Simultaneous Abstraction and Semantic Theories

Peter Ruhrberg

Declaration

This thesis has been composed by myself and it has not been submitted in any previous application for a degree. The work reported within was executed by myself, unless otherwise stated.

March 1996
Acknowledgements

My first, and foremost thank goes to Robin Cooper. His influence on my work is evident for almost every part of this thesis. His use of Aczel and Lunnon’s system of simultaneous abstraction as a framework for semantic theories, in particular for his Situation-Theoretic Discourse Representation Theory, has been largely my inspiration for the work presented here. Without his interest and support I would not have carried out this work, and neither would I have had the opportunity to do so, for financing a PhD abroad is no trivial matter, even for citizens of a wealthy, allegedly Euro-enthusiast country, like Germany.

My second “thank you” goes to the people at the Centre for Cognitive Science in Edinburgh. I feel privileged to have been one of them for so long, and I am continually amazed by their level of enthusiasm and commitment. Measured by such standards, one can only hope not to disappoint too severely.

Furthermore, I am grateful for the interest and helpful comments I received for this work, from Peter Aczel, Lawrence Cavendish, Susanna Kuschert, Rachel Lunnon, David Milward, Massimo Poesio, Paul Schweizer, Colin Stirling, Henk Zeevat, the members of the MsC working group at the CCS, the people at the IMS in Stuttgart and the Computerlinguistik in Saarbrücken, as well as my examiners Lex Holt and Hans Kamp. I thank my various office mates over the years, in particular Isatomi Fujinami and Matthias Paul. For financial support I thank the DAAD, ESRC, ESPRIT BR 8852 (DYANA-2, special thanks to Marc Moens), and LRE 62-651 (PmCaS).

Abstract

I present a simple Simultaneous Abstraction Calculus, where the familiar λ abstraction over single variables is replaced by abstraction over whole sets of them. Terms are applied to partial assignments of objects to variables. Variants of the system are investigated and compared, with respect to their semantic and proof-theoretic properties.

The system overcomes the strict ordering requirements of the standard λ calculus, and is shown to provide the kind of ‘non-selective’ binding needed for Dynamic Montague Grammar and Discourse Representation Theory. It is closely related to a more complex system, due to Peter Aczel and Rachel Lunnon, and can be used for Situation Theory in a similar way. I present versions of these theories within an axiomatic, proof-theoretic framework, based on Aczel Frege Structures.

The aim of this work is to provide the means for integrating various semantic theories within a formal framework, so that they can share what is common between them, and adopt from each other what is compatible with them.

1Earlier versions of parts of this work have been previously published by the author under the title A Simultaneous Abstraction Calculus and Theories of Semantics, in [Cooper & Groenendijk 1994] and slightly updated in [Seligman et al. 1995]. The complete draft, before corrections, has also been published under the current title, in [Cooper et al. 1999].
Contents

Acknowledgements ii
Abstract iii

1 Introduction 1
1.1 Basic Assumptions 1
1.2 Why Simultaneous Abstraction? 4
1.2.1 Variables: Bound yet Free 8
1.2.2 Axiomatic Semantics 9
1.2.3 Bealer’s Puzzle 10
1.3 Other Frameworks 11

2 Simultaneous Abstraction 13
2.1 Simple Simultaneous Abstraction 13
2.1.1 Syntax for A 13
2.1.2 Semantics for A 15
2.1.3 Basic Properties of A 17
2.1.4 Partial Application 20
2.2 Aczel-Lennon Abstraction 21
2.2.1 Relating A, AAL, and λ 23
2.2.2 Proof Theory for AAL 25
2.2.3 The Church-Rosser Theorem 27
2.3 Combinatory Aczel-Lennon Logic 31
2.3.1 Completeness for CAL 33
2.3.2 Abstraction in CAL-models 34

2.3.3 Completeness for AAL 37
2.4 Structured Objects and Systems of Equations 39

3 Semantic Theories 41
3.1 Frege Structures 41
3.1.1 Consistency 46
3.2 Dynamic Montague Grammar 48
3.2.1 Some DMG Translations 50
3.3 Discourse Representation Theory 53
3.3.1 Zeevat’s DRT 53
3.3.2 w-Properties 55
3.3.3 Kamp Structures 56
3.3.4 Consistency 58
3.4 Situation Theory 60
3.4.1 States of Affairs 61
3.4.2 Barwise Structures 63
3.4.3 Consistency 64

4 Further Directions 66
4.1 True Dynamics 66
4.2 A Typed System 68
4.3 A-DRT 72
4.3.1 Consistency 77
4.3.2 Some A-DRT Translations 78
4.4 A Note on Appropriateness and Restrictions 80
4.5 Conclusion 81

References 83

Index 86
Chapter 1

Introduction

Linguistic semantics seems to have become a true Babel of newly invented and ever-changing logical formalisms, where much research is dedicated to the business of translating between them. I will add another two or three formalisms to the list. Yet my aim is not to increase the confusion, but rather to reduce it, by giving not so much a particular semantic formalism but a formal framework in which various semantic theories can be cast, in a way that allows them to share explicitly what they have in common. My faint hope is to regain some of the unity in semantics that has been lost over the last decade or so, during which the authority of Montague's Higher Order Intensional Logic (IL) was so badly eroded. The proposed systems contain no really new ideas, but rather steal elements of existing formalisms and put them together so that their common roots are clearly exposed. The most prominent common feature of the theories discussed in this work is their use of some form of simultaneous abstraction. I will give (variations of) a formal system of simultaneous abstraction and application, as the backbone for a set of theories which are obtained by the addition of axioms of truth, following the lead of Property Theory.

1.1 Basic Assumptions

Let me briefly outline the underlying philosophical assumptions of this work. I won't attempt to defend them here, as that would be another thesis.

---

One might say that meanings are fixed by the interpretation \( | \cdot | \) in a particular A-model, while matters of fact are reflected in the extension \( T \) of the truth predicate in a Frege Structure, based on that A-model.
only on the basis of these. The most prominent theory which follows these course
grained lines, by identifying propositions with sets of logically possible worlds, has
long been known to suffer from fundamental inadequacies in the analysis of
propositional (and other) attitudes. I take it as a guiding principle of semantics
that propositions and other contents of utterances must be plausible candidates
for being the objects of our attitudes. The content of an assertion is something
a sincere and competent speaker believes, and wishes to convey to his audience,
thus hopefully contributing to the spread of knowledge in his community.

The links between the notions of content, propositions, truth, and attitudes,
as I have outlined them, are largely subscribed to by the particular semantic
theories I will discuss in this work, or they come at least close to it, if we push
them a little in this direction.

Montague Grammar (MG) for example, [Montague 1974], in most of its forms
subscribes to the identification of propositions with sets of possible worlds. These
are the objects of propositional attitudes, and they are truth bearers in the sense
that they may or may not contain the actual world. Versions of MG which
do not make this identification, and operate along property theoretic lines in
stead, have been proposed, for example in [Thomason 1980] and the more recent
[Chierchia 1994]. The latter work will concern me here in particular, as it presents
a dynamic version of MG. It follows Dynamic Montague Grammar (DMG) of
[Greendijk & Stokhof 1990], but takes intensionality at face value and appears
less baroque in its treatment of dynamic binding.

Discourse Representation Theory (DRT), [Kamp 1981], appears to regard
propositions as rather syntactic objects, the Discourse Representation Structures
(DRSs), whose truth is a matter of anchoring its Discourse Referents with respect
to an appropriate model. I propose to regard DRSs as particular kinds of rela-
tions, which are easily turned into propositions by existential quantification, or
by application (anchoring) to individuals. They are not to be confused with their
notations, although one may argue that they are almost as highly structured as
the latter. I believe this view can make sense of the kind of theory of attitudes
envisioned in [Kamp 1990], at least to some extent, without identifying the ob-
jects of attitudes with syntactic structures. I have no quarrel with a syntactic
theory of mind as such, but I do not think it is embodied in the meaning of our
intensional idiom.

Situation Theory (ST) [Barwise & Perry 1983] can be taken to promote the
idea that propositions have two ingredients: a situation and a state of affairs
(infon, fact). Truth depends upon the ‘support’ of the state of affairs by the
situation, and the actuality of the situation, if abstract ‘possible’ situations are
admitted into the ontology. Not all versions of ST take such a narrow view of
propositions, see [Barwise & Cooper 1991], and neither do I. With regards
to attitudes ST tends to stress the role of situations, yet it is generally agreed,
following the critique of ST in [Sommers 1985], that they do not allow us to dispense
with a highly structured view of propositions and states of affairs.

I stand in some distance to theories of ‘dynamic’ semantics that stress the
changes in information states brought about by the processing of sentences. There
can be no doubt that such changes are of utmost importance, and that a theory
of meaning worth its money must say something insightful about such matters,
but this does not, on my view, relieve us from coming up with a serious theory
of propositional contents as relata of attitudes. It seems to me to be in fact a
precondition for any realistic dynamic theory. DMG, DRT, and ST all seem to
conform fairly well to this traditional view, despite occasional rhetoric to the
contrary.

1.2 Why Simultaneous Abstraction?

Since Montague’s pioneering use of IL for the semantics of a fragment of English,
λ-abstraction has continued to play a key role in linguistic semantics. Yet, while
(repeated) unary abstraction may be all we need for mathematical logic, it looks

Robin Cooper takes a similar view in [Cooper 1993]. In contrast to him, I do not think that simultaneous abstraction by itself provides an analysis of what Kamp calls the ‘sharing’ of discourse referents in [Kamp 1990].

For a more radical, syntactic approach to attitudes, see [Asher 1993], who does not only regard DRSs as the “language of thought”, but also takes attitude reports to refer to these formal structures.
less satisfactory when we encounter linguistic applications, in particular to languages with more freedom in their word order than English. It is awkward, if not genuinely inadequate, to impose a strict order on the denotations of predicative expressions, which then has to be undone by sophisticated means, to cope with words coming in the ‘wrong’ order. Moreover it seems slightly puzzling how easily logicians with Realist views on properties and relations seem to extrapolate the order of words of their preferred language into the structure of those independently existing entities. Yet, my main reason for embarking into simultaneous abstraction is that it is already being put to good use, if often in disguised forms, in linguistic semantics. Forms of simultaneous abstraction seem to be at work in the three prominent semantic approaches of the post Montagovian era to be discussed in this thesis, namely in Discourse Representation Theory, Dynamic Montague Grammar, especially when it is based on the Dynamic Property Theory (DPT) of [Chierchia 1994], and modern Situation Theory, as described in [Barwise & Cooper 1991], with its underlying theory of abstraction (AL) developed in [Aczel & Lunnon 1991].

The basic principle of simultaneous abstraction I take to be that the \( \lambda \) operator takes a whole set of variables, or something to that effect, rather than just one at a time (or a list that is easily ‘carried’ into a sequence of singular abstractions). To give a concrete example, we may write something like

\[
\lambda \{x,y\}.likes(x,y)
\]

in a system like the language \( \Lambda \) of the next chapter.\(^3\) The absence of any order in \( \{x,y\} \) creates the problem of defining a sensible operation of application. We may want to express the fact that Robert Parker likes Dominus. So we need a way to unambiguously determine which of the objects \( rob \) and \( dom \) fill which of the positions abstracted over. The solution to this difficulty comes in the form of application to assignments. I write the application in question as

\[
(\lambda \{x,y\}.likes(x,y))(\{\text{rob},\text{y,dom}\})
\]

which then reduces to \( \text{likes(rob, dom)} \). Our abstracts denote, or correspond to, functions from assignments to objects of some kind; in this case a function from variable assignments to propositions.\(^4\) It is instructive to regard the free variables in an open formula as a kind of paradigm case: the denotation of an open formula is precisely a function from variable assignments to the contents of closed expressions. Other examples are not hard to find either.

DRSs for example are usually taken to be pairs \( (U,C) \) consisting of a set \( U \) of variables (the ‘universe of discourse markers’ of the DRS) and a conjunctive set \( C \) of ‘conditions’. In giving truth conditions to DRSs the variables in \( U \) are read as being ‘non-selectively’ existentially quantified, but only those of the outermost level DRS. Embedded DRSs receive all sorts of quantification over their universes, most prominently a non-selective universal reading of the discourse markers in \( U \) in a conditional of the form \( (U,C) \Rightarrow (U',C') \). Abstraction over its universe seems thus a plausible mechanism for giving a denotation to DRSs, which leaves the quantification force to be fixed explicitly or implicitly by the context in which they occur. The standard semantics for DRT does not give an explicit denotation to DRSs, and thus seems to avoid needless complications involving abstraction, but once we widen our horizons and allow DRSs to occur as arguments of relations, in particular those involving attitudes, the complications have to be faced in one way or another. My inclination is to regard a DRS as a term \( \lambda U.C \), where we abstract over an unordered set of variables \( U \), rather than in the one by one fashion of IL, and \( C \) expresses a fine grained conjunction of propositions, rather than a truth value or a set of possible worlds. The content of a DRS is thus a property in as many arguments as there are variables in its universe.

DMG was developed to introduce ideas from DRT into the compositional

\(^3\) The reader may suspend concerns regarding the order in \( \text{likes}(x,y) \) until the coming chapter. Notice though how application of \( \text{likes} \) to \( x \) and \( y \) turns an ordered relation into an unordered one, where the variables act as identifiers for the roles of the relation.

\(^4\) Variables are then used rather like key words in records. Case marking might be regarded in this way for example, when the language in question employs case rather than order for the identification of the roles of a predicate.
framework of IL, by means of a new basic type of “states” which are essentially assignments of entities to discourse markers. Chierchia’s DPT gets rid of the states as primitives, in favor of explicit assignments which are manipulated by means of two operators: \( \alpha \) for abstraction, and \( \iota \) for application. In fact the \( \alpha \)-operator can be understood as abstracting over the infinite set of discourse markers, while \( \iota \) applies a term to the current discourse marker assignment. There is no need to be quite so restrictive in what may be abstracted over and applied to, since the underlying idea is a perfectly general one. In our system any set of variables can be abstracted over, and any assignment can be applied to. This may dissolve any remaining doubts about the difference between variables and discourse markers and simplifies matters by eliminating unary abstraction and application as primitive operations from the system.

AL finally is a system of simultaneous abstraction in which \( \lambda \) takes one-one mappings from labels, called “role indices”, to a kind of variables, called “parameters”, as input to abstract over an unordered set of variables in one go. Application is performed on assignments of objects to those labels. The above example would look more like this in AL:

\[
(\lambda (r_1 \mapsto X, r_2 \mapsto Y). \text{likes}(X, Y)) (r_1, \text{rot}, r_2, \text{dom}).
\]

The essentially simple core of simultaneous abstraction and application in AL is obscured by some less obvious features of the system, which derive from its historical role in modeling Situation Theory via generalizations of the non well-founded set theory of [Aczel 1988]. In particular the theory is developed not as a formal language, but within a highly general setting of a theory of structured objects. Some of these objects are of an especially suspect looking “indeterminate” kind, namely the “parameters”. \( \lambda \) is an operation “in the model”, which takes role-indexed sets of parameters and parametric objects into things called “abstracts”. Much concern is devoted to the task of finding structured universes containing such abstracts, in which systems of equations involving those parameters have unique solutions in terms of anchoring the parameters to non-parametric objects. Given the other different concerns of Linguistics, it is little wonder then that AL has not found widespread popularity among semanticists,\(^5\) who remain to be convinced that such sophisticated machinery is needed to achieve their modest ambitions. This thesis offers a hopefully more accessible system of simultaneous abstraction, which shares many — though perhaps not all — of AL’s virtues, despite its comparatively simple minded, traditional design features.

1.2.1 Variables: “Bound yet Free”

A peculiar hurdle on the way to acceptance of the very idea of simultaneous abstraction is the resulting (partial) loss of α-equivalence. For example we have

\[
\lambda\{x, y\}. \text{likes}(x, y) \neq \lambda\{y, x\}. \text{likes}(y, x).
\]

This loss is an inevitable consequence of the way in which application has to work for such terms. A renaming of variables in an abstraction term will normally change the object, as the two supposedly equivalent terms will yield different results when applied to the same assignment. For example, applying the two terms above to \((r.\text{rot}, y, \text{dom})\) yields two quite different results. This should come as no surprise. It is of course well known that free variables cannot be renamed. Neither can the abstracted discourse markers of a DRS be changed without consequences: any subsequent anaphoric link depends on the very identity of the variables in the universe of a DRS. In the same vein no renaming is allowed for DPT discourse markers, may they be free or in the scope of a \( \alpha \)-abstraction. This is easily demonstrated by an application of \( \iota \), which frees those discourse markers again by \( \alpha \)-cancellation. Parameters in AL can be renamed, if abstracted over, but role indices cannot without changing the object.

I will not embark here into an argument about whether it is wrong to speak of “abstraction” when there is no (full) \( \alpha \)-equivalence, though it seems legitimate to me to use the term even in the case of free variables. It is more important to see that it is not a virtue in itself to be able to rename bound variables, as this means that computational work has to be done to establish a trivial identity of objects. Its virtue, as we shall see, lies in obtaining a total substitution operation, which

\(^5\)Robin Cooper is the notable exception.
is essential for the full exploitation of $\beta$-equivalences.

The idea that variables may be essential to the identity of a semantic object is certainly now fairly accepted, due to DRT and its "dynamic" children. From a more philosophical perspective I would argue that using variables in this way is no more a contamination of semantic objects by non semantic concepts than the common view that the order of arguments to a many place function should matter. From my perspective the charge of a lack in $\alpha$-equivalence exposes the corresponding lack of permutability for the curried or otherwise ordered systems.

1.2.2 Axiomatic Semantics

Given the apparent similarities between the semantic theories mentioned, it seems worthwhile to look for a common framework which brings out the similarities as well as the differences between theories like DRT, DMG, and ST more clearly and with formal precision. The core of this framework is my Simultaneous Abstraction Calculus (SAC). Semantic theories can be obtained from it in very much the same way as from the Lambda Calculus. I will pursue here an axiomatic approach that follows the model of Property Theory and Situation Theory in taking the fundamental semantic notions of properties, propositions, and truth pretty much at face value, and describes them by explicit axiomatic theories. One advantage of this approach lies in the possibility to incrementally accumulate such axioms as more aspects of the structure of propositions get treated, and more fine grained distinctions between propositions are introduced. One can thus be formally explicit about what is common to different semantic theories and what is not. DMG and DRT for example share a common theory of identity and propositional logic, but diverge in the kinds of quantification they employ, as well as in the syntax-semantics interface. Some form of ST can be obtained by developing a logic of situation types within such unsituated theories of propositions. Furthermore one can obtain results about the consistency of the various assumptions behind different theories, allowing for safe transfer of solutions from one to the other, such as for example the use of Situation Theoretic ideas within DRT or DMG.

My approach is minimalist in its spirit, introducing assumptions only when necessary into ones theories, as explicit claims rather than implicit into the definition of the formalism itself. A good example of this is the lack of any typing of our expressions. I do not want to exclude self-application and the like right from the start from my framework, and prefer to allow for different explicit theories that restrict the formation of meaningful propositions.

The topic of relating semantic objects to natural language utterances will be left largely in the background. I believe that given the spirit of axiomatic specification of this work a congenial approach would be to formalize the syntax-semantics interface in a way that leaves as much as possible open about the precise nature of the semantic objects involved.

1.2.3 Bealer's Puzzle

In [Bealer 89] George Bealer has argued that any theory which identifies properties with abstractions from propositions, as I am pursuing it, is not fine grained enough for dealing with attitudes. To quote his example, from the proposition "Jane Fonda follows Rajneesh" $(follow(f, r))$ one can form the properties to "fondalee" $(\lambda x. follow(f, x))$ and to "rajneesh" $(\lambda x. follow(x, r))$, so that by functional application we get that (i) 'Jane Fonda rajneeshes' and (ii) 'Rajneesh fondalees' express the very same proposition $(\lambda x. follow(x, r))((f) = follow(f, r) = (\lambda x. follow(f, x))(r))$. If one agrees that it is possible to believe (i) without (ii), or at least accepts that a plausible example might be constructed along these lines, then one has to admit that functional application, governed by the principle of $\beta$-reduction, is guilty of erasing too many distinctions.

Bealer's solution is to use combinators for filling and reordering argument places, which do not generate equalities like the one between (i) and (ii) above, but only logical equivalences between distinct propositions. For example instead of using abstraction to permute a relation $R$ into $\lambda x \lambda y. Rx y$, with the resulting equality $Rab = (\lambda x \lambda y. Rx y)ba$, he introduces basic operations like $\text{case}$ and $\text{pred}$, such that $\text{pred} \text{predcase}(R, b) \neq \text{pred}(\text{pred}(R, a), b)$, but one is true whenever the other is.

One might envisage a transfer of this approach into an unordered setting, yet I find it plausible that the meanings of incomplete phrases, including words like
‘foundational’, are learned by grasping what claim they produce in which context of other meaningful words. Sticking to some device of abstraction, it thus seems necessary to invoke a structure preserving notion of predication of propositional functions to their arguments, besides the structure erasing application operation, in order to cope with this puzzle. Bealer’s introduction of two kinds of propositions in [Bealer 1982], namely unstructured ‘conditions’ and highly structured ‘Cambridge propositions’ seems to be very much in line with the above distinction. I thus prefer to follow Peter Aczel, [Aczel 1989], and distinguish application from predications in chapter 3.

1.3 Other Frameworks

My approach is very much inspired by the work of Robin Cooper who uses the Situation Theoretic language EKN as a framework in which various semantic theories can be expressed, as in [Cooper 1993]. One problem with his approach is that one seems to be committed to take a formidable number of Situation Theoretic assumptions on board in order to achieve ones often rather limited aims. In [Cooper & Poesio 1994] an attempt is made to deal with this charge by arranging the axioms into convenient packages, under the banner of ‘algebraic specification,’ in order to allow for a more minimalist use of the tools. Still, the theory looks quite complex, mainly due to the use of parametric objects with AL style abstraction, and the rather fundamental role that is played by situations and ‘infons’. My hope is that my humbler system of abstraction and more prominent use of propositions can make things easier to understand and to use in applications. In fact as far I have carried out this work the number axioms has remained so small that the use of specification methods would seem out of proportion.

Another notable effort to unify the three semantic theories mentioned above is found in Reinhard Muskens’ work. He does not temper with the basic aspects of the rigidly typed lambda calculus, but prefers to modify it to cope with the new challenges. Thus he introduces a new basic type of ‘states’ which by means of some axioms are in one-one correspondence with discourse marker assignments. This allows him to formulate systems of DMG, [Muskens 1992], and DRT, [Muskens 1994], if in a somewhat indirect manner. The systems are ‘dynamic’ in using DPL-style relational composition, following [Groenendijk & Stokhof 1991], which allows to capture discourse markers to the right of the syntactic scope of a quantifier. Otherwise he sticks to the identification of propositions with functions from worlds to truth values, but in adopting a four values lattice of truth values in [Muskens 1989] he effectively turned possible worlds into (im)possible situations. This created a version of ST, which is very close to the ‘classical’ theory of Barwise & Perry 1983. I find my system as lying in between the in some sense more conservative approach of Muskens’, and Cooper’s embrace of the unfamiliar.
Chapter 2

Simultaneous Abstraction

In this part two slightly different formalisms for simultaneous abstraction are introduced. First I look at the conceptually simple, if somewhat unorthodox language $\Lambda$, and then at the proof theoretically more convenient $\Lambda_{UL}$, which is equivalent to the original $\Lambda$, modulo some additional variables. I give a proof theory for $\Lambda_{UL}$ and show the Church-Rosser property for reduction. I investigate $\Lambda_{UL}$ from a Combinatory Logic perspective, establishing a completeness result for it. Finally I discuss structured objects, and the solving of equations, which may give us non well-founded objects.

For the full understanding of chapter three the proof of theoretic sections, particularly from 2.2.3 to the end of 2.3, are not essential.

2.1 Simple Simultaneous Abstraction

2.1.1 Syntax for $\Lambda$

Instead of abstracting over single variables, as in standard $\lambda$-calculus, we allow $\lambda$-abstraction over any set of them. Terms are applied to records that assign terms to variables. Ways of introducing unitary abstraction and application into such a system will be discussed later on. A purified version of the earlier ontological example, where all applications are now performed uniformly to variable assignments, would now look like this. Let us take $z_1$ and $z_2$ to function as the 'likes' and 'liking' roles of our predicate 'likes'.

\[
(\lambda [x, y]. \text{likes}(z_1, x, z_2, y)) \langle x, \text{red}, y, \text{dom} \rangle
\]

By convention differently indexed variables are always assumed to be non-identical.

Definition 1 The language $\Lambda$ consists of terms $t, t' \ldots$ built up from basic constants $c, c', \ldots$ and $\#$ for "undefined," and variables $x, y, \ldots \in \text{Var}$ by means of abstraction $\lambda M.t$, for $M \subseteq \text{Var}$, and application $t(x, t')_{\# t}$.

Notice that terms of the language $\Lambda$ are fairly abstract objects, containing sets of variables and unordered records as constituents. We in fact allow infinite sets of variables $\{x_i\}_{# t}$ to be abstracted over and also infinite records $(x_i, t_i)_{# t}$ to be applied to.\(^1\) Infinite terms are useful for theoretical purposes, even if in applications one will have to restrict oneself to finitary $\Lambda$, possibly extended by short hand versions of a few selected infinitary expressions.

In writing an expression like $\lambda [x_1, \ldots x_n] \cdot t'$ or simply $\{x_1, \ldots, x_n\} \cdot t'$ I therefore do not mean to suggest that the terms denoted by these expressions have any order in their set of abstracted variables, as one might suspect from the indexing, and neither do I mean to imply that $n$ is a finite number. An similar remark holds for application terms: all indexes are part of the meta language, and may be attached freely to terms, subject only to the convention that different variables should be indexed differently. To give a concrete example, I take $(x_1, x_2) \cdot t$ to be the very same term as $(x_2, x_1) \cdot t$, and in the same vein $(x_1, x_2, y, t_b)$ and $(y, x_2, x_1) \cdot t$ to be identical $\Lambda$-expressions. Thus $\Lambda$ is not a language that can be written, but a convenient abstraction over any systems of the appropriate structure which allow for free permutation within any particular abstraction or application. Order is only significant for multiple uses of abstraction or application, for example syntactically (and semantically) $(x_1)\{x_2\} \cdot t \neq \{x_2\} \{x_1\} \cdot t$ and $(y, x_2) (y, t_2) \neq (y, t_2, x_1) \cdot t_1$.

\(^1\)The notation $\{x_i\}_{# t}$ stands here for the set $\{x_i \in t\}$, not for a function from indices to variables, as it is sometimes used by other authors, such as [Aczel & Lambek 1993].
2.1.2 Semantics for $\lambda$-calculus

The semantics for $\lambda$ will be a fairly straightforward variation on the semantics of the untyped lambda calculus. The crucial difference is that in our system the objects denoted by our terms apply not to objects, but rather to partial assignments of objects to variables. We thus need a set $D$ of objects and a way to apply those objects to assignments: with every object $d \in D$ we need to associate a function $\Psi(d)$, which I call $d$’s applicative behavior, telling us which object in $D$ is the value of applying $d$ to which assignment. Thus $\Psi$ is a mapping from $D$ to $(\text{Var} \to D) \to D$. To deal with partial assignments we assume that $D$ comes with a distinguished element $\bot_D$ that plays the role of the “undefined object”. We write $(x \mapsto \xi_i)_{i \in I}$ for the function $f$ such that $f(x_i) = \xi_i$ for $i \in I$ and $f(x) = \bot$ otherwise, and let $\mathsf{dom}(f) = \{x \mid f(x) \neq \bot\}$. This is enough to make sense of application for our language: ignoring free variables, we get $[(t(x_1, x_2)]_{i \in I} = \Psi([t]_i)((x_1 \mapsto [t]_i)_{i \in I})$.

To interpret an abstraction $\lambda t. t$ we look at the interpretation of $t$ under variation of the objects that may be assigned to the variables in $t$. In an obvious generalization from standard usage we define variants $\varphi^y_t$ of an assignment $y$ for a set of variables $M$ and an assignment $f$ by $\varphi^y_t(x) = f(x)$ if $x \in M$, otherwise $\varphi^y_t(x) = g(x)$, for all $x \in \text{Var}$. Thus, given $t, M$ and $g$ there is a function $\lambda_{t \in M} \varphi^y_t \in (\text{Var} \to D) \to D$, from assignments $f$ to the interpretation of $t$ under $\varphi^y_t$. This function defines the applicative behavior of $\lambda t. t$. For the denotation of this term all that is needed is a member of $D$ which has precisely this applicative behavior. We thus assume a mapping $\Psi$ from $(\text{Var} \to D) \to D$ to $D$ which maps any applicative behavior $\mu$ to an object $\Psi(\mu)$ which has this behavior, which is to say that $\Psi(\Phi(\mu)) = \mu$. We are thus dealing with a retraction between two domains, written $D \leftrightarrow_D C$, which is a pair of mappings $\Phi : D \to C$ and $\Psi : C \to D$ such that $\Psi \circ \Phi = id_C$.

One final point has to be observed though: $\Phi \circ \Phi = id_C$ implies that $\Phi$ is injective, hence $|D| \geq |C|$, which means there cannot be a retraction between $a$ set $D$ and the full function space $(\text{Var} \to D) \to D$. One way to cut down the size of the function space without throwing away some vital functions is to work in the category CPO of complete partial orders with continuous functions between them. So a domain will be understood to be a cpo with a least element $\bot$, a mapping between cpo’s to be continuous, and the operation $\to$ to form the space of continuous functions, ordered pointwise. The two mappings of a retraction are in addition required to be strict, that is, to preserve the bottom element. Later, in section 2.3, I will use a combinatory logic to restrict the space of applicative behaviors without the use of domain theory. I will regard that version of the semantics as the official one.

**Definition 2** A $\lambda$-model consists of a retraction $D \leftrightarrow_D (\text{Var} \to D) \to D$, and an interpretation $\mathfrak{I}$ which maps constants into $D$, with $\mathfrak{I}(\bot) = \bot_D$. The denotation function $[[\cdot]]$ for terms, under a variable assignment $\mathfrak{g}$, is given by:

- $[[\mathsf{id}]_\mathfrak{g}] = \mathfrak{I}(\mathsf{id})$,
- $[[x]]_\mathfrak{g} = \mathfrak{g}(x)$,
- $[[\lambda t. t]]_\mathfrak{g} \equiv \Phi(\lambda_{t \in \text{Var}} [[t]]_\mathfrak{g})$,
- $[[t(x_1, x_2)]_{i \in I}]_\mathfrak{g} \equiv \Psi ([[t]_i]((x_1 \mapsto [[t]_i]_{i \in I})_{i \in I})$.

A model is non-trivial if $|D| > 1$. Working in CPO the denotations are only well-defined because the following holds.

**Theorem 1** $[[\lambda t. t]]^\mathfrak{g}$ is continuous for all terms $t$.

**Proof**: Notice that $[[t]] = [[\lambda t. t]]^\mathfrak{g}$ is an instance of this. We show by induction that lambda $h^\mathfrak{g} = \bigcup_{\mathfrak{g} \in \mathfrak{G}} h_\mathfrak{g}$ of ascending chains of assignments are preserved. I leave the proof of monotonicity to the reader.

- $[[\lambda t. t]]^\mathfrak{g}(h^*) = \mathfrak{I}(\mathfrak{g}) = \bigcup_{\mathfrak{g} \in \mathfrak{G}} [[\lambda t. t]]^\mathfrak{g}(h_\mathfrak{g})$.

---

2See [Hindley & Seldin 1986] for an introduction into models for the untyped lambda calculus.

3The ‘$\lambda$’ in the following expression is used as a symbol of the meta language, with the standard meaning.

4For more background information on these notions see also [Barendregt 1981]. Notice that every partial assignment function is continuous (if strict), given that Var is a flat domain with an added $\bot$, as I will assume.
CHAPTER 2. SIMULTANEOUS ABSTRACTION

2.1.3 Basic Properties of \( \lambda \)

Definition 3 The free variables of a term are defined by

- \( FV(x) = \{ x \} \);
- \( FV(c) = \emptyset \);
- \( FV(\lambda M.t) = FV(t) \setminus M \);
- \( FV(t(x,t)_{\emptyset \emptyset}) = FV(t) \cup \{ x \} \).

Lemma 2 If \( x \notin FV(t) \) then \( \| t \|_{t'} = \| t \|_{t''} \) for all \( g \) and \( d \).

So if \( x \) is not free in a term \( t \) then the denotation of \( t \) is independent of what we assign to \( x \), but this does not mean one could rename bound occurrences of \( x \) by some \( y \) which is fresh for \( t \). For example, the denotation of \( \{ x \} \) is independent of \( x \), but \( \| \{ x \} .x \|_{t'} \neq \| \{ y \} .y \|_{t'} \). The latter becomes clear when we apply these two functions to suitable assignments e.g., \( \| \{ x \} x (x,y) \|_{t'} = \| a \|_{t'} \neq \| b \|_{t'} = \| \{ y \} y (x,a,y,b) \|_{t'} \), presuming our model is non-trivial.

\[ * \]

Definition 4 Simultaneous substitution \([s/y]_{\emptyset \emptyset} \) (dropping the index set when no confusion can arise) is a partially defined operation given by:

- \([s/y]_{\emptyset \emptyset} \{ x \} s = s \), if \( x = y \) for some \( j \in J \), otherwise \( x \);
- \([s/y]_{\emptyset \emptyset} c = c \);
- \([s/y]_{\emptyset \emptyset} \lambda M.t = \lambda M.([s/y]_{\emptyset \emptyset} t) \), where \( J = J \setminus \{ j \} \); \( y \in M \), if \( M \cap \{ s/y \} FV(s) = \emptyset \), otherwise undefined\(^4\).

An example of an undefined substitution would be \([y/x]_{\emptyset \emptyset} \{ x \} t \), where replacing \( x \) in \( t \) by \( y \) would give undesired results and renaming \( y \) is not a possibility. To avoid such trouble we require a fresh variable for a term to occur neither free nor bound in it.

Let us say that \( \models t = t' \) iff for every \( \lambda \)-model and assignment \( \| t \|' = \| t' \|' \).

Theorem 3 There are terms \( t \) such that:

1. \( \not\models \{ x \} = \{ y : y \} \{ x/y \} \),
2. \( \not\models \{ x \} \{ x \} \{ x \} \{ x \} \).

This is obvious if non-trivial models exist. For the \( \Lambda\)-models of section 2.3.2 their existence follows from the non triviality of the term model via the Church-Rosser Theorem.

I will now present a few equations, where the list is not meant to give the axioms of a complete calculus for \( \Lambda \). The fact that substitution is not always defined is responsible for the difficulties in giving such a list. It means that \( \beta \) reduction cannot be exploited in many cases where one would usually expect it to be useful. Fact 5.3, which allows for some amount of renaming of variables, is just one example of an equation that cannot be derived from the limited \( \beta \) reduction allowed by fact 5.2, because substitution is partial. It seems a good idea to look for a complete calculus for \( \Lambda\)-model, to be defined later, rather than for \( \Lambda \), since substitution can be totally defined in that system.

\(^4\)Being undefined is meant here in the sense of not producing any result, rather than resulting in \( \emptyset \).
Lemma 4 \( \|\cdot\|_{\text{M}} = \{x_i\}_{\text{set}} \|\cdot\|_{\text{M}} \) if \( \gamma = \{x_i \mapsto \|\cdot\|_{\text{M}}\} \) and \( \gamma = \{x_i \mapsto \|\cdot\|_{\text{M}}\} \).

**Proof:**

- obvious for atomic terms;
- \( \|N_{x,v}^{\text{M}}\|_{\text{M}} = \Phi \lambda x\|\cdot\|_{\text{M}}^{\text{M}} = \Phi \lambda x\|\cdot\|_{\text{M}}^{\text{M}} = \Phi \lambda x\|\cdot\|_{\text{M}}^{\text{M}} \)
- \( \|N_{x,v}^{\text{M}}\|_{\text{M}} = \{x_i\}_{\text{set}} \|\cdot\|_{\text{M}} \)
- \( \gamma = \{x_i \mapsto \|\cdot\|_{\text{M}}\} \)
- \( \gamma = \{x_i \mapsto \|\cdot\|_{\text{M}}\} \)

Theorem 5 \( I = \lambda N.t = \lambda N.t \) for \( N = M \cap FV(t) \);

\[\|t\|_{\text{M},\gamma} = \{x_i\}_{\text{set}} \|t\|_{\text{M},\gamma} \]

where \( J' = J \cap I \) and \( I' = I \setminus J \);

3. \( \|\{x_i\}_{\text{set}} \{x_i\}_{\text{set}} \{x_i\}_{\text{set}}\|_{\text{M}} = \{x_i\}_{\text{set}} \|t\|_{\text{M},\gamma} \)

**Proof:**

1. From Lemma 2.

2. By Lemma 2 and 4 for every \( y \) in all-\( \lambda \)-models:

\[ \|t\|_{\text{M},\gamma} = \{x_i\}_{\text{set}} \|t\|_{\text{M},\gamma} \]

where \( M = \{x_i\}_{\text{set}} \), \( \gamma = \{x_i \mapsto \|\cdot\|_{\text{M}}\} \).

3. Again by Lemma 2 and 4:

\[ \|\{x_i\}_{\text{set}} \{x_i\}_{\text{set}} \{x_i\}_{\text{set}}\|_{\text{M}} = \{x_i\}_{\text{set}} \|t\|_{\text{M},\gamma} \]

where \( M = \{x_i\}_{\text{set}} \), \( \gamma = \{x_i \mapsto \|\cdot\|_{\text{M}}\} \) and \( \gamma = \{x_i \mapsto \|\cdot\|_{\text{M}}\} \).

2.1.4 Partial Application

It is very natural to ask for an operation for partially filling the argument roles of a relation, leaving some of them simply open for more to come. We use the notation \( t[x_1, \ldots, x_m, t_0] \) for partial application. The semantics is as follows:

\[ \|t[x_1, t_1]_{\text{set}}\|_{\text{M}} = \Phi \gamma \|\cdot\|_{\text{M}} \]

The following properties of partial application are of particular interest, letting application associate to the left.

**Theorem 6**

\[ t[x_1, t_1]_{\text{set}} = t[x_1, t_1]_{\text{set}} \]

where \( J' = J \cap I \) and \( I' = I \setminus J \);

2. \( t[x_1, t_1]_{\text{set}} = t[x_1, t_1]_{\text{set}} \)

where \( J' = J \setminus \{ \exists y_2 \mid y_2 \}

3. \( t[x_1, t_1]_{\text{set}} = t[x_1, t_1]_{\text{set}} \)

where \( J' = J \setminus \{ \exists y_2 \mid y_2 \}

4. \( t[x_1, t_1]_{\text{set}} = t[x_1, t_1]_{\text{set}} \)

where \( J' = J \setminus \{ \exists y_2 \mid y_2 \}

Notice that the \( \beta \)-conversion of partial application, as expressed in equation No. 1, does not reduce the set of abstracted variables, if the substitution is defined, but never removes the abstraction altogether, even if the set \( \{x_i\}_{\text{set}} \) is empty.

Notice also that we do not have \( \|t\| = t \) unless \( t \) is an abstraction term.

To have this equation hold in full generality would require our models to be extensional. For that, \( \Phi \) would have to be an injection, which turns our restriction into an isomorphism between \( D \) and \( \langle \text{Var} \rangle \). But our semantics is only weakly extensional in that any two abstracts with the same applicative behavior denote the same object, as they will both denote the value of \( \Phi \) for this particular behavior.

There seems to be a way to define partial application by means of the fourth equation, or some variant of it which makes equally sure that any further rule of \( t \) can be filled by some further application. But this cannot be ensured because
such roles/variables might occur free in some $t$, and thus the side condition of clause 4 will be violated. In $\Lambda_{AL}$ there will be no such problem for defining partial application any more.

### 2.2 Aczel–Lunnon Abstraction

The language $\Lambda$ is closely related to Barwise and Cooper's Extended Kamp Notation (EKN), see [Barwise & Cooper 1981], which is based on Aczel and Lunnon's theory of abstraction. In EKN $\lambda$-abstraction operates on injective functions from a set of role indices to parameters. In the case of $\Lambda$ those two notions are merged into one, the variables, which allows us to replace the injections by simple sets.

Variables thus play a double role in our system: as role identifiers and as means for directing the values for such roles into the intended slots. We have seen that this creates problems, as by virtue of their former function, variables can often not be renamed and thus certain substitutions, such as $[x/y][x]t_{ab}$, cannot be performed. A similar problem is encountered when we try to find an abstraction term $t'$ which uses the same role $x$ on two levels of abstraction, so that $t'(x.a)(x.b) = t_{ab}$. The simple idea of two abstractions $\{x\}$ $\{x\}t_{ab}$ fails as the outer role $x$ cannot be linked to the correct position in $t$ for reasons of interference with $x$'s use as a role identifier in the inner abstraction. Using different variables instead creates the problem that application can no longer be to assignments to $x$ in both arguments.

The following system revises $\Lambda$ by separating the two uses of variables, using $\Lambda_{AL}$-style syntax. The old variables $x, y, \ldots \in \text{Var}$ will now be used as roles only, while upper case variables, called "parameters" go into the positions of terms, to assign values into their intended positions. Our standard example becomes now

$$\lambda(X \rightarrow x, Y \rightarrow y).\text{likes}(x, X, y, Y) \subseteq x, y, \text{dom},$$

where I have turned the mapping $x \rightarrow X, y \rightarrow Y$, as it would appear in $\Lambda_{AL}$.

The term $t$ we were looking for above can now be constructed as $\lambda(Y \rightarrow x)\lambda(Z \rightarrow x).t_{ab}$. It thus seems possible that such a move may increase the expressivity for our system. I will show that the new language $\Lambda_{AL}$ is in fact equivalent to the old $\Lambda$, modulo some additional variables, but the proof of theory is sufficiently simpler for the new language to warrant its introduction.

**Definition 5** The terms $t, t', \ldots$ of $\Lambda_{AL}$ are built up from constants $c, \# \ldots$ and parameters $X, Y, \ldots$ by means of abstraction $\lambda(X)\$ over injective functions $\chi$ from parameters to variables (=roles), and application $t[x, t']\text{get}$.

Semantically we stay in the same space of domains $D \equiv D (\text{Var} \rightarrow D) \rightarrow D$ as for $\Lambda$, with the only difference being that we now interpret under assignments of objects to parameters, called anchors, and that these have to be obtained from assignments to roles when we define the meaning of abstraction.

**Definition 6** A $\Lambda_{AL}$-model is a $\Lambda$-model, where the denotation function $[\cdot]$ for terms, under an anchor $g$, is given by:

- $[c] = \exists(x)$;
- $[X] = g(x, X)$;
- $[\lambda \chi \cdot \theta]^g = \phi_{\lambda \chi \cdot \theta}^g$;
- $[[t[x, t']]\text{get}] = \psi([t][t'][x \rightarrow g(x), t''[t'']\text{get}])$.

Given that we always have enough fresh parameters at our disposal, which can be insured by requiring $|\text{Par}| > |\text{Var}|$, we can define partial application by

$$t[x, t']\text{get} = \text{get} \lambda(x) \rightarrow x; t[x, t', x, X]_{\text{get}, \text{get}, t}.$$

In $\Lambda_{AL}$, the mappings in abstractions take roles to parameters injectively. In turning them around we could be more general by allowing several parameters to be mapped to the same role, which is equivalent to replacing those parameters with one of them, which then gets mapped to the role in question. Notice that associating one parameter with many roles in one abstraction clearly makes no sense. My semantics for $\Lambda_{AL}$ will not presuppose injectivity, but I will assume it here nevertheless for convenience.
where \( \{x_j \mid j \in J\} = \text{Var} \setminus \{x_i \mid i \in I\} \) and the \( X_j \) are fresh. The resulting semantics for partial application is the same as for the basic operation we added to \( \Lambda \). Notice that infinitary abstraction is used to achieve such expressivity.

The move to \( \Lambda_{AL} \) solves the problems with substitution in \( \Lambda \), as substitution can now be totally defined by means of renaming parameters, so that for example
\[
\frac{\lambda (X \to z)_{TV} = \lambda (Z \to z)_{TGY}}{X/V}.
\]

2.2.1 Relating \( \Lambda, \Lambda_{AL} \), and \( \lambda \)

The question that needs to be addressed is whether we gain more from the switch from \( \Lambda \) to \( \Lambda_{AL} \) than just convenience. There might be terms in the new language which have no equivalent in the original language \( \Lambda \). To show that no serious gain in expressive power is made we have to face the earlier difficulty for \( \Lambda \) again, namely that of allowing multiple uses of roles in different levels of abstraction, without clash of variables. Such ‘role recycling’ is made easy in \( \Lambda_{AL} \) by using different parameters in a term like \( \lambda (Y \to x)_{TV} \lambda (Z \to x)_{TGY} \). In creating a \( \Lambda \)-term with the same role \( x \) used twice we faced the problem of the outer abstraction in \( \{x\} \{x\}_{TXY} \) being vacuous, thus failing to connect \( x \) to the intended position. The solution is to inject an intermediate role \( y \) to arrive at \( \{x\} \{y\}_{TXY} \{y,x\} \). To have such roles available at all times we expand our language \( \Lambda \) to \( \Lambda' \) with its new set of variables \( x, y, \ldots \in \text{Var}' = \text{Var} \cup \text{Par} \). The idea can now be applied in full generality.

**Definition 7** The translation \( [x_i/X_i]^{\text{Par}}_{\text{AL}} : \Lambda_{AL} \to \Lambda' \) under a substitution is defined by:

- \( [x_i/X_i]^{\text{Par}}_{\text{AL}} x_i = x_i \)
- \( [x_i/X_i]^{\text{Par}}_{\text{AL}} c = c \)
- \( [x_i/X_i]^{\text{Par}}_{\text{AL}} \lambda (Y_j \to y_j)_{\text{AL}} = \lambda (X_i \to x_i)_{\text{AL}} \lambda (Y_j \to y_j)_{\text{AL}} (X_i, x_i)_{\text{GyP}} \)

where \( K = I \setminus \{Y_j \mid j \in J\} \) and \( P' = K \cap \{y_j \mid j \in J\} \).

- \( [x_i/X_i]^{\text{Par}}_{\text{AL}} \lambda (Y_j \to y_j)_{\text{AL}} = [x_i/X_i]^{\text{Par}}_{\text{AL}} \lambda (Y_j \to y_j)_{\text{AL}} (x_i, x_i)_{\text{GyP}} \)

**Theorem 8** \( \Lambda \) can be embedded in \( \Lambda_{AL} \) and \( \Lambda_{AL} \) can be embedded in \( \Lambda' \).

Let us turn to \( \lambda \), the untyped \( \lambda \)-calculus, see [Hindley & Seldin 1986]. It is clear that we can define unary abstraction \( \lambda X.t \) and application \( t u \) for \( \Lambda_{AL} \) in
terms of a designated role \( z \), used solely for that purpose, as \( \lambda(x \rightarrow z).t \) and \( t(z/t) \) respectively. By the translations between \( \Lambda_{\text{AL}} \) and \( \Lambda \) we see that unary abstraction/application for \( \lambda \)-terms can be encoded in \( \Lambda \), allowing us to use expressions such as \( \lambda x.t \) and \( t(t') \) as shorthands for \( \Lambda_{\text{AL}} \)- or \( \Lambda \)-terms, according to our preferences.

**Corollary 9** \( \lambda \)-abstraction and application for \( \lambda \)-terms can be encoded in \( \Lambda \),\(^9\)

I will from now on take \( \Lambda_{\text{AL}} \) as my official language for the Simultaneous Abstraction Calculus, and regard the use of \( \Lambda \) as well as standard \( \lambda \)-terms as syntactic sugar. Occurrences of small variables \( x \) as terms, or in abstractions \( \lambda x.t \) and \( \lambda z \ldots t \), can be taken as shorthands for their corresponding \( \lambda \)-transitions into \( \Lambda_{\text{AL}} \), based on the ‘upper-case’ mapping from variables to parameters, and the encoding of \( \lambda \).

### 2.2.2 Proof Theory for \( \Lambda_{\text{AL}} \)

The complications in clause three of the translation from \( \Lambda_{\text{AL}} \) to \( \Lambda \) indicate why a complete and reasonably simple set of equations for \( \Lambda \) will be hard to find. The slightly more complex language \( \Lambda_{\text{AL}} \) seems more suitable for this task.

The notions of free parameters and substitution are assumed to be defined in straightforward adaption of these notions from the \( \lambda \)-calculus, to be found in [Hindley & Seldin 1986].

---

\(^9\)It is obvious that terms, for example \( \lambda x.y.z \), cannot be encoded in \( \Lambda \) in the naive way, as \( \lambda(x/y)(x/y) \) or something of similar simplicity. If we follow the path suggested, via \( \Lambda_{\text{AL}} \), the example term becomes \( \lambda(x \rightarrow z)(y \rightarrow z)x(y) \). Translation into \( \Lambda \) proceeds as follows:

\[
\begin{align*}
\left[ f \right](\lambda(x \rightarrow z)(y \rightarrow z)x(y)) &= \left[ f \right]\left[ \left[ \lambda(x \rightarrow z)(y \rightarrow z)x(y) \right] \right] \\
&= \lambda(x \rightarrow z)(y \rightarrow z)x(y)
\end{align*}
\]

\[\lambda(x \rightarrow z)(y \rightarrow z)x(y) = \lambda(x \rightarrow z)(y \rightarrow z)[x(y) = z] \]

Notice that the equivalence of \( \lambda x.y.z \) and \( \lambda x.y.z \) becomes, under this encoding, an instance of Theorem 5.5, which allows us to rename \( X \) by \( X' \) in the resulting \( \lambda \)-term.

\[^{10}\]The infinitary nature of \( \Lambda \), and the treatment of application to over- and under-defined assignments rule out a similarly straightforward encoding in \( \Lambda \). I have not investigated the matter any further. Notice that \( \lambda \) is encoded in finitary \( \Lambda_{\text{AL}} \).

---

**Definition 9** We write \( Ax v \psi t = \ell \) iff \( \psi t = \ell \) can be derived when the set of equations \( Ax \) is added to the calculus below. \( t =_\psi \ell \) stands for \( v \psi t = \ell \), and \( t =_\psi \ell \) iff \( v \psi t = \ell \) can be proved without using \( R1 \).

\[\begin{align*}
A1 & \quad t = t; \\
A2 & \quad t(x,t,y) #set \xi \ell = t(x,t,y) #set \\
A3 & \quad \ell \lambda x.t = \ell \lambda x.t \xi = \ell \chi
\end{align*}\]

\[\begin{align*}
A4 & \quad \ell \lambda(x,y) #set = \ell \lambda(x,y) #set [t(Y, X) #set t] \quad \text{for fresh } Y, X;
A5 & \quad \ell \lambda(x,y) #set (x, t) #set = [t/x, Y, #] # set t \quad \text{with } J' = J \cap t \text{ and } J' = J \setminus J;
A6 & \quad \ell # set = \ell #
\end{align*}\]

\[\begin{align*}
A7 & \quad \lambda x. # = #; \\
R1 & \quad t = t / t = t; \\
R2 & \quad t = t, t = t' / t = t'; \\
R3 & \quad t = t, s = s'/ (i \in I) / t(x, s) #set = t'(x, s') #set; \\
R4 & \quad t = t / \lambda x. t = \lambda x. t
\end{align*}\]

Given infinitary terms, we are stretching the notion of ‘proof’ somewhat with this system: “Being a proof” is not recursive, as termhood is not, and we also have a potentially infinitely rule \( R3 \). If we restrict the calculus to equations between finite terms we get a notion of proof which is much better behaved. I will call that system the **FINITARY \( \beta \)-CALCULUS**.\(^{11}\) Notice that \( \beta \)-reduction of a finite term never produces an infinite one, even in the infinitary calculus, as there is no Axiom or rule other than \( R1 \) by which to obtain an equation \( t = s \) between a finite \( t \) and an infinite \( s \) from equations that are not of this form.

**Theorem 10 (Soundness)** \( t =_\psi \ell \Rightarrow t = \ell \).

\[^{11}\]It is clear that “having a proof” is not going to be recursive even for the finitary calculus, by the embedding of \( \lambda \) in finitary \( \Lambda_{\text{AL}} \).
2.2.3 The Church Rosser Theorem

The reduction relation $\rightarrow$ is of crucial importance for computational concerns, as it provides us with a way of simplifying terms, possibly to a normal form, which is a term that cannot be further reduced. By the embedding of the $\lambda$-calculus, given earlier, it is obvious though that reduction does not always lead to smaller terms, or terminate at some normal form.\footnote{For example take $t = \lambda x . f (x z z) (\lambda x . f (x z z))$, which reduces to $t \rightarrow f (t) \rightarrow f (f (t)) \rightarrow \ldots$ without ever finding a normal form.} In such a situation it is highly desirable to be sure that reduction cannot lead into blind alleys, from which a possibly existing normal form cannot be found. It is also critical to have a unique one, if existent, at least up to the renaming of bound parameters. Both these facts are consequences of the Church-Rosser property, or confluence, of the calculus. This is called a one step reduction of $t$ over redex $r$.

A reduction $t \rightarrow t'$ can thus be split into a sequence of changes of bound parameters and one step reductions. From one step reduction of $t$ to $t'$ we get various residuals $r'$ of occurrences of terms $r$ in $t$. Let $s$ be the redex contracted to $s'$ in a one step reduction:

- if $r$ is not part of $s$ or vice versa then the unchanged $r$ is $r$'s residual;
- if $r = s$ then $r$ has no residual;

\footnote{The reduction order may not be fully determined by this.}
CHAPTER 2. SIMULTANEOUS ABSTRACTION

For case (i) where \( t' = p'(y_j, y_j')_t \), \( p \cdot \Delta \text{mod} q', q \Delta \text{mod} y_j \) for \( j \in J \) and \( J = J' \) or \( t 's \) residual was \( A3 \)-contracted in the last step, we have \([s_j/X_j]_t = [s_j/X_j]_p \cdot \Delta \text{mod} [y_j, y_j']_t \) or \( \Delta A3 [s_j/X_j]_t \), which is an mcd.

Case (ii) has \( t' = \#_j \), from using \( A5 \) on \( t 's \) residual, with \( p \cdot \Delta \text{mod} q' \) for \( j \in J \). Then \([s_j/X_j]_t = [s_j/X_j]_p \cdot \Delta \text{mod} [y_j, y_j']_t \) or \( \Delta A5 [s_j/X_j]_t \), which is an mcd.

Case (iii) in \( t' = \#_j \) from \( A6 \) applied to \( t 's \) residual. Then \([s_j/X_j]_t = [s_j/X_j]_p \cdot \Delta \text{mod} [y_j, y_j']_t \) or \( \Delta A6 \), which is an mcd.

\( \square \)

Lemma 13 \( t \cdot \Delta \text{mod} u = t \cdot \Delta \text{mod} v \Rightarrow \exists w \cdot \Delta \text{mod} w = t \cdot \Delta \text{mod} w \)

PROOF: We may again assume the given mcd's not to contain any \( a \)-steps.

By induction on \( t 's \):

- \( t = X, t = e \) are trivial.
- \( t = \lambda \xi x \)

Let \( u = \lambda \xi y \) and \( e = \lambda \xi y p' \), both possibly derived by an \( A3 \)-contraction of \( t 's \) residual in the last step. By induction hypothesis there is a \( p + \) with \( p' \cdot \Delta \text{mod} p' \) and \( p' \cdot \Delta \text{mod} p + \). Chose \( w = \lambda \xi x p' \), where \( \xi' = \xi \cdot p' \). If \( u = \# \), using \( A7 \) in the last step, choose \( w = \# \). The mcd is obtained by \( p' = \# \) using \( A7 \).

\( \square \)

Theorem 14 (Church–Rosser Theorem) \( t =_{\beta} t' \Rightarrow \exists s t \rightarrow s \land t \rightarrow s \)

PROOF: If \( t =_{\beta} t ' \) then there are \( t = t_1, t_2, \ldots, t_n = t ' \) such that \( t \cdot \Delta \text{mod} t_1 \) or \( t_1 \cdot \Delta \text{mod} t_2 \) or \( \ldots \) or \( t_n \cdot \Delta \text{mod} t_1 \). The claims follows by induction on \( n \) using the previous lemma. \( \square \)
2.3 Combinatory Aczel–Lunnan Logic

In this section I will develop Combinatory Aczel–Lunnan Logic (CALL), which relates to the SAC in the way that standard Combinatory Logic relates to the \( \lambda \)-calculus. It provides us with a more general notion of a model for the SAC than the previous domain-theoretic one. I will obtain a completeness result for the proof theory of \( \Lambda_{AC} \) relative to these models.

This section is quite technical, and presupposes some familiarity with Combinatory Logic and its relation to the \( \lambda \)-calculus. It may be skipped without much loss for the later chapters.

**Definition 12** The language \( C_{AL} \) consists of terms \( T, U, T', \ldots \), built up from constants \( \#, \), \( \ell, \), \( S \) and \( K_x \) for each \( x \in \text{Var} \), and parameters \( X, Y, \ldots \) by means of application \( T(x, T, k) \).

We use \( z \) as a distinguished variable for unary application, writing \( T(T') \) for \( T(z, T') \). For the interpretation of \( C_{AL} \) we do not use \( c_p \)'s but flat sets \( D_L \) with an added \"undefined\" object \( \perp \). Hence functions are no longer required to be continuous either. The free parameters \( F(\Gamma) \) of a term \( T \) are simply the parameters occurring in \( T \), as we have no abstraction mechanism in \( C_{AL} \).

**Definition 13** A \( C_{AL} \)-algebra consists of a set \( D_L \), an operation \( \bullet : (D_L \times (\text{Var} \rightarrow D_L)) \rightarrow D_L \), and an interpretation \( \exists : \text{Const} \rightarrow D_L \), such that \( \exists(\#) = \perp_D \), \( \exists(S) = S \), and \( \exists(K_x) = k_x \) for all \( x \in \text{Var} \), where for all \( f, h, k \in (\text{Var} \rightarrow D_L) \):

1. \( k_x \bullet f \bullet h = f(x) \) for \( x \in \text{Var} \);
2. \( S \bullet f \bullet h \bullet k = f(x) \bullet k \bullet (x \rightarrow h(x) \bullet k) \in \text{Var} \);
3. \( \perp_D \bullet f = \perp_D \).

The denotation of terms under an anchor \( g : \text{Var} \rightarrow D_L \) is inductively obtained by:

1. \( \exists g^* = \exists (g) \);
2. \( X^* = g(X) \);
3. \( T(x, T, k)_{\text{eff}} = T^* \bullet (x \rightarrow T^* k)_{\text{eff}} \).

**Lemma 15** \( X \not\in F(\Gamma) \) \( \Rightarrow \) \( T^* \not\in F(\Gamma) \) in every \( C_{AL} \)-algebra. The proof theory for CALL looks like this.

**Definition 14** We write \( T \equiv_{C_{AL}} T' \) if \( T^* = T'^* \) in every \( C_{AL} \)-algebra. The proof theory of CALL looks like this.

**A1** \( T = T \);

**A2** \( T(x, T, \#)_{\text{eff}} = T(x, T, k)_{\text{eff}} \);

**A3** \( k_x (x, T, k)_{\text{eff}} = T, \) if \( j \in I \), otherwise \#;

**A4** \( S(x, T, k)_{\text{eff}} (y, T, k)_{\text{eff}} (z, T, k)_{\text{eff}} = T, \) if \( z = x_i \) for some \( i \in I \), otherwise \#;

**A5** \( \#(x, T, k)_{\text{eff}} = \# \);

**R1** \( T = T' \Rightarrow T' = T \);

**R2** \( T = T', T' = T'' \Rightarrow T = T'' \);

**R3** \( T = T', U_i = U'_i \) \( (i \in I) \) \( \Rightarrow T(x, U_i)_{\text{eff}} = T'(x, U'_i)_{\text{eff}} \);

**Theorem 16** (Soundness) \( T \equiv_{C_{AL}} T' \Rightarrow T \equiv_{C_{AL}} T' \).
2.3.1 Completeness for $C_{AL}$

**Definition 15** Let $[T] = \text{q} \{ T \} = T_{C_{AL}}$. We define the $C_{AL}$-term algebra $T_{C_{AL}}$ by $D_{AL} = \text{q} \{ T \} \in C_{AL}$, $\bot = \text{q} \{ \# \}$, $\exists (c) = \text{q} \{ c \}$ for all constants, and $[T] \cdot [x_i \to T_j]_{\text{q}} = T_j$ for any $x_i \to T_j$.

**Lemma 17** $T_{C_{AL}}$ is a $C_{AL}$-algebra.

**Proof:** By A2 and R3 the operation $\bullet$ is well-defined for $T_{C_{AL}}$. We need to check the axioms.

1. $[K_v] \bullet (x_i \to [T_j]_{\text{q}}) = K_v(x_i, T_j)_{\text{q}} = [T_j]_{\text{q}}$ if $i \in I$, else $\#$, where $K_v$ is a $C_{AL}$-algebra.

2. $[S] \bullet (x_i \to [T_j]_{\text{q}}) = S(x_i, T_j)_{\text{q}} = [T_j]_{\text{q}}$ for any $x_i \to T_j$.

3. $[\#] \bullet (x_i \to [T_j]_{\text{q}}) = [\#]_{\text{q}} = [\#]$.

$\square$

**Lemma 18** $\|T\| = T$, where $g(X) = [X]$ for $X \in \text{Par}$.

**Proof:** $\|T\| = [T]$ by definition. By induction, $\|T(x_i, T_j)_{\text{q}}\| = \|T\| \cdot (x_i \to [T_j]_{\text{q}})_{\text{q}} = T(x_i, T_j)$.

$\square$

**Theorem 19** (Completeness) $T \models_{C_{AL}} T' \Rightarrow T = T'$.

**Proof:** If $T \not\models_{C_{AL}} T'$ then $[T] \not\models [T']$ and thus $T_{C_{AL}} \not\models T$ by the previous lemma.

$\square$

2.3.2 Abstraction in $C_{AL}$-models

The operations $\bullet : (D_{AL} \times (\text{Var} \to D_{AL})) \to D_{AL}$ and $\Psi : D_{AL} \to ((\text{Var} \to D_{AL}) \to D_{AL})$ are interdefinable by $d \bullet f = \Psi(d)(f)$. The range of $\Psi$ are the representable functions of a given $C_{AL}$-algebra, written $((\text{Var} \to D_{AL}) \to D_{AL})$. For $\Lambda$-models these are just the continuous functions w.r.t. a given order. Now we have $s$ and the $K_v$ to make sure we have enough representable functions in order to be able to simulate abstraction. A function $h : D_{AL} \to D_{AL}$ is represented by any $d \in D_{AL}$ such that $\forall f \in \text{Var} \rightarrow D_{AL} \quad d \bullet f = h(f(c))$. We define $d \bullet h$ for $d \bullet (z \to b)$ if no confusion will arise.

Let $K = K_v$ and $\Pi_s = S \cdot K_v$. This gives us projections: $s \cdot (x_i, T_j)_{\text{q}} = [T_j]_i$ if $j \in I$, otherwise $\#$. We can now define terms that mirror abstractions as follows.

**Definition 16**

1. $\lambda^*(X_i \to x_i)_{\text{q}}(X_j) = \Pi_s$ if $j \in I$;
2. $\lambda^*(T) = K(T)$, if $d m(\{T\}) \cap FP(T) \equiv \emptyset$;
3. $\lambda^*(T) \cdot (x_i, T_i)_{\text{q}} = S(\lambda^*(T), (x_i, \lambda^*(T_i)))_{\text{q}}$,

if $d$ doesn’t apply.

**Lemma 20** $\|\lambda^*(T)\| = \|T\|^{\Pi_s}$.

**Lemma 21** $X \in dm(\{T\}) \Rightarrow X \not\in FP(\lambda^*(T))$.

Let applicative structures $\langle D_{AL}, \bullet \rangle$ be defined like $C_{AL}$-algebras without mentioning $s$ or any $K_v$, and $V$ range over terms not containing $S$ and the $K_v$. I call such expressions pure terms.

**Definition 17** An applicative structure is combinatorially complete if for any pure term $V$ and $\xi_1, \ldots, \xi_n$ with $FP(V) \subseteq \bigcup_{i=1}^n dm(\xi_i)$, there is a $d \in D_{AL}$ such that for all $f_1, \ldots, f_n$ and arbitrary $g$:

\[ d \bullet f_1 \bullet \ldots \bullet f_n = \|V\|^{\Pi_s} \cdot f_1 \bullet \ldots \bullet f_n.\]

**Theorem 22** Every $C_{AL}$-algebra is combinatorially complete.
CHAPTER 2. SIMULTANEOUS ABSTRACTION

PROOF: From the previous lemmata by taking $d =_{df} \| \lambda^* \xi_1 \ldots \lambda^* \xi_n V \|^p$. □

We now have ways to translate from $\Lambda_{AL}$ to $C_{AL}$ and vice versa.

Definition 18 $(\cdot)^{c_{AL}} : \Lambda_{AL} \rightarrow C_{AL}$ is defined by:

1. $X^{c_{AL}} =_{df} X$ for all parameters;
2. $c^{c_{AL}} =_{df} c$ for all constants;
3. $(t \langle x_1, t \rangle)^{c_{AL}} =_{df} t \langle x_1, (t \langle x \rangle)^{c_{AL}} \rangle$;
4. $(\lambda \langle x \rangle)^{c_{AL}} =_{df} \lambda^* \langle x \rangle^{c_{AL}}$.

Definition 19 $(\cdot)^{\Lambda_{AL}} : C_{AL} \rightarrow \Lambda_{AL}$ is defined by:

1. $X^{\Lambda_{AL}} =_{df} X$ for all parameters;
2. $(\langle x \rangle)^{\Lambda_{AL}} =_{df} \lambda(X \rightarrow x) \lambda() \cdot X$;
3. $s_{\langle x \rangle}^{\Lambda_{AL}} =_{df} \lambda Z \lambda(X \rightarrow x) \lambda(Y \rightarrow y) \lambda Z(y, Y) \cdot \langle x \rangle \langle y \rangle$, where $Z : \text{Var} \rightarrow \text{Var}$.
4. $e_{\langle x \rangle}^{\Lambda_{AL}} =_{df} e$ for all other constants;
5. $(T \langle x_1, T \rangle)^{\Lambda_{AL}} =_{df} T_{\langle x_1 \rangle}^{\Lambda_{AL}}$.

Theorem 23 $T^{c_{AL}} = T_{\Lambda_{AL}}$.

As in the case of the $\lambda$-calculus and standard CL we do not have $\tau = \tau' \Rightarrow \tau^{c_{AL}} = \tau_{\Lambda_{AL}}$. Given the completeness for $c_{AL}$, we would therefore not have soundness for $\beta$-equality if we interpreted $\Lambda_{AL}$-terms naively via their $C_{AL}$ translations. To interpret $\Lambda_{AL}$ by means of $C_{AL}$-algebras we need a canonical representative to be singled out for each representable function.

Definition 20 A $C_{AL}$-MODEL is a $C_{AL}$-algebra with an additional strict map $\Phi$ such that $D_{\perp}^{\perp} =_{df} \frac{1}{2} (\text{Var} \rightarrow D_{\perp}) \rightarrow D_{\perp}$, is a retraction and $\Phi \circ \Psi$ is represented by some $e \in D_{\perp}$.

CHAPTER 2. SIMULTANEOUS ABSTRACTION

Theorem 24 For any $C_{AL}$-model there is a unique, well-defined interpretation function $\| - \| : \Lambda_{AL} \rightarrow D_{\perp}$ such that

- $\| e \|^p = \exists(t)$;
- $\| X \|^p = \emptyset(X)$;
- $\| t(x_1, t)^{c_{AL}} \|^p = \emptyset(\| t \|^p) \cdot \langle x_1 \rangle \|^p$;
- $\| \lambda \langle x \rangle \|^p = \emptyset \lambda \cdot \| t \|^p$.

PROOF: We show by induction on $t$ that the following hold simultaneously:

1. $\| t \|^p$ is well-defined for all $t$;
2. $\| t \|^p = \| t \|^p$ if $X \notin FP(t)$;
3. For all $\xi_1 \ldots \xi_n$ with $FP(t) \subseteq \bigcup_{m=1}^n d_{m}(\xi_i)$, there is a $d \in D_{\perp}$ such that for all $f_1 \ldots f_n$ and arbitrary $y$:
   
   $d \cdot f_1 \ldots f_n = \| d \|^p \cdot f_1 \ldots f_n$.

- For pure terms, in particular for constants and parameters, 1 and 2 are trivial, or follow from combinatory completeness.

- For applications $t(x_1, t)^{c_{AL}}$ 1 and 2 are trivial. Let $d_1$ and $d_2$ satisfy 3 for $t$ and $t_1$. By the previous item the following object $b$ exists, where we write $\xi$ for $(x(X), Y)$, and suppose that $\tau \eta \langle x \rangle = \text{Var}$ and $d_{m}(\langle x \rangle) \cap d_{m}(\langle y \rangle) = \emptyset$ for $i \neq j = 1$ and the $X_i$ are fresh as well:

   \[ b =_{df} \big\| \lambda Y \lambda \langle x_1 \rangle \langle x_2 \rangle \ldots \lambda x_n \cdot X_1 \ldots X_n (x_1 X_2 \ldots x_n) \big\|^{\perp}. \]

Then $d = d \cdot d_1 \cdot (x_1 \rightarrow d_2)^{c_{AL}}$ satisfies 3.

- For abstractions $\lambda \langle x \rangle t$ let $d_1$ fulfill 3 for $t$ and $\xi_1 \ldots \xi_n \lambda x$, and let $f_i(x) = \emptyset(x_i)^p$ for $i = 1 \ldots n$. Then

   \[ a =_{df} d_1 \cdot f_1 \ldots f_n \]
represents $\lambda f[\|f\|^2]$. Hence $e \cdot a = \|\lambda x.t\|^2$. This proves 1. 2. For $3$ let $\chi_1 \ldots \chi_n$ be as above and $X, Y$ fresh. Let

$$t = \text{df} \|\lambda X.\lambda Y.\lambda \chi_1 \ldots \lambda \chi_n.X(Y(\chi_1 \ldots \chi_n))\| \cdot e \cdot d \cdot$$

Hence $t \cdot f_1 \cdot \ldots \cdot f_n = e \cdot (d_t \cdot f_1 \cdot \ldots \cdot f_n) = \|\lambda X, Y.t\|^2$.

\[\square\]

2.3.3 Completeness for $\Lambda_{AL}$

We write $A \vdash \phi$ if $\phi$ is true in every $C_{AL}$-model that satisfies all the equations in $A$. 

**Theorem 25 (Soundness)** $A \vdash t \Rightarrow A \vdash C_{LM} \ t = t'$. We define the $\Lambda_{AL}$-TERM model for $A$. $T_{AL}(A)$ is the $\Lambda_{AL}$-TERM model for $A$. $T_{AL}(A)$ is the $\Lambda_{AL}$-TERM model for $A$.

**Definition 21** Let $\phi \in \Lambda_{AL}$, then $\phi$ is valid in $\Lambda_{AL}$, if $\phi$ is true in all $C_{AL}$-models $\mathfrak{M}$ for all constants, and

- $\Psi([\phi]) = df [\phi] \cdot \forall x \in \mathfrak{M} \ \phi(x) = df \phi(x)$;
- $s = df [S]$;
- $\lambda x.\chi_1 \lambda x.\chi_n.X(Y(\chi_1 \ldots \chi_n))\| \cdot e \cdot d \cdot$;
- $\Phi(f) = df \cdot [f]$, if $\phi$ represents $f$.

**Lemma 26** $T_{AL}(A)$ is a $C_{AL}$-model satisfying $A$.

**Proof:** By $A4$ and $R3$ the operation $\Psi$, and thus $\Phi$, is well-defined on $T_{AL}(A)$. $\Phi$ is also well-defined: let $\phi$ and $\phi'$ both represent $f$, then $A \vdash \phi \Rightarrow A \vdash \phi'$.

We need to check the axioms for $C_{AL}$-algebras.

---

1. $k_{AL} \ (x_i \rightarrow [t_i])_A \ (y_j \rightarrow [t_j])_A \ (K_{AL})_A \ (x_i, t_i)_A = \ (y_j, t_j)_A$

2. $s \ (x_i \rightarrow [t_i])_A \ (y_j \rightarrow [t_j])_A \ (a) \ (c) = \ (d(a, c))_A$, for $a \in A$, and

3. $e \ (x_i \rightarrow [t_i])_A$, where $e = \ [t_i]$.

The equations in $A$ obviously hold. For the additional conditions on $C_{AL}$-models we have by definition that $\Phi \circ \Psi$ is represented by $e$. For strictness of $\Phi$ we note that $\phi \models [\phi]$ represents $A_{AL} \vdash A_{LM} \ t = t'$. Hence $\Phi(\phi \models [\phi]) = e \cdot [t]$, $\phi$ is well-defined for all $\phi$. Therefore, $A_{AL} \vdash A_{LM} \ t = t'$.

---

1. $k_{AL} \ (x_i \rightarrow [t_i])_A \ (y_j \rightarrow [t_j])_A \ (K_{AL})_A \ (x_i, t_i)_A = \ (y_j, t_j)_A$

2. $s \ (x_i \rightarrow [t_i])_A \ (y_j \rightarrow [t_j])_A \ (a) \ (c) = \ (d(a, c))_A$, for $a \in A$, and

3. $e \ (x_i \rightarrow [t_i])_A$, where $e = \ [t_i]$.

The equations in $A$ obviously hold. For the additional conditions on $C_{AL}$-models we have by definition that $\Phi \circ \Psi$ is represented by $e$. For strictness of $\Phi$ we note that $\phi \models [\phi]$ represents $A_{AL} \vdash A_{LM} \ t = t'$. Hence $\Phi(\phi \models [\phi]) = e \cdot [t]$, $\phi$ is well-defined for all $\phi$. Therefore, $A_{AL} \vdash A_{LM} \ t = t'$.

---

1. $k_{AL} \ (x_i \rightarrow [t_i])_A \ (y_j \rightarrow [t_j])_A \ (K_{AL})_A \ (x_i, t_i)_A = \ (y_j, t_j)_A$

2. $s \ (x_i \rightarrow [t_i])_A \ (y_j \rightarrow [t_j])_A \ (a) \ (c) = \ (d(a, c))_A$, for $a \in A$, and

3. $e \ (x_i \rightarrow [t_i])_A$, where $e = \ [t_i]$.
CHAPTER 2. SIMULTANEOUS ABSTRACTION

39

Theorem 28 (Completeness) $Ax \vdash_{CLM} t = t' \Rightarrow Ax \vdash_{\beta} t = t'$.

PROOF: If $Ax \vdash_{\beta} t = t'$ then $[t] \neq [t']$ in $T_\beta(Ax)$ and thus $\|t\| \neq \|t'\|$, setting $\beta(X) = [X]$ for all $X \in Par$. Therefore $Ax \vdash_{CLM} t = t'$.

Corollary 29 The finitary $\beta$-calculus is complete for finitary $\Lambda_{k\ell}$.

PROOF: If $\vdash_{CLM} t = t'$ then $\vdash_{\beta} t = t'$, and hence by Church-Rosser $t \rightarrow s \rightarrow t'$ for some $s$. But reduction of finite terms cannot produce infinite terms. Thus, there is a proof of $t = t'$ by final application of R1 and R2, which involves no infinite terms.

\[ \square \]

2.4 Structured Objects and Systems of Equations

Structured objects are distinguished from ordinary objects in that they can be uniquely decomposed into the component parts, from which they were put together. Ordered pairs are typical examples, forming an object from two objects $a, b$, such that there are projection functions $\pi_1, \pi_2$ with $\pi_1((a, b)) = a$ and $\pi_2((a, b)) = b$. The definition of ordered pairs in the $\lambda$-calculus will be generalised for our purposes here. I use the following operation, which allows us to build structured objects with any number of components.

\[ \llbracket (x_1, x_2, \ldots, x_n) \rrbracket_{k\ell} = \lambda y. (y(x_1, x_2, \ldots, x_n)) \quad (y \text{ fresh}) \]

Theorem 30

\[ \llbracket (x_1, x_2, \ldots, x_n) \rrbracket_{k\ell} = \llbracket (x_1, x_2, \ldots, x_n) \rrbracket_{k\ell} \]

PROOF: The $x_i$-component of the $i$th member is recovered by the projection

\[ \pi_{ni} = \lambda y. \lambda \{ \lambda (\ldots) \lambda (\ldots) \} \lambda (\ldots) \lambda (\ldots) \lambda (\ldots) \ldots \lambda (\ldots). \]

\[ \square \]

For ordered $n$-tuples we write $\llbracket (x_1, x_2, \ldots, x_n) \rrbracket_{k\ell}$, where $x_i$ is as usual the designated role to encode unary abstraction/application. By $s = 1$ one can encode assignments, which we write as $\llbracket (x, t_i) \rrbracket_{s\ell}$. The values of an encoded assignment are retrieved by projection functions $\pi_{x_i} = \lambda y. y(\llbracket (x_i, x) \rrbracket_{s\ell})$. We have

\[ \pi_{x_i}(\llbracket (x, t_i) \rrbracket_{s\ell}) = t_i \text{ for } i \in J. \]

Notice that for any $x \neq y$, if $\pi_x = \pi_y$ then for all $t, t'$: $t = \pi_y(\llbracket (x, t, t') \rrbracket_{s\ell}) = \pi_y(\llbracket (x, t, t') \rrbracket_{s\ell}) = t'$. Hence, in a non trivial model all projections $\pi_{x_i}$ must be different. We conclude:

Theorem 31 Any non trivial SAC model is infinite.

A particular demand of ST, at the focus of [Aczel and Lunnun 1991], is to have non well-founded objects, obtained from cyclic sets of equations.\(^2\) To solve systems of equations $(x_i = t_i)_{s\ell}$ we use the above encoding of assignments. We say that an (encoded) assignment $s$ SOLVES A SYSTEM OF EQUATIONS $(x_i = t_i)_{s\ell}$ iff for all $i \in I$ $\pi_{x_i}(s) = \llbracket \pi_{x_i}(s) / x_i \rrbracket_{s\ell} t_i$.

Theorem 32 Every system of equations has a solution.

PROOF: Recall that $Y = \lambda f. (x, f(x))((x, f(x)))$ is a fixed point combinator with $t(Tt) = (Tt)$ for any term $t$. Define a system of equations $(x_i = t_i)_{s\ell}$ the function $F = \lambda z. \llbracket x_i [\pi_{x_i}(z) / x_i]_{s\ell} t_i \rrbracket_{s\ell}$. Then $(T F)$ solves the system, as $\pi_{x_i}((T F)) = \pi_{x_i}(F(T F)) = \pi_{x_i}(\llbracket x_i [\pi_{x_i}(F(T F) / x_i]_{s\ell} t_i \rrbracket_{s\ell}) = [\pi_{x_i}(T F) / x_i]_{s\ell} t_i$.\(\square\)

As things stand we do not necessarily have unique solutions to such equations. It may be consistent to assume uniqueness at least for some kinds of systems, to get closer to the original aims of Aczel and Lunnun.

\(^2\)Notice that this holds for finitary versions of the SAC as well.

\(^3\)An example of a cyclic equation would be $x = (x, x)$. Any solution of which must provide a non well-founded pair whose first component is $x$ itself.

In the following, we need to assume that at least three variables do not occur in such a system, to allow us to form the terms below without accidentally binding some of the $x_i$ in them.
Chapter 3

Semantic Theories

I now show how the SAC can provide a framework for specifying semantic theories, especially recent ones, which do not easily fit into the traditional $\lambda$-calculus. Perhaps the key idea for semantics, as conceived of in this thesis, is to think of properties as functions from assignments to propositions. I will use the term $\omega$-properties to distinguish them from ordinary functions which take entities into propositions.

In the following I will freely mix the languages $\Lambda_{AL}$, $\Lambda$, and $\lambda$, according to whatever notation seems most convenient. The reader is expected to switch between the systems according to the translations given in section 2.2.1, though it will not be necessary to mentally translate every term into $\Lambda_{AL}$, I hope. To the contrary, the use of $\lambda$ and $\lambda$ should help to avoid tedious complications, which are the price we pay for $\Lambda_{AL}$'s simple proof theory.

The following section on Frege Structures will be presupposed in all three semantic theories to be discussed, but otherwise those sections can be read relatively independently.

3.1 Frege Structures

For semantics, as I understand it, we need a theory of truth for the propositions that we hope to denote by our terms. The usual way to approach such a theory is to give an inductive definition of truth-in-a-model for terms of the right type, where the notion of a model is suitably enriched to fix the base cases. In $\lambda$ and similar systems for example we find a basic type of truth values (usually two), and the theory of truth is given as conditions that determine which truth value is denoted by a complex formula, depending on its parts. The matter is complicated by the presence of a number of parameters with respect to which a formula is evaluated, most prominently an index for possible worlds. The whole setup leans towards building theoretic assumptions into the system via modeling choices and implicit consequences of certain definitions, rather than taking things at face value and saying explicitly what we assume to hold about them.

My aim here is to provide a flexible framework in which different semantic assumptions can be expressed and tried out. It is thus a good idea to look elsewhere, namely to Property Theories, for ways to avoid prejudging too many issues before we can even begin to put any of the ideas we have produced into practice. The kinds of things I want to assume for the envisaged framework are these: some terms denote propositions, and propositions are the things that are true or false. $\omega$-properties are things that form propositions when applied to assignments of objects to certain variables, which we think of as argument roles for those $\omega$-properties. The notion of truth does not primarily apply to formal expressions, in relation to various parameters of evaluation, but to propositions, which are things that need no further fixing of any such parameters. This does not prevent us from introducing such parameters into our framework, but rather allows us to do so without changing the theory of truth: if an expression has to be evaluated with respect to certain entities, worlds, situations, time intervals etc., it simply does not express a proposition, but an incomplete object, which needs some open argument places filled.

I do not claim that this approach is the only possible way of making use of the theory of simultaneous abstraction developed in the preceding chapter, but merely that this way of proceeding delivers reasonable versions of (the underlying logics of) DMG, DRT and ST in a simple, economical, and philosophically sound way. The addition of new structure in this framework, such as for example a theory of situations and states of affairs into a Montagovian or Kamp style semantics, does not require us to revise the basic axioms of truth for DMG or DRT, which in turn are perfectly compatible extensions of the theory of Proposition
CHAPTER 3. SEMANTIC THEORIES

Structures. What has to be revised in switching from one system to another are axioms about the lexical meaning of words, and possibly other aspects of the syntax-semantics interface, but not the fundamental apparatus of abstraction, application, propositions, and truth.

Let me then introduce a formal first order language in which to state such a theory of truth. One might do the job in an informal model theoretic way along the lines of [Aczel 1988] but I will make things look more like the formal treatment of [Turner 1999], without attaching any philosophical importance to this. In line with our convention to use $\text{AL}$ as our official term language I regard its parameters as variables $v, v_1, v_2, \ldots$ of $\text{FOL}_{\text{AL}}$, defined below.\(^3\)

**Definition 22** We define WFFs of $\text{FOL}_{\text{AL}}$ over a set of $n$-ary predicates $R_i$:

- if $t_1, \ldots, t_n, t \in \text{AL}$ then $R_i(t_1, \ldots, t_n), t = t \in \text{WFF}$,
- if $\psi, \phi \in \text{WFF}$ then $\neg \psi, \phi \land \psi, \exists \psi \in \text{WFF}$.

Other connectives and quantifiers are defined as usual. A model $M$ for $\text{FOL}_{\text{AL}}$ is a $\text{AL}$-model with an interpretation $\exists$ for the predicates, such that the usual conditions obtain:

- $M, \beta \models R_i(t_1, \ldots, t_n)$ iff $\langle t_1^\beta, \ldots, t_n^\beta \rangle \in \exists(R_i^\beta)$;
- $M, \beta \models t = t'$ iff $\langle t^\beta \rangle = \langle t'^\beta \rangle$;
- $M, \beta \models \phi \land \psi$ iff $M, \beta \models \phi$ and $M, \beta \models \psi$;
- $M, \beta \models \neg \phi$ iff $M, \beta \models \phi$;
- $M, \beta \models \exists \psi$ iff $\exists \psi$.

\(^3\) It has technical advantages though, as this allows us to use FOL's well-understood proof theory to show that a (finite) term $t$ implies another (finite) term $t'$ relative to a particular theory of truth. For this, one would have to derive $T(t')$ in FOL from $T(t)$ plus the axioms of truth and ($\exists$-) identity. It seems possible to prove a completeness theorem for this.

We single out a unary relation $T$ as a truth predicate. We define $F(t) =_T T(L(t)$ (falsity) and $P(t) =_T T(t) \lor F(t)$ (being a proposition). For dealing with quantification we define $PTY^n(t) =_T \forall v_1, \ldots, v_n F(t(v_1, \ldots, v_n))$ (being an $n$-place property).

Logical operations in our calculus are treated as $\text{AL}$-constants whose semantic behavior is captured by a theory of truth.\(^3\) The following logical constants are singled out: $\top, \bot, \supset, \land, \lor, \neg, \forall, \exists, \Delta$, intended to be the operations of conjunction, disjunction, negation, implication, existential and universal quantification, equality, and structured predication.\(^4\) I will take $\top, \bot, \forall, \exists, \Delta$ as primitive, and the others to be defined in terms of these.\(^5\) I assume the logical combinations, other than $\Delta$, to have one or two distinguished argument roles $z_1, z_2$. One can of course carry them for standard Montaguean techniques to run smoothly.\(^6\) In any case I write them in the conventional form, $t \supset t$ for $\forall (z_1, t, z_2, t)$, $\Sigma (z, \lambda x, t)$ etc. and $(x_i t_i)_{i\in I}$ for $\Delta (z, t, (x_i t_i)_{i\in I})$, to enhance readability.

**Definition 23** 1. A **Proposition Structure** is a model of $\text{FOL}_{\text{AL}}$ satisfying

- $P(t \equiv t')$ \land \langle $T(t \equiv t')' \equiv t \equiv t' \rangle$\)
- $P(t) \rightarrow (P(t \equiv t') \land \langle T(t \equiv t') \equiv \neg T(t) \rangle)$\)
- $P(t) \land P(t') \rightarrow (P(t \equiv t') \land \langle T(t \equiv t') \equiv t \equiv t' \rangle)$\).

2. A **Proposition Structure with Predication** has the additional truth-axiom

- $P(t(x_i t_i)_{i\in I}) \rightarrow (P((x_i t_i)_{i\in I}) \land \langle T(x_i t_i)_{i\in I} \equiv T(t(x_i t_i)_{i\in I}) \rangle)$.\)

\(^3\)These constants are not to be confused with the logical symbols of $\text{FOL}_{\text{AL}}$, which we use to express our theories of truth.

\(^4\)The SAC, in contrast to Aczel and Lunnon's theory, does not take the idea of structured objects as a starting point. I take the notions of abstraction and application as fundamental here, rather than structure preserving predication. It is crucial that these two notions of application/predication are kept distinct. Occasionally forms of Situation Theory have been guilty of conflating them, when they assumed "infons" to be composed out of a unique relation and assignment plus relations to be in-situabstracts which apply to assignments to form structured infons. This implies that $r(c, t, t_i) = (c, t_i, c')$ for all relations $r$, which quickly leads to the triviality of the model, by the argument given in [Aczel 1989].

\(^5\)There may be reasons not to do this, such as to avoid unintended identities in attitude contexts, or to be able to use the less symmetric notion of "strong" implication in [Aczel 1988].

\(^6\)Avoiding all forms of currying is useful if the predicative axioms is to be fully exploited.
plus the axiom of structure preservation

\[ ([x, t] \in \Delta t) = ([x, t'] \in \Delta t') \rightarrow t = t' \land \bigwedge_{i \in \Delta t} t_i = t_i'. \]

3. A Frege Structure is a Proposition Structure satisfying

\[ PTY^3(t) \rightarrow \left( P\left[\Sigma t\right] \land (T\left[\Sigma t\right] \leftrightarrow \exists v T(t(v))) \right). \]

4. I speak of a strong Proposition / Frege Structure (with predication) if we have biconditionals for propositionhood, that is P(t) \iff P(\bar{t}), P(t) \land P(t') \iff P(t \cap t'), P(t_i t_i t_i) \iff P((i, \bar{t}) \Delta t), and PTY^3(t) \iff P(\Sigma t).

The class of propositions guaranteed to exist by these axioms is rather restricted, in that the relations which have an internal representation by some property only involve equality and logical constants. In semantic applications one needs a much larger stock of basic properties. So it is useful to consider the addition of further predicates to FOLSA, whose extensions may be fixed in arbitrary ways. The question arises whether these extensions also correspond to properties that are internally represented in the Frege Structure, and thus are denotable by a (possibly newly introduced) term.

**Definition 24** A term \( t \) internally defines a predicate \( \phi_{t_1 \ldots t_n} \), with free variables \( t_1, \ldots, t_n \notin \text{FV}(t) \), in some FOLSA-model, if:

\[ PTY^m(t) \land \forall t_1 \ldots t_n \ T(t_1 \ldots t_n) \iff \phi_{t_1 \ldots t_n}. \]

We may apply this terminology to denotations as well, and say in such cases that a particular relation is internally defined by an \( n \)-place property. As Aczel pointed out in [Aczel 1988] there are limits to what can be internally defined in a Proposition Structure. For example the truth predicate \( T \) cannot be internally defined, for it would lead to the construction of a liar proposition.\(^6\)

### 3.1.1 Consistency

I more or less repeat the proof of the existence of Frege Structures from [Aczel 1988] here, in a somewhat sketchy manner.

Take any non-trivial model for the SAC. To get a model for our axioms we need a set \( T \) that interprets \( T \), such that the axioms are true. The other predicates, like \( P, F \), are defined in terms of \( T \), but it is more convenient to work with pairs \((P, T)\) of subsets \( T \subseteq P \), called PROPOSITIONS and TRUTHS, to interpret \( P \) and \( T \). Such pairs of subsets form a cpo as follows, where \( F = \{P \setminus T \}\) are the FALSES:

\[ (P, T) \subseteq (P', T') \iff T \subseteq T' \text{ and } F \subseteq F'. \]

Lub's of ascending chains are given by set union. Any monotone operator over a cpo has a least fixed point. We define a monotone operator \( \delta \) whose fixed points give us FOLSA-models for the Frege Structure axioms when \( \exists \{T\} = T \).

As any non-trivial model is infinite we can select for each logical constant \( \delta \) semantically distinct terms \( g_k \) such that we can use structured objects to avoid possible overlaps between the results of applying different logical constants.\(^6\)

For binary \( \delta \) we assign the following interpretation:

\[ d_\delta = \delta \| \lambda \{x, y\}. \| \{r; x, y\} \| \}

Similarly for unary ones. Predication is encoded as

\[ d_\Delta = \Delta \lambda Z \lambda X \lambda Y \lambda Z \lambda X. \lambda y. \| \{r; \Delta; Z, Y, X, y\} \]

\( ^6 \)Let \( \delta \) be a term that solves the equation \( \delta = \lambda x. \lambda y. \delta \), and assume that true internally defines \( T \). Then \( \delta \) and \( \delta \) are not properties, as long as for all \( \delta \) which \( P(\delta) \) is true, \( \delta(\delta) \) is a proposition.
CHAPTER 3. SEMANTIC THEORIES

Let \( d_{i,j,k} \) stand for \( d_i \cdot f \cdot h \) if \( i = \Delta \), otherwise \( d_i \cdot f \). The truth-axioms \( \text{Ax}_T \) for a logical operator \( d_i \) can be expressed semantically with reference to a proposition condition \( \text{PC}(T, f, h) \) and a truth condition \( \text{TC}(T, f, h) \) as

\[
(\text{Ax}_T) \quad \text{PC}(T, f, h) \Rightarrow (d_{i,j,k} \in \mathcal{P} \land (d_{i,j,k} \in T \Rightarrow \text{TC}(T, f, h)))
\]

For example, the proposition condition for \( \mathbf{\&} \) is \( f(z_1) \in \mathcal{P} \) and \( f(z_2) \in \mathcal{P}' \), and the truth condition for \( \mathbf{\&} \) is \( f(z_1) \notin \mathcal{T} \).

**Definition 25** \( \theta_P (\mathcal{P}, T) =_{df} (\mathcal{P}', T') \), where for the logical constants \( i \) of Frege Structures:

- \( \mathcal{P}' \) is the smallest set such that for each \( i, f, h \):
  \[ \text{PC}(\mathcal{P}, f, h) \Rightarrow d_{i,j,k} \in \mathcal{P}' \]

- \( T' \) is the smallest set such that for each \( i, f, h \):
  \[ \text{PC}(\mathcal{P}, f, h) \land \text{TC}(T, f, h) \Rightarrow d_{i,j,k} \in T' \]

**Lemma 33** Any fixed point for \( \theta_P \) satisfies \( \text{Ax}_F \).

**Proof:** Let \( \theta_P (\mathcal{P}, T) = (\mathcal{P}, T) \).

(i) If \( \text{PC}(\mathcal{P}, f, h) \) then \( d_{i,j,k} \in \mathcal{P} \).

(ii) If \( \text{PC}(\mathcal{P}, f, h) \) and \( \text{TC}(T, f, h) \) then \( d_{i,j,k} \in T \).

(iii) Given \( \text{PC}(\mathcal{P}, f, h) \) and \( d_{i,j,k} \in T \) there exist \( \mathcal{P}', f', h' \) with \( d_{i,j,k} = d_{i,j,k} \), and \( \text{TC}(T, f', h') \). But then \( \eta = \eta \) and \( f = f' \), and in case \( i = \Delta \) also \( h = h' \), and hence \( \text{TC}(T, f, h) \).

By the fact that \( \text{Ax}_F \) is satisfied, we see that \( \mathcal{P} \) is the extension of the predicate \( P \) as it was defined, and thus the axioms \( \text{Ax}_F \) hold if and only if the formal axioms for Frege Structures are true. Notice that the reverse direction of (i) is true as well, giving us in fact a strong Frege Structure. The fact that predication is structure preserving is obvious. So we conclude:

**Theorem 34** Every non trivial SAC-model can be extended to a strong Frege Structure with Predication.

---

As Aczel has shown in [Aczel 1980], it is consistent to add infinitely many axioms for the internal definition of relations to the theory of Frege Structures, as long as the extensions of the possibly newly introduced predicates \( \phi, \ldots \), in the SAC-model are given in advance, without reference to \( T \). In the construction of the Frege Structures they are treated just as we treated identity, which is already internally defined.

**Theorem 35** Every non trivial SAC-model can be extended to a strong Frege Structure with Predication in which a countable number of previously given relations are internally defined by no-place properties.

### 3.2 Dynamic Montague Grammar

There are several versions of DMG on the market, which use different 'dynamic' relatives of Montague's original logic II. One such logic, namely the Dynamic Property Theory (DPT) of [Chierchia 1994] is particularly close to the system \( A \) of the preceding chapter. The additions we have to make to obtain DPT are very small indeed, as \( A \) was first developed as a generalization of DPTs limited means of simultaneous abstraction and application.

In DPT, the variables are divided into two kinds: the **discourse markers** \( DM \), and the remaining ones, which I call **meta variables** \( MV \). I use Greek letters \( \alpha, \beta, \ldots \) for meta variables, dotted letters \( x, y, \ldots \) for discourse markers, and \( x, y, \ldots \) for variables that can be of either kind. The sorting of the variables is not a matter of the kinds of denotations they take. It is rather a matter of getting the two crucial operators of the system, \( " \) abstraction and \( " \) application, to operate on a certain set of variables, namely the discourse markers only. We need to protect the other variables, i.e. the meta variables, from the influence of \( " \) abstraction so that we can bind variables inside the scope of \( " \), as will become evident from the examples later on.

The "dynamic" part of DMG can be captured easily by abstracting and applying to the infinite set of discourse markers via the following definitions:
CHAPTER 3. SEMANTIC THEORIES

- $^n t = \lambda D M. t$, and
- $^u t = \lambda (x . x^2)_{EDM}$.

All other abstractions and applications will be the traditional unary ones.\(^{10}\)

We use them to get a fairly standard system that does not require us to rethink the setup of DMG. Let me call DPT-terms any expressions built from unary abstraction and application and the $^u$ and $^n$ operators.

Convenient as DPT-terms may be for the purposes of DMG, one would expect them to make trouble in the proof theory, given the incompatibilities of $\Lambda$ in that respect. On the positive side, notice that every discourse marker is bound in a term of the form $^n t$, and occurs free in $^u t$. We thus can $\beta$-convert a term $\lambda x. t (^n x)$ even if some discourse markers of $t$ end up in the scope of an abstraction in $t$. Discourse markers are not free in $^u t$ and hence do not become bound by a new operator inside $t$. Another reduction we use in $\Lambda$ beside the more standard $\beta$-reductions is $^u n t = t$.

On the negative side, there are DPT-terms which cannot be reduced to simpler DPT-expressions, although their translations into $\Lambda_{AL}$ may be contractable into new $\Lambda_{AL}$-terms which are not translations of any DPT expressions. An example would be a term $\lambda x. t(\langle x \rangle)$ which translates to $\lambda (x^2 \rightarrow \langle x \rangle) \lambda (x^2 \rightarrow \langle x \rangle)_{EDM}. (x^2) (\alpha . x^2)$. As a DPT-term $\lambda x. t(\langle x \rangle)$ is not reducible. In particular it is not semantically equivalent to $^n x$, which means the same thing as the $\Lambda$-term $\langle x \rangle x$. The translated term on the other hand boils down to $[x^2/\alpha] \lambda (x^2 \rightarrow \langle x \rangle)_{EDM}. (x^2) \lambda (\langle x \rangle)_{EDM}$, which is the translation of the $\Lambda$-term $\langle x \rangle x$, but not of any DPT expression. Luckily the terms actually used in DMG do not belong to the problematic sort. I will think of DPT-expressions as convenient shorthands for the underlying $\Lambda_{AL}$-terms, for which I have already given a complete proof theory in 2.2.2, and a calculus for DPT as being obtained via the standard translation given in section 2.2.1. For computational applications this stance is of course impractical, as the terms involving $^u$ and $^n$ will be infinite, and thus in real life something closer to actual DPT-expressions will have to be used. But for an understanding of how these expressions have to be handled the translation into $\Lambda_{AL}$ still is instructive.

The sentence ‘A man walks in’ can now be rendered in our notation as

$$\lambda x. \Sigma x (\langle x \rangle \cap \text{walkin}(\langle x \rangle) \cap \langle x \rangle)$$

which can applied to $^n \text{whistle}(\langle x \rangle)$ (‘he whistles’) to yield

$$\Sigma x (\langle x \rangle \cap \text{walkin}(\langle x \rangle) \cap \text{whistle}(\langle x \rangle))$$

The important action happens at the level of conjunction of $\omega$-properties, which is achieved by abstracting over ‘possible continuations’ of the discourse.\(^{11}\)

### 3.2.1 Some DMG Translations

I only give a short sketch here and refer the reader to [Chierchia 1994] and [Chierchia 1992] for a more complete treatment. Sentence denotations in DMG are not propositions, but context change potentials (ccp’s), where a ccp is a function from $\omega$-properties to propositions. As a notational convention I write $^n t$ as $[t]$. Two operations are handy to convert propositions into ccps: and $\downarrow t = \lambda (t \cap \langle x \rangle)$ and $\uparrow t = \lambda (t \cap \langle x \rangle)$. I assume the scope of operators to be narrowest for $^u$, and increasingly wider for application (bracketing to the left), $\cap$, $\uparrow$, $\downarrow$, and unary $\lambda$-abstraction.

Ccps allow for clever ways of leaving holes within quantified structures to be filled by later material in the discourse so that one can use functional composition $t; t' = \lambda x. t ([t' x])$ for discourse sequencing:

\[ ^n [a \text{ man}] \text{ walks in.} \] $^u [\text{ he}]$ $\text{whistles.}$

\[ (\lambda x. \Sigma x (\langle x \rangle \cap \text{walkin}(\langle x \rangle) \cap \langle x \rangle)) \cup \rightarrow x. \Sigma x (\langle x \rangle \cap \text{walkin}(\langle x \rangle) \cap \text{whistle}(\langle x \rangle)) \]

\[ ^{10} \text{As an alternative to taking these unary operations as defined in terms of } \Lambda_{AL} \text{ or } \Lambda \text{-expressions, one could introduce them as basic operations, and give them a semantics in an expanded domain } D ^{\alpha} ((D \times D) \rightarrow D) + (D \rightarrow D), \text{ in the way [Chierchia 1994] does.} \]

\[ ^{11} \text{Lambda conversion on } ^n \text{whistle}(\langle x \rangle) \text{ is only one way to make the conjunction happen. Universalization of } \text{whistle}(\langle x \rangle) \text{ with the unabstracted variable } \alpha, \text{ as in UCG, see [Zeevat 1991], is just as intuitive for this. Hence the choice of the term 'meta variables', reminiscent of its use in UCG.} \]
The flexibility of CCG’s is shown by various possibilities to define dynamic kinds of universal quantification for donkey sentences. Here is an example derivation of the “strong” reading for them. To make things concrete I use a basic, purely applicative Categorial Grammar, operating on sequences Γ of (indexed) strings, and category-meaning pairs A : t . We have lexical axioms, plus the slash elimination rules (\langle E \rangle) and (\langle E \rangle)^{12}$, interpreted by “intensional” function application.

$$(E) \text{ if } \Gamma_{1} \Rightarrow A \mid B : t_{1} \text{ and } \Gamma_{2} \Rightarrow B : t_{2} \text{ then } \Gamma_{1}\Gamma_{2} \Rightarrow A : t_{1} \mid t_{2}$$

$$($$E$)$ if $\Gamma_{1} \Rightarrow B : t_{1} \text{ and } \Gamma_{2} \Rightarrow B \mid A : t_{2} \text{ then } \Gamma_{1}\Gamma_{2} \Rightarrow A : t_{2} \mid t_{1}$

$^{12}$[\every] $\Rightarrow \ (s/f(n\backslash x)) / en : \lambda \alpha \lambda \beta . \uparrow \Pi \alpha \downarrow ( (\alpha \downarrow \downarrow (\beta \gamma \gamma) )$ $\text{‘man’} \Rightarrow \ (en \backslash en) / (s/f(n \backslash x)) : \lambda \alpha \lambda \beta \lambda \delta \lambda \gamma \gamma . (\uparrow \alpha \uparrow \lambda \gamma \uparrow (\delta \gamma)) \ (\gamma)$$

$^{12}$[\every] $\Rightarrow \ s/f(n \backslash x) : \lambda \beta \lambda \delta \lambda \gamma \gamma . (\uparrow \lambda \gamma \uparrow (\delta \gamma)) \ (\gamma)$$

$^{12}$[\every] $\Rightarrow \ s/f(n \backslash x) : \lambda \beta \lambda \delta \lambda \gamma \gamma . (\uparrow \lambda \gamma \uparrow (\delta \gamma)) \ (\gamma)$

$^{12}$[\every] $\Rightarrow \ s/f(n \backslash x) : \lambda \beta \lambda \delta \lambda \gamma \gamma . (\uparrow \lambda \gamma \uparrow (\delta \gamma)) \ (\gamma)$

An alternative to the above “strong” reading of the universal, would assign the translation:

$^{12}$[\every] $\Rightarrow \lambda \lambda \lambda \beta \beta \beta \beta \lambda \beta \beta \beta \beta \lambda \gamma \gamma \gamma \gamma . (\uparrow \lambda \gamma \uparrow (\delta \gamma)) \ (\gamma)$

resulting in the “weak” reading:

$^{12}$[\every] $\Rightarrow \lambda \lambda \lambda \beta \beta \beta \beta \lambda \beta \beta \beta \beta \lambda \gamma \gamma \gamma \gamma . (\uparrow \lambda \gamma \uparrow (\delta \gamma)) \ (\gamma)$
3.3 Discourse Representation Theory

The SAC mirrors very closely the use of discourse referents in DRSs. The system I present in this section is inspired Robin Cooper's Situation Theoretic DRT, [Cooper 1993] where he uses EKN with Aczel-Lummon abstraction to similar effect. As in Cooper's treatment, this version of DRT is not quite faithful to standard versions, as it imposes a need for some additional abstractions at the right level. This is because only abstracted variables are linked to their antecedents, while free variables cannot in this system be 'captured' in the way one has come to expect. I won't go into the resulting complications for the syntax-semantics interface here, as I will explore alternatives which use free variables more prominently in the next chapter.

3.3.1 Zeevat's DRT

First, let me briefly sketch Henk Zeevats compositional semantics for DRT, see [Zeevat 1989], which is an important reference point for the versions to be presented in this and the next chapter.

For the usual DRT syntax, we start with atomic conditions $\alpha$, which are just atomic first order formulae. DRSs and complex conditions are defined recursively. DRSs are pairs $(U, C)$, where $U$ is a set of variables, and $C$ a set of conditions; the connective $\supset$ forms implicational conditions $(U, C) \supset (U', C')$ from two DRSs.13

Atomic conditions have their expected semantics. DRSs, and complex conditions are treated as standardly in DRT, where $g \Leftrightarrow f$ means $g$ and $f$ disagree on most of the variables in $M$.

- $\llbracket C \rrbracket =_{df} \{ \forall x \in C : x \in \llbracket \alpha \rrbracket \}$
- $\llbracket (U, C) \supset (U', C') \rrbracket =_{df} \{ \forall f \Leftrightarrow g : f \land g \in \llbracket C \rrbracket \rightarrow \exists h : h \Leftrightarrow f \land h \in \llbracket C' \rrbracket \}$

This system Zeevat adds a syntactic operation $\circ$, for "merging" two DRSs into one:

$\langle \bullet \rangle$ \quad $(U, C) \circ (U', C') =_{df} (U \cup U', C \cup C')$.

DRSs are given a denotation in terms of pairs,

$\llbracket (U, C) \rrbracket =_{df} \llbracket U \rrbracket \llbracket C \rrbracket$,

whose first and second component are noted as $\llbracket U \rrbracket_1$ and $\llbracket U \rrbracket_2$ respectively. He has shown that this makes $\llbracket \cdot \rrbracket$ a homomorphism between the syntactic and the semantic algebra of DRSs. In particular, with respect to $\circ$ we have:

$\langle \ast \rangle$ \quad $\llbracket (U, C) \circ (U', C') \rrbracket = \llbracket (U, C) \rrbracket_1 \llbracket (U', C') \rrbracket_2 \llbracket (U, C) \rrbracket_2 \llbracket (U', C') \rrbracket_1$.

Notice that there is no unique decomposition of DRSs with respect to $\circ$. In a more 'linear' format, similar to Zeevats BL,14 one can have expressions of the form $D R S \supset D R S$, which mention $\supset$ explicitly and thus uniquely decompose along this operation. The fact $\langle \ast \rangle$ then inductively defines the denotation of such DRSs. My systems will operate along the latter lines. Given such a syntax, interpreted by $\langle \ast \rangle$, we get the following "merging principle" to hold:

$\langle MP \rangle$ \quad $\llbracket (U, C) \circ (U', C') \rrbracket = \llbracket (U \cup U', C \cup C') \rrbracket$,

which semantically reflects Zeevat's earlier syntactic definition in $\langle \bullet \rangle$.

Let me try to bridge the gap to our framework a bit further. Syntactically, sets of conditions may be replaced by conjunctions. Semantically, the sets of assignments $\llbracket C \rrbracket$ can be turned into functions from assignments to truth, and finer grained propositions should be put in the place of truth values in order to get closer to our highly intentional approach.

There is of course an obvious difference to be observed. The SAC has an abstraction operation, which gives us terms $\lambda u.C'$ rather than pairs $(U, C)$, denoting

---

13I will ignore negation here.

14See section 1.3, in [Zeevat 1989].
functions (under assignments), rather than pairs. In contrast to DRS formation, we can iterate abstractions indefinitely, and we have a simple notion of application at our disposal. The downside is a certain loss of structure, to be faced in the SAC, as the denotation \( \llbracket M \rrbracket \) may be identical to \( \llbracket UC \rrbracket \), although \( U \neq \llbracket U \rrbracket \).

There is no way of recovering the set of variables \( U \) that was abstracted over from the denotation of a term, while it can be read off from Zeev’s \( \llbracket (U, C) \rrbracket \) as the first component. This fact allows him to generate the correct interpretations for \( \exists \) and \( \& \) in a compositional way. Thus, he has to find other ways to semantically identify certain sets of variables as the universes in the denotations of DRSs, or we need to introduce more structure of the suitable kind into the semantics for our calculus. I explore the first option here, using a suitable notion of \( \omega \)-properties. In the next chapter I pursue the route of changing the semantics of the SAC itself, towards a more fine-grained meaning for abstraction, which draws on Zeev’s ideas even more. The resulting system of \( \Lambda \)-DRT does seem to be sufficiently attractive to justify such a move.

### 3.3.2 \( \omega \)-Properties

Let us regard the universe of a DRS as the set of argument roles of an \( \omega \)-property. We start by defining a propositional function in the argument roles \( x_1, \ldots, x_n \) as follows:

\[
P F(x_1, \ldots, x_n)(t) = \forall x_{1:n} \bigwedge_{i=1}^n v_i \not= \# \rightarrow P(t(x_1, x_2, \ldots, x_n, v_n)) \quad (v_i \ fresh).
\]

This is not a sufficiently strict notion of \( \omega \)-properties to render Kamp Structures consistent. We need a notion of \( PTY^M \) such that the set of argument roles \( M \) is uniquely determined for any given \( \omega \)-property. The predicate \( PF^M \) does not possess this property as abstracts are insensitive to overdefined assignments:

\[
\models \{x_1, x_2\} t(x_1, x_2, l, k) = \{x_1, x_2\} t(x_1, x_2, l_{\{\}} k_{\{\}}).
\]

How strict a notion of \( \omega \)-property one ultimately wants remains to be seen. For now I take it to imply that the set of argument roles \( M \) of an \( \omega \)-property is the smallest set of roles in which it is a propositional function, and that it is a propositional function in every set of roles containing \( M \).

**Definition 26** We define the notion of \( \omega \)-property in the argument roles \( M \neq \emptyset \):

\[
PTY^M(t) =_{df} \bigwedge_{M \in \mathbb{C}} PF^M(t) \land \bigwedge_{M \in \mathbb{C}} \neg PF^M(t).
\]

We introduce a new notion of internal representation of relations, which reflects the definedness requirements for the argument roles of \( \omega \)-properties. A relation \( \phi \) is internally defined by the \( \omega \)-property \( t \) if for some \( x_1, \ldots, x_n \):

\[
PTY^M(x_1, \ldots, x_n)(t) \land \forall v_1, v_n \bigwedge_{i=1}^n v_i \not= \# \rightarrow T(t(x_1, v_1, \ldots, x_n, v_n) \oplus \phi_{v_1, v_n}).
\]

### 3.3.3 Kamp Structures

Let us now try to give a theory of \( \omega \)-properties and propositions and truth that is capable of dealing with basic ideas of DRT. We add to the basic logical operations \( \bigvee, \bigwedge, \exists \) a new quantificational operator, the binary conditional \( \bigtriangledown \).

**Definition 27** A Kamp Structure is a Proposition Structure which satisfies the following axiom, where the \( v_i \) are all fresh:

\[
PTY^M(x_1, \ldots, x_n)(t) \land PTY^M(x_1, \ldots, x_n)(s) \rightarrow P(t \bigtriangledown s) \land T(t \bigtriangledown s) \oplus \exists v_{1:n} T(t(x_1, v_1, \ldots, x_n, v_n) \bigtriangledown s) \land \forall t \bigtriangledown s)
\]

The notion of a strong Kamp Structure is as expected. Letting \( \Xi =_{df} \lambda X. \bigwedge \{p \in \Lambda p \mid \text{for some proposition } p, \text{obtain the unselective existential quantifier that is needed to reflect the truth conditions of non-embedded DRSs:} \}

\[
PTY^M(x_1, \ldots, x_n)(t) \rightarrow P(\Xi t) \land T(\Xi t) \oplus \exists v_{1:n} T(t(x_1, v_1, \ldots, x_n, v_n)).
\]

1I use an infinite version of \( \text{FO}_{\text{logic}} \) to express this definition. This may seem unreasonable, in particular as we only need finitary \( \Lambda \)-terms for DRT. I do not regard this as problematic, as I attach no philosophical significance to the use of a formal language in expressing the theory. The system \( \Lambda \)-DRT in section 4.3 will be axiomatised informally.
 CHAPTER 3. SEMANTIC THEORIES

The resulting theory is distinct from standard treatments of DRT in several aspects, of which one appears to be somewhat worrying: in the axiom of the conditional only bound variables in $s$ are quantified, in contrast to standard DRT. This means for example that in $\{x, y\} \cup s \not\supseteq \{x\} \cup s$, only the $x$ in $s_{xy}$ will be bound by universal quantification, while the occurrence of $y$ in $s_{xy}$ remains free, and thus unaffected by $\mathcal{G}$. The same need to abstract over anaphoric variables reappears when we introduce an operation of ‘merging’ of DRSs $\otimes$, taken from [Cooper 1983] into the system, usually written in infix notation, where $Y$, $Z$ are not among the $X_i$.

\[ \begin{align*}
  \mathcal{G} & \equiv \lambda Y \lambda Z \lambda (X_i \rightarrow x_i)_{u \in \mathcal{V}_w} \cdot (Y(x_i, X_i)_{u \in \mathcal{V}_w} \cap Z(x_i, X_i)_{u \in \mathcal{V}_w}).
\end{align*} \]

Item 3 of the following theorem approximates the ‘merging principle’ of the previous section. It exposes the limitations of $\otimes$, which in contrast to Zeev’s $\otimes$ allows no capturing of free variables/parameters. Ways of overcoming this restriction will be discussed in the next chapter.

**Theorem 36**

1. $PTY_M(t) \land PTY_N(t') \iff PTY_{M \cup N}(t \otimes t')$, if the Proposition Structure is strong.

2. $t \equiv t'(x_1, \ldots, x_n, l_0) \iff t(x_1, \ldots, x_n, l_0) \cap t'(x_1, \ldots, x_n, l_0).

3. $\lambda (X_i \rightarrow x_i)_{u \in \mathcal{V}_w} \cdot (x_1, \ldots, x_n, l_0) \in \mathcal{F}(x_1, \ldots, x_n, l_0)$

   \begin{align*}
   \mathcal{F}(X_i, x_i)_{u \in \mathcal{V}_w} & \equiv \lambda (X_i \rightarrow x_i)_{u \in \mathcal{V}_w} \cdot (x_1, \ldots, x_n, l_0) \in \mathcal{F}(x_1, \ldots, x_n, l_0).
   \end{align*}

**Proof:**

1. Let $M \cup N \subseteq L$, then $PF^L(t)$ and $PF^L(t')$. Hence $PF^L(t \equiv t')$.

2. By A5.

3. We may assume no $X_i$ to be vacuously bound. Then

   \[ \begin{align*}
   \mathcal{G} & \equiv \lambda (X_i \rightarrow x_i)_{u \in \mathcal{V}_w} \cdot (x_1, \ldots, x_n, l_0) \in \mathcal{F}(x_1, \ldots, x_n, l_0).
   \end{align*} \]

**3.3.4 Consistency**

To prove the existence of Kamp Structures we just need to check that the proof in section 3.1.2 for Frege Structures goes through for an operator $\hat{\mathcal{G}}$ which is like $\mathcal{G}$ but replaces $\mathcal{G}$ with $\mathcal{G}_{\otimes}$ (or just adds $\otimes$). The only problem in doing so stems from the quantification over sets of variables in the axiom:

\[ \begin{align*}
  (Ax_{\otimes}) \quad \forall M, N \subseteq \mathcal{V}_w: \quad PC_i(P, f, h, M, N) \Rightarrow (d_t \land f \in \mathcal{P} \land d_t \land f \in \mathcal{T} \Rightarrow T_{G}(T, f, h, M, N)).
\end{align*} \]

We thus need to adapt the definition of $\hat{\mathcal{G}}$ to this format.

**Definition 28**

$\hat{\mathcal{G}}(P, T) \equiv (P', T')$, where for the logical constants $i$ of Kamp Structures:

- $P'$ is the smallest set such that for each $i, j, h, M, N$:
  \[ PC_i(P, f, h, M, N) \Rightarrow d_t \land f \in P' \land d_t \land f \in T'. \]

- $T'$ is the smallest set such that for each $i, j, h, M, N$:
  \[ PC_i(P, f, h, M, N) \land TC_i(T, f, h, M, N) \Rightarrow d_t \land f \in T'. \]

**Lemma 37** Any fixed point for $\hat{\mathcal{G}}$ satisfies $Ax_{\otimes}$.
PROOF: Only part (iii) of the proof for Frege Structures needs to be examined. The crucial fact about \( \omega \)-properties that we need here is that

\[
(s) \quad PTY^H(d) \land PTY^H(d) \Rightarrow M = M'.
\]

Let \( PC_\varphi(P, f, h, M, N) \) and \( d_\varphi \cdot f \in T \). Then there are \( i, f', h', M', N' \) with \( d_\varphi \cdot f = d_\varphi \cdot f' \) such that \( PC_\varphi(P, f', h', M', N') \) and \( TC_\varphi(T, f', h', M', N') \). But then by \( i = \emptyset \) and \( f = f' \) we have \( PC_\varphi(P, f, h, M', N') \), which implies that \( M = M' \), and \( N = N' \). Therefore \( TC_\varphi(T, f, h, M, N) \). \( \square \)

The other axioms are obeyed as before. Again the result of the construction is in fact a strong Kamp Structure, as can be easily proved from (s).

**Theorem 38** Every non-trivial \( SAC \)-model can be extended to a Kamp Structure, where infinitely many relations are internally defined by \( \omega \)-properties.

**PROOF:** For each given formula \( \phi \) on \( x \), we take a semantically distinct term \( r_\phi \), plus a term \( r_\phi \) distinct from all \( r_\iota \) and \( x_1, \ldots, x_n \) and add the following axiom to the theory of Kamp Structures

\[
\forall x_1, \ldots, x_n \bigwedge_{i=1}^n x_i \neq x_j \Rightarrow P\left(\langle r_\iota; r_\phi, \langle x_i, x_i \rangle \rangle \rangle \right) \land T\left(\langle r_\iota; r_\phi, \langle x_i, x_i \rangle \rangle \rangle \right) \Rightarrow \phi_\iota, \phi_\iota.
\]

Given these proposition conditions, the least fixed point for \( \delta_K \) makes \( \langle r_\iota; r_\phi, \langle x_i, x_i \rangle \rangle \rangle \) a proposition if and only if for all \( j \neq i \) iff \( x_j = x_i \) for some \( i \in \{1, \ldots, n\} \). Therefore we have

\[
PTY^{(x_1, \ldots, x_n)}(\delta_\{x_1, \ldots, x_n\}, \langle r_\iota; r_\phi, \langle x_i, x_i \rangle \rangle \rangle).
\]

Together with the truth conditions, \( \phi \) is therefore internally defined by an \( \omega \)-property. \( \square \)

### 3.4 Situation Theory

Situation Theory is different from DMG and DRT in that it is not a semantic theory that by itself provides at least the outlines of a syntax-semantics interface. Providing such an interface for ST is the task of ‘Situation Semantics’.\(^{16}\)

Situation Theory concentrates on the (ontological) side of the enterprise, as I do in this thesis. It fits less well into this piece of work though, because simultaneous abstraction is not really at the focus of ST’s attention. AL is important for modelling forms of ST, but the central contributions of ST lie in the theory of situations and their informational content, which has very little to do with simultaneous abstraction.\(^{17}\)

My aim here is not to give a definitive version of ST, but rather to show how some rather mild assumptions about the part-of structure of events\(^{18}\) buys us rather cheaply a fair amount of ST, if we make use of the theory of Proposition Structures. My aim is to get a theory that is close to that of [Barwise & Cooper 1991], but rests on rather few axiomatic assumptions over and above those already introduced in previous sections.

We add to \( FOL_{SKC} \) the predicates \( Sit(x) \) (being a situation) and \( \leq \) (part of), assuming that \( \leq \) is a partial order on situations. It seems reasonable to assume them to form a (complete) lattice, but I just add the weaker condition of directedness here. Let me use \( e, e' \ldots \) as variables for events, writing ‘\( \forall e \ldots \)’ instead of ‘\( \forall e \ Sit(e) \to \ldots \)’

**Definition 29** the following axioms constitute our Theory of Situations:

- \((\text{reflexivity})\) \( \forall e, e' \leq e \)
- \((\text{transitivity})\) \( \forall e, e', e'' \leq e \land e' \leq e'' \Rightarrow e \leq e'' \)
- \((\text{antisymmetry})\) \( \forall e, e' \leq e' \leq e \Rightarrow e = e' \)
- \((\text{directedness})\) \( \forall e, e' \exists e'' \leq e' \land e'' \leq e \)

\(^{16}\)There are of course countless ways of going about this task, which may make use of ST to varying degrees.\(^{17}\)No unique, generally accepted, Situation Theory has emerged from these efforts. In [Barwise 1989] a number of “choice points” have been isolated, which provide a map for the various versions of ST one might encounter. I leave most of these choices open, to be fixed in the light of the semantic application or particularly strong metaphysical intuitions.\(^{18}\)Use event and situation interchangeably.
CHAPTER 3. SEMANTIC THEORIES

I take these axioms to be fairly uncontroversial among those who believe that events have a role to play in semantic theories. One might want to restrict the last axiom to events which are part of the same possible world, if one assumes the existence of non-actual "possible situations", for it would lead to 'impossible situations' otherwise. I will take an actualist stance here. Non-realised types of events will hopefully compensate for the absence of other worlds, though the matter is not of central concern here.

3.4.1 States of Affairs

Typically in Situation Theory one classifies situations $\varepsilon$ by certain "states of affairs" (soa's) $\gamma$ which they "support", written $\varepsilon \models \gamma$. Soa's have traditionally been taken as fundamental entities of ST, with situations being regarded as something rather like sets of them. Even when situations are taken as primitives, and support as the ST correlate to truth, they still usually are in close correspondence with the sets $\{\gamma | \varepsilon \models \gamma\}$.19 and their part-of structure is defined in terms of the subset relation on those sets.

I do not want to pursue an extreme version of ST here where all propositions are "Austinian", that is of the form $'e \models \gamma'$. Given a more liberal view of propositions, it seems most natural to take part-of as basic, together with truth, and define support of soa's in terms of these: soa's are persistent types of events, and a situation supporting a soa is just another case of an object having a certain property. This brings ST closer in line with our versions of DMG and DRT, and makes consistency a trivial matter: Our Theory of Situations poses no particular consistency problem, and so long as $Sit$ and $\leq$ are given in advance it is also consistent for them to be internally defined in a Proposition-Proposition-Structure./. This is in contrast to Muskens' work, cf. [Muskens 1988], who uses a typed logic in which weak negation is unproblematic.

Definition 30 We define

- $T \varepsilon t = \forall \varepsilon \mathcal{P}(t(\varepsilon))$ ("$t$ is a type of events"),
- $Soa(t) = \forall \varepsilon, t' (\varepsilon \models t \land \varepsilon \models t' \rightarrow t' \models t)$ ("$t$ is a state of affairs"),
- Soa's are referred to by $\sigma, \sigma'$, $\ldots$. Toe's can be conjoined, disjoined, and weakly negated. We define
- $\Pi = \forall \lambda \chi \gamma \chi Z. X(Z) \cap Y(Z)$;
- $\cup = \forall \lambda \chi \gamma \chi Z. X(Z) \cup Y(Z)$;
- $\sim = \forall \lambda \chi \gamma Z. \bot X(Z)$.

Theorem 39 In every Proposition Structure obeying the Theory of Situations holds:

1. $Soa(\sigma) \land Soa(\sigma') \rightarrow Soa(\sigma \land \sigma')$;
2. $Soa(\sigma) \land Soa(\sigma') \rightarrow Soa(\sigma \lor \sigma')$;
3. $T \varepsilon t \rightarrow T \varepsilon t'$;
4. $e \models \tau \land t' \leftrightarrow e \models \tau \land e \models t'$;
5. $e \models \tau \lor \tau' \leftrightarrow e \models \tau \lor e \models \tau'$;
6. $e \models \tau \leftrightarrow e \models \tau$.

19It is thus often assumed that $\{\varepsilon | \varepsilon \models \gamma\} = \{\gamma | \varepsilon \models \gamma\}$ implies $\varepsilon = \varepsilon'$.
3.4.2 Barwise Structures

We saw above that our standard notion of negation is too weak to capture the partiality of negation in Situation Theory. The intuition for strong negation seems to be that a situation supports a negative fact, \( e \models \varphi \), if some fact in \( e \) precludes \( \varphi \) from holding, and does so in a persistent way. Using strong implication and a property \( \text{sit} \) which internally defines the predicate \( \text{Sit} \), one might try a definition like \( \varphi =_{df} \lambda X. \text{sit}(Y) \supset \bot \circ \text{sit}(Y) \). But this would make negative son’s downward persistent, and thus unnecessarily complicated. Perhaps an internal entailment relation \( \models \) between propositions, roughly similar to the ‘constraints’ of ST, could be used, together with an internally defined \( \leq \), to express the intuition as \( \varphi =_{df} \lambda X. \text{sit}(Y) \wedge X \leq Y \models \bot \circ \text{sit}(Y) \). This would presumably be genuinely partial and not downward persistent, but I have no results on how appropriate entailment relations could be internalized, and such a notion of negation may fail to derive the full set of properties of strong negation, as they are usually assumed in ST.\(^2\)

Let me thus follow the standard ST account more closely, and assume we are given a list \( u_1, \pi_1, u_2, \pi_2, \ldots, u_n, \pi_n \) of symbols for basic situated relations which apply to an assignment and a situation to form propositions.

**Definition 31** A **Barwise Structure** is a Proposition Structure satisfying the Theory of Situations such that for each \( j \in \{1, \ldots, n\} \) there are roles \( \{x_i\}_{i \in \mathcal{I}_j} \) such that

\[
\forall e \models (\sum_{i \in \mathcal{I}_j} \varphi(x_i, v_i) \land \sum_{i \in \mathcal{I}_j} \neg \varphi(x_i, v_i)) \iff \forall e \models (\sum_{i \in \mathcal{I}_j} \varphi(x_i, v_i) \land \sum_{i \in \mathcal{I}_j} \neg \varphi(x_i, v_i)) \iff \exists e \models (\sum_{i \in \mathcal{I}_j} \varphi(x_i, v_i) \land \sum_{i \in \mathcal{I}_j} \neg \varphi(x_i, v_i))
\]

\( 2 \)One may of course wonder whether these assumptions are well motivated if propositions are taken to be more fundamental than sons, as I am inclined to do.

**Theorem 40** \( \exists e \models \varphi \iff \forall e \not\models \varphi \).

3.4.3 Consistency

**Theorem 41** Any \( \text{SAC-model} \) can be turned into a Barwise Structure.

**Proof:** Let \( \mathbf{R}_n \), for \( i = 1, \ldots, n \), be a set of \( n \)-place relations satisfying \( \forall v_1, \ldots, v_n \neq \# \):

1. \( \forall e, e' R(e', e_1, \ldots, e_n) \land e \leq e' \rightarrow R(e', e_1, \ldots, e_n) \);
2. \( \forall e, e' \neg R(e', e_1, \ldots, e_n) \land e \leq e' \rightarrow \neg R(e', e_1, \ldots, e_n) \);
3. \( \forall e R(e, e_1, \ldots, e_n) \rightarrow \forall e' R(e', e_1, \ldots, e_n) \);
4. \( \exists e R(e, e_1, \ldots, e_n) \lor \neg R(e, e_1, \ldots, e_n) \).

\( \)(It would be nice to be able to show consistency for a theory that has strong negation as an internal operation, with appropriate axioms of support, rather than using the following metasynthetic definitions.)
Such relations clearly exist, and can be internally defined by $n$-place properties $t_i, \overline{t}_i$. We then take

$$u_i = \lambda \{ x_1, \ldots, x_n \} \lambda \overline{x}. t_i x_1 \ldots x_n$$

$$\overline{t}_i = \lambda \{ x_1, \ldots, x_n \} \lambda \overline{x}. \overline{t}_i x_1 \ldots x_n$$

The Barwise axioms are then obeyed as required.

It is clear from the fixed point construction that the additional quantificational axioms of Frege and Kamp-Structures are all consistent with this theory too. So we can throw them all together without any worries.

**Theorem 42** Any SAC-model can be turned into a Frege-Kamp-Barwise Structure with Predication.

---

**Chapter 4**

**Further Directions**

In this chapter I will be mainly exploring variations of the SAC that might bring it closer to more popular versions of DRT. After a look at dynamic semantics, and a typed version of the SAC, I will pursue what seems to me a rather promising direction of giving a slightly more structured semantics to the SAC, inspired by [Zeevat 1988]. Under this semantics the set of abstracted variables is a defining ingredient to the meaning of $\Lambda$-terms. This allows us to define versions of the conditional, and the merge-operator, which are closer to the originals than those of the preceding chapter. Finally I briefly discuss the notions of restriction and appropriateness, which figure prominently in [Barwise & Cooper 1991].

An understanding of sections 3.1 to 3.3 will be presupposed for most of this chapter.

**4.1 True Dynamics**

The system of DRT presented in chapter 3.3 did not use any non-standard kind of binding of variables beyond the scope of an appropriate operator. But such binding is prominent in so called ‘dynamic’ forms of semantics, which thus stand in some contrast to my approach. It is time to ask what the relation is between the SAC and systems like the Dynamic Predicate Logic (DPL) of [Groenendijk & Stokhof 1991], which are based on relational interpretations. DPL replaces the non-standard DRT syntax by FOL formulas, but they have
shown how to give the relational interpretation for DRSs too. Let us view relations between assignments as functions from assignments to sets of assignments. Call sets of assignments \( \omega\)-sets, and take conditions to denote just \( \omega\)-sets \( [[\phi]] \).

We then get the following denotation for a DRS:

\[
[[x_1\ldots x_n] \phi]_{\text{DPL}} = \{ h \mid \exists \sigma \in h \land \sigma \in [\phi] \}.
\]

This should be compared with the extensions of such terms as they are interpreted in Kamp Structures. Let \( \text{ext}(d) = \{ \{ d \} \cdot h \mid h \in T \} \). We then also get \( \omega\)-sets as denotations under an assignment, but they are different from the DPL ones:

\[
\text{ext}([[x_1\ldots x_n] \phi]_{\Lambda}) = \{ h \mid \exists \sigma \in h \in \text{ext}([\phi]_{\lambda}) \}.
\]

The point is that the \( h\)'s in our set can assign anything they like to the variables outside \( \{ x_1\ldots x_n \} \), while for the DPL style of denotation they have to agree with the incoming \( g \).

A natural development of DPL, carried further in [Dekker 1993], is to view ‘information states’ as such \( \omega\)-sets, and then lift denotations of DRSs to \( \omega\)-sets of \( \omega\)-sets, such as:

\[
[[\phi]_{\text{DPL}}(G)] = \bigcup_{\phi \in G} [\phi]_{\text{DPL}}
\]

which is quite different from the trivial lifting from \( \phi \) to \( \lambda p . [p \in \phi] \) that might be used in our system to make our meanings less static.

The lack of a requirement for assignments to agree on certain variables has consequences for semantics. In contrast to DPL for example we had to abstract over variables in order to link them anaphorically to preceding discourse by our discourse sequencing operator \( \# \). I will discuss other kinds of sequencing, which may allow the capturing of free variables, in due course. What is clear though is that relational composition, the hallmark of DPL-style dynamics, is not an option for SAC-style meanings. If we would use DPL style composition for discourse conjunction the referential link between identical variables could not be sustained over more than two sentences, because in the SAC any ‘outgoing’ assignment is free to disagree with the ‘incoming’ one. So I opt for a version of Zeev's merging operator to link anaphora to their antecedents. In contrast to composition, merging is commutative, thus allowing us to capture cataphora in the same way.

A similar divergence between the two approaches is found in the treatment of the conditional. Again we had to put anaphoric variables into the abstracted set of the consequence, but only those which are anaphoric to the antecedent, not those which refer further back.

There is something of a philosophical gap between the approach of this paper and the ‘truly dynamic’ ones, which is brought out by these considerations. On the SAC view the functions from assignments to \( \omega\)-properties are not regarded as ways of ‘updating’ information states by incoming material. It makes no sense on this view to talk of ‘input’ and ‘output’ in reference to the assignments. My guiding intuition is rather that semantic objects should be plausible as things that have attitudes towards. These objects, I take it, are rather static creatures with reference to which such notions as ‘information change’ should be defined. So if we are to introduce some way of capturing free variables into our system it should be done in a way that preserves the fundamentally static view of semantic objects that underlies the proposed framework.

4.2 A Typed System

I will sketch a typed version of the SAC here, with some operators added to formalize DRT in relatively simple, extensional fashion. The purpose is to illustrate how a simple way of typing the SAC can work, and to provide a system of DRT which looks more familiar to standard versions than the previous one. In particular it provides a way to bind free variables outside of the scope of their abstracted antecedents, just as it was done in traditional versions of the DRT conditional, and with Zeev's merge-operator \( @ \), described in chapter 3.3.1. By using the universe of a DRS as part of the typing information we will able to prove
as sound the merging principle of DRSs (MP) in full generality, which states that

$$(MP) \quad \lambda M.t \otimes \lambda M'.t' = \lambda M \cup M'.t \cap t'$$

even if some variables in $M$ occur free in $t'$, or some in $M'$ occur free in $t$. This means that free variables can be captured to the left and to the right, via merging with DRSs in which they are abstracted over.

It is clear that we cannot get the equation in full generality for $\lambda$, under its original interpretation, when we look at vacuous abstractions. We would need to prove for example that $\{x\}.true \otimes \{x\}.F(x)$ equals $\{x\}.true \cap F(x)$, and $\lambda M.t \otimes \lambda M'.t$ equals $\lambda (M \cup M').t \cap t'$, which is clearly not correct.

It should be noted that the capturing of free variables as such is not incompatible with the original semantics of $\lambda$. One could for example capture all discourse markers in a merger of two terms by defining

$$\| t \otimes t' \| = \lambda \| t \| \uparrow \langle \| t' \| \rangle \lambda \| t \| \uparrow \langle \| t' \| \rangle$$

where $\lambda$ is the semantic correlate of conjunction. The problem would be that we capture too much to validate the merging principle: if $M, M' \subseteq DM$ we would get $\| \lambda M.t \otimes \lambda M'.t' \| = \lambda DM.t \cap t'$. What we need is a semantic notion of the roles of abstracts, corresponding to the sets of their abstracted variables, here $M$ and $M'$. This allows us to restrict our catch to $M \cup M'$, leaving the rest free to be caught by other abstractions. The types, with the corresponding restricted denotations of $\lambda_{DRS}$, provide such a notion. Another way to provide it will be explored in due course.

In order to type SAC expressions one has to build function-types that reflect the fact that terms apply to variable assignments. So on the argument side of such a type we introduce an indication of what type the objects have to be that are assigned to which variables. A set of typed variables provides this kind of information. Hence a function type will be a pair $(M, \tau)$, where $M$ is a set of typed variables and $\tau$ is a type. We thus define the set of types and the set of variables simultaneously.

**Definition 33** Let $Type$ and $Var$ be the smallest set and family of sets such that

- $\epsilon, t \in Type$;
- $\tau \in Type \Rightarrow \epsilon^\tau \in Var$;
- $t \in Type$ and $x_i \in Var$ for $i \in I \Rightarrow \{(x_i \in I), \tau \} \in Type$.

Let $Var = \bigcup \{ \tau \in Type \mid \epsilon \tau \}$. This allows us to restrict our catch to $M \cup M'$, leaving the rest free to be caught by other abstractions.

**Definition 34** The set of Terms of $\lambda_{DRS}$ are defined by:

- $x \in Var \Rightarrow x \in Term$;
- $\epsilon, t \in Type \Rightarrow \epsilon, t \in Term$;
- $t \in Term$, and $M \subseteq Var \Rightarrow \lambda M.t \in Term_{\lambda DM}$;
- $t \in Term_{\lambda DM}$ and $M \subseteq Var \Rightarrow \lambda M.t \in Term_{\lambda DM}$;
- $t \in Term_{\lambda DM}$ and $t' \in Term_{\lambda DM}$, and $t \Rightarrow t' \in Term_{\lambda DM}$;
- $t, t' \in Term$, and $t \Rightarrow t' \in Term$;
- $t, t' \in Term$, and $t \Rightarrow t' \in Term$.

Notice that two DRSs $\lambda M.t$ and $\lambda M'.t'$ are of different types if $M \neq M'$. The operations $\otimes$ and $\Rightarrow$ thus have to be polymorphic ones, which cannot be expressed by typed constants. Such a simple-minded typing scheme is too restrictive to be of much practical use, as it prevents abstraction over DRSs with varying universes. It seems desirable to investigate conceptions of polymorphism and subtyping for the SAC to make further progress in this direction.

For illustrative purposes I will nevertheless give a simple, extensional semantics to the system. Let $U$ be a given set of objects, $D \rightarrow D'$ be the full set of total functions from $D$ to $D'$, and we define the sets $D_t$ and type correct variable assignments $M \rightarrow D$ as follows.
Definition 35 Given a universe $U$ we define typed domains by:

- $D_0 = U$;
- $D_1 = \{0, 1\}$;
- $D_{1(|M|)} = (M \to \bigcup_{x \in M} D_{|x|}) \to D_r$, where $f \in M \to \bigcup_{x \in M} D_{|x|}$ iff $dm(f) = M$ and $f(x) \in D_{|x|}$ for all $x \in M$.

Let $D = \bigcup_{i \geq 0} D_i$, and $\lambda, \gamma$ be boolean conjunction and negation on $D_i$. The notation $(x_i \to d_i)_{i \in I}$ stands for (truly) partial assignment functions, and $f \upharpoonright M$ is the restriction of $f$ to $M$.

Definition 36 The notation $\| \cdot \|$ of $\Lambda_{DRT}$-terms in a model $\langle D, \emptyset \rangle$ under a type correct variable assignment $g$ are defined by:

- $\|a\|^p = \emptyset(c_r)$;
- $\|e_i\|^p = \gamma(e_i^r)$;
- $\|t\|^p = 1$ if $\|t\|^p = \|t\|^p$, otherwise $\emptyset$;
- $\|t \cap t'\|^p = \|t\|^p \land \|t'\|^p$;
- $\|t \lor t'\|^p = \|t\|^p \lor \|t'\|^p$;
- $\|\lambda M.t\|^p = \lambda_{f \circ w} \|t\|^p$;
- $\|f(x_i, t_i)_{i \in I}\|^p = \|t_i\|^p ((x_i \to \|t\|^p)_{i \in I})$;
- $\|h_{M \circ f} \|^p = 1$ if $\forall h_{M \circ f} \|t\|^p (f) = 1 \Rightarrow \exists h_{M \circ f} \|t\|^p (h) = 1$, otherwise $0$;
- $\|h_{M \circ f} \|^p = \lambda h_{M \circ f} \|t\|^p (f(h))$.

Notice in particular the last clause: we conjoin the conditions $s$ and $s'$ when evaluating $\lambda M.t \circ \lambda M'.t'$, while making sure that variables in $M \cup M'$ are interpreted by $f$ wherever they are free in $s'$ or $s$. This is achieved by not only applying the denotations of $\lambda M.t$ and $\lambda M'.t'$ to suitable restrictions of $f$, but also overwriting $g$ by $f$ on the variables to be captured. Only free variables which are not in the universe of either of the two merged DRSs receive their values from the assignment $g$, and thus behave as ordinary placeholders whose values are fixed by the context of evaluation.

Theorem 43 (Merging Principle) $\models_{DRT} \lambda M.t \circ \lambda M'.t' = \lambda (M \cup M').t \cap t'$

Proof: $\|\lambda M.t \circ \lambda M'.t\|^p = \lambda \|M.t\|^p \land \|M'.t\|^p (f) = \lambda \|\lambda M'.t\|^p (f) = \lambda \|\lambda M'.t\|^p \land \|\lambda M'.t\|^p (f) = \|\lambda M'.t\|^p$.

In a type free, or more liberally typed system, one may need to introduce additional structure into the semantics to achieve the same result. This is taken up in the next section.

4.3 $\Lambda$-DRT

In this section I will introduce a modified semantics for $\Lambda$, enriched with the basic operation $\circ$, which integrates the earlier presented ideas of Henk Zeevat into the type free SAC. It is also closely related to the system of $\Lambda$-DRT [Millies & Pinkal 1992], which syntactically integrates ordinary typed lambda-abstraction with DRS-formation over variables of type $e$. An official semantics for their system has not yet been published, and there is reason to think that giving one is not an entirely straightforward task.\footnote{By the time of writing, a semantics for $\Lambda$-DRT, which relies on a somewhat non-standard interpretation of lambda-abstraction, had been proposed in [Kohlhase et al. 1996]. This interpretation seemed somewhat problematic, and has been improved upon, in [Kohlhase et al. 1996]. A comparison between this semantics and my system should therefore be of some interest.}

Definition 37 The language $\Lambda_{DRT}$ consists of terms $t, f, \ldots$ built up from basic constants $e, e', \ldots, #$, and variables $x, y, \ldots \in \text{Var}$ by means of abstraction $\lambda M.t$, for $M \subseteq \text{Var}$, application $f(x, t)_{i \in I}$, and merge $t \circ t'$.

In contrast to the original SAC semantics we associate not only a particular applicative behaviour with each object of the domain $D$ but also a unique set of
variables. This makes our domain even more intensional, in the sense that for two abstraction terms to denote the same object, they not only have to have the same applicative behaviour but also must abstract over exactly the same set of variables.\footnote{Translated into \( \mathbb{S} \), this means that only Axioms A3 and A7 of the proof theory are no longer sound in this system. Notice that Rule 4 still holds, giving us a weakened form of extensionality. The change seems to have no further dramatic impact on the theory developed in chapter 2, as far as I can see. In particular, all encodings of \( \lambda \)-terms have the same roles, namely \( \sigma \) or whatever we choose, and thus weak extensionality for \( \lambda \) is preserved.} For this to work, we need to turn the power set of \( \text{Var} \) into a domain, in a way that allows us to obey the continuity requirements of our semantics. Unfortunately the naive approach of using the subset relation to form a \( \text{cpo} \) won’t work. Instead I assume that \( \text{Pow}(\text{Var}) \) is a flat domain of sets of variables, with an added \( \perp_{\text{PV}} \) element, rather than a true power-domain. We need to extend the notions of set-union and assignment updating to the cases involving \( \perp_{\text{PV}} \). The cases involving proper (possibly empty) sets of variables remain as before, and we define \( M \cup \perp_{\text{PV}} = \perp_{\text{PV}} \cup M \equiv\| \) \( \perp_{\text{PV}} \) and \( g|_{\text{PV}} \equiv\| \perp_{\text{PV}} \rightarrow D \).

The retraction by means of which abstraction and application are interpreted is thus slightly more complicated in mapping between objects \( d \) and pairs \( (M, \mu) \), where \( M \in \text{Pow}(\text{Var}) \) is the set of roles of \( d \) (on the outermost level), and \( \mu \in (\text{Var} \rightarrow D) \) describes \( d \)’s applicative behaviour. We thus use domains \( D \) such that

\[
\begin{align*}
D & \equiv \{ \text{Pow}(\text{Var}) \times ((\text{Var} \rightarrow D) \rightarrow D) \}.
\end{align*}
\]

From this we define \( \Psi(d) \) as the first component of \( \Psi(\langle d \rangle) \), also called the \text{ROLES} of \( d \), and \( \Phi(d) \) as the second component, called the \text{APPLICATIVE BEHAVIOUR} of \( d \). We write \( d \cdot f \) for \( (\Phi(d))(f) \).

\begin{definition}
A DRT-\text{MODEL} consists of a retraction \( D \equiv \{ \text{Pow}(\text{Var}) \times ((\text{Var} \rightarrow D) \rightarrow D) \} \) and an interpretation \( \mathfrak{I} \) which maps constants into \( D \), with \( \mathfrak{I}(\#) = \perp_{\mathbb{D}} \). The denotation function \( \mathfrak{I} \cdot \) for terms, under a variable assignment \( \gamma \), is given by:

\begin{itemize}
  \item \( \mathfrak{I}( \lambda x. t ) = \mathfrak{I}(\lambda x. t) \);
  \item \( \mathfrak{I}( t \cdot f ) = \mathfrak{I}( t \cdot f) \);
\end{itemize}
\end{definition}

\[\theta \]
Capturing free variables is clearly possible in this way, but we cannot hope to increase the catch by applications to terms containing such free variables. For example, we cannot derive an equation where the VP content \( \lambda z\{.Gzx\} \) contains an anaphoric reference 'x' to the content \( \lambda P\{[x].Fx \circPx\}\) of the subject NP, as in

\[ ([x]) (\lambda P\{[x].Fx \circPx\}) (\lambda z\{.Gzx\}) = \{x\}.Fx \circ \{\}, Gxx. \]

It has been assumed in \( \lambda\)-DRT that such equations hold, without having given a semantic justification for them. But this is not a problem that can't be surmounted. We simply use the \( \eta \) and \( \lambda \) operators of DMG, from chapter 3.2., to get a cleaned-up version of this form of variable capture. Recall their definition as \( \lambda DM \) and \( (x, \lambda ) \) from that section, where \( DM \) is some infinite subset of the variables. The following is valid in our \( \lambda\)-DRT, given \( x \in DM\):

\[ (\lambda P\{[x].Fx \circPx\}) ^\eta (\lambda z\{.Gzx\}) \in \{x\}.Fx \circ (\{\}, Gxx. \]

which by the MP reduces to \( \{x\}(Fx \cap Gxx). \)

Notice that this kind of system uses three apparently different kinds of abstraction: ordinary unary \( \lambda \)-abstraction, abstraction over the universes of DRSs, and \( \eta \)-abstraction to "intensionalize" over discourse marker assignments. Yet these three operations are in fact, as I have shown, instances of one and the same form of abstraction, namely simultaneous abstraction, as it is formalised in \( \Lambda \) and \( \Lambda AL \). The present modification in the semantics of the SAC in this section does not affect that result.

I thus propose DMG style 'intensional' application \( t/f \) as a general mode of semantic composition for \( \lambda\)-DRT, in order to extend the capturing of variables beyond the scope of \( \otimes \). Having the extra level of abstraction \( \eta \) in the representation \( \eta \lambda UC \) of DRSs is also useful for giving a truth-axioms for \( \otimes \). If we would insist on using free variables, as in standard DRT, we would be unable to regard \( \otimes \) as an operator, or constant, which takes two terms to form a complex one, whose denotation under assignment \( g \) is a function of the denotations of its arguments under \( g \). The additional abstraction \( \eta \) over all discourse markers allows us to view \( \otimes \) as a constant, as we did in the earlier system of Kamp Structures. Instead of \( PTV \) \( \eta \) we therefore need a similar predicate \( DRSH \), which refers to the second level of argument roles of \( \eta \)-lifted DRSs. Given two lifted DRSs \( \eta \lambda UC \) and \( \eta \lambda UC', \) we get a condition \( \eta \lambda UC \otimes \eta \lambda UC' \) which is also lifted, in the sense that it needs to be applied to a discourse marker assignment to form a proposition.

I will state the new axiom of truth in semantic terms here, rather than in some variant of FOL SAC. Let \( \otimes \) be the operation that interprets the constant \( \otimes \).

**Definition 39**

1. \( DRSH (d) \otimes_{d f} \forall f dm(f) \supset DM \Rightarrow \)

\[
\Phi_{d f} = M \wedge \forall h dm(h) \supset M \Rightarrow d f \in \mathcal{T}.
\]

2. \( CNDSH (d) \otimes_{d f} \forall f dm(f) \supset DM \Rightarrow \)

\[
d f \in \mathcal{T}.
\]

3. A Kamp Structure for \( \lambda\)-DRT is a \( \Lambda\)-model with subsets \( \mathcal{T}, \mathcal{P} \) obeying the Proposition Structure axioms, such that for all \( d, \delta \in \mathcal{D}:
\]

\[
DRSH (d) \land DRSH (d) \Rightarrow CNDSH (d) \land \forall f dm(f) \supset DM \Rightarrow (d f \delta) \in \mathcal{T} \land
\]

\[
\forall h dm(h) \supset M \wedge (d f \delta) \in \mathcal{T} \Rightarrow \exists k dm(k) \supset M \land (d f \delta) \in \mathcal{T}.
\]

\(\eta\) Notice that \( \otimes \) is not treated as a term; here for this reason \( t \otimes f \) is not a function of \( t f \) and \( t f \), but of \( t \), \( t f \) and \( g \), similar to \( \Lambda M \) being a function of \( t \) and \( g \), not of \( t f \) and \( M \).
One might want to always abstract over all discourse markers of a DRS, thus uniformly representing DRSs \((U, C)\) as \(\lambda U. C\) rather than \(U \land C\). The merge operator could be redefined in a suitable way, or simply be lifted to \(\lambda (U \land U)\), with the expected result. The conjunction of conditions could be lifted similarly. Here are some simple facts about lifted DRSs and conditions.

**Theorem 46**

1. \(DRS^{L}_{DM}(t) \land DRS^{L}_{DM}(\theta) \Rightarrow M = M'\).

2. \(DRS^{L}_{DM}(t) \land DRS^{L}_{DM}(\theta) \Rightarrow DRS^{L}_{DM}(\lambda (U \land U)\).

3. \(CN_{DM}(t) \land CN_{DM}(\theta) \Rightarrow CN_{DM}(\lambda (U \land U)\).

Yet, a sweeping use of abstractions seems unnecessary to me, and I prefer to make only minimal use of caps and caps, along the lines of the earlier DMG fragment. I hope the resulting variation in the representation of DRSs, sometimes using \(\lambda\), sometimes not, will not confound too much here. Perhaps parsimony should be sacrificed for generality here.

### 4.3.1 Consistency

Let me show how to adapt the fixed point construction for Kamp Structures to the new axiom. We need to make sure that \((d \Rightarrow f) \cdot f\) cannot be equal to the application of any other logical constant to its arguments, or of the same logical constant to different arguments. We thus set

\[
[d \Rightarrow f] = [\lambda (z_1, z_2) \cdot \lambda (x_1) \cdot x_1 \in v \in \lambda (t \Rightarrow (z_1, z_2)) \cdot (x_1, x_2, x_3) \in v]\]

where \(v\) is the proposition condition \(DRS^{L}_{DM}(f(z_1)) \land DRS^{L}_{DM}(f(z_2)) \land dm(k) \supset DM\), and the truth condition \(\forall k \ dm(k) \supset M \land (f(z_1) \cdot h \cdot k) \in T \Rightarrow \exists k' \ dm(k') \supset N \land (f(z_2) \cdot k' \cdot k') \in T\).

The new axiom may be reformulated as:

\[
(A_{\lambda}) \quad PC_{\lambda}(P, f, h, M, N) \Rightarrow (d_{\lambda} \cdot f) \cdot h \in P \text{ and } (d_{\lambda} \cdot f) \cdot f \in T \iff

TC_{\lambda}(T, f, h, M, N))
\]

with the proposition condition \(DRS^{L}_{DM}(f(z_1)) \land DRS^{L}_{DM}(f(z_2)) \land dm(k) \supset DM\), and the truth condition \(\forall k \ dm(k) \supset M \land (f(z_1) \cdot h \cdot k) \in T \Rightarrow \exists k' \ dm(k') \supset N \land (f(z_2) \cdot k' \cdot k') \in T\).

We keep the definition of \(d_{\lambda}\) from 3.3.3, letting \(d_{\lambda} \cdot f\) for \(i = \Delta, \Rightarrow\) and \(d_{\lambda} \cdot f\) otherwise.

**Lemma 47** Any fixed point for \(d_{\lambda}\) satisfies \(A_{\lambda}\).

**Proof:** Again only case (iii) of the original proof is of interest. (iii) Let \(PC_{\lambda}(P, f, h, M, N)\) and \(d_{\lambda} \cdot f \cdot h \in T\). Then there are \(i, j, f', h', M', N'\) such that \(PC_{\lambda}(P, f', h', M', N')\) and \(TC_{\lambda}(T, f', h', M', N')\). But then \(i = \Delta, j = f(z_1), f(z_2) = f(z_2), h = h'\). Hence \(PC_{\lambda}(P, f, h, M', N')\), which implies by theorem 45.1 that \(M = M'\) and \(N = N'\). Therefore \(TC_{\lambda}(T, f, h, M, N)\).

The other logical constants behave as before. So we conclude:

**Theorem 48** Every non trivial model can be extended to a Kamp Structure for \(\Lambda_{DRT}\).

### 4.3.2 Some \(\Lambda\)-DRT Translations

I present the same Categorial Grammar fragment as for DMG here, where I try to give translations that are very close to DMG yet result in appropriate DRSs. We use the same categories, and slash elimination rules. I delete the tautological assertion 'true' from any conjunction, to make things more readable. To save brackets I assume the scope of operators to be narrowest for \(\land\), and increasingly wider for application (bracketing to the left), \(\lor\), universe-abstraction, \(\Rightarrow\), and unary \(\lambda\)-abstraction.

\[
([E]) \text{ if } \Gamma_1 \Rightarrow A \land B : t_1 \text{ and } \Gamma_2 \Rightarrow B \land A : t_2 \text{ then } \Gamma_1 ; \Gamma_2 \Rightarrow A : t_1 \land t_2
\]

\[
([E]) \text{ if } \Gamma_1 \Rightarrow B : t_1 \text{ and } \Gamma_2 \Rightarrow B \land A : t_2 \text{ then } \Gamma_1 ; \Gamma_2 \Rightarrow A : t_2 \land t_1
\]
CHAPTER 4. FURTHER DIRECTIONS

4.4 A Note on Appropriateness and Restrictions

The Situation Theory of [Barwise & Cooper 1991] contains a rather sophisticated theory of appropriateness of assignments for abstracts. Any object carries appropriateness conditions with it, which have to be met by its arguments if application is to be defined. One particular way of introducing such appropriateness constraints on abstracts is by using a restriction operator $\mathbf{R}$. It links truth to identity in a presuppositional way:

\[
\text{(1) } T(t') \rightarrow (t \uparrow t') = t \land \neg T(t') \rightarrow (t \uparrow t') = \emptyset.
\]

With such an axiom one can impose additional presuppositions $\mathbf{A}$, $\mathbf{X}$, into an abstract, to form $\lambda(X, X) \rightarrow X_{\mathbf{A} \mathbf{X}}$, $(t, X, X, X, X, X)$. This applies to $(x, t \uparrow t')$, to result in $[t, X_{\mathbf{A} \mathbf{X}}] = t$ only if $[t, X_{\mathbf{A} \mathbf{X}}] = t'$ is true.

Unfortunately the axiom (1) is inconsistent with Proposition Structures: take some term $t$ which does not denote $x$, then the above version implies that $\lambda x (t \uparrow t') \uparrow t' \uparrow t')$ internally defines truth, which is inconsistent, as we saw earlier. One might try to use a weaker version instead:

\[
\text{(*) } T(t') \rightarrow (t \uparrow t') = t \land \neg T(t') \rightarrow (t \uparrow t') = \emptyset.
\]

I do not know whether (**) is consistent with respect to Proposition Structures. What is demonstrably consistent, is to have just restricted propositions.

*To prove this, one just has to account for the fact that the proposition condition $PC$, is
following [Plotkin 1990], such that:

\[(A\forall)\quad P(t) \land T(t) \rightarrow P(t[t]t) \land T(t[t]t) \leftrightarrow T[t](t)\].

This axiom is certainly a less powerful tool for introducing presuppositions than (*), which might motivate us to change our relatively naive treatment of undefinedness. As things stand, I have no notion of appropriateness for the SAC beyond the rather weak principles governing \#, and those obtained from the conditions for the formation of propositions.

4.5 Conclusion

I have shown that a minor change from the strictly ordered \(\lambda\)-calculus to a theory of simultaneous abstraction, which would seem well motivated by word order considerations alone, is enough to recreate three of the currently most influential semantic theories. I have established some fundamental results concerning variants of this form of abstraction, including equivalences between the basic SAC, an Aczel-Lanham style system, and a Combinatory Logic version. Completeness and confluence theorems were obtained.

I have used Frege Style style axioms to obtain Dynamic Montague Grammar, Discourse Representation Theory, and Situation Theory. DMG was obtained via Chierchia's Dynamic Property Theory. ST was shown to be easily obtained by adding a simple theory of events and strong negation to the basic Property Theory. A simple minded Property Theoretic DRT was obtained as well, by adding an axiom for conditionals. The disadvantages of this system motivated a slight change in the semantics of the SAC, which allowed for a more satisfactory formulation of DRT. The resulting system can be considered as a semantically kosher, highly intensional relative of \(\lambda\)-ERT, with only one underlying form of abstraction, appearing in three different guises.

I believe the SAC together with Property Theoretic axioms provides a powerful framework for linguistic semantics. It provides an elegant machinery of high generality and simplicity, and appears to make good philosophical as well as methodological sense. I feel that I have only scratched the surface of the logical and linguistic issues involved, yet I hope this may suffice to suggest that such a framework is worth having, and that the effort to obtain it is relatively modest.
Bibliography


[BIBLIOGRAPHY]  


Index

\[ \Lambda, 43 \]
\[ \lambda, 71, 74 \]
\[ \alpha, 44 \]
\[ \pi, 62 \]
\[ u, 44 \]
\[ \mu, 62 \]
\[ \approx, 43 \]
\[ \approx, 44 \]
\[ \equiv, 27 \]
\[ \equiv, 26 \]
\[ \equiv_{CL}, 32 \]
\[ \models, 44 \]
\[ \models, 53, 56, 70, 76 \]
\[ \models, 76 \]
\[ \ll, 18, 22, 71, 73 \]
\[ (\ldots), 39 \]
\[ \leq, 60 \]
\[ \models t = t', 18 \]
\[ \models_{CL} T = T', 32 \]
\[ \models_{CLM} t = t', 37 \]
\[ \varepsilon \models r, 62 \]
\[ \gamma, 43 \]
\[ \gamma, 71 \]
\[ \perp, 44 \]
\[ \sim, 62 \]
\[ \Theta, 57 \]
\[ \Theta, 54, 70, 72 \]
\[ \Delta, 26 \]
\[ \Delta_{max}, 27 \]
\[ \Delta_{mod}, 28 \]
\[ 1, 88 \]
\[ \Rightarrow, 51 \]
\[ \Rightarrow, 70 \]
\[ \uparrow \downarrow, 50 \]
\[ \bullet, 1, 4, 22, 31 \]
\[ \#, 14, 22, 31 \]
\[ \perp, 1, 5 \]
\[ \alpha, \beta, \ldots, 48 \]
\[ f, 26 \]
\[ c, c', \ldots, 14, 22, 31 \]
\[ C_{AL}, 31 \]
\[ (\cdot)_{CL}, 35 \]
\[ \Lambda, 44 \]
\[ \Theta, 46 \]
\[ LM, 48 \]
\[ dm(f), 15 \]
\[ D_{m}, 70 \]
\[ c, c', \ldots, 60 \]
\[ 3, 43 \]
\[ e, 33 \]
\[ e, 70 \]
\[ F, 44 \]
\[ FOL_{SAC}, 43 \]
\[ FP(T), 31 \]
\[ FV(t), 17 \]
\[ g_{T}, 15 \]
\[ \exists, 16, 31 \]
\[ k_{x}, 31 \]
\[ K_{x}, 31 \]
\[ \lambda, 24 \]
\[ \Lambda, 14 \]
\[ \Lambda', 23 \]
\[ \Lambda_{AL}, 22 \]
\[ \Lambda_{DRT}, 70 \]
\[ (\cdot)_{\lambda}, 35 \]
\[ \lambda^{*}, 34 \]
\[ \lambda x, t, 22 \]
\[ \lambda M_{t}, 14, 72 \]
\[ \lambda X_{t}, 24 \]
\[ \lambda x_{i}, 25 \]
\[ MV, 48 \]
\[ P, 44 \]
\[ P, 46 \]
\[ PC_{t}(P, f, h), 47 \]
\[ PP_{t}: \rightarrow, 55 \]
\[ \Phi, 15 \]
\[ \phi_{m}, \phi_{n}, 45 \]
\[ \Pi, 44 \]
\[ \pi_{n}, 40 \]
\[ P\omega(U)_{\omega}, 73 \]
\[ \Psi, 15, 34 \]
\[ \Psi_{S}, 73 \]
\[ \Psi_{S}, 15 \]
\[ \Psi Y^{n}, 56 \]
\[ (Q_{x})^{n}, 58 \]
\[ (Q_{x})^{n}, 58 \]
\[ R^{e}, 43 \]
\[ S, 31 \]
\[ s, 31 \]
\[ \sigma, \sigma', \ldots, 62 \]
\[ \tau, 64 \]
\[ \Sigma, 44 \]
\[ \Sigma, 60 \]
\[ \Sigma, 62 \]
\[ t, \tau, \ldots, 14, 22 \]
\[ T, T', 31 \]
\[ T(x, T), 31 \]
\[ T(x, T), 31 \]
\[ (\varepsilon, 25 \]
\[ T_{(T')}, 31 \]
\[ T, 44 \]
\[ \theta, 46 \]
\[ \theta, 33 \]
\[ \Theta_{x}, 37 \]
\[ T_{(T')}, 47 \]
\[ T_{(T')}, 47 \]
\[ T_{(Ax)}, 37 \]
INDEX

\[ \tau, \tau', \ldots, 62 \]
\[ \tau, 70 \]
\[ \tau, 70 \]
\[ \theta_F, 47 \]
\[ \theta_K, 58 \]
\[ \theta(x), 70 \]
\[ T_{\alpha e}, 62 \]
\[ \eta_t, \alpha_t, 49 \]
\[ \partial_t, 24 \]
\[ [x/X_i]_{\theta t}, 23 \]
\[ [y_2/y_1]_{\theta t}, 18 \]
\[ V, 34 \]
\[ n, v, \ldots, 43 \]
\[ (\text{Var} \rightarrow D) \rightarrow D, 15 \]
\[ ((\text{Var} \rightarrow D) \rightarrow D)_{\alpha x}, 34 \]
\[ X, Y, \ldots, 22, 31 \]
\[ x, y, \ldots, 14 \]
\[ x, y, \ldots, 23 \]
\[ x, y, \ldots, 48 \]
\[ \Xi, 56 \]
\[ \{x_i\}_{\theta t}, 14 \]
\[ \langle x_i \rightarrow \xi_i \rangle_{\theta t}, 15 \]
\[ z, 25, 31 \]
\[ z, z_2, 44 \]