interpretations are in the candidate set. If emptiness' were part of the set of soft constraints, the exception it models would have to follow from an interaction between constraints.

Although the principle in (3.65) is clearly in need of justification, I briefly explore the potential of the optimality theoretic analysis further, since it is an ideal way of modelling the interaction of constraints on reference and will thus enable us to clarify the paradigm.

3.3.6.1 Avoid Contradiction

An unresolved issue concerning the presented paradigm is how to derive the potential cancelling of the preference for refset. One way to account for

		EMPTY'	AVOIDC	FORWD
Most(A)(B). They ¬B	☞refset		*	
Most(A)(B). They ¬B	compset	*		*
Most(A)(B). They C_{neutral}	☞refset			
Most(A)(B). They C_{neutral}	compset	*		*
Less than half the (A)(B). They ¬B	refset		*	
Less than half the (A)(B). They ¬B	☞compset			*
Less than half the (A)(B). They C_{neutral}	☞refset			
Less than half the (A)(B). They C_{neutral}	compset			*
Less than ten (A)(B). They ¬B	☞refset		*	
Less than ten (A)(B). They ¬B	compset	*		*
Less than ten (A)(B). They C_{neutral}	☞refset			
Less than ten (A)(B). They C_{neutral}	compset	*		*

Figure 3.3: Tableau for the paradigm

this is to add the high-ranked constraint *avoid contradiction*.

Avoid Contradiction seems appropriate for accounting for the last-resort strategy necessary for resolving pronoun reference to compset. That is, in cases of complement anaphora, the refset interpretation violates *avoid contradiction*, while complement set reference does not. But what about cases like (3.26), repeated here as (3.66)?

(3.66) Most students went to the party.
They went to the beach instead.

We have to account for the fact that avoid contradiction does not force us to choose the unproblematic complement set interpretation of the plural pronoun. This is where we need emptiness'. I already hinted at the fact, that this constraint is part of the generator and thus ranks over all other constraints. It thus dictates when compset is available as an antecedent. In cases like (3.66), the fact that a complement set interpretation of the pronoun violates emptiness' leaves no choices left for avoid contradiction. In fact, we are forced to interpret (3.66) in its contradictory reading, which explains the infelicity of the continuation. (See de Hoop 2001 for more on optimality theory and unintelligibility.)

The effects of this ranking are illustrated in the tableau in 3.3. The subscripts *neutral* indicate that predicates C and B do not get a disjoint interpretation in sensible models. Contradictory continuations are indicated by predication $\neg B$. Emptiness' rules out all references to compset where compset is possibly empty. This only leaves the proportional downward entailing quantifiers. In general, refset reference is preferred, but this can be overruled in the non-neutral cases, where the predication in the continuation contradicts the predication in the antecedent sentence.

This optimality theoretic account will be the blue-print for the analysis to follow. It gives us the opportunity to separate some important issues. First of all, it illustrates that semantic considerations overrule a powerful pragmatic preference for anaphoric relations involving reference sets. Second, anaphoric reference following increasing or non-proportional quantifiers can only be resolved to the reference set or the maximal set even if this results in an inconsistent interpretation. The analysis should therefore not only account for the distribution of compset and refset reference; it should especially account for their different status in discourse.

3.3.7 Analysis

So far, we saw that there is a preference for interpreting pronouns as referring to the refset over interpreting them as referring to the compset. Semantics can force us to override this principle, but only if the complement set is available for reference in the first place. That is, while the reference set is always generated as one of the possible resolutions of plural pronouns, the complement set has to obey the non-emptiness requirement. In this section, we will focus on this contrast and propose that complement anaphora is a case of pronominal anaphora with a non-salient antecedent, subject to specific semantic and pragmatic constraints.

3.3.7.1 Pronouns and entailment

Since in cases of reference to quantificational NPs there is no coreference involved, it is difficult to predict the accessibility or the *cognitive status* (e.g. what Moxey and Sanford have called 'focus') of refset and compset. There is widespread consensus that subjects have a higher attentional status than direct objects and that direct objects in turn are more prominent than indirect objects, etc. However, neither compset nor refset was explicitly mentioned or predicated over, so no syntactic strengthening of salience applies. There is, however, an important semantic difference between accessing refset and compset using a pronoun. Reference to the reference set allows entailments which prior to the pronominal reference were not valid.

(3.67) Few of the students went to the party.
$\not\Rightarrow$
Some students went to the party.

(3.68) Few of the students went to the party.
They had a wonderful time.
$\Rightarrow$
Some students went to the party.

Support for (3.67) comes from the felicity of continuations expressing that there were no students at the party as in (3.69), this in contrast to comparable increasing examples like (3.70).

(3.69) Few of the students went to the party. In fact, none did.

(3.70) A few students went to the party. # In fact, none did.

A situation of no students attending the party is no longer a possibility after pronominal reference to the reference set, leading to the entailment in (3.68). Similar entailments fail when the pronoun attempts to refer to the complement set. Still, similar to the possibility of no students attending the party in the decreasing case, a sentence like *Most students went to the party* does not exclude the possibility of every student attending the party. A continuation like *In fact, all did* is felicitous, hence (3.71). But since pronominal reference to the complement set is out in increasing contexts, there is no case of pronominal reference to enforce the entailment that is impossible in (3.71). This is illustrated (once again) by (3.66).

(3.71) Most students went to the party.
$\not\Rightarrow$
Not all students went to the party.

(3.66) Most students went to the party.
They stayed at home instead.

In contrast to pronominal reference to the complement set, the definite description *the others* does enforce the entailment pattern missing from (3.71).

(3.72) Most students went to the party.
The others went to the beach instead.
$\Rightarrow$
Not all students went to the party.

3.3.7.2 Pronominal complement anaphora and inference

It is often claimed that pronouns are anaphorically *weak*. For instance, Ariel (1990) claims they are relatively *high accessibility markers*, meaning that pronouns signal to the addressee that their antecedent is easily accessible. In work on the givenness hierarchy (Gundel, Hedberg, and Zacharski 1993) pronouns are thought to need antecedents in *focus*, which according to the hierarchy is the highest attentional state a referent of a referring expression can be in. In general then, there is thought to be a one-to-one correspondence between a hierarchy of linguistic anaphoric forms and a scale of salience of the antecedent. Pronouns map to an antecedent with relatively high salience. In other words, the referent set of a quantificational structure is a salient antecedent. Pronouns referring to the reference set are ordinary pronouns whose antecedent has a high degree of salience. Pronominal complement anaphora are extra-ordinary pronouns. Their antecedent is not salient and the acceptability of the anaphoric link they contribute is both semantically and pragmatically constrained.

I will argue that complement anaphora link to their antecedents by inference. This explains why they are semantically constrained: only from quantificational structures that guarantee the non-emptiness of their complement set can we infer the existence of individuals satisfying the restrictor but not the scope. Furthermore, there is a striking similarity between pronominal complement anaphora and other pronouns with non-salient reference. I will justify the inference account of compset reference by drawing a parallel with these other uses of pronouns.

Let us consider some cases of non-salient reference, i.e. anaphoric relations which do not involve some linguistically available prominent antecedent, but which involve some kind of inference in order to derive an antecedent from the linguistic context. I first concentrate on examples where the anaphoric link obeys the correspondence between linguistic form and salience of the antecedent, meaning that the anaphor is a definite description. I will then show that it is possible to break with the correspondence and use pronouns instead of definite descriptions under some specific conditions.

The most straightforward example is subsectional anaphora.

(3.73) The children were having a lot of fun in the park. The boys played hide and seek and the girls were picking flowers.

(3.74) Tom, Susan and Mary went to a party. The girls went home early.

In (3.73), the definite descriptions *the boys* and *the girls* both refer to a subset of their (syntactic) antecedent *the children*. In (3.74), *the girls* refers to the two female conjuncts of the coordinated antecedent. These are fairly straightforward inferences, based on inclusion relations. A more complicated form of inference plays a role in the well-known bridging examples (Clark 1977).

(3.75) John's car doesn't run. He blew the engine.

(3.76) John walked into the room. The chandelier sparkled brightly.

In (3.75), *the engine* is associated with John's car. This anaphoric relation can be explained by making use of the world-knowledge (cars have engines). In (3.76), a more sophisticated inference is needed, since not every room has a chandelier, so the inference needs to be supported by the fact that chandeliers are *typically* associated with rooms.

Another example of anaphoric reference without explicit antecedents is reference to implicit arguments.

(3.77) John got a haircut.
The barber used a razor-blade. (ter Meulen 1999)

Although there is no explicit agent in the first sentence of (3.77), it supports an inference of the fact that *someone cut John's hair*. The definite description *the barber* can link to this inferred antecedent.

All these examples have in common that the antecedent involved in the anaphoric relation was not explicitly mentioned. This means that the use of definite descriptions in these examples was predicted by the hypothesis that non-pronominal anaphora are able to pick up antecedents with a low degree of salience. In general, the examples are unacceptable if we use pronouns instead of definite descriptions.

(3.78) Peter$_i$ and Harold$_j$ met in Utrecht. He$_{\#i/\#j}$ loves that town.

(3.79) Peter$_i$ met Harold$_j$ in Utrecht. He$_{i/j}$ loves that town.

(3.80) The children had a lot of fun in the park. They picked flowers.

(3.81) We cannot use John's car. # All four of them are flat.

(3.82) # After John entered the room, it came crashing down on him.

(3.83) # John got a hair-cut. He used a razor-blade.

The pronoun in (3.78) is infelicitous, while the pronoun in (3.79), with Peter and Harold both available as singular arguments is ambiguous. The plural pronoun in (3.80) cannot refer to the subset of girls in the set of children having fun in the park. Although there is an impressive amount of semantic support, the second sentence in (3.81) is out. The same applies to (3.82). Finally, pronominal reference to implicit arguments is also impossible. The continuation in (3.83) is incoherent, for the pronoun can only be resolved to John.

Interestingly though, there are special cases where pronominal reference to inferred antecedents is felicitous. Consider the following examples.[15]

(3.84) Peter and Susan met in Utrecht. He loves the town, but she thinks it's too small.

(3.85) My right-door neighbours make a lot of noise. He plays the drums and she keeps on shouting at him.

(3.86) John kept on staring at the newly-wed couple. She resembled a childhood sweetheart of his.

(3.87) The priest was tortured for days. They wanted him to reveal where the insurgents were hiding out. (Mauner 1996)

[15] Hendriks and Dekker (1995) argue that the stress which is obligatory on the pronouns in (3.84) and (3.85) is the result of the fact these are cases of non-monotone anaphora: "Linkhood (marked by L+H* accent in English) serves to signal non-monotone anaphora. If an expression is a link, then its discourse referent Y is anaphoric to an antecedent discourse referent X such that $X \not\subseteq Y$" (Hendriks and Dekker 1995, p. 353.) In this respect, it is interesting to note that the pronoun in (3.86) does not get stressed at all.

In (3.84), both conjuncts are pronominally accessible. The example in (3.85), based on an example from Hendriks and Dekker (1995), contains two singular pronominal subsectional anaphora. In (3.86), there is a bridge to the bride and in (3.87), the implicit argument *the torturers* is realised by the pronoun *they*.

There are a number of conditions which need to be satisfied in order for these extraordinary uses of pronouns to be licenced. The first and most straightforward necessary condition is that the antecedent has to be *inferable*. This is of course a general condition which also applies to the examples with anaphoric relations involving definite descriptions. In (3.88), although there is an implicit argument, the fact that the negation operator scopes over it makes the inference *someone cut a student's hair* impossible. The definite description in the second sentence of (3.88) has to be accommodated globally.

(3.88) No student got his hair cut.
The barber was on vacation. (ter Meulen 1999)

Another necessary condition is uniqueness: there should be no other similarly inferable entities with the same semantic features. For singular pronouns this means uniqueness of gender for the inferred object. In the plural case, this translates to maximality. Since English does not specify gender on plural pronouns, the prediction is that plural subsectional anaphora in English will never be pronominalised. There would be no unique way of identifying which part of the antecedent is referred to.[16] The same goes for reference to parts of conjunctions (which is a special case of subsectional anaphora) and bridging to parts (like (3.86)). Pronominal reference to implicit arguments always obeys the uniqueness condition.

Apart from the uniqueness of the inference, there is a condition on the effort involved. In fact, it seems that only inferences using semantic information are allowed. The examples in (3.84)–(3.87), I claim, have in common that they allow the inference necessary for the pronominal reference to occur on the basis of the semantic representation of the antecedent sentence only. In (3.84), both Peter and Susan are part of the semantic representation, even though the infelicity of pronominal reference to either of the conjuncts in (3.78) implies that they are not realised as *discourse referents*. The example (3.84), therefore, seems to suggest that semantic representation of an entity as a part of a plurality makes this entity inferable.

This begs the question, however, why (3.86) is felicitous, since it does not semantically represent a plurality: the definite *the newly-wed couple* is singular. It is clear, however, that at some semantic level this NP is represented as a plurality, witness the plural agreement it sometimes triggers on coreferring pronouns. Consider the following examples.

[16] It also predicts that languages that *do* specify gender on plural pronouns allow for pronominal plural subsectional anaphora. An informal inquiry into the status of such examples in French and Spanish suggests this is indeed the case.

(3.89) The newly-weds were both nervous about the wedding night.

(3.90) *The newly-wed couple were/was both nervous about the wedding night.

(3.91) Grain is a symbol of both fertility and prosperity. Symbolically, throwing rice (or other grains) at the newly wed couple is a means of wishing them a lifetime of those blessings.[17]

The implicit semantic encoding of plurality in bunch-denoting singulars is well-known. While ordinary plurality tests like distributivity fail (see (3.89)/(3.90)), there are some traces of plurality on other levels, as is illustrated in (3.91); see Schwarzschild 1996, chapter nine, for discussion.

Whenever the inference involved is slightly more complicated and in need of information which is not semantically present, pronominal reference is no longer an option. For instance in the bridging examples, although there is a unique jockey associated with the race horse in (3.92), this needs to be inferred using more than the information semantically available from the first sentence.

(3.92) The race horse suffered a lot. During the race the jockey/# he whipped it constantly.

Alternatively, one could argue that the inference in (3.92) is not unique, since *the jockey* is not the unique masculine individual to be associated with the race horse (there is e.g. *its trainer*, *its owner*, *its veterinarian*.) Descriptions, then, narrow all the associations down to a single possible one. This suggests that since the application of world-knowledge is in principle unbounded, the semantic specification that comes with a pronoun is necessarily too weak to satisfy the uniqueness requirement. The condition asking for minimal effort thus follows from uniqueness.

Another condition for pronominal inferable reference is that the anaphoric link supports discourse coherence.[18] The examples above all have a natural coherence relation, improving the link between the first and second sentence. But now consider:

(3.93) # My right-door neighbours make a lot of noise and I met her yesterday.

(3.94) # John kept on staring at the newly-weds. She was thinking about the flowers.

(3.95) # The priest was tortured for days. They went for beers everyday.

---

[17]This is a naturally occurring example from the world-wide-web.

[18]See Asher and Lascarides 1999 for an explicit claim that bridging is directly linked to rhetorical relations.

In (3.93), the link between the pronoun *her* and one of the right-door neighbours is not an option, since there is no support for discourse coherence. The same applies to (3.94) and (3.95). Notice that there is an intuitive difference between an example like (3.93) and (3.96), below. Both are incoherent, but in contrast to (3.93), in (3.96) the plural pronoun *is* resolved to refer to the right-door neighbours.

(3.96) My right-door neighbours make a lot of noise and I met them yesterday.

Summarising, there are four important conditions on inferred pronominal anaphoric reference: (i) inferability, (ii) uniqueness, (iii) use of semantically available information only and (iv) support of discourse coherence by the anaphoric link.

The parallel with complement anaphora is straightforward. Pronominal reference to the complement set is, in general, infelicitous. The exceptional cases are semantically constrained by inferability (condition (i)). The emptiness' constraint from the optimality theory approach translates to checking whether or not the following inference holds: $D(A)(B) \rightarrow \exists X(A(X) \wedge \neg B(X))$. Such an inference is impossible in examples like (3.54), repeated here.

(3.54) More than 10% of the students went to the party.
They went to the beach instead.

Furthermore, complement anaphora are only *semantically* inferred. That is, the effort involved in the inference is minimal (condition (iii)). The pronoun in (3.54) cannot refer to the complement set as a consequence of the implicature that not all students went to the party invoked by the first sentence (as was noted in section 3.3.3) and/or the fact that this set would be a plausible candidate for the predication in the second sentence.

Uniqueness (condition (ii)) is ensured with pronominal complement set reference by the fact that, just like pronouns referring to refset, they are maximal.[19] The continuation in (3.97) cannot mean that some of the MPs who did not attend the meeting were too busy, and some other non-attending MPs had other excuses.

(3.97) Few of the MPs attended the meeting. They were too busy.

The final condition which parallels (other) antecedentless anaphora is support for discourse coherence. Moxey and Sanford already reported a link to rhetorical structure. Moreover, we concluded in section 3.3.5 that semantic support is a necessary condition for pronominal complement anaphora. In

[19]Note that uniqueness should not be interpreted here as uniqueness of antecedent, since both the reference set and the maximal set remain alternative referents for the pronoun to pick up. Rather, once an anaphoric link is made, this link should point to a unique individual or, in the case of plurals, to a maximal one.

fact, we could interpret the *avoid contradiction*constraint of section 3.3.6 as modelling the fact that the complement set (once inferable) is only used as a resolution for pronominal reference once there is a semantic reason to do so. If there is no such reason, the resolution process will pick the reference set which is the default reference for plural pronouns linked to quantificational structures, as modelled by *forward directionality* in the optimality theoretic analysis and the triggered abstraction rule presented above.

In sum, I conclude that pronominal complement anaphora should not be viewed as alternative e-type pronouns, but as extra-ordinary pronouns whose antecedent is to be inferred. Inferability accounts for the strict semantic nature of the distribution of complement anaphora, while pronominal reference to the reference set is furthermore constrained by conditions that apply to other anaphora with inferred antecedents as well.

3.3.7.3 A note on 'The others'

In this section, I discuss non-pronominal complement anaphora. Note that the possibility of explicit reference to the compset using a definite description in all contexts, does not involve a referent for compset. In fact, a closer look at the semantics of *the others* shows us that it contains two anaphoric parts, a domain (the usual context variable which gives definite descriptions their anaphoricity) and an anaphoric description *others*. In (3.72), repeated here as (3.98), *the others* is anaphoric to the maximal set, while its description complements the reference set.

(3.98) [[Most$_j$ [students]$_i$] went to the party].
The others$_{i,j}$ went to the beach instead.
$\Rightarrow$
Not all students went to the party.

This explains why a definite description making explicit reference to the complement set *can* and a pronoun can*not* accommodate the existence of a set of individuals satisfying restrictor but not scope. Use of *the others* as in (3.98) thus results in a DRS as in figure 3.4. The referent Z is introduced by abstraction which was triggered by *the others* being coindexed with the reference set.[20] The description *others* collects the maximal difference between Z and Y, the referents corresponding to the two indices of *the others*. Since the referent X is existentially closed, the complement set is accommodated (accounting for the entailment in (3.72)). The same story applies to a description as *the students who did not go to the party* which is also anaphoric to the maximal set, but here the description itself is not anaphoric. For other definite descriptions there is no option for complement set

[20] Alternatively, the referent Y could be derived in the same way, if one does not choose to have this referent introduced by the quantification itself.

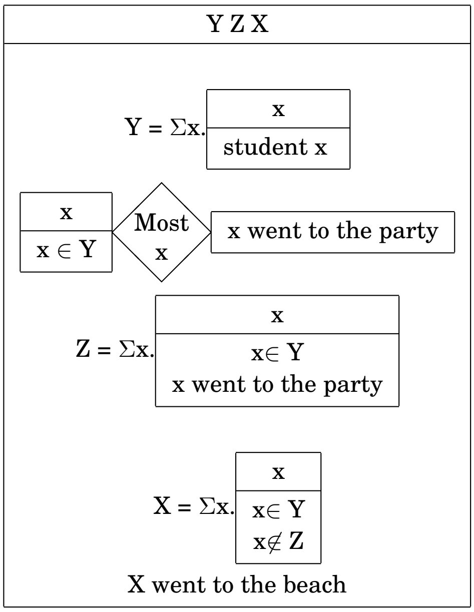


Figure 3.4: DRS for (3.98)

reference if the antecedent is not right monotone decreasing and proportional. This is due to the general condition of inferability on complement anaphora. This is illustrated below. The use of an epithet is felicitous in the downward entailing case (3.99), but not in (3.100).

(3.99) Less than half of the students went to the party.
The bastards went to the beach instead.

(3.100) More than half of the students went to the party.
The bastards went to the beach instead.

3.3.8 Taking stock

I have shown that plural pronominal reference, although extremely flexible, ordinarily refers to the reference set of quantification, since this set is made most salient by a determiner, independent of what set is focused on by the determiner.

I have argued that from the interpretation perspective several principles govern the distribution of complement anaphora. Pragmatically, there is a preference for reference set reference over complement set reference. This preference can be overruled by semantic considerations. Crucial, however, is the fact that the possibility of complement set reference follows from inferability. Once the existence of individuals satisfying restrictor but not scope cannot be inferred from the antecedent quantificational sentence, pronominal reference to compset is not an option, regardless of the semantic or pragmatic context. In contrast to pronouns referring to the reference set, pronominal complement anaphora belong to the class of cases in which pronouns select an antecedent which is not salient, but has to be inferred. As such, they are subject to the strict constraints that govern such marked usages of pronouns.

Finally, let me suggest that the present approach to complement anaphora seems to explain why so many people have been hesitant to accept Moxey and Sanford's experimental results. Many have thought complement anaphora to be not quite felicitous. The account presented here, where complement anaphora are considered to be an extra-ordinary case of anaphora, might (at least partly) explain where this discomfort with pronominal reference to the complement set comes from.

3.4 Context and pronominal reference to sets

In sum, in this chapter I showed the following. First, every determiner imposes high salience on its referent set. Second, the complement set is, in principle inaccessible for pronouns, but may form the antecedent of a pronoun when it serves some salient communicative purpose and no other antecedents for the pronoun result in a consistent discourse. Third, the maximal set is presupposed to be contextually given in the case of strong, i.e. non-intersective determiners. Weak noun phrases do not carry such a presupposition and, since the intersectivity property makes the truth-conditions independent of the full quantificational domain this set will not be accessible in discourse.

I take the salience of the reference set to mean that every determiner introduces a referent for this set. Similarly, the general inaccessibility of the complement set suggests that no referent exists for this set. This means that the marked case of pronominal complement anaphora is modeled by a marked operation on context, inferring the existence of a complement set and introducing a referent for it. In the cases the complement set is available as an antecedent this is due to an inference and a last resort choice of the hearer and not due to the interpretation of the antecedent quantification structure. In what follows, I therefore assume that the existence of complement anaphora has to be accommodated locally: it should not follow from a semantics which expresses how context is built up in discourse. As for the maximal set, the presupposition that comes with strong quantifiers is interpreted as a presupposition of the existence of a referent for this set in context. Weak quantifiers have no such presupposition, nor do they introduce a referent for their domain by themselves.

Chapter 4

Dynamic Semantics for Plurals

In the previous chapter, we saw that following a sentence $(D(A))(B)$, the reference set, i.e. the set of individuals satisfying both restrictor and scope, is made salient enough to guarantee unconditional pronominal anaphora. We focus now on what is desired of a semantic formalism in order for it to deal with quantification and anaphora with quantificational antecedents, keeping in mind that the e-type pronoun is a regular pronoun, picking up a salient antecedent.

As we have seen, the reference set is sometimes involved in dependencies with other sets that were created during quantification. The standard example is repeated here as (4.1), where the reference of the singular pronoun co-varies with the quantification over the three students.

(4.1) Three students (each) wrote a paper. They (each) sent it to L&P.

In DRT, such examples lead to the proposal of a copy-mechanism that gives the singular pronoun access to the paper written by each student. One of the advantages that dynamic semantic theories have over DRT with respect to plural anaphora is that, given a sufficiently rich notion of context, they are able to account for the intuition that quantificational sentences do not only change the context by adding maximal plural individuals but also by recording dependencies between such individuals. That is, apart from keeping track of subjects that were introduced in discourse, context can contain other kinds of anaphoric resources like, for instance, a functional dependence that exists between two subjects.

In this chapter, I discuss three proposals which have in common that they use a more fine-grained notion of context than do dynamic semantic analyses of singular anaphora. I concentrate on just these three, since they all focus on accounting for maximal reference anaphora as well as the dependencies they are involved in. The proposals are: van den Berg's *dynamic plural logic* which uses *information states for plurals* (see e.g. van den Berg 1990, 1993, 1996b, 1996a); Krifka's semantics for plural anaphora based on *parametrised sum individuals* (see Krifka 1996a); and Elworthy's *theory of anaphoric information* which uses *discourse sets* (see Elworthy 1992, 1995). After an introduction to these frameworks, I compare them with respect to the specific notion of context they use. I discuss how the data-structures involved in these proposals interact with the semantics and how they compare. The result of the discussion will be that the information states for plurals in van den Berg's dynamic plural logic appear to be the most expressive way of representing dependencies in context. Moreover, apart from sharing some common successes all three proposals leave part of the data unaccounted for.

4.1 Dynamic Plural Logic

As far as I know, van den Berg (1990) was the first paper to propose a dynamic semantics with an alternative notion of context for the purpose of dealing with plural anaphora. Van den Berg refines and extends his semantics in a number of papers (e.g. van den Berg 1993), but I take it that the detailed presentation of a *dynamic plural logic* in his dissertation (1996b) and in the paper van den Berg 1996a is his final proposal. I will abbreviate dynamic plural logic as 'DPlL'.

The most important features of van den Berg's work are, in my opinion, the following.

(4.2)
a. The interpretation of a formula in dynamic plural logic yields a relation between information states.
b. Information states are sets of assignment functions.
c. Assignment functions are partial.
d. Assignment functions only have atomic individuals in their range.
e. Dependencies encoded in information states are due to distributivity.

Dynamic plural logic was inspired by Groenendijk and Stokhof's DPL, but instead of having the interpretation function '$[\![\;]\!]$' expressing a function from formulae to relations between assignments, van den Berg lets it map formulae to relations between sets of partial assignment functions. These sets are called *information states for plurals*. For instance, a possible information state in dynamic plural logic is the following:

(4.3) $\{\{\langle x, j\rangle, \langle y, t\rangle\}, \{\langle x, s\rangle, \langle y, d\rangle\}, \{\langle y, h\rangle\}\}$

I will represent such structures as matrices, where each row corresponds to an assignment function and each column to a variable. The information state made up by the three functions in (4.3) is depicted in (4.4).

(4.4)

	x	y
f_1	j	t
f_2	s	d
f_3		h

In this information state, x is 'distributively' assigned the plurality $\{j, s\}$ (say, John and Susan) and y the plurality $\{t, d, h\}$ (say, Tom, Dick and Harry). At the same time, it expresses that there is some sort of relation between John and Tom (as expressed by f_1), and between Susan and Dick (as expressed by f_2). The third assignment function, f_3, is not defined for x. Following Van den Berg, we will henceforth represent undefinedness using the value '$\star$'. Each variable for which a partial assignment is undefined will be assigned the undefinedness value, as in (4.5). This way, we can treat partial assignments as if they were total.[1]

(4.5)

	x	y
f_1	j	t
f_2	s	d
f_3	$\star$	h

Van den Berg has several reasons for choosing partial assignments for DPlL. The most important is that it allows for a more intuitive notion of information increase. The initial state is undefined for all variables. Once noun phrases are encountered, more and more variables are assigned values. In other words, as the discourse proceeds, the domain of the information state increases. Another reason for choosing partial assignments, more relevant to plurality, is that it allows a notion of subset which is sensitive to dependencies in the state. This is best explained by an example. Consider a state in which x and y functionally depend on one another.

(4.6)

	x	y
f_1	j	t
f_2	s	d
f_3	m	h

Say now that we are interested in a subset of the values assigned to x, for instance, $\{j, s\}$ and we want the variable z to range over this subset. Two possible ways to go would be the states in (4.7) and (4.8).

(4.7)

	x	y	z
f_1	j	t	j
f_2	s	d	s
f_3	m	h	s

[1] Notice that consequently our method of graphically depicting information states *is* partial, since it only reveals the relevant (i.e. defined) snapshot of a state.

(4.8)

	x	y	z
f_1	j	t	j
f_2	s	d	s
f_3	m	h	s
f_4	m	h	j

Although both these state assigns the set $\{j, s\}$ to z they also create dependencies which were not in the original state. For instance, (4.7) registers a dependence between Susan and Harry and Susan and Mary, which was not in the original state. The state in (4.8) has an additional dependence between John and Harry and Mary. A much more realistic option is to use a dummy value (in this case, the undefinedness value '$\star$') to express the lacking individual:

(4.9)

	x	y	z
f_1	j	t	j
f_2	s	d	s
f_3	m	h	$\star$

The range of partial assignment functions inside an information state contains singulars only. The pros and cons of this design choice will be discussed below. For now, it suffices to illustrate how pluralities are built up from the atoms and how this kind of information state is capable of storing (functional) dependencies between variables. In all the example states above, x and y are said to be *dependent*. This is a formal notion and it expresses that the way collections are spread over the assignments in the state is not arbitrary.

Definition 4.2

Dependence

In information state G, y is dependent on x if:
$\exists d, e \in G(x) : G|_{x=d}(y) \neq G|_{x=e}(y)$

□

Definition 4.3

Projection

$$G(x) \ := \ \{g(x) | g \in G \ \& \ g(x) \neq \star\}$$

□

Definition 4.4

Sub-states

$$G|_{x=d} \;:=\; \{g \in G | g(x) = d\}$$

□

The definition of dependence makes use of special operations on information states: $G(x)$ collects the values assigned to x in the assignment functions in G and $G|_{x=d}$ is the sub-state of G wherein all assignment functions assign d to x. Dependence expresses that there exists a pair of sub-states assigning different values to x which also differ at y.

Let us now turn to how interpretation interacts with dependence. In van den Berg's work, the logic itself is also partial, that is, it is three-valued. The interpretation function relates to a pair of information states either by truth, falsity or undefinedness ('$\star$'). This is necessary, since the assignment functions are partial: they might not always provide a value for all the relevant variables for the interpretation of a formula. A state like the one in (4.9) is irrelevant for some formulae containing 'v', since it provides no value for v. In those cases, interpretation is undefined.

The notation $G\,[\![\varphi]\!]^{\,d}H$ means that the relation corresponding to φ is defined with respect to input G and output H. We use $G\,[\![\varphi]\!]^{\,+}H$ to express that the relation is true for G and H. Finally, falsehood is written as $G\,[\![\varphi]\!]^{\,-}H$.

The notion of dependence given above is interesting with respect to the existential quantifier van den Berg defines. The conditions for the existential quantifier are:[2]

(4.10)
$$\begin{array}{lll} G\,[\![\epsilon_x]\!]^{\,d}H & \text{iff} & G(x) = \emptyset \;\&\; \exists X : H = \{g[x/d] | g \in G \;\&\; d \in X\} \\ G\,[\![\epsilon_x]\!]^{\,+}H & \text{iff} & G\,[\![\epsilon_x]\!]^{\,d}H \\ G\,[\![\epsilon_x]\!]^{\,-}H & \text{iff} & \bot \end{array}$$

This definition is similar to the random assignment relation in DPL. It introduces all values in X as values for x without introducing any dependencies on other variables that may occur in G. In fact, this defines ϵ_x as randomly assigning pluralities to x. In a domain with three individuals (say, t, d and h) and a language with just one variable (namely x), the following seven relations are covered by the interpretation. (Here and in forthcoming representations, we do not give names to the individual functions for notational convenience.)

[2]The clause '$G(x) = \emptyset$' ensures that ϵ_x always *increases* information. The interpretation can only be true once, in the input set G, there is no assignment function which provides a proper value for x. Thus, existent values for x can never be overwritten. See chapter 5 for more on such considerations.

(4.11)

$$\begin{array}{|c|}\multicolumn{1}{c}{x}\\\hline \star\\\hline\end{array}\;[\![\epsilon_x]\!]^{+}\begin{array}{|c|}\multicolumn{1}{c}{x}\\\hline t\\\hline\end{array}\qquad\begin{array}{|c|}\multicolumn{1}{c}{x}\\\hline \star\\\hline\end{array}\;[\![\epsilon_x]\!]^{+}\begin{array}{|c|}\multicolumn{1}{c}{x}\\\hline d\\\hline\end{array}\qquad\begin{array}{|c|}\multicolumn{1}{c}{x}\\\hline \star\\\hline\end{array}\;[\![\epsilon_x]\!]^{+}\begin{array}{|c|}\multicolumn{1}{c}{x}\\\hline h\\\hline\end{array}$$

$$\begin{array}{|c|}\multicolumn{1}{c}{x}\\\hline \star\\\hline\end{array}\;[\![\epsilon_x]\!]^{+}\begin{array}{|c|}\multicolumn{1}{c}{x}\\\hline t\\\hline d\\\hline\end{array}\qquad\begin{array}{|c|}\multicolumn{1}{c}{x}\\\hline \star\\\hline\end{array}\;[\![\epsilon_x]\!]^{+}\begin{array}{|c|}\multicolumn{1}{c}{x}\\\hline t\\\hline h\\\hline\end{array}\qquad\begin{array}{|c|}\multicolumn{1}{c}{x}\\\hline \star\\\hline\end{array}\;[\![\epsilon_x]\!]^{+}\begin{array}{|c|}\multicolumn{1}{c}{x}\\\hline d\\\hline h\\\hline\end{array}$$

$$\begin{array}{|c|}\multicolumn{1}{c}{x}\\\hline \star\\\hline\end{array}\;[\![\epsilon_x]\!]^{+}\begin{array}{|c|}\multicolumn{1}{c}{x}\\\hline t\\\hline d\\\hline h\\\hline\end{array}$$

In other words, following ϵ_x the possible values for x are the elements in the powerset of the domain of entities, excluding the empty value. Arguably, 'X' could be the empty set in which case, an additional pair is in the relation namely where ϵ_x leaves x undefined.

The introduction of variables in dynamic plural logic does by itself never introduce functional dependencies. This can be understood by studying the way the ϵ-operator builds on the input state. Each assignment in the input state branches into a set of new assignment functions that collectively assign the introduced entity X. For instance, interpreting ϵ_x in a language with two variables, namely x and y, one pair in the relation is one in which x is assigned the plurality $\{t, h\}$.

(4.12)

$$\begin{array}{|c|}\multicolumn{1}{c}{y}\\\hline d\\\hline h\\\hline\end{array}\;[\![\epsilon_x]\!]\;\begin{array}{|c|c|}\multicolumn{1}{c}{y}&\multicolumn{1}{c}{x}\\\hline d&t\\\hline d&h\\\hline h&t\\\hline h&h\\\hline\end{array}$$

It is easy to see that in the output state in (4.12), x is not dependent on y. For both d and h are paired with both t and h.

The default mode of interpretation is collective. That is, predication is evaluated with respect to the collection of values assigned to variables in the assignment functions collected in a state. Just as in DPL, predicates are tests. Note the definedness condition. A predication over x_i is defined even if x_i is undefined in the incoming state. The reason is that the plural projection of this variable $G(x_i)$ is *not* undefined. It is simply the empty set.

(4.13)
$$\begin{aligned}
&G\,[\![Px_1,\ldots,x_n]\!]^{d}H \;\Leftrightarrow\; G = H\\
&G\,[\![Px_1,\ldots,x_n]\!]^{+}H \;\Leftrightarrow\; G = H \;\&\; \langle G(x_1),\ldots,G(x_n)\rangle \in I(P)\\
&G\,[\![Px_1,\ldots,x_n]\!]^{-}H \;\Leftrightarrow\; G\,[\![Px_1,\ldots,x_n]\!]^{d}H \;\&\; \neg G\,[\![Px_1,\ldots,x_n]\!]^{+}H
\end{aligned}$$

Since the existential quantifier in van den Berg's formalism is by itself not capable of introducing dependence relations, collective interpretation will not cause any dependencies at all. Distributivity, however, *is* a source of dependencies. Van den Berg uses an operator on formulae to break with the default collective interpretation. Dependencies arise since whatever is in the scope of a distributivity operator is evaluated with respect to subsets of the original input state.

(4.14)
$$\begin{array}{lcl} G\,[\![\delta_x(\varphi)]\!]^{\,d}H & \Leftrightarrow & \forall d \in G(x) : G|_{x=d}\,[\![\varphi]\!]^{\,d}H|_{x=d} \&\ G|_{x=\star} = H|_{x=\star} \\ G\,[\![\delta_x(\varphi)]\!]^{\,+}H & \Leftrightarrow & G\,[\![\delta_x(\varphi)]\!]^{\,d}H\ \& \\ & & \forall d \in G(x) : G|_{x=d}\,[\![\varphi]\!]^{\,+}H|_{x=d} \\ G\,[\![\delta_x(\varphi)]\!]^{\,-}H & \Leftrightarrow & G\,[\![\delta_x(\varphi)]\!]^{\,d}H\ \&\ \neg G\,[\![\delta_x(\varphi)]\!]^{\,+}H \end{array}$$

Recall that the sub-state $G|_{x=d}$ is the set of assignments in G that assign d to x. For each (atomic) individual in the collective assignment of G to x, the scope of the distributivity operator (φ) is evaluated. This means that if in φ new values are assigned to variables, these may differ from sub-state to sub-state. In other words, variables introduced in the scope of δ_x may be dependent on x.

A sentence like 'Every man loves a woman' produces an information state wherein each assignment f is such that $f(x)$ is a man loving woman $f(y)$. Once such a context is created, pronouns can profit from this by re-entering the structure using distributivity again.

For instance, the formula in (4.15) can only be true if the pairs of atoms that satisfy R also satisfy S. In other words, R and S have to contain the same pairs (with respect to the values for x and y) in order to verify (4.15). This is in contrast to (4.16) which merely expresses they involve the same sets.

(4.15) $\epsilon_x \wedge \delta_x(\epsilon_y \wedge R(x,y)) \wedge \delta_x(S(x,y))$

(4.16) $\epsilon_x \wedge \epsilon_y \wedge R(x,y) \wedge S(x,y)$

A critical note is in order if we want to seriously consider using van den Berg's system. There is a peculiarity concerning the distribution operator. What do the conditions in (4.14) for two sets of assignment G and H to truthfully stand in the relation $[\![\delta_x(\varphi)]\!]^{\,+}$ tell us about H? Well, with respect to the respective values $d \in G(x)$, subsets $H|_{x=d}$ are outputs of processing φ with respect to $G|_{x=d}$. Moreover, we know that $H|_{x=\star} = G|_{x=\star}$. No other constraints are given and thus the above definition does not exclude H to contain values for x which are not in $G(x)$. For instance, distributing over three boys assigned to x in G could result in an output state H, wherein $H(x)$ contains the three boys as a proper subset.

In order to constrain the output state H, we thus need to state that the values for x in G are exactly the values in $H(x)$ (given that the undefinedness value is not included in projection). So, I redefine distributivity as follows:

(4.17)
$$\begin{array}{lcl} G\,[\![\delta_x(\varphi)]\!]^{\,d}H & \Leftrightarrow & G(x) = H(x)\ \&\ G|_{x=\star} = H|_{x=\star}\ \& \\ & & \forall d \in G(x) : G|_{x=d}\,[\![\varphi]\!]^{\,d}H|_{x=d} \\ G\,[\![\delta_x(\varphi)]\!]^{\,+}H & \Leftrightarrow & G\,[\![\delta_x(\varphi)]\!]^{\,d}H\ \&\ \forall d \in G(x) : G|_{x=d}\,[\![\varphi]\!]^{\,+}H|_{x=d} \\ G\,[\![\delta_x(\varphi)]\!]^{\,-}H & \Leftrightarrow & G\,[\![\delta_x(\varphi)]\!]^{\,d}H\ \&\ \neg G\,[\![\delta_x(\varphi)]\!]^{\,+}H \end{array}$$

Let me illustrate the merits of dynamic plural logic using our running example in (4.18). The DPlL translation is in (4.19)[3].

(4.18) Three students each wrote a paper. They each sent it to L&P.

(4.19) $\epsilon_x \wedge \text{STUDENT}(x) \wedge 3(x) \wedge \delta_x(\epsilon_y \wedge 1(y) \wedge \text{PAPER}(y) \wedge \text{WROTE}(x, y))$
$\wedge \delta_x(\text{SENT\_TO\_LP}(x, y))$

Say now that given an initial state, an update with $\epsilon_x \wedge \text{STUDENT}(x) \wedge 3(x)$ will result in a state like in (4.20)[4].

(4.20)

x
s_1
s_2
s_3

The distributivity operator will now consider sub-states of this state for the interpretation of its scope. The VP translation '$\epsilon_y \wedge 1(y) \wedge \text{PAPER}(y) \wedge \text{WROTE}(x, y)$' will be interpreted three times. This results in three output states that form the sub-states of the output for the distribution. Say that s_1 wrote paper p_1, that s_2 wrote p_2 and s_3 wrote p_3, then:

(4.21)

$$\begin{array}{c} x \\ \boxed{s_1} \end{array} [\![\epsilon_y \wedge 1(y) \wedge \text{PAPER}(y) \wedge \text{WROTE}(x, y)]\!]^+ \begin{array}{cc} x & y \\ \hline \multicolumn{1}{|c|}{s_1} & \multicolumn{1}{c|}{p_1} \\ \hline \end{array}$$

$$\begin{array}{c} x \\ \boxed{s_2} \end{array} [\![\epsilon_y \wedge 1(y) \wedge \text{PAPER}(y) \wedge \text{WROTE}(x, y)]\!]^+ \begin{array}{cc} x & y \\ \hline \multicolumn{1}{|c|}{s_2} & \multicolumn{1}{c|}{p_2} \\ \hline \end{array}$$

$$\begin{array}{c} x \\ \boxed{s_3} \end{array} [\![\epsilon_y \wedge 1(y) \wedge \text{PAPER}(y) \wedge \text{WROTE}(x, y)]\!]^+ \begin{array}{cc} x & y \\ \hline \multicolumn{1}{|c|}{s_3} & \multicolumn{1}{c|}{p_3} \\ \hline \end{array}$$

and the output state after processing the first sentence is then:

(4.22)

x	y
s_1	p_1
s_2	p_2
s_3	p_3

The second sentence in (4.18) now forces a new distribution over x (if we recognise that the plural pronoun is co-indexed with 'three students'). The sub-states that are considered now, however, contain not only the atomic values for x, but also those for y. So, in case each of the students indeed sent his or her paper to Linguistics and Philosophy:

[3]Here, I assume that 'three' is analysed as a simple predicate '3' which counts the atoms in a set.

[4]Actually, such a formula results in a collection of states containing three students, given the indefiniteness of 'three students'. For simplicity, we will consider only one of these potential output states.

$$(4.23)\quad \begin{array}{l} \begin{array}{cc} x & y \\ \hline \multicolumn{1}{|c|}{s_1} & \multicolumn{1}{c|}{p_1} \\ \hline \end{array} \; [\![\text{SENT\_TO\_LP}(x,y)]\!]^{+} \; \begin{array}{cc} x & y \\ \hline \multicolumn{1}{|c|}{s_1} & \multicolumn{1}{c|}{p_1} \\ \hline \end{array} \\ \begin{array}{cc} x & y \\ \hline \multicolumn{1}{|c|}{s_2} & \multicolumn{1}{c|}{p_2} \\ \hline \end{array} \; [\![\text{SENT\_TO\_LP}(x,y)]\!]^{+} \; \begin{array}{cc} x & y \\ \hline \multicolumn{1}{|c|}{s_2} & \multicolumn{1}{c|}{p_2} \\ \hline \end{array} \\ \begin{array}{cc} x & y \\ \hline \multicolumn{1}{|c|}{s_3} & \multicolumn{1}{c|}{p_3} \\ \hline \end{array} \; [\![\text{SENT\_TO\_LP}(x,y)]\!]^{+} \; \begin{array}{cc} x & y \\ \hline \multicolumn{1}{|c|}{s_3} & \multicolumn{1}{c|}{p_3} \\ \hline \end{array} \end{array}$$

The output state for distribution is again (4.22).

Independent existential quantification, standard collective predication and a distributivity operator creating and accessing dependencies form the backbone of dynamic plural logic's treatment of plural anaphora and are my main focus in this chapter. Let me briefly discuss van den Berg's approach to quantificational noun phrases, however.

Van den Berg defines some more operators needed to come to an analysis of quantificational mechanisms. Since Van den Berg assumes that quantification is externally dynamic, he needs to make sure that quantifiers introduce only the sets suitable for subsequent anaphoric reference. Most of the work is done by a maximality operator 'M'. It has often been observed that maximality plays an important role in quantification, not least because of the 'maximal' interpretation of e-type pronouns. In dynamic plural logic the interpretation of a formula is maximised relative to a variable x by finding an output state H such that no alternative output state exists which assigns a superset of $H(x)$ to x. I will not discuss the complex truth and falsity conditions for such a relation, here. (See van den Berg 1996b, p. 141 for details).

In a quantificational structure, both the referent for the maxset and for the reference set are introduced using the maximalisation operator. The maxset is moreover presupposed. Van den Berg defines an operator '+' to express this presupposition. That is, '$+(\varphi)$' is only defined for an input state and an output state if φ is true relative to these states. The blue-print for quantification is then:

(4.24) $\epsilon_{x'} \wedge \epsilon_x \wedge +\mathsf{M}(\varphi) \wedge \mathsf{M}(x \leq x' \wedge \psi) \wedge Q(x', x)$

Here, φ represents the restrictor clause and ψ the nuclear scope of quantification, while x' is the maximal set and x is the reference set. 'Q' is the quantificational relation expressed by the determiner. The condition $x \leq x'$ expresses that the value for x' is a subset of the value of x. This way, conservativity is explicit in the representation.

The final ingredient for quantification is a notion of domain restriction. I remarked earlier that an important reason for van den Berg's choice for partial assignment functions is because of the need of a 'safe' notion of subset. The dependency preserving nature of this notion is important for quantification. Van den Berg follows Westerståhl (1984) in assuming that quantification always occurs relative to some context set which establishes the global domain of quantification. That is, if $Q(A)(B)$ is interpreted with respect to context set C, it is interpreted as $Q(A \cap C, B)$. Also, definite descriptions can take such sets as domains. After introducing a set of children with an NP like ‘six children’, a definite description like ‘the boys’ can access the boys among the children. Moreover, it also accesses the dependencies the children partake in. For instance, in (4.25), the NP ‘the boys’ takes the maximal set of boys among the children and claims that each of them took the dollar he found home.

(4.25) Six children each found a dollar in the park. The boys took it home.

To deal with these observations, Van den Berg relativises his ϵ-operator and his maximality operator to a referent (thought to express the context set). For instance, $\epsilon_{x \subseteq y}$ is the state which assigns identical values to x and y on just a subset of the incoming set of assignment that provide a value for y. A similar transformation is available for the maximality operator[5].

Given these tools, a standard quantificational structure looks like this:

(4.26) $\mathsf{Q}^y x(\varphi, \psi) := \epsilon_{x' \subseteq y} \wedge \epsilon_{x \subseteq x'} \wedge +\mathsf{M}_{x' \subseteq y}(\varphi[x/x']) \wedge \mathsf{M}_{x \subseteq x'}(\psi) \wedge Q(x', x)$

This concludes, for now, my discussion of van den Berg’s dynamic plural logic. I now turn to an alternative proposed by Manfred Krifka.

4.2 Parametrised sum individuals

Like dynamic plural logic, Krifka’s proposal involving parametrised sum individuals is set in the ‘tradition’ of dynamic predicate logic[6]. Krifka follows the line of DPL and other dynamic approaches by interpreting a sentence as a relation between assignment functions. It is the way in which these functions assign pluralities which distinguish Krifka 1996a from standard proposals. Here are the important aspects of his proposal.

(4.27)
a. The meaning of a sentence is a relation between assignment functions
b. Assignment functions are functions from variables to *parametrised sum individuals*
c. Parametrised sum individuals are plural individuals paired with an assignment function
d. Parametrisations are due to distributivity.

[5]For discussion and formalisation see van den Berg 1996b, p. 137-142.

[6]In what follows I changed a lot of Krifka’s notation in order to compare it to van den Berg’s work more easily. For instance, Krifka does not make use of variables, but instead uses index-like *discourse entities*. Ignoring such choices is harmless with respect to our current purpose, namely that of comparison of the basic quantificational and anaphoric mechanisms.

The notion of a *parametrised individual* was introduced in Barwise 1987. Since parametrised individuals are individuals which have a variable assignment associated with them, they are an ideal medium for modeling dependencies. The basic idea is that if y is dependent on x, then each atom in the group assigned to x comes with its own assignment for y.

Krifka builds up his ontology carefully. Let VAR be a set of variables. Let E be the domain of atomic entities. The domain of plural individuals is taken to be $\wp^+(E)$, where $\wp^+$ expresses the powerset operation excluding the empty set. Parametrised sum individuals or P-individuals are now defined recursively in tandem with the definition of the assignments.

(4.28) $\mathbb{P}_0 := \wp^+(\wp^+(D_e) \times \{\emptyset\})$

This is the base set of P-individuals: the set of plural individuals paired with empty assignments. We will write a P-individual made up from a (sum-)individual d and a function f as $d : f$. So, for instance, the individual $\{j, m\}$ is represented as $\{j : \emptyset, m : \emptyset\}$. The base set for assignments is based on $\mathbb{P}_0$: it is the set of partial functions from variables to base case P-individuals.

(4.29) $\mathbb{G}_0 := \text{VAR} \hookrightarrow \mathbb{P}_0$

Next, a first induction step defines $\mathbb{P}_1$ as the set of non-empty sets of pairs of plural individuals and base case assignments. $\mathbb{G}_1$ is defined as is to be expected:

(4.30)
$$\begin{array}{lcl} \mathbb{P}_1 & := & \wp^+(\wp^+(D_e) \times \mathbb{G}_0) \\ \mathbb{G}_1 & := & \text{VAR} \hookrightarrow \mathbb{P}_1 \end{array}$$

An example of an assignment in $\mathbb{G}_1$ is the following:

(4.31) $g = \{\langle x, \{j : \{\langle y, t : \emptyset\rangle\}, s : \{\langle y, d : \emptyset\rangle\}, m : \{\langle y, h : \emptyset\rangle\}\}\rangle\}$

This is roughly the same assignment as van den Berg's (4.6), except that the variable y is subordinated to x. It is only accessible through x. So, if the above assignment is called g, then $g(x)$ corresponds to the plural individual $\{j, s, m\}$ parametrised for y. That is, $g(x)(j)(y)$ yields t, $g(x)(s)(y)$ yields d etc.

The definition in (4.30) can be generalised to parametrised individuals and assignment functions which have assignments to complex P-individuals inside a P-individual itself.

(4.32)
$$\begin{array}{lcl} \mathbb{G}_n & := & \text{VAR} \hookrightarrow \mathbb{P}_n \\ \mathbb{P}_n & := & \wp^+(\wp^+(D_e) \times \cup\{\mathbb{G}_i | 0 \leq i < n\}) \end{array}$$

Finally, the full set of parametrised sum individuals and assignments to them is derived by:

(4.33)
$$\begin{array}{lcl} \mathbb{P} & := & \cup\{\mathbb{P}_i \mid 0 \leq i\} \\ \mathbb{G} & := & \cup\{\mathbb{G}_i \mid 0 \leq i\} \end{array}$$

To make the inhabitants of $\mathbb{P}$ and $\mathbb{G}$ somewhat readable, I propose the following notation. An assignment $\langle x, \{a_1, \ldots, a_n\}\rangle$ is represented as

$$\begin{array}{c} x \\ \boxed{\begin{array}{c} a_1 \\ \vdots \\ a_n \end{array}} \end{array}$$

I'll abbreviate the inhabitants of $\mathbb{P}_0$, $d : \emptyset$ as simply d. The example in (4.31) is written as:

(4.34)
$$\begin{array}{c} x \\ \boxed{\begin{array}{c} j : \begin{array}{c} y \\ \boxed{t} \end{array} \\ s : \begin{array}{c} y \\ \boxed{d} \end{array} \\ m : \begin{array}{c} y \\ \boxed{h} \end{array} \end{array}} \end{array}$$

Given an assignment g in $\mathbb{G}$, its value with respect to v is simply $g(v)$. Embedded assignments are not visible for ordinary application, so if the figure above depicts g, then $g(y)$ is undefined. In order to allow for parametrised individuals in the extension of a predicate as well as ordinary individuals, Krifka stipulates a lexical rule.

(4.35) $P(\{\langle a, g\rangle\}) \Leftrightarrow P(\{a\})$

There is a notion of *extending* an assignment which is closely related to the random assignment of DPL.

(4.36) g is extended to h by x or $g <_x h$ iff
$x \notin \text{RDOM}(g)\ \&\ \exists d \in \mathbb{P}_0 : h = g \cup \{\langle x, d\rangle\}$.

This is a relation which expresses a random assignment of a non-parametrised sum individual. It adds to g a fresh index x, which is assigned an individual in $\mathbb{P}_0$. The freshness is ensured by $x \notin \text{RDOM}(g)$ which expresses that x is not part of the so-called recursive domain of g, the ordinary domain plus all the indices in the domains of arbitrarily deeply embedded assignments.

Krifka's formalism is typed. Sentences correspond to relations between $\mathbb{G}$-assignments. One-place predicates correspond to ternary relations between two assignments and a variable. Krifka assumes that determiners are responsible for adding individuals to the context. For instance, he analyses "two students" as "$\exists_x$(two(student))," where '$\exists_x$' is a morphologically empty determiner, interpreted as the existential quantifier introducing index x. Here are some details:

(4.37)

$$\begin{array}{rcl} \text{two(student)} & = & \lambda v.\{\langle g,g\rangle | \text{2(student)}(v)\} \\ \exists_x & = & \lambda P.\{\langle g,h\rangle | \exists k : g <_x k \wedge \langle k,h\rangle \in P(k(x))\} \\ \exists_x(\text{two(student)}) & = & \{\langle g,h\rangle | \exists k : g <_x k \wedge \langle k,h\rangle \in \text{2(student)}(k(x))\} \\ & = & \{\langle g,h\rangle | g <_x h \wedge \text{2(student)}(h(x))\} \end{array}$$

The final line shows that "[two students]$_x$" is interpreted as introducing a sum individual containing 2 students for x in the context. This relation is combined with an operator indexed with the relevant variable (x in this case) which dictates how it combines with a verbal predicate. For instance, the subject operator takes a relation and a one-place predicate and feeds the value for this index to the predicate. Object operators behave similarly, but take two-place instead of one-place predicates.

Let us now focus on Krifka's treatment of dependence. The default mode of interpreting predicates is collective, just as in dynamic plural logic. And just as in that formalism, while there are no dependencies created or accessed with collective predication, distributivity acts as a major source of dependency. The definition of distribution is not straightforward:

(4.38) $\models g\,[\![\triangle_x\,(\pi)]\!]\,h \Leftrightarrow$ h is like g, except that every parametrised individual $d : f$ in $g(x)$ is replaced by $d : (f + i)$ such that: $\langle g + f, g + f + i\rangle \in \pi(\langle d, f\rangle)$.
where, $f + g = f \cup g$ provided that $\text{RDOM}(f) \cap \text{RDOM}(g) = \emptyset$

Here, π is a (possibly complex) one-place predicate, such as, for instance:

$$\lambda d.\{\langle g,h\rangle \mid g <_y h \wedge \text{article}(h(y)) \wedge \text{wrote}(d, h(y))\},$$

which corresponds to 'wrote an$_y$ article.' Distributing over this property with respect to x results in adding parametrisation to the individuals assigned to x in the incoming assignment. This parametrisation is such that it enables satisfaction of π.[7]

Using our representation of assignments in $\mathbb{G}$, a more schematic version of the definition can be given.

(4.39)
$$\left\{\ldots, \begin{array}{c} x \\ \boxed{\begin{array}{c} d_1 : f_1 \\ \vdots \\ d_n : f_n \end{array}} \end{array}, \ldots\right\} [\![\triangle_x\,(\pi)]\!] \left\{\ldots, \begin{array}{c} x \\ \boxed{\begin{array}{c} d_1 : f_1 + i_1 \\ \vdots \\ d_n : f_n + i_n \end{array}} \end{array}, \ldots\right\}$$
$$\Leftrightarrow$$
$$\forall 1 \leq j \leq n : \ g + f_j\,[\![\pi(d : f_j)]\!]\,g + f_j + i_j$$

[7] Unfortunately, this seems to be an oversimplification. The assignment $g + f$ can never be defined, since the domain of f occurs in the recursive domain of g.

Notice the resemblance to van den Berg's distribution operator. For each atomic individual d_j in $g(x)$, the scope of the distribution operator (the one-place predicate π) is interpreted with respect to its own parametrisation f_j and the input assignment g. Any extension of these assignments i_j due to the interpretation of π with $d : f_j$ as its argument end up in the parametrisation of d_j as well.

Here is an example. Say we have an assignment which assigns to x the parametrised sum individual 'the$_x$ students who each read a$_z$ book' and, say, this looks like (4.40).

(4.40)
$$\begin{array}{c} x \\ \boxed{\begin{array}{l} s_1 : \begin{array}{c} z \\ \boxed{b_1} \end{array} \\ s_2 : \begin{array}{c} z \\ \boxed{b_2} \end{array} \\ s_3 : \begin{array}{c} z \\ \boxed{b_3} \end{array} \end{array}} \end{array}$$

If we now interpret $\triangle_x\ (\lambda d.\{\langle g, h\rangle \mid g <_y h \wedge \textsf{article}(h(y)) \wedge \textsf{wrote}(d, h(y))\})$ relative to the input assignment in (4.40) (call it g), in a model in which s_1 wrote article a_1 etc., we first check whether each individual $d : f$ in $g(x)$ satisfies the predicate. This means the following three conditions hold.

(4.41) $\langle g + \{\langle z, b_1\rangle\}, g + \{\langle z, b_1\rangle\} + \{\langle y, a_1\rangle\}\rangle \in$
$\{\langle g, h\rangle \mid g <_y h \wedge \textsf{article}(h(y)) \wedge \textsf{wrote}(s_1, h(y))\}$

(4.42) $\langle g + \{\langle z, b_2\rangle\}, g + \{\langle z, b_2\rangle\} + \{\langle y, a_2\rangle\}\rangle \in$
$\{\langle g, h\rangle \mid g <_y h \wedge \textsf{article}(h(y)) \wedge \textsf{wrote}(s_2, h(y))\}$

(4.43) $\langle g + \{\langle z, b_3\rangle\}, g + \{\langle z, b_3\rangle\} + \{\langle y, a_3\rangle\}\rangle \in$
$\{\langle g, h\rangle \mid g <_y h \wedge \textsf{article}(h(y)) \wedge \textsf{wrote}(s_3, h(y))\}$

The output state corresponds to (4.44).

(4.44)
$$\begin{array}{c} x \\ \boxed{\begin{array}{l} s_1 : \begin{array}{cc} z & y \\ \boxed{b_1} & \boxed{a_1} \end{array} \\ s_2 : \begin{array}{cc} z & y \\ \boxed{b_2} & \boxed{a_2} \end{array} \\ s_3 : \begin{array}{cc} z & y \\ \boxed{b_3} & \boxed{a_3} \end{array} \end{array}} \end{array}$$

Just like in dynamic plural logic, distributivity here not only adds a dependency relation, it also accesses one. For each sub-individual of the entity assigned to x in the incoming assignment function, the predicate is evaluated with respect to the parametrisation of that sub-individual only. This means that if the predicate distributed over makes reference to a variable for which the individual assigned to x is parametrised, this variable is interpreted distributively as well. For instance, the sentence in (4.45) can be represented using distribution over the one-place predicate in (4.46).

(4.45) The$_x$ students who each read a$_z$ book each liked it$_z$.

(4.46) $\lambda d.\{\langle g, g\rangle \mid \text{like}(d, g(z))\}$

This way, each student in (4.40) will have to like the book it read in order for (4.45) to be true.

I have only discussed Krifka's treatment of numeral NPs. Here is a glance at Krifka's treatment of quantificational noun phrases. These use a maximality operator on dynamic propositions, like dynamic plural logic does. Quantificational determiners introduce two maximal values for its indices, which correspond to the restrictor and the intersection of the restrictor and scope. For instance, here is, for illustration only, the entry of (the subject quantifier) 'most.'

(4.47) $$\text{most}_{x,y} : \lambda P'.\lambda P.\{\langle g, h\rangle \mid \exists k, f : \langle g, k\rangle \in \text{MAX}(\{\langle g, k\rangle | \exists i : g <_x i \;\&\; \langle i, k\rangle \in P'(i(x))\}) \;\&\; k <_y f \;\&\; f(x) \subseteq k(x) \;\&\; \langle f, h\rangle \in \text{MAX}(\{\langle f, h\rangle | \langle f, h\rangle \in P(f(y))\}) \;\&\; |h(x)|/|h(y)| > 1/2\}$$

In many respects, Krifka's proposal resembles that of van den Berg, especially with respect to the notion of extension/existential introduction and that of distributivity. The key difference seems to be that parametrised sum individuals come with a notion of subordination. For instance, in (4.44), y and z are subordinated to x. In an information state for plurals representing the same context there would be no hierarchy of variables: x, y and z would simply be dependent on one another.

In the appendix to this chapter, we will focus in more detail on the relationship between the data structures involved in the formalisms. Let us now turn to another related framework.

4.3 A theory of anaphoric information

The key features of Elworthy's theory of anaphoric information (Elworthy 1995) are:

(4.48)
a. There are two interpretation functions, one constraining the model-theoretic interpretation, one constraining anaphoric interpretation.
b. Interpretation is relative to a *discourse set* (and a model).
c. Discourse sets are sets of tuples of possibly plural entities.

d. Dependent (discourse) pronouns differ from independent ones in the way they constrain the discourse set.

Interpretation in Elworthy's theory of anaphoric information (or TAI) proceeds with respect to a model and a *discourse set* "to give the truth conditions and the conditions on the anaphoric information." The interpretation function in TAI is split into two separate functions for constraints on the model and constraints on the discourse set. This means that there are two interacting semantics on formulae. The anaphora/discourse component expresses which discourse sets are allowed, while the truth-conditional component is a classical static semantics.

Although Elworthy's discourse sets look a lot like van den Berg's information states, there are important differences between Elworthy's work and that of van den Berg. To start with, *discourse sets* (or DSs), arguably the notion of information state used in TAI, is a set of tuples of *sum*-individuals. This means DSs are subsets of $\{\oplus X | X \subseteq D_e\}^n$. The zero individual ($\oplus\emptyset$, written '$\perp$') is included. If we recognise the fact that such objects can be seen as partial functions from indices to sum individuals –i.e. a tuple $\langle x_0, x_1, x_2\rangle$ corresponds to the function $\{\langle 0, x_0\rangle, \langle 1, x_1\rangle, \langle 2, x_2\rangle\}$– and if we furthermore confuse '$\perp$' with '$\star$', then it is obvious that the crucial difference between DSs and information states for dynamic plural logic is the fact that the former, but not the latter, allows pluralities in its states. As we did in the discussion of van den Berg's work, we will postpone the evaluation of the effects of such a decision to section 4.4.

I will for now also refrain from giving a detailed account of how the interpretation functions in TAI operate and interact. Let us jump straight to how quantificational sentences are interpreted, since that will enable us to understand how dependencies in TAI are accounted for. The following example is from Elworthy 1995.

(4.49) Most^i donkeys bray. = Elworthy 1995 (ex. 16)

(4.50) *(most donkey$_s$)(i)bray$_s$* = Elworthy 1995 (ex. 16a)

(4.51) *most'(donkey*,bray*)* = Elworthy 1995 (ex. 16.T)

(4.52) $(W \Downarrow i) \in \mathrm{Max}_{D_e}(\mathit{donkey}^* \cap \mathit{bray}^*)$ = Elworthy 1995 (ex. 16.A)

The example in (4.49) is enriched with information about indices. That is, the set introduced by the determiner *most* is positioned in slot i of the discourse set. In (4.50) we find the logical representation of (4.49), using the language L(GQA) (based on the language 'logic with generalised quantifiers' from L(GQ) in Barwise and Cooper 1981). This formula receives two interpretations: the truth-conditional one is in (4.51) and the anaphoric one is in (4.52). The latter expresses that the set of entities located at slot i in discourse set W (that is, '$W \Downarrow i$') should be a member of the set of maximal sets of braying donkeys.

As the discourse continues more and more conditions on the model and the discourse set W can be collected. Anaphora can make use of the constraints on W by relating to a certain slot. That is, the slot i in a discourse set W obeying the condition in (4.52) corresponds to the maximal set of braying donkeys. This is e-type information to be used in the subsequent discourse.

Elworthy has a single mechanism which derives all readings for quantificational sentences. Since predicate extensions are lifted by closure under the summation operator, all readings are allowed. For instance, Elworthy gives the truth-conditions in (4.54) for (4.53), where P^* is the closure under summation of P's extension.

(4.53) Two$_i$ boys buy three$_j$ roses.

(4.54) $2(\text{boy}^*, \oplus\{X \in \oplus E | 3(rose^*, buy^*(X))\}$

The numeral quantifier symbols 2 and 3 are ambiguous between 'exactly n', 'at least n' and an existential interpretation (see Elworthy 1995, p. 326, for details). The condition in (4.53) is not sensitive to how the 'buy'-relation is actually structured. It only uses the sum of individuals involved.

The anaphoric condition, however, *does* make use of, what Elworthy calls, the *essential* extension of a predicate. Part of it establishes the dependence relation by requiring that each value in the i slot (the (groups of) boys) buys the roses in the corresponding j slot (see Elworthy 1995, p. 307, for details). Discourse sets *always* mirror the model.

Anaphora introduce their own slots in discourse sets, but establish a relation with the slot of their antecedent. A plural pronoun 'They$_i^k$' takes the sum of individuals to be found in slot i and puts these individuals on slot k in whatever configuration that satisfies the predicate. A special case are pronouns that have two identical indices. A pronoun 'They$_i^i$' has to satisfy its predicate in exactly the way its antecedent did. This way it is possible to force the dependencies of which the antecedent is part of upon the anaphor.

For instance, consider (4.55) in a cumulative rendering resulting in the discourse set in (4.56).

(4.55) Four boys bought five roses.

(4.56)
$$\begin{array}{lclcl} & i & & j & \\ \langle & b_1 \oplus b_2 & , & r_1 & \rangle \\ \langle & b_2 \oplus b_3 & , & r_2 \oplus r_3 & \rangle \\ \langle & b_4 & , & r_4 \oplus r_5 & \rangle \end{array}$$

According to Elworthy, a subsequent continuation like 'They sniff them' is ambiguous between a dependent reading and an independent one. If we allow the two pronouns to introduce their own slot in the discourse set, then we interpret the continuation as meaning that the four boys sniff the five roses in whatever configuration. For instance:

They$_i^k$ sniff them$_j^l$

$$
(4.57)\quad \begin{array}{lcccccccccl}
 & i & & j & & k & & l & \\
\langle & b_1 \oplus b_2 & , & r_1 & , & b_1 & , & r_5 & \rangle \\
\langle & b_2 \oplus b_3 & , & r_2 \oplus r_3 & , & b_2 & , & r_4 & \rangle \\
\langle & b_4 & , & r_4 \oplus r_5 & , & b_3 & , & r_3 & \rangle \\
\langle & \perp & , & \perp & , & b_4 & , & r_2 \oplus r_1 & \rangle
\end{array}
$$

The more readily available reading, however, is one in which the boys sniff the roses they bought. This reading is enforced whenever the pronouns do not introduce a slot of their own, but simply add more conditions to the slot of their antecedent. This means that 'They$_i^i$ sniff them$_j^j$' is true whenever the boys collected at i sniff the roses collected in j in exactly the way i and j are structured within the discourse set, (4.56).

4.4 Comparison of data structures

The previous sections showed how the three proposals under discussion each give a semantics defined on contexts containing functional information concerning dependence relations. In this section, I want to focus more on the nature of such relations. What relations are predicted to occur in discourse by the proposals and which of these do we encounter empirically? Van den Berg 1996b gives the following examples to support his choice for having a distributive notion of assignment.

(4.58) Every^1 man loves a^2 woman. van den Berg 1996b (p. 126)

a. They$_1$ bring them$_2$ [..] flowers to prove this.
b. He$_1$ brings her$_2$ flowers to prove this.
c. Every$_1$ old man brings her$_2$ a present to prove this.
d. The$_1$ old men bring them$_2$ presents to prove this.
e. And they$_2$ love them$_1$ right back.
f. And she$_2$ loves him$_1$ right back.
g. Yesterday, several dogs dug holes in my garden.

Van den Berg distills four desiderata from this example. First, there is the trivial demand that information states should somehow assign pluralities to discourse referents (from (a)). Second, specific relations between the atoms of pluralities should be encoded (from (b), (c), (d), (f)).[8] Third, these relations should remain intact when subsets of the pluralities are

[8] I disagree with van den Berg that example (4.58b) and (4.58f) are prototypical cases of inter-sentential anaphora. They are cases of what is called *telescoping* in Roberts 1987. The fact that such examples is not generally acceptable indicates that one should be careful to use it to distill general properties of anaphora. (See also Poesio and Zucchi 1992; Carminati, Frazier, and Rayner 2002.)

addressed (from (c), (d)). Fourth, only explicitly expressed relationships should be encoded (from (g)).

Surely, these are basic desiderata both Elworthy and Krifka would agree with. Recall that Krifka's choice for parametrised sum individuals originated from very similar examples with a distributive antecedent sentence. So, it is interesting to look at more involved cases of dependency covering the whole spectrum of readings for quantificational sentences. In order to evaluate the three proposals under discussion, we will have to decide which interpretations are generated and what anaphoric effects these cause in discourse.

4.4.1 Distributivity, cumulativity and correspondence

Let us start with distributive readings and study in more detail how they behave in discourse and how this relates to the three formalisms discussed above.

Krifka and van den Berg have similar distributivity operators which evaluate its scope with respect to respective sub-states. They therefore have no problem with deriving just the single interpretation for (4.59).

(4.59) Three students each wrote an article. They each sent it to L&P.

The second sentence in this example cannot mean that each student sent any other article than the one she wrote to L&P. This is explained by Krifka and van den Berg by interpreting the pronoun 'it' relative to a substate or a subordinate assignment. We might wonder, however, how a pronoun inside the scope of a distributivity operator escapes the dependent reading. This is necessary to get the proper reading of (4.60).

(4.60) Three students each wrote an article. They each sent them to L&P.

The second sentence (4.60) tells us that each student sent all articles written by the three students to L&P. That is, the (same) three articles reached the L&P editor three times.

In dynamic plural logic, if a variable y is dependent on the ranging variable x of a distribution δ_x and y occurs in the scope of δ_x, then y is bound by the distributor. That is, it never returns the full collection that was assigned to it earlier, but rather a dependent value. In Krifka's proposal, once a variable is subordinated relative to another variable, its values can only be accessed through the latter variable. Distributivity enters the parametrisations value per value and is never able to retrieve a single collective value from the original input information state. Both Krifka and van den Berg are unable to derive the reading we described for (4.60).

Elworthy provides a solution to this problem by making pronouns ambiguous between structure sensitive and structure insensitive ones. The former class of pronouns are those which do not introduce a new slot in the discourse set, but simply add conditions to the slot of its antecedent.

Elworthy's analysis of the first sentence in (4.59) does not differ from van den Berg's very much. An example of a possible discourse set which is in accordance with this example is:

(4.61)
$$\begin{array}{cccc} & i & j & \\ \langle & s_1, & a_1 & \rangle \\ \langle & s_2, & a_2 & \rangle \\ \langle & s_3, & a_3 & \rangle \end{array}$$

The two ways the plural pronoun in (4.59) can be represented in Elworthy's system is either as 'They$_i^k$', or as a pronoun which creates extra conditions on slot i: 'They$_i^i$'. For Elworthy there are two renderings of (4.59):

(4.62) Three^i students each wrote an^j article. They$_i^k$ each sent it$_j^l$ to L&P.

(4.63) Three^i students each wrote an^j article. They$_i^i$ each sent it$_j^j$ to L&P.

The interpretation for (4.63) roughly says that the contents on slot i and the contents of slot j both satisfy the 'write' and the 'sent to L&P' relation and are represented in a discourse set which respects these relations. In other words, the extension of 'writing' and 'sending to L&P' should be identical with respect to the three students and the articles. In contrast to this, (4.62) only relates the summed content of i to that of k. Its truth-conditions are that whatever is collected in i sent, in one way or another, the content of j to L&P. One of those readings is any doubly distributed reading we can think of, for instance, one in which all the articles end up at L&P and all the students sent one, but none sent her own article, as illustrated in (4.64). Thus, (4.62) seems to wrongly predict that a non-dependent reading is available for the distributive case.

(4.64)
$$\begin{array}{ccccc} & i & j & k & \\ \langle & s_1, & a_1, & a_2 & \rangle \\ \langle & s_2, & a_2, & a_1 & \rangle \\ \langle & s_3, & a_3, & a_3 & \rangle \end{array}$$

This shows a serious defect in Elworthy's system. The flexible way pronouns can be indexed leads to a multitude of readings for (4.59), while only one seems to exist.

Something similar happens with the example in (4.60). The reading described above is derived by Elworthy by indexing the object plural pronoun as 'them$_j^l$'. But nothing stops us from deriving another reading by indexing it as 'them$_j^j$', rendering it the same as the dependency in (4.59). In contrast to Elworthy's analysis of (4.59), this might actually be a welcome prediction, since quite some native speakers seem to allow a reading for (4.60) in which each student sent her (and only her) paper to L&P.[9]

[9]This illustrates the vague principles for number agreement in discourse. Elworthy comments: "[A] fully adequate account of antecedent-anaphor agreement must draw on more

In sum, although all three proposals acknowledge the importance of the example in (4.59), their empirical coverage is otherwise limited. Dynamic plural logic and Krifka's semantics with parametrised sum individuals make too strong predictions, since their distributivity operation enforces dependency. Elworthy's ambiguity approach derives dependent and independent readings all of the time.

More issues arise when the antecedent sentence involves a more complicated quantificational *mode*, for instance cumulativity. Dependent readings are naturally available if we do not force distributivity on the continuation as well. One of the situations verifying the second sentence in (4.65) is one in which the student-article pairs that were involved in the writing relation according to the first sentence are also involved in the sending relation.

(4.65) Three students wrote four articles. They sent them to L&P.

Krifka follows Elworthy in calling the interpretation of the second sentence in (4.65) which parallels the dependencies in the verifying situation for the first sentence a *correspondence interpretation*. According to this reading, the students that cooperated in writing one or more articles also cooperated in sending these (and no other) article(s). The status of this interpretation is very much unclear. How do we know this is not simply a possible model for the cumulative interpretation of the second sentence? Krifka claims that "the correspondence interpretation can be marked by *each* [,] which shows that it is nothing else but the distributive interpretation" (Krifka 1996a, p. 577). Krifka acknowledges the difficulty of getting clear intuitions on the possibility of marking these interpretations with a distributive floating quantifier, which is why he turns to sentences where the number of the subject and the object are the same. These sentences, if interpreted cumulatively, are most likely understood as corresponding to a one-to-one situation. The continuation, then, follows this mapping.

(4.66) Three students wrote three articles. They each sent them to L&P.

This shows that the first sentence in (4.66) recorded a dependency between students and articles, since the distributivity operator is capable of accessing this relation, at least, if we acknowledge that the plural inflection on the object pronoun is due to the plurality of its antecedent[10].

This is not to say that correspondence readings can be marked with overt distributivity operators in general. Whenever pluralities are involved in the dependencies created by cumulativity, 'each' can no longer

than one kind of information, and [..] –given the variability in intuitions– it is not possible to specify one source of information as being the definitive one." However, some hard principles seem to exist. While (4.60) shows that plural pronouns can refer to both semantically and syntactically singular antecedents, a singular pronoun as the one in (4.59) can never be used with a semantically plural antecedent.

[10]Some people actually prefer a singular object pronoun in (4.66).

access them. Although it is unclear how Krifka stops his distributivity operator from ranging over pluralities, I'll assume that overt 'each' quantifies over atomic individuals. This is supported by examples like (4.67).

(4.67) During Live Aid, sixty bands performed on two locations.
They each wrote a song for the occasion.

Were 'each' capable of accessing the dependency between groups of bands and locations in (4.67), then we expect a reading for (4.67) in which there were two songs written, one by each of the two groups of bands. Instead, however, we only get a reading in which each band wrote a song.

Krifka's example (4.66) shows that cumulativity, like distributivity, involves some quantificational mechanism which stores dependencies. Still, in contrast to dependent readings under a distributivity operator, correspondence readings do not seem to be enforced by a cumulative interpretation. Consider (4.68) and (4.69).

(4.68) Tom, Dick and Harry each wrote an article. They (each) sent it to L&P. # To be precise, Tom sent the article Dick wrote, Dick sent the article Harry wrote and Harry sent the article Tom wrote.

(4.69) Tom, Dick and Harry wrote three articles. They sent them to L&P. To be precise, Tom sent the article Dick wrote, Dick sent the article Harry wrote and Harry sent the article Tom wrote.

While in (4.68) one has to interpret the continuation as being dependent on its antecedent, this is not the case in (4.69). The correspondence 'interpretation' is the most salient reading for (4.69), but the continuation shows that other readings are just as acceptable. If cumulativity involves the storage and subsequent accessing of dependencies, the question becomes how a cumulative interpretation can be verified by a model in accordance with the dependency as well as with an arbitrary model that is not.

Krifka's cumulativity operation potentially explains this. It involves making a cover of the group denoted by the subject. For instance, in a cumulative interpretation a cover is formed from the set of three students. Each of the cells in this cover is expected to write something. That what is written is expected to amount to three articles. The quantification over cells in the cover takes care of storing the dependencies.

The cover of the group expressed by the subject is associated with a different variable than the original group is. So if the example is indexed as 'Three$_x$ students wrote five$_y$ articles' than a successful interpretation leads to an assignment to parametrised sum individuals g in which some variable x' has a cover of $g(x)$ assigned to it. The articles are assigned in the parametrisations of the cells at x'. Consequently, the original group (at x) can be involved in subsequent cumulative predications without addressing the cells at x'. Alternatively, however, the cover stored in x' can be used to access the dependencies. This explains the dependency observed in (4.66) and the potential independence observed in (4.69).

A detailed analysis of cumulativity and its relation to anaphora are beyond the scope of this dissertation. If Krifka's example (4.66) indeed shows that cumulativity involved the storing of dependencies in discourse, then we will be able to accommodate this in an analysis by assuming that cumulativity involves a dependency-creating operation very much like distribution. As we mentioned above, however, it is difficult to keep apart dependent readings from possible verifying models of a cumulative interpretation. In fact, it seems to me that we only have evidence of the possibility of accessing of a dependency under the scope of distributivity. (This intuition is certainly supported by (4.66) and (4.69)). Studying distributive sentences will therefore give a much clearer picture of the constraints on dependent reference.

One question is left open for now. Namely, how all these structures can be set to work in such a way as to at the same time have access to both the dependent values and the original group denoted by the antecedent. The second sentence in (4.70) is ambiguous between Tom, Dick and Harry each sending the two papers they wrote to L&P, or each of them sending copies of the six papers they wrote altogether to L&P.

(4.70) Tom, Dick and Harry each wrote exactly two articles. They each sent them to L&P.

As we saw, both Krifka and van den Berg cannot predict these possibilities. For now, I will leave this an open problem, but it will play a central role in chapter 5.

4.4.2 Subordination

Finally, a few words on Krifka's notion of subordination. Krifka sees a problem for theories which do not assume some sort of subordination relation between indices dependent of one another. If two respective sentences introduce two dependency relations, then how are these combined? With respect to Elworthy's discourse sets, Krifka gives the following example.

(4.71) Every$_1$ student wrote an$_2$ article.

(4.72) $\{\langle s, a\rangle, \langle s', a'\rangle\}$

(4.73) Every$_3$ professor read a$_4$ book.

(4.74) $\{\langle \ldots, \ldots, p, b\rangle, \langle \ldots, \ldots, p', b'\rangle, \langle \ldots, \ldots, p'', b''\rangle\}$

In case (4.71) results into the discourse set (4.72) and (4.73) into (4.74), then what discourse set represents the complex text which combines (4.71) and (4.73)? Krifka considers making use of Elworthy's '$\perp$'-element and deriving a discourse set as in (4.75).

(4.75) $\{\langle s, a, p, b\rangle, \langle s', a', p', b'\rangle, \langle \perp, \perp, p'', b''\rangle\}$

He rejects such an approach since "it introduces unwarranted dependencies between individuals, like between s and p, as they happen to be instantiated in the same tuple" (Krifka 1996a, p. 596). He concludes: "A framework [..] with subordination of discourse referents [..] does not run into such problems" (Krifka 1996a, p. 596).

The problem is a very interesting one, since it addresses a fundamental issue in dynamic semantics. If meaning is to be interpreted as context change potential, then it should not matter *what* context is being changed: formulae should be uniformally interpreted and these interpretations should be uniformally combined. Krifka's complaint is that Elworthy is not specific about how his interpretation of sentences combines with interpretations of other sentences without leading to new dependencies.

In the appendix[11] of Elworthy 1995, we find some definitions for interpretation of sequences. First, a formula F is interpreted relative to a set of discourse sets V as in (4.76).

(4.76) $I(F, V) = \{W \in V | \, [\![F]\!]_t^W \wedge [\![F]\!]_a^W\}$

The set V is a set of discourse sets which have not been discarded yet. The interpretation of a closed formula F takes the largest subset of V which is truth-conditionally and anaphorically consistent with F. Elworthy gives a number of options for the interpretation of a sequence of formulae '$F_1; F_2$'.

(4.77) $I(F_1; F_2, V) = I(F_1, V) \cap I(F_2, V)$

(4.78) $I(F_1; F_2, V) = I(F_2, I(F_1, V))$

Returning now to Krifka's example, given a model in which s wrote a and s' wrote a', (4.71) results in a subset X of V such that (say) slot 1 and slot 2 form either the pair (s, a) or the pair (s', a'). Similarly, (4.73) results in a subset Y of V such that (say) slot 3 and slot 4 form either the pairs (p, b) or (p', b') or (p'', b''). Sequencing (4.71) and (4.73) either results in $X \cap Y$ or in an interpretation of (4.73) not relative to V but rather to X. These are equivalent in this case and result in the set of those tuples in V, which have a student in slot 1 who wrote the article in slot 2 and which have at the same time a professor in slot 3 reading the book in slot 4.

It seems that Krifka's worry about unwarranted dependencies are justified. A third sentence F_3 in the sequence will be interpreted relative to the intersection of X and Y. That is, the sequence is interpreted as the set of tuples such that they each contain the aforementioned relations between slot 1 and 2 and slot 3 and 4 plus obey the truth-conditional and anaphoric constraints expressed by F_3.

Consider now a model in which the students happened to have read all the books which were read by the professors. This means that for each

[11] See also Elworthy 1995, p. 310, and Elworthy 1992, section 4.3.5.

tuple in $\Gamma \cap \Gamma'$ it is accidentally the case that the student in slot 1 read the book in slot 4. Yet in order to describe such a situation, a sentence like (4.79) seems highly inappropriate after (4.71) and (4.73), since it refers to a dependence relation which was never expressed. (Notice that the plural variant 'they each read them as well' *is* appropriate.)

(4.79) They$_1$ each read it$_4$ as well.

Still, interpreting (4.79) relative to $\Gamma \cap \Gamma'$ would succeed (it would return $\Gamma \cap \Gamma'$ again). The reason is that inside $\Gamma \cap \Gamma'$ the relation between slot 1 and 2 is represented in exactly the same way as is the *in*dependence between slots 1 and 4.

This does not mean, however, that subordination is a necessary ingredient to cure the unwarranted dependencies. Krifka does not discuss van den Berg's work relative to this issue, but his approach differs from Elworthy's. For instance, remember that the existential quantifier never introduced dependencies. In fact, just like Krifka's work, the only source of dependency in dynamic plural logic is distributivity.

It is unclear, whether Krifka has reasons for assuming a subordination structure other than the avoidance of unwarranted dependencies. Natural language does not seem to indicate the need for subordination. Surely, in (4.80), the set of papers is available as an antecedent without first accessing the students.

(4.80) Every student turned in a paper. They were all identical.

This strongly suggests that the element of variable subordination in the parametrised sum individual framework is redundant.

4.5 Conclusions

Let us focus on the common achievements of the three proposals discussed here.

All three authors acknowledged that the accessibility of dependent antecedents was the most basic fact they had to account for. This is mainly why I focused on these authors, since I know of no other analyses which do so in this much detail. As discussed above, all three successfully deal with the crucial example (4.59). Unfortunately, none of the proposals is able to express the relation between distributivity and dependence in much detail and all of them make wrong predictions somewhere along the line. In the end, I have a strong preference for van den Berg's data structures which in the deepest level involve nothing but atoms. Moreover, because of their fine-grainedness and flexibility, information states for plurals are the most *general* proposal. The appendix to this chapter stresses this formally. There, it is shown that there exists a mapping from P-sum structures to information states for plurals which preserves the representation of dependence and independence from one formalism to the other.

In the remaining chapters of this thesis, I therefore use van den Berg's proposal as a basis for mine. In the next chapter, I will focus on a problem we encountered in the discussion above, namely how a sentence containing a pronoun can be ambiguous between a dependent and an independent reading.

4.A Formal comparison of data structures

In this appendix, we set out to obtain a better understanding of how the different data structures we encountered in this chapter relate formally. In order to achieve this we will translate the structures in Krifka's formalism into discourse sets, which in turn we will map to information states for plurals. Then, we will show that given these mappings two key features of Krifka's and van den Berg's formalisms, namely dependency and independent existential introduction, correspond.

I will not attempt to give a translation the other way around, from information states for plurals to assignments to parametrised individuals. This is because of the subordination of variables in Krifka's framework. A single information state from van den Berg's work would correspond to a multitude of assignments to P-individuals. Also, since in the coming chapters we take dynamic plural logic as our starting point, it suffices to show that information states for plural are a more general medium for storing dependencies and that Krifka's results can be transferred to DPlL.

Let us agree on some set of variables $\mathbb{V}$. Let us also simplify Elworthy's *discourse sets* as sets of functions of the same domain, namely some subset of $\mathbb{V}$.

Definition 4.5

From sum-parametrisations to discourse sets

Let $g \in \mathbb{G}$ (an assignment to P-individuals)
Let $f : V \to \wp(D_e)$ for some $V \subseteq \mathbb{V}$ (a function in a discourse set)

$$f \sim g \quad :\Leftrightarrow \quad \text{rdom}(g) = \text{dom}(f)\ \& \\ \forall x \in \text{rdom}(g) : \exists W \subseteq \mathbb{V} : \exists h : W \to \wp(D_e) : \\ \exists x'_1 \ldots x'_n : n \geq 0\ \&\ g(x'_1)(f(x'_1)) \ldots (x'_n)(f(x'_n))(x)(f(x)) = h$$

□

This says that f is a function in a discourse set corresponding to an assignment to P-individuals g if and only if for each variable x in the recursive domain of g: $g(x)$ and $f(x)$ agree on the individual they assign. This can happen in two ways: either $\exists h : g(x)(d) = h$ and $f(x) = d$, that is the value that f assigns to x is (one of the) the parametrised sum(s) assigned to x in

g or $f(x)$ is a parametrised sum assigned to x somewhere deeply embedded in g. In this latter case such an assignment to x is reached through a series of other variables and variable-assignments. Crucial is that reaching this value happens through values assigned to this series of variables in f.

Here is an example. Consider the P-individual assignment in (4.34), repeated here, and call it g.

(4.81) $$g = \begin{array}{c} x \\ \boxed{\begin{array}{r} j : \begin{array}{c} y \\ \boxed{t} \end{array} \\ s : \begin{array}{c} y \\ \boxed{d} \end{array} \\ m : \begin{array}{c} y \\ \boxed{h} \end{array} \end{array}} \end{array}$$

Call f the function $\{\langle x, j\rangle, \langle y, t\rangle\}$. According to the definition above it holds that $f \sim g$, since $g(x)(f(x)) = \{\langle y, t : \emptyset\rangle\}$ and $g(x)(f(x))(y)(f(y)) = \emptyset$.

We transform an arbitrary P-individual assignment function g into a discourse set $g^{\sim}$ as follows: $g^{\sim} = \{f | f \sim g\}$.

For (4.81) $g^{\sim}$ results into the following discourse set.

(4.82) $$g^{\sim} \;=\; \begin{array}{l} \{\{\langle x, j\rangle, \langle y, t\rangle\}, \\ \;\{\langle x, s\rangle, \langle y, d\rangle\}, \\ \;\{\langle x, m\rangle, \langle y, h\rangle\}\} \end{array}$$

The mapping '$\sim$' takes away the element of subordination central to Krifka's proposal and missing in Elworthy's framework. Whereas in g dependencies are stored in parametrisations of individuals, in $g^{\sim}$ they come about in the form of functions. The relation between discourse sets and information states for plurals is more straightforward. In contrast to Elworthy's system, dynamic plural logic does not have any means to assign pluralities within single functions.

Definition 4.6

From discourse sets to information states for plurals

Let $V \subseteq \mathbb{V}$ (a set of variables)
Let $S \subseteq \{f | f : V \to \wp(D_e)\}$ (a discourse set)
$S^{\bullet} := \{g | \text{dom}(g) = V \;\&\; \exists f \in S : \forall x \in V : g(x) \in f(x)\}$

□

$S^{\bullet}$ collects those functions that share their domain with the functions in S and that assign atomic individuals that are part of the plural values given in one of the functions in S to the variables. Here, I assume pluralities to be sets, so I use the membership relation '$\in$', but nothing hinges on this and a similar definition could be given using i-sums and the corresponding atomic part relation 'Π' (Link 1983).[12]

In order to demonstrate that these mappings are meaningful, I show that both *independence* (existential introduction) and *dependence* in van den Berg's proposal correspond to similar notions in Krifka's framework, according to the translations defined here. That is, from $f \leq_x g$ it follows that $(f^\sim)^\bullet \llbracket \epsilon_x \rrbracket (g^\sim)^\bullet$ and x and y are dependent in $(f^\sim)^\bullet$ if they are also in f.

Notice, first of all, the following:

$$(I): \quad \begin{array}{lcl} f \leq_x g & \Leftrightarrow & x \notin \mathrm{rdom}(f) \ \& \ \exists d \in \mathbb{P}_0 : g = f \cup \{\langle x, d\rangle\} \\ f \leq_x g & \Rightarrow & \neg\exists h \in f^\sim : x \in \mathrm{dom}(h) \ \& \\ & & \exists d \subseteq \wp(D_e) : g^\sim = \{h \cup \{\langle x, d\rangle\} | h \in f^\sim\} \end{array} \tag{4.36}$$

If g extends f by x, then the discourse set corresponding to g simply adds the assignment of some plurality d to x to each function in the set $f^\sim$. There exists such a plurality, since we can take d such that $g(x) = d : \emptyset$. Since, if $f \leq_x g$, it holds that $g = f \cup \{\langle x, d : \emptyset\rangle\}$, it is easy to see that $g^\sim$ can be derived by taking the functions in $f^\sim$ and adding $\langle x, d\rangle$ to each of them.

With respect to the mapping from discourse sets to information states for plurals, the following holds for each discourse set S:

$$(II): \quad \begin{array}{l} \neg\exists h' \in S : x \in \mathrm{dom}(h') \ \& \ \exists d \in \wp(D_e) : S' = \{h \cup \{\langle x, d\rangle\} | h \in S\} \\ \Rightarrow \\ S^\bullet(x) = \emptyset \ \& \ \exists d' \in \wp(D_e) : S'^\bullet = \{h \cup \{\langle x, e\rangle\} | e \in d' \ \& \ h \in S^\bullet\} \end{array}$$

Of course, we take $d' = d$. (II) follows straightforwardly from how we defined '$\bullet$'. From the conclusion of (II) it follows that $S^\bullet \llbracket \epsilon_x \rrbracket S'^\bullet$. Consequently, from (I) and (II) it follows that:

$$f \leq_x g \ \Rightarrow (f^\sim)^\bullet \llbracket \epsilon_x \rrbracket (g^\sim)^\bullet$$

Interestingly, it does not hold that if $(f^\sim)^\bullet \llbracket \epsilon_x \rrbracket (g^\sim)^\bullet$ it follows that $f \leq_x g$. Firstly, the reverse of (II), given here as $(II)'$, is not valid.

$$(II)' \quad \begin{array}{l} \text{(invalid)} \\ \neg\exists h' \in S : x \in \mathrm{dom}(h') \ \& \ \exists d \in \wp(D_e) : S' = \{h \cup \{\langle x, d\rangle\} | h \in S\} \\ \Leftarrow \\ S^\bullet(x) = \emptyset \ \& \ \exists d' \in \wp(D_e) : S'^\bullet = \{h \cup \{\langle x, e\rangle\} | e \in d' \ \& \ h \in S^\bullet\} \end{array}$$

The culprit is the empty set. Since any plural individual, including the empty set may occupy a position in a discourse set, an information state for plurals which does not assign a value to some variable x could either be derived from a discourse set which does not have a slot x, or from a discourse set which assigns $\emptyset$ at x. For instance, the two discourse sets below, S and S', map to the same state. That is $S^\bullet = S'^\bullet$.

[12]This also means that in a notation as in (4.82) I take atoms like j to correspond to the singleton set $\{j\}$. Schwarzschild (1992) refers to such an assumption as *Quine's innovation*.

(4.83) $S = \{\{\langle x, \emptyset\rangle, \langle y, d\rangle\}\}$ $\qquad S' = \{\{\langle y, d\rangle\}\}$

Notice that if we ignore the empty set or assume that assigning the empty set to a variable leaves the assignment undefined for that variable that (II') *is* valid. In the coming two chapters, we will elaborate on how undefinedness relates to empty sets.

The reverse of (I) is not valid either. It *does* hold that if $f^{\sim}$ does not contain a function defined on x, then the original assignment to p-individuals cannot have x in its recursive domain. However, if $g^{\sim}$ is derived from $f^{\sim}$ by adding the assignment $\langle x, d\rangle$ to all of $f^{\sim}$'s functions, then this can mean two things. Either $\langle x, d : \emptyset\rangle$ was not embedded in g (which is the case described by $f \leq_x g$) or the same p-individual $d : \emptyset$ occurred in the parametrisation of all the parts of some other individual. Here is an example.

(4.84) $g = \begin{array}{c|c} x & y \\ \hline d & \{e_1, e_2, e_3\} \end{array}$ $\qquad g' = \begin{array}{c} y \\ \boxed{\begin{array}{l} e_1 : \overset{x}{\boxed{d}} \\ e_2 : \overset{x}{\boxed{d}} \\ e_3 : \overset{x}{\boxed{d}} \end{array}} \end{array}$

Both assignment functions to P-individuals in (4.84) map to the same discourse set via '$\sim$'. That is, $g^{\sim} = g'^{\sim}$. This means that we cannot prove that $f \leq_x g \;\Leftarrow\; (f^{\sim})^{\bullet} [\![\epsilon_x]\!] (g^{\sim})^{\bullet}$. Nevertheless, notice that the problematic cases here, are cases in which the same individual occurs in all the parametrisation of an individual assigned to some other value. That is, with respect to the values that are assigned, x, is independent from y in g' in (4.84). This indicates where the subtle differences between the proposals lie. Nevertheless, we have seen that, given the mappings $\sim$ and $\bullet$, the notion of extension in Krifka's formalism corresponds to the independent existential introduction of van den Berg's framework. Now for the notion of dependency.

First of all, we have to agree what it means for two variables to be in a dependency relation in Krifka's formalism.

Definition 4.7

Dependency in assignments to P-individuals

y is dependent on x in g
iff
$\exists d, e \in g[x] : g[x](d)[y] \neq g[x](e)[y]$

$\Box$

Here, $g[x]$ is the cumulative value of x in g: the set of parametrised sum individuals assigned to x at an arbitrary deep level in g.

In order to show that the notion of dependency is not only defined to serve our goal of demonstrating the close relation between the parametrised sum individual framework and dynamic plural logic, we first focus on how the above definition of dependency interacts with distributivity.

Notice that the above definition does not reduce to subordination. Although in g' in (4.84) x is subordinate to y, x is *not* dependent on y. In fact, with respect to distribution the two representations in (4.84) do not differ. With respect to both g and g', an interpretation of 'They$_y$ each V it$_x$' as '$\triangle_y\ (V(x))$' leads to the same result. (That is, $\exists h : g\, [\![\triangle_y\ (V(x))]\!]\, h$ if and only if $\exists h' : g'\, [\![\triangle_y\ (V(x))]\!]\, h'$.)

We would like to show that if x is introduced while distributing over y, then x is dependent on y. This, however, does not hold. Cases like (4.84), where x is assigned a constant value, form the exception. But as we saw above, structures like g' in (4.84) are best seen as objects in which x and y are independent.

If a new variable x is introduced using extension within the scope of distribution over y, then x is always subordinated to y. This follows from the fact that the extension is performed with respect to the parametrisations of values assigned to y. If x is subordinated to y, then (trivially so) either y's value is constant in each P-individual assigned to x or there are two values assigned to y with a different parametrisation with respect to x. Consequently, distributivity over a scope containing an extension always leads to subordination but only leads to dependency in case the different parametrisations are distinguishable.

Define for a discourse set S, $S(x)$ as the set of assignment to x made by some function $f \in S$: $S(x) = \{f(x)|f \in s\}$. Also, $S|_{x=d}$ is that subset of S containing only functions $f \in S$ that assign d to x: $S|_{x=d} = \{f \in S|f(x) = d\}$. These definitions are parallel to van den Berg's definitions (see page 4.1 of this chapter), with the key difference that $S(x)$ is a set of sets (instead of a set) and the d in $S|_{x=d}$ is a set (instead of an atom). We now have to show:

$$(III) \qquad \begin{array}{l} \exists d, e \in g[x] : g[x](d)[y] \neq g[x](e)[y] \\ \Rightarrow \\ \exists d', e' \in g^{\sim}(x) : g^{\sim}|_{x=d'}(y) \neq g^{\sim}|_{x=e'}(y) \end{array}$$

If there is a P-individual $d' : h'$ such that for some P-individual $d : h$, $g[x](d)[y](d') = h$, then there will be a function f in $g^{\sim}$ such that $f(x) = d$ and $f(y) = d'$. According to the premise there is also a $f' \in g^{\sim}$ for which there is a $e : i$ and a $e' : i'$ such that $f'(x) = e$, $f'(y) = e'$ and $e' \neq d'$. Consequently, the two subsets of the discourse set, $g^{\sim}|_{x=d}$ and $g^{\sim}|_{x=e}$ give different projections for y. This proves (III). Now, toward independence in information state for plurals:

$$(IV) \quad \exists d', e' \in S(x) : S|_{x=d'}(y) \neq S|_{x=e'}(y)$$
$$\Rightarrow$$
$$\exists d'', e'' \in S^{\bullet}(x) : S^{\bullet}|_{x=d''}(y) \neq S^{\bullet}|_{x=e''}(y)$$

Say S, contains at least two functions: f and f'. According to the premise two different assignments to x in S lead to different projections to y. Schematically:

$$S: \quad f = \begin{array}{|c|c|} \multicolumn{1}{c}{x} & \multicolumn{1}{c}{y} \\ \hline \{\ldots, a, \ldots\} & \{\ldots, b, \ldots\} \\ \hline \end{array}$$
$$f' = \begin{array}{|c|c|} \multicolumn{1}{c}{x} & \multicolumn{1}{c}{y} \\ \hline \{\ldots, c, \ldots\} & \{\ldots, d, \ldots\} \\ \hline \end{array}$$

The d' and e' witnessing independence in discourse set (and the d and e in P-sum structures) are non-identical (since, otherwise, $S|_{x=d'}$ and $S|_{x=e'}$ would be the same set of functions). So, in the schema, above, $a \neq c$ (or, rather, the sets assigned there should contain an atom a and c respectively such that $a \neq c$.) Since the projections to y with respect to the function(s) assigning d' to x and those assigning e' to x differ, there should be two atoms b and d, as above, which are non-identical as well. Consequently, to verify the consequent-clause of (IV), we take $d'' = a$ and $e'' = c$. According to the definition of $^{\bullet}$ there should be at least one function $f \in S^{\bullet}$ such that $f(x) = a$ and $f(y) = b$ and a $f' \in S^{\bullet}$ such that $f(x) = c$ and $f(y) = d$.

Although the possibility of representing a subordination relation between variables in an assignment to parametrised sum individuals makes Krifka's data structures more expressive than discourse sets or information states for plurals, the formal discussion above shows that the elementary representation of (in)dependence in the three data structures are closely related.

Chapter 5

Plurality and Indices

Following our discussion about what type of information state is needed to come to a model of the anaphoric possibilities of plural discourse pronouns, we now turn to the question of how exactly the anaphoric relations themselves are represented. In (5.1), the antecedent quantifier *few senators* will have to introduce new information in the sufficiently structured context. This information will have to be stored under some sort of label, a referent or a variable name, a file card, an index, etc. The plural pronoun will have to retrieve that information using the label in question. To express such a relation, we often decorate examples with indices.

(5.1) Few$_i$ senators admire Kennedy$_j$; and they$_i$ are very junior.

What (5.1) represents is a possible understanding of the utterance 'few senators admire Kennedy and they are very junior.' Other readings, where 'they' does not refer to the senators that admire Kennedy are also possible albeit rather unlikely.

In (5.1), the choice of the indices is arbitrary and with respect to the clarification method of example decoration this is harmless. Instead of giving 'few' and 'they' index i we could just have used k without changing the reading we wanted to express. On the logical representation level, however, it is of much more importance which label we choose. This is because the utterance in question could occur in a context in which the label in question is already in use and consequently we could unwantedly access and update information which is independent from the current utterance. A related problem occurs with the anaphoric dependencies we discussed in the previous chapter. Dependent uses of pronouns and ordinary uses of pronouns can have the same antecedent, so mere co-indexation will not be a sufficient way of representing these cases of anaphora. The current chapter focuses on a solution to index-related problems. It is structured as follows. First, I will discuss the two problematic issues related to index or variable choice that occur in a dynamic semantics of plurality. In a more general formulation they involve: (i) the potential of standard dynamic semantics to overwrite information present in context and (ii) the existence of (discourse) anaphoric relations not expressible in terms of simple coindexation at some level of syntactic representation. After sketching a solution for the first of these problems, I will introduce a dynamic framework called *incremental dynamics* (van Eijck 2001) in section 5.2. The particular variable-free way of dynamic binding proposed in this system ensures that the choice for a label is made *in context* and it thus circumvents the first problem. In section 5.3, I will present a reformulation of van den Berg's dynamic plural logic in the framework of incremental dynamics. Finally, in section 5.4, its merits will be demonstrated by defining an alternative to van den Berg's distributivity operator which shows much more anaphoric flexibility. Finally, in section 5.5, I will discuss how pronouns receive an underspecified interpretation, which allows us to account for the problematic cases of anaphora in (ii).

5.1 Indices and the syntax-semantic interface

5.1.1 Variable choice

Let us start with asking a question which at first sight might seem simple to answer: where do the variables we use to model pronouns in the semantic representations come from? I think there are at least two possible answers to this question. First, we could think of the relation between pronoun and antecedent as something independent from semantics: coindexation relations are already given in the input to the interpretation mechanism. That is, together with a syntactic structure (say, a 'logical form') comes information of how elements in that structure relate anaphorically, for instance by expressing indexation patterns on the structure (as e.g. in (5.1)). During interpretation, the variable corresponding to the pronoun will be whatever variable we choose to associate with the antecedent. The ambiguity in reference of the pronoun is expressed on the level of the structure which is the input to the interpretation procedure. For instance, the same syntactic structure could come with different possible indexations.

This is the standard approach of, for instance, Dynamic Montague Grammar (Groenendijk and Stokhof 1990): "[W]e do not translate sentences as such, but indexed structures, i.e., sentences in which determiners, pronouns and proper names, are marked with indices" Groenendijk and Stokhof 1990 (p. 22).

A second option, however, is to wait with pronominal interpretation until the utterance it occurs in is (partially) interpreted itself. Whereas in the first strategy antecedents were syntactic objects, in the second strategy the semantic level provides the antecedents for pronouns, namely a variable or a referent, i.e. a semantic entity.

An example of the second strategy is DRT. Kamp and Reyle remark: "We will analyse anaphora not as a relation between pronouns and other NPs, but as one between pronouns and discourse referents that are already present in the semantic representation under construction" Kamp and Reyle 1993 (p. 67). The way this works in DRT is that in the construction of a DRS pronouns introduce their own referent whose value is then equated with that of some accessible other referent.

But even in a more compositional setup a semantic approach to antecedent choice is possible. One could for instance have expressions containing pronouns derive underspecified interpretations, in which a dummy object corresponds to the pronoun. In order to come to a specific interpretation of such a representation the dummy object has to be replaced by an (accessible) variable. For instance, an example like 'if a dog sees a cat, it runs away' could compositionally be mapped to the formula:

(5.2) $(\exists x_i(\text{DOG}(x_i) \wedge \exists x_j(\text{CAT}(x_j) \wedge \text{SEE}(x_i, x_j))) \Rightarrow \text{RUN-AWAY}(x_?))$

Here, $x_?$ is a place-holder for the pronoun. Once it is replaced by one of the 'accessible' variables x_i or x_j (depending on the outcome of pronoun resolution) the formula in (5.2) becomes interpretable.

The question where the variables corresponding to pronouns come from is thus a question about whether the resolution mechanism makes its choice with respect to a set of syntactic antecedents or if it operates on semantic representations. Some problems pointed out in Heim, Lasnik, and May 1991a and Heim, Lasnik, and May 1991b are relevant to the first strategy, where indexation is thought to be part of the input to the semantic module. In these problems the syntactic nature of such indexations gets in the way. Heim, Lasnik and May extensively discuss a set of problems which show up in the analysis of reciprocity and plural pronouns, some of which appear to be due to the fact that the traditional logical forms associated with the examples in question are not fine-grained enough to represent all referential options. In (5.3), for instance, there are three possible resolutions for the plural pronoun. Still, only two (suitable) antecedents are represented at LF. Even worse, ideally, these two expressions are coindexed.

(5.3) John and Mary told each other that they should leave.

a. John and Mary both said: 'I should leave.'
b. John and Mary both said: 'You should leave.'
c. John and Mary both said: 'We should leave.'

The solution to this problem is to decompose the reciprocal into a distributor 'each' and a reciprocator 'other', which both carry an index. The three readings of (5.3) can then be expressed as:

(5.4) [John and Mary]$_i$ each$_j$ told [the other]$_k$ that they$_{i/j/k}$ should leave.

Choosing i as the index for 'they' results in the 'we'-reading, choosing j results in the 'I'-reading and k corresponds to the 'you'-reading.

The problem is independent of reciprocity. The distributor 'each' always needs to carry an index, whether it is explicit or implicitly given. For instance, (5.5), which contains the distributive floating quantifier 'each,' has a reading in which the first, but not the second, plural pronoun is interpreted dependent on the distributivity over John and Mary. That is, in this reading John claims that only he loves the son he has with Mary and Mary claims that only she loves this child. The second pronoun in that reading is interpreted as coreferring with the group John and Mary.[1,2]

(5.5) John and Mary each claim that only they love their son.

If only the noun phrase 'John and Mary' were allowed to carry an index, we would not be able to represent the co-indexation patterns of this example. Assuming that distributivity operators (no matter whether they are overt of covert) carry their own index, as in (5.6), enables us to express the anaphoric relations.

(5.6) [John and Mary]$_i$ each$_j$ claim that only they$_j$ love their$_i$ son.

Such problems do not occur if one assumes that indexation is not part of the input to interpretation. For instance, specifying the interpretation of (5.3) necessarily involves having three variables.

[1]Such readings are, of course, independent of an overt occurrence of 'each'. In what follows, when discussing distributive readings, I'll often explicitly enforce them using the floating quantifier 'each'.

[2]The example in (5.5) is reminiscent of the example in (i). (See Heim 1993 and Reinhart 2000.) However, the purpose of the example in (5.5) is different.

(i) Every wife thinks only she respects her husband.

Apart from a deictic reading for 'her', there are two other readings that should be noticed. First of all, (i) could be interpreted as meaning that every wife x thinks that apart from x, no other wife y has the property 'respect(y,y's husband)'. In other words, every wife thinks she is the only good wife. In a second reading, however, every wife thinks that her husband is not respected: for every wife x, x thinks that x's husband is not respected by anyone except x. The example in (5.5) also shows these readings, but they are irrelevant to my present purpose. What is at stake here is the possibility of the two pronouns to refer either to the group 'John and Mary' or to the parts. That is, given the plurality involved in this example, even more readings than those available for (i) surface. The operator 'only' is useful to get better intuitions about such examples, but the phenomenon is independent from the presence of such an operator. For instance, (ii) also has a reading wherein the first pronoun covaries with John and Mary, while the second refers to the group John and Mary. The example in (iii) provides yet another example.

(ii) John and Mary each claim that they love their son.
(iii) John and Mary each claim that their mother spoiled their son.

(5.7) $\exists x_k(x_k = \{j, m\} \wedge \forall x_l \in x_k : \forall x_m \in x_k :$
$x_m \neq x_l \Rightarrow \text{TOLD}(x_l, x_m, \text{SHOULD-LEAVE}(x_?)))$

A resolution step completes this representation by substituting a suitable variable (one leading to an interpretable representation) for '$x_?$'. The potential 'antecedents' in (5.7) are the variables 'x_k', 'x_m' and 'x_l' and, consequently, the three possible resolutions for (5.7) correspond to the three reported readings.

In sum, assuming that indexation is part of the input to semantic interpretations makes it necessary to express all referentially relevant entities in a syntactic form. In discourse, however, it is unclear how such a syntactic form can be made sufficiently fine-grained. In the previous chapter, I pointed out a problem which in many respects resembles the problem of fine-grainedness of LF of Heim, Lasnik and May. In the scope of a distributivity operator, both global values and local *dependent* values need to be accessible. For instance, in (5.8), the pronoun 'them' should (at least) have two possible indexations, one which results in the reading in which each student sends all the written papers to L&P and one which results in a reading in which the students just send their own papers to L&P.

(5.8) Three students each wrote exactly two papers. They each sent them to L&P.

The two readings for the second sentence in (5.8) we are focusing on would have to involve different co-indexation patterns, but clearly in both cases the object plural pronoun *them* has the *same* linguistic antecedent, namely *exactly two papers*.

A solution like Heim et.al.'s seems the right way to go for intra-sentential cases of anaphora like (5.6). But as we can conclude from examples like (5.8), in discourse the facts are more complicated. Distributivity operators allow for access to not only the respective atoms in the group distributed over, but also to values dependent on them, while at the same time leaving intact the accessibility of the groups these dependent values are taken from. Distributivity thus opens up all sorts of referential possibilities and, as examples like (5.8) show, each of these possibilities should come with a unique index. Clearly, representing the ambiguity in (5.8) in syntax is not going to be easy.

An obvious solution would be to claim that the plural pronoun is ambiguous: it either takes the currently active functional value of some index or it takes the collective value of that same index. However, given that we are looking for a uniform interpretation for pronouns, we have no room for ambiguity and we will discard this option.

I conclude that in order to cover the full range of pronominal plural discourse anaphora, we will have to assume a strategy to reference resolution which is based on a semantic level. The input to the interpretation mechanism comes without indexation information. Notice, for instance, that DRT does not have the problems discussed here. After processing the first sentence of (5.8), two antecedents are formed, namely the set of three students and the set of six papers written by these students. In the second sentence, quantification over the three students takes place, triggering the accessibility of yet another set, namely the set of papers written by each of the students. The dependent and the independent value for 'the papers' are thus both represented in the semantics. This way, the pronoun has a choice between the local value and the global value for 'the papers'.

5.1.2 Indices and Assignments

We can conclude from the previous section that discourse semantics should assume that discourse pronouns choose their antecedent on the semantic level. However, dynamic formats like dynamic predicate logic and dynamic plural logic force any semantic theory based on them to make a striking assumption, namely that the input-form for interpretation is enriched with information on (co-)indexation. This is necessary because of the so-called *destructive assignment* problem. In DPL, '$\exists x$' randomly assigns a value to x. That is, if, in the incoming information state, there is already some knowledge about x available, '$\exists x$' randomly overwrites this information.Indefinites should therefore always quantify over 'fresh' variables. In a lexical approach to meaning derivation this is especially worrisome. The entry for the determiner 'a', for instance, could naively be expressed as $\lambda P.\lambda Q.\exists x(Px \wedge Qx)$. However, due to the problem just described, we will have to assume that 'a' comes with an index which is used in the interpretation. So, 'a$_i$' translates as $\lambda P.\lambda Q.\exists x_i(Px_i \wedge Qx_i)$.[3]

In DPlL, part of the destructive assignment problem is solved by assuming that an existential quantification action is undefined once the incoming information state already provides a (defined) value for the variable in question. That is, destructive assignment never really occurs, but is replaced by undefinedness. With respect to, for instance, the indefinite determiner 'a', whenever a non-fresh variable is chosen to be the ranging variable of the existential quantification, the derivation results in undefinedness. This means that in DPlL we still need a specification of the index of the determiner, this time not to protect us against loss of information, but rather to make sure the interpretation is defined.

In sum, frameworks like DPL will have to turn to decorated lexical entries once they are used for bottom-up meaning derivation. Moreover, one will have to assume that the input is properly indexed, such that, on the *logical* level, destructive assignment will not appear. This clearly goes

[3]Notice that DRT does not have the problems described here. The construction rule for indefinites, for instance, specifies that a referent should be chosen which is distinct from any referent in the universe of the DRS in which the indefinite is interpreted. This option is available due to the top down architecture of DRT. Linguistic forms are interpreted inside the current DRS, so information about the current context is always available.

against what we remarked in the previous subsection where we argued that antecedent choice should be a semantic process.

The problem is also clear in van den Berg's approach to dependent pronouns. Recall that in DPlL, the distributivity operator interprets its scope in reduced information states: only dependent values are accessible in the scope. That this is problematic follows already from example (5.5) where the global value of the group (John and Mary) and the local value of the group (first John, then Mary) are both accessed by a pronoun. Van den Berg's distributor cannot handle such examples, since there is only one index available, namely the index of the distributivity operator. The atoms ranged over in the scope of 'each' (namely the individual John and the individual Mary) have the same variable associated with them as the group referent which acts as the domain of the distributivity (namely the group John and Mary). In the scope of distributivity only the first of these is activated, outside its scope the variable takes on its global guise.

In a bottom-up approach, we will have to worry about where the multitude of indices needed for dependencies comes from. We need an operator which freely introduces new indices, as do the principles of DRT.[4] Unfortunately, *unlike* DRT, dynamic plural logic has to deal with the destructive assignment problem at the same time. The destructive assignment problem is thus directly relevant for a dynamic semantics of plurality.

In sum, what I have aimed to show is that the well-known (technical) problem of destructive assignment bears an interesting relationship with plural co-reference patterns in discourse. Some of these patterns are not expressible using (co-)indexation. So whereas destructive assignments call for index patterns to be defined on the input to the interpretation process, there are empirical clues that interpretation should not take indexation for granted.

[4]It is possible though, to come up with an operator within DPlL which would generate a sufficient amount of indices. What is needed for such an operator is a specific way of extending the context. Say that F is a set of assignments, defined for x. Distributivity with respect to x now creates a range of temporary extensions of F relating to specific values d for x in F. These extensions, $F^{\bullet}_{x=d}$, have new variables in them pointing to the local values of $\{f \in F | fx = d\}$.

$$F^{\bullet}_{x=d} := \{\{\langle v^{\bullet}, a\rangle | \langle v, a\rangle \in f \ \&\ fx = d\} | f \in F\}$$

Here $v^{\bullet}$ acts as the variable name for the zoomed-in value of global variable v. Of course, we do not have any information on whether $v^{\bullet}$ is already defined or not.

$$F\, [\![\delta'_x(\Gamma)]\!]\, G \quad :\Leftrightarrow \quad \forall d \in \{fx | f \in F\} : F \cup F^{\bullet}_{x=d}\, [\![\Gamma[x/x^{\bullet}]]\!]\, F \cup G^{\bullet}_{x=d}$$

See section 5.4, for a development of this idea in a setting which does not display the destructive assignment problem.

5.2 Eliminating destructive assignment

5.2.1 Index and Context

Van den Berg does not attempt a system for a bottom up derivation of his formulae and that is why the undefinedness conditions of the existential quantifier solve the destructive assignment problem for his purposes; a formula which potentially destroys information is simply uninterpretable. In order to successfully randomly assign an individual to x, the input state has to provide the information that no value for x has yet been defined.

An alternative to van den Berg's solution can be found in Bekki 2000b and Bekki 2000a. In Bekki's typed dynamic logic (TDL), the initial context is one containing all possible (sets of) complete assignment functions. Furthermore, the language does not really have an existential quantifier action, but rather contains two special predicates: 'FREE' and 'BOUND'. The predicate 'FREE' tests whether an argument index x is still associated with the full range of possibilities according to the incoming context. In other words, '$\text{FREE}(x)$' tests whether x is assigned the full domain of entities. Its counterpart 'BOUND' is successful whenever 'FREE' is not. Anaphoric noun-phrases thus carry a condition that the index associated with it is bound, while non-anaphoric ones carry a freeness condition. The indefinite 'a man', for instance, is derive as $\lambda P.\text{FREE}(x) \wedge \text{MAN}(x) \wedge P(x) \wedge 1(x)$, where x is an index which is 'selected properly so that it succeeds in the free(x) test' (Bekki 2000b, lexicon 3). That is, once the wrong index is selected, the free-test will not succeed and the interpretation will result in the empty-set.

Just like van den Berg's partial definition of random assignment trades destructive assignment for undefinedness, Bekki's proposal simply trades it in for unsatisfiability. Both van den Berg and Bekki, however, seem to realise that in meaning derivation, the choice of indices should depend on the context of interpretation. The 'FREE' predicate signals in any context which variables are fresh and which are not, just like variables which are defined in an information state in DPlL are never fresh, while undefined ones are.

Van Eijck (2001) offers a solution to the destructive assignment problem which fully incorporates the intuition that indexation is context-driven. In what follows, it is van Eijck's solution I will adopt. Roughly, his proposal is as follows. Say that the set of variable names is linearly ordered. Say, for instance, that they are the set of natural numbers. Now we construct actions which not just pick any fresh variable, but rather the lowest natural number that is not yet in the incoming context. This way, the first fresh variable will be 0, the next one 1, etc. It ensures us of a clean and economical handling of variable names. If we now assume that utterances denote context transitions, we have the relevant information for variable choice always available, since there is a systematic way of choosing the *next* fresh variable. The following section will present such a system in much more detail.

5.2.2 Incremental Dynamics

Incremental dynamics (van Eijck 2001) is a dynamic logic for incremental dynamic interpretation[5] which has as its main feature an essentially variable free way of dynamic binding. In a series of papers (see for instance, van Eijck 2000b and van Eijck 2000a), formal aspects and linguistic applications of this logic are discussed. In what follows, my particular interest is its suitability to be incorporated in a bottom-up approach to natural language interpretation.

I start with a loose presentation of the formalisms that play a crucial role in the analysis to follow. The goal is to communicate the general idea behind incremental dynamics and its type theoretical implementation, so that the reader can get accustomed to the (from her point of view perhaps slightly unorthodox) system of variable free dynamic binding.

Incremental dynamics can be seen as a semantics operating on stacks of individuals[6]. A stack of individuals can be seen as a special kind of assignment function. It maps positions in the stack to the individual in this position. For instance, a stack $\boxed{a \mid b \mid c}$ is a function f mapping positions to individuals. If we count positions from left to right and the left-most position is called 0, then $f(0) = a$, $f(1) = b$ and $f(2) = c$. But unlike variable labels in ordinary assignment functions, the position associated with an individual is flexible. For instance, if we combine two stacks f and g, by pushing g on to f (i.e. attaching it on the right edge of f), then the individuals in g will have new positions associated to them. If $g(i)$ returns some individual d, then in the combined stack this individual can be found in position $|f| + i$, where $|f|$ is the length of stack f, as illustrated in (5.9). This shows that stacks enable us to combine two assignments without danger of variable clashes. The labels associated with entities are contextual, they depend on what the stack looks like.

(5.9)

$$\begin{array}{ccccc} 0 & 1 & 2 & \cdots & n \\ \hline \multicolumn{1}{|c|}{e_0} & \multicolumn{1}{c|}{e_1} & \multicolumn{1}{c|}{e_2} & \multicolumn{1}{c|}{\cdots} & \multicolumn{1}{c|}{e_n} \\ \hline \end{array} \wedge \begin{array}{ccccc} 0 & 1 & 2 & \cdots & m \\ \hline \multicolumn{1}{|c|}{e'_0} & \multicolumn{1}{c|}{e'_1} & \multicolumn{1}{c|}{e'_2} & \multicolumn{1}{c|}{\cdots} & \multicolumn{1}{c|}{e'_m} \\ \hline \end{array}$$

$$=$$

$$\begin{array}{cccccccccc} 0 & 1 & 2 & \cdots & n & n+1 & n+2 & n+3 & \cdots & n+1+m \\ \hline \multicolumn{1}{|c|}{e_0} & \multicolumn{1}{c|}{e_1} & \multicolumn{1}{c|}{e_2} & \multicolumn{1}{c|}{\cdots} & \multicolumn{1}{c|}{e_n} & \multicolumn{1}{c|}{e'_0} & \multicolumn{1}{c|}{e'_1} & \multicolumn{1}{c|}{e'_2} & \multicolumn{1}{c|}{\cdots} & \multicolumn{1}{c|}{e'_m} \\ \hline \end{array}$$

[5] To avoid confusion, it might be wise to remark immediately that in incremental dynamic interpretation, it is the dynamics, rather than the interpretation, which is incremental.

[6] The use of the term *stack* rather than queue, or list, or sequence is not crucial. The formal differences between these different notions of data structure do not really matter here. In general, *stacks* are lists where the last added element is the first one to be retrieved. *Queues*, on the other hand, push new material to the back of the list. Since, as we will see, we push individuals to the top of the list, *stack* would be a more appropriate name for our structures, than queue.

Using stacks, it is possible to define an existential quantifier which does not randomly assign, but which extends the state with another individual. Given what we have said so far about stacks it will be clear that the label for this new individual follows from the stack it is pushed on to. That is, if we know what the input stack looks like, then we will automatically know the label for the individual we are introducing. If g is a stack and d an individual, pushing d to g will result in a stack g', where d is on the $|g|$th position, since if we start counting at zero, the top position of g is labeled $|g| - 1$. A form of existential quantification is then possible which is guaranteed to be non-destructive. Rather than assigning an individual to a label from an independently defined domain of labels it *increments* the context, hence the name incremental dynamics.

This way, quantification has become variable free. The predicate logical formula $\exists x(Px)$ becomes $\exists; Pn$, once we know that the length of the input stack is n[7]. We start with the details of incremental dynamics by defining a variable free predicate logical language.

Definition 5.8

The language $\mathcal{L}_{\text{ID}}$ is the smallest set such that:

$$\begin{array}{rcll} \exists & \in & \mathcal{L}_{\text{ID}} & \\ Pv_1 \ldots v_n & \in & \mathcal{L}_{\text{ID}} & \text{if } \forall 1 \leq i \leq n : v_i \in \mathbb{N} \\ \neg(\varphi) & \in & \mathcal{L}_{\text{ID}} & \text{if } \varphi \in \mathcal{L}_{\text{ID}} \\ \varphi; \psi & \in & \mathcal{L}_{\text{ID}} & \text{if } \varphi \in \mathcal{L}_{\text{ID}} \ \&\ \psi \in \mathcal{L}_{\text{ID}} \end{array}$$

□

We interpret this simple language using the function $[\![\cdot]\!]$ from formulae in $\mathcal{L}_{\text{ID}}$ to relations between information states. We assume a fixed model $M = \langle D_e, I \rangle$, which is left implicit in the interpretation notation. An information state is a pair (n, f), where n is a natural number and f is a function from $\{0, \ldots, n\text{-}1\}$ to D_e. The intuition is that n represents the universe of discourse by specifying its size. The function f collects the information known about this universe. In the stack terminology discussed above, f is a stack and n specifies its length. In this light, predicate-argument structures are interpreted in the way familiar from dynamic predicate logic: they perform a test on the information state. A condition should be made with respect to the arguments of the predicate, however. A predication $Pv_1 \ldots v_k$ is undefined in context n if any of its arguments v_i exceeds n.

(5.10)
$$\begin{array}{rcll} (n, f)\,[\![Pv_1 \ldots v_k]\!]\,(m, g) & = & \uparrow & :\Leftrightarrow\ \exists 1 \leq i \leq k :\ v_i \geq n \\ (n, f)\,[\![Pv_1 \ldots v_k]\!]\,(m, g) & = & 1 & :\Leftrightarrow\ (h, f)\,[\![Pv_1 \ldots v_k]\!]\,(m, g)\ \neq \uparrow\ \& \\ & & & \quad n = m\ \&\ f = g\ \& \\ & & & \quad \langle f(v_1), \ldots, f(v_k) \rangle \in I(P) \end{array}$$

[7]Variable free formalisms have a rich history. The key work is Quine 1960.

Negation is also a test: it returns an incoming information state as soon as its scope cannot be interpreted successfully.

$$
\text{(5.11)}\quad
\begin{array}{lcll}
(n,f)\,[\![\neg(\varphi)]\!]\,(m,g) & = & \uparrow & :\Leftrightarrow\ \exists m',g' : (n,f)\,[\![\varphi]\!]\,(m',g') = \uparrow \\
(n,f)\,[\![\neg(\varphi)]\!]\,(m,g) & = & 1 & :\Leftrightarrow\ (n,f)\,[\![\neg(\varphi)]\!]\,(m,g) \neq \uparrow\ \& \\
 & & & \quad n = m\ \&\ f = g\ \& \\
 & & & \quad \neg\exists n'\exists f' :\ (n,f)\,[\![\varphi]\!]\,(n',f')
\end{array}
$$

The existential quantifier in $\mathcal{L}_{\text{ID}}$ is bare. It does not come with a scope, nor does it carry an index. The action $\exists$ is nothing more than a push operator, extending the context with an extra position.

$$
\text{(5.12)}\quad
\begin{array}{lcll}
(n,f)\,[\![\exists]\!]\,(m,g) & = & \uparrow & :\Leftrightarrow\ \bot \\
(n,f)\,[\![\exists]\!]\,(m,g) & = & 1 & :\Leftrightarrow\ m = n+1\ \&\ \exists d \in D :\ g = f \cup \{\langle n,d\rangle\}
\end{array}
$$

In an empty information state $(0,\emptyset)$, an $\exists$-action will result in a context of length one, with a single assignment (to slot 0). In general, the existential quantifier extends the context and adds a new individual to the information state. Combinations of formulae are interpreted by simple relation composition.

$$
\text{(5.13)}\quad
\begin{array}{lcllcll}
(n,f)\,[\![\varphi]\!]\,;\,[\![\psi]\!]\,(m,g) & = & \uparrow & :\Leftrightarrow \neg\exists l\exists h : & (n,f)\,[\![\varphi]\!]\,(l,h) & \neq & \uparrow\ \& \\
 & & & & (l,h)\,[\![\psi]\!]\,(m,g) & \neq & \uparrow \\
(n,f)\,[\![\varphi]\!]\,;\,[\![\psi]\!]\,(m,g) & = & 1 & :\Leftrightarrow\ \exists l\exists h : & (n,f)\,[\![\varphi]\!]\,(l,h) & = & 1\ \& \\
 & & & & (l,h)\,[\![\psi]\!]\,(m,g) & = & 1
\end{array}
$$

$$
\text{(5.14)}\quad [\![\varphi;\psi]\!] \;:=\; [\![\varphi]\!]\,;\,[\![\psi]\!]
$$

Since pairs (n,f) can be seen as sequences of individuals of length n, we may want to exploit this in the specification of the semantics. The function f in a pair (n,f) is a set of pairs $\{\langle 0,d_0\rangle,\ldots,\langle n-1,d_{n-1}\rangle\}$. Let $c,c',\ldots$ range over sequences of individuals. We write $c^\wedge c'$ for the stack resulting from appending the stack c' to c. That is, if c corresponds to (n,f) and c' corresponds to (m,g):

Definition 5.9

Append

$$
c^\wedge c' := f \cup \{\langle i+|c|,d\rangle \mid \langle i,d\rangle \in g\}
$$

□

We will write $c^\wedge d$, where $d \in D_e$, for the stack resulting from pushing d to the top of c. That is, $c^\wedge d$ is shorthand for $c^\wedge\{\langle 0,d\rangle\}$. Given such an operation, we can rephrase the existential quantifier as a simple push action.

$$
\text{(5.15)}\quad
\begin{array}{lcll}
c\,[\![\exists]\!]\,c' & = & \uparrow & :\Leftrightarrow\ \bot \\
c\,[\![\exists]\!]\,c' & = & 1 & :\Leftrightarrow\ \exists d \in D_e :\ c' = c^\wedge d
\end{array}
$$

The relational semantics of existential quantification as a push operation satisfies the requirement of quantifying over a fresh variable in every context. Such a semantics should be supplemented with a theory of how predicates are combined with the right indices. For instance, 'a man' should receive translation $\exists; \text{MAN}(0)$ if uttered in an empty information state, but $\exists; \text{MAN}(23)$ in a context counting 23 individuals. In order to establish the link between the index supplying the argument of the predicate and the index contributed by the determiner we need to specify the derivation procedure. Van Eijck (2000b) develops an extension of incremental dynamic semantics with a flexible typing mechanism which enables a Montague style semantics.[8]

Definition 5.10

The set of types $\mathcal{T}$

$$\begin{array}{rcll} \mathcal{C} & := & \{[e]_i \mid i \in \mathbb{N}\} & \text{types of contexts} \\ \mathcal{BT} & := & C \cup \{t, \iota\} & \text{basic types} \\ \mathcal{T} & := & \mathcal{BT} \cup \{\langle \alpha, \beta \rangle \mid \alpha, \beta \in \mathcal{T}\} & \end{array}$$

The set of basic types consists of the type of indices, ι, the type of truth-values, t, and a family of types for contexts. That is, a context c is of type $[e]_{|c|}$, the type of stacks of length $|c|$. This means that linguistics expressions will generally be associated not with a single type, but rather with a family of types. For instance, a noun (or any other predicational element) will be interpreted as a function from indices to context transitions, so it will have the type family $\langle \iota, \langle [e]_i, \langle [e]_i, t \rangle \rangle \rangle$. In what follows we will mostly leave the size of a context implicit in its type and write $[e]$ to generalise over the types for information states.

With the move toward a typed semantics, I will also shift from the relational perspective on incremental dynamics to a functional one. That is, saturated expressions denote update functions of type $\langle [e], \langle [e], t \rangle \rangle$: they take a stack and return a set of stacks. We will abbreviate this type by T, the type of *context transitions*.

In order to give a typed fragment using functional incremental dynamics, I define functional counterparts for ';' and '$\exists$', unary predicates P and binary predicates R. We use $\gamma :: \tau$ to express that γ is of type τ.

[8]There is a rich history of compositional approaches to dynamic semantics and compositional variants of DRT. See, for instance, Zeevat 1989; Asher 1993; Muskens 1994; Muskens 1996; Bos, Mastenbroek, McGlashan, Millies, and Pinkal 1994; Kuschert 1995; Kohlhase, Kuschert, and Pinkal 1996.

Definition 5.11

Functional semantics for Incremental Dynamics:

$$\begin{array}{rcll} \varphi \cdot \psi & := & \lambda c.\lambda c'.\exists c''(c'' \in \varphi c \wedge c' \in \psi c'') & :: \; T \\ \exists^{\diamond} & := & \lambda c.\lambda c'.\exists d(c' = c\hat{\ }d) & :: \; T \\ P^{\diamond} & := & \lambda i \lambda c.\lambda c'.c = c' \wedge P(c(i)) & :: \; \langle \iota, T \rangle \\ R^{\diamond} & := & \lambda i \lambda j.\lambda c.\lambda c'.c = c' \wedge R(c(j), c(i)) & :: \; \langle \iota, \langle \iota, T \rangle \rangle \end{array}$$

□

We have now circumvented any representational level. $\exists^{\diamond}$, $P^{\diamond}$ and $R^{\diamond}$ are simply functions: operations on functions. The function $P^{\diamond}$ is the functional counterpart of the model theoretic predicate symbol P. That is, whereas P is interpreted by the model to yield a set of individuals $I(P)$, $P^{\diamond}$ tests whether the argument index i points to an individual in $I(P)$ in the incoming context c.

Notice that given the abbreviation T, the types begin to resemble the Montagovian types more. For instance, nouns will be of type $\langle \iota, T \rangle$, while determiners are of type $\langle \langle \iota, T \rangle, \langle \langle \iota, T \rangle, T \rangle \rangle$, etc.

Here is an example derivation for the noun phrase 'a man':

Derivation 5.12

1. man $\rightsquigarrow \lambda i.\lambda c.\lambda c'.c = c' \wedge \text{MAN}(c(i)) \quad = \text{MAN}^{\diamond} \quad :: \langle \iota, T \rangle$
2. a $\rightsquigarrow \lambda P.\lambda Q.\lambda c.((\exists^{\diamond} \cdot P(|c|) \cdot Q(|c|))c) \quad :: \langle \langle \iota, T \rangle, \langle \langle \iota, T \rangle, T \rangle \rangle$
3. a man $\rightsquigarrow \lambda Q.\lambda c.((\exists^{\diamond} \cdot \text{MAN}^{\diamond}(|c|) \cdot Q(|c|))c) \quad :: \langle \langle \iota, T \rangle, T \rangle$

□

Notice how the lexical entry for 'man' is directly translated into the function '$\text{MAN}^{\diamond}$'. Thus, it takes an index u and returns a stack transition function which returns the input stack c if and only if the value c assigns to u is a man. The determiner 'a' takes two functions P and Q of type $\langle \iota, T \rangle$ and returns a stack transition function.

Figure 5.1 exemplifies a small fragment (compare with van Eijck 2000a, p. 5). Here, the left-hand column represents the syntax, while the right-hand column represents the interpretation. In general a syntactic structure P formed by combining $C_1, \ldots, C_n$ is interpreted as X, which corresponds to some combination of the interpretations X_1 of C_1 to X_n of C_n.

Given this fragment, it is easy to derive an interpretation for a simple sentence like 'a man loved a woman.'

Derivation 5.13

1. loved a woman $\rightsquigarrow \lambda j.(\lambda Q.\lambda c'.(\exists^{\diamond} \cdot \text{WOMAN}^{\diamond}(|c'|) \cdot Q(|c'|))c' \; (\lambda i.\text{LOVE}^{\diamond} ij))$
2. $= \lambda j.\lambda c'.(\exists^{\diamond} \cdot \text{WOMAN}^{\diamond}(|c'|) \cdot \text{LOVE}^{\diamond}(|c'|)j)c'$

S	::=	S.S	X	::=	$X_1 \cdot X_2$
S	::=	NP VP	X	::=	$(X_1 X_2)$
NP	::=	Det CN	X	::=	$(X_1 X_2)$
VP	::=	IV	X	::=	X_1
VP	::=	TV NP	X	::=	$\lambda i.(X_2(\lambda j.(X_1 j)i))$
TV	::=	love	X	::=	$\text{LOVE}^\diamond$
CN	::=	man	X	::=	$\text{MAN}^\diamond$
CN	::=	woman	X	::=	$\text{WOMAN}^\diamond$
DET	::=	a	X	::=	$\lambda P^\diamond.\lambda Q^\diamond.\lambda c.(\exists^\diamond \cdot P^\diamond(\lvert c\rvert) \cdot Q^\diamond(\lvert c\rvert))c$

Figure 5.1: A small fragment

3. a man loved a woman $\rightsquigarrow$
   $\lambda Q.\lambda c.((\exists^\diamond \cdot \text{MAN}^\diamond(\lvert c\rvert) \cdot Q(\lvert c\rvert))c)\ (\lambda j.\lambda c'.(\exists^\diamond \cdot \text{WOMAN}^\diamond(\lvert c'\rvert) \cdot \text{LOVE}^\diamond(\lvert c'\rvert)j)c')$
4. $= \lambda c.((\exists^\diamond \cdot \text{MAN}^\diamond(\lvert c\rvert) \cdot \lambda c'.(\exists^\diamond \cdot \text{WOMAN}^\diamond(\lvert c'\rvert) \cdot \text{LOVE}^\diamond(\lvert c'\rvert)(\lvert c\rvert))c')c)$
5. $= \lambda c.((\exists^\diamond \cdot \text{MAN}^\diamond(\lvert c\rvert) \cdot \exists^\diamond \cdot \text{WOMAN}^\diamond(\lvert c\rvert{+}1) \cdot \text{LOVE}^\diamond(\lvert c\rvert{+}1)(\lvert c\rvert))c)$

□

The final step deserves some explanation. It illustrates that even without indices in the input, logical syntax will take care of the distribution of indices to arguments of predication. A context transition like $\lambda c.((\exists^\diamond \cdot P^\diamond(\lvert c\rvert))c)$ makes sure that the predicate P takes the individual introduced by the quantifier as its argument. In context, however, the appearance of this index might change. For instance, $\lambda c'.((\exists^\diamond \cdot [\lambda c.(\exists^\diamond \cdot P^\diamond(\lvert c\rvert))c]\ )c')$ equals $\lambda c'.((\exists^\diamond \cdot (\exists^\diamond \cdot P^\diamond(\lvert c{+}1\rvert))c')$, since with respect to the c' context, the quantifier in question is preceded by another quantifier.

In sum, we saw how the shift from contexts as assignment functions to contexts as stacks allows for an interpretive mechanism which distributes indices 'on the go.' Remember that using DPL makes it impossible to formulate a unique lexical entry for an expression that introduces discourse entities. This is due to the fact that the choice of label for this entity would be informative in one but destructive in another context. In a framework like incremental dynamics, such a problem has disappeared, since it makes the index-choice fully context-dependent.

5.3 Incremental dynamics and plurality

The concise guide to incremental dynamics in the previous section showed us the basic tools: a push operator on a stack and several tests defined on stacks. van Eijck (2001) shows that there is a natural translation from (the non-typed version of) ID to dynamic predicate logic. In fact, this translation results in a version of dynamic predicate logic which lacks the destructive assignment property. This close link to DPL suggests that a reformulation of Van den Berg's dynamic plural logic in ID-style should be straightforward. In this section I present such a reformulation.

5.3.1 Data Structure

In *incremental dynamic plural logic* (IDPlL), the incremental variation of DPlL we are about to construct, van den Berg-style data structures come down to sets of stacks of the same length. Recall that van den Berg uses sets of partial assignments as data structure. It is not straightforward, however, how van den Berg's use of partiality translates to incremental dynamics. I will follow van den Berg's choices as closely as possible. However, apart from the notion of partiality which is inherent in stacks, in contrast to van den Berg, I assume no undefinedness dummy value '$\star$'. That is, in IDPlL, we will use a data structure which is based on stacks which are total: a stack c is defined on all positions $i < |s|$ and undefined on all other positions.

Before presenting the data structure involved in IDPlL, let us slightly digress and review the need for partial assignments. This is a digression, since it requires us to look ahead at our treatment of quantificational noun phrases in chapter 6.

Van den Berg (1996b) gives three reasons for working with partial rather than total assignments. First, it provides a natural treatment of information increase where as the discourse continues more and more variables (standing for 'topics') are defined in the state. Second, it provides a way to represent the empty set by assigning no individual to a variable. Third, it provides a sound way to define a notion of subset.

In incremental dynamics, there are two kinds of partiality we should keep apart. Notice first of all that, clearly, any stack c fails to be defined on indices $i \geq |c|$. There is, however, a second way the stacks in ID (and in parallel the states based on sets of these stacks) could involve a notion of partiality, namely when there is undefinedness with respect to some $i < |s|$.

Van den Berg's first argument for partiality plays no role in the formalism presented in this chapter. Information increase follows from the very philosophy of incremental dynamics, namely that context is *incremented* by existential quantification rather than rewritten. Consequently, we do not need to assume stacks with holes for this reason.

Related to this is the issue of emptiness. Anaphoric reference is never empty, so the storage of an empty value for some discourse marker amounts to the unaccessibility of that marker. In incremental dynamics this is even clearer. The difference between the stacks c and $c^\wedge\star$ is only that their length seems to differ. But surely both $c(|c|)$ and $(c^\wedge\star)(|c|)$ are undefined. It is then a question of ontology (with respect to the nature of '$\star$') whether $c^\wedge d$ equals $c^\wedge\star^\wedge d$. In chapter 6, where we give a semantic for quantificational noun phrases, we will show that the assumption that the introduction of an empty set in discourse does not change the context at all yields correct results for anaphoric reference and entailment patterns.

The vital reason for an undefinedness value *within* stacks thus seems to be the notion of subset it provides us. Let us review this argument once more. The problem noted in chapter 4 consisted of the fact that when introducing a subset X of say Y, the original superset Y could be in a dependency relation with some other set (say, Z). Some of the atoms in Z, however, are then unrelated to any of the atoms in X. For instance, if Tom admires John F. Kennedy, Dick admires Robert Kennedy and Harry admires Ted Kennedy, then for the subset Tom and Dick there are initially two options. Either they each are paired with the three Kennedys, or they are only paired with the one Kennedy the admire. In this latter case we will need a dummy value to be inserted in the function containing Robert Kennedy, since he is paired with neither Tom nor Dick.

So why do subsets need to preserve dependencies in the first place? The reason is that should we mangle with the dependencies of the atoms, we make wrong prediction about dependent pronouns. The example in (5.16), for instance, shows that each student should be paired with his or her own paper only.

(5.16) Every student wrote a paper.
Each advanced student submitted it to L&P.

The question is, however, why the above example should introduce a subset in the first place. In chapter 6, we will develop the tools to analyse this example as follows. First, 'every student wrote a paper' introduces a context containing the relevant student-paper pairs, just like it would in DPlL. The second sentence, however, can be seen as a quantification of a domain parasitic on the first sentence. So, we are quantifying over the student-paper pairs but we ignore all non-advanced students. That is, we extend the context of student-paper pairs with another student-paper pair if and only if the student is advanced and the paper was submitted to L&P by him or her. The result will be that the context now contains not only student-paper pairs, but that it is independently extended with advanced student-submitted paper pairs. (See chapter 6 for details.) The upshot is that, in our analysis of (5.16), the second sentence does not only (independently) introduce a set of advanced students, it moreover (independently) introduces a corresponding set of papers. With respect to accessibility this means that a pronoun may either refer to *all the papers* or it covaries with the advanced students. This way, no unwanted dependencies arise and the need for a '$\star$'-value disappears.

End of digression. We may now turn to the definition of the data structure we will use in IDPlL. A *state* or a *possibility* is a set of stacks of equal length. That is, the set of states of length n is the set of sets of stacks of length n consisting of atomic individuals. Members of this set will be written as s, s', etc., while the individual stacks in such states are referred to as c, c', etc. The type of a set of stacks of length n is $[E]_n$. In (5.17), I illustrate such states.[9]

Definition 5.14

Possibilities in IDPlL

$$s :: [E]_n \;\Leftrightarrow\; s \subseteq \{\langle d_1, \ldots, d_n\rangle \mid \forall i : 1 \le i \le n \;\rightarrow d_i \in D_e\}$$

□

(5.17)

s	0	1	2	$\cdots$		$\lvert s\rvert$-1
c				$\cdots$		
c'				$\cdots$		
c''				$\cdots$		
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\cdots$	$\vdots$	$\vdots$

As in ordinary incremental dynamics, the stacks are built up from left to right. That is, the left-most position is called 0, containing the 'oldest' material. With respect to such states, we will write '$s[i]$' for the collection of all the values at position i in the stacks in s. So, $s[i] = \{c(i) | c \in s \;\&\; c(i) =\downarrow\}$. (Of course, $s[i]$ is undefined whenever i exceeds the size of s.) Also, we abuse cardinality notation by using '$|s|$' to refer to the (common) length of the stacks in s. Note that we will assume all stacks to have the same length.

It is difficult to compare a plural version of ID with DPlL, since there are different languages involved (there are no quantified variables in ID) and the nature of ID's index control makes some of the partiality definitions of DPlL unnecessary. I will therefore start not by constructing a proper 'translation', but rather compare DPlL atomic formulae with potential ID counterparts. See the appendix 5.A for some remarks on a more formal relationship.

5.3.2 Incremental dynamic plural logic

Remember that existential quantification in DPlL was defined partially in order to avoid variable clashes. This is not necessary in incremental dynamics, since the quantifier is bare.

(5.18)
$$\begin{array}{llll} s[\![\epsilon]\!]s' & =\uparrow & :\Leftrightarrow & \bot \\ s[\![\epsilon]\!]s' & = 1 & :\Leftrightarrow & \exists X \in \wp^+(D_e) : \; s' = \{c^\frown d | d \in X \;\wedge c \in s\} \end{array}$$

Just like 'ϵ_x', the action ϵ is *independent*. That is, introduction of a plurality by ϵ does not result in any functional dependencies. For instance, in a state

[9]Notice that according to this definition s is allowed to be the empty set. This means that we should have empty sets of multiple types. I leave the formal exposition of this typing implicit.

in which we only have the plurality Tom and Harry, introduction of another plurality results in a state in which all the possible pairings between the atomic individuals in the two pluralities are represented. So, (5.19a) is an instance of the ϵ relation, since it introduces no dependencies, but (5.19b) is not, since the representation of the two pluralities is not arbitrary.

(5.19) (a) $\begin{array}{|c|}\hline t\\\hline h\\\hline\end{array}\ [\![\epsilon]\!]\ \begin{array}{|c|c|}\hline t & t\\\hline t & d\\\hline t & h\\\hline h & t\\\hline h & d\\\hline h & h\\\hline\end{array}$ *(b) $\begin{array}{|c|}\hline t\\\hline h\\\hline\end{array}\ [\![\epsilon]\!]\ \begin{array}{|c|c|}\hline t & t\\\hline t & d\\\hline h & h\\\hline\end{array}$

In order to say more about dependence and distributivity, we need a way of accessing specific parts of a set of stacks. The set of stacks '$s\upharpoonright_{i=d}$' collects those stacks from s that have 'd' in position i.

Definition 5.15

Sub-states

$$s\upharpoonright_{i=d} := \begin{cases} \{c \in s | c(i) = d\} & i < |s| \\ \uparrow & \text{otherwise} \end{cases}$$

□

The formal notion of dependence is comparable with van den Berg's definition (see page 84).

(5.20) In information state s, j is dependent on i if:
$\exists d, e \in s[i] :\ s\upharpoonright_{i=d} [j] \neq s\upharpoonright_{i=e} [j]$

Turning now to simple (static) negation and distributivity, we copy the DPlL definition and apply it to sets of stacks.

(5.21) $s[\![\neg(\varphi)]\!]s' = \uparrow \quad :\Leftrightarrow \quad \exists s' : s\,[\![\varphi]\!]\,s' = \uparrow$
$s[\![\neg(\varphi)]\!]s' = 1 \quad :\Leftrightarrow \quad s = s' \ \&\ \neg\exists s'' : s\,[\![\varphi]\!]\,s''$

(5.22) $s[\![\delta_i(\varphi)]\!]s' = \uparrow \quad :\Leftrightarrow \quad i \geq |s| \text{ or } \exists d \in s[i] : s\upharpoonright_{i=d}\ [\![\varphi]\!]\ s'\upharpoonright_{i=d} = \uparrow$
$s[\![\delta_i(\varphi)]\!]s' = 1 \quad :\Leftrightarrow \quad s[i] = s'[i]\ \&\ s[\![\delta_i(\varphi)]\!]s' \neq \uparrow\ \&$
$\forall d \in s[i] : s\upharpoonright_{i=d}\ [\![\varphi]\!]\ s'\upharpoonright_{i=d}$

Given the discussion concerning indexation and variable clashes at the start of this chapter, the index on the distributivity operator might come as a surprise. However, we may assume that syntax will take care of the choice for i. Basically, i is the index which represents the group the distributivity operator takes the atomic values from. Roughly, this is the index introduced by the NP the operator is attached to. Or, if we insist the operator modifies the VP, the index corresponds to the index of the open slot of the verb phrase.[10]

The case of predication is more difficult. Remember that in van den Berg's dynamic plural logic, a predication $P(x)$ in a context in which x was not defined is interpreted as being false. This is because $F\,[\![P(x)]\!]\,F$ is valid if and only if the set $\{f(x)|f \in F \ \&\ f(x) \neq \star\}$ is in the extension of P. In case F is undefined for x, all functions f in F will assign $\star$ to x and therefore $F\,[\![P(x)]\!]\,F$ amounts to testing whether the empty set is in $I(P)$.

In incremental dynamics, undefinedness is likely to be interpreted rather differently. Given some stack s and some $i > |s|$, intuitively, $s[i]$ will be undefined rather than return the empty set. This view will become important in resolution (see section 5.5), since it gives us a way of identifying accessibility with definedness. The idea is that in a context represented by a stack (or a set of stacks) s, only indices $i < |s|$ are accessible since $P(i)$ is only defined in those cases.[11] Consequently, we need a partial definition, since we need to ensure there are no arguments that refer to inexistent slots.

(5.23)
$$\begin{array}{lcl} s[\![P(n,\ldots,m)]\!]s' = \uparrow & :\Leftrightarrow & \max(\{n,\ldots m\}) \geq |s| \\ s[\![P(n,\ldots,m)]\!]s' = 1 & :\Leftrightarrow & s = s' \ \&\ s[\![P(n,\ldots,m)]\!]s' \neq \uparrow \ \& \\ & & \langle s[n],\ldots,s[m]\rangle \in I(P) \end{array}$$

As we will soon see, we need not worry about this partiality. Reasonable natural language fragments using the functional counterpart of the IDPlL formalism will never derive functions that in combination with some context return undefinedness. A quick way of deriving function variants of the relational definitions above is by defining the (update) functions in terms of $[\![\cdot]\!]$.

(5.24) $$\Gamma^{\spadesuit} := \lambda s.\lambda s'.s[\![\Gamma]\!]s'$$

These functions are partial as well. For instance, $(P(i))^{\spadesuit}(s)$ is undefined as soon as $i \geq |s|$. By defining some of the functions separately, we can illustrate that the need for partiality is special to predication. (We ignore distributivity for now.) That is, given definedness:

(5.25)
$$\begin{array}{rcll} \exists^{\star} & := & \lambda s.\lambda s'.\exists X \in \wp^{+}(D_e) : s' = \{c^\frown d | c \in s \ \&\ d \in X\} & :: T \\ \neg^{\star}\varphi & := & \lambda s.\lambda s'.s = s' \ \&\ \varphi s = \emptyset & :: T \\ \varphi \cdot \psi & := & \lambda s.\bigcup\{\psi s' | s' \in \varphi s\} & :: T \end{array}$$

Here, in accordance with the change in data structure, we use T to abbreviate $\langle [E], \langle [E], t\rangle\rangle$, the type of state transitions. It is easy to see that for any s of type $[E]$, these definitions result in something of type $\langle [E], t\rangle$, a set of states. The update function for predication will stay partial, however, since the predicate's arguments might not be suitable to some of the states we apply the function to.

[10] In section 5.4, I show dropping the index is straightforward.

[11] This, once more, shows the attractiveness of doing without a $\star$-value.

(5.26)

$$P^\star \quad := \quad \lambda i \ldots \lambda j.\lambda s. \begin{cases} \uparrow & \max(\{i, \ldots j\}) \geq |s| \\ \lambda s'.s = s' \wedge P(s[j], \ldots, s[i]) & \text{otherwise} \end{cases}$$

This partiality is harmless, however, if we can show that in the fragments we build such a situation can never occur. This is quite straightforward since, in practice, the indices of the arguments of a predicate are determined by the length of the context. Whether or not our semantics is total thus relies solely on the fragment we built.

Let me stress that as in our presentation of incremental dynamics in section 5.2.2, we again have done away with any level of representation. We only have contexts and operations on contexts to deal with. $\exists^\star$, e.g., does not *represent* a function, it *is* a function.[12] This lack of representational level does not stop us from defining truth, validity and entailment in the normal way.[13]

Definition 5.16

Truth, validity and entailment. For $\varphi, \psi :: T$

$$\begin{array}{rcll} \models_s \varphi & \Leftrightarrow & \varphi(s) \neq \emptyset & \text{(truth)} \\ \models \varphi & \Leftrightarrow & \forall s : \varphi(s) =\downarrow\rightarrow \varphi(s) \neq \emptyset & \text{(validity)} \\ \varphi \models \psi & \Leftrightarrow & \forall s, s' : s' \in \varphi(s) \ \&\ \psi(s') =\downarrow\rightarrow \psi(s') \neq \emptyset & \text{(entailment)} \end{array}$$

□

Before presenting an example fragment, it is useful to define a special unary predicate, which enables us to count the atoms of pluralities. So for $n \in \mathbb{N}$:

(5.27) $$n^\star \quad := \quad \lambda i.\lambda s. \begin{cases} \uparrow & i \geq |s| \\ \lambda s'.s = s' \ \&\ |s[i]| = n & \text{otherwise} \end{cases}$$

Figure 5.2 represents a simple plural fragment. We derive 'two men love a woman' with it.

Derivation 5.17

1. loved a woman $\rightsquigarrow$
   $\lambda j.(\lambda Q.\lambda s'.(\exists^\star \cdot 1^\star(|s'|) \cdot \text{WOMAN}^\star(|s'|) \cdot Q(|s'|))s' \ (\lambda i.\text{LOVE}^\star ij))$
2. $= \lambda j.\lambda s'.(\exists^\star \cdot 1^\star(|s'|) \cdot \text{WOMAN}^\star(|s'|) \cdot \text{LOVE}^\star(|s'|)j)s'$
3. two men loved a woman $\rightsquigarrow$ $\lambda Q.\lambda s.(\exists^\star \cdot 2^\star(|s|) \cdot \text{MAN}^\star(|s|) \cdot Q(|s|))s$
   $[\lambda j.\lambda s'.(\exists^\star \cdot 1^\star(|s'|) \cdot \text{WOMAN}^\star(|s'|) \cdot \text{LOVE}^\star(|s'|)j)s']$
4. $= \lambda s.(\exists^\star \cdot 2^\star(|s|) \cdot \text{MAN}^\star(|s|) \cdot [\lambda s'.(\exists^\star \cdot 1^\star(|s'|) \cdot \text{WOMAN}^\star(|s'|) \cdot \text{LOVE}^\star(|s'|)(|s|))s'] \ )s$
5. $= \lambda s.(\exists^\star \cdot 2^\star(|s|) \cdot \text{MAN}^\star(|s|) \cdot \exists^\star) \cdot 1^\star(|s|{+}1) \cdot \text{WOMAN}^\star(|s|{+}1) \cdot \text{LOVE}^\star(|s|{+}1)(|s|)s$

[12] Or, to be more precise, it *is* a *family of functions*.

[13] Using these definitions we may come to a deduction theorem: $\models \neg(\varphi \cdot \neg(\psi)) \Leftrightarrow \varphi \models \psi$. The proof is left to the reader.

S	::=	S.S	X	::=	$X_1 \cdot X_2$
S	::=	NP VP	X	::=	$(X_1 X_2)$
NP	::=	Det CN	X	::=	$(X_1 X_2)$
VP	::=	IV	X	::=	X_1
VP	::=	TV NP	X	::=	$\lambda u.(X_2(\lambda v.(X_1 v)u))$
TV	::=	love	X	::=	$\text{LOVE}^\star$
CN	::=	man	X	::=	$\text{MAN}^\star$
CN	::=	woman	X	::=	$\text{WOMAN}^\star$
DET	::=	a	X	::=	$\lambda P.\lambda Q.\lambda s.(\exists^\star \cdot 1^\star(\|s\|) \cdot P(\|s\|) \cdot Q(\|s\|))s$
DET	::=	two	X	::=	$\lambda P.\lambda Q.\lambda s.(\exists^\star \cdot 2^\star(\|s\|) \cdot P(\|s\|) \cdot Q(\|s\|))s$

Figure 5.2: A simple plural fragment

□

We assume that, in the model, distributive predicates will be interpreted in the plural. So, for any P: $P(a) \wedge P(b) \rightarrow P(\{a, b\})$. Starting with an empty information state, after processing 'two men loved a woman' we produce output states containing two stacks $\{\langle m_1, w_1\rangle, \langle m_2, w_2\rangle\}$, such that: (i) $\{m_1, m_2\}$ is a plurality of two men, (ii) both w_1 and w_2 are women, (iii) $\{m_1, m_2\}$ is in the pluralised love relation with $\{w_1\}$ and $\{w_2\}$ and (iv) $\{w_1, w_2\}$ has cardinality 1. In other words: $w_1 = w_2$. As in van den Berg's system, without quantification we arrive at a collective reading.

It is easy to see that the above fragment will never result in undefinedness. The only entries that hand out indices to predicates are the determiners and they only refer to $|s|$ after extending s with one position. Intuitively, this will be the general case with determiners.

5.4 Distributivity and dependency revisited

Recall that in section 5.1 I discussed two related problems. One of them, the destructive assignment problem, was solved by adopting van Eijck's incremental dynamics framework. The second problem, the impossibility of expressing dependent anaphora in terms of co-indexation of syntactic material, will be dealt with in the current section. I claim that the contextualisation of variable/index choice solves this problem as well. To show this I turn back to the question of what the semantics of a distributivity operator should look like in order to account for the full range of anaphoric possibilities in distributive sentences including cases of dependency. Remember that there are two relevant observations. First, in the scope of a distributivity operator, reference to both the group the operator is applied to and to the respective atomic entities in that group is possible. So, (5.28) is ambiguous between a reading wherein Tom thinks that he is the best, Dick thinks that he is the best and Harry thinks that he is the best and one in which they each think that they as a group are the best. In order to account for this ambiguity we need to assume that the distributivity operator introduces an extra potential antecedent.

(5.28) [Tom, Dick and Harry]$_i$ each$_j$ think they$_{i/j}$ are the best.

The second observation regarding distributivity is that the distributivity operator does not only introduce the possibility of reference to the currently evaluated member of a group, but also to any (singular or plural) individual which (in discourse) has proved to be functionally dependent on that individual, as in one of the readings of the second sentence of (5.29). In order to account for this possibility we need to assume that distributivity operators evaluate their scope relative to specific parts of the current information state.

(5.29) [Tom, Dick and Harry]$_i$ each$_j$ have [two sons]$_k$ in the competition.
They$_i$ each$_l$ think they$_?$ are excellent sportsmen.

The problem was how to combine both observations, since there is no index available for the pronoun in the second sentence of (5.29). This is because the pronoun is able to express both the functionally dependent value of 'the two sons', and the six sons together. But in both these readings the antecedent is the same (i.e. in both cases the pronoun should be co-indexed k). The extra antecedents introduced by the distributivity operators are of no help either.

In DRT, the problem is solved by applying two principles of antecedent formation: (i) quantificational structures allow for set abstraction and (ii) quantification over a domain gives access to how this domain is created. Application of the second principle, however, changes nothing with respect to the accessibility of the original abstracted set itself. This is because DRT has the general property that referents introduced in some DRS α are always accessible in a subordinated DRS β. In DPlL, although, x in the scope of δ_x has a value, the local context no longer gives access to the original value of x. The same goes for any variable that is dependent on x. Just as in DRT, for a non-representational semantics we would like the independent value and the dependent values to be accessible at the same time. In DPlL, this was problematic because of the destructive assignment problem.

In IDPlL, however, the solution to this problem is straightforward now that we have freed ourself from the burden of having to come up with an indexation for the readings of (5.29). All we have to do is have the distributivity operator (temporarily) expand the information state with the dependent values. That is, the specific substate used in van den Berg's operator is not the actual context of interpretation but is added to (or, rather, in DRT terms, *subordinated* to) the incoming state (compare with footnote 4).

In a state s, distributivity with respect to index i makes accessible the substates like $s\restriction_{i=d}$, where d is an atomic part of $s[i]$. Such a state is of course a set of stacks itself. Assuming we have a way to combine such states with the input state into a new set of stacks $s^\sqcap s'$, we end up with a state wherein both functional and global values are present. Pronouns in the scope of the distributivity operator have now their full anaphoric potential. Ignoring definedness conditions for now, the operator is given as:

(5.30) $$s[\![\delta_i(\varphi)]\!]s' \;:\Leftrightarrow\; \begin{array}{l} s[i] = s'[i] \;\& \\ \forall x \in s[i] : \; s^\sqcap s\restriction_{i=x} [\![\varphi[i/i{+}|s|]]\!] s^\sqcap s'\restriction_{i=x} \end{array}$$

Let me clarify this definition. First of all, the states in pairs $\langle s, s'\rangle$ in the interpretation of $\delta_i(\varphi)$ differ at most in that the output context s' can contain more values than s due to potential existential introductions in φ. However, with respect to the indices in s, the states s and s' are equal in the values and dependencies they contain. For each atomic value assigned to i in the input state s, the scope φ is interpreted with respect to the input state s extended with a substate $s\restriction_{i=x}$. In the definition we need to make sure that occurences of i in φ are replaced by $i + |s|$, the slot for the respective values x.

The question now, of course, is how to define the mode of combination expressed by '$^\sqcap$'. The combination $s^\sqcap s'$ should contain exactly the dependencies that s and s' contained. That is, no dependencies between values of s and those in s' should arise. This means that every stack in s' should be appended to every stack in s:

Definition 5.18

Append for sets of stacks

$$s^\sqcap s' \;:=\; \{c^\wedge c' | c \in s \;\&\; c' \in s'\}$$

□

Notice a similarity to the DRT approach. The dependent values are only made accessible in the scope of the distributor, just as in DRT these values are only accessible in the duplex condition triggered by distributivity. An important difference, however, is that whereas in DRT the availability of such values is due to a stipulated principle of accessibility, in IDPlL, these values are produced by the distributor itself.

The functional version of the operator is a predicate modifier of type $\langle\langle\iota, T\rangle, \langle\iota, T\rangle\rangle$.[14] This enables us to explicitly show that the index of the

[14] I am indebted to Paul Dekker for pointing out a flaw in an earlier proposal, which attempted to define δ as a state-modifier. The problem that occurs with such a definition is that there is no form to perform the necessary substitution operation on. One may avoid substitution altogether by appending s onto $s\restriction_{i=x}$ instead of the other way around, thus ensuring that x is associated with i rather than with $i + |s|$. Unfortunately, this means that entities introduced in the scope of δ become separated from the output state and that consequently there is no natural way of handling state extensions in the scope of the distribution operator.

distributivity operator is taken care of by syntax. As a VP operator, the distributor is in no need of an index.

(5.31) $\delta^\star \;:=\; \lambda P.\lambda i.\lambda s.\lambda s'.\, s[i] = s'[i]\ \&$
$\forall x \in s[i] :\ s^\sqcap s'\restriction_{i=x} \in P(i+|s|)(s^\sqcap s\restriction_{i=x})$

The $\delta^\star$ function is subject to a definedness condition, which is closely related to that for predication: the function $\delta^\star(P)(i)(s)$ is defined if and only if $i < |s|$. It is time for an example. Let us first derive the semantics for our central example (5.32).

(5.32) Three students each wrote a paper.

Derivation 5.19

1. a $\rightsquigarrow$ $\lambda P.\lambda Q.\lambda s'.\,(\exists^\star \cdot 1^\star(|s'|) \cdot P(|s'|) \cdot Q(|s'|))\, s'$
2. paper $\rightsquigarrow$ $\text{PAPER}^\star$
3. a paper $\rightsquigarrow$ $\lambda Q.\lambda s'.(\exists^\star \cdot 1^\star(|s'|) \cdot \text{PAPER}^\star(|s'|) \cdot Q(|s'|))s'$
4. wrote $\rightsquigarrow$ $\text{WROTE}^\star$
5. wrote a paper $\rightsquigarrow$ $\lambda j.(\lambda Q.\lambda s'.(\exists^\star \cdot 1^\star(|s'|) \cdot \text{PAPER}^\star(|s'|))s')\ (\lambda i.(\text{WROTE}^\star i)j)$
   $= \lambda j.\lambda s'.(\exists^\star \cdot 1^\star(|s'|) \cdot \text{PAPER}^\star(|s'|) \cdot \text{WROTE}^\star(|s'|)j)s'$
6. each $\rightsquigarrow$ $\delta^\star$
7. each wrote a paper $\rightsquigarrow$ $\delta^\star(\lambda j.\lambda s'.(\exists^\star \cdot 1^\star(|s'|) \cdot \text{PAPER}^\star(|s'|) \cdot \text{WROTE}^\star(|s'|)j)s')$
8. three $\rightsquigarrow$ $\lambda P.\lambda Q.\lambda s.\,(\exists^\star \cdot 3^\star(|s|) \cdot P(|s|) \cdot Q(|s|))\, s$
9. students $\rightsquigarrow$ $\text{STUDENT}^\star$
10. three students $\rightsquigarrow$ $\lambda Q.\lambda s.(\exists^\star \cdot 3^\star(|s|) \cdot \text{STUDENT}^\star(|s|) \cdot Q(|s|))s$
11. (5.32) $\rightsquigarrow$ $\lambda s.(\exists^\star \cdot 3^\star(|s|) \cdot \text{STUDENT}^\star(|s|)\cdot$
    $[\delta^\star(\lambda j.\lambda s'.(\exists^\star \cdot 1^\star(|s'|) \cdot \text{PAPER}^\star(|s'|) \cdot \text{WROTE}^\star(|s'|)j)s')](|s|))s$

□

The resulting function in 11. takes a state and performs a series of state transitions on it, including the distribution with respect to index $|s|$ over the $\langle\iota, T\rangle$-function corresponding to the VP.

Let us apply the resulting state transition in 11. to an empty state $\{\langle\rangle\}$. Consider the verb-phrase $\langle\iota, T\rangle$ function in (5.33), call it V.

(5.33) $V = \lambda j.\lambda s'.(\exists^\star \cdot 1^\star(|s'|) \cdot \text{PAPER}^\star(|s'|) \cdot \text{WROTE}^\star(|s'|)j)s'$

Say that b_1 and b_2 and b_3 are three students, then one possible input for the distributivity $\delta^\star(V)(0)$ is the set of three one-place stacks containing exactly these three individuals: $\{\langle b_1\rangle, \langle b_2\rangle, \langle b_3\rangle\}$. Say furthermore that b_1 wrote the paper p_1, b_2 wrote p_2 and b_3 wrote p_3 and that no other papers were written by these three students. Then:

(5.34) $\delta^\star(V)(0)\begin{array}{c} 0 \\ \hline \multicolumn{1}{|c|}{b_1} \\ \hline \multicolumn{1}{|c|}{b_2} \\ \hline \multicolumn{1}{|c|}{b_3} \\ \hline \end{array} = \left\{ \begin{array}{cc} 0 & 1 \\ \hline \multicolumn{1}{|c|}{b_1} & \multicolumn{1}{c|}{p_1} \\ \hline \multicolumn{1}{|c|}{b_2} & \multicolumn{1}{c|}{p_2} \\ \hline \multicolumn{1}{|c|}{b_3} & \multicolumn{1}{c|}{p_3} \\ \hline \end{array} \right\}$

The scope is processed three times each time with one of the atomic students pushed to the top of the context. So, applying V, in (5.33) to index 1 and the three extended states yields the following results:

(5.35) $$V(1)\begin{array}{|c|c|} \multicolumn{1}{c}{0} & \multicolumn{1}{c}{1} \\ \hline b_1 & b_1 \\ \hline b_2 & b_1 \\ \hline b_3 & b_1 \\ \hline \end{array} = \left\{ \begin{array}{|c|c|c|} \multicolumn{1}{c}{0} & \multicolumn{1}{c}{1} & \multicolumn{1}{c}{2} \\ \hline b_1 & b_1 & p_1 \\ \hline b_2 & b_1 & p_1 \\ \hline b_3 & b_1 & p_1 \\ \hline \end{array} \right\}$$

(5.36) $$V(1)\begin{array}{|c|c|} \multicolumn{1}{c}{0} & \multicolumn{1}{c}{1} \\ \hline b_1 & b_2 \\ \hline b_2 & b_2 \\ \hline b_3 & b_2 \\ \hline \end{array} = \left\{ \begin{array}{|c|c|c|} \multicolumn{1}{c}{0} & \multicolumn{1}{c}{1} & \multicolumn{1}{c}{2} \\ \hline b_1 & b_2 & p_2 \\ \hline b_2 & b_2 & p_2 \\ \hline b_3 & b_2 & p_2 \\ \hline \end{array} \right\}$$

(5.37) $$V(1)\begin{array}{|c|c|} \multicolumn{1}{c}{0} & \multicolumn{1}{c}{1} \\ \hline b_1 & b_3 \\ \hline b_2 & b_3 \\ \hline b_3 & b_3 \\ \hline \end{array} = \left\{ \begin{array}{|c|c|c|} \multicolumn{1}{c}{0} & \multicolumn{1}{c}{1} & \multicolumn{1}{c}{2} \\ \hline b_1 & b_3 & p_3 \\ \hline b_2 & b_3 & p_3 \\ \hline b_3 & b_3 & p_3 \\ \hline \end{array} \right\}$$

Take, for instance, the output set of stacks in (5.37). According to the definition of $\delta^\star$, this state is $s^\frown s'\restriction_{0=b_3}$. That is, $s'\restriction_{0=b_3}$ equals a singleton set of one stack assigning with b_3 on position 0 and p_3 on position 1: i.e., $\{\langle b_3, p_3\rangle\}$. Similarly, the output states in (5.35) and (5.36) tell us what s' looks like with respect to the value b_1 on 0 and b_2 on 0 respectively.

Crucially, the equations in (5.35) to (5.37) tell us something about the accessibility features of the $\delta^\star$-operation. They show that the scope is interpreted with respect to a state which contains both the original set over which the distribution takes place (i.e. the set in index 0) and the respective atoms (i.e. the atomic individual in 1). This shows that, in contrast to the distributivity operator of DPlL, pronouns have access to both the set-antecedent and to the respective atoms of that set. Outside the scope of distributivity, only sets and, potentially, dependencies remain.

Nevertheless, it is important to be more specific about pronominal interpretation and resolution. So far, we tacitly assumed some kind of underspecification of pronominal reference at the semantic level. We have remained silent about how to integrate such an approach to pronominal interpretation in the incremental dynamic plural logic. In the coming section, we will finally turn to how pronouns interact with the semantics.

5.5 Pronouns and accessibility

We observed before that in order to be able to resolve discourse pronouns in such a way that both dependent readings and independent readings show up, the pronoun needs to take its antecedent from the semantic level. This means that not a single interpretation but a set containing all possible (pronominally) disambiguated interpretations should be derived. In other words, a sentence needs to be (pronominally) disambiguated at the semantic rather than at the syntactic level.

I will discuss two alternative ways of generating possible resolutions of pronoun reference at the semantic level. The first is a simple Montagovian approach, the second uses an underspecified interpretation for pronouns. In both cases, all and only the reported cases of anaphora are generated.

5.5.1 A Montagovian interpretation for pronouns

The Montague-style approach to pronoun interpretation boils down to an ambiguity in indexation. A pronoun 'they' is (in principle infinitely) ambiguous. Its reference may correspond to that associated with any index.

In the current IDPlL setup, this means that the lexical entry for 'they' is really of list of entries of the form 'they $\rightsquigarrow$ $\lambda P.P(i)$' where i is in $\mathbb{N}$. Different entries lead to different resolutions of the pronoun and many choices will not lead to a defined interpretation at all since not all indices are represented in the context.

By adopting a Montagovian interpretation for pronouns, we can show how the above theory of distributivity within IDPlL accounts for the kinds of anaphora which triggered the move to incremental dynamics.

Derivation 5.20

1. they $\rightsquigarrow$ $\lambda P \lambda s.P(i)(s)$ where $i \in \mathbb{N}$
2. each $\rightsquigarrow$ $\delta^\star$
3. sent to L&P $\rightsquigarrow$ $\text{SENT\_TO\_L\&P}^\star$
4. it $\rightsquigarrow$ $\lambda P \lambda s.(1^\star(j) \cdot P(j))(s)$ where $j \in \mathbb{N}$
5. sent it to L&P $\rightsquigarrow$ $\lambda u((\lambda P.\lambda s.(1^\star(j) \cdot P(j))(s))(\lambda v.(\text{SENT\_TO\_L\&P}^\star v)u))$
   $= \lambda u.\lambda s.(1^\star(j) \cdot \text{SENT\_TO\_L\&P}^\star(u,j))(s)$
6. each sent it to L&P $\rightsquigarrow$ $\delta^\star(\lambda u.\lambda s.(1^\star(j) \cdot \text{SENT\_TO\_L\&P}^\star(u,j))s)$
   $= \delta^\star(\lambda u.1^\star(j) \cdot \text{SENT\_TO\_L\&P}^\star(u,j))$
7. they each sent it to L&P
   $\rightsquigarrow$ $\lambda P \lambda s.P(i)(s)\ (\delta^\star(\lambda u.1^\star(j) \cdot \text{SENT\_TO\_L\&P}^\star(u,j)))$
   $= \delta^\star(\lambda u.1^\star(j) \cdot \text{SENT\_TO\_L\&P}^\star(u,j))(i)$

□

This is not really a single, but rather an infinity of derivations. 'They sent it to L&P' is interpreted as the set of state transitions in (5.38).

(5.38) $\{\delta^\star(\lambda u.1^\star(j) \cdot \text{SENT\_TO\_L\&P}^\star(u,j))(i) \mid i,j \in \mathbb{N}\}$

The functions in this set are only defined on states that are of a sufficient size with respect to the choices made for i and j. Take, for example, the state that resulted from our example (5.32) and apply individual functions in (5.38) to it.

(5.39) $\delta^\star(\lambda u.1^\star(j) \cdot \text{SENT\_TO\_L\&P}^\star(u,j))(i)$

0	1
b_1	p_1
b_2	p_2
b_3	p_3

If defined, this will result into a set of output states. It is easy to see that the term above is only defined if $i \in \{0,1\}$ and $j \in \{0,1,2,3\}$, since outside the scope of the distributor the size of the context is two and inside the scope it is doubled to four. Among the eight possible resolutions, there are two intelligible ones. (In the others, students are sent by papers, etc.) If we choose 0 for i and 1 for j, then we get the reading where the three papers were sent to the L&P editor three times. Choosing 0 for i and 3 for j returns the dependent reading where the individual students sent the papers they wrote.

The above proposal, however, is counter some motivations behind ID-PlL. By making indexation context dependent, we expect that, in principle, there should be a way of deriving defined interpretations only. The infinite ambiguity of the pronoun goes against the intuition that in ID-PlL, the referential options of a pronoun that lead to definedness should be recoverable from context. I will therefore now turn to an approach to pronominal interpretation which assumes a pronoun can only be resolved to antecedents that are (in some way still to be defined) accessible. This will come at the price of having to introduce some level of representation.

5.5.2 Underspecification

So far, we have insisted that the interpretation of an expression in no way needs to be mediated by a level of representation. In the previous section, expressions were mapped to lambda terms which ultimately reduce into state transitions: functions from states to sets of states. So, when we said that 'each' corresponded to '$\delta^\star$' we expressed that 'each' translates to a function of type $\langle\langle\iota,T\rangle,\langle\iota,T\rangle\rangle$ abbreviated by $\delta^\star$. In principle, it was unnecessary to map 'each' to an intermediate representation which interpreted to the function in question. The question is whether such an approach is still tenable once we try to incorporate the underspecified interpretation of pronouns in our fragment.

In order to devise an underspecification-alternative to the Montagovian approach discussed above, a derivation results in an underspecified representation. These representations correspond to sets of meanings. So, one could argue that in fact these representations are dispensable, and that instead of deriving a function which maps states to representations, a function mapping a state to a set of possible output states is derived. For example, say that our underspecified semantics maps a VP like 'sent it to L&P' to a function taking an index i and a state s and returning a set of possible disambiguated VP interpretations. That is, combining the

VP with a subject like 'John', it could result in a set $\{\{s\}, \{\emptyset\}\}$. Here the singleton set containing s is due to all the resolutions of 'it' such that they were sent to L&P by John and $\{\emptyset\}$ is due to all the resolutions to objects that were not sent to L&P by John.

An immediate problem becomes apparent. There is quite some psycholinguistic evidence that pronouns are not resolved immediately when they are encountered. Corbett and Chang (1983), for instance, show that if a pronoun has more than one (suitable) antecedents, it remains unresolved until the sentence is completed. At the sentence level, however, the underspecified interpretation does not contain any information about which anaphoric links are potential resolutions (nor into which interpretation they result). There is just the choice between an absurd and a non-absurd output context. One could argue that this is a good result, since it correctly represents the options that are available in terms of the semantic effect of these options. (Additionally, one could argue that the resolution mechanism operates on more than just semantic information and we should consequently not be worried about not representing each resolution in context separately.) Examples containing a quantified pronoun, however, show that using some level of representation is inevitable. Say, we give the VP 'sent it to L&P' the underspecified interpretation as we did above and say we combine it with a quantifier like 'every student'. What happens now is that we are not collecting resolutions of 'every students sending x to L&P', but rather, we are quantifying over resolutions. That is, 'every student sent it to L&P' could be predicted to have a resolution verifying the sentence even if no single antecedent was sent by all the students. (For instance, the sentence could be verified by resolving 'it' to a salient set of papers for one half of the students and to a salient set of letters for the others, in a situation in which half the students sent the paper they wrote to L&P, while the other half sent the letter they wrote.)

In sum, in order to deal with the underspecification of pronominal interpretation we need a level of representation which enables us to represent open sentences. We now turn to the definition of such a representation language. We will simply use the abbreviations of functions we introduced above for this representational level. But now, instead of being identified with functions, these symbols map to functions via an interpretation function.

The representation language is typed. We have three basic types: E, T and $\mathtt{i}$. If α and β are types, then so is $\langle \alpha, \beta \rangle$. Assume the following variables:

- a set $V_{\mathsf{E}} = \{s, s', \ldots\}$ of variables of type E
- a set $V_{\mathtt{i}} = \{i, i', \ldots, j, j', \ldots, u, u', \ldots\}$ of variables of type $\mathtt{i}$
- a set $P_{\mathtt{i}\mathsf{T}} = \{P, P', \ldots, R, R', \ldots\}$ of variables of type $\langle \mathtt{i}, \mathsf{T} \rangle$

The set V is the combine of these sets $V_{\mathtt{E}} \cup V_{\mathtt{i}} \cup V_{\mathtt{iT}}$. Assume the following constants:

- a set $C_{\mathtt{i\_T}} = \{\text{MAN}, \text{WOMAN}, \ldots, 1, 2, \ldots, \text{SENT\_TO\_L\&P}, \ldots\}$ of constants of type $\langle \mathtt{i}, \mathsf{T}\rangle$, or $\langle \mathtt{i}, \langle \mathtt{i}, \mathsf{T}\rangle\rangle$, or $\langle \mathtt{i}, \langle \mathtt{i}, \langle \mathtt{i}, \mathsf{T}\rangle\rangle\rangle$ etc.
- a set $C_{\mathtt{i}} = \mathbb{N}$ of constants of type $\mathtt{i}$
- a set $C'_{\mathtt{i}} = \{\mathsf{x}_n | n \in \mathbb{N}\}$ also of constants of type $\mathtt{i}$

The intuition behind these definitions is that type E corresponds to states, type T to state transitions and type $\mathtt{i}$ to indices.[15] Moreover, the two sets of constants of type $\mathtt{i}$ correspond to a set of ordinary indices, $C_{\mathtt{i}}$, and a set $C'_{\mathtt{i}}$ containing the pronominal place holders 'x_i'.

The set $C_{\mathtt{i\_T}}$ is the set of predicate symbols of whatever arity and includes symbols for cardinality predicates as well as symbols for ordinary predicates in the model. A function Ar is defined on the elements in the set $C_{\mathtt{i\_T}}$ and returns the corresponding arity. (So, for instance, $Ar(2) = 1$ and $Ar(\text{SENT\_TO\_L\&P}) = 2$.) Now we define the language L as follows. We abbreviate '$\alpha \in L$ & $\alpha :: x$' as '$\alpha \in L :: x$'.

Definition 5.21

The language L

$$\begin{array}{rcl}
v \in L :: \mathtt{i} & \Leftrightarrow & v \in V_{\mathtt{i}} \cup C_{\mathtt{i}} \cup C'_{\mathtt{i}} \\
\exists \in L :: \mathsf{T} & & \\
P(v_n), \ldots, (v_1) \in L :: \mathsf{T} & \Leftrightarrow & Ar(P) = n \;\&\; v_1, \ldots, v_n \in L :: \mathtt{i} \;\& \\
 & & P :: \langle \mathtt{i}, \ldots \langle \mathtt{i}, \mathsf{T}\rangle\rangle \\
\neg(\varphi) \in L :: \mathsf{T} & \Leftrightarrow & \varphi \in L :: \mathsf{T} \\
\varphi \bullet \psi \in L :: \mathsf{T} & \Leftrightarrow & \varphi \in L :: \mathsf{T} \;\&\; \psi \in L :: \mathsf{T} \\
\delta(P)(v) \in L :: \mathsf{T} & \Leftrightarrow & P \in L : \langle \mathtt{i}, \mathsf{T}\rangle \;\&\; v \in L :: \mathtt{i} \\
\lambda a.\psi \in L :: \langle x, y\rangle & \Leftrightarrow & a \in V \;\&\; a :: x \;\&\; \psi \in L :: y \\
\varphi(\psi) \in L :: y & \Leftrightarrow & \varphi \in L :: \langle x, y\rangle \;\&\; \psi \in L :: x
\end{array}$$

We do not use standard β-reduction of lambda terms. The reason is that while combining two terms, we need to make sure that the pronominal constants x_i are renamed in order to avoid clashes (which would eventually lead to unwanted co-reference between pronouns). We do this by having pronouns always introduce the constant x_0 and by replacing such a constant with x_{1+n} when the form containing it is applied to a term containing place holders with n as their maximal index.

[15] In fact, the interpretation of objects of type E, T and $\mathtt{i}$ will be states, transitions and indices respectively. It might seem that these three types are odd choice for *basic* types, since, obviously, T could have been defined in terms of E and a type for truth-values and E in turn in terms of ι and a type for entities. However, this choice is meant to reflect the purpose of the language. Since nothing in the language expresses a truth-value or an entity, there are no such types.

Definition 5.22

β-reduction

$$\lambda\gamma.\varphi(\psi) \;=_\beta\; \varphi[\gamma/\psi[\mathsf{x}_i/\mathsf{x}_{i+m(\varphi)}]_{\mathsf{x}_i\in\pi(\psi)}]$$

□

Definition 5.23

The set π of place holders in a term in L

$$\begin{array}{rcll}
\pi(\exists) & = & \emptyset & \\
\pi(P) & = & \emptyset & \\
\pi(P(v_n)\ldots(v_2)(\mathsf{x}_i)) & = & \{\mathsf{x}_i\}\cup\pi(P(v_n)\ldots(v_2)) & \\
\pi(P(v_n)\ldots(v_2)(v_1)) & = & \pi(P(v_n)\ldots(v_2)) & \Leftrightarrow\; v_1\in C_{\mathtt{i}} \\
\pi(\neg(\varphi)) & = & \pi(\varphi) & \\
\pi(\delta(P)(\mathsf{x}_i)) & = & \{\mathsf{x}_i\}\cup\pi(P) & \\
\pi(\delta(P)(v)) & = & \pi(P) & \Leftrightarrow\; v\in C_{\mathtt{i}} \\
\pi(\varphi\bullet\psi) & = & \pi(\varphi)\cup\pi(\psi) & \\
\pi(\lambda\gamma.\varphi) & = & \pi(\varphi) &
\end{array}$$

□

Definition 5.24

The minimum index not associated with one of the place holders in φ in L

$$m(\varphi) = \begin{cases} 0 & \pi(\varphi)=\emptyset \\ \max_i(\mathsf{x}_i\in\pi(\varphi))+1 & \text{otherwise} \end{cases}$$

□

Finally, we define the language L' which is a proper subset of L and which is generated using the set $C_{\mathtt{i}}$ as the only set of constants of type $\mathtt{i}$.

The language L is the language of underspecified representations. A form $P(\mathsf{x}_i)$ expresses that the argument of P needs to be resolved. The language L' is the target-language for resolution. Expressions in L' will eventually be interpreted into state transitions.

Now we turn to the resolution process. Formulae of type T in L map to sets of formulae of L' of the same type by replacing the dummy argument fillers 'x_i' with an index from $C_{\mathtt{i}}$.

Definition 5.25

Let $\varphi\in L :: \mathsf{T}$ and $\varphi'\in L' :: \mathsf{T}$.

$$\varphi\trianglerighteq\varphi' \;\Leftrightarrow\; \exists X\subseteq C_{\mathtt{i}} : |X|\leq|\pi(\varphi)|\;\&\;\exists f : X^{\pi(\varphi)} : \varphi'=\varphi[\mathsf{x}_i/f(\mathsf{x}_i)]_{\mathsf{x}_i\in\pi(\varphi)}$$

□

This says that there is a set of indices X such that there is a function f mapping the place holders in φ to the indices in X. This function, as it were, represents the key to the resolution of φ: it maps the pronoun occurrences to their antecedent. The form φ' results from substituting each occurrence of x_i in φ with $f(\mathsf{x}_i)$.

Of course, many of such disambiguated forms that result from a specific choice of f will not be *proper*, in the sense that they will contain unbound indices. In IDPlL, predication over an index which exceeds the size of the context resulted in undefinedness. Consequently, we can use the potential undefinedness of IDPlL formulae to establish whether a resolution is proper or not. The semantics of L', it will be no surprise, is therefore largely identical to IDPlL.

Definition 5.26

Semantics for L'

$$\begin{array}{rcll}
[\exists] & = & \exists^\star & \\
[P(v_n)\ldots(v_1)] & = & \lambda s.(P^\star(|s|\text{-}v_1,\ldots,|s|\text{-}v_n))s & \text{where } P(v_n)\ldots(v_1) \in L' :: \mathsf{T} \\
[\neg\varphi] & = & \neg^\star([\varphi]) & \text{where } \varphi \in L' :: \mathsf{T} \\
[\delta(P)(v)] & = & \delta^\star([P])(|s|\text{-}v) & \text{where } \varphi \in L' :: \mathsf{T} \\
[\varphi \bullet \psi] & = & [\varphi]\cdot[\psi] & \text{where } \varphi \in L' :: \mathsf{T}\ \&\ \psi \in L' :: \mathsf{T} \\
[\varphi(\psi)] & = & [\varphi]([\psi]) & \text{where } \varphi(\psi) \in L \\
[\lambda v.\varphi] & = & \lambda i.\lambda s.[\varphi(i+|s|)](s) & \text{where } v \in L' :: \mathtt{i}
\end{array}$$

□

The most important point here is that forms in L' use an indexation method which is reminiscent of the *de Bruijn notation* for (variable free) lambda calculus (de Bruijn 1972). This is necessary, since in the forms in L' we do not have information about the size of the context. That is, given a variable s of type E, '$|s|$' does not mean anything; it certainly does not correspond to an index. In the de Bruijn-style notation an argument i will be interpreted as the i-th position to the left of the top of the stacks. So, a determiner like 'a' takes two forms P and R of type $\mathtt{i}\mathsf{T}$ to results in a form '$\exists \bullet P(0) \bullet R(0)$' of type T, expressing that the predicates take the newly introduced slot as their argument. In general then, at any point, 0 corresponds to the top of the stack, 1 to the position to the left of that, etc.

Given this semantics, we can check whether an unambiguated form φ in L' is proper with respect to some state s.

Definition 5.27

Properness: Let s be a set of stacks and $\varphi \in L'$:

$$\varphi \text{ is proper w.r.t. } s \text{ iff } [\varphi]s \neq\uparrow$$

S	::=	NP VP	X	::=	$(X_1 X_2)$
NP	::=	Det CN	X	::=	$(X_1 X_2)$
VP	::=	IV	X	::=	X_1
VP	::=	TV NP	X	::=	$\lambda u.(X_2(\lambda v.(X_1 v)u))$
VP	::=	MOD VP	X	::=	$(X_1 X_2)$
TV	::=	sent to L&P	X	::=	SENT_TO_L&P
NP	::=	they	X	::=	$\lambda P.(P(\mathsf{x}_0))$
NP	::=	it	X	::=	$\lambda P.(1(\mathsf{x}_0) \bullet P(\mathsf{x}_0))$
CN	::=	man	X	::=	MAN
CN	::=	woman	X	::=	WOMAN
DET	::=	a	X	::=	$\lambda P.\lambda Q.(\exists \bullet 1(0) \bullet P(0) \bullet Q(0))$
DET	::=	three	X	::=	$\lambda P.\lambda Q.(\exists \bullet 3(0) \bullet P(0) \bullet Q(0))$
MOD	::=	each	X	::=	$\lambda P.\lambda i.(\delta(P)(i))$

Figure 5.3: Fragment for deriving underspecified forms

□

A resolution of a formula φ in L of type T, given some state s, is now the collection of formulae ψ in L' of type T such that $\varphi \trianglerighteq \psi$ and ψ is proper in s.

Definition 5.28

Resolution

$$\texttt{resolve}_s(\varphi) = \{\psi | \varphi \trianglerighteq \psi \ \& \ \psi \text{ is proper in } s\}$$

□

Here is an example. Consider the fragment in figure 5.3. Note the following minute differences with the fragment in figure 5.2. First of all, instead of deriving state transitions, on sentence level a contextualised formula is derived: a function from states to underspecified representations. Second, the fragment only deals with sentence interpretation. This is because at some point resolution will have to be applied in order to have the resolved and subsequently interpreted representation update the context.

Let us turn to our running example, (5.40).

(5.40) They each sent it to L&P.

Derivation 5.29

1. they $\rightsquigarrow$ $\lambda P.(P(\mathsf{x}_0))$
2. each $\rightsquigarrow$ $\lambda P.\lambda i.(\delta(P)(i))$
3. sent to L&P $\rightsquigarrow$ SENT_TO_L&P
4. it $\rightsquigarrow$ $\lambda P.(1(\mathsf{x}_0) \bullet P(\mathsf{x}_0))$

5. sent it to L&P $\rightsquigarrow$ $\lambda u.(\lambda P.(1(\mathsf{x}_0) \bullet P(\mathsf{x}_0))(\lambda v.(\text{SENT\_TO\_L\&P } v)u))$
   $=_\beta$ $\lambda u.(1(\mathsf{x}_0) \bullet \text{SENT\_TO\_L\&P}\mathsf{x}_0)(u))$
6. each sent it to L&P
   $\rightsquigarrow$ $\lambda P.\lambda i.(\delta(P)(i))$ $(\lambda u.(1(\mathsf{x}_0) \bullet \text{SENT\_TO\_L\&P}(\mathsf{x}_0)(\mathsf{x}_2 u)))$
   $=_\beta$ $\lambda i.\delta(\lambda j.1(\mathsf{x}_0) \bullet \text{SENT\_TO\_L\&P}(\mathsf{x}_0)(j))(i)$
7. they each sent it to L&P
   $\rightsquigarrow$ $\lambda P.(P(\mathsf{x}_0))$ $(\lambda i.(\delta(\lambda j.1(\mathsf{x}_0) \bullet \text{SENT\_TO\_L\&P}(\mathsf{x}_0)(j))(i)))$
   $=_\beta$ $(\lambda i.(\delta(\lambda j.1(\mathsf{x}_1) \bullet \text{SENT\_TO\_L\&P}(\mathsf{x}_1)(j))(i))(\mathsf{x}_0))$
   $=_\beta$ $\delta(\lambda j.1(\mathsf{x}_1) \bullet \text{SENT\_TO\_L\&P}(\mathsf{x}_1)(j))(\mathsf{x}_2)$

□

As becomes clear from this derivation, the substitution of place holders during beta reduction is sometimes redundant (as in the final step). Still, better safe than sorry. Moreover, this redundancy seems harmless.

The resulting formula, an underspecified representation of the interpretation of (5.40), needs to be resolved in order to be interpreted. That is, the form can be disambiguated by mapping it to a set of forms in L'.

(5.41) $\texttt{resolve}_s(\delta(\lambda j.1(\mathsf{x}_1) \bullet \text{SENT\_TO\_L\&P}(\mathsf{x}_1)(j))(\mathsf{x}_2))$

This returns a set of representations of the following form:

(5.42) $\delta(\lambda k.1(j) \bullet \text{SENT\_TO\_L\&P}(j)(k))(i)$

For a state s, i will be restricted to values not exceeding $|s| - 1$, while j will not exceed $(|s| - 1) \times 2$. These values have been established by checking whether they result in an interpretable formula with respect to s. As in the Montagovian approach definedness is used as a criterion for accessibility.

5.5.3 Pronouns and Entailment

What can we say about entailment patterns involving underspecified representations? Following Beaver (1999), we will define the notion of *possible entailment*. The idea is that φ possibly entails ψ, notation: '$\varphi|\!\rightsquigarrow \psi$', if and only if some resolution of φ entails some resolution of ψ.

Definition 5.30

Possible entailment. Let $\varphi, \psi \in L$:

$$\varphi|\!\rightsquigarrow \psi$$
$$\Leftrightarrow$$
$$\forall s : (\exists \gamma \in \texttt{resolve}_s(\varphi) \;\&\; \exists s' \in [\gamma]s \;\rightarrow\; \exists \gamma' \in \texttt{resolve}_{s'}(\psi) : \gamma \models_{L'} \gamma')$$

□

The intuition behind this definition is the following. A form φ possibly entails a form ψ if for each state s if there are any resolutions, then at least one of the resolutions of φ entails a resolution of ψ. Here, $\models_{L'}$ is the notion of entailment defined for L', which is based on entailment as defined on type T state transitions.

Definition 5.31

Entailment. Let $\varphi, \psi \in L'$:

$$\varphi \models_{L'} \psi \Leftrightarrow [\varphi] \models [\psi]$$

□

Also, for forms in L we have a notion of necessary entailment between two forms which says that no matter what resolutions we choose they will entail each other (at the level of L').

Definition 5.32

Necessary entailment. Let $\varphi, \psi \in L$:

$$\varphi \models_L \psi$$
$$\Leftrightarrow$$
$$\forall s : \forall \gamma \in \texttt{resolve}_s(\varphi) : \forall s' \in [\gamma]s : \forall \gamma' \in \texttt{resolve}_{s'}(\psi) : \gamma \models_{L'} \gamma'$$

□

Of course, $\varphi \models_L \psi$ always equals $\varphi \models_{L'} \psi$ in case φ and ψ contain no pronominal dummies, since in that case there is only one resolution for φ, namely φ itself, while, similarly, ψ is the only resolution for ψ.

The rather straightforward approach to reasoning with open sentences has adequate results. For instance,

(5.43) They are mad men $\mid\!\leadsto$ They are men

For each context in which a resolution for 'they are mad men' results in truth, there will be a resolution for which 'they are men' results in truth (namely, the *same* resolution). But, of course,

(5.44) They are mad men $\not\models_L$ They are men

since 'They are men' could be construed as referring to a different set of entities as the pronoun in front of the entailment sign.

Possible entailment is not transitive, just as the consequence relation in dynamic predicate logic is not transitive, as can be shown using the example in (5.45) from van Benthem 1987.

(5.45)
If a man owns a house, he owns a garden.
If a man owns a garden, he sprinkles it.
———
?⇒ If a man owns a house, he sprinkles it.

In van Eijck 2001, the transitivity of consequence is discussed in detail. Van Eijck is of the opinion that the intransitivity of $\models$ in DPL and DRT is *logically* speaking unintuitive. From $\varphi \models \psi$ and $\psi \models \gamma$ it should follow that $\varphi \models \gamma$. One might argue, however, that *linguistically* speaking transitivity does not hold for consequence, as is illustrated by (5.45).

Clearly, our linguistic intuition tells us that the only possible resolution for the consequence sentence boils down to men sprinkling their house, if they own one. Van Eijck's notion of consequence, however, is not totally at odds with this linguistic intuition. What he derives is namely a form $\exists \cdot \text{MAN}(0) \cdot \exists \cdot \text{HOUSE}(0) \cdot \text{OWN}(1,0) \models \text{SPRINKLE}(2,0)$. The index 2 is the result of the premise that owning a house leads to owning a garden. This garden is supposed to take the top position in the stack, so that the man in the context is now 2 positions removed from the top of the stack. But in the consequence, the index 2 does not correspond to anything anymore. Van Eijck proposes to deal with formulae like this by means of *existential padding*. Since the context only provides two indices for $\text{SPRINKLE}(2,0)$, namely the man and the house, an extra existential quantifier is applied to increment the context and provide a value for slot 2. This way, the third line in (5.45) reads (in its incremental dynamic interpretation): if a man owns a house, there is something he sprinkles.

In our current underspecification approach, possible entailment is not transitive, due to the fact that no antecedent for 'it' other than the house is accessible in the third line of (5.45). A form like $\text{SPRINKLE}(2,0)$ can not be generated by resolution, since it would result in undefinedness. Possible entailment, i.e. entailment *before* reference resolution, therefore follows the linguistic intuition that the reasoning in (5.45) is false simply because there is no resolution of the pronouns that makes it true.

If we focus instead on necessary entailment or on the entailment notion defined for disambiguated forms in L', we seem to get a similar result. According to this:

$$\begin{array}{rl}
\exists \cdot \text{MAN}(0) \cdot \exists \cdot \text{HOUSE}(0) \cdot \text{OWN}(0,1) \models_{L'} & \exists \cdot \text{GARDEN}(0) \cdot \text{OWN}(2,0) \\
\exists \cdot \text{GARDEN}(0) \cdot \text{OWN}(2,0) \models_{L'} & \text{SPRINKLE}(2,0) \\
\hline
\exists \cdot \text{MAN}(0) \cdot \exists \cdot \text{HOUSE}(0) \cdot \text{OWN}(1,0) \not\models_{L'} & \text{SPRINKLE}(2,0)
\end{array}$$

simply because

$$\exists \cdot \text{MAN}(0) \cdot \exists \cdot \text{HOUSE}(0) \cdot \text{OWN}(0,1) \cdot \text{SPRINKLE}(0,2)$$

is undefined in every context. But if we follow van Eijck and apply existential padding to avoid the undefinedness, then we do derive the conclusion:

$$\exists \cdot \text{MAN}(0) \cdot \exists \cdot \text{HOUSE}(0) \cdot \text{OWN}(1,0) \models_{L'} \exists \cdot \text{SPRINKLE}(2,0)$$

In sum, the level of underspecified representation given by L with its notion of possible entailment and the level of disambiguated forms given by L' allow for a reconciliation of linguistic and logical intuitions concerning the consequence relation. In a language without pronouns (and consequently without pronominal ambiguity) entailment is transitive conform logical intuition. In the underspecification language, however, transitivity does not hold for consequence, which does justice to the linguistic intuitions concerning (5.45).

$$
\begin{array}{rcl}
G\,[\![\epsilon_x]\!]\,^d H & \Leftrightarrow & G(x)=\emptyset \,\&\, \exists X : H = \{g[x/d] | g \in G \,\&\, d \in X\} \\
G\,[\![\epsilon_x]\!]\,^+ H & \Leftrightarrow & G\,[\![\epsilon_x]\!]\,^d H \\
G\,[\![\epsilon_x]\!]\,^- H & \Leftrightarrow & \bot \\
G\,[\![Px_1,\ldots,x_n]\!]\,^d H & \Leftrightarrow & G = H \\
G\,[\![Px_1,\ldots,x_n]\!]\,^+ H & \Leftrightarrow & G = H \& \,\langle G(x_1),\ldots,G(x_n)\rangle \in I(P) \\
G\,[\![Px_1,\ldots,x_n]\!]\,^- H & \Leftrightarrow & G\,[\![Px_1,\ldots,x_n]\!]\,^d H \,\&\, \neg G\,[\![Px_1,\ldots,x_n]\!]\,^+ H \\
G\,[\![\delta_x(\varphi)]\!]\,^d H & \Leftrightarrow & G(x) = H(x) \,\&\, G|_{x=\star} = H|_{x=\star} \,\& \\
 & & \forall d \in G(x) : G|_{x=d}\,[\![\varphi]\!]\,^d H|_{x=d} \\
G\,[\![\delta_x(\varphi)]\!]\,^+ H & \Leftrightarrow & G\,[\![\delta_x(\varphi)]\!]\,^d H \,\&\, \forall d \in G(x) : G|_{x=d}\,[\![\varphi]\!]\,^+ H|_{x=d} \\
G\,[\![\delta_x(\varphi)]\!]\,^- H & \Leftrightarrow & G\,[\![\delta_x(\varphi)]\!]\,^d H \,\&\, \neg G\,[\![\delta_x(\varphi)]\!]\,^+ H
\end{array}
$$

Figure 5.4: Dynamic Plural Logic, van den Berg 1996

5.6 Conclusion

In this chapter, we focused on the representation of dependent reference of discourse anaphora. We saw that in order to account for all possible kinds of anaphora, it was necessary to assume that pronoun resolution takes place on the semantic level. Moreover, in order to avoid variable clashes, we adopted the incremental dynamics framework and presented a variation on van den Berg's distributivity operator within an incremental dynamic plural logic. We showed that this formalism allows us to predict the reported accessibility patterns, which we illustrated using two possible tactics of enriching the formalism with pronoun resolution.

5.A Translating IDPlL forms into DPlL forms

Here we show that the formalisms presented here are close relatives to the dynamic plural logic of van den Berg.

I will show here that with respect to some essential actions IDPlL behaves the same as DPlL. The version of DPlL I'm focusing on is in figure 5.4 (cf. with section 4.1). Notice that we ignore negation for now. I will mainly be concerned with issues of undefinedness. I will compare this version of DPlL with the semantics defined in section 5.3.2, that is, with respect to a semantics which includes a distributivity operator which cannot handle the accessibility of non-functional material within its scope. I repeat the relevant definitions.

(5.46)
$$
\begin{array}{llll}
s[\![\epsilon]\!]s' & =\uparrow & :\Leftrightarrow & \bot \\
s[\![\epsilon]\!]s' & = 1 & :\Leftrightarrow & \exists X \in \wp^+(D_e) :\ s' = \{c^\frown d | d \in X\ \wedge c \in s\}
\end{array}
$$

$$\begin{array}{llll}(5.47) & s[\![P(n,\ldots,m)]\!]s' = \uparrow & :\Leftrightarrow & \max(\{n,\ldots m\}) \geq |s| \\ & s[\![P(n,\ldots,m)]\!]s' = 1 & :\Leftrightarrow & s = s' \;\&\; s[\![P(n,\ldots,m)]\!]s' \neq\uparrow \;\& \\ & & & \langle s[n],\ldots,s[m]\rangle \in I(P)\end{array}$$

$$\begin{array}{llll}(5.48) & s[\![\delta_i(\varphi)]\!]s' = \uparrow & :\Leftrightarrow & i \geq |s| \text{ or } \exists d \in s[i] : s\upharpoonright_{i=d} [\![\varphi]\!]\, s'\upharpoonright_{i=d} = \uparrow \\ & s[\![\delta_i(\varphi)]\!]s' = 1 & :\Leftrightarrow & s[i] = s'[i] \;\&\; s[\![\delta_i(\varphi)]\!]s' \neq \uparrow \;\& \\ & & & \forall d \in s[i] : s\upharpoonright_{i=d} [\![\varphi]\!]\, s'\upharpoonright_{i=d} = 1\end{array}$$

I assume a set of variables $V = \{x_0, x_1, x_2, \ldots, x_v\}$ for DPlL for some fixed $v \in \mathbb{N}$. The sets of stacks involved in IDPlL can then easily be mapped onto the states for plurals of DPlL.

Definition 5.33

From sets of stacks to information states for plurals:

For $c \in D_e^*$:
$c^\dagger := c \cup \{\langle i, \star\rangle \mid |c| \leq i \leq v\}$

Let s be a set of n-tuples:
$s^\ddagger \;:=\; \{\; \{\langle x_i, d\rangle \mid \langle i, d\rangle \in c^\dagger\;\} \mid c \in s\}$

□

For any stack c, $c^\dagger$ returns a stack of v positions, where the positions not defined in c are given the $\star$-value. That is, $c^\dagger$ is a total function from $\{0,\ldots,v\}$ to $D_e \cup \{\star\}$. For a set of stacks s, $s^\ddagger$ has each stack in s replaced by a totalised version in which natural numbers are replaced by their corresponding variable in V. Notice that $s\upharpoonright_{i=d}$ corresponds to $s^\ddagger|_{x_i=d}$ only in case $i < |s|$, since $s\upharpoonright_{i=d}$ is undefined whenever $i \geq |s|$. Furthermore: $s[i] = s^\ddagger(x_i)$ only in case $i < |s|$. In case i exceeds the number of positions in s, $s[i]$ will be undefined whereas $s^\ddagger(x_i)$ will return the empty set.

For formulae we also need to replace the indices of IDPlL with variables. While doing so, however, we should keep track of which variables we have used during the translation process. The following translation procedure does exactly that, provided that we define a way to retrieve the number of existential quantifiers in a sentence.

Definition 5.34

Translation from IDPlL to DPlL:

$$\begin{array}{rcl}(\epsilon)^n & := & \epsilon_{x_n} \\ (Pi\ldots j)^n & := & Px_i\ldots x_j \\ (\delta_i(\varphi))^n & := & \delta_{x_i}((\varphi)^n) \\ (\alpha;\beta)^n & := & (\alpha)^n;(\beta)^{n+e(\alpha)}\end{array}$$

where :

$$
\begin{array}{rcl}
e(\epsilon) & = & 1 \\
e(Pi \ldots j) & = & 0 \\
e(\delta_i(\varphi)) & = & e(\varphi) \\
e(\varphi;\psi) & = & e(\varphi) + e(\psi)
\end{array}
$$

□

In a context s, we may translate the IDPlL formula φ as $(\varphi)^{|s|}$.

Given these definitions it now becomes interesting to show that φ in context s corresponds to $(\varphi)^{|s|}$ in context $s^\dagger$. In other words:

Lemma 5.35

$$s[\![\varphi]\!]s' = 1 \;\Rightarrow\; s^\dagger \,[\![(\varphi)^{|s|}]\!]^+ s'^\dagger$$

□

Let us start with the crucial atomic formulae, namely 'ϵ' versus 'ϵ_{x_i}'. Notice that if $s[\![\epsilon]\!]s' = 1$, it holds that $s^\ddagger(x_{|s|}) = \emptyset$. Also, $s'[|s|] \in \wp^+(D_e)$. Since $s'^\ddagger(x_{|s|}) = s'[|s|]$, it holds that $\exists X$ such that $s'^\ddagger = \{g[x/d] \mid g \in s^\dagger \;\&\; d \in X\}$. (Compare with figure 5.4.)

For predication things are different. Van den Berg does not need to check incoming assignments for undefinedness, since in a state G, $G(x)$ returns the empty set whenever all assignment functions are undefined with respect to x. In IDPlL, we assumed that undefinedness never occurs *within* a stack. Indices exceeding the size of the stack, however, make predication uninterpretable and hence predication is only defined if the argument-indices do not exceed this size. We did show, however, that in natural language fragments, where determiners determine the indices of the arguments of the predicates, no undefinedness will actually occur. The translation, thus has to insist on proper indices for the arguments of predicates. Then, it is easy to see that if $s[\![Pi,\ldots,j]\!]s'$ is defined and true then $s^\ddagger\,[\![Px_i,\ldots,x_j]\!]^d s'^\ddagger$, since $s = s'$ and hence $s^\ddagger = s'^\ddagger$. Since $i,\ldots,j$ were all defined in s, $x_i,\ldots,x_j$ will be defined in $s^\ddagger$. And thus $s[n] = s^\ddagger(x_n)$ for all n in $\{i,\ldots,j\}$. Thus, when $s[\![Pi,\ldots,j]\!]s'$ is both defined and true, $s^\ddagger\,[\![Px_i,\ldots,x_j]\!]^+ s'^\ddagger$.

With respect to distribution, notice first that $(s\restriction_{i=d})^\ddagger$ equals $s^\ddagger|_{x_i=d}$, whenever $s\restriction_{i=d}$ is defined (that is, when $i < |s|$.) It follows then that if $s[\![\delta_i(\varphi)]\!]s'$ is defined, $s^\ddagger\,[\![(\delta_i(\varphi))^{|s|}]\!]\,s'^\ddagger$ will be defined as well. Since we assumed that if $i < |s|$ there is no $c \in s$ such that $c(i) =\uparrow$, the condition $G|_{x=\star} = H|_{x=\star}$ is vacuously true for DPlL formulae which are translations of IDPlL forms. Consequently, from $s[\![\delta_i(\varphi)]\!]s' = 1$ it follows that $s^\ddagger\,[\![(\delta_i(\varphi))^{|s|}]\!]\,s'^\ddagger$.

Chapter 6

Quantifier Dynamics

Let us take stock of what we have achieved so far. In the last two chapters, we have focused on what is demanded of a dynamic semantic theory of discourse anaphora in accounting for some aspects of plurality and quantification. We saw the need for a structured notion of context and argued that careful variable management is in order. So far, we have restricted our attention to distributivity (in particular, by focusing on the floating distributive quantifier 'each'), while the noun phrases in our examples were simple indefinites interpreted as existential quantifiers. In this chapter, we will turn to quantificational noun phrases and their dynamic semantics.

Recall from chapter 2, that the DRT analysis of quantificational NPs (QNPs) successfully modelled the following aspects: (i) QNPs are distributive;[1] (ii) QNPs introduce their full reference set; (iii) QNPs introduce maximal antecedents that correspond to indefinites in their scope; (iv) QNPs cannot antecede pronouns in their scope, without leading to a bound variable interpretation; and (v) a QNP may introduce dependencies in discourse.

DRT established this success by three theoretical proposals. First of all, QNPs introduce static, distributive quantificational structures (duplex conditions). Second, these structures in turn trigger the introduction of referents identified with abstractions over conditions in the triggering duplex condition. Finally, subsequent quantification over such a referent identified with an abstraction triggers the accessibility of referents involved in the abstraction. This latter principle was introduced only to account for (v) above. Duplex conditions account for (i) and in tandem with abstraction, they explain (iv). Abstraction by itself takes care of (ii) and (iii).

The aim of this chapter is to cover the same aspects of quantification

[1]Recall, however, that this is a simplifying generalisation. See footnote 14 of chapter 2. I leave the data mentioned there for further research.

using just a single proposal: quantificational noun phrases involve a distributive evaluation of their restrictor and scope. The specific semantics of this distributive interpretation will take care of all referential effects.

Let us briefly look ahead and give a preview of the proposal, by abstracting away from the IDPlL framework and the use of states and stacks. Van den Berg's approach to distributivity applied to the running example in (6.1) can roughly be paraphrased as in (6.2).

(6.1)

(a). Three students each wrote a paper.
(b). They each sent it to L&P

(6.2)

(a). Call r a set of three students
Call r' the set of student-paper pairs in the 'write' relation containing a student in r
Every student in r should be in r'.
(b). Call r'' the set of pairs in r' that are also in the relation 'sent-to-L&P'
Every student in r' should be in r''.

In the paraphrases, r' and r'' denote relations or sets of pairs. Alternatively, in the terminology of chapter 5, r' and r'' are states of size two. The set r is a special kind of relation, namely a *unary* relation, or in terms of chapter 5, a set of one-place stacks. In the paraphrases in (6.2), distributivity is modelled by considering the student-paper pairs in the relation instead of considering the set of students and the set of papers separately. The idea of the approach in the current chapter is to develop this treatment of distributivity into an account of distributive quantification. Consider, for instance, parallel to (6.1) and (6.2), the example in (6.3) and its paraphrase in (6.4).

(6.3)

(a). Most students wrote a paper.
(b). Exactly three students sent it to L&P.

(6.4)

(a). Call r the set of salient students
Call r' the set of student-paper pairs in the 'write' relation where the student is in r
r' should contain a majority of the students in r
(b). Call r'' the set of pairs in r' that are also in the 'sent-to-L&P' relation
r'' should contain exactly three of the students in r'.

In these paraphrases, r, r' and r'' correspond to contexts. Notice the following: in (6.4), r' contains both the students that wrote a paper and the papers that were written by a student. Moreover, r'' contains the students that sent *their*(!) paper to Linguistics and Philosophy. We can give similar paraphrases if an indefinite is embedded in the restrictor of a quantificational structure.

(6.5) Most students who read a book enjoyed it.

(6.6) Call r the set of student-book pairs in the 'read' relation
Call r' the set of pairs in r that are also in the 'enjoy' relation
r' should contain a majority of the students in r.

The structure of this chapter is as follows. Section 6.1 develops a strategy for the interpretation of quantificational noun phrases; this is inspired by van den Berg-style distributive interpretation, using the framework developed in chapter 5. Next, in section 6.2, we focus on the maxset and the interpretation of the restrictor. This is necessary because of two complications, namely the conditions on the accessibility of the maximal set and the well-known weak/strong distinction in the interpretation of donkey sentences. In section 6.3, we show that this proposal correctly handles entailment patterns involving decreasing quantifiers and pronouns. In 6.4, the findings of chapter 5 and the current chapter are combined in an exposition of the full proposal.

6.1 Quantification and distributivity

The work of van den Berg, discussed in chapter 3, shows us that maximalisation of reference to quantificationally embedded indefinites and dependency phenomena are really two sides of the same coin. Once context keeps track of functional dependencies, it will have to specify all values of the indefinite. That is, a van den Bergian representation of 'the boys who wrote a paper', necessarily also represents 'the papers written by boys', since each boy will be paired with *his* paper. The key to constructing these representations of context was distributivity, which combines updates at the level of atomic entities to a single update involving pluralities.

As we mentioned above, quantificational noun phrases have much in common with this view on distributivity: they are distributive, they access and create dependencies and they trigger maximal reference to embedded indefinites. It therefore seems a good idea to model the dynamics of quantificational noun phrases on the basis of van den Berg-style distributivity.

Recall how the δ operator changed the context. Say we have a state s and say we are interested in the set assigned to slot i. Distributing over this set creates contexts which are formed by appending a state $s\restriction_{i=d}$ to s for a succession of d's in $s[i]$. Each such an extended context is used to interpret the scope of δ and results in an output state which is also an extension of s. The combine of all these output extensions is the output for the distribution operation.

In distributive quantification, something similar can be observed. The restrictor sets up a set (the maxset) and the nuclear scope considers the atoms of the set one by one. The only difference is that in distributive quantification it is not always necessary for all the atoms to satisfy the scopal predicate. In fact, the output to distributive quantification is the collection of parts of the restrictor interpretation that comply with the nuclear scope.

What we are interested in, then, is a way of collecting successful updates. Say that we focus, for instance, on a VP 'wrote a paper', which corresponds to the function '$F = \lambda i.\lambda s.(\exists^\star \cdot \text{PAPER}^\star(|s|) \cdot \text{WROTE}^\star(i, |s|))(s)$' and, say, that S is a salient set of students. Consider now the following:

(6.7) $t = \lambda s. \cup \{u \mid \exists d \in S : s^\sqcap u \in F(|s|)(s^\sqcap\{\langle d\rangle\})\}$

Given a state s, this function unifies sets of stacks u which are potential extensions of s with respect to F. In any state s, $t(s)$ is a set of stacks of two positions, representing student-paper pairs such that the student wrote the paper. It collects those extensions of s that form the output of $F(|s|)$ relative to s extended with one of the students in S. In this case, $t(s)$ is the same set of stacks no matter what the input context is like. This is because 'wrote a paper' is completely context independent. Had we chosen a VP like 'sent it to L&P' for F, then different input contexts would have resulted in different states.

We may generalise this way of collecting successful updates as follows:

(6.8) $\sigma := \lambda S_{\langle e,t\rangle}.\lambda P_{\langle \iota,T\rangle}.\lambda s. \cup \{u \mid \exists d \in S : s^\sqcap u \in F(|s|)(s^\sqcap\{\langle d\rangle\})\}$

For instance, in an empty state and a world wherein s_1, s_2 and s_3 wrote the papers p_1, p_2 and p_3 respectively and no other student wrote a paper, $\sigma(I(\text{STUDENT}))(F)$ results in three student-paper pairs such that the student wrote the paper. This is because,

$$
\begin{array}{ccc}
F\left(\begin{array}{c}0\\ \boxed{s_1}\end{array}\right) & = & \begin{array}{cc}0 & 1\\ \hline \multicolumn{1}{|c|}{s_1} & \multicolumn{1}{c|}{p_1}\\ \hline\end{array}\\
F\left(\begin{array}{c}0\\ \boxed{s_2}\end{array}\right) & = & \begin{array}{cc}0 & 1\\ \hline \multicolumn{1}{|c|}{s_2} & \multicolumn{1}{c|}{p_2}\\ \hline\end{array}\\
F\left(\begin{array}{c}0\\ \boxed{s_3}\end{array}\right) & = & \begin{array}{cc}0 & 1\\ \hline \multicolumn{1}{|c|}{s_3} & \multicolumn{1}{c|}{p_3}\\ \hline\end{array}\\
F\left(\begin{array}{c}0\\ \boxed{s_4}\end{array}\right) & = & \emptyset\\
\vdots & & \vdots\\
\text{etc.} & & \text{etc.}
\end{array}
$$

This operation can already provide the tools for the analysis of simple quantificational expressions. Take for instance ‘most students wrote a paper’. Its interpretation will be:

(6.9) $\lambda s.\lambda s'.s' = s^{\sqcap}\sigma(I(\textsc{student}))(F)\ \&\ \langle I(\textsc{student}), s'[|s|]\rangle \in I(\textsc{most})$

The state s' results from extending s with all the relevant student-paper pairs. Such a state is an output only if the set of students it contains is in the ‘most’-relation with the total set of students, $I(\textsc{student})$. That is, the set $s'[|s|]$ returns the reference set. Of course, this oversimplifies the role of the restrictor. Most importantly, the restrictor is not simply a set; it may have dynamic effects itself. For instance, the N′ ‘students who read a book’ presents not only a set of students to the scope, but actually student-book pairs. This is because, for each student, the VP has access to the book he or she read. Moreover, restrictors may contain anaphoric material which needs to find an antecedent in the incoming context. We may then use σ for the restrictor interpretation as well. For instance, consider ‘students who read a book’, interpreted as the $\langle\iota, T\rangle$-function ‘$G = \lambda i.\lambda s.(\exists^\star \cdot \textsc{book}^\star(|s|) \cdot \textsc{read}^\star(i, |s|))(s)$’. As a restrictor it sets up the state $\sigma(D_e)(G)(s)$, namely the set of student-book pairs such that the student read the book. For instance,

(6.10) $r = \sigma(D_e)(G)(\{\langle\rangle\}) =$

0	1
s_1	b_1
s_2	b_2
s_3	b_3
s_4	b_4

This state can subsequently be used as a base for the interpretation of the scope. Take ‘wrote a paper’ again and its interpretation as F, above. The VP considers student-book pairs one by one. However, we cannot use σ to this end, since it takes sets, not sets of stacks. We therefore need a means of collecting successful updates which is even more like δ, in that it considers states extended with a sub-state:

(6.11) $\varsigma\ :=\ \lambda t.\lambda i.\lambda P.\lambda s. \cup \{u \mid \exists d \in t[i] : s^{\sqcap}u \in P(i + |s|)(s^{\sqcap}t\restriction_{i=d})\}$

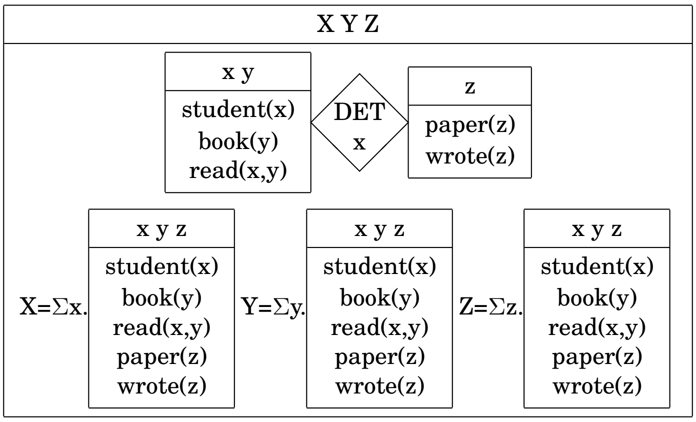
Let me explain what this says. The function ς takes a state t and an index i and a function of type $\langle\iota, T\rangle$ to return a state modifier. In a state s, it extends s with parts of t that assign a single atomic value d at i. The extensions of s, u, that result from interpreting P with respect to this atomic individual are collected. For instance, for ‘DET students who read a book wrote a paper’ we can now analyse the scope interpretation as $\varsigma(r)(0)(F)$, which in an empty state results in:

(6.12) $r' = \varsigma(r)(0)(F)(\{\langle\rangle\}) =$

0	1	2
s_1	b_1	p_1
s_2	b_2	p_2
s_3	b_3	p_3

We can now compare the set of students in r at 0 with that in r' at 0. These two sets are supposed to be in accordance with the relation between sets expressed by the determiner. If this is so, r' seems a likely candidate for an output state of the quantificational structure. Not only is $r'[0]$ the full reference set (the set of students that read a book and wrote a paper), r' also has exhaustively collected the values of the indefinites embedded in the quantificational structure. In fact, compare r' with the interpretation of the following DRS in the same situation. A duplex condition corresponding to 'DET students who read a book wrote a paper' triggers three abstraction procedures, as shown in (6.13). In this DRS, X corresponds to $r'[0]$, Y to $r'[1]$ and Z to $r'[2]$.

(6.13)



In the dynamic setup we are pursuing, however, there is no need for a separate procedure for quantification (duplex conditions in DRT) and antecedent formation (abstraction in DRT). In constructing the state r' above, using the state r we already have all the future antecedents we need. The truth-conditions of quantification, moreover, boil down to nothing more than a comparison of r' and r.

Recall, however, that DRT needed to stipulate a third principle governing plural anaphora, which said that if an abstracted set is quantified over, the material used in the abstraction becomes accessible. In our construction of the set of stacks r (the interpretation of the restrictor), merely sets are taken into account, so we cannot expect to be able to predict the effects of this principle as yet.

For instance, in an example like (6.14), the second sentence needs access to the individual papers that were introduced in the first sentence.

(6.14) Every student wrote a paper. Most students sent it to L&P.

The restrictor 'students' in the second sentence should therefore take its set of students from the context, so that the VP 'sent it to L&P' is able to consider student-paper pairs. This means, that instead of 'σ', we should use 'ς' to set up the restrictor as well. Say, for instance that the first sentence in (6.14) results in a state with the three student-paper pairs we mentioned above (t_1 in (6.15)). Say now that only the students s_1 and s_2 (which qualify as a majority) sent their paper to L&P. Given an obvious resolution for the object pronoun, the following states are thus involved in the interpretation of (6.14).

(6.15) $t_1 =$

0	1
s_1	p_1
s_2	p_2
s_3	p_3

(6.16) $t_2 = \varsigma(t_1)(0)(\text{STUDENT}^\star)(t_1) = t_1$

(6.17) $t_3 = \varsigma(t_2)(0)(\text{SENT\_TO\_L\&P}^\star(2,3))(t_1) =$

0	1
s_1	p_1
s_2	p_2

The restrictor 'students' uses the input state t_1 as a resource for collecting the salient set of student plus the sets dependent on it. In effect, it returns the input state, so $t_1 = t_2$. The scope 'sent it to L&P' collects those student-paper pairs in t_2 such that the student sent the paper to L&P. The state t_2 is thus reduced to containing only two pairs.

We can now come to a general interpretation scheme for determiners. Say that Q is interpreted in the model as a relation between two sets, then:

(6.18) $$Q_i^\star := \lambda R.\lambda S.\lambda s.\lambda s'.\exists r,r' : s' = s^\sqcap r' \ \&\ \langle r[i], r'[i]\rangle \in I(Q)\ \&\\ r = \varsigma(s)(i)(R)(s)\ \&\\ r' = \varsigma(r)(i)(S)(s)$$

The function $Q_i^\star$ is of a type which is to be expected: $\langle\langle\iota.T\rangle,\langle\langle\iota,T\rangle,T\rangle\rangle$. The index i supplies the so-called context set for quantification (Westerståhl 1984). It makes sure that the domain of quantification is contextually restricted with $s[i]$. Let us for now, for sake of simplicity, assume that determiner functions are magically decorated with such an index. (We will address the issue when we turn to the strong/weak distinction in section 6.2.)

In (6.18), '$r[i]$' corresponds to the maximal set. The state r is formed taking atoms x from $s[i]$ and testing whether they satisfy the restrictor condition R in s extended with $s\restriction_{i=x}$. The reference set is $r'[i]$. It is formed by taking atoms x from $r[i]$ and testing whether they satisfy the scope condition S in s extended with $r\restriction_{i=x}$. Since r', rather than r, is appended to s to yield the output state, only the reference set is introduced in discourse.

Here is an example. The function in (6.19) is the interpretation of 'most students wrote a paper'. It is a function taking a state s and incrementing it with a reference set of students (at slot $|s|+i$) and a set of papers parallel to this reference set.

(6.19) $\text{MOST}^\star_i(\text{STUDENTS}^\star)(\lambda i.\lambda s.((\exists^\star \cdot \text{PAPER}^\star(|s|) \cdot \text{WROTE}^\star(i,|s|))(s)))$

Say we apply this function to some state s. This is what happens: first we take atoms d from $s[i]$ and increment s with $s\lceil_{i=d}$, yielding (say) s'. Then we form the state $(\lambda i.\text{STUDENTS}^\star(i))(i+|s|)(s')$. This returns either s' or $\emptyset$, depending on whether or not d is a student. If we collect all such updates into a set r, this state represents the distributive update potential of the restrictor. For the interpretation of the scope we now increment s with the atomic parts of r. These updates result in incrementations of s containing a student in slot $|s|+1$ and they are collected in a state r'. We can now compare how many students there are in r with how many there are in r'. If this comparison is conform the relation expressed by 'most', then we output s incremented with r'.

As another example consider the input context (6.20) to the function in (6.21) and (6.22) respectively.

(6.20) $s = $

0	1
s_1	p_1
s_2	p_2
s_3	p_3
s_4	p_4

(6.21) $\text{MOST}^\star_0(\text{STUDENT}^\star)(\lambda i.\text{SENT\_TO\_L\&P}^\star(i,3))$

(6.22) $\text{MOST}^\star_0(\text{STUDENT}^\star)(\lambda i.\text{SENT\_TO\_L\&P}^\star(i,1))$

Whereas (6.21) corresponds to a likely resolution of 'most students sent *it* to L&P', the function in (6.22) is likely to originate from 'most students sent *them* to L&P' (i.e. describing the rather odd situation wherein most students sent all four papers to L&P). Both forms are indexed with 0, that is, the set of students in (6.20) is taken as the context set. In both forms, then, the restrictor expands the input with a pair of a student and a paper in s and checks whether in that expanded state the position $2(=|s|+0)$ is filled by a student. In effect, this returns s again. These same pairs are considered for the nuclear scope. That is, the function in the second argument position of $\text{MOST}^\star$ is applied to a context which results from appending $s\lceil_{0=d}$ to s for students d. In such contexts, the position 3 is the paper that paired with the student d in s, while the position 1 is the total set of papers in s. The successful extensions $s\lceil_{0=d}$ are collected in a state r'. In case students in r' are a majority of the students in s, we get an output, namely $s^\frown r'$.

The procedural nature of this exposition of the semantics of distributive quantification is not for presentation purposes only. It illustrates a key contrast with the interpretation of referential noun phrases. Quantificational NPs introduce a set, distributively present this set as an argument for the VP and then check whether the determiner relation is satisfied. Referential noun phrases, on the other hand, introduce a set, count and then interpret the VP with respect to this set. The two strategies can be paraphrased as follows: for QNPs– 'introduce atoms x ; predicate over the atoms ; check whether the successful x's form a witness set' and for RNPs – 'introduce a set X ; check whether X is a witness ; predicate over X'. Notice how these paraphrases explain the differences between RNPs and QNPs mentioned above. In the quantificational case, maximality, distributivity, dependency and the inaccessibility of the reference set within the scope follow directly. (Compare with Szabolcsi 1997 for a statement of similar strategies within DRT terminology. See also Winter 1998.)

There is another contrast between the quantificational and the referential strategy of introducing entities in the discourse. RNPs range over non-empty subsets of the domain of entities, while QNPs include the empty set. This means that all downward monotone NPs will have to be analysed as being quantificational, since they, by definition, include the empty set in their denotation. This is not surprising, however, given that with respect to many properties, downward monotone NPs systematically pair with QNPs. For instance, they are distributive and their refset is accessible only outside of their scope.

A noun phrase like 'no students,' then, is analysed as the $\langle\langle\iota, T\rangle, T\rangle$ function '$\lambda P.\text{NO}_i^\star(\text{STUDENT}^\star)(P)$'. Notice that this function for some verb phrase P and some state s either returns the empty set or returns s. This is because the relation NO can only be satisfied when there is no extension of s that satisfies the conditions in the nuclear scope. In other words, a sentence '[[No N′] VP]' is (correctly) analysed as a test.

Let us evaluate the merits of the definition in (6.18). We already established that ς constructs states in such a way that both the refset and sets depending on the refset are exhaustively represented in context. Moreover, ς assures a purely distributive interpretation by considering sets atom by atom only. Note, moreover, that the full refset is only represented in context once the interpretation of the quantificational structure is completed.

In sum, the above proposal reaches results comparable to those of DRT, using a single interpretational strategy of determiners. Let us, now, turn to other aspects of quantificational structures. We begin with discussing the maximal set.

6.2 The maxset

We have remained silent about the accessibility of the maximal set. The interpretation of a quantificational noun phrase as given above predicts the following: the maximal set is at no point accessible. The reason is that at no point in the evaluation is the full set of successful updates with the restrictor (systematically called 'r' above) taken as an input context. The scope only considers atoms from this set and a successful quantificational structure only increments the input state with states due to the scope.

This begs the question as to how the maxset can sometimes function as an antecedent. However, given that we have so far assumed that restrictors take their value from a salient set in the context, it follows that the (potential) accessibility of maxset is due to the fact that it already occurs in context. In other words, its existence was presupposed.

We will now consider two details of how restrictions are interpreted which deserve more attention. Both deal with a (different) weak/strong distinction in quantification. The first concerns the weak and strong readings of donkey sentences. The second addresses the weak/strong distinction in the sense of intersective/non-intersective quantifiers.

6.2.1 Weak versus strong readings

The interpretation of a restrictor collects every possible extension of the current information state that satisfies the restrictor clause. However, in the case of complex restrictors containing indefinites, there is no straightforward way of choosing how to collect these extensions. For instance, the N′ 'boys who wrote a paper' extends the input state with relevant boy-paper pairs. However, when one of the boys wrote more than one paper, there are two options: either we collect all the pairs, entering some boys more than once into the extension (call it the 'A'-option), or we allow extensions which do not exhaustively collect all pairs, as long as all boys that wrote a paper are in the extension (call it 'B'). The choice between these two ways of building up restrictions is not straightforward.

For example, in a situation with three boys writing a paper, wherein b_1 wrote p_1 and b_2 wrote p_2 and b_3 wrote p_3 and p_4, (6.23a) gives the A-option for extending the context, while (6.23b) and (6.23c) give two possible B-style extensions.

$$\text{(6.23)}\quad \text{a.}\ \begin{array}{cc} 0 & 1 \\ \hline \multicolumn{1}{|c|}{b_1} & \multicolumn{1}{c|}{p_1} \\ \hline \multicolumn{1}{|c|}{b_2} & \multicolumn{1}{c|}{p_2} \\ \hline \multicolumn{1}{|c|}{b_3} & \multicolumn{1}{c|}{p_3} \\ \hline \multicolumn{1}{|c|}{b_3} & \multicolumn{1}{c|}{p_4} \\ \hline \end{array} \quad \text{b.}\ \begin{array}{cc} 0 & 1 \\ \hline \multicolumn{1}{|c|}{b_1} & \multicolumn{1}{c|}{p_1} \\ \hline \multicolumn{1}{|c|}{b_2} & \multicolumn{1}{c|}{p_2} \\ \hline \multicolumn{1}{|c|}{b_3} & \multicolumn{1}{c|}{p_3} \\ \hline \end{array} \quad \text{c.}\ \begin{array}{cc} 0 & 1 \\ \hline \multicolumn{1}{|c|}{b_1} & \multicolumn{1}{c|}{p_1} \\ \hline \multicolumn{1}{|c|}{b_2} & \multicolumn{1}{c|}{p_2} \\ \hline \multicolumn{1}{|c|}{b_3} & \multicolumn{1}{c|}{p_4} \\ \hline \end{array}$$

The scope is interpreted by extending the input context with parts of one of the above states. For instance, for the A-option, in a state s a nuclear scope 'sent it to L&P' will be interpreted by considering the state s, extended with (6.23a)$\upharpoonright_{0=b_i}$ for i ranging over 1, 2, and 3. With respect to the third boy, this means the scope is interpreted as follows (assuming 'it' to be resolved to be dependent on the papers):

$$\text{(6.24)}\quad \text{SENT\_TO\_L\&P}(|s|, |s|+1)\left(s^\sqcap \begin{array}{cc} 0 & 1 \\ \hline \multicolumn{1}{|c|}{b_3} & \multicolumn{1}{c|}{p_3} \\ \hline \multicolumn{1}{|c|}{b_3} & \multicolumn{1}{c|}{p_4} \\ \hline \end{array}\right)$$

In other words, according to the A-option, in order for the sentence 'D boys who wrote a paper sent it to L&P' to be true, it should hold that D boys sent *all* the papers they wrote to L&P.

In the 'A' set-up, pronominal reference to the indefinite inside the restrictor clause will be interpreted as exhaustive. That is, for each boy it will retrieve all the papers he wrote. This is not the case for the 'B' option, where the nuclear scope may be false for one of the papers written by a particular boy as long as there exists another one he did sent to L&P.

All this is closely connected to the well-known distinction between weak (or existential) versus strong (or universal) readings of quantificational sentences. For instance, while (6.25) is preferably interpreted weakly as being true when every guest owning one or more credit cards uses one of his cards to pay the bill, (6.26) seems to lead to a stronger reading in which every farmer is required to beat all the donkeys he owns. The weak paraphrases (indicated as $\exists$) and the strong paraphrases (indicated as $\forall$) are as in (6.27) and (6.28), after Kanazawa 1994.

(6.25) Every guest who owned a credit card, used it to pay the bill.

(6.26) Every farmer who owns a donkey, beats it.

(6.27) Q farmer who owns a donkey, beats a donkey he owns. ($\exists$)

(6.28) Q farmer who owns a donkey, beats every donkey he owns. ($\forall$)

In absence of conflicting information, it seems that 'every' triggers a strong reading. The knowledge that it takes only one credit card to pay a bill, weakens the reading for (6.25). Not all determiners show similar preference patterns for weak and strong readings however. For instance, (6.29) is only verified when no farmer beats *any* donkey he owns, not just some of them. In fact, the strong reading seems impossible to obtain.

(6.29) No farmer who owns a donkey, beats it.
#*"No farmer who owns a donkey beats every donkey he owns"* ($\forall$)
"No farmer who owns a donkey beats a donkey he owns" ($\exists$)

Some downward entailing environments, however, do not seem to require a similarly weak reading. Kanazawa (1994) points out that an example like (6.30) seems to favor a strong reading. Krifka (1996b), however, shows that this might be due to the fact that 'not every' is not a determiner, since the semantically similar "less than 100%" pairs with other determiners which are downward entailing in their second argument and requires an existential interpretation.

(6.30) Not every farmer who owns a donkey beats it.
"Not every farmer who owns a donkey beats every donkey he owns" ($\forall$)
?*"Not every farmer who owns a donkey beats a donkey he owns"* ($\exists$)

(6.31) Less than 100% of the farmers who own a donkey beat it.
?*"Less than 100% of the farmers who own a donkey beat every donkey they own"* (∀)
"Less than 100% of the farmers who own a donkey beat a donkey they own" (∃)

In our current approach, the distinction between weak and strong comes out rather different. What we have called the 'A' option, where all values are bundled together and consequently pronouns in the scope have to exhaustively refer to an indefinite in the restrictor clause, will provide the strong reading for upward entailing quantifiers, but the weak reading for downward monotone ones. Consider, for instance (6.32), below, and its paraphrase in (6.33).

(6.32) No farmer who owns a donkey beats it.

(6.33) Call r the set of farmer-donkey pairs in the own relation.
Call r' the set of pairs in r that are also in the 'beat' relation.
r' should be the empty set.

This is in accordance with the ∃-paraphrase in (6.29). The weakness of (6.33) is due to the fact that the quantificational relation (and hence its monotonicity properties) only has its effect after processing the scope. Notice that the approach to the interpretation of quantificational noun phrases set out above represents the A-approach. For the 'B' option, upward quantifiers are interpreted as being weak, while downward ones receive a universal reading. It seems then that although all determiners prefer the 'A' option, sometimes a shift is made toward generating multiple environments.[2]

Let me briefly turn to a related issue. One would expect that the weak/strong distinction at the sentence level also occurs on discourse level. That is, it would not be surprising if the discourse in (6.34) were ambiguous between a reading wherein all papers written by the students were not very good and one wherein any set of papers, each of them written by one of the boys, were not good.

(6.34) Every student wrote a paper. They weren't very good.

Although it is difficult to get clear judgments, it seems that strong readings are by far the more preferred ones. In fact, it is questionable whether weak readings really exist. Intuitively, the second sentence in (6.34) is false when one of the three students wrote one bad and one good paper. (See van der Does 1993 for a similar speculation.)

Nevertheless, an example like (6.35) might complicate things.

[2]See, especially, Kanazawa 1994 and Geurts 1999 for two interesting discussions of factors involved in the availability of and preferences for certain readings.

(6.35) That day, all the guests wanted to leave the hotel. Unfortunately, the banks were closed. Luckily, every guest owned a credit card. They used them to pay their bills.

As in (6.25), (6.35) does not require a guest owning more than one credit card to use all his cards to pay the bill. One could argue, however, that the strong reading is not that bad for this example. It simply says that every guest used his credit cards to pay the bill. That is, he payed by credit card and it does not matter which of his credit cards he used. In other words, the guest uses his credit resources to pay the bill.

The proposal in this chapter predicts that only strong readings exist in discourse. This is again due to the way the ς-operation collects successful updates. It does so exhaustively, so following 'every guest owned a credit card', the refset is paired with all the cards owned by the guests. Although my remarks above seem to strongly suggest that this is a welcome prediction, I leave it an open question as to whether or not apparent counter-examples like (6.35) can be accounted for by other means.

6.2.2 Parasitic domains

Let us deal with the index-decoration of determiners. We suggested that a structure $Q(A)(B)$ is to be translated into the function $Q_i^\star(A^\star)(B^\star)$. Here, the purpose of the index i was to supply the contextual restriction of the domain of quantification (following, again, Westerståhl (1984)).

Of course, such an index decoration is not compositionally given. In effect, the quantifier is anaphoric. Or, to be more precise, in cases where the quantifier is strong and therefore presupposes its maxset, this presupposition may be anaphorically resolved. So instead of the function $Q_i^\star(A^\star)(B^\star)$, quantificational structures should correspond to underspecified representations: $Q_{\mathsf{X}_0}(A)(B)$.

Note however that weak quantifiers (that is, intersective ones) do not need a contextually supplied context set. In existential-there constructions, for instance, clearly no domain restriction is needed.

(6.36) There are four boys in the garden.

This example is intelligible even if no set of salient boys exist in context. In fact, it simply reports the existence of such a salient set. Another example is the well-known observation that the object of verbs like 'have' may only be weak (see, e.g., de Hoop 1992). In (6.37), then, no salient set of windows is assumed. The obligatorily strong determiner 'most' is banned from such positions, since it requires such an antecedent.

(6.37) This house has many windows.

(6.38) *This house has most windows.

Consequently, weak quantifiers are in need of a different interpretation, one which has no anaphoric properties. The following provides a straightforward solution. Instead of taking a set from a slot in the input context, the weak quantifier takes its values from a state $\{\langle d\rangle | d \in D_e\}$, the state containing only the universe.

(6.39) $Q^\star_{\mathsf{W}} := \lambda R.\lambda S.\lambda s.\lambda s'.\, \exists r, r' : s' = s^\sqcap r' \,\&\, \langle r[i], r'[i]\rangle \in I(Q) \,\&$
$r = \varsigma(\{\langle d\rangle | d \in D_e\})(0)(R)(s) \,\&$
$r' = \varsigma(r)(i)(S)(s)$

According to this definition, in the interpretation of a weak quantifier, the maximal set (given by r) serves only as a resource for the interpretation of the scope. It is not supposed to be given by the context and consequently is never available for anaphora.

6.3 Entailment and emptiness

Recall from chapter 2 and 3 that pronominal reference to the refset of a downward entailing quantifier demonstrates the non-emptiness of that set, whereas this is not guaranteed by the antecedent quantification by itself.

(6.40) Few senators admire Kennedy.
$\not\Rightarrow$ Some senator admires Kennedy.

(6.41) Few senators admire Kennedy. They are very junior.
$\Rightarrow$ Some senator admires Kennedy

Let us, for the sake of simplicity, assume that 'few' is interpreted as weak, here. (Although nothing hinges on this). That is, 'Few senators admire Kennedy' is interpreted as the function in ('6.42')

(6.42) $\text{FEW}^\star_{\mathsf{W}}(\text{SENATOR}^\star)(\text{ADMIRE\_K}^\star)$

Consider now a world wherein none of the senators admire Kennedy. Then, call r the set of one-place stacks comprising the set of senators. Then r' is the set of states $r\restriction_{0=d}$ such that d admires Kennedy. Since there is no such d, it follows that $r' = \emptyset$. Clearly now, r and r' satisfy the 'few'-relation, so the output to (6.42) in this world for any s is s again. In other words, (6.42) is true in this world, but given that no senator admiring Kennedy was found, the input state was not incremented. (Or, in fact, the input state has been incremented with the empty set.) This shows that (6.40), in the form (6.43) below, holds.

(6.43) $\text{FEW}^\star_{\mathsf{W}}(\text{SENATOR}^\star)(\text{ADMIRE\_K}^\star) \not\models$
$\lambda s.\exists^\star \cdot \text{SENATOR}^\star(|s|) \cdot \text{ADMIRE\_K}^\star(|s|)(s)$

Turning now to the influence of the pronoun, the sentence ‘they are very junior’ is represented as ‘$\text{VERY\_JUNIOR}(\mathsf{x}_0)$’. In a world as described above, i.e. one wherein no senator admires Kennedy, and following (6.42) in context s, the resolutions of this representation are those wherein an index $i < |s|$ is substituted for x_0. That is, there simply is no possible resolution to the refset. Consequently, when the pronoun *is* resolved to the refset of (6.42), we know that we are in a world wherein some senators admire Kennedy. The entailment in (6.41) follows.

More formally,

(6.44) $\lambda s.(\text{FEW}^{\star}_{\mathsf{W}}(\text{SENATOR}^{\star})(\text{ADMIRE\_K}^{\star}) \cdot \text{VERY\_JUNIOR}^{\star}(|s|))(s) \models$
$\lambda s.\exists^{\star} \cdot \text{SENATOR}^{\star}(|s|) \cdot \text{ADMIRE\_K}^{\star}(|s|)(s)$

The proof is easy. There is only a single possible output state resulting from combining the function in front of ‘$\models$’ with some state s. In models containing few, but some, Kennedy admiring senators, this function returns the input state s incremented with the set of these admirers. In any other model, the function is undefined since there is no slot $|s|$ at the point where $\text{VERY\_JUNIOR}^{\star}(|s|)$ is encountered. So, in those worlds, this function would have never resulted from resolution. It follows then that all contexts resulting from applying this function to some input state yield an output after applying $\lambda s.\exists^{\star} \cdot \text{SENATOR}^{\star}(|s|) \cdot \text{ADMIRE\_K}^{\star}(|s|)(s)$ to it.

Given these observations, note, however, the following: the text ‘*Few senators admire Kennedy. They are very junior.*’ does not possibly-entail ‘*Some senators admire Kennedy*’ (put more formally in (6.45). This is because in worlds without Kennedy-admiring senators, there will be no resolution to the refset and ‘they’ will be forced to take a completely different antecedent. In sum, only the resolved case, where ‘they’ picks up the reference set, allows us to infer the non-emptiness of this set.

(6.45) $\text{FEW}_{\mathsf{W}}(\text{SENATOR})(\text{ADMIRE\_K}) \cdot \text{VERY\_JUNIOR}(0) \not\mid\!\!\leadsto$
$\exists \cdot \text{SENATOR}(0) \cdot \text{ADMIRE\_K}(0)$

The following two patterns are left to the reader to check.

(6.46) Most students are bachelors $\mid\!\!\leadsto$ They are unmarried

(6.47) Few students are bachelors $\not\mid\!\!\leadsto$ They are unmarried

We have shown here that the proposal predicts the presuppositional nature of a pronoun, by the fact that empty values are not stored in context. Pronoun resolution to the refset is therefore only possible in worlds wherein this set is not empty.

6.4 The full proposal

In this section, we will give an overview of the full account of quantification and anaphora. Some definitions will be repeated from earlier chapters.

In this section, we will collect all manipulations we have defined on our notion of context, namely that of a set of stacks. Our ontology is as follows. D_e is the domain of entities of type e. The set of indices of type ι is $\mathbb{N}$.

Definition 6.36

States

$$s :: \langle\langle\iota, e\rangle, t\rangle \text{ is a state } \Leftrightarrow \exists n \in \mathbb{N} : \ s \subseteq \{\langle d_1, \ldots, d_n\rangle \mid d_1, \ldots, d_n \in D_e\}$$

$$\begin{array}{rcl} s[i] & := & \{c(i) \mid c \in s \ \& \ c(i) \neq \uparrow\} \\ |s| & := & \iota n.\forall c \in s : |c| = n \\ s\restriction_{i=d} & := & \left\{ \begin{array}{ll} \{c \in s \mid c(i) = d\} & i < |s| \\ \uparrow & \text{otherwise} \end{array} \right. \end{array}$$

□

This defines the states and the operations of states we have been using in chapter 5. As we did there, we abbreviate the type $\langle\langle\langle\iota, e\rangle, t\rangle, \langle\langle\iota, e\rangle, t\rangle\rangle$ as T. Let $\mathcal{Q}$ be a set of determiner symbols and $\mathcal{P}$ be a set of predicate symbols. We consider a model $M = \langle D, I\rangle$ which is such that $I(R)$ returns an element in $D_e \times \ldots \times D_e$ for any R in $\mathcal{Q} \cup \mathcal{P}$. We define the following collection of functions:

Definition 6.37

State manipulation

$$\begin{array}{rcll} \varsigma & := & \lambda t.\lambda i.\lambda P.\lambda s. \cup \{u \mid \exists d \in t[i] : s^\sqcap u \in P(i + |s|)(s^\sqcap t\restriction_{i=d})\} & \\ Q_i^\star & := & \lambda R.\lambda S.\lambda s.\lambda s'. \exists r, r' : s' = s^\sqcap r' \ \& \ \langle r[i], r'[i]\rangle \in I(Q) \ \& & \text{for } Q \in \mathcal{Q} \\ & & \qquad r = \varsigma(s)(i)(R)(s) \ \& & \\ & & \qquad r' = \varsigma(r)(i)(S)(s) & \\ Q_{\mathsf{W}}^\star & := & \lambda R.\lambda S.\lambda s.\lambda s'. \exists r, r' : s' = s^\sqcap r' \ \& \ \langle r[i], r'[i]\rangle \in I(Q) \ \& & \text{for } Q \in \mathcal{Q} \\ & & \qquad r = \varsigma(\{\langle d\rangle | d \in D_e\})(0)(R)(s) \ \& & \\ & & \qquad r' = \varsigma(r)(i)(S)(s) & \\ P^\star & := & \lambda i_1, \ldots, i_n.\lambda s.\lambda s'.s = s' \ \& \ \langle s[i_1], \ldots, s[i_n]\rangle \in I(P) & \text{for } P \in \mathcal{P} \\ \exists^\star & := & \lambda s.\lambda s'.\exists X \in \wp^+(D_e) : s' = \{c^\wedge d | c \in s \ \& \ d \in X\} & \\ \neg^\star & := & \lambda s.\lambda s'.s = s' \ \& \ \varphi s = \emptyset & \\ \delta^\star & := & \lambda P.\lambda i.\lambda s.\lambda s'. \, s[i] = s'[i] \ \& & \\ & & \qquad \forall x \in s[i] : \ s^\sqcap s'\restriction_{i=x} \in P(i+|s|)(s^\sqcap s\restriction_{i=x}) & \end{array}$$

$$(\varphi \cdot \psi) \quad := \quad \lambda s. \bigcup\{\psi s' | s' \in \varphi s\}$$

□

Some of these functions are actually partial. Here are the definedness conditions:

Definition 6.38

Definedness

$$\begin{array}{rcl} Q^\star_i(R)(S)(s) &=& \downarrow \;\Leftrightarrow\; i < |s| \;\&\; R(i)(s) = \downarrow \;\&\; S(i)(s) = \downarrow \\ Q^\star_{\mathsf{W}}(R)(S)(s) &=& \downarrow \;\Leftrightarrow\; i < |s| \;\&\; R(i)(s) = \downarrow \;\&\; S(i)(s) = \downarrow \\ P^\star(i_1)\ldots(i_n)(s) &=& \downarrow \;\Leftrightarrow\; \forall 1 \leq j \leq n : i_j < |s| \\ \delta^\star(P)(i)(s) &=& \downarrow \;\Leftrightarrow\; i < |s| \end{array}$$

□

Truth, validity and entailment are defined on functions of type T:

Definition 6.39

Truth and entailment. For $\varphi, \psi :: T$:

$$\begin{array}{rcll} \models_s \varphi &\Leftrightarrow& \varphi(s) \neq \emptyset & \text{(truth)} \\ \models \varphi &\Leftrightarrow& \forall s : \varphi(s) = \downarrow \rightarrow \varphi(s) \neq \emptyset & \text{(validity)} \\ \varphi \models \psi &\Leftrightarrow& \forall s, s' : s' \in \varphi(s) \rightarrow \psi(s') \neq \emptyset & \text{(entailment)} \end{array}$$

□

The representation language L' for context manipulation is extended to include the atomic formulae Q, for $Q \in \mathcal{Q}$. In L', such a form Q is of type T. Similarly, the underspecification language L is extended to include the same atomic forms. (The details for these languages can be found in section 5.5.2.) Resolution and possible entailment are as before.

Figure 6.1 gives a fragment. It only derives sentence meanings, namely underspecified representation in the language L of type T. Interpretation of multi-sentence discourses need a mediating resolution mechanism. Disambiguated sentence meanings, i.e. state transitions, are composed using '·'.

6.5 Conclusion

In this final chapter, we argued that three principles used in the DRT approach to distributive quantification and anaphora can, in the dynamic framework, be replaced by a single interpretative strategy for quantificational noun phrases. By further making a distinction between strong quantifiers, which are anaphoric, and weak quantifiers, which can be interpreted without antecedent, we predict that the former, but not the latter group of NPs license subsequent reference to the maximal set. Finally, since in the proposal quantificational noun phrases increment the input context with the successful output states of the interpretation of their scope, there is no guarantee that a downward entailing NP increments the context at all. Consequently, context will never store an empty value. This accounts for why pronominal reference to the reference set presupposes the non-emptiness of the antecedent.

S	::=	NP VP	X	::=	$(X_1 X_2)$
NP	::=	Det CN	X	::=	$(X_1 X_2)$
VP	::=	IV	X	::=	X_1
VP	::=	TV NP	X	::=	$\lambda u.(X_2(\lambda v.(X_1 v)u))$
VP	::=	MOD VP	X	::=	$(X_1 X_2)$
NP	::=	they	X	::=	$\lambda P.(P(\mathsf{x}_0))$
NP	::=	it	X	::=	$\lambda P.(1(\mathsf{x}_0) \bullet P(\mathsf{x}_0))$
CN	::=	man	X	::=	MAN
CN	::=	woman	X	::=	WOMAN
TV	::=	sent to L&P	X	::=	SENT_TO_L&P
MOD	::=	each	X	::=	$\lambda P.\lambda i.(\delta(P)(i))$
DET	::=	a	X	::=	$\lambda P.\lambda Q.(\exists \bullet 1(0) \bullet P(0) \bullet Q(0))$
DET	::=	three	X	::=	$\lambda P.\lambda Q.(\exists \bullet 3(0) \bullet P(0) \bullet Q(0))$
DET	::=	many	X	::=	$\lambda P.\lambda Q.(\text{MANY}_{\mathsf{W}}(P(0))(Q(0)))$
DET	::=	few	X	::=	$\lambda P.\lambda Q.(\text{FEW}_{\mathsf{W}}(P(0))(Q(0)))$
DET	::=	at_least_three	X	::=	$\lambda P.\lambda Q.(>3_{\mathsf{W}}(P(0))(Q(0)))$
DET	::=	at_most_three	X	::=	$\lambda P.\lambda Q.(<3_{\mathsf{W}}(P(0))(Q(0)))$
DET	::=	many	X	::=	$\lambda P.\lambda Q.(\text{MANY}_{\mathsf{x}_0}(P(\mathsf{x}_0))(Q(\mathsf{x}_0)))$
DET	::=	few	X	::=	$\lambda P.\lambda Q.(\text{FEW}_{\mathsf{x}_0}(P(\mathsf{x}_0))(Q(\mathsf{x}_0)))$
DET	::=	at_least_three	X	::=	$\lambda P.\lambda Q.(>3_{\mathsf{x}_0}(P(\mathsf{x}_0))(Q(\mathsf{x}_0)))$
DET	::=	at_most_three	X	::=	$\lambda P.\lambda Q.(<3_{\mathsf{x}_0}(P(\mathsf{x}_0))(Q(\mathsf{x}_0)))$
DET	::=	most	X	::=	$\lambda P.\lambda Q.(\text{MOST}_{\mathsf{x}_0}(P(\mathsf{x}_0))(Q(\mathsf{x}_0)))$
DET	::=	less_than_half	X	::=	$\lambda P.\lambda Q.(<.5_{\mathsf{x}_0}(P(\mathsf{x}_0))(Q(\mathsf{x}_0)))$

Figure 6.1: Fragment

Concluding Summary

The starting point for this thesis was the observation that plural pronominal anaphora is a very versatile phenomenon. It would be a gross oversimplification to identify the relation between a plural pronoun and its antecedent as a simple coreference relation. In discourse, plural pronouns can have quantified or quantificational, that is, not strictly referential antecedents. Gareth Evans' famous example in (1), for instance, illustrated that some plural pronouns are neither bound by, nor corefer with their antecedents.

(1) Few senators admire Kennedy; and they are very junior.

The question was why the plural pronoun in (1) refers to the set of *all* senators that admire Kennedy. Apparently, the pronoun refers to a set constructed from sets that play a role in the first sentence, namely to the intersection of the denotation of the restrictor (the senators) and the denotation of the verb phrase (the Kennedy-admirers). So, if $(D(A))(B)$ is a quantificational sentence, where D corresponds to the meaning of a determiner, A to that of the restrictor and B to that of the VP, then $A \cap B$ is a potential resolution of a subsequent (plural) pronoun. We called this set $A \cap B$ the *reference set*.

Apart from reference to the reference set, however, it turned out that there are quite a few other types of plural pronominal anaphora. Apart from addressing examples like (1), this dissertation focused on two such other types of anaphora: (i) pronominal reference to the set-theoretical difference between the restrictor denotation and the denotation of the VP, the so-called *complement set* $A - B$, as in example (2), and (ii) pronominal reference to sets directly associated with the reference set, as in the example in (3). (Other forms of plural anaphora which I discussed were non-maximal reference to the reference set and reference to the restrictor denotation, A.)

(2) Few MPs attended the meeting. They went to the beach instead.

(3) Every student wrote a paper. They weren't very good.

Additionally, there was a use of pronouns which is strongly related to the form of anaphora in (3). The detailed relation between students and papers, or, to put it slightly differently, the dependence of papers on students turned out to be accessible in discourse, too.

(4) Every student wrote a paper. They submitted it to L&P.

The pronoun 'it' in (4) takes the indefinite 'a paper' as its antecedent. This pronoun, however, is dependent on the subject pronoun 'they'. The second sentence in (4) means that each student who wrote a paper submitted *his* or *her paper* to L&P. The students-paper pairs described by the first sentence are accessed in the second sentence.

Apparently then, there are several different types of plural anaphora. The most important goal of this thesis was to express this versatility in a semantic theory. By far the most complete semantic theory of plural reference to date is that presented in Kamp and Reyle 1993. In this book, the discourse representation theoretic framework is extended to deal with plural pronouns. I discussed two especially noteworthy aspects of that theory: (i) e-type anaphora, as in (1), is treated using a principle which is independent of other forms of discourse anaphora, and (ii) the accessibility of dependence information in discourse (as illustrated by (4)) is analysed using a stipulated principle that allows the copying of representations under certain specific circumstances.

Plural anaphora is thus accounted for by DRT by using two principles which are assumed on top of DRT's mechanism for 'standard' anaphora. Representations play a crucial role in these principles. Plural e-type anaphora is analysed using an operation called *abstraction*. This operation allows representational material from a (representation of a) quantificational sentence which is to be used to form a plural antecedent. The accessibility of dependence information, moreover, also boils down to a constrained license to re-use representational material.

In the literature, DRT's emphasis on representations has lead to the question of whether the same empirical coverage could not be achieved without positing a level of representation. Moreover, due to the specific top-down architecture used by DRT to create these representational structures, the need for alternative accounts which were more faithful to the successful (bottom-up) Montague Grammar became clear. With the invention of dynamic predicate logic (Groenendijk and Stokhof 1991), an important step was made to make such alternatives possible. This set off a programme of 'dynamic semantics' which had the non-representational and compositional interpretation of natural language expressions in discourse as its goal. In dynamic semantics, the meaning of an expression is identified as its context-change potential. 'Context', however, is not used here as referring to some sort of representation. Context is a simple model-theoretical object, namely a (set of) assignment function(s).

Given DRT's successful account of plural reference, I found it interesting to see what a non-representational dynamic semantic theory which focuses on the same topic should look like. One specific aim of this dissertation was to discard the two DRT principles discussed above by moving to a dynamic semantic analysis.

The existence of examples like (2), however, formed a complicating factor. Before I started to model plural pronominal reference, I therefore first needed to investigate these examples.

The phenomenon illustrated by (2) was called *complement anaphora*. It has been (and still is) studied in detail in the psycholinguistic research of Linda Moxey, Anthony Sanford and co-authors. Their experiments show that subjects tend to refer to the complement set following a negative (that is, roughly, a downward monotone) quantificational expression. Within the semantic community, the reality of complement anaphora has been doubted several times. The likely root of this doubt is a generalisation made by Kamp and Reyle (1993). According to them, the operation of set-theoretical difference (like the complement set $A - B$) is not an admissible operation for antecedent-formation, in contrast to constructing the union of two sets.

There exists, however, an alternative interpretation of (2) which does justice to Kamp and Reyle's generalisation, namely one in which the pronoun does not refer to the complement set, but rather to the complete restrictor set, A. The second sentence in (2) would then be paraphrasable by "generally speaking, the MPs went to the beach instead."

Moxey and Sanford argue against such an approach. In chapter 3 of this thesis, I followed their argumentation and showed that the idea that complement anaphora do not involve reference to the complement set is indeed very implausible. The existence of pronominal reference to the complement set, however, should not be seen as a clue that complement anaphora is a phenomenon comparable to cases of e-type anaphora as in (1). This became evident from two observations. First of all, whenever a pronoun can both take the reference set and the complement set as its antecedent to yield a consistent interpretation, the resolution to the reference set will 'win'. In other words, complement anaphora is only possible once other forms of anaphora are excluded. Second, in contrast to reference to the complement set, reference set reference is not constrained by the expression anteceding it. Complement anaphora can only occur once the antecedent guarantees that the complement set is non-empty. Such a condition does not apply to the reference set. Monotone decreasing expressions, which do not guarantee the non-emptiness of the reference set, are fine antecedents for pronominal reference set reference.

I concluded from this that the reference set is made *salient* by a quantificational expression. An example like (1) is therefore an ordinary case of pronominal reference, for pronouns refer to individuals or groups of individuals which are contextually highly salient. In order to refer to the complement set, its non-emptiness will first have to be inferred. This necessary inference step makes the complement set principally unfit for pronominal reference. Still, the experiments of Moxey and Sanford did show that apart from a preference for complement set reference following negative expressions, subjects also displayed a rhetorical preference for giving an explanation of the negativity of the antecedent. This effect makes the complement set a tempting antecedent. In conclusion, then, pronominal complement anaphora is a marked form of anaphora which owes its existence to a rhetorical preference of speakers.

This conclusion justified a reduction in the empirical coverage of the dynamic semantics. A model of unmarked pronominal reference is a model that describes the introduction of salient sets in context. Pronominal reference to the complement set, however, is in need of an intervening inference process and, consequently, does not belong to such a model. Chapters 4, 5 and 6 focused on the construction of a theory with this reduced empirical coverage.

Chapter 4 discussed three studies from the literature: the dynamic plural predicate logic of van den Berg (1996b), the semantics for plural anaphora based on parametrised sum individuals from Krifka (1996a) and the theory of anaphoric information from Elworthy (1995). These studies have in common that they all address the dependent type of anaphora exemplified in (4) and, moreover, use a structured notion of context to deal with examples involving dependence phenomena. The chapter showed that the proposals do not differ a lot and that the proposals of Krifka and van den Berg, in particular, are very similar with respect to important (formal) notions like dependence, *in*dependence and distributivity. These notions are responsible for the treatment of dependence phenomena and, as such, an elegant non-representational alternative to the DRT principles.

With respect to the data, I focused mainly on examples like (4) in chapter 4, and concluded that the theory of Elworthy is too strong. It generates a reading for (4) wherein each student submits a paper to L&P irrespective of whether or not this is his or her own paper. The theories of van den Berg and Krifka, however, turned out to be too weak, since they cannot explain how, in the scope of distributivity, pronouns can refer to both sub-individuals in a group and to the original total group itself. The example in (5), for instance, is problematic. The second sentence is ambiguous between a reading in which each student submits the two papers he or she wrote to L&P and (an odd) one wherein each student submits all the written papers to L&P.

(5) Every student wrote exactly two papers. They each submitted them to L&P.

The example in (5) played a crucial role in chapter 5. The question was how anaphora should be represented given that in examples like (5), the same syntactic antecedent ('two papers') antecedes two different resolutions of the pronoun, namely a dependent and an independent one. In theories like those of van den Berg and Krifka, there is one variable which is associated with both readings. But only one of those can be realised at one time. I argued that the choice of the label which is associated with an individual introduced into the context should be contextualised.

This conclusion is not new. It is related to the so-called problem of destructive assignment, which was noticed as being an important shortcoming of dynamic predicate logic. The cause of the problem is the specific definition of existential quantification as random assignment. That is, existential quantification over some variable x rewrites the value assigned to x previously. Consequently, languages like dynamic predicate logic allow values to be overwitten and information to be lost.

As a solution for the destructive assignment problem, I turned to the framework of *incremental dynamics* (van Eijck 2001). In incremental dynamics, the labelling of values is contextualised. This is realised by replacing assignment functions (sets of variable-individual pairs) with 'stacks', sets of position-individual pairs. Since the position of an individual in a stack can change when this stack is combined with another stack, the index which goes with an introduced individual is contextualised. This is because the position of a value is always relative to the stack (i.e. the context) it appears in.

Incremental dynamics is a variation on dynamic predicate logic. In chapter 5, I presented a variation on the dynamic plural predicate logic of van den Berg along the lines of incremental dynamics. This enabled us to devise a distributivity operator which gives access to dependence information without overwriting 'old' antecedents. Using an underspecified representation of the interpretation of examples like (5), we can then show that a pronoun can have two plausible resolutions, completely in accordance with the intuitions given for these sentences.

In chapter 6, I returned to DRT and the principles it has to stipulate in order to show that our non-representational approach may do without such additional principles. I showed that the semantics for distributivity which was developed in chapter 5 can be generalised to a semantics for quantificational determiners, which at the same time captures the introduction of the (maximal) reference set in context and the potential introduction of dependence information. Moreover, this approach was shown to give a straightforward account of the dynamics of monotone decreasing quantifiers. Most importantly, however, all anaphoric effects were lexically governed and, consequently, no additional principles needed to be assumed to account for the range of resolutions of plural pronouns.

Samenvatting

Dit proefschrift heeft als uitgangpunt de observatie dat meervoudige pronominale anafora een bijzonder veelzijdig fenomeen is. Het zou een grove oversimplificatie zijn om de relatie tussen een meervoudig voornaamwoord en haar antecedent te karakteriseren als een coreferentie-relatie. Een meervoudige pronomen kan in discourse namelijk gekwantificeerde en kwantificationele en aldus niet strikt referentiële antecedenten nemen. Het beroemde voorbeeld van Gareth Evans in (1) is een scherpe illustratie dat sommige meervoudige pronomina noch gebonden zijn door, noch coreferen met hun antecedent.

(1) Few senators admire Kennedy; and they are very junior.
"Er zijn weinig senatoren die Kennedy bewonderen; en ze zijn erg jong."

Wat uitgelegd moet worden is hoe in (1) het meervoudige voornaamwoord verwijst naar de verzameling van *alle* senatoren die Kennedy bewonderen. Het pronomen verwijst dus naar een verzameling die is geconstrueerd uit verzamelingen die een rol spelen in de eerste zin, namelijk de doorsnede van de denotatie van de restrictor en de denotatie van het bereik (dat is, het gezegde). Dus, als $(D(A))(B)$ een kwantificationele zin is, waarbij D met de betekenis van een determinator, A met die van de restrictor en B met die van het gezegde correspondeert, dan is $A \cap B$ een potentiële resolutie voor een volgend pronomen. Deze verzameling $A \cap B$ wordt de *referentie-verzameling* genoemd.

Behalve verwijzing naar de referentie-verzameling is er echter een grote variatie aan andere soorten meervoudige pronominale anafora. Naast aandacht voor voorbeelden als (1), concentreer ik me in dit proefschrift op twee van die variaties: (i) pronominale referentie naar het verschil tussen de restrictor-denotatie en de denotatie van het gezegde, de zogenaamde *complement-verzameling* $A - B$, zie bijvoorbeeld (2) en (ii) pronominale verwijzing naar verzamelingen die direct geassocieerd zijn met de referentie-verzameling, zie (3). (Andere soorten meervoudige anafora waar ik aandacht aan besteed, zij het in mindere mate, zijn niet-maximale verwijzing naar de referentie-verzameling en verwijzing naar de verzameling A, de denotatie van de restrictor.)

(2) Slechts weinig parlementariërs woonden de bijeenkomst bij. Ze hadden het veel te druk.

(3) De studenten schreven elk een artikel. Ze waren stuk voor stuk veel te lang.

Naast de vorm van anafora in (3) is er een gebruik van discourse pronomina die sterk aan dit voorbeeld gerelateerd is. De gedetailleerde relatie tussen de studenten en de artikelen, dat is de afhankelijkheid van artikelen van studenten in (3), is namelijk eveneens toegankelijk in discourse.

(4) De studenten schreven elk een artikel. Ze stuurden het bovendien elk naar een tijdschrift.

Het voornaamwoord 'het' in (4) neem het indefiniet 'een artikel' als zijn antecedent. De verwijzing van dit pronomen is echter afhankelijk van de verwijzing van pronomen 'ze', dat het onderwerp vormt. De tweede zin in (4) betekent namelijk dat elke student die een artikel schreef, *zijn* of *haar artikel* naar een tijdschrift stuurde. De koppeling tussen artikelen en studenten in de eerste zin wordt gebruikt door de pronomina in de tweede zin.

Meervoudige anafora blijkt dus een zeer veelzijdig fenomeen te zijn. De belangrijkste doelstelling van dit proefschrift is om die veelzijdigheid uit te drukken in een semantische theorie. De meest volledige bestaande semantische theorie van meervoudige verwijzing is zonder twijfel die in Kamp en Reyle 1993. In dat boek wordt het raamwerk van discourse representatie theorie (DRT) uitgebreid met (onder andere) een behandeling van meervoudige pronomina. Naar mijn mening springen twee aspecten van die theorie er uit: (i) e-type anafora zoals in (1) wordt behandeld met behulp van een principe dat los staat van de analyse van andere vormen van discourse anafora en (ii) de toegankelijkheid van afhankelijkheids-informatie in discourse (zoals te zien in (4)) wordt geanalyseerd met behulp van een gestipuleerd principe dat onder speciale omstandigheden het kopiëren van representaties toe staat.

Meervoudige anafora wordt in DRT dus aangepakt met twee principes die los staan van DRT's analyse van 'standaard' anafora. Bij deze principes speelt het bestaan van een representationeel niveau een cruciale rol. Meervoudige e-type anaphora worden in DRT geanalyseerd met behulp van abstractie. Deze operatie houdt in dat representationeel materiaal van een (representatie van een) kwantificationele zin gebruikt kan worden om tot een meervoudig antecedent te komen. De toegankelijkheid van afhankelijkheids-informatie wordt geanalyseerd door bovendien toe te staan dat dergelijk representationeel materiaal eveneens opgenomen mag worden in nieuwe kwantificationele structuren.

In de literatuur heeft het sterke leunen van DRT op representaties de vraag opgeroepen of hetzelfde empirische bereik niet behaald kan worden zonder gebruik te maken van een representationeel niveau. Aangezien DRT's representaties via een top-down architectuur gecreëerd worden, kwam er bovendien behoefte aan alternatieven die beter passen in de succesvolle traditie van Montague Grammatica. Met de ontwikkeling van dynamische predikaat logica (Groenendijk and Stokhof 1991) werd de belangrijkste stap gezet om dergelijke alternatieven mogelijk te maken en begon een programma van 'dynamische semantiek' dat als doel had de semantiek van natuurlijke taalexpressies te analyseren aan de hand van hun potentieel om de context te veranderen. 'Context' is hier nadrukkelijk geen representatie maar een simpel modeltheoretisch object, namelijk een (verzameling) toekenningsfunctie(s).

Gezien DRT's succes op het gebied van meervoudige verwijzing, is het interessant om te zien hoe een niet-representationele dynamische semantiek die zich op hetzelfde onderwerp richt eruit moet zien. Een specifiek doel van het proefschrift is aldus om de twee representationele DRT principes die ik hierboven besprak overbodig te maken door een niet-representationele dynamische analyse aan te gaan.

Het bestaan van voorbeelden als in (2) blijkt echter een bemoeilijkende factor. Voordat we aan de modellering van meervoudige verwijzing beginnen, moeten we dus eerst hieraan onze aandacht wijten.

Het fenomeen dat met het voorbeeld in (2) geïllustreerd wordt, wordt *complement-anafora* genoemd. Dit fenomeen werd en wordt uitgebreid bestudeerd in het psycholinguistische onderzoek van Linda Moxey, Anthony Sanford en co-auteurs. Hun experimenten laten zien dat proefpersonen een voorkeur hebben om te refereren aan de complement-verzameling na een negatieve (dat is, ruwweg, een monotoon dalende) kwantificationele expressie. Binnen de semantische gemeenschap is een aantal keer getwijfeld aan de realiteit van complement-anafora. Ten grondslag aan die twijfel ligt een generalisatie die in Kamp en Reyle 1993 gedaan wordt. Volgens Kamp en Reyle is het vormen van een verzamelingtheoretisch verschil (zoals in de complement-verzameling $A - B$) niet een geoorloofde methode om (meervoudige) antecedenten te vormen, dit in scherp kontrast met het bij elkaar voegen van twee verzamelingen.

Toch, een alternatieve interpretatie van voorbeelden zoals (2) die wel voldoet aan Kamp en Reyle's generalisatie zou er een zijn waarin het pronomen niet naar de complement-verzameling verwijst maar naar de gehele restrictorverzameling A. De tweede zin in (2) betekent dan zoiets als "De parlementariërs hadden het *over het algemeen* te druk." Moxey en Sanford argumenteren zelf tegen zo een analyse. In hoofdstuk 3 van dit proefschrift volg ik hun argumentatie en toon ik aan dat het idee dat complement-anafora geen verwijzing naar de complement-verzameling in-houdt implausibel is. Het bestaan van pronominale verwijzing naar de complement-verzameling moet echter niet gezien worden als aanwijzing dat complement-anafora eenzelfde soort fenomeen is als het typische geval van e-type anafora in (1). Dit blijkt uit twee dingen: (i) wanneer in een contekst een pronomen zowel naar de referentie-verzameling als naar de complement-verzameling zou kunnen verwijzen en allebei die verwijzingen een consistente interpretatie opleveren, dan 'wint' de referentie-verzameling, (ii) in tegenstelling tot verwijzing naar de complement-verzameling, stelt verwijzing naar de referentie-verzameling geen eisen aan het antecedent. Wat (i) laat zien is dat complement-anafora alleen mogelijk is op het moment dat meer gangbare soorten anafora *on*mogelijk zijn. Verder kunnen complement-anafora alleen voorkomen als het antecedent zodanig is dat het de garantie geeft dat de complement-verzameling niet leeg is. Voor verwijzing naar de referentie-verzameling is er niet zo een dergelijke voorwaarde; ook monotoon dalende expressies (die niet garanderen dat de referentie-verzameling elementen bevat) kunnen als antecedent dienen.

Ik concludeer hieruit dat de referentie-verzameling een *saillant* antecedent is en dat een voorbeeld als (1) aldus een gangbaar geval van pronominale verwijzing is. Voornaamwoorden verwijzen immers naar individuen of groepen individuen die in de contekst in hoge mate geactiveerd zijn. Het bestaan van een complement-verzameling moet echter eerst geïnfereerd worden wil er naar verwezen kunnen worden. Dat maakt de complement-verzameling in principe ongeschikt als antecedent voor voornaamwoorden. Echter, de experimenten van Moxey en Sanford toonden niet alleen aan dat proefpersonen in negatieve conteksten naar de complement-verzameling willen verwijzen, maar ook dat die proefpersonen een retorische voorkeur laten zien om de negativiteit van het antecedent te verklaren. Dit zorgt ervoor dat de complement-verzameling een verleidelijke antecedent is. Complement-anafora is aldus een gemarkeerde variant van anafora die haar bestaan te danken heeft aan retorische voorkeuren van sprekers.

Deze conclusie rechtvaardigt een verkleining van het empirische bereik van de dynamische semantiek die we willen construeren. Een model van ongemarkeerde verwijzing door pronomina is een model dat de introductie van saillante verzamelingen in contekst beschrijft. Pronominale verwijzing naar de complement-verzameling, daarentegen, gebeurt met tussenkomst van een inferentie-proces en hoort zodanig niet tot een dergelijk model. Hoofdstuk 4, 5 en 6 richten zich op de constructie van een dynamische semantiek met een bereik dat zich beperkt tot saillante verzamelingen.

Hoofdstuk 4 behandelt drie studies uit de literatuur, namelijk de dynamische meervouds-predikaatlogica uit van den Berg 1996b, de semantiek gebaseerd op geparametriseerde pluraliteiten uit Krifka 1996a en de theorie van anaforische informatie uit Elworthy 1995. Deze drie voorstellen hebben met elkaar gemeen dat ze alle drie een gestructureerde datastructuur gebruiken als context. Het hoofdstuk laat zien dat de verschillen tussen deze systemen niet erg groot zijn en dat met name de theorieën van van den Berg en van Krifka erg op elkaar lijken wat betreft belangrijke (formele) noties als afhankelijkheid, onafhankelijkheid en distributiviteit. Dergelijke noties zijn verantwoordelijk voor de behandeling van afhankelijkheidsfenomenen en vormen op die manier een elegant alternatief op de DRT principes.

Wat betreft empirisch bereik wordt er in het vierde hoofdstuk vooral aandacht besteed aan voorbeelden als (4), hierboven. Er wordt beargumenteerd dat Elworthy's theorie te sterk is. Het genereert bijvoorbeeld een lezing voor (4) waarin elke student een paper naar een tijdschrift stuurt ongeacht of het om zijn of haar eigen paper gaat. De theorieën van van den Berg en Krifka, daarentegen, zijn te zwak in de zin dat ze niet kunnen verklaren hoe de verwijzing van een pronomen zowel naar subindividuen van een groep als naar de oorspronkelijke totale groep kan verwijzen. Problematisch is bijvoorbeeld (5). Dit voorbeeld is ambigue tussen een lezing waarin elke student de twee artikelen die hij of zij geschreven heeft aan een tijdschrift stuurt en een lezing waarin elke student alle geschreven artikelen naar een tijdschrift stuurt.

(5) De studenten hebben elk precies twee artikelen geschreven. Elk van hen heeft ze vervolgens aan een tijdschrift aangeboden.

Het zojuist omschreven probleem voor de theorieën van van den Berg en Krifka speelt vervolgens een centrale rol in hoofdstuk vijf. Hier wordt de vraag gesteld hoe anafora gerepresenteerd dient te worden gegeven het feit dat hetzelfde meervoudige antecedent zowel bij een afhankelijke als bij een onafhankelijke vorm van referentie betrokken kan zijn. Er wordt beargumenteerd dat een (mogelijk meervoudig) individu dat in de context geïntroduceerd is aan een label (variabele, index) toegekend moet worden dat afhankelijk is van de context zelf.

Een dergelijke conclusie is niet nieuw. Het relateert aan het zogenaamde probleem van destructieve toekenning, dat met name bij dynamische predikaat logica als een problematische complicatie is opgemerkt. Dit probleem ontstaat door de interpretatie van de existentiële kwantor als de willekeurige toekenning van een waarde aan een variabele. Bij deze toekenning wordt de 'oude' waarde van die variabele overschreven. Talen als dynamische predikaat logica staan in principe dus toe dat bestaande waardes overschreven worden en dat er verlies van informatie optreedt.

Als oplossing voor de toekenningsproblematiek wordt in hoofdstuk vijf het raamwerk van incrementele dynamiek (van Eijck 2001) geïntroduceerd. In incrementele dynamiek (ID) is de indexering van waardes gecontekstualiseerd. Dit wordt bewerkstelligd door toekenningsfuncties (verzamelingen van paren variabelen en waardes) te vervangen door zogenaamde stacks ('stapeltjes'), dat is, verzamelingen van paren posities en waardes. Aangezien de positie die een individu inneemt kan veranderen als we twee stapeltjes combineren, is daarmee de indexering van waardes gecontekstualiseerd. De positie van een waarde is immers altijd relatief aan de stack (lees, contekst) waarin de waarde zich bevindt.

Incrementele dynamiek is een variant op dynamische predikaat logika. In hoofdstuk 5 presenteer ik een variant van de dynamische meervouds predikaatlogika van van den Berg, geheel volgens de ideeën van incrementele dynamiek. Dit stelt ons in staat om een distributiviteits-operator te definiëren die afhankelijkheids-informatie toegankelijk maakt zonder 'oude' antecedenten te overschrijven. Aan de hand van een ondergespecificeerde representatie van de interpretatie van voorbeelden als (5), kunnen we vervolgens zien dat er twee plausibele resoluties voor het pronomen mogelijk zijn, geheel volgens de intuities voor deze voorbeelden.

In het zesde hoofdstuk keer ik terug naar DRT en de in die theorie gestipuleerde principes voor meervoudige anaforische referentie en laat ik zien dat onze niet-representationele aanpak zonder dergelijke principes kan. Er wordt aangetoond dat de semantiek voor distributiviteit die in hoofdstuk 5 is geïntroduceerd gegeneraliseerd kan worden om zo tot een semantiek voor kwantificationele determinatoren te komen die aan twee eisen voldoet. Ten eerste bevat een contekst na interpretatie van een kwantificationele zin de (maximale) referentie-verzameling. Ten tweede bevat een dergelijke contekst de afhankelijkheidsinformatie van die zin. Bovendien wordt aangetoond dat deze aanpak op eenvoudige wijze het dynamische karakter van monotoon dalende kwantoren kan beschrijven. Belangrijker echter is het feit dat alle anaforische effecten lexicaal gestuurd zijn. Als gevolg daarvan is er geen extra principe nodig dat de verscheidenheid aan resoluties voor meervoudige pronomina verklaart.

References

Ariel, M. (1990). *Accessing Noun-Phrase Antecedents*. London; New York: Routledge.

Asher, N. (1993). *Reference to Abstract Objects in Discourse*. Dordrecht: Kluwer.

Asher, N. and A. Lascarides (1999). Bridging. *Journal of Semantics 15*, 83–113.

Barwise, J. (1987). Noun Phrases, Generalized Quantifiers, and Anaphora. In *Generalized Quantifiers, Linguistic and Logical Approaches*, pp. 237–268. Dordrecht: Reidel.

Barwise, J. and R. Cooper (1981). Generalized quantifiers and natural language. *Linguistics and Philosophy 4*(2), 159–219.

Beaver, D. (1999). The logic of anaphora resolution. In P. Dekker (Ed.), *Proceedings of the Twelfth Amsterdam Colloquium*, Amsterdam, pp. 61–66. ILLC Publications.

Bekki, D. (2000a). *Typed Dynamic Logic for Compositional Grammar*. Ph. D. thesis, The University of Tokyo.

Bekki, D. (2000b). Typed dynamic logic for e-type link. In *Third International Conference on Discourse Anaphora and Anaphor Resolution (DAARC2000)*, Lancaster University, U.K., pp. 39–48.

Bennet, M. (1974). *Some extensions of a Montague fragment of English*. Ph. D. thesis, UCLA, Los Angeles.

Bos, J., E. Mastenbroek, S. McGlashan, S. Millies, and M. Pinkal (1994). A compositional DRS-based formalism for NLP-applications. In *Proceedings of the international workshop on computational semantics*, Tilburg, pp. 21–31.

Brisson, C. (1997). On definite plural NPs and the meaning of *all*. In A. Lawson (Ed.), *SALT VII*, Ithaca, New York: Cornell University, pp. 55–72.

Carminati, M., L. Frazier, and K. Rayner (2002). Bound variables and C-Command. *Journal of Semantics 19*, 1–34.

Chierchia, G. (1992). Anaphora and dynamic binding. *Linguistics and Philosophy 15*, 111–183.

Church, A. (1941). *The calculi of Lambda Conversion*. Princeton: Princeton University Press.

Clark, H. (1977). Bridging. In P. Johnson-Laird and P. Wason (Eds.), *Thinking: readings in cognitive science*, pp. 411–420. Cambridge university press.

Cooper, R. (1979). The interpretation of pronouns. In F. Heny and H.Schnelle (Eds.), *Syntax and Semantics 10*. New York: Academic Press.

Corbett, A. and F. Chang (1983). Pronoun disambiguating: accessing potential antecedents. *Memory and Cognition 11*, 283–294.

Corblin, F. (1996a, avril). Peut-on anaphoriser le complémentaire d'un ensemble? In J. Moeschler and M. Béguelin (Eds.), *Référence temporelle et nominale. Actes du 3e cycle romand de Sciences du langage*, Cluny.

Corblin, F. (1996b). Quantification et anaphore discursive: la référence aux complémentaires. *Langages 123*, 51–74.

de Bruijn, N. G. (1972). Lambda calculus notation with nameless dummies, a tool for automatic formula manipulation, with application to the church-rosser theorem. *Indagationes Mathematicae 34*, 381–392.

de Hoop, H. (1992). *Case Configuration and noun phrase interpretation*. Ph. D. thesis, University of Groningen.

de Hoop, H. (2001). The Problem of Unintelligibility in OT Semantics. In G. van der Meer and A. ter Meulen (Eds.), *Making Sense: from Lexeme to Discourse*, Volume 44 of *GAGL*. Groningen: CLCG.

Dekker, P. (1993). *Transsentential Mediations*. Ph. D. thesis, ILLC, Department of philosophy, University of Amsterdam.

Dekker, P. (1994). Predicate Logic with Anaphora. In L. Santelmann and M. Harvey (Eds.), *Proceedings of the Fourth Semantics and Linguistic Theory Conference*, pp. 17. DMLL publications, Cornell University.

Dowty, D. (1987). A note on collective predicates, distributive predicates, and 'all'. In F. Marshall (Ed.), *Proceedings of the Third Eastern States Conference on Linguistics (ESCOL 86)*, Columbus, OH, pp. 97–115. The Ohio State University.

Elworthy, D. (1992, october). *The semantics of noun phrase anaphora*. Ph. D. thesis, Darwin College, University of Cambridge.

Elworthy, D. (1995). A theory of anaphoric information. *Linguistics and Philosophy 18*, 297–332.

Fernando, T. (1992). Transition systems and dynamic semantics. In D. Pearce and G. Wagner (Eds.), *Logics in AI, LNCS 633 (subseries LNAI)*. Berlin: Springer-Verlag.

Frege, G. (1884). *Die Grundlagen der Arithmetik. Eine logisch-mathematische Untersuchung über den Begriff der Zahl*. Breslau: Köbner. Reprint (1961) by Georg Olms, Hildesheim.

Geurts, B. (1997). Book review of Linda M. Moxey and Anthony J. Sanford. Communicating Quantities. 1993. *Journal of semantics 14*(1), 87–94.

Geurts, B. (1999). Quantifying on empty. Unpublished manuscript, Department of philosophy, University of Nijmegen.

Gillon, B. (1987). The readings of plural noun phrases in english. *Linguistics and Philosophy 10*, 199–219.

Groenendijk, J. and M. Stokhof (1990). Dynamic Montague Grammar. In L. Kálmán and L. Pólos (Eds.), *Papers from the 2nd Symposium on Logic and Language*. Budapest: Akadémiái Kiadó.

Groenendijk, J. and M. Stokhof (1991). Dynamic Predicate Logic. *Linguistics and Philosophy 14*, 39–100.

Groenendijk, J. and M. Stokhof (1996). Changing the context: dynamics and discourse. In *IATL 3: Proceedings of the 11th annual conference and of the workshop on discourse*, Jerusalem, pp. 104–128. Israel Association for Theoretical Linguistics.

Gundel, J., N. Hedberg, and R. Zacharski (1993). Cognitive status and the form of referring expressions in discourse. *Language 69*, 274–307.

Hausser, R. (1974). *Quantification in an extended Montague grammar*. Ph. D. thesis, University of Texas, Austin.

Heim, I. (1982). *The Semantics of Definite and Indefinite Noun Phrases*. Ph. D. thesis, University of Massachusetts at Amherst.

Heim, I. (1990). E-type pronouns and donkey anaphora. *Linguistics and Philosophy 13*(2), 137–138.

Heim, I. (1993). Anaphora and semantic interpretation: a reinterpretation of reinhart's approach. SfS-Report-07-93, University of Tübingen.

Heim, I., H. Lasnik, and R. May (1991a, Winter). Reciprocity and plurality. *Linguistic Inquiry 22*(1), 63–101.

Heim, I., H. Lasnik, and R. May (1991b, Winter). Reply: On 'reciprocal scope'. *Linguistic Inquiry 22*(1), 173–192.

Hendriks, H. and P. Dekker (1995). Links without locations: information packaging and non-monotone anaphora. In P. Dekker and M. Stokhof (Eds.), *Proceedings of the 10th Amsterdam colloquium*, pp. 339–358. Universiteit van Amersterdam.

Hendriks, P. and H. de Hoop (2001, February). Optimality Theoretic Semantics. *Linguistics and Philosophy 24*(1), 1–32.

Kadmon, N. (1990). Uniqueness. *Linguistics and Philosophy 13*, 273–324.

Kamp, H. (1981). A theory of truth and semantic representation. In J. A. G. Groenendijk, T. M. V. Janssen, and M. J. B. Stokhof (Eds.), *Formal Methods in the Study of Language*. Amsterdam: Mathematical Centre.

Kamp, H. and U. Reyle (1993). *From Discourse to Logic*. Dordrecht: D. Reidel.

Kanazawa, M. (1994). Weak vs. strong readings of donkey sentences and monotonicity inference in a dynamic setting. *Linguistics and Philosophy 17*, 109–158.

Keenan, E. L. (1987). A semantic definition of "indefinite NP". In E. Reuland and A. ter Meulen (Eds.), *The representation of (in)definiteness*, Chapter 12, pp. 286–317. Cambridge, MA: MIT Press.

Kibble, R. (1997a). Complement anaphora and dynamic binding. In A. Lawson (Ed.), *SALT VII*, Ithaca, NY: Cornell University.

Kibble, R. (1997b). Complement anaphora and monotonicity. In G. Morrill, G.-J. Kruijff, and R. Oehrle (Eds.), *Formal Grammar*, pp. 125–136.

Kohlhase, M., S. Kuschert, and M. Pinkal (1996). A type-theoretic semantics for λ-DRT. In P. Dekker and M. Stokhof (Eds.), *Proceedings of the 10th Amsterdam Colloquium*, Amsterdam, pp. 479–498. ILLC.

Krahmer, E. and R. Muskens (1995). Negation and Disjunction in discourse representation theory. *Journal of Semantics 12*, 357–376.

Krahmer, E. and K. van Deemter (1998). Anaphoric noun phrases: toward a full understanding of partial matches. *Journal of semantics 15*, 355–392.

Kratzer, A. (1986). Conditionals. *Chicago Linguistic Society 22*, 1–15.

Krifka, M. (1991). How to get rid of groups, using DRT: A case for discourse-oriented semantics. *Texas Linguistic Forum 32*, 71–110.

Krifka, M. (1996a). Parametrized sum individuals for plural reference and partitive quantification. *Linguistics and Philosophy 19*, 555–598.

Krifka, M. (1996b). Pragmatic strengthening in donkey sentences and plural predications. In J. Spence (Ed.), *Proceedings of Semantics and Linguistic Theory 6*, Cornell University, pp. 136–153.

Kuschert, S. (1995). Eine Erweiterung des λ-Kalküls um Diskursrepresentationsstrukturen. Master's thesis, Universität des Saarlandes.

Landman, F. (1989). Groups. Part I, II. *Linguistics and Philosophy 12*, 559–605;723–744.

Lappin, S. (1989). Donkey pronouns unbound. *Theoretical Linguistics 15*, 263–289.

Lasersohn, P. (1995). *Plurality, conjunction and events*, Volume 55 of *Studies in linguistics and philosophy*. Kluwer academic publishers.

Le Fournier, J. (1967). *De Selby - l'Énigme de l'Occident*.

Lewis, D. (1975). Adverbs of Quantification. In E. L. Keenan (Ed.), *Formal semantics of natural language*, pp. 3–15. Cambridge: Cambridge University Press.

Link, G. (1983). The logical analysis of plurals and mass terms: A lattice- theoretical approach. In R. Bäuerle, C. Schwarze, and A. von Stechow (Eds.), *Meaning, Use and Interpretation of Language*, pp. 302–323. Berlin: Walter de Gruyter.

Link, G. (1987). Generalised quantifiers and plurals. In P. Gårdenfors (Ed.), *Generalised Quantifiers: Linguistic and Logical Approaches*, Volume 31 of *Studies in Linguistics and Philosophy*, pp. 151–180. Dordrecht: D. Reidel Publishing Company.

Link, G. (1991). Plural. In D. Wunderlich and A. Von Stechow (Eds.), *Semantik. Ein internationales Handbuch der zeitgenoessischen Forschung.*, pp. Chapter 2. 418–440. de Gruyter. Berlin.

Lønning, J. T. (1991). Among readings. In J. van der Does (Ed.), *Quantification and anaphora*, Volume R2.2b, pp. 37–52. Edinburgh: Dyana deliverable.

Mauner, G. (1996). *The role of implicit arguments in sentence processing*. Ph. D. thesis, University of Rochester.

Milsark, G. (1974). *Existential Sentences in English*. Ph. D. thesis, MIT, Cambridge, MA.

Moltmann, F. (1996). Domain-related dynamic semantics and the weak-strong distinction among quantifiers. ms. CUNY Graduate center.

Montague, R. (1974). *Formal Philosophy: Selected papers of Richard Montague,* Edited by R.H. Thomason. Yale University Press.

Moxey, L. and A. Sanford (1987). Quantifiers and focus. *Journal of semantics 5*, 189–206.

Muskens, R. (1994). A compositional discourse representation theory. In P. Dekker and M. Stokhof (Eds.), *Proceedings of the 9th Amsterdam Colloquium*, Amsterdam, pp. 467–486. ILLC.

Muskens, R. (1996). Combining Montague semantics and discourse representation. *Linguistics and Philosophy 19*, 147–183.

Neale, S. (1988). *Descriptions*. Ph. D. thesis, Stanford University.

Neale, S. (1990). *Descriptions*. Cambridge, MA: MIT Press.

Nouwen, R. (2003). Complement anaphora and interpretation. *Journal of Semantics 20*(1), 73–113.

Poesio, M. and S. Zucchi (1992). On telescoping. In *Proceedings of SALT 2*.

Prince, A. and P. Smolensky (1997). Optimality: From neural networks to universal grammar. *Science 275*, 1604–1610.

Quine, W. (1960). Variables explained away. In *Proceedings of the American Philosophical Society*, Volume 104:3, pp. 343–347.

Reinhart, T. (2000). Strategies of anaphora resolution. In H. Bennis, M. Everaert, and E. Reuland (Eds.), *Interfact Stragegies*, pp. 295–324. North Holland Amsterdam: Royal Academy of Arts and Sciences, KNAW.

Roberts, C. (1987). *Modal Subordination, anaphora and distributivity*. Ph. D. thesis, University of Massachussets, Amherst.

Rooth, M. (1987). Noun phrase interpretation in Montague grammar, file change semantics and situation semantics. In P. Gärdenfors (Ed.), *Generalized Quantifiers*, pp. 237–269. Dordrecht: Reidel.

Sanford, T. and L. Moxey (1993). *Communicating quantities. A psychological perspective*. Laurence Erlbaum Associates.

Sanford, T., L. Moxey, and Paterson (1994). Psychological studies of quantifiers. *Journal of semantics 10*, 153–170.

Scha, R. (1981). Distributive, collective, and cumulative quantification. In J. Groenendijk, T. Janssen, and M. Stokhof (Eds.), *Formal methods in the study of language*, pp. 483–512. Mathematical Centre.

Schwarzschild, R. (1990). Against groups. In M. Stokhof and L. Torenvliet (Eds.), *Proceedings of the seventh Amsterdam colloquium*, Amsterdam, pp. 475–493. ITLI.

Schwarzschild, R. (1991). *On the meaning of definite plural noun phrases*. Ph. D. thesis, University of Massachussets, Amherst.

Schwarzschild, R. (1992). Types of plural individuals. *Linguistics and Philosophy 15*, 641–675.

Schwarzschild, R. (1996). *Pluralities*, Volume 61 of *Studies in Linguistics and Philosophy*. Dordrecht: Kluwer academic publishers.

Szabolcsi, A. (1997). Strategies for scope taking. In A. Szabolsci (Ed.), *Ways of scope taking*, pp. 109–154. Kluwer Academic Publishers.

ter Meulen, A. (1999). Binding implicit arguments. In P. Dekker (Ed.), *Proceedings of the 13th Amsterdam Colloquium*, Amsterdam. ILLC.

van Benthem, J. (1987). Meaning: interpretation and inference. *Synthese 73*, 451–470.

van den Berg, M. (1990). A dynamic predicate logic for plurals. In M. Stokhof and L. Torenvliet (Eds.), *Proceedings of the Amsterdam Colloquium*, ILLC, UvA, Amsterdam.

van den Berg, M. (1993). Full dynamic plural logic.

van den Berg, M. (1996a). Dynamic generalised quantifiers. In J. van der Does and J. van Eijck (Eds.), *Quantifiers, logic and language*, pp. 63–94. Stanford: CSLI.

van den Berg, M. (1996b). *Some aspects of the internal structure of discourse: the dynamics of nominal anaphora*. Ph. D. thesis, ILLC, Universiteit van Amsterdam.

van der Does, J. (1992). *Applied Quantifier Logics. Collective, naked infinitives*. Ph. D. thesis, University of Amsterdam, Amsterdam.

van der Does, J. (1993). The dynamics of sophisticated laziness. ms. ILLC, Department of Philosophy, University of Amsterdam.

van Eijck, J. (2000a). Context semantics for NL. Unpublished manuscript.

van Eijck, J. (2000b). On the proper treatment of context. In *Proceedings of CLIN99*, Utrecht.

van Eijck, J. (2001, Summer). Incremental dynamics. *Journal of Logic Language and Information 10*(3), 319–351.

van Rooy, R. (1997). Desciptive pronouns in dynamic semantics. In *Proceedings of the 11th Amsterdam Colloquium*, ILLC.

Veltman, F. (1991). Defaults in update semantics. In H. Kamp (Ed.), *Conditionals, defaults and belief revision*, Volume R2.5A. Edinburgh: Dyana Deliverable.

Verkuyl, H. and J. van der Does (1996). The semantics of Plural Noun Phrases. In J. van der Does and J. van Eijck (Eds.), *Quantifiers, Logic and Language*, Volume 54 of *CSLI Lecture Notes*, pp. 337–374. Amsterdam: CSLI Publications.

Vermeulen, K. (1993). Sequence semantics for dynamic predicate logic. *Journal of Logic, Language and Information 2*, 217–254.

Vermeulen, K. (1994). *Explorations of the dynamic environment*. Ph. D. thesis, Onderzoeksinstituut voor Taal en Spraak, Utrecht.

Visser, A. (1998). Contexts in Dynamic Predicate Logic. *Journal of Logic, Language and Information* 7(1), 21–52.

Westerståhl, D. (1984). Determiners and context sets. In J. van Benthem and A. ter Meulen (Eds.), *Generalized Quantifiers in Natural language*, pp. 45–71. Foris Dordrecht.

Winter, Y. (1998). *Flexible boolean semantics: coordination, plurality and scope in natural language*. Ph. D. thesis, UiL-OTS, Universiteit Utrecht.

Zeevat, H. (1989). A compositional approach to DRT. *Linguistics and Philosophy 12*, 95–131.

Zwarts, F. (1996). Facets of negation. In J. van der Does and J. van Eijck (Eds.), *Quantifiers, logic and language*, pp. 385–421. Stanford: CSLI.

Name Index

Subject Index

Curriculum Vitae

Rick Nouwen was born in Heerlen on the 12th of September, 1974. He completed grammar school ('atheneum' at the Bernardinus College in Heerlen) in 1992 and then went to study knowledge technology ('kennistechnologie') in Hasselt (Belgium) in 1992/93. After this, he studied cognitive artificial intelligence ('cognitieve kunstmatige intelligentie') at the faculty of philosophy at Utrecht University. He obtained his masters degree ('doctoraal') in 1999.

From 1999 until 2003, Rick held a position as an AiO (PhD researcher) at the Utrecht Institute of Linguistics OTS. The present dissertation is the result of the work he did there.