INSTANTIAL LOGIC

An Investigation into Reasoning with Instances

W.P.M. Meyer Viol

ILLC Dissertation Series 1995-11

Instantial Logic

ILLC Dissertation Series 1995-11

institute *for* logic, language *and* computation

For further information about ILLC-publications, please contact

Institute for Logic, Language and Computation
Universiteit van Amsterdam
Plantage Muidergracht 24
1018 TV Amsterdam
phone: +31-20-5256090
fax: +31-20-5255101
e-mail: illc@fwi.uva.nl

Instantial Logic

An Investigation into Reasoning with Instances

Instantiële Logica

Een Onderzoek naar Redeneren met Voorbeelden

(Met een Samenvatting in het Nederlands)

Proefschrift

ter verkrijging van de graad van doctor aan de Universiteit Utrecht op gezag van de rector magnificus, Prof. dr. J.A. van Ginkel ingevolge het besluit van het college van decanen in het openbaar te verdedigen op dinsdag 20 juni 1995 des voormiddags te 10.30 uur

door

Wilfried Peter Marie Meyer Viol

Promotoren:

Prof. dr. D.J.N. van Eijck
Prof. dr. J.F.A.K. van Benthem

CIP-GEGEVENS KONINKLIJKE BIBLIOTHEEK, DEN HAAG

Meyer Viol, Wilfried Peter Marie

Instantial Logic: an investigation into reasoning with
instances / Wilfried Peter Marie Meyer Viol. - Utrecht:
LEd. - (ILLC dissertation series, 1995-11)
Proefschrift Universiteit Utrecht. - Met lit. opg.
ISBN 90-5434-041-X
NUGI 941
trefw.: logica/taalkunde

Contents

Acknowledgments

It is rare that the initial cause of an event can be pinpointed as accurately as that of the origin of this thesis. In 1977 a young Johan van Benthem visited the department of philosophy of the University of Groningen to talk about his PhD thesis. I entered the lecture hall a dissatisfied student who read philosophy "wildly in all directions". The conclusion of the lecture saw me a changed person. For better or worse, my life took a different course that day and this thesis is one of the results. Both as my teacher and my supervisor Johan has been an inspiration to me. His good company and stubborn support, surviving even the most unrelenting resistance, leave me deeply indebted.

Jan van Eijck and I go back an even longer way. We know each other from the time I was half-heartedly intending to graduate in ethics, and he was still active in faculty politics and sported a beard. Together with Johan and Ed Brinksma we formed a reading group of which I still have fond memories. A decade later he was instrumental in getting me to the OTS. Four years ago he handed me Kit Fine's "Arbitrary Objects" with the admonition to read it. This thesis grew out of the study of that book. I want to thank him for putting up with my inadequacies as a PhD student, for his contributions to the content and form of the thesis, for our discussions about life and logic, and for his dinner parties.

One must be considered fortunate to be presently a student of logic in the Netherlands. A more stimulating environment for a PhD student is hard to imagine. Of the people who formed my personal part of that environment I can only mention a few.

Besides being great fun, my weekly meetings with Patrick Blackburn, Claire Gardent and Jennifer taught me a lot about linguistics. I thank Patrick for his inspiration, friendship and outstanding cooking.

For significantly influencing the direction of the thesis I am indebted to Jaap van der Does, Tim Fernando, Dov Gabbay, Ruth Kempson, Michiel van Lambalgen, Diana Santos, Anne Troelstra, Albert Visser, Fer-Jan de Vries.

For contributions to the layout of this thesis beyond the call of duty I would like to thank Nataša Rakić and the amazing Maarten de Rijke.

At the OTS, CWI and FWI it was a pleasure to work, eat and party with Elena Marciori, Jan Rutten, Frank de Boer, Daniel Turi, Krzysztov Apt, Yde Venema, Kees Vermeulen, Marcus Kracht, Vera Stebletsova, Michael Moortgat, Annius Groenink, Natasha Alechina, Jan Jaspars, Dorit Ben-Shalom, Marianne Kalsbeek.

Finally, I would like to thank my parents Peter Meyer Viol and Françoise Meyer Viol for showing me that love can have more than one face, and Nataša for giving it mine.

Chapter 1

What is Instantial Logic?

1.1 Introduction

In introductions to mathematics which take a logical perspective on their subject matter, students tend to be treated to warnings against using talk about 'arbitrary objects' in their proofs, the party line among logicians being that arbitrary object talk is dangerous for mental health, if not morally wrong then at least highly misleading, and that it should therefore at all costs be avoided. Doets [Doe94] constitutes a nice example of this attitude. Still, there is a wide gap between theory and practice. Arbitrary object talk abounds in mathematical discourse. Apparently, and maybe sadly, the warnings do not have much effect.

This is how the Dutch engineer and mathematician Simon Stevin reasons about the center of gravity of a triangle.

> *Theorem II. Proposition II*
> The center of gravity of any triangle is in the line drawn from the vertex to the middle point of the opposite side.
> *Supposition.* Let ABC be a triangle of any form ...
> *Conclusion.* Given therefore a triangle, we have found its center of gravity, as required.
> (Quoted from Struik [Str86, p. 189–191].)

In informal mathematics, when we have shown of an 'arbitrary triangle' that its center of gravity is in the line drawn from the vertex to the middle point of the opposite side, we have established that this holds for *all* triangles. The commonly held view among mathematicians has been for a long time that in addition to individual triangles, there are 'arbitrary' triangles, in addition to individual obtuse triangles, 'arbitrary' obtuse triangles, and so on.

Similarly, in traditional logic there has been a time when the grammatical similarity between sentences like "John owns a donkey" and "A farmer owns a donkey" was taken to show that the phrase 'A farmer' denotes an entity called the 'arbitrary farmer'. Not only the realm of mathematics, but the world of ordinary life as well was taken to be inhabited by a mix of individual objects and arbitrary objects. An overview of the history of this belief in arbitrary objects, spiced with heavy criticism to the effect that this view almost inevitably leads to moral degradation, is given in Barth [Bar74].

The reason the notion of an arbitrary object has fallen into almost total disrepute has to do with some fundamental problems concerning the principle of *generic attribution*, which states, bluntly, that an arbitrary object has those properties common to the individual objects in its range. In its informal formulation, the principle of generic attribution leads in a straightforward way to contradictions for *complex properties*. Take an arbitrary triangle. Then it is obtuse or acute-angled, since each individual triangle is either obtuse or acute-angled. But it is not obtuse, since there are individual triangles which are not obtuse. Similarly, it is not acute-angled, since there are individual triangles which are not acute-angled. Therefore it is obtuse or acute-angled, yet it is not obtuse and it is not acute-angled: a contradiction. This problem has brought many logicians to the conclusion that arbitrary objects belong to the "dark ages of logic" (see e.g. Lewis [Lew70]).

In linguistics also, there are cases where reference to arbitrary objects seems very natural indeed. Consider the following text:

If a farmer owns a donkey, he beats it regularly. (1.1)

Discourse Representation Theory (Kamp [Kam81]) takes it that pronouns refer to objects that have in some sense been introduced by the previous text. Donkey examples like this one are the stock- in-trade of Discourse Representation, but the theory has trouble with this particular example. If one admits only individuals in one's ontology, the example leads to problems; for to what individual farmer and individual donkey can the pronouns 'he' and 'it' be said to refer? It seems natural to have them refer to the arbitrary farmer and the arbitrary donkey he owns, respectively.

Common to these examples from informal mathematics and formal linguistics is the use of *instances* to witness existential, indefinite, information. In both cases the instances chosen are intended to be interpreted as *arbitrary* representatives of the concept in question.

Now, *instantial logic* is the general name for logical frameworks that formalize reasoning which proceeds by introducing, or choosing, instances to deal with indefinite information in order to draw general conclusions. Consequently, in this dissertation our main interest will be in the logic of reasoning with indefinite information. That is, the existential quantifier will play the leading role.

Various traditional logical frameworks have something to contribute to instantial logic . The following three are the most important, and they therefore constitute three main ingredients of the subject matter of this thesis:

- natural deduction,
- epsilon calculus,
- arbitrary object theory.

1.2 Overview of the Thesis

Chapter 2 will introduce the three main ingredients of the thesis. *Natural deduction* was first proposed as a format for the proof theory of classical and intuitionistic logic, and we shall introduce it in the setting of classical logic, after a review of the semantics of classical predicate logic. In this proof theoretic set-up of classical predicate logic (CPL) and intuitionistic predicate logic (IPL) an existentially quantified formula $\exists x\varphi$ is dealt with by introducing a so-called *proper term* and continuing the derivation with $\varphi[a/x]$: the proper term is taken to be an instance or representative which supposedly denotes an arbitrary individual satisfying the matrix of the quantified formula. In general, variables or individual constants are used as proper terms. These terms get their meaning as representatives of φ-ers through the proof theoretic context but, as individual constants, they lack syntactic structure expressing what they are to be representatives of.

In the *epsilon calculus* this situation changes. The epsilon calculus was introduced as an extension for a Hilbert-style axiomatization of classical logic. After a presentation of the semantics of first order logic with epsilon terms, we shall discuss axiomatic deduction and natural deduction for this language. In the latter proof theoretic framework an existentially quantified formula $\exists x\varphi$ gives rise to the introduction of an *epsilon term* $\epsilon x : \varphi$. This instantial term has enough syntactic structure to identify it uniquely as representing an arbitrary φ-er.

Next we move on to *arbitrary object theory*, which has a natural link to a natural deduction calculus. This theory supplies a semantics intimately tied to derivations in natural deduction calculi and their use of proper terms. Its main focus is on *dependence* between proper terms arising in the course of a natural deduction proof.

After these individual introductions, some connections between epsilon calculus and arbitrary objects theory will be charted.

The three methods of reasoning with instances introduced in Chapter 2 do not exceed the bounds of classical logic. It is well-known that first-order logic can be completely formalized by an appropriate natural deduction calculus. A famous result by Hibert states the conservativity of the epsilon calculus over first-order logic. And the standard calculi formalizing reasoning with arbitrary objects are conservative over classical logic as well. But this is no longer the case if we consider instantial reasoning in IPL.

In the third chapter we shall investigate the mechanisms for instantial reasoning of Chapter 2 within the context of IPL. In particular, we shall consider epsilon extensions of intuitionistic logic. This turns out to be interesting, because such extensions are *not* conservative. Reasoning with instances in IPL allows us to derive principles that are not valid in standard IPL. We shall locate the increase in strength of intuitionistic instantial logic in *Plato's Principle* ([Bet69]):

$$\exists x(\exists x\varphi \rightarrow \varphi).$$

The addition of this principle to IPL gives us an *intermediate* logic. Plato's principle spawns a host of weaker and related principles all of which determine their own intermediate logic. All of these intermediate logics are interesting from a logical point of view. In this chapter we shall first determine these intermediate logics proof theoretically. Next we determine the classes of frames they define and show that all but one are *incomplete.*

Are there ways to restrict the epsilon calculus in such a way as to achieve conservative extensions of intuitionistic predicate logic? This question will occupy us in Chapter 4. The answer to this question will lead us to a deeper analysis of process of producing and using instances. In instantial reasoning, instances generally are introduced in an ordered way. In this process of introduction *dependencies* arise between the terms involved. These dependencies have to be taken into account in order to create conservative epsilon extensions of IPL. Indefinite information, for instance a logical formula of the form $\exists x\varphi$, leads us to introduce a representative, for instance a, and continue reasoning with this representative, that is, with the formula $\varphi[a/x]$. In such a case the term a *depends* on the existential formula $\exists x\varphi$. Now CPL does not respect this dependence. That is, it may *discharge* the assumption $\exists x\varphi$ while retaining a as a representative. On the other hand, in IPL this move is prohibited. There the dependence must be preserved if the integrity of the logic is to be maintained.

But having introduced 'dependence', this notion fully merits an analysis in its own right. A large part of the *process* of instantial reasoning essentially involves dependencies between formulas, between formulas and terms, and between terms among themselves. Chapters 4 and 5 will investigate these notions. In Chapter 4 we shall concentrate on dependencies between *formulas* in derivations. In particular we shall study dependencies arising between formulas used as *assumptions.* We shall discover important differences between CPL and IPL in their treatment of assumptions. To be more precise, the assumptions used in a derivation can be viewed as constituting a *stack* where, at any point in the derivation, only a top element may be discharged. Now CPL allows arbitrary permutation of this stack, thus making any assumption available for discharge at any point, while IPL has to treat the stack as given. This analysis will be used to create an intuitionistic epsilon calculus which is conservative over IPL. This chapter will also consider the question of conservative epsilon extensions of IPL in general.

Chapter 5 will explore the area of dependencies arising in a proof theoretic context between *terms* and the issue of 'management of individuals' in general. Here we shall introduce the general notion of a *choice process* to interpret dependencies. In the course of a derivation indefinite information is used by making arbitrary *choices*. These choices are made to satisfy logical *conditions* and refer to previous choices in their use of choice values. Dependence is an abstract property of such processes. This chapter will deal with various proof theoretical issues concerning the *substructure* of the classical quantifiers in terms of choice processes. A language will be introduced in which proof theoretic dependencies can be explicitly represented. Possible semantics for the substructural logics introduced in this chapter will be briefly discussed.

All in all, no definite results are presented in this exploratory chapter. The notion of term-dependence is a multi-faceted one and very much the subject of work in progress, not only of the present author, but also of other investigators working in Amsterdam and Budapest.

In Chapter 6, instantial logic will be used for the analysis of natural language to deal with pronominal resolution and plurals in their generic and non-generic use. In natural language analysis working in the tradition of Kamp [Kam81] and Heim [Hei82] indefinite noun phrases lead to the introduction of so-called *markers* which are interpreted as arbitrary representatives of the noun phrase in question. Here the introduction of markers is part of the construction algorithm which produces semantic representations from natural language sentences. In instantial logic, indefinite information leads to the introduction of an arbitrary representative. In this logic the introduction of representatives is a logical move. Chapter 6 will describe a small fragment of English which uses this logical move of instantial logic to create possible antecedents for inter- and intrasentential pronouns. Various regimes to explain the distribution of pronouns and their interpretation will be discussed. But instantial logic also has a natural way to deal with *plural* noun phrases, both in their generic and specific interpretation. This chapter will conclude by a discussion of the interpretation of plural noun phrases within instantial logic.

Chapter 2

A Brief History of Instantial Logic

This chapter first sets the stage by a brief presentation of the semantics of first order logic, with a matching natural deduction calculus. The natural deduction format in proof theory is identified as the first ingredient of instantial logic. Next, the history of instantial logic is traced by presenting two other main ingredients, David Hilbert's epsilon calculus and Kit Fine's theory of arbitrary objects. At the end of the chapter we draw attention to some obvious links between epsilon calculus and arbitrary object theory.

2.1 Introduction

In this chapter we introduce the historical components of instantial logic as they have been developed within the framework of first-order logic. We start with a quick review of the semantics of classical predicate logic, agreeing on some notation in the process. We introduce the proof theoretic set-up that will constitute the core of our investigations of the subsequent chapters, namely *natural deduction* with its introduction and elimination rules for logical connectives and operators. In the course of a natural deduction derivation quantifiers are eliminated by introducing proper terms, that is, arbitrary representatives. In instantial logic we have the choice to interpret these representatives in the standard way, that is *locally*, by mapping them to a domain element satisfying the matrix of the quantified formula, or *generically*, by mapping them to the set of all such elements. We shall discuss proof calculi for both interpretations.

The proper terms introduced by the quantifier rules in natural deduction can be supplied with internal structure by considering *Hilbert's epsilon calculus*. This calculus will be the subject of the second section of this chapter. We shall discuss various semantic and proof theoretic aspects.

We conclude this chapter by discussing Kit Fine's theory of *arbitrary objects.* This is essentially a semantic theory of instantial terms which can accomodate a local as well as a generic interpretation. We shall discuss its most appropriate proof theory and connect it to the epsilon calculus. This theory will only occasionally play a role in the subsequent chapters.

2.2 Classical Predicate Logic

2.2.1 Semantics

The non-logical vocabulary of a predicate logical language L consists of a set

$$\mathrm{C} = \{c_0, c_1, c_2, \ldots\}$$

of *names* (or *individual constants*), for each $n > 0$ a set

$$\mathrm{P}^n = \{P_0^n, P_1^n, P_2^n, \ldots\}$$

of *n-place predicate constants* and for each $n > 0$ a set

$$\mathrm{f}^n = \{f_0^n, f_1^n, f_2^n, \ldots\}$$

of *n-place function constants.*

In practice, C and most of the P^n and f^n may be finite or even empty. To mention an example, the predicate logical language that is used in axiomatic set theory has one individual constant ($\emptyset$), no function constants and just one predicate constant: a two-place symbol R for the relation $\in$.

The logical vocabulary of a predicate logical language L consists of parentheses, the connectives $\neg$ and $\rightarrow$, the quantifier $\forall$, the identity relation symbol $=$ and an infinitely enumerable set V of *individual variables* $x_1, x_2, \ldots, y_1, y_2, \ldots, v$.

If the non-logical vocabulary (a set of individual constants, predicate constants and function constants) is given, the language L is defined in two stages. Here is a BNF definition of the set of terms (assume $c \in C$, $v \in V$, $f \in \mathrm{f}^n$):

terms $t ::= c \mid v \mid ft_1 \cdots t_n$.

This definition says that terms are either individual variables or constants, or results of writing n terms in parentheses after an n-place function constant. The second stage is the definition of formulas. Assume that indexed terms t range over terms, $P \in \mathrm{P}^n$, and $v \in V$.

formulas $\varphi ::= \bot \mid Pt_1 \cdots t_n \mid t_1 = t_2 \mid \neg\varphi \mid (\varphi_1 \wedge \varphi_2) \mid (\varphi_1 \vee \varphi_2) \mid (\varphi_1 \rightarrow \varphi_2) \mid \forall v\varphi \mid \exists v\varphi$.

Note that every collection of individual constants, predicate constants and function constants determines a different language L. A predicate logical language

L is often called a *first order language*, because predicate logic allows quantification over *entities of the first order*—in a classification of objects due to Bertrand Russell—that is to say over *individual objects.*

Identity statements are made by means of a special two-place predicate $=$; for convenience we use infix notation here, writing $a = b$ instead of $= ab$. Because of its fixed interpretation, the symbol for identity is called a *logical predicate constant.*

It is convenient to introduce a further sentential connective by way of abbreviation:

- $(\varphi \leftrightarrow \psi)$ abbreviates $((\varphi \rightarrow \psi) \wedge (\psi \rightarrow \varphi))$.

For convenience, we shall often omit outer parentheses. Similarly for parentheses between conjuncts or disjuncts in cases where there is no danger of ambiguity.

2.1. DEFINITION. (Free variables) An occurrence of a variable v is *free* in φ if it is not in the scope of a quantifier $\forall v$ or $\exists v$. Let variable v occur free in φ. Then term t is *free for v in φ* if v does not lie in the scope of a quantifier binding a variable that occurs in t.

For example, $f(y, a)$ is free for x in $Px \rightarrow \forall x Px$, but the same term is not free for x in $\forall y Rxy \rightarrow \forall x Rxx$. If a term t is free for a variable v in a formula φ we can substitute t for all free occurrences of v in φ without worrying about variables in t getting bound by one of the quantifiers in φ. If t is not free for v in φ we can always rename the bound variables in φ to ensure that substitution of t for v in φ has the right meaning. Although $f(y, a)$ is not free for x in $\forall y Rxy \rightarrow \forall x Rxx$, the term is free for x in $\forall z Rxz \rightarrow \forall x Rxx$, an alphabetic variant of the original formula.

It is customary to write $\varphi(v)$ to indicate that φ has at most the variable v free.

> The result of uniform substitution of t for free occurrences of v in $\varphi(v)$, with suitable renaming of bound variables in φ if the need arises, is written $\varphi[t/v]$.

This notational convention automatically staves off the danger of accidental capture of variables from t by quantifiers in φ. The convention will become important when we formulate natural deduction rules for quantification, in the next section.

It is useful to extend the notation to $\varphi[t/t']$. This means the result of uniform substitution of t for free occurrences of t' in φ (where a term t' is free in φ if all of its variables are free in φ), with suitable renaming of bound variables in φ if the need arises to avoid accidental capture of variables in t' in the substitution process. Note that according to these conventions, $Rxa[a/x]$ is the formula Raa, $Rxa[b/a]$ is Rxb, and $Rxa[a/x][b/a]$ is Rbb. Also, $\forall y Ryfy[x/y]$ equals $\forall y Ryfy$, $\forall x Rxfy[gx/fy]$ equals $\forall z Rzgx$, and $\forall x Rxfy[x/y]$ equals $\forall z Rzfx$.

We write $\varphi(v_1, \ldots, v_n)$ to indicate that φ has at most $v_1, \ldots, v_n$ free. The result of simultaneous uniform substitution of $t_1, \ldots, t_n$ for free occurrences of $v_1, \ldots, v_n$, respectively, in φ, with renaming of bound variables in φ as the need arises, is written $\varphi[t_1/v_1, \ldots, t_n/v_n]$. Again, this notational convention allows us to think about substitution without worrying about accidental capture of variables.

Assume L is a first order language based on a particular set of individual constants, predicate constants and function constants. Then a *model* for L is a pair $M = \langle \text{dom}\,(M), \text{int}\,(M)\rangle$, with dom (M) a non-empty set and int (M) a function with the following properties:

- int (M) maps every $c \in C$ to a member of dom (M).
- For every $n > 0$, I maps each member of P^n to an n-place relation R on dom (M).
- For every $n > 0$, int (M) maps each member of f^n to an n-place operation g on dom (M).

The set dom (M) is called the *domain* of M, int (M) its interpretation function. If a language L has a finite number k of non-logical constants it is convenient to fix an order for these and to present a model M for L as a $k + 1$-tuple $\langle \text{dom}\,(M), \ldots\rangle$, where the interpretations for the non-logical constants are listed in the same order.

Sentences involving quantification generally do not have sentences as parts but open formulae. As it is impossible to define truth for open formulae without making a decision about the interpretation of the free variables occurring in them, it is customary to employ infinite *assignments* of values to the variables of L, that is to say functions with domain V and range $\subseteq$ dom (M). The assignment function s enables us to define a function that assigns values in dom (M) to all terms of the language. We shall use aM for the set of all assignments in M. If M is a model for L and $s \in aM$, then a *value* function for terms $V_{\mathcal{M},s} : \mathbf{term}_L \to M$ is defined as follows.

2.2. DEFINITION. (Values for terms)

- $V_{M,s}(c) = \text{int}\,(M)(c)$.
- $V_{M,s}(v) = s(v)$.
- $V_{M,s}(ft_1 \cdots t_n) = \text{int}\,(M)(f)(V_{M,s}(t_1), \ldots, V_{M,s}(t_n))$.

Next, this value function is used in Tarski's truth definition. To handle the quantifiers, we need the concept of an assignment which is like a given assignment s except for the fact that it may assign a different value to some variable v:

$$s(v|d)(w) = \begin{cases} s(w) & \text{if } w \neq v \\ d & \text{if } w = v. \end{cases}$$

This piece of notation allows us to handle the quantifier case (see the clauses for $\forall v\varphi$ and $\exists v\varphi$ in the following definition).

2.3. DEFINITION. (Truth under s)

1. $M, s \models \bot$ never.
2. $M, s \models Pt_1 \cdots t_n$ iff $\langle V_{M,s}(t_1), \ldots, V_{M,s}(t_n) \rangle \in \mathrm{int}\,(M)(P)$.
3. $M, s \models t_1 = t_2$ iff $V_{M,s}(t_1) = V_{M,s}(t_2)$.
4. $M, s \models \neg\varphi$ iff not $M, s \models \varphi$.
5. $M, s \models (\varphi_1 \wedge \varphi_2)$ iff both $M, s \models \varphi_1$ and $M, s \models \varphi_2$.
6. $M, s \models (\varphi_1 \vee \varphi_2)$ iff either $M, s \models \varphi_1$ or $M, s \models \varphi_2$.
7. $M, s \models (\varphi_1 \rightarrow \varphi_2)$ iff either not $M, s \models \varphi_1$, or $M, s \models \varphi_2$.
8. $M, s \models \forall v\varphi$ iff for all $d \in \mathrm{dom}\,(M)$, $M, s(v|d) \models \varphi$.
9. $M, s \models \exists v\varphi$ iff for some $d \in \mathrm{dom}\,(M)$, $M, s(v|d) \models \varphi$.

It is easy to see that only the finite parts of the assignments that provide values for the free variables in a given formula are relevant (one can prove this so-called finiteness lemma by induction on the structure of a formula).

Note that we have been liberal in our choice of logical constants. It is well known that there are smaller complete sets of constants: by taking, e.g., $\neg$, $\rightarrow$ and $\forall$ as primitive one can define the other constants. Choosing sparse sets of constants is useful in the study of Hilbert style axiom systems. As our main concern in this thesis will be with natural deduction proof systems, we can get away with our generous assumption that all of $\neg, \wedge, \vee, \rightarrow, \forall, \exists$ are primitive.

2.4. DEFINITION. (Truth) A formula φ of L is true in M, notation $M \models \varphi$, iff for every assignment $s \in aM$, $M, s \models \varphi$.

2.5. DEFINITION. (Validity) A formula φ of L is logically valid, notation $\models \varphi$, iff for every M for L, $M \models \varphi$.

For logical consequence we have two options. Let Γ be a set of formulas of L. Then *case-to-case* consequence is defined as follows.

2.6. DEFINITION. (Case-to-case consequence) $\Gamma \models_c \varphi$ iff for all M, all $s \in aM$, if $M, s \models \gamma$ for all $\gamma \in \Gamma$, then $M, s \models \varphi$.

If Γ has the form $\{\gamma_1, \ldots, \gamma_n\}$, then we shall write $\Gamma \models_c \varphi$ as $\gamma_1, \ldots, \gamma_n \models_c \varphi$. *Truth-to-truth* consequence is defined as follows.

2.7. DEFINITION. (Truth-to-truth consequence) $\Gamma \models_t \varphi$ iff for all M, if $M \models \gamma$ for all $\gamma \in \Gamma$, then $M \models \varphi$.

For finite Γ we use the same notational convention as above.

2.8. REMARK. In case-to-case consequence free variables are interpreted as concrete individuals. They can be seen as denoting a fixed but unspecified object. In truth-to-truth consequence on the other hand variables are interpreted *generically*. They denote the *arbitrary* individuals: whatever holds for an arbitrary individual holds for all individuals. This difference in interpretation of free variables leads to differences in validities. Most typically, $Px \models_t \forall x Px$, but $Px \not\models_c \forall x Px$. Moreover, the *deduction theorem* is only generally valid for case-to-case consequence. From $\Gamma, \varphi \models_c \psi$ we can conclude to $\Gamma \models_c \varphi \rightarrow \psi$. But $\Gamma, \varphi \models_t \psi$ does not allow us to conclude $\Gamma \models_t \varphi \rightarrow \psi$. For instance, we have $Px \models_t \forall x Px$, but we do not have $\models_t Px \rightarrow \forall x Px$.

The reason why many textbooks of logic do not mention the distinction between case-to-case and truth-to-truth consequence is that they consider only premise sets Γ consisting of sentences (closed formulas). If γ is a closed formula, then $M, s \models \gamma$ for some s iff $M \models \gamma$, so under the assumption of closed premise sets case-to-case consequence coincides with truth-to-truth consequence. In the next section we shall give proof systems for both of these relations.

2.2.2 Natural Deduction for Classical Predicate Logic

Natural deduction is a perspective on proof theory due to Gentzen [Gen34], and streamlined by Prawitz [Pra65]. The key idea is that the role that each logical symbol (connective or quantifier) plays in reasoning can be captured by two sets of rules: one set for *introducing* the symbol, and one for *eliminating* it (see Zucker and Tragesser [ZT87, Zuc87]). We shall discuss the rules of natural deduction in three stages. First we deal with the rules for the Boolean connectives, next we deal with the rules for the quantifiers, and finally we shall supply rules for identity.

Rules for the Boolean connectives

We start with the case of *conjunction*. How does one prove a conjunction? That is, how can we arrive at a conclusion of the form $\varphi \wedge \psi$? Informally speaking, by noticing that both conjuncts are premises. This gives the following rule:

$$\frac{\varphi \qquad \psi}{\varphi \wedge \psi}\,\wedge\mathrm{I}$$

To eliminate a conjunction from a proof, one either focuses on the left conjunct or on the right conjunct. This gives the rules:

$$\frac{\varphi \wedge \psi}{\varphi}\,\wedge_l\mathrm{E} \qquad\qquad \frac{\varphi \wedge \psi}{\psi}\,\wedge_r\mathrm{E}$$

The rules for *disjunction* introduction derive the disjunction from the presence of one of the disjuncts as a premise. Disjunction elimination is slightly more involved. What it says is this: if we have a disjunction as a premise, and we can derive a certain conclusion from either disjunct, then we can derive that conclusion from the disjunction.

$$\frac{\varphi}{\varphi \vee \psi}\vee_l\mathrm{I} \qquad\qquad \frac{\psi}{\varphi \vee \psi}\vee_r\mathrm{I} \qquad\qquad \frac{\varphi \vee \psi \qquad \begin{array}{c}\varphi\ [i]\\ \vdots\\ \xi\end{array} \qquad \begin{array}{c}\psi\ [j]\\ \vdots\\ \xi\end{array}}{\xi}\vee\mathrm{E}_{i,j}$$

Note that the disjunction elimination rule carries two indices i, j, to indicate that the assumptions with labels $[i]$ and $[j]$ have been discharged by the application of the rule.

For *implication* introduction, the story is similar. If we can derive ψ from φ, then we have established $\varphi \rightarrow \psi$, and the introduction of the implication cancels the assumption ψ. This is again indicated by the index i of the rule which matches label $[i]$ on the discharged premise. Implication elimination is in fact the familiar rule Modus Ponens.

$$\frac{\begin{array}{c}\varphi\ [i]\\ \vdots\\ \psi\end{array}}{\varphi \rightarrow \psi}\rightarrow\mathrm{I}_i \qquad\qquad \frac{\varphi \rightarrow \psi \qquad \varphi}{\psi}\rightarrow\mathrm{E}$$

For handling *negation*, we employ the formula $\perp$. This formula expresses an absurd conclusion, so negation introduction consists of noting that $\perp$ is derived from φ, discharging the assumption φ, and drawing the conclusion $\neg\varphi$. Negation elimination takes the form of drawing the absurd conclusion $\perp$ from the pair of premises φ and $\neg\varphi$. Here are the rules:

$$\frac{\begin{array}{c}\varphi\ [i]\\ \vdots\\ \perp\end{array}}{\neg\varphi}\neg\mathrm{I}_i \qquad\qquad \frac{\varphi \qquad \neg\varphi}{\perp}\neg\mathrm{E}$$

The following rule expresses that $\perp$ denotes an absurd conclusion, from which anything follows:

$$\frac{\perp}{\varphi}\perp$$

All this together does not yet completely specify the use of negation in classical logic. For this, we also need to throw in the law of double negation:

$$\frac{\neg\neg\varphi}{\varphi}\,DN$$

Combining DN with the negation introduction rule, we effectively get a new negation *elimination* rule:

$$\dfrac{\dfrac{\begin{matrix}\neg\varphi\ [i]\\ \vdots \\ \bot\end{matrix}}{\neg\neg\varphi}\,\neg\mathrm{I}_i}{\varphi}\,DN$$

It is convenient to summarize this combination of $\neg I$ and DN in the following extra rule for (classical) negation:

$$\dfrac{\begin{matrix}\neg\varphi\ [i]\\ \vdots \\ \bot\end{matrix}}{\varphi}\,\neg\mathrm{E}_i$$

The above set of rules (i.e, with $\neg E$ and without DN) is a sound and complete set of introduction and elimination rules for the classical propositional connectives (see e.g. Van Dalen [Dal83]). To see precisely what this means one has to know what counts as a derivation in natural deduction. The best way to explain this is by example. Here is a derivation of p from $\neg\neg p$ (demonstrating that nothing is lost by leaving out DN from the set of rules):

$$\dfrac{\dfrac{\neg p\ [1] \qquad \neg\neg p}{\bot}\,\neg\mathrm{E}}{p}\,\neg\mathrm{E}_1$$

The proof tree exemplifies the process of indirect reasoning. Assume $\neg\neg p$, and also assume $\neg p$. First draw the conclusion $\bot$, and then draw the conclusion p from this, while discharging $\neg p$. The conclusion depends only on the premise $\neg\neg p$, and we have indeed derived p from $\neg\neg p$. Here is another example:

$$\dfrac{\dfrac{p \lor \neg p \qquad p\ [1] \qquad \dfrac{\neg p\ [2] \qquad \neg\neg p\ [3]}{p}\,\bot}{p}\,\lor\mathrm{E}_{1,2}}{\neg\neg p \to p}\,\to\mathrm{I}_3$$

This proof derives $\neg\neg p \to p$ from the principle of excluded middle $p \lor \neg p$ (the only undischarged assumption in the proof tree).

To make all this a bit more formal, a *proof tree* is a finite tree with the root stating the conclusion drawn at the bottom and the premises or assumptions drawn at the leaf nodes. The assumptions labeled with an index which matches a rule application below are the discharged assumptions. The proof tree establishes that the conclusion follows from the set of those assumptions that have not been discharged. If this set is given by $\{\gamma_1, \dots, \gamma_n\} \subseteq \Gamma$ (where Γ may be infinite) and the conclusion is φ, then the proof tree establishes that $\Gamma \vdash \varphi$.

Rules for the quantifiers

To get at a natural deduction system for classical predicate logic, we have to add rules for the quantifiers. Here are the rules for the *universal* quantifier (conditions on the rules will be stated below):

$$\frac{\varphi}{\forall v(\varphi[v/t]}\,\forall\mathrm{I} \qquad\qquad \frac{\forall v\varphi}{\varphi[t/v]}\,\forall\mathrm{E}$$

The term t is called the *proper term* of the application of $\forall I$.

Note that it is possible to switch to an alphabetic variant of a universally quantified formula by combining the two rules:

$$\frac{\dfrac{\forall v\varphi}{\varphi[w/v]}\,\forall\mathrm{E}}{\forall w(\varphi[w/v])}\,\forall\mathrm{I} \tag{2.1}$$

It should be noted, however, that this piece of reasoning may lead us astray, for in the first step of (2.1) we substitute w for v in φ without any guarantee that w does not already occur free in $\forall v\varphi$. It is for precisely this reason that application of the rule $\forall I$ is subject to the following condition:

2.2.1. CONDITION. (Term condition on ($\forall$I)) *The proper term t of an application of $\forall I$ should not occur free in any undischarged assumption above φ.*

This condition staves of the danger in (2.1): in case w occurs free in $\forall v\varphi$, w is not free for v in φ, but then also the application of $\forall I$ is blocked by Condition 2.2.1.

The presence of Condition 2.2.1 on the rule also makes the following applications of $\forall I$ incorrect:

$$\frac{Py}{\forall xPx}\,\forall\mathrm{I} \qquad\qquad \frac{\dfrac{Py \qquad \dfrac{\forall xQx}{Qy}\,\forall\mathrm{E}}{Py \wedge Qy}\,\wedge\mathrm{I}}{\forall x(Px \wedge Qx)}\,\forall\mathrm{I}$$

The reason why the first of these is not an acceptable instance of $\forall I$ is that Py is itself introduced by assumption; the variable y occurs in it, and y is the proper term of the application, so this derivation violates Condition 2.2.1 on $\forall I$. In the second example, the proper term of the application, y, occurs in the assumption Py on which the premise $Py \wedge Qy$ depends. This is also forbidden by Condition 2.2.1.

Here are the rules for the existential quantifier:

$$\frac{\varphi[t/v]}{\exists v\varphi}\,\exists\mathrm{I} \qquad\qquad \frac{\exists v\varphi \qquad \begin{array}{c}\varphi[t/v]\ [i]\\ \vdots \\ \psi\end{array}}{\psi}\,\exists\mathrm{E}_i$$

Note the reason for stating the premise of $\exists I$ in the form $\varphi[t/v]$: we want to allow the rule to apply non-uniformly, i.e., even to a proper subset of the occurrences of t in φ, and this is precisely what the premise allows us to do. For a specific example of this, recall that Raa has the form $Rxa[a/x]$, so the rule allows us to existentially generalize only over the first argument position of R. But Raa is also of the form $Rax[a/x]$, so the rule allows us to existentially generalize only over the second argument position of R. Finally, Raa is of the form $Rxx[a/x]$ so the rule also allows us to existentially generalize over both argument positions of R at once.

Note that the use of the substitution $[t/v]$ in the premise of $\exists I$ rules out the possibility that the variable bound in the conclusion of $\exists I$ occurs free in the premise. The reason is that $\varphi[t/v]$ in the premise denotes the uniform substitution of t for v, so $\varphi[t/v]$ does not have any free occurrences of v by definition. Thus, the following application of $\exists I$ is ruled out because the premise, Rxa is not of the required form $Rxx[a/x]$.

$$\frac{Rxa}{\exists x Rxx}\,\exists\mathrm{I}$$

Note that our convention about the meaning of $\varphi[t/v]$ implies that for instance $\forall x Rxf(x)$ is not of the form $\forall x Rxy[f(x)/y]$, because $f(x)$ is not free for y in $\forall x Rxy$ (see Section 2.2.1). Thus, the following application of $\exists I$ is, rightly, ruled out:

$$\frac{\forall x Rxf(x)}{\exists y \forall x Rxy}\,\exists\mathrm{I}$$

The term t is called the *proper term* of the application of $\exists E$. The existential quantifier elimination rule has the following condition imposed on it:

2.2.2. CONDITION. (Term condition on ($\exists$E)) *The proper term t of the application $\exists E$ should not occur in φ, in ψ, or in any assumption on which the occurrence of ψ above the line depends, other than $\varphi[t/x]$.*

This rules out:

$$\frac{\exists y Ray \qquad \dfrac{Rab\ [1]}{\exists x Rxb}\,\exists\mathrm{I}}{\exists x Rxb}\,\exists\mathrm{E}_1$$

The reason is that the proper term b of the application of $\exists E$ does occur in premise $\exists x Rxb$.

Also forbidden by Condition 2.2.2:

$$\dfrac{\exists x Px \qquad \dfrac{Pa \qquad Qa\ [1]}{Pa \wedge Qa}\,\wedge\mathrm{I}}{\exists x(Px \wedge Qx)}\,\exists\mathrm{E}_1$$

This proof tree pretends to establish the unsound principle $\exists x Px, Qa \vdash \exists x(Px \wedge Qx)$, and it is indeed fortunate that it is ruled out by Condition 2.2.2. Note that $\exists E$ cancels an assumption. Indeed, $\exists E$ looks a lot like $\vee E$.

A term t is a *proper term* in a proof tree if it is the proper term of some application of $\forall I$ or $\exists E$ in that proof tree. If t is the proper term of an application of $\forall I$ or $\exists E$ in a proof tree $\mathcal{D}$, then t must satisfy *global* constraints in $\mathcal{D}$; in other words, not any term will do as a proper term. This contrasts with the case for terms used to introduce the existential quantifier in $\exists I$ and the terms introduced upon elimination of a universal quantifier occurrence in $\forall E$. Such terms do not have to satisfy global restrictions. In other words, the correctness of an application of these rules can be established by considering *only* the premises and the conclusion, without inspecting the rest of the proof tree. Consequently, with respect to the introduction of proper terms, the pairs of quantifier rules cannot be considered as converses of each other.

It is customary to refer to the premise in an application of an elimination rule from which the connective or quantifier is eliminated as the *major premise* of the rule. If there are other premises, these are called *minor premises* of the rule.

As an example of correct quantifier reasoning, we derive $\exists y \exists x Rxy$ from $\exists x \exists y Rxy$.

$$\dfrac{\exists x\exists y Rxy \qquad \dfrac{\exists y Ray\ [1] \qquad \dfrac{\dfrac{Rab\ [2]}{\exists x Rxb}\,\exists\mathrm{I}}{\exists y\exists x Rxy}\,\exists\mathrm{I}}{\exists y\exists x Rxy}\,\exists\mathrm{E}_2}{\exists y\exists x Rxy}\,\exists\mathrm{E}_1$$

Rules for identity

Finally, we need rules for handling identities. These do not fit as nicely in the introduction versus elimination pattern, but the reflexivity, symmetry and transitivity of the identity relation have to be stated somewhere. The following set of rules express precisely this (the first rule has no premises, i.e., it is in fact an axiom):

$$\frac{}{t = t} \qquad \frac{t_1 = t_2}{t_2 = t_1} \qquad \frac{t_1 = t_2 \qquad t_2 = t_3}{t_1 = t_3}$$

Finally, we need a rule stating that identities allow us to perform substitutions:

$$\frac{\varphi[t_1/v_1, \dots, t_n/v_n] \qquad t_1 = t_1' \qquad \cdots \qquad t_n = t_n'}{\varphi[t_1'/v_1, \dots, t_n'/v_n]}$$

Again, the formulation of the premise as $\varphi[t_1/v_1, \dots, t_n/v_n]$ has the effect of making the rule handle non-uniform substitution (compare the remark on $\exists I$ above).

Note that these rules handle identity by brute force, by formulating the principles for terms of arbitrary complexity. An alternative formulation would state the principles for variables only, then rely on the substitution principle to derive the identity laws for functions with variables as arguments, and finally on the quantifier rules to derive the general format given here.

The calculus consisting of the rules for Boolean connectives, quantifiers and identity, will be referred to as CPL.

2.9. DEFINITION. (CPL derivability) Let $\Sigma \cup \{\varphi\}$ be a set of predicate logical formulas. Formula φ can be *derived* from Σ by CPL, notation $\Sigma \vdash_c \varphi$, if there is a CPL prooftree with all non-discharged assumptions in Σ and conclusion φ.

2.10. THEOREM. (CPL completeness)

$$\Sigma \vdash_c \varphi \iff \Sigma \models_c \varphi.$$

For a proof, see van Dalen [Dal83]. CPL is sound and complete for the case-to-case notion of logical consequence we have defined. It is obvious that CPL is not sound for the truth-to-truth notion of logical consequence. For instance, by the rule (→I) we may conclude from $\Sigma, \varphi \vdash_c \psi$ to $\Sigma \vdash_c \varphi \to \psi$. This does not generally hold if we replace $\vdash_c$ by $\models_t$.

2.2.3 Natural Deduction for Classical Generic Consequence

For a sound and complete natural deduction system for truth-to-truth consequence we have to modify the rules of CPL. Before discussing the modifications in a systematic fashion, let us note that derivations such as the following should be allowed now:

$$\frac{Ax}{\forall x Ax}\,\forall\mathrm{I} \qquad \frac{Rxa}{\exists x Rxx}\,\exists\mathrm{I}$$

On the other hand, inferences such as the following should be blocked:

$$\frac{\dfrac{Ax\ [1]}{\forall x Ax}\,\forall\mathrm{I}}{Ax \to \forall x Ax}\,{\to}\mathrm{I}_1 \qquad \frac{\dfrac{Rxa\ [1]}{\exists x Rxx}\,\exists\mathrm{I}}{Rxa \to \exists x Rxx}\,{\to}\mathrm{I}_1$$

If these arguments are read truth-to-truth (i.e., from the universal closure of the premise to the universal closure of the conclusion) they are invalid. Let us try to see what goes wrong in the reasoning here. In the first example, it is the interaction of the rules $\forall I$ and $\rightarrow I$ that causes trouble. First the statement Ax is read as a universal truth, next Ax is used as a hypothesis about a particular case, and these are two essentially different uses of the same assumption. Similarly, the relaxation of the condition on $\exists I$ which allows $Rxa \vdash \exists x Rxx$ reads the premise Rxa as a general statement, but the next rule assumes Rxa is a hypothesis about a particular case. Again, two essentially different uses of the same assumption Rxa. Thus, the examples suggest that the difference between case-to-case readings and truth-to-truth readings of the proof trees has something to do with the interaction between the rule $\forall I$ and the discharge of assumptions in hypothetical reasoning.

The following proposal works. Let us make a distinction between two kinds of assumptions in a proof tree, assertions and hypotheses:

- An assertion is an assumed formula which is taken to hold generally.
- A hypothesis is what it says: material for hypothetical reasoning.

As the examples show, a formula that is to be discharged later on (a hypothesis in the proper sense) cannot be read generally.

The above proof system is to be modified as follows:

- Every assumption in a proof tree which is not meant to be discharged later on is given a $\surd$ mark, to indicate that it is an assertion.
- Hypothesis discharging rule applications cannot cancel hypotheses bearing the $\surd$ mark.
- Condition 2.2.1 on rule $\forall I$ should be replaced by the following:

  > If the proper term of an application of $\forall I$ is a variable, then it is only allowed to occur free in those assumptions upon which φ depends which bear the $\surd$ mark.

After completing a proof tree, the $\surd$ marks can be erased, as they only serve for bookkeeping while the tree is under construction. A proof tree establishes that the conclusion follows truth-to-truth from the set of those assumptions that have not been discharged. If this set is given by $\{\gamma_1, \ldots, \gamma_n\} \subseteq \Gamma$ (where Γ may be infinite) and the conclusion is φ, then the proof tree establishes that $\Gamma \vdash_t \varphi$.

Note that the condition on $\exists E$ remains in place. Reasoning from Rxa to $\exists x Rxx$ is still possible, however:

$$\cfrac{\cfrac{\cfrac{Rxa\ \surd}{\forall x Rxa}\ \forall\mathrm{I}}{Raa}\ \forall\mathrm{E}}{\exists x Rxx}\ \exists\mathrm{I}$$

2.11. PROPOSITION. *The generic proof system for classical logic is sound:* $\Gamma \vdash_t \varphi$ *implies* $\Gamma \models_t \varphi$.

Proof: Let $\forall\Gamma$ denote the universal closure of all the formulas in Γ, and $\forall\varphi$ the universal closure of φ. We note that, by definition of the consequence relations,

$$\Gamma \models_t \varphi \iff \forall\Gamma \models_c \forall\varphi.$$

So, given soundness and completeness of the case-to-case proof system with respect to case-to-case consequence, *soundness* of the generic proof system with respect to generic consequence follows if we can show

$$\Gamma \vdash_t \varphi \Rightarrow \forall\Gamma \vdash_c \forall\varphi.$$

For then: $\Gamma \vdash_t \varphi \Rightarrow \forall\Gamma \vdash_c \forall\varphi \iff \forall\Gamma \models_c \forall\varphi \iff \Gamma \models_t \varphi$. So suppose $\Gamma \vdash_t \varphi$. Consider a derivation $\mathcal{D}$ of φ from Γ in the generic proof system. We are going to transform $\mathcal{D}$ in a derivation $\mathcal{D}'$ in the case-to-case proof system. Consider all non-discharged assumptions ψ of $\mathcal{D}$. Without loss of generality we may assume that all these assumption are accompanied by a tick $\surd$. We create $\mathcal{D}'$ by placing above each such assumption $\psi \in \Gamma$ the formula $\forall\psi$ and below the conclusion φ of $\mathcal{D}$ the formula $\forall\varphi$. Now, every $\forall\psi/\psi$ is a correct application of ($\forall$E) in $\mathcal{D}'$. Because all ψ have been ticked in $\mathcal{D}$ none is discharged, so in $\mathcal{D}'$ no conflict arises with the discharge rules. Moreover, in the derivation of φ from $\forall\Gamma$ every *generic* application of the rule ($\forall$I) in $\mathcal{D}$ has been turned in a standard application in $\mathcal{D}'$, for no undischarged assumption of $\mathcal{D}'$ has free variables. Finally, the proof step $\varphi/\forall\varphi$ is a correct application of the standard rule ($\forall$I), for if φ has free variables, then these do not occur free in the assumptions in $\forall\Gamma$. So $\mathcal{D}'$ is a derivation of $\forall\varphi$ from $\forall\Gamma$ in the case-to-case proof system. ⊠

2.12. PROPOSITION. *The generic proof system for classical logic is complete:* $\Gamma \models_t \varphi$ *implies* $\Gamma \vdash_t \varphi$.

Proof: This time, by $\Gamma \models_t \varphi \iff \forall\Gamma \models_c \forall\varphi$ and the soundness and completeness of the case-to-case proof system, completeness of the generic proof system with respect to generic consequence follows if we can show

$$\forall\Gamma \vdash_c \forall\varphi \Rightarrow \Gamma \vdash_t \varphi.$$

For then: $\Gamma \models_t \varphi \iff \forall\Gamma \models_c \forall\varphi \iff \forall\Gamma \vdash_c \forall\varphi \Rightarrow \Gamma \vdash_t \varphi$. So suppose $\forall\Gamma \vdash_c \forall\varphi$. Every derivation in the case-to-case sytem is a derivation in the generic proof system. So, $\forall\Gamma \vdash_c \forall\varphi$ implies $\forall\Gamma \vdash_t \forall\varphi$. Now consider a derivation $\mathcal{D}$ of $\forall\varphi$ from elements of $\forall\Gamma$. We are going to transform this into a derivation $\mathcal{D}'$ deriving φ from Γ in the generic proof system. We get $\mathcal{D}'$ by placing above every leaf of $\mathcal{D}$ labeled by an undischarged assumption $\forall\psi \in \forall\Gamma$, the assumption $\psi\ \surd$ and by placing φ below the conclusion $\forall\varphi$ of $\mathcal{D}$. Now every $\psi\ \surd/\forall\psi$ is a correct application of the generic rule ($\forall$I): because $\forall\psi$ is not discharged in $\mathcal{D}$, ψ is not discharged in $\mathcal{D}'$, so the tick $\surd$ is respected in $\mathcal{D}'$. Moreover $\forall\varphi/\varphi$ is a correct application of the rule ($\forall$E). Consequently $\Gamma \vdash_t \varphi$. ⊠

Interestingly, Van Dalen [Dal83] explicitly states that the natural deduction universal quantifier rules provide a connection between universal reasoning and 'generic reasoning':

> The reader will have grasped the technique behind the quantifier rules: reduce a $\forall x\varphi$ to φ and reintroduce $\forall$ later, if necessary. Intuitively, one makes the following step: to show "for all x $\ldots x \ldots$" it suffices to show "$\ldots x \ldots$" for an arbitrary x. The latter statement is easier to handle. Without going into fine philosophical distinctions, we note that the distinction "for all x $\ldots x \ldots$" — "for an arbitrary $x \ldots x \ldots$" is embodied in our system by the distinction "quantified statement" — "free variable statement". [Dal83, p. 95]

Here is a further connection between natural deduction and 'generic reasoning': the variables in the $\surd$ assumptions of a generic proof tree can be said to denote arbitrary individuals. The assertion that an arbitrary individual has property P can be viewed as a statement Px, with an indication (by means of the $\surd$ mark) that this assertion is meant as a universal statement.

2.13. REMARK. There is an obvious connection between the two notions of consequence discussed here, and the notions of a *global* and *local* consequence in *Modal Logic*. An inference Σ/φ is globally valid on a possible worlds model if, whenever Σ is true on every world, then so is φ. On the other hand, Σ/φ is locally valid on such a model if at every world where Σ is true, φ is also true. Global consequence corresponds to our truth-to-truth consequence: if φ is true on all worlds in a model, then $\Box\varphi$ is true on every world. So $\varphi \models_{\text{global}} \Box\varphi$. Again, this does not allow us to conclude $\models_{\text{global}} \varphi \to \Box\varphi$.

2.3 Intensional Epsilon Logic

2.3.1 Background

Epsilon terms were introduced by Hilbert and his collaborators (Ackermann, Bernays) in order to provide explicit definitions of the existential and universal quantifiers. This move was part of the formalistic program, with as its ultimate goal to legitimate non-constructive techniques in logic and mathematics by means of founding the whole edifice on a provably consistent basis. Of course, as we know now, this attempt has failed, but epsilon terms have turned out to be interesting in their own right.

2.3.2 Language

Languages of first order logic with epsilon terms (ϵ-terms) are defined as for the case of first order logic, with the difference that a new set of terms is added which gets defined recursively over formulas:

terms $t ::= c \mid v \mid ft_1 \cdots t_n \mid \epsilon v : \varphi$.
formulas $\varphi ::= \bot \mid Pt_1 \cdots t_n \mid t_1 = t_2 \mid \neg\varphi \mid (\varphi \wedge \psi) \mid (\varphi \vee \psi) \mid (\varphi_1 \rightarrow \varphi_2) \mid \forall v\varphi \mid \exists v\varphi$.

Note the simultaneous recursion on terms and formulas in the definition. If L is a language of first order logic, L^ϵ is the language with epsilon terms over the same non-logical vocabulary.

Abbreviations and notational conventions are as before, with one important difference. Without ϵ-terms, all variables occurring in a term t of the language occur free in t. In ϵ-terms however, variables can occur bound or free: the ϵ-symbol is a variable binding operator. So the notions of a variable *occurring free* in a formula, and of a term *occurring free for a variable in a formula* have to take the possibility into account that variables may occur in a term within the scope of an ϵ-symbol.

2.3.3 Semantics

To find a semantics for the language of the ϵ-calculus we shall start out with a first-order semantics and extend this by an interpretation for ϵ-terms. Every first-order model can be extended with such an interpretation. Epsilon terms can be viewed as a means for naming *Skolem functions*. Given a formula $\varphi(\overline{x}, y)$, with $\overline{x}$ a list of free variables, and y a designated free variable, we can look at the set of all objects b in a model which satisfy φ, given a choice $\overline{a}$ of values for the parameters $\overline{x}$. What a Skolem function does is to pick a particular object in this set as the value. The Skolem function corresponding to $\varphi(\overline{x}, y)$ can be given an arbitrary name. If we want to stress its connection with the formula it derives from, we can denote it as $F_\varphi(\overline{x})$. The epsilon calculus can be viewed as a proposal to name all Skolem functions explicitly, for now we write $F_\varphi(\overline{x})$ simply as $\epsilon y : \varphi(\overline{x}, y)$.

The general procedure to expand a predicate logical language with Skolem functions to interpret the epsilon terms is this. Start out with the original predicate logical language L, and then expand the language in layers:

1. $L = L_0$,
2. $L_{k+1} = L_k \cup \{F_\varphi(\overline{x}) \mid$ for some $z: \ \varphi(\overline{x}, y)[z/y] \in L_k\}$,
3. $L' = \bigcup_{k=0}^{\infty} L_k$.

So the language L_{k+1} is constructed from L_k by introducing for any formula $\varphi(\overline{x}, y)$ of L_k a new function symbol $F_\varphi(\overline{x})$ (or epsilon term $\epsilon y : \varphi(\overline{x}, y)$), and letting L_{k+1} consist of L with these new function symbols added. Note that formulas $\varphi(x_1, \ldots, x_n, y)$ and $\varphi(z_1, \ldots, z_n, z)$ give rise to the same Skolem function, as the variables serve only to identify argument slots.

Because the Skolem function names correspond one to one with epsilon terms, we have that $L' = L^{\epsilon}$.

To get from M to an expanded model M', a *Skolem expansion* of M, which interprets L', we proceed as follows. Start with $M = M_0$. This model interprets L_0. Now, given that we have defined M_k interpreting L_k, we choose for every formula $\varphi(\overline{x}, y)$ of L_k with exactly n free variables an interpretation $I_{\varphi(\overline{x})}$ extending the domain of $int(M_k)$ with $F_{\varphi(\overline{x})}$ such that $I_{\varphi(\overline{x})}(F_{\varphi(\overline{x})}) : dom(M)^n \mapsto dom(M)$ and

$$M_k, I_{\varphi(\overline{x})} \models \forall\overline{x}(\exists y\varphi(\overline{x}, y) \to \varphi(\overline{x}, y)[F_{\varphi(\overline{x})}(\overline{x})/y]).$$

We then set $int(M_{k+1}) = int(M_k) \cup \{I_{\varphi(\overline{x})} \mid \varphi(\overline{x}, y)\ inL_k\}$. We get the eventual model M' by gathering all construction stages together: $M' = \bigcup_{k=0}^{\infty} M_k$.

It is clear that in the extended model M' the interpretations of formulas in the original L are not affected. Therefore, for all $\psi \in L$: $M, s \models \psi$ if and only if $M', s \models \psi$.

2.14. REMARK. Actually, Skolem used his functions to construct *smaller* models on the basis of a given model. See Hodges [Hod93, Chapter 3] for more information. In the standard Skolem expansion argument it is usually assumed that the list of parameters $\overline{x}$ of a Skolem function F_{φ} is non-empty, i.e., that the Skolem function is a function rather than a constant. This is because Skolem constants are not so useful for reducing sizes of models. Skolem constants correspond to closed epsilon terms, i.e., epsilon terms of the form $\epsilon y : \varphi$, where φ has only y free. If epsilon terms are used to analyze quantification, then closed epsilon terms are quite useful.

2.15. REMARK. In a sense, our interpretation of Hilbert's epsilon terms commits a historical injustice, for our model-theoretic reasoning about Skolem expansions takes the whole concept of quantification for granted, and Hilbert's very reason for putting forward the epsilon rule was to provide a proof-theoretical analysis of the fine-structure of the use of the quantifiers in mathematical discourse. In fact, Hilbert did not care about the semantics of epsilon terms, and it is easy to see why. Hilbert was primarily interested in proof theory, as it was there that he hoped to find a firm foundation for the edifice of mathematics. Hilbert's proof of the conservativity of the epsilon rule over classical logic (his so-called 'Second Epsilon Theorem') proceeds by way of a detailed inspection of the use of epsilon terms in actual proofs and a proposal for a procedure for eliminating epsilon terms from proofs by means of proof transformation. For a clear presentation of this proof-theoretic means to arrive at the conservativity result we refer the reader to Leisenring [Lei69], which is in fact a logic textbook based on the epsilon calculus.

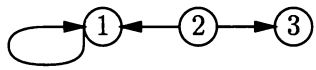

Figure 2.1: A two-place relation R.

Let us look at a concrete example of a formula and the Skolem functions it gives rise to in a given model. Consider the structure pictured in Figure 2.1. Assume that the arrow interprets the relation symbol R and consider the formula Rxy. Viewing x as parameter and y as designated variable this gives rise to the Skolem function or epsilon term $\epsilon y : Rxy$. Viewing y as parameter and x as designated variable we get $\epsilon x : Rxy$. Consider the epsilon term $\epsilon y : Rxy$. To check the leeway that we have for interpreting it, it is useful to look at the one place predicate given by $\lambda y.Rxy$, for all possible values of x. This gives:

$$\begin{array}{ll} x \mapsto 1 & \lambda y.Rxy = \{1\} \\ x \mapsto 2 & \lambda y.Rxy = \{1,3\} \\ x \mapsto 3 & \lambda y.Rxy = \emptyset. \end{array}$$

What this means is that in the case $x \mapsto 1$ we have no choice: we have to interpret $\epsilon y : Rxy$ as 1. In the case $x \mapsto 2$ we can either interpret $\epsilon y : Rxy$ as 1 or as 3. In case $x \mapsto 3$ the choice doesn't matter, for 3 has no outgoing arrows at all.

For the epsilon term $\epsilon x : Rxy$ we get a similar array:

$$\begin{array}{ll} y \mapsto 1 & \lambda x.Rxy = \{1,2\} \\ y \mapsto 2 & \lambda x.Rxy = \emptyset \\ y \mapsto 3 & \lambda x.Rxy = \{2\}. \end{array}$$

This means that in the case $y \mapsto 1$ we can either interpret $\epsilon x : Rxy$ as 1 or as 2, for these are the two objects with an outgoing arrow pointing to 1. In case $y \mapsto 2$ the choice doesn't matter, as 2 has no incoming arrows. In case $y \mapsto 3$ we are forced to interpret $\epsilon x : Rxy$ as 2, for this is the only object with an arrow pointing to 3.

Now consider the epsilon term $\epsilon x : R(x, \epsilon y : Rxy)$. This is a closed epsilon term, so its interpretation should not depend on the variable assignment. On the other hand it is not independent of our choice for the interpretation of the embedded epsilon term $\epsilon y : Rxy$. Suppose the interpretation of $\epsilon y : Rxy$ element is 1 if $x \mapsto 1$, is 1 if $x \mapsto 2$ and is 1 if $x \mapsto 3$. Then we know that the denotation of $\lambda x.R(x, \epsilon y : Rxy)$ equals the set $\{1,2\}$. In this case, the interpretation of $\epsilon x : R(x, \epsilon y : Rxy)$ will have to pick out a member of this set. Suppose on the other hand that the interpretation of $\epsilon y : Rxy$ is given by 1 for $x \mapsto 1$, 3 for $x \mapsto 2$, and 1 for $x \mapsto 3$. In this case the interpretation of $\epsilon x : R(x, \epsilon y : Rxy)$ can only be 1. The example illustrates the dependencies between interpretations of epsilon terms that may arise.

Given an expanded model M' based on M, we can separate out the difference between int (M') and int (M) and call it Φ.

2.16. DEFINITION. (Intensional Choice Functions) Let M be a model for language L, let M' be a Skolem expansion of model M. Then Φ is the mapping assigning Skolem functions in M' to ϵ-terms over L, given by

$$\Phi(\epsilon y : \varphi(\overline{x}, y)) = int(M')(F_{\varphi(\overline{x})}).$$

What Φ does is to provide interpretations for the epsilon terms, by mapping $\epsilon v : \varphi(v_1, \ldots, v_n, v)$ to an appropriate Skolem function F on dom (M). The mapping Φ over M is 'choice function' because it chooses appropriate values in dom (M) for an epsilon term with suitable arguments, it is 'intensional' because the values do not only depend on the extension of the epsilon formula $\varphi(v_1, \ldots, v_n, v)$ in the model M. The Skolem expansion argument in fact proves the existence of intensional choice functions. We shall use iM for the set of all intensional choice functions for M.

2.17. DEFINITION. (Valuation of Terms) Let M be a first-order model, s a variable assignment for M, and Φ an intensional choice function for M. Then the *term valuation* $V_{M,\Phi,s}$ *in* M *based on* Φ *and* s is given by the following clauses:

- $V_{M,\Phi,s}(c) = \text{int}\,(M)(c)$.
- $V_{M,\Phi,s}(v) = s(v)$.
- $V_{M,\Phi,s}(ft_1 \cdots t_n) = \text{int}\,(M)(f)(V_{M,s}(t_1), \ldots, V_{M,s}(t_n))$.
- $V_{M,\Phi,s}(\epsilon v : \varphi(v_1, \ldots, t_n, v)) = \Phi(\epsilon v : \varphi)(s(v_1), \ldots, s(v_n))$.

The term valuation is used to define the relation $M, \Phi, s \models \varphi$ in the standard way; the relation depends on Φ because it uses $V_{M,\Phi,s}$.

By inspecting the construction of the Skolem expansion we can check that the function Φ has the following *choice property*:

- If $N = \{d \in \text{dom}\,(M) \mid M, \Phi, s(v|d) \models \varphi(v_1, \ldots, v_n, v)\} \neq \emptyset$, then $\Phi^s(\epsilon v : \varphi)(s(v_1), \ldots, s(v_n)) \in N$.
- if $N = \{d \in \text{dom}\,(M) \mid M, \Phi, s(v|d) \models \varphi(v_1, \ldots, v_n, v)\} = \emptyset$, then $\Phi^s(\epsilon v : \varphi)(s(v_1), \ldots, s(v_n)) \in dom(M)$.

It follows that, for the interesting special case of the set of *closed* epsilon terms E_c, every first order model M can be supplied with a function Φ from E_c to the domain of M satisfying the choice property: for all $\epsilon x : \varphi \in E_c$,

$$\text{if } N = \{d \in \text{dom}\,(M) \mid M, \Phi \models \varphi(d)\} \neq \emptyset, \text{ then } \Phi(\epsilon v : \varphi) \in N.$$

Interestingly, a choice function Φ does for closed epsilon terms what a variable assignment function does for free variables. To get at the interpretation of an open epsilon term, such as $\epsilon x : Rxy$, we need both a variable assignment function and a choice function. A counterexample to the validity of an open formula of first order logic takes the form of a first order model plus an assignment function. Analogously, a counterexample to the validity of an open formula of first order logic with epsilon terms takes the form of a triple consisting of a first order model M, a choice function Φ, and an assignment function s. For instance, to refute $Rx(\epsilon y : y = y)$ it suffices to give a model M, an assignment s with $s(x) \mapsto a$, and a choice function $\Phi \in iM$ with $\Phi(\epsilon y : y = y) = b$ with the property that $\langle a, b\rangle \notin R^M$.

Quantifier-Free First-Order Logic

By their interpretation by Skolem functions, epsilon terms of the form $\epsilon x : \varphi$ and existentially quantified formulas of the form $\exists x\varphi$ are intimately connected. In fact, a first-order language with ϵ-terms may do without quantifiers altogether. Let φ be a formula of a first-order language with ϵ-terms, and let φ' be the result of replacing an occurrence of $\exists v\psi$ in φ, *not within the scope of an ϵ operator*, by an occurrence of $\psi(\epsilon v : \psi)$, and an occurrence of $\forall v\psi$ in φ, *not within the scope of an ϵ operator*, by an occurrence of $\psi(\epsilon v : \neg\psi)$. Use $\varphi \mapsto \varphi'$ to express that φ' is obtained from φ by such a replacement step. Then we have:

2.18. PROPOSITION. *If $\varphi \mapsto \varphi'$ then for all M, all $\Phi \in iM$, all $s \in aM$:*

$$M, \Phi, s \models \varphi \textit{ iff } M, \Phi, s \models \varphi'.$$

This proposition is a direct consequence of the truth definition and the interpretation of ϵ-terms. For,

$$M, \Phi, s \models \exists xRxy \text{ iff } M, \Phi, s \models R(\epsilon x : Rxy)y$$

and

$$M, \Phi, s \models \forall xRxy \text{ iff } M, \Phi, s \models R(\epsilon x : \neg Rxy)y.$$

For the latter equivalence, note that

$$\forall xRxy \leftrightarrow \neg\exists x\neg Rxy \leftrightarrow \neg\neg R(\epsilon x : \neg Rxy)y \leftrightarrow R(\epsilon x : \neg Rxy)y.$$

The reason for the restriction that the subformulas to be replaced should not occur in the scope of an epsilon operator is the following. Because the choice function Φ is intensional, there is no guarantee that $\epsilon y : \exists xRxy$ and $\epsilon y : R(\epsilon x : Rxy, y)$ have the same value under Φ.

Proposition 2.18 allows us to show that every first-order formula free of ϵ-terms has a quantifier-free equivalent. That is, an equivalent in which quantifiers do not occur on the level of the formula nor within the scope of an ϵ-operator (note that the formula $P(\epsilon x : \exists yQxy)$ is term-logical formula without being

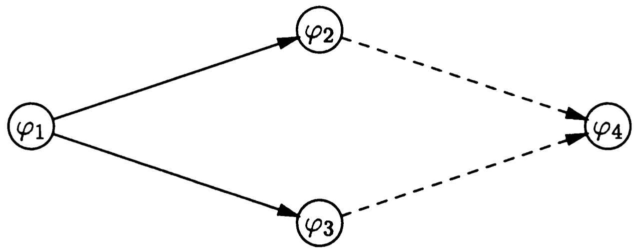

Figure 2.2: Church-Rosser property of minimal $\mapsto$ reduction.

strictly quantifier free). Let φ be a formula of first order logic without epsilon terms, and let φ^ϵ be the result of applying $\mapsto$ steps for subformulas $\exists v : \psi$ or $\forall v : \psi$ with ψ quantifier free (call such a move a minimal $\mapsto$ move), until the result does not contain subformulas of the form $\exists v\psi$ or $\forall v\psi$ anymore. Note that by working 'inside out' in this way, no quantifiers will ever end up in the scope of an epsilon operator as a result of a $\mapsto$ move. The procedure of iterating minimal $\mapsto$ steps allows the reduction to a unique 'normal form', because minimal $\mapsto$ reduction steps applied to different subformulas are independent and need never be repeated. That is, minimal reduction has the Church-Rosser property, in the sense that the products φ_2 and φ_3 of two minimal $\mapsto$ steps from φ_1 can always be reduced to a formula φ_4 by a minimal reduction step (see Figure 2.2). Also, the process of minimal $\mapsto$ reduction obviously terminates. It follows that φ^ϵ is a normal form under minimal $\mapsto$ reduction.

From the previous proposition we now get:

2.19. PROPOSITION. *If φ is a formula without epsilon terms and $\varphi \overset{*}{\mapsto} \psi$, then for all M, all $\Phi \in iM$, all $s \in aM$:*

$$M, s \models \varphi \textit{ iff } M, \Phi, s \models \psi.$$

By this proposition all first–order formulas have quantifier free equivalents in the ϵ-calculus.

But the language with epsilon terms is more expressive than an epsilon free first-order language. It follows from Proposition 2.19 that for all φ, ψ with φ epsilon free and $\varphi \overset{*}{\mapsto} \psi$, for all models M and assignments $s \in aM$:

- if $M, s \models \varphi$, then for *all* choice functions $\Phi \in iM$, $M, \Phi, s \models \psi$.
- if there is a choice function $\Phi \in iM$ for which $M, \Phi, s \models \psi$, then $M, s \models \varphi$.

2.3.4 Generic Truth

Now that we have a proper definition for $M, \Phi, s \models \varphi$, for $\varphi \in L^\epsilon$, we can define *generic* truth of a formula of the epsilon calculus.

2.20. DEFINITION. (Generic Truth) Formula $\varphi \in L^\epsilon$ is *generically true* on a model M with respect to variable assignment s, notation $M, s \models \varphi$, if

$$M, \Phi, s \models \varphi \text{ for all } \Phi \in iM.$$

This interpretation will be needed for the comparison with arbitrary object theory in Section 2.6.

We observe that the notions of Generic truth and truth simpliciter are different, even for closed formulas. Note that

$$M \models P(\epsilon x : Qx) \text{ implies } M \models_g \forall x(Qx \to Px),$$

while

$$M, \Phi \models P(\epsilon x : Qx) \text{ does not imply } M, \Phi, \models_l \forall x(Qx \to Px).$$

If Γ is a set of formulas (possibly with epsilon terms), then $M, \Phi, s \models \Gamma$ holds by definition iff $M, \Phi, s \models \gamma$ holds for all $\gamma \in \Gamma$, and similarly for $M, s \models \Gamma$.

2.3.5 Logical Consequence

Recall that for first order logic we did have the choice between case-to-case and truth-to-truth notions of consequence. This distinction was due to two different ways of quantifying over the assignment functions. Similarly, we can quantify in different ways over choice functions.

In fact, the standard and the generic interpretation of epsilon formulas suggest the following two possibilities for a consequence notion for such formulas.

2.21. DEFINITION. (Local consequence) $\Gamma \models_l \varphi$ iff for all M, all Φ, and all s: $M, \Phi, s \models \Gamma$ implies $M, \Phi, s \models \varphi$.

2.22. DEFINITION. (Generic consequence) $\Gamma \models_g \varphi$ iff for all M: $M \models \Gamma$ implies $M \models \varphi$.

The previous example can also be used to illustrate that local consequence and generic consequence are different, for we have:

$$P(\epsilon x : Qx) \models_g \forall x(Qx \to Px),$$

versus

$$P(\epsilon x : Qx) \not\models_l \forall x(Qx \to Px).$$

Let us unravel the two consequence notions a bit, in order to compare them. Local consequence says in fact the following:

$$\Gamma \models_l \varphi \text{ iff } \forall M \ \forall \Phi \in iM \ \forall s \in aM(\forall \gamma \in \Gamma : M, \Phi, s \models \gamma \Rightarrow M, \Phi, s \models \varphi).$$

It turns out that the quantification over choice functions is similar to that over assignment functions in the case-to-case notion of first order consequence.

The notion of local validity derives from this as the special case where Γ is empty. We get:

$$\models_l \varphi \text{ iff } \forall M\ \forall\Phi \in iM\ \forall s \in aM : M, \Phi, s \models \varphi.$$

As an example of a formula which is not locally valid, consider

$$\varphi := \exists x \exists y Rxy \to R(\epsilon x : \exists y Rxy, \epsilon y : \exists x Rxy).$$

It is obvious that for every M there is a $\Phi \in iM$ with $M, \Phi \models \varphi$, but note that this is not enough for local validity. The requirement for that is much stronger: all pairs M, Φ have to satisfy the formula. For a counterexample, consider a model M with $R^M = \{\langle d_1, d_2\rangle, \langle d_3, d_4\rangle\}$, and a choice function Φ with:

$$(\epsilon x : \exists y Rxy) \mapsto d_1, (\epsilon y : \exists x Rxy) \mapsto d_4.$$

Generic consequence says something quite different from local consequence:

$$\Gamma \models_g \varphi \quad \text{iff} \quad \forall M(\forall\gamma \in \Gamma\ \forall\Phi \in iM\ \forall s \in aM : M, \Phi, s \models \gamma \Rightarrow \\ \forall\Phi \in iM\ \forall s \in aM : M, \Phi, s \models \varphi).$$

Observe that generic consequence quantifies over choice functions in precisely the way in which truth-to-truth consequence for classical first order logic quantifies over assignment functions.

Again, the notion of generic validity derives from this as the special case where Γ is empty:

$$\models_g \varphi \text{ iff } \forall M\ \forall s \in aM\ \forall\Phi \in iM : M, \Phi, s \models \varphi.$$

Note that it follows from Proposition 2.19 that for closed formulas of the fragment L° of an epsilon language L^ϵ the notions of local consequence and generic consequence coincide. And if we confine attention to closed premise sets without epsilon terms, the two notions collapse again to first order consequence without further ado.

A Proof System for Generic Consequence

For a proof system for generic consequence we take our cue from the notion of classical generic consequence (Section 2.2.3). Again we use the device $\surd$ to mark assumptions in a derivation that are not to be discharged. Now consider the following rule:

$$\frac{\psi[\epsilon x : \varphi/x]}{\forall x(\varphi \to \psi) \wedge (\neg\exists x\varphi \to \forall x\psi)}$$

provided all assumptions containing $\epsilon x : \varphi$ above $\psi[\epsilon x : \varphi/x]$ are marked by $\surd$.

We shall denote this rule by $(\forall I_g)$. First of all, the rule $(\forall I_g)$ is generically sound. For suppose $M, s \models_g \psi[\epsilon x : \varphi/x]$. Assume $M, s \not\models_g \forall x(\varphi \to \psi)$. That is, there is an $m \in dom(M)$ such that $M, s(m|x) \models_g \varphi \wedge \neg\psi$. Then we can take m as the value of some Φ on $\epsilon x : \varphi$, so $M, s, \Phi \not\models_g \psi[\epsilon x : \varphi/x]$: a contradiction. Consequently $\psi[\epsilon x : \varphi/x] \models_g \forall x(\varphi \to \psi)$.
Now assume $M, s \models_g \neg\exists x\varphi$. Then for every $m \in dom(M)$ there is a choice function Φ such that $V_{M,s,\Phi}(\epsilon x : \varphi) = m$. But for all Φ: $M, s, \Phi \models_l \psi[\epsilon x : \varphi]$, so $M, s \models_g \forall x\psi$. Consequently $\psi[\epsilon x : \varphi/x] \models_g \neg\exists x\varphi \to \psi[\epsilon x : \varphi/x]$.

We shall denote the derivability notion arising by adding the rule $(\forall I_g)$ to the natural deduction calculus for intensional epsilon logic by $\vdash_g$.

2.23. PROPOSITION. (Soundness of generic consequence) *The generic proof system for generic epsilon logic is sound:* $\Sigma \vdash_g \varphi \Rightarrow \Sigma \models_g \varphi$.

Proof: Note that $\Sigma \vdash_l \varphi \Rightarrow \Sigma \models_g \varphi$. This follows by the soundness of $\vdash_l$ with respect to $\models_l$ and the definition of $\models_g$. So, if we can reduce a generic epsilon derivation $\mathcal{D}$ to a derivation free of applications of the new rule $(\forall I_t)$, then we have shown soundness.

Let $(\psi[\epsilon x : \varphi/x])^*$ be the formula $\forall x(\varphi \to \psi) \wedge (\neg\exists x\varphi \to \forall x\psi)$, and let $\mathcal{D}$ be a generic epsilon derivation. Suppose $(\psi[\epsilon x : \varphi/x])^*$ is the conclusion of a *highest* application, of the rule $(\forall I_t)$ in $\mathcal{D}$. If we show we can eliminate this application, then we have shown we can eliminate all applications. We proceed as follows. Above all ticked assumptions on which $\psi[\epsilon x : \varphi/x]$ depends of the form $\chi[\epsilon x : \varphi/x]$ we place their $\vdash_l$-derivation from the starred translations. These derivations exist.

2.24. PROPOSITION. $(\chi[\epsilon x : \varphi/x])^* \vdash_l \chi[\epsilon x : \varphi/x]$.

Proof: The first conjunct of $(\chi[\epsilon x : \varphi/x])^*$ derives $\exists x\varphi \to \chi[\epsilon x : \varphi/x]$ (by $(\forall\text{E})$ and $(\exists E_\epsilon)$), the second derives $\neg\exists x\varphi \to \chi[\epsilon x : \varphi/x]$. So, by excluded middle, $(\chi[\epsilon x : \varphi/x])^* \vdash_l \chi[\epsilon x : \varphi/x]$. ⊠

This gives $\mathcal{D}'$, a case-to-case derivation of $\psi[\epsilon x : \varphi/x]$ from ϵx : *φ-free assumption* set Σ'. Now, we have in general

2.25. PROPOSITION. *If Σ has no formula of the form $\chi[\epsilon x : \varphi/x]$, then*

$$\Sigma \vdash_l \psi[\epsilon x : \varphi/x] \Rightarrow \Sigma \vdash_l \forall x(\varphi \to \psi) \wedge (\neg\exists x\varphi \to \forall x\psi).$$

Proof: We reason semantically and use completeness of $\vdash_l$ with respect to $\models_l$. Suppose $M, s, \Phi \models \Sigma$, but $M, s, \Phi \not\models \forall x(\varphi \to \psi)$. The counterexample m such that $M, s(x|m), \Phi \models \varphi \wedge \neg\psi$ can be chosen as the value of a $\epsilon x : \varphi$-variant Φ' of Φ which agrees with Φ on all ϵ-terms in Σ. Thus $M, s, \Phi' \models \Sigma$ but $M, s, \Phi' \not\models \psi[\epsilon x : \varphi/x]$. So, by completeness $\Sigma \not\vdash_c \psi[\epsilon x : \varphi/x]$. The proof for the second conjunct proceeds analogously. ⊠

So now we know that in derivation $\mathcal{D}'$ we can replace the application of $(\forall I_t)$ with premise $\psi[\epsilon x : \varphi/x]$ by a local derivation. In the resulting derivation we have no application of $(\forall I_t)$ and we have $\Sigma' \models_t (\psi[\epsilon x : \varphi/x])^*$. But, by the truth definition, $M, s \models_t \Sigma$ iff $M, s \models_t \Sigma'$, so we have shown soundness for $\mathcal{D}$. ⊠

2.26. PROPOSITION. (Completenes of generic epsilon consequence) *The generic proof system is complete with respect to generic epsilon logic:* $\Sigma \models_g \varphi \Rightarrow \Sigma \vdash_g \varphi$.

Proof: Suppose $\Sigma \not\vdash_g \varphi$. As usual we construct a model for the consistent set $\Sigma \cup \{\neg\varphi\}$. By standard means we extend $\Sigma \cup \{\neg\varphi\}$ to a maximally consistent set Σ'. This set is witnessing, due to the presence of the ϵ-rule. The closed terms in Σ constitute the domain of the model $M_{\Sigma'}$. The interpretation function $int(M_{\Sigma'})$ is read of from Σ'. This give the first-order model. Now for the choice functions. These we build by induction on the level of embedding of closed ϵ-terms. An ϵ-term $\epsilon x : \varphi$ has *depth* 1 if the ϵ-symbol does not occur in φ. Otherwise it has the maximal depth of ϵ-terms occurring in φ, plus 1.

Let Φ^n, $1 \leq n \leq \omega$, be a function with domain E^n, the set closed ϵ-terms of depth at most n, and range $dom(M_{\Sigma'})$, such that, for $\epsilon x : \varphi \in E^n$:

if $\Phi^n(\epsilon x : \varphi) = t$ then *either* $\varphi[\Phi^n(t_1)/t_1 \ldots \Phi^n(t_k)/t_k, t/x] \in \Sigma'$,
or $\neg\exists x\varphi[\Phi^n(t_1)/t_1 \ldots \Phi^n(t_k)/t_k] \in \Sigma'$.

Here, $t_1, \ldots, t_k$ are all closed ϵ-terms occurring in φ not within the scope of an ϵ-symbol.

Let C^n be the set of all such choice functions with domain E^n. We create the set C of choice functions over $M_{\Sigma'}$ by considering $C = \overline{\bigcup}_{n<\omega} E^n$, where $\overline{\cup}$ denotes 'functional' union. That is, $\Phi_1 \overline{\cup} \Phi_2$ is $\Phi_1 \cup \Phi_2$ if this is a function, otherwise $\Phi_1 \overline{\cup} \Phi_2$ is undefined.

We now show the basic step of the *truth lemma*. All other steps proceed completely standard. Let A be an atomic formula. Then

$$M_{\Sigma'}, s, \models_g A \iff A \in \Sigma'.$$

Proof: Suppose A is of the form $B[\epsilon x : \varphi/x]$ for closed $\epsilon x : \varphi$ where, for simplicity, B has no ϵ-terms.
$\Rightarrow$: Let $M, s \models B[\epsilon x : \varphi/x]$. This means that for all $\Phi \in C$: $M, s, \Phi \models B[\epsilon x : \varphi/x]$ and so, by definition, for all $\Phi \in C$: $A[\Phi(\epsilon x : \varphi)/\epsilon x : \varphi] \in \Sigma'$. But there is always a Φ such that $\Phi(\epsilon x : \varphi) = \epsilon x : \varphi$. For *either* $M, s, \Phi \models \exists x\varphi$ and then $\varphi[\epsilon x : \varphi/x] \in \Sigma'$, so $B[\epsilon x : \varphi] \in \Sigma'$, *or* $M, s, \Phi \models \neg\exists x\varphi$ and $\Phi(\epsilon x : \varphi)$ can be arbitrarily chosen, so also as $\epsilon x : \varphi$.
$\Leftarrow$: Now suppose $B[\epsilon x : \varphi/x] \in \Sigma'$, so, by the rule $(\forall I_t)$ and deductive closure of Σ', $(B[\epsilon x : \varphi/x])^* \in \Sigma'$. Now take any $\Phi \in C$. *Either* $\exists x\varphi[\Phi(t_1/t_1 \ldots \Phi(t_k)/t_k] \in \Sigma'$ and the first conjunct of the starred formula (and the ϵ-rule) give $B[\epsilon x :$

$\varphi/x] \in \Sigma'$, *or* $\neg\exists x\varphi[\Phi(t_1/t_1 \ldots \Phi(t_k)/t_k] \in \Sigma'$ and the second conjunct gives $B[\epsilon x : \varphi/x] \in \Sigma'$. Consequently $M, s, \Phi \models B[\epsilon x : \varphi/x]$ for this arbitrary function Φ. ⊠

2.3.6 Expressivity

The matter of expressivity of the epsilon language we can address as follows. Let L^ϵ be a first order language L extended with epsilon terms. Let L^* be some epsilon free extension of L, possibly higher order, that is, possibly allowing quantification over function or predicated variables. L^ϵ is *at least* as expressive as L^* if for all $\varphi \in L^*$, there is a $\psi \in L^\epsilon$ such that for all $\chi \in L^*$, $\varphi \models \chi$ iff $\psi \models_l \chi$. And L^ϵ is *at most* as expressive as L^* if for all $\varphi \in L^\epsilon$, there is a $\psi \in L^*$ such that for all $\chi \in L^*$, $\varphi \models_l \chi$ iff $\psi \models \chi$. Now L^ϵ has the expressiveness of L^* if L^ϵ is at least, and at most as expressive as L^*.

2.27. DEFINITION. (Epsilon Free Equivalents) For $\varphi \in L^*$ a formula free of epsilon terms, the set $EQ(\varphi)$ of φ-equivalents is given by

$$EQ(\varphi) = \{\psi \in L^\epsilon \mid \forall M, \forall s(M, s \models \varphi \iff \exists\Phi : M, \Phi, s \models \psi)\}.$$

Notice that $\psi \in EQ(\varphi)$ implies $\psi \models_l \varphi$. The converse is generally not the case. The relevance of this definition for expressiveness is shown by the following proposition.

2.28. PROPOSITION. *Let $\varphi, \chi \in L^*$ be epsilon free and let $\psi \in EQ(\varphi)$. Then:*

$$\varphi \models \chi \iff \psi \models_l \chi.$$

Proof: From right to left: suppose $\varphi \not\models \chi$. So there is a model M and a variable assignment s such that $M, s \models \varphi$ and $M, s \not\models \chi$. Because $\psi \in EQ(\varphi)$, there is a choice function Φ such that $M, \Phi, s \models \psi$ and because χ has no ϵ-terms, $M, \Phi, s \not\models \chi$. Consequently $\psi \not\models_l \chi$.

From left to right. Suppose $\psi \not\models_l \chi$. Thus there is a model M, a choice function Φ and a variable assignment s such that $M, \Phi, s \models \psi$ and $M, \Phi, s \not\models \chi$. Again, because $\psi \in EQ(\varphi)$ and φ and χ contain no ϵ-terms, we have $M, s \models \varphi$ and $M, s \not\models \chi$. Consequently $\varphi \not\models \chi$. ⊠

By this simple proposition we can determine when a formula ψ from the epsilon calculus has the same epsilon free consequences as an epsilon free formula φ.

2.29. EXAMPLE. (Some epsilon free equivalents) First we shall consider some examples in which L^* is the first-order language L. In all these cases, the fact that $\psi \in EQ(\varphi)$ is easy to see.

1. $R(\epsilon x : \exists yRxy)(\epsilon y : \exists xRxy) \in EQ(\exists x\exists yRxy)$. So for any epsilon free χ: $R(\epsilon x : \exists yRxy)(\epsilon y : \exists xRxy) \models \chi \iff \exists x\exists yRxy \models \chi$. Moreover $R(\epsilon x : \exists yRxy)(\epsilon y : \exists xRxy) \models_l \exists x\exists yRxy$ but $\exists x\exists yRxy \not\models_l R(\epsilon x : \exists yRxy)(\epsilon y : \exists xRxy)$.
2. $\exists xPx \wedge Q(\epsilon x : Px), \exists xQx \wedge P(\epsilon x : Qx) \in EQ(\exists x(Px \wedge Qx))$. Again these formulas have the same epsilon free first-order consequences. Note $\exists x(Px \wedge Qx) \not\models_l \exists xPx \wedge Q(\epsilon x : Px)$.
3. $\forall x(\varphi \to \psi[\epsilon x : \neg(\varphi \wedge \neg\psi)/x]) \in EQ(\forall x(\varphi \to \psi))$. Also here, these formulas have the same epsilon free first-order consequences and $\forall x(\varphi \to \psi) \not\models_l \forall x(\varphi \to \psi[\epsilon x : \neg(\varphi \wedge \neg\psi)/x])$.

Now a case where the language L^* is actually an extension of the language L.

4. $\forall x\forall yRx(\epsilon u : \forall y\exists vRxuyv)y(\epsilon v : \forall x\exists uRxuyv) \in EQ(\varphi)$ where φ is the formula
$$\left.\begin{matrix}\forall x\exists u\\ \forall y\exists v\end{matrix}\right\rangle R(x,u,y,v).$$
Here we need some argumentation that the formulas are equivalents with respect to epsilon free consequences. Note that the term $(\epsilon u : \forall y\exists vRxuyv)$ has only x free, while $(\epsilon v : \forall x\exists uRxuyv)$ has only y free. Thus, given some choice function Φ, the epsilon terms denote one-argument functions F, G. Also, if
$$\forall x\forall yRx(\epsilon u : \forall y\exists vRxuyv)y(\epsilon v : \forall x\exists uRxuyv)$$
is true in some model M (for some choice function Φ), then:
$$M, x \mapsto a, u \mapsto F(a) \models \forall y\exists vRxuyv$$
will be true, and so will
$$M, y \mapsto b, v \mapsto G(b) \models \forall x\exists uRxuyv.$$
But this is precisely what φ asserts. Indeed, it is well known that formula φ is equivalent to the following second order sentence (cf. Barwise [Bar79]):
$$\exists F\exists G\forall x\forall yRx(Fx)y(Gy).$$
The epsilon version we have given above merely spells out a recipe for the two functions F and G.

From the procedure in the last example we can extract a recipe to transform any formula with a finite partially ordered quantifier prefix into an equivalent L^ϵ formula. This leads to our conjecture about the expressive power of the epsilon calculus.

If L is a first order language, then the fragment of second order logic L^{sk} over the same non-logical vocabulary is defined as follows:

terms $t ::= c \mid v \mid ft_1 \cdots t_n \mid Ft_1 \cdots t_n$.
formulas $\varphi ::= Pt_1 \cdots t_n \mid t_1 = t_2 \mid \neg\varphi \mid (\varphi \wedge \psi) \mid (\varphi \vee \psi) \mid (\varphi_1 \rightarrow \varphi_2) \mid \forall v\varphi \mid \exists v\varphi$.
Skolem-existentials $S ::= \varphi \mid \exists FS$.

All branching patterns are expressible as closed formulas of this fragment.

Conjecture *For every $\varphi \in L^{sk}$ there is a $\psi \in L^{\epsilon}$ such that $\psi \in EQ(\varphi)$ and for every $\varphi \in L^{\epsilon}$ there is a $\psi \in L^{sk}$ such that $\varphi \in EQ(\psi)$.*

2.3.7 Proof Theory for the Intensional Epsilon Calculus

In this section we shall introduce various proof theories for the intensional epsilon calculus. We start with the most common approach, the axiomatic one.

Axiomatic Deduction

Hilbert proposed to add the following *axiom schema* to the axioms of classical first order logic.

$$\exists v\varphi \rightarrow \varphi[\epsilon v : \varphi/v].$$

We shall call this the *epsilon axiom.* Intuitively, the term $\epsilon v : \varphi$ denotes an arbitrary object a in the domain of discourse which has property φ, if there are such objects at all, and an arbitrary object of the domain *tout court* if there aren't such objects.

By means of the introduction rule for the existential quantifier the converse direction of the epsilon rule is easily derived. This gives the following equivalence:

$$\exists v\varphi \leftrightarrow \varphi[\epsilon v : \varphi/v].$$

Hilbert's proposal was to view this equivalence as a definition of existential quantification (see Hilbert and Bernays [HB39] for additional motivation and for a presentation in a Hilbert style axiomatic framework).

Adding a rule to the schemata for classical first order logic is extending the logic. The first natural question which arises is: How does the result of adding the epsilon rule to the calculus of first order logic relate to standard classical first order logic? The answer is given by the following theorem.

2.30. THEOREM. (Hilbert's second ϵ-theorem) *Adding the epsilon rule is a* conservative *extension of classical predicate logic.*

Proof: We shall argue semantically. Suppose a formula is not derivable in predicate logic, i.e., suppose we have $\not\vdash \psi$. Then by the completeness of predicate logic, there exist a model M and assignment function s which refute ψ, i.e., $M, s \not\models \psi$. The Skolem expansion argument from the previous section shows that M can be expanded with an intensional choice function Φ which does not affect the formulas without epsilon terms. So $M, s \not\models \psi$ implies that there is a Φ with $M, \Phi \not\models \psi$. Because the extension of the proof system for first order logic with the epsilon rule is obviously sound, this yields that $\nvdash_\epsilon \psi$. In other words, the epsilon calculus does not allow the derivation of any new (epsilon-free) formulas. ⊠

For sake of completeness we shall also mention Hilbert's first ϵ-theorem. This theorem, closely related to Herbrand's Theorem, does not mention ϵ-terms, but its proof was formulated by Hilbert within the context of the ϵ-calculus. Let $\Sigma \vdash_p \varphi$ denote the fact that φ can be derived from Σ by means of only *propositional rules*.

2.31. THEOREM. (Hilbert's first ϵ-theorem) *If Σ is any set of first-order formulas in prenex form and φ a first-order formula in prenex form such that $\Sigma \vdash \varphi$, then there is a set Σ' and a disjunction $\varphi_1 \vee \ldots \vee \varphi_n$ where each member of Σ' is a substitution instance of the matrix of some member of Σ and each φ_i is a substitution instance of the matrix of φ and $\Sigma' \vdash_p \varphi_1 \vee \ldots \vee \varphi_n$.*

The addition of the epsilon rule to classical logic axiomatizes the interpretation of epsilon logic by models with respect to one choice function and variable assignment. That it is complete with respect to the standard interpretation has been shown by Leisenring [Lei69]. This proof follows the standard Henkin construction with some additions to account for the interpretation of ϵ-terms. By the standard construction, any consistent set of closed epsilon formulas can be extended to a maximally consistent such set Γ. Notice that such a set is always witnessing by the epsilon axiom. The domain of the model for the consistent set is constructed from equivalence classes of closed terms. To interpret *closed* ϵ-terms, let a subset N of the domain be *representable* if there is a formula $\varphi(x)$ with only x free, such that $N = \{t \mid \varphi[t/x] \in \Gamma\}$. In this case we call $\varphi(x)$ a formula *representing* N. If $\varphi(x)$ is a formula representing N, and $\varphi[t/x] \in \Gamma$, then the equivalence class of $\epsilon x : \varphi$ is an element of N. This follows by existential generalization and the epsilon axiom. So we can define a function mapping all representable sets to the (equivalence classes of) closed ϵ-terms of some representing formula. Thus we have given an interpretation of all closed ϵ-terms. The interpretation of ϵ-terms in general by Skolem functions can be defined from this in a straightforward way.

The addition of the epsilon axiom to a standard first-order axiomatization is not complete with respect to the *generic* notion of consequence. The following *generic rule*, for instance, is not derivable in Hilbert's epsilon calculus.

$$P(\epsilon x : Qx) \rightarrow \forall x(Qx \rightarrow Px).$$

In fact, any model M with a choice function Φ mapping $\epsilon x : \psi$ outside of $\{m \in dom(M) \mid M, \Phi, s(x|m) \models \varphi\}$ provides a counterexample, so it follows from the soundness of the calculus that the principle cannot be derived.

What we have seen is that *adding* the epsilon rule to classical first order logic is conservative (in Chapter 3 we shall see that adding it to intuitionistic logic is not). In fact, it can *replace* the usual rule for the existential quantifier

$$\frac{\varphi[t/x] \rightarrow \psi}{\exists x\varphi \rightarrow \psi}$$

(in axiomatic calculi which have this rule).

Proof Theory for Quantifier-Free First-Order Logic

In the epsilon calculus we can do without quantifiers altogether. Hilbert observed that we can add an axiom (schema) to the quantifier-free epsilon calculus which gives a proof theory deriving all first-order theorems in their quantifier-free form. This schema is the following.

2.32. DEFINITION. (Epsilon term rule)

$$\varphi[t/v] \rightarrow \varphi[(\epsilon v : \varphi/v)].$$

To get at the first-order correspondents of quantifier-free epsilon formulas we define $\mathcal{E}x\varphi$ as $\varphi[\epsilon x : \varphi/x]$ and $\mathcal{A}x\varphi$ as $\varphi[\epsilon x : \neg\varphi/x]$. This definition ensures that the operator $\mathcal{E}x$ ($\mathcal{A}x$) binds precisely the same occurrences of x in $\mathcal{E}x\varphi$ ($\mathcal{A}x\varphi$) as $\exists x$ does in $\exists x\varphi$ (and $\forall x$ in $\forall x\varphi$). We shall show that $\mathcal{E}$ coincides with $\exists$ and $\mathcal{A}$ with $\forall$. Let us use $\vdash_{\epsilon t}$ for derivability in the epsilon term calculus, i.e., in the calculus which has the classical propositional tautologies plus modus ponens and the epsilon term rule.

2.33. PROPOSITION. *Let $\varphi(\exists, \forall)$ be some first-order formula and $\varphi(\mathcal{E}, \mathcal{A})$ be the same formula with all occurrences of $\exists$ replaced by $\mathcal{E}$ and all occurrences of $\forall$ replaced by $\mathcal{A}$. Then*

$$\vdash \varphi(\exists, \forall) \iff \vdash_{\epsilon t} \varphi(\mathcal{E}, \mathcal{A}).$$

Proof: To check the direction from left to right, we have to show that the quantifier axioms of classical predicate logic are derivable in the epsilon term calculus. Take for instance the quantifier axiom forms from Enderton [End72]. Enderton takes $\forall$, $\neg$ and $\rightarrow$ as his primitive constants, and introduces the other logical constants by abbreviation. This allows him to get by with the following simple set of axiom schemes:

1.	$\forall x\varphi \rightarrow \varphi[t/x]$	t substitutable for x in φ
2.	$\forall x(\varphi \rightarrow \psi) \rightarrow (\forall x\varphi \rightarrow \forall x\psi)$	
3.	$\varphi \rightarrow \forall x\varphi$	x not free in φ.

To derive the first, note that, in classical logic, $\mathcal{A}x\varphi \rightarrow \varphi(t)$ is just the contraposition of $\neg\varphi(t) \rightarrow \mathcal{E}x\neg\varphi$, which in turn is an abbreviation of $\neg\varphi(t) \rightarrow \neg\varphi(\epsilon x : \neg\varphi)$, which is an instance of the epsilon term rule.

To derive the second, note that the epsilon term calculus inherits the deduction theorem from propositional logic. Therefore, it is enough to show the following:

$$\mathcal{A}x(\varphi \to \psi), \mathcal{A}x\varphi \vdash_{\epsilon t} \mathcal{A}x\psi.$$

Assume $\mathcal{A}x(\varphi \to \psi)$ and $\mathcal{A}x\varphi$. Use the axiom schema $\mathcal{A}x\varphi \to \varphi(t)$ applied for the term $\epsilon x : \neg\psi$ to derive:

$$(\varphi \to \psi)[\epsilon x : \neg\psi/x].$$

By the principles of substitution this is the same as:

$$\varphi[\epsilon x : \neg\psi/x] \to \psi[\epsilon x : \neg\psi/x].$$

Applying the same schema with the same term to $\mathcal{A}x\varphi$ and using modus ponens gives $\psi[\epsilon x : \neg\psi/x]$, which is the unabbreviated form of the desired result $\mathcal{A}x\psi$.

To see that the third schema is derivable, note that if x does not occur in φ, $\varphi[\epsilon x : \neg\varphi/x]$ is in fact the same as φ, so we can derive $\varphi \to \mathcal{A}x\varphi$ from the tautology $\varphi \to \varphi$.

These are the only three quantifier forms of classical predicate logic, so this establishes the direction from left to right.

For the other direction, we can reason semantically. Assume $\nvdash \varphi(\exists, \forall)$. Then, because of the completeness of predicate logic, there exists a counterexample for φ, i.e., a model M and an assignment s with $M, s \not\models \varphi(\exists, \forall)$. Construct a Skolem expansion M' for M as before. Note that the construction guarantees for every formula ψ of the expanded language L' (which has names for all the Skolem functions, or, equivalently, which has all the epsilon terms thrown in) that:

$$M', s \models \exists x\psi \leftrightarrow \psi(\epsilon x : \psi).$$

But this means that in φ we can safely replace each occurrence of $\exists$ by an occurrence of $\mathcal{E}$, and each occurrence of $\forall$ by an occurrence of $\mathcal{A}$. This proves that $M', s \not\models \varphi(\mathcal{E}, \mathcal{A})$, and thus, as the epsilon term calculus is obviously sound, $\nvdash_{\epsilon t} \varphi(\mathcal{E}, \mathcal{A})$. We conclude that $\vdash_{\epsilon t} \varphi(\mathcal{E}, \mathcal{A})$ implies $\vdash \varphi(\exists, \forall)$. This establishes the direction from right to left. ⊠

2.34. REMARK. The instances of the epsilon term schema are known as *critical formulas.* They play an essential role in the transformation of first-order deductions into deductions in a quantifier free format. Any set of first-order formulas can be translated into a set of quantifier free epsilon formulas, and any first-order derivation can be transformed in a quantifier free ϵ-derivation in which a finite number of critical formulas are used. To show that any use of a critical formula can be eliminated is to show that quantifiers can be eliminated (see Hilbert [HB39], Mints [Min94], Tait [Tai65]).

2.3.8 Natural Deduction for Intensional Epsilon Logic

Stated in natural deduction format, the epsilon axiom schema becomes the rule:

$$\frac{\exists v\varphi}{\varphi[\epsilon v : \varphi/v]}\ \exists\mathrm{E}_\epsilon$$

We also need a mechanism for for renaming bound variables:

$$\frac{}{(\epsilon v : \varphi) = (\epsilon w : \varphi[w/v])}$$

(Notice that by our definition of substitution, no clash of variables can occur.) These rules are simply added to the our natural deduction system for first-order logic. However, the presence of epsilon terms in deductions calls for an extra condition on the rule $\forall I$:

2.3.1. CONDITION. *If the proper term t of an application of $\forall I$ is an epsilon term, then t should not be the result of an application of $\exists E_\epsilon$.*

The condition rules out derivations like the following:

$$\frac{\dfrac{\exists xPx}{P(\epsilon x : Px)}\ \exists\mathrm{E}_\epsilon}{\forall xPx}\ \forall\mathrm{I}$$

2.35. EXAMPLE. (Permutation of Existential Quantifiers) We repeat our previous example of a deduction of $\exists y\exists xRxx$ from $\exists x\exists yRxy$, this time using the $\exists E_\epsilon$ rule instead of $\exists E$:

$$\frac{\dfrac{\dfrac{\dfrac{\exists x\exists yRxy}{\exists yR(\epsilon x : \exists yRxy)y}\ \exists\mathrm{E}_\epsilon}{R(\epsilon x : \exists yRxy)(\epsilon y : R(\epsilon x : \exists yRxy)y)}\ \exists\mathrm{E}_\epsilon}{\exists xRx(\epsilon y : R(\epsilon x : \exists yRxy)y)}\ \exists\mathrm{I}}{\exists y\exists xRxy}\ \exists\mathrm{I}$$

Lemma 3.14 in Chapter 3 will show that this calculus is conservative over the natural deduction format CPL of this chapter. Moreover, this lemma shows that the rule ($\exists E_\epsilon$) can even *replace* the rule ($\exists$E). As in the previous Section, we can also formulate the epsilon rule as a term rule:

$$\frac{\varphi[t/v]}{\varphi[(\epsilon v : \varphi)/v]}\ \epsilon\mathrm{I} \tag{2.2}$$

The reason for stating the premise in the form $\varphi[t/v]$ was explained in Section 2.2.2. The conclusion $\varphi[(\epsilon v : \varphi)/v]$ denotes the result of substituting $\epsilon v : \varphi$

for v in φ. Note again that our convention about substitution assumes that proper care is taken to prevent accidental capture of free variables in $\epsilon v : \varphi$ by quantifiers in φ. The formulation of the rule also ensures that the ϵv operator cannot accidentally bind occurrences of v that are in the scope of another binding operator, simply because the epsilon term is only substituted for occurrences of v that were free in the first place.

Some examples should make all this clearer.

2.36. EXAMPLE. (Correct and incorrect derivations) Here is a correct application of ϵI:

$$\frac{Rx(\epsilon y : Rxy)}{R(\epsilon z : Rz(\epsilon y : Rxy))(\epsilon y : Rxy)}\ \epsilon\mathrm{I}$$

To see that this is correct, note that the premise $Rx(\epsilon y : Rxy)$ is of the form

$$Rz(\epsilon y : Rxy)[x/z],$$

and the conclusion squares with this, for it is indeed of the form

$$Rz(\epsilon y : Rxy)[(\epsilon z : Rz(\epsilon y : Rxy))/z],$$

as it should be according to the rule.

Contrast this with the following application, which is incorrect:

$$\frac{Rx(\epsilon y : Rxy)}{R(\epsilon x : Rx(\epsilon y : Rxy))(\epsilon y : Rxy)}\ \epsilon\mathrm{I} \tag{2.3}$$

The problem is that the conclusion indicates that the premise was taken to be $Rx(\epsilon y : Rxy)[x/x]$, but the substitution that ϵI prescribes was not performed uniformly. The result should have been:

$$R(\epsilon x : Rx(\epsilon y : Rxy))(\epsilon y : R(\epsilon x : Rx(\epsilon y : Rxy))y).$$

It is also easy to see intuitively what went wrong in application (2.3): in the premise, both occurrences of x are free, but in the conclusion x occurs both bound (twice, by the epsilon operator) and free (once, in the second argument place of the main relation symbol).

It is instructive to go through the reasoning for the left to right direction of Proposition 2.33 again, but now in natural deduction style. Using $\vdash_{en}$ for the calculus which results from adding the ϵI rule to the set of introduction and elimination rules for the propositional connectives we can state the following:

2.37. PROPOSITION. *If* $\vdash_c \varphi(\exists, \forall)$ *then* $\vdash_{en} \varphi(\mathcal{E}, \mathcal{A})$.

Proof: We show that the quantifier introduction and elimination rules, in their $\mathcal{E}, \mathcal{A}$ guise, are admissible rules in $\vdash_{en}$.

For $\mathcal{E}I$, note that this is the epsilon term rule itself. This proves that the following rule is derivable:

$$\frac{\varphi[t/v]}{\mathcal{E}v\varphi}\,\mathcal{E}\mathrm{I}$$

For $\mathcal{A}E$, note that the following is a derivation in $\vdash_{en}$:

$$\dfrac{\dfrac{\dfrac{\neg\varphi[t/v][i]}{\neg\varphi[\epsilon v : \neg\varphi/v]}\,\epsilon\mathrm{I} \qquad \varphi[\epsilon v : \neg\varphi/v]}{\bot}\,\neg\mathrm{E}}{\varphi[t/v]}\,\neg\mathrm{E}_i$$

This proves that the following rule is derivable:

$$\frac{\mathcal{A}v\varphi}{\varphi[t/v]}\,\mathcal{A}\mathrm{E}$$

For $\mathcal{A}I$, assume that $\mathcal{D}$ is a proof tree with conclusion φ. Assume that v does not occur free in any assumption of $\mathcal{D}$ (note that this is Condition 2.2.1 from Section 2.2.2). Let $\mathcal{D}'$ be the result of substituting $\epsilon v : \neg\varphi$ for v where-ever this term occurs in $\mathcal{D}$. Then $\mathcal{D}'$ is a proof tree with the same set of undischarged assumptions as $\mathcal{D}$, and moreover $\mathcal{D}'$ derives $\varphi(\epsilon v : \neg\varphi)$. This gives: on Condition 2.2.1, from the premise φ the conclusion $\varphi[(\epsilon v : \neg\varphi)/v]$ is derivable. In other words, under these conditions, the following application of rule $\mathcal{A}I$ is admissible:

$$\frac{\varphi}{\mathcal{A}v\varphi.}\,\mathcal{A}\mathrm{I}$$

Note that we only state that a particular application of the rule is admissible; this is different from the statement that the rule itself is a derivable rule of the system. The latter is not the case for $\mathcal{A}I$ (although it is for $\mathcal{E}I$ and $\mathcal{A}E$, as we have just seen).

For $\mathcal{E}E$, assume $\mathcal{D}$ is a proof tree with conclusion ψ, using an assumption $\varphi[t/v]$. Let Σ be the set of assumptions in $\mathcal{D}$ on which ψ depends, other than $\varphi[t/v]$. Assume t does not occur in either ψ or Σ (note that this is Condition 2.2.2 from Section 2.2.2). Let $\mathcal{D}'$ be the result of substituting $\epsilon v : \varphi$ for t wherever t occurs in $\mathcal{D}$. Then $\mathcal{D}'$ is a proof tree for ψ with assumptions Σ and $\varphi[\epsilon v : \varphi/v]$. This shows that on Condition 2.2.2 we have:

$$\dfrac{\varphi[\epsilon v : \varphi/v] \qquad \begin{matrix}\varphi[t/v]\ [i]\\ \vdots\\ \psi\end{matrix}}{\psi}\,i$$

So we have established (under Condition 2.2.2) that the following application of $\mathcal{E}E$ is admissible:

$$\dfrac{\mathcal{E}v\varphi \qquad \begin{array}{c}\varphi[t/v]\ [i] \\ \vdots \\ \psi\end{array}}{\psi}\ \mathcal{E}\mathrm{E}_i$$

Note that, as in the case of $\mathcal{A}I$, we have proved something about a particular application of the rule. But this is enough. We have proved the proposition. ⊠

If we use the $\mathcal{A},\mathcal{E}$ abbreviation conventions, then proof trees for epsilon logic look exactly like proof trees for classical predicate logic:

$$\dfrac{\mathcal{E}x\mathcal{E}yRxy \qquad \dfrac{\mathcal{E}yRay\ [1] \qquad \dfrac{\dfrac{Rab\ [2]}{\mathcal{E}xRxb}\ \mathcal{E}\mathrm{I}}{\mathcal{E}y\mathcal{E}xRxy}\ \mathcal{E}\mathrm{I}}{\mathcal{E}y\mathcal{E}xRxy}\ \mathcal{E}\mathrm{E}_2}{\mathcal{E}y\mathcal{E}xRxy}\ \mathcal{E}\mathrm{E}_1$$

A subtle point that should be noted here is that the existence of the above proof tree implies that a direct derivation with the same premise and conclusion, but using only ϵI and the propositional rules, also exists. However, this connection tells us very little about how this classical look-alike relates to the most obvious 'unabbreviated' proof tree.

2.38. EXAMPLE. (Permutation of Existential Quantifiers) Here is an example of the concrete reasoning that is involved in deriving $\mathcal{E}y\mathcal{E}xRxy$ from $\mathcal{E}x\mathcal{E}yRxy$ in unabbreviated form. First note that the following derives the desired conclusion from Rab:

$$\dfrac{\dfrac{Rab}{R(\epsilon x : Rxb)b}\ \epsilon\mathrm{I}}{R(\epsilon x : Rx(\epsilon y : R(\epsilon x : Rxy)y))(\epsilon y : R(\epsilon x : Rxy)y)}\ \epsilon\mathrm{I} \tag{2.4}$$

Next, substitute $\epsilon y : Ray$ for b in this proof tree, and observe that the result (2.5) is still a correct proof tree.

$$\dfrac{\dfrac{Ra(\epsilon y : Ray)}{R(\epsilon x : Rx(\epsilon y : Ray))(\epsilon y : Ray)}\ \epsilon\mathrm{I}}{R(\epsilon x : Rx(\epsilon y : R(\epsilon x : Rxy)y))(\epsilon y : R(\epsilon x : Rxy)y)}\ \epsilon\mathrm{I} \tag{2.5}$$

To see that (2.5) is still correct, observe that $Ra(\epsilon y : Ray)$ is taken to be of the form $Rx(\epsilon y : Ray)[a/x]$ for the first application of ϵI, and $R(\epsilon x : Rx(\epsilon y :$

$Ray))(\epsilon y : Ray)$ is taken to be of the form $R(\epsilon x : Rxy)y[(\epsilon y : Ray)/y]$ for the second application of ϵI.

Finally, substitute $\epsilon x : Rx(\epsilon y : Rxy)$ for a in proof tree (2.5). This gives:

$$\cfrac{\cfrac{R(\epsilon x : Rx(\epsilon y : Rxy))(\epsilon y : R(\epsilon x : Rx(\epsilon y : Rxy))y)}{R(\epsilon x : Rx(\epsilon y : R(\epsilon x : Rx(\epsilon y : Rxy))y))(\epsilon y : R(\epsilon x : Rx(\epsilon y : Rxy))y)}\,\epsilon\mathrm{I}}{R(\epsilon x : Rx(\epsilon y : R(\epsilon x : Rxy)y))(\epsilon y : R(\epsilon x : Rxy)y)}\,\epsilon\mathrm{I} \tag{2.6}$$

To see that proof tree (2.6) is correct as well, note that the initial premise

$$R(\epsilon x : Rx(\epsilon y : Rxy))(\epsilon y : R(\epsilon x : Rx(\epsilon y : Rxy))y)$$

is taken to be of the form

$$Rx(\epsilon y : R(\epsilon x : Rx(\epsilon y : Rxy))y)[(\epsilon x : Rx(\epsilon y : Rxy))/x]$$

for the first application of ϵI, and the result of this application,

$$R(\epsilon x : Rx(\epsilon y : R(\epsilon x : Rx(\epsilon y : Rxy))y))(\epsilon y : R(\epsilon x : Rx(\epsilon y : Rxy))y),$$

is taken to be of the form

$$R(\epsilon x : Rxy)y[(\epsilon y : R(\epsilon x : Rx(\epsilon y : Rxy))y)/y]$$

for the second application of ϵI. Thus, in this particular case, the unabbreviated proof tree which derives $\mathcal{E}y\mathcal{E}xRxy$ from $\mathcal{E}x\mathcal{E}yRxy$ consists of just two applications of ϵI, and that's all. Of course, we have to pay for this by a considerable rewriting effort (the mountain of rice gruel we have to eat ourselves through to arrive in the epsilon land of Cockayne, so to speak). In many cases we shall prefer to work with the abbreviated forms to avoid unwanted epsilon term explosions.

The semantic reasoning to prove the converse of Proposition 2.37 goes through as before. This means that in the natural deduction calculus for epsilon logic $\exists$ coincides with $\mathcal{E}$, and $\forall$ coincides with $\mathcal{A}$ (just like in the Hilbert axiomatization).

2.4 Extensional Epsilon Logic

2.4.1 The Extensionality Principle for Epsilon Terms

Leisenring [Lei69] has something interesting to add to Hilbert's version of the epsilon calculus, namely an analysis of the addition of the following extensionality principle for epsilon terms:

$$\forall z(\varphi[z/x] \leftrightarrow \psi[z/y]) \rightarrow \epsilon x : \varphi = \epsilon y : \psi.$$

This principle expresses informally that two formulas defining the same subset of the domain, give rise to identical epsilon terms. This principle is mentioned by Hilbert himself, but only in passing. Leisenring's treatment of it is based largely on Asser [Ass57].

2.4.2 Semantics

It is clear that throwing in the extensionality principle for epsilon terms imposes an extra restriction on the way we build Skolem expansions. The recipe for building the Skolem expansion M' of an arbitrary first order model M that was sketched above will not guarantee without further ado that the principle holds. The extra condition we need turns out to be easy to formulate, however. What we need is that if $\lambda y.\varphi(\overline{x}, y)$ equals $\lambda y.\psi(\overline{x}, y)$, then the Skolem functions for the two formulas should assign the same element of the domain.

2.39. DEFINITION. (Extensional Choice Functions) An *extensional choice function* for first-order model M is a mapping $\Phi : \mathcal{P}(\text{dom}\,(M)) \to \text{dom}\,(M)$, satisfying the conditions:

1. If $N \subseteq \text{dom}\,(M), N \neq \emptyset$, then $\Phi(N) \in N$.
2. If $N \subseteq \text{dom}\,(M), N = \emptyset$, then $\Phi(N) \in dom(M)$.

Note the difference with intensional choice functions, which are defined on the epsilon terms themselves. We let eM denote the set of all extensional choice functions for M.

2.40. DEFINITION. (Valuation of Terms) Let M be a first-order model, s a variable assignment for M, and Φ an extensional choice function for M. Then the *term valuation* $V_{M,\Phi,s}$ *in* M *based on* Φ *and* s is given by the following clauses:

- $V_{M,\Phi,s}(c) = \text{int}\,(M)(c)$.
- $V_{M,\Phi,s}(v) = s(v)$.
- $V_{M,\Phi,s}(ft_1 \cdots t_n) = \text{int}\,(M)(f)(V_{M,s}(t_1), \ldots, V_{M,s}(t_n))$.
- $V_{M,\Phi,s}(\epsilon y : \varphi(\overline{x}, y)) = \Phi(\{m \in dom(M) \mid M, \Phi, s(x|m) \models \varphi(\overline{x}, y)\})$

The Skolem function $F_{\varphi(\overline{x})}$ corresponding to the ϵ-term $\epsilon y : \varphi(\overline{x}, y)$ can then be defined to be the function $V_{M,\Phi}(F_{\varphi(\overline{x})})$ satisfying for all sequences of elements of the domain $\overline{m}$ of the arity of $\overline{x}$

$$V_{M,\Phi}(F_{\varphi(\overline{x})})(\overline{m}) = V_{M,\Phi,s(\overline{x}|\overline{m})}(\epsilon y : \varphi(\overline{x}, y)).$$

Take M' to be the model with interpretations for all Skolem functions in L', assigned in accordance with this strategy. Then it is clear, again, that the interpretations of formulas in the original language L are not affected. Because the extension of the proof system for first order logic with the epsilon rule and the extensionality principle for epsilon terms is obviously sound for models with choice functions, we get that the epsilon calculus with the extensionality principle for epsilon terms is a *conservative* extension of classical predicate logic, too.

Note that in fact we have defined the notion $M, \Phi, s \models \varphi$, for a first order model M, an extensional choice function Φ and an arbitrary φ from a language with epsilon terms.

For extensional epsilon logic, we can strengthen Proposition (2.18). Let $\rightsquigarrow$ be a binary relation such that $\varphi \rightsquigarrow \varphi'$ holds if φ' is the result of replacing an occurrence of $\exists v\psi$ in φ, possibly within the scope of an ϵ operator, by an occurrence of $\psi(\epsilon v : \psi)$, or an occurrence of $\forall v\psi$ in φ, possibly within the scope of an ϵ operator, by an occurrence of $\psi(\epsilon v : \neg\psi)$. Then:

2.41. PROPOSITION. *If $\varphi \rightsquigarrow \varphi'$ then for all M, all $\Phi \in eM$, all $s \in aM$:*

$$M, \Phi, s \models \varphi \textit{ iff } M, \Phi, s \models \varphi'.$$

Because of the fact that the choice functions are extensional, substitution within an epsilon term will do no harm. From the previous proposition we get:

2.42. PROPOSITION. *If φ, ψ is a pair of formulas such that φ has no epsilon terms and $\varphi \overset{*}{\rightsquigarrow} \psi$, then for all M, all $\Phi \in eM$, all $s \in aM$:*

$$M, s \models \varphi \textit{ iff } M, \Phi, s \models \psi.$$

For defining the truth of a formula in the extensional epsilon calculus which may contain epsilon terms we again may define a *generic* interpretation.

2.43. DEFINITION. (Generic e-interpretation)

$$M, s \models^e_g \varphi \textit{ iff } M, \Phi, s \models \varphi \textit{ for all } \Phi \in eM.$$

Again, it depends on our choice what a formula like $E(\epsilon x : x = x)$ is going to mean. Suppose N is the model of the natural numbers, and E denotes the property of being an even number. Then there certainly is a choice function Φ with $M, \Phi \models Ex(\epsilon x : x = x)$, so under the non-generic interpretation the formula is true on N, given that choice function. Under the generic interpretation, the formula is false, however. This squares with the intuition that if one 'arbitrarily' picks a number from the domain N, then there is of course no guarantee that this number will be even.

As above, the distinction between local interpretation and generic interpretation engenders two notions of extensional consequence: *local e-consequence* which will be denoted $\models^e_l$ and *generic e-consequence*, for which we shall use the relation symbol $\models^e_g$.

2.4.3 Natural Deduction for Extensional Epsilon Logic

The extensionality principle is expressed in natural deduction format in the following rule:

$$\dfrac{\begin{matrix}\varphi(v_1,\ldots,v_n,v)\ [i] \\ \vdots \\ \psi(v_n,\ldots,v_n,v)\end{matrix} \qquad \begin{matrix}\psi(v_1,\ldots,v_n,v)\ [j] \\ \vdots \\ \varphi(v_1,\ldots,v_n,v)\end{matrix}}{\epsilon v\varphi = \epsilon v\psi}\ \epsilon\mathrm{I}_{i,j}$$

Adding this rule to the calculus for $\vdash_h$ yields an axiomatization of local e-consequence which is sound and complete. Soundness is straightforward. For completeness, again we extend a consistent set of closed formulas in standard way to a maximally consistent set Γ. This set is witnessing because the epsilon axiom is present. We construct the domain of the desired model from equivalence classes of closed terms. As in our discussion below Proposition 2.30 we consider subsets N of the domain which are represented by formulas $\varphi(x)$. We have argued that the equivalence class of $\epsilon x : \varphi$ is in N, if $\varphi(x)$ represents N. For the extensional epsilon calculus we have to show moreover, that if $\varphi(x)$ and $\psi(x)$ both represent N, then $\epsilon x : \varphi = \epsilon x : \psi \in \Gamma$. Let $t = \epsilon x : \neg(\varphi \leftrightarrow \psi)$. Because $\varphi(x)$ and $\psi(x)$ both represent N, we have $\varphi[t/x] \in \Gamma$ iff $\psi[t/x] \in \Gamma$. Consequently $\varphi[t/x] \leftrightarrow \psi[t/x] \in \Gamma$. But then, by the derivability of $\xi(\epsilon x : \neg\xi) \rightarrow \forall x \xi$, we have $\forall x(\varphi \leftrightarrow \psi) \in \Gamma$. By extensionality $\epsilon x : \varphi = \epsilon x : \psi \in \Gamma$.

2.5 Arbitrary Object Theory

2.5.1 Background

In a series of papers resulting in the book [Fin85], Kit Fine has set out to rehabilitate arbitrary objects by formulating a coherent account of the principle of generic attribution, and by constructing formal models for interpreting languages with constants denoting arbitrary objects. Fine argues convincingly that there are various areas of research where the introduction of arbitrary objects is well motivated. This holds in particular for the analysis of informal mathematical reasoning and for the semantics of natural language.

The heart of Fine's theory of arbitrary objects consists of a reformulation of the principle of generic attribution. According to Fine, the argument showing that the notion of an arbitrary object leads to contradictions for complex properties, depends upon the failure to distinguish two basically different formulations of this principle: one is merely a rule of *equivalence* and is stated in the material mode; the other is a rule of *truth* and is stated in the formal mode.

Fine claims that there are two versions of the principle of generic attribution, of which only one leads to unsurmountable difficulties. To formulate the two versions, let $\overline{a}$ be the name of an arbitrary object a and let x be a variable that ranges over the individuals in the range of a.

The 'equivalence formulation' of the principle of generic attribution now takes the following form:

$$\varphi(a) \leftrightarrow \forall x \varphi(x).$$

This is the formulation which leads to contradiction. For let $A(x)$ be the statement that triangle x is acute-angled, and $O(x)$ the statement that triangle x is obtuse. Because $\forall x(A(x) \vee O(x))$ it follows that $A(a) \vee O(a)$ for the arbitrary triangle a. But because $\neg\forall x A(x)$ and $\neg\forall x O(x)$ it follows that $\neg A(a) \wedge \neg O(a)$, and we have arrived at a contradiction.

Another formulation of the principle, which Fine calls the 'truth formulation', has a completely different form:

$$\textit{Sentence } \varphi(a) \textit{ is true iff sentence } \forall x\varphi(x) \textit{ is true.} \tag{2.7}$$

In this formulation the argument leading to contradiction is blocked, and we get a coherent principle. It should be noted, however, that there is a price to pay. In general, formulas containing names for arbitrary objects do not decompose truth–functionally.

2.5.2 Semantics

To formalize his theory of arbitrary objects, Fine defines the notion of an *Arbitrary Object Model* (AO model). An AO model is a standard first-order model to which a second domain A of arbitrary objects is added. The objects in this domain are related to standard individuals by a set V of *value assignment functions.* On the set of arbitrary objects a relation of dependence $\prec$ is defined. Fine assumes that he has names for all of his arbitrary objects. Assuming $\dot{a}$ ranges over a set of names $\dot{A}$, where $\dot{a}$ names $a \in A$, we can take it that languages for arbitrary objects have the following set of terms:

terms $t ::= c \mid \dot{a} \mid v \mid f(t_1 \cdots t_n)$.

Formulas are built from these terms as in the first order case.

Truth of a sentence with constants denoting arbitrary objects now consists of truth of this sentence on the underlying first–order model with respect to all value assignment functions. In this way the truth formulation of the principle of generic attribution is given a rigorous semantics.

In all interesting applications of this theory, the dependency relation $\prec$ plays an essential part. The notation $a \prec b$ is to be read as "object a depends on object b". The dependency relation represents the *order*, in some sense, in which the generic superstructure is created over the first–order model. This entails certain natural requirements on the relation: $\prec$ is transitive and conversely well-founded (i.e., it does not admit infinite chains $a \prec a_1 \prec a_2 \prec \ldots$). Thus, the arbitrary object part of an AO model (its 'generic extension') can be seen as the result of a stage-by-stage process of construction arbitrary objects over the model.

2.44. DEFINITION. (Arbitrary Object Models) An *arbitrary object model* (AO model) M^* is a quadruple $\langle M, A, \prec, V\rangle$ such that

1. M is a first order model.
2. A is a non-empty set disjoint from dom (M).
3. $\prec$ is a relation on A that is transitive and conversely well-founded.
4. V is a non-empty set of partial functions from A into dom (M) such that

(a) V is closed under restriction to subsets of its domain.
(b) V is partially extendible in the sense that if $v \in V$, then there is $v^+ \in V$ with $v^+ \supseteq v$ and dom $(v^+) \supseteq [\text{dom}\,(v)]$.
(c) V is closed under 'piecing'. That is, if $\{v_i \mid i \in I\}$ is a non-empty closed subset of V such that for each v_i, dom $(v_i) = [\text{dom}\,(v_i)]$ (each v_i has a domain closed under dependency), and the union $\bigcup v_i$ is a function, then $\bigcup v_i \in V$.

Here $[\text{dom}\,(v)] = \text{dom}\,(v) \cup \{b \in A \mid \exists a \in \text{dom}\,(v) : a \prec b\}$, i..e., the closure of dom (v) under dependency.

We shall use M^* to denote an arbitrary object model over a classical model M. The addition $\langle A, \prec, V\rangle$ to M is called the *generic superstructure* of M^*.

Let $A(\varphi)$ be the set of arbitrary object names occurring in φ. If $B \subseteq A$, then we use V_B for the set of members of V of which the domain includes B. Some further notational conventions will be useful later on. For every $a \in A$, let $|a|$ be the set of all elements of A on which a depends, i.e., $|a| = \{b \in A \mid a \prec b\}$, and let $[a]$ be the set consisting of a together with all elements of A on which a depends, i.e., $[a] = |a| \cup \{a\}$. The *value range* of an arbitrary object a, VR_a, is the set

$$\{v(a) \in \text{dom}\,(M) \mid v \in V\}.$$

If $a \in A, B \subseteq A$, then the *value dependence* of a upon B, $\mathrm{VD}(a, B)$, is the function $f : V_B \mapsto \mathcal{P}(dom(M))$ defined by:

$$f(v) = \{d \in \text{dom}\,(M) \mid v \cup \{\langle a, d\rangle\} \in V\}.$$

The value dependence VD_a is $\mathrm{VD}(a, |a|)$.

There are two notions of truth for formulas with arbitrary object names in them:

2.45. DEFINITION. (Local truth in AO models) Let $\overline{a}$ be the sequence all names of arbitrary objects occurring in φ and $\overline{x}$ be a sequence of variables of the length of $\overline{a}$ not occurring in φ. Then

$$\langle M, A, \prec, V\rangle, s, v \models \varphi \text{ iff } M, s(\overline{x}|v(\overline{a})) \models \varphi[\overline{x}/\overline{a}].$$

2.46. DEFINITION. (Global truth in AO models)

$$\langle M, A, \prec, V\rangle, s \models \varphi \text{ iff for all } v \in V_{A(\varphi)} : \langle M, A, \prec, V\rangle, s, v \models \varphi.$$

Figure 2.3 illustrates that, according to the definition of global truth, arbitrary objects may lack both a property and its negation.

Valid case-to-case consequence is defined in terms of (local) truth as follows:

2.47. DEFINITION. (Case-to-case AO consequence) $\Gamma \models_c \varphi$ iff for all M^*, s, v: If $v \in V$ and v is defined on all AO parameters in Γ, φ, then $M^*, s, v \models \Gamma$ implies $M^*, s, v \models \varphi$.

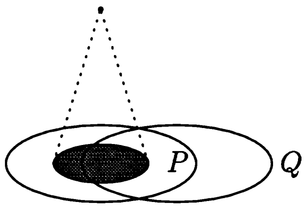

Figure 2.3: An arbitrary object with property P, and with neither Q nor $\neg Q$.

Valid truth-to-truth consequence is defined in terms of (global) truth, as follows:

2.48. DEFINITION. (Truth-to-truth AO consequence) $\Gamma \models_t \varphi$ iff for all M^*: if $M^* \models \Gamma$, then $M^* \models \varphi$.

2.5.3 Natural Deduction with Arbitrary Objects

There is a straightforward connection between natural deduction and arbitrary object theory. The paradigmatic use of arbitrary objects in natural deduction is exemplified in the following proof step in a mathematical argument:

$$\textit{There exists a bisector to the angle } \alpha\textit{. Call } \text{it } a. \tag{2.8}$$

This proof step can be formalized in a natural deduction set–up in either of two ways. The most common one uses the *elimination rule* for the existential quantifier ([Pra65]). Here the presence of an existential formula $\exists v\varphi$ allows us to *assume* $\varphi(t)$ for some fresh term t, the *proper term* of the application. This assumption may then be discharged upon reaching a conclusion that does not contain the proper term. If the assumption $\varphi(t)$ is itself an existential form $\exists y\psi(t, y)$, then a second proper term t' is introduced in the assumption $\psi(t, t')$, with the marginal remark that t' depends on t. This dependency has to be respected as the proof unfolds. The assumption $\varphi(t)$ may not be discharged before we have discharged $\psi(t, t')$. This dependency can be modeled by the interpreting the proper terms of ($\exists$E) applications on arbitrary objects in a generic superstructure of a first–order model for the assumptions of $\mathcal{D}$.

There is also a less familiar formalization of the use of existential information. This takes the form of a rule for *existential instantiation*: from $\exists v\varphi$ *conclude* $\varphi(a)$ for some appropriate term a. In a proof system with this rule restrictions have to be put on the use of the proper term to prevent some obviously incorrect inferences (most patently, the inference $\exists v\varphi/\varphi(a)/\forall v\varphi$ has to be blocked). If one spells out the restrictions involved it becomes clear that the proper terms used in such derivations must map to arbitrary objects related by dependency on a generic superstructure of a first–order model.

We shall work out the second proposal in some detail (this is in the Copi-Kalish system of Natural Deduction, in AO guise — see Fine [Fin85]). The system is like our standard system for classical first order logic, but with subtly different quantifier rules.

$$\frac{\varphi}{\forall v\varphi[v/t]}\,\forall\mathrm{I} \qquad\qquad \frac{\forall v\varphi}{\varphi[a/v]}\,\forall\mathrm{E}$$

The rule (∀I) is our old rule (but see the condition stated below). In the rule $\forall E$, the proper term of the application, a, is an arbitrary object term. We define $AO(\varphi)$ to be the set of proper terms upon which the derivation of φ depends. This set is defined recursively with $AO(\varphi) = \emptyset$ for φ occurring at a leaf node in a proof tree, and $AO(C) = AO(P)$ if C is the result of an application other than $\exists E$ or $\forall E$ with P as premise, and $AO(C) := AO(P) \cup \{a\}$ for C the result of an application of $\exists E$ or $\forall E$ with P as premise and a as proper term.

$$\frac{\varphi[t/v]}{\exists v\varphi}\,\exists\mathrm{I} \qquad\qquad \frac{\exists v\varphi}{\varphi[a/v]}\,\exists\mathrm{E}\ a \prec AO(\varphi)$$

In $\exists E$, the proper term a is again an arbitrary object term. The proper terms resulting from applications of $\exists E$ are called e-proper terms, the proper terms resulting from applications of $\forall E$ are called a-proper terms.

The dependency relation $\prec$ between proper terms which occur in a derivation is defined in accordance with the annotation of the $\exists E$ rule: e-proper terms introduced at some level depend on all AO terms of the premise.

The conditions on the rules are the following:

2.5.1. CONDITION. *The variable bound in the conclusion of* $\forall I$ *should not occur free in the premise.*

2.5.2. CONDITION. *The proper term of an application of* $\forall I$ *should not occur as an e-proper term in the proof tree of the premise.*

2.5.3. CONDITION. *The proper term* t *of an application of* $\forall I$ *should not occur in the conclusion or in any assumption on which the premise* φ *depends, nor should any term* b *with* $b \prec t$.

Note that Condition 2.5.3 is a strengthening of Condition 2.2.1 on the rule $\forall I$ in the classical system of Section 2.2.2.

2.5.4. CONDITION. *The proper term of an application of* $\exists E$ *is fresh (i.e., it should not occur in the proof tree of the premise).*

2.5.5. CONDITION. *The variable bound in the conclusion of* $\exists I$ *should not occur free in the premise.*

Here are some example deductions in this calculus.

$$\dfrac{\dfrac{\dfrac{\dfrac{\exists x\forall y Rxy}{\forall y Ray}\ \exists\mathrm{E}\ a \prec \emptyset}{Rab}\ \forall\mathrm{E}}{\exists x Rxb}\ \exists\mathrm{I}}{\forall y\exists x Rxy}\ \forall\mathrm{I}
\qquad
\dfrac{\dfrac{\dfrac{\dfrac{\exists x\exists y Rxy}{\exists y Ray}\ \exists\mathrm{E}\ a \prec \emptyset}{Rab}\ \exists\mathrm{E}\ b \prec a}{\exists x Rxb}\ \exists\mathrm{I}}{\exists y\exists x Rxy}\ \exists\mathrm{I}$$

Here is an example of how the strengthened condition on $\forall I$ blocks the derivation of $\exists y\forall x Rxy$ from $\forall x\exists y Rxy$:

$$\dfrac{\dfrac{\dfrac{\forall x\exists y Rxy}{\exists y Rax}\ \forall\mathrm{E}}{Rab}\ \exists\mathrm{E}\ b \prec a}{\text{application of } \forall I \text{ to } a \text{ blocked because } b \prec a}$$

2.6 Comparison of Epsilon Logic and AO Theory

It should be clear from the discussion so far that arbitrary objects are creatures in limbo, living in the shadowy realm between syntax and semantics. Fine tries to pull them over into the area of semantics altogether, but the dependency relation $\prec$ continues to give off a distinctly syntactic smell. This becomes even clearer when we look at concrete examples of the way in which arbitrary objects are created as a result of a definition. When we say 'let a be an arbitrary even number', and a bit later, in the course of the argument, 'let b be the successor of a' then we have in fact focussed on two arbitrary objects a, b with $b \prec a$. Suppose we are talking about the domain of integers. Then we are in fact talking about all pairs of the form $2n, 2n+1$. But we could also have proceeded the other way around: 'let b be an arbitrary odd number', followed by 'let a be the predecessor of b'. Then we are talking about the same relation between numbers of the form $2n, 2n+1$, but the dependency relation now is different, for we have $a \prec b$. So the two ways of introducing the pair a, b give rise to different generic superstructures of the model of the integers.

2.6.1 Semantic Comparison

The reader may have a feeling that arbitrary object models and epsilon calculus should be related, and indeed they are. If we take closed epsilon terms as arbitrary objects, and if we assume that they name themselves, then a formula φ which has only closed epsilon terms becomes an AO formula.

2.49. PROPOSITION. *Every first order model M for a language L can be turned into a arbitrary object model M^* where the arbitrary objects are the closed epsilon terms of L^+ (the expansion of L with epsilon terms). For every formula φ which contains only closed epsilon terms we have:*

$$M \models_g \varphi \textit{ iff } M^* \models \varphi.$$

Proof: Let a first order model M for L be given, and let L^+ be the expansion of that language with epsilon terms. Let A be defined as follows:

$$A = \{(\epsilon x : \varphi(x)) \mid \varphi(x) \in L^+, \varphi(x) \text{ has only } x \text{ free}\}.$$

In other words, A consists of all closed epsilon terms of L^+. Let $\prec$ be the dependency relation on epsilon terms given by $a \prec b$ iff $b \in A$ and b occurs in a. Thus, e.g., $\epsilon x : R(x, \epsilon y : Ryy) \prec \epsilon y : Ryy$, but there is no $a \in A$ with $\epsilon x : R(x, \epsilon y : Rxy) \prec a$, the reason being that $\epsilon y : Rxy$ is not a closed epsilon term.

It is clear that every intensional choice function Φ for M determines a valuation function $V_{M,\Phi}$ for the members of A. This is because closed epsilon terms can be evaluated in the model M once we have a choice function available, independently of a variable assignment. Let $\mathcal{F}$ be the set of intensional choice functions for M, and let V be defined as follows:

$$V = \{V_{M,\Phi} \upharpoonright B \mid B \subseteq A, B \neq \emptyset, \Phi \in \mathcal{F}\}.$$

In other words, the members of V are all non-empty restrictions of intensional choice functions for M. We claim that $\langle M, A, \prec, V\rangle$ is an AO-model.

To check this claim, all that is needed is a perusal of Fine's eight conditions on quadruples $\langle M, A, \prec, V\rangle$. We shall only discuss the closure of V under piecing. If $\{v_i \mid i \in I\}$ is a non-empty closed subset of V such that for each v_i, dom $(v_i) =$ [dom (v_i)] (each v_i has a domain closed under dependency), and the union $\bigcup v_i$ is a function, then $\bigcup v_i \in V$. This holds, for if $\bigcup v_i$ is a function satisfying the requirements, then it can be extended to a function Φ which assigns values to *all* closed epsilon terms of the language. Then $\Phi \in \mathcal{F}$, and so $\bigcup v_i \in V$ by definition.

To complete the proof, let φ be a formula with only closed epsilon terms. We then have: $M, s \models_g \varphi$ iff for all $\Phi \in iM$: $M, \Phi, s \models \varphi$. This is the case iff for all $v \in V$ defined over the closed epsilon terms occurring in φ, $M, v \cup s \models \varphi$, which is the case iff $M^*, s \models \varphi$. ⊠

We may also ask the converse question.

2.50. QUESTION. Given an arbitrary AO model $\langle M, A, \prec, V\rangle$, is there always a set of choice functions $\mathcal{F}$ (possibly intensional) over the same model M and a set of epsilon term replacements for the arbitrary objects such that for formulas with closed epsilon terms, generic truth in the sense of epsilon logic coincides with truth in the AO sense?

Of course, this question presupposes restriction to the language which has only those epsilon terms thrown in which correspond to some element of A. Otherwise, the answer to the question would be a trivial 'no', for there are AO models M^* with $M^* \models \exists x\varphi$ while for no $a \in A$, $M^* \models \varphi(\dot{a})$. On the other hand, $M, \Phi \models \exists x\varphi$ implies $M, \Phi \models \varphi(\epsilon x : \varphi)$. So the question only makes sense if we are prepared to leave such $\epsilon x : \varphi$ out of the language.

Given this further restriction, it is not immediately clear anymore what the answer to the question should be. There may be an independent $a \in A$ for which $\{v(a) \mid v \in V\}$ is not definable in M without parameters, or a dependent $a \in A$ for which $\{v(a) \mid v \in V\}$ is not definable in M with parameters. If this is the case, there will be no $\varphi(x)$ with the property:

$$\{b \mid M, s(x|b) \models \varphi(x)\} = \mathrm{VR}_a.$$

Prima facie, it seems that there is no suitable formula φ for building an epsilon term $\epsilon x : \varphi$ to represent a. But maybe a clever choice of $\mathcal{F}$ can remedy this. After all, by suitably restricting $\mathcal{F}$, we can make sure that the set

$$X = \{\Phi(\mathrm{dom}\,(M)) \mid \Phi \in \mathcal{F}\}$$

is a non-definable set in M, and still, $\epsilon x : x = x$ now gets represented by X. We leave this question open for now.

But we may wish to put the case of undefinable AO objects aside, taking our cue from the following quote from Chapter 3 of Fine [Fin85]:

> In many of the applications, all of the A-objects that are required may, in a certain sense, be defined within a previously specified language.

2.51. DEFINITION. (Representable arbitrary objects) If M is a model for L, and M^* is an AO-model based on M, then we say that an arbitrary object a, not depending on any object, in M^* is *representable* if there is a $\varphi(x) \in L$ such that:

$$\mathrm{VR}_a = \{d \mid M, \{\langle x, d\rangle\} \models \varphi(x)\}.$$

A dependent arbitrary object a is *representable* if $[a] = \{a_1, \ldots, a_n\}$, and there is a $\varphi(x_1, \ldots, x_n, x) \in L$ such that $\mathrm{VD}_a = f$ is given by:

$$f(v) = \{d \mid M, v \cup \{\langle x, d\rangle\} \models \varphi(\dot{a}_1/x_1, \ldots, \dot{a}_n/x_n, x)\},$$

where $v \in V_{[a]}$.

For arbitrary object models where every arbitrary object is representable we can get a straightforward connection with epsilon logic by translating the arbitrary object names into their obvious epsilon term translations. Using ϵa as an abbreviation for the epsilon term translation of $\dot{a}$ we get the following:

2.52. PROPOSITION. *If every arbitrary object of M^* is representable, then:*

$$M^* \models \varphi(\dot{a}_1, \ldots, \dot{a}_n) \textit{ iff } M \models_g \varphi(\epsilon a_1/\dot{a}_1, \ldots, \epsilon a_n/\dot{a}_n),$$

for every φ in the language of M.

Propositions 2.49 and 2.52 suggest that epsilon logic can be considered as the logic of representable arbitrary objects, where the closed epsilon terms are the arbitrary objects which wear their definitions on their sleeves.

Among various possible conditions on AO models Fine discusses the following requirement of 'identity':

2.53. DEFINITION. (Identity) M^* satisfies *identity* iff the following hold:

- For any two independent elements a and b of A: $\mathrm{VR}_a = \mathrm{VR}_b$ implies $a = b$.
- For any two dependent elements $a, b \in A$, $|a| = |b|$ and $\mathrm{VD}_a = \mathrm{VD}_b$ together imply $a = b$.

The following propositions gives the obvious connection with extensionality in epsilon logic:

2.54. PROPOSITION. *Every first order model M for a language L can be turned into an arbitrary object model M^* which satisfies 'identity' and where the arbitrary objects are constructed from the extensional equivalence classes of closed epsilon terms of L^+ (the expansion of L with epsilon terms).*

If we assume that the object $[\epsilon x : \varphi]$ is named by $\epsilon x : \varphi$, we have, for every formula φ which contains only closed epsilon terms:

$$M \models_g^e \varphi \textit{ iff } M^* \models \varphi.$$

Proof: Construct M^* as in Proposition 2.49, only this time take equivalence classes of closed epsilon terms as arbitrary objects, as follows.

If E is the set of closed epsilon terms of L, define $t_1 \prec t_2$ iff t_2 occurs as a proper subterm in t_1. Let $|t|$ be the set $\{t' \in E \mid t \prec t'\}$. Next, for $(\epsilon v : \varphi), (\epsilon v : \psi) \in E$, set $(\epsilon v : \varphi) \sim (\epsilon v : \psi)$ iff (i) $\models \forall v(\varphi \leftrightarrow \psi)$, and (ii) $|\epsilon v : \varphi| = |\epsilon v : \psi|$. Let $[\epsilon v : \varphi]$ be the $\sim$ equivalence class of φ. Note that because of (ii) in the definition of $\sim$, we can set $[t] \prec [\epsilon v : \varphi]$ iff $t \prec (\epsilon v : \varphi)$, so $|[\epsilon v : \varphi]|$ equals $\{[t] \mid t \prec (\epsilon v : \varphi)\}$. It is obvious from this that $\prec$ on $E_\sim$ inherits the properties of irreflexivity, transitivity and converse well-foundedness from $\prec$ on E.

The remainder of the proof is just a check that the extensional equivalence classes of the closed epsilon terms satisfy the requirements of the 'identity' definition. ⊠

2.55. PROPOSITION. *If every A object of M^* is representable and M^* satisfies 'identity' then:*

$$M^* \models \varphi(\dot{a}_1, \ldots, \dot{a}_n) \textit{ iff } M \models_g^e \varphi(\epsilon a_1/\dot{a}_1, \ldots, \epsilon a_n/\dot{a}_n),$$

for every φ of the language of M.

Propositions 2.54 and 2.55 suggest that extensional epsilon logic can be considered as the logic of representable arbitrary objects in AO models satisfying 'identity', where the extensional equivalence classes of closed epsilon terms are the arbitrary objects which wear their definitions on their sleeves.

With our link between epsilon calculus and arbitrary object theory established, let us briefly reconsider the principle of generic attribution, applied to an unrestricted arbitrary object a which does not depend on any other arbitrary object. Given these assumptions, an appropriate epsilon term for a is $\epsilon x : x = x$. Let us see what the principle of generic attribution now says, in both of its formulations. Here is the equivalence formulation:

$$\varphi(\epsilon x : x = x) \leftrightarrow \forall x \varphi(x).$$

Now is this true in every model M, for all choice functions Φ on M? Obviously not, for given a particular choice for $\epsilon x : x = x$, the object that is picked out will definitely have lots of properties which are not shared with all objects in the domain. Indeed, the equivalence formulation of the principle of generic attribution is as hopelessly wrong in its epsilon guise as in the guise discussed above. But now look at the 'truth formulation' of the principle:

$$\textit{Sentence } \varphi(\epsilon x : x = x) \textit{ is true iff sentence } \forall x \varphi(x) \textit{ is true.} \tag{2.9}$$

Is this true in every model M, for all choice functions Φ on M? Yes, precisely because of the definition of truth for sentences of the form $\varphi(\epsilon x : x = x)$ in terms of universal quantification over choice functions. To be true for an epsilon formula means to be true under all possible choice functions Φ. So, obviously, the truth formulation of the principle of generic attribution, in its epsilon guise, is correct.

2.6.2 Proof Theoretic Comparison

For the proof theoretic comparison of AO theory and epsilon logic, we proceed by example. Compare the following two derivations. First an AO proof tree:

$$\cfrac{\cfrac{\cfrac{\cfrac{\exists x \forall y Rxy}{\forall y Ray}\ \exists\mathrm{E}\ a \prec \emptyset}{Rab}\ \forall\mathrm{E}}{\exists x Rxb}\ \exists\mathrm{I}}{\forall y \exists x Rxy}\ \forall\mathrm{I}$$

Here is a proof tree in epsilon logic, using $\exists E\epsilon$ in addition to the rules of classical natural deduction.

$$\dfrac{\dfrac{\dfrac{\dfrac{\exists x \forall y Rxy}{\forall y R(\epsilon x : \forall y Rxy)y}\ \exists \mathrm{E}_\epsilon}{R(\epsilon x : \forall y Rx(\epsilon y : \neg R(\epsilon x : \forall y Rxy)y))(\epsilon y : \neg R(\epsilon x : \forall y Rxy)y)}\ \forall \mathrm{E}}{\exists x Rx(\epsilon y : \neg Rxy)}\ \exists \mathrm{I}}{\forall y \exists x Rxy}\ \forall \mathrm{I}$$

Note that the only difference between the trees is that in the second tree the AO terms have been replaced by suitable epsilon terms.

Note also that the condition on $\forall I$ from the AO version of the rules has become visible in the proper (epsilon) term of the application. In the example: $(\epsilon y : \neg Rxy)$, the proper term of $\forall I$ in the proof tree, does not depend on an e-term of the tree, for the only e-term of the proof tree is $(\epsilon x : \forall y Rxy)$, and this term does not occur in $(\epsilon y : \neg Rxy)$.

2.7 Conclusion

In this chapter, we have demonstrated that reasoning about arbitrary objects can be based on sound principles. Also, we have drawn some important distinctions: between local consequence and generic consequence (or in Fine's terminology, between case-to-case and truth-to-truth reasoning with arbitrary objects), and between intensional and extensional epsilon logic (or in Fine's way of putting it, between working with models which don't satisfy 'identity' and working with models which do). Finally, we have seen that if we confine attention to arbitrary objects a which are representable by formulas $\varphi(\dot{a}_1, \ldots, \dot{a}_n, v)$ of a first order language L, where $\dot{a}_1, \ldots, \dot{a}_n$ refer to arbitrary objects which are themselves representable (with the obvious requirement of non-circularity in the representations), then epsilon logic is equivalent to Fine's theory of arbitrary objects.

Chapter 3

Intuitionistic Instantial Logic

This chapter explores the topic of instantial intuitionistic reasoning by investigating what happens when we add the epsilon rule to a natural deduction calculus for intuitionistic predicate logic. As it turns out, this extension is not conservative, but leads to an interesting intermediate logic. This logic, as well as some weaker and related logics, will be investigated proof theoretically as well as semantically. All but one of the logics we describe will be frame incomplete.

3.1 Introduction

Hilbert invented epsilon terms as a weapon against Brouwer's intuitionism; but this should not deter us from considering the result of adding the epsilon rule to intuitionistic logic.

Our interest in the strength of the epsilon rule *per se* pushes us in this direction, for we can view the conservativity of the epsilon rule over classical predicate logic as an indication that the classical quantifiers have such strong structural properties that the contribution of the epsilon rule gets swamped out by these effects. To find out more about the epsilon rule we therefore have to add it to a system with weaker quantification principles. Intuitionistic predicate logic is an obvious choice.

3.2 Intuitionistic Predicate Logic

3.2.1 Semantics

Assume the same language definition as in Section 2.2 of the previous chapter. Intuitionistic logic models the process of growth of knowledge of an ideal mathematical reasoner, who on the one hand does not know every mathematical truth that there is to know (s/he may not know whether the decimal expansion of π contains a row of nine consecutive nines, to mention an example from Brouwer), but on the other hand never makes mistakes. In other words, mathematical knowledge can only grow, as more and more mathematical conjectures get proved or disproved. A statement φ which is found out to be true at some stage in the process of knowledge acquisition will never become false later on, and a statement φ which is found out to be false at some stage will never become true later on. Using $\varphi \rightarrow \psi$ for "in every stage extending the present stage where φ is true, ψ is true", and $\neg\varphi$ for "in no stage extending the present stage φ is true", we can express this 'monotonicity principle' as follows: $\varphi \rightarrow \neg\neg\varphi$ holds, and $\neg\varphi \rightarrow \neg\neg\neg\varphi$ holds. On the other hand, there are principles which may never get proved nor disproved, in other words, the principle of double negation, $\neg\neg\varphi \rightarrow \varphi$ does not hold in intuitionistic logic.

To formally define models for intuitionistic logic we need the concept of classical first-order model extending another one. Let M, M' be models for the same first-order language. M is *extended by* M', notation $M \leq M'$, iff

- $\mathrm{dom}\,(M) \subseteq \mathrm{dom}\,(M')$,
- for every $c \in \mathrm{C}$, if $\mathrm{int}\,(M')(c) \in \mathrm{dom}\,(M)$, then $\mathrm{int}\,(M')(c) = \mathrm{int}\,(M)(c)$,
- for every $P \in \mathrm{P}^n$, $\mathrm{int}\,(M)(P) \subseteq \mathrm{int}\,(M')(P)$,
- for every $f \in \mathrm{f}^n$, $\mathrm{int}\,(M)(f) \subseteq \mathrm{int}\,(M')(f)$.

Note that the final two requirements boil down to the condition that the interpretation of a relation or function symbol in $M' \restriction \mathrm{dom}\,(M)$ agrees with the interpretation of that symbol in M.

Identity Special care needs to be taken to interpret the identity symbol '=' over models for intuitionistic logic. Usually, growth of knowledge about *identities* is allowed. To this end, we provide each model M with a special two place relation $\sim^M$ with the property that $\sim^M$ is an equivalence relation on $\mathrm{dom}\,(M)$ (i.e., reflexive, symmetric and transitive), and that the predicates and functions of M respect $\sim^M$ in the following sense:

- If $d_1 \sim^M d'_1, \ldots, d_n \sim^M d'_n$, then $\langle d_1, \ldots, d_n\rangle \in \mathrm{int}\,(M)(P)$ iff $\langle d'_1, \ldots, d'_n\rangle \in \mathrm{int}\,(M)(P)$.
- If $d_1 \sim^M d'_1, \ldots, d_n \sim^M d'_n$, then $\mathrm{int}\,(M)(f)(d_1, \ldots, d_n) \sim^M \mathrm{int}\,(M)(f)(d'_1, \ldots, d'_n)$.

In other words, $\sim^M$ is a *congruence* relation. The idea is that $d_1 \sim^M d_2$ expresses that in M the two individuals d_1, d_2 are identified. Of course, things which where found out to be identical should remain identical under growth of knowledge. In other words, $M \leq M'$ should imply $\sim^M \subseteq \sim^{M'}$ (see Troelstra and Van Dalen [TD88, 2.5] for more clarification).

An *intuitionistic Kripke model* for a language L of intuitionistic predicate logic is a triple $K = \langle W, \leq, D\rangle$ such that:

- W is a non-empty set,
- $\leq$ is a partial order on W, i.e., $\leq$ is reflexive, transitive, and anti-symmetric,
- D is a function mapping every $\alpha \in W$ to a classical model for L in such manner that
$$\alpha \leq \beta \Rightarrow D(\alpha) \leq D(\beta).$$

Blurring the distinction between 'worlds' in W and the first-order models to which they are mapped, we can say that a Kripke model consists of a set of first-order models, partially ordered such that if we move upwards along $\leq$ from one 'world' to the next, the structures we encounter can only grow.

It is convenient to define variable assignments globally and term valuation functions locally, as follows. Let s be a mapping from the set V of variables to $\mathbf{D} = \bigcup\{\text{dom}\,(M) \mid M \in \text{rng}\,(D)\}$ (the union of all the domains). Let α be a member of W. Then $V_{\alpha,s}$ is the partial mapping from the set of terms to $\mathbf{D}$ defined as follows:

- $V_{\alpha,s}(v) = \begin{cases} s(v) & \text{if } s(v) \in \text{dom}\,(D(\alpha)) \\ \uparrow & \text{otherwise,} \end{cases}$
- $V_{\alpha,s}(c) = \begin{cases} \text{int}\,(D(\alpha))(c) & \text{if } c \in \text{dom}\,(\text{int}\,(D(\alpha))) \\ \uparrow & \text{otherwise,} \end{cases}$
- $V_{\alpha,s}(ft_1 \cdots t_n) = \begin{cases} \text{int}\,(D(\alpha))(f)(V_{\alpha,s}(t_1), \ldots, V_{\alpha,s}(t_n)) & \begin{array}{l}\text{if } V_{\alpha,s}(t_1) \neq \uparrow, \\ \ldots, V_{\alpha,s}(t_n) \neq \uparrow\end{array} \\ \uparrow & \text{otherwise.} \end{cases}$

It is easy to see that $V_{\alpha,s}(t) \neq\uparrow$ and $\alpha \leq \beta$ together imply $V_{\beta,s}(t) \neq\uparrow$.

The intuitionistic interpretation of the language L now proceeds from the vantage point of a member $\alpha \in W$ and takes possible 'growth of knowledge' into account. Let $K = \langle W, \leq, D\rangle$ be a Kripke model. We shall use aK for the set of all variable assignments for K. Let $s \in aK$ and let $\alpha \in K$ ("α is a node of K", i.e., α is a member of W). We first define the relation $K, s, \alpha \Vdash \varphi$ (*"node α of K forces formula φ under assignment s"*).

3.1. DEFINITION. (Forcing)

1. $K, s, \alpha \Vdash \bot$ never.
2. $K, s, \alpha \Vdash Pt_1 \cdots t_n$ iff $V_{\alpha),s}(t_1) \neq \uparrow, \ldots, V_{\alpha,s}(t_n) \neq \uparrow$, and $\langle V_{\alpha,s}(t_1), \ldots, V_{\alpha,s}(t_n)\rangle \in \text{int}\,(D(\alpha))(P)$.
3. $K, s, \alpha \Vdash t_1 = t_2$ iff $V_{\alpha,s}(t_1) \neq \uparrow$, $V_{\alpha,s}(t_2) \neq \uparrow$, and $V_{\alpha,s}(t_1) \sim^{D(\alpha)} V_{\alpha,s}(t_2)$.
4. $K, s, \alpha \Vdash \neg\varphi$ iff for all $\beta \geq \alpha$, $K, s, \beta \not\Vdash \varphi$.
5. $K, s, \alpha \Vdash (\varphi_1 \wedge \varphi_2)$ iff both $K, s, \alpha \Vdash \varphi_1$ and $K, s, \alpha \Vdash \varphi_2$.
6. $K, s, \alpha \Vdash (\varphi_1 \vee \varphi_2)$ iff either $K, s, \alpha \Vdash \varphi_1$ or $K, s, \alpha \Vdash \varphi_2$.
7. $K, s, \alpha \Vdash (\varphi_1 \rightarrow \varphi_2)$ iff for all $\beta \geq \alpha$, if $K, s, \beta \Vdash \varphi_1$ then $K, s, \beta \Vdash \varphi_2$.
8. $K, s, \alpha \Vdash \forall v\varphi$ iff for all $\beta \geq \alpha$, for all $d \in \text{dom}\,(D(\beta))$, $K, s(v|d), \beta \Vdash \varphi$.

9. $K, s, \alpha \Vdash \exists v\varphi$ iff for some $d \in \text{dom}\,(D(\alpha))$, $K, s(v|d), \alpha \Vdash \varphi$.

Next, we define $\models$ in terms of $\Vdash$:

3.2. DEFINITION. $\langle W, \leq, D\rangle, s \models \varphi$ iff for all $\alpha \in W$, $\langle W, \leq, D\rangle, s, \alpha \Vdash \varphi$.

3.3. DEFINITION. $K \models \varphi$ iff for all assignments s for K, it holds that $K, s \models \varphi$.

3.4. DEFINITION. (Intuitionistic consequence) $\Gamma \models_{IPL} \varphi$ iff for all intuitionistic Kripke models K, all $s \in aK$, if $K, s \models \gamma$ for all $\gamma \in \Gamma$, then $K, s \models \varphi$.

Note that the existential and the universal quantifier are not duals anymore under the semantic regime of intuitionistic logic. We have $\exists v\varphi \models_{IPL} \neg\forall v \neg\varphi$, but not the other way around.

But first, let us move to a still more general level.

3.5. DEFINITION. A *Kripke frame* is a triple $F = \langle W, \leq, \mathcal{O}\rangle$ where

1. $\langle W, \leq\rangle$ is a partial order.
2. $\mathcal{O}$ assigns to every $\alpha \in W$ a non–empty set $\mathcal{O}_\alpha$ such that, if $\alpha \leq \beta$ then $\mathcal{O}_\alpha \subseteq \mathcal{O}_\beta$.

We say that $K = \langle W, \leq, D\rangle$ is a *model over* Kripke frame $F = \langle W, \leq, \mathcal{O}\rangle$ if for all $\alpha \in M$: $\text{dom}\,(D(\alpha)) = \mathcal{O}_\alpha$.

3.6. DEFINITION. (Frame validity) For F a Kripke frame, φ is *frame valid* on F, notation $F \models \varphi$, if $K \models \varphi$ for every Kripke model K over F.

3.7. DEFINITION. (Frame definition by a formula) Let A be a class of Kripke frames: φ *defines* A if $F \in A$ if and only if $F \models \varphi$.

These definitions will be used in Section 3.4.1.

3.2.2 Natural Deduction for Intuitionistic Predicate Logic

A natural deduction system for IPL can be got from the natural deduction system for classical predicate logic in Section 2.2.2 by leaving out the following rule:

$$\frac{\begin{array}{c}\neg\varphi\ [i]\\ \vdots\\ \bot\end{array}}{\varphi}\ \neg\mathrm{E}_i$$

As we have seen in Section 2.2.2, this boils down to leaving out the law of double negation:

$$\frac{\neg\neg\varphi}{\varphi}\ DN$$

Recall that EM is the principle of excluded middle:

$$\varphi \vee \neg\varphi. \qquad \text{EM}$$

3.8. PROPOSITION.

$$\vdash_{IPL+EM}\varphi \iff \vdash_{IPL+DN}\varphi \iff \vdash_c \varphi$$

Proof: Assuming EM, the law of double negation can be derived:

$$\dfrac{\dfrac{\varphi\vee\neg\varphi \qquad \varphi\ [2] \qquad \dfrac{\neg\varphi\ [3] \qquad \neg\neg\varphi\ [1]}{\varphi}\,\bot}{\varphi}\,\vee\mathrm{E}_{2,3}}{\neg\neg\varphi\to\varphi}\,\to\mathrm{I}_1$$

Conversely, given DN, we can derive the principle of excluded middle:

$$\dfrac{\dfrac{\dfrac{\dfrac{\neg(\varphi\vee\neg\varphi)\ [3] \qquad \dfrac{\varphi\ [1]}{(\varphi\vee\neg\varphi)}\,\vee I}{\bot}\,\neg I_1}{\neg\varphi}\,\neg E \qquad \dfrac{\dfrac{\neg(\varphi\vee\neg\varphi)\ [3] \qquad \dfrac{\neg\varphi\ [2]}{(\varphi\vee\neg\varphi)}\,\vee I}{\bot}\,\neg I_2}{\neg\neg\varphi}\,\neg E}{\dfrac{\bot}{\dfrac{\neg\neg(\varphi\vee\neg\varphi)}{\varphi\vee\neg\varphi}\,DN}\,\neg I_3}\,\neg E$$

Because DN allows us to get $\neg$ E_i as a derived rule, all classical inference rules are present in IPL+DN. ⊠

What the proof trees above show is that the principle DN and the schema EM are rule equivalent, and that adding either principle to intuitionistic logic yields classical logic.

It is interesting to check what the weakening of the system entails for the quantifier interaction principles. The following intuitionistic proof tree shows that we still have $\neg\exists v\varphi \vdash_{IPL} \forall v\neg\varphi$:

$$\dfrac{\dfrac{\dfrac{\neg\exists v\varphi \qquad \dfrac{\varphi\ [1]}{\exists v\varphi}\,\exists\mathrm{I}}{\bot}\,\neg\mathrm{E}}{\neg\varphi}\,\neg\mathrm{I}_1}{\forall v\neg\varphi}\,\forall\mathrm{I}$$

Also, we still have $\forall v \neg\varphi \vdash_{IPL} \neg\exists v\varphi$:

$$
\begin{prooftree}
\AxiomC{$\exists v\varphi$ [1]}
\AxiomC{$\forall v \neg\varphi$}
\RightLabel{$\forall$E}
\UnaryInfC{$\neg\varphi(t/v)$}
\AxiomC{$\varphi(v/t)$ [2]}
\RightLabel{$\neg$E}
\BinaryInfC{$\bot$}
\RightLabel{$\exists\mathrm{E}_2$}
\BinaryInfC{$\bot$}
\RightLabel{$\neg\mathrm{I}_1$}
\UnaryInfC{$\neg\exists v\varphi$}
\end{prooftree}
$$

Another principle which we have still got is $\exists v \neg\varphi \vdash_{IPL} \neg\forall v\varphi$:

$$
\begin{prooftree}
\AxiomC{$\exists v \neg\varphi$}
\AxiomC{$\forall v\varphi$ [2]}
\RightLabel{$\forall$E}
\UnaryInfC{φ}
\AxiomC{$\neg\varphi(t/v)$ [1]}
\RightLabel{$\neg$E}
\TrinaryInfC{$\bot$}
\RightLabel{$\neg\mathrm{E}_2$}
\UnaryInfC{$\neg\forall v\varphi$}
\RightLabel{$\exists\mathrm{E}_1$}
\UnaryInfC{$\neg\forall v\varphi$}
\end{prooftree}
$$

We can summarize these facts as follows:

$$\vdash_{IPL} \exists v \neg\varphi \rightarrow \neg\forall v\varphi. \tag{3.1}$$

$$\vdash_{IPL} \neg\exists v\varphi \leftrightarrow \forall v \neg\varphi. \tag{3.2}$$

But we have lost the fourth quantifier interaction principle. As Kolmogorov [Kol67] already observed, $\neg\forall v\varphi \rightarrow \exists v\neg\varphi$ can only be derived with the principle of excluded middle (or equivalently, with the negation elimination rule that turns intuitionistic logic into classical logic).

Here is a list of further quantifier principles which are intuitionistically derivable (from Van Dalen [Dal86]):

$$\vdash_{IPL} \varphi \vee \forall v\psi \rightarrow \forall v(\varphi \vee \psi) \ \text{ provided } v \notin FV(\varphi). \tag{3.3}$$

$$\vdash_{IPL} \forall v(\varphi \rightarrow \psi) \leftrightarrow (\varphi \rightarrow \forall v\psi) \ \text{ provided } v \notin FV(\varphi). \tag{3.4}$$

$$\vdash_{IPL} \forall v(\varphi \rightarrow \psi) \leftrightarrow (\exists v\varphi \rightarrow \psi) \ \text{ provided } v \notin FV(\psi). \tag{3.5}$$

$$\vdash_{IPL} \exists v(\varphi \rightarrow \psi) \rightarrow (\varphi \rightarrow \exists v\psi) \ \text{ provided } v \notin FV(\varphi). \tag{3.6}$$

$$\vdash_{IPL} \exists v(\varphi \rightarrow \psi) \rightarrow (\forall v\varphi \rightarrow \psi) \ \text{ provided } v \notin FV(\psi). \tag{3.7}$$

$$\vdash_{IPL} \neg\neg\forall v\varphi \rightarrow \forall v\neg\neg\varphi. \tag{3.8}$$

Van Dalen [Dal86] gives the following list of quantifier interaction principles that we have lost:

$$\not\vdash_{IPL} \neg\forall v\varphi \rightarrow \exists v\neg\varphi. \tag{3.9}$$

$$\not\vdash_{IPL} \forall v\neg\neg\varphi \rightarrow \neg\neg\forall v\varphi. \tag{3.10}$$

$$\nvdash_{IPL} \forall v((\varphi \vee \psi) \to (\varphi \vee \forall v \psi)) \text{ provided } v \notin FV(\varphi). \tag{3.11}$$

$$\nvdash_{IPL} (\varphi \to \exists v \psi) \to \exists v(\varphi \to \psi) \text{ provided } v \notin FV(\varphi). \tag{3.12}$$

$$\nvdash_{IPL} (\forall v \varphi \to \psi) \to \exists v(\varphi \to \psi) \text{ provided } v \notin FV(\psi). \tag{3.13}$$

$$\nvdash_{IPL} (\forall v(\varphi \vee \neg\varphi) \wedge \neg\neg\exists v \varphi) \to \exists v \varphi. \tag{3.14}$$

$$\nvdash_{IPL} \neg\neg\forall v \forall w(v = w \vee v \neq w). \tag{3.15}$$

$$\nvdash_{IPL} \neg\neg\forall v \forall w(\neg v \neq w \to v \neq w). \tag{3.16}$$

In the next section we shall investigate how this situation changes when we add the epsilon rule to intuitionistic logic.

IPL has the classical rules for identity. However, this does not force the identity symbol to be interpreted by real identity. Let DE, the schema for *decidable identity*, be given by

$$\forall x \forall y(x = y \vee \neg(x = y)). \tag{DE}$$

Under interpretation of '=', by real identity, $K \models \forall x \forall y(x = y \vee \neg(x = y))$ for all Kripke models K, but DE is not an IPL theorem (see [Dal86]).

IPL is sound and complete for the class of Kripke models.

3.9. THEOREM. (Completeness of IPL) *For all sets of formulas* $\Gamma \cup \{\varphi\}$:

$$\Gamma \models_{IPL} \varphi \Rightarrow \Gamma \vdash_{IPL} \varphi.$$

A more general notion of completeness is given in the following definition.

3.10. DEFINITION. (Model completeness) Formula φ is *complete* for model class A if $\vdash_{IPL+\varphi} \psi \iff \forall K \in A : K \Vdash \psi$.

For example, the formula

$$(\varphi \to \psi) \vee (\psi \to \varphi) \tag{LIN}$$

is complete for the class of linear models (a model K is linear if $\alpha \leq \beta$ or $\beta \leq \alpha$, for all α, β in the set of nodes of K). As another example, the formula schema DE is complete for the class $A_=$ of Kripke models where '=' is interpreted by real identity (see Van Dalen [Dal86], or Gabbay [Gab81]).

3.3 Intensional Intuitionistic Epsilon Logic

Proof theoretically, it is no problem to add the epsilon rule to intuitionistic logic. The *intensional intuitionistic epsilon calculus* is the result of adding the epsilon rule

$$\frac{\exists v \varphi}{\varphi[\epsilon v : \varphi / v]} \exists E_\epsilon \tag{ϵ}$$

plus the alphabetic variant rule

$$\frac{}{(\epsilon v : \varphi) = (\epsilon w : \varphi(w/v))} \tag{AV}$$

to the intuitionistic rules. Or equivalently, intensional intuitionistic epsilon logic is the result of deleting the rule for negation elimination from the calculus for classical epsilon logic. Let us use $\vdash_{IPL+\epsilon}$ for derivability in this logic. We are now going to explore some of its properties.

Very little knowledge of the proof theory for intuitionistic predicate logic suffices to see that adding the epsilon rule to IPL is not conservative. The epsilon rule allows us to derive Plato's principle (abbreviated P∃ in the sequel) in intuitionistic logic:

$$\exists v(\exists v\varphi \to \varphi). \tag{P∃}$$

Here is a derivation of this principle in the calculus IPL+ϵ.

$$\dfrac{\dfrac{\dfrac{\exists x\varphi\ [1]}{\varphi[\epsilon x : \varphi/x]}\ \exists\mathrm{E}_\epsilon}{\exists x\varphi \to \varphi[\epsilon x : \varphi]}\ {\to}\mathrm{I}_1}{\exists x(\exists x\varphi \to \varphi)}\ \exists\mathrm{I}$$

The principle P∃ will be shown to be central to the IPL+ϵ calculus. Therefore we shall also consider a well-know equivalent form.

3.11. PROPOSITION. *P∃ is equivalent to the rule called* Independence of Premise IP∃*:*

$$\frac{\varphi \to \exists x\psi}{\exists x(\varphi \to \psi)},\ \textit{provided } x \textit{ does not occur free in } \varphi. \tag{IP∃}$$

Proof: Two simple derivations show this[1].
From *IP* to P∃:

$$\dfrac{\dfrac{\exists x\varphi(x)}{\exists x\varphi(x) \to \exists x\varphi(x)}}{\exists x(\exists x\varphi(x) \to \varphi(x))}$$

From P∃ to *IP*:

$$\dfrac{\exists x(\exists x\varphi(x) \to \varphi(x)) \qquad \dfrac{\dfrac{\dfrac{\dfrac{\dfrac{[\psi] \quad \psi \to \exists x\varphi(x)}{\exists x\varphi(x)} \quad [\exists x\varphi(x) \to \varphi(x)]}{\varphi(x)}}{\psi \to \varphi(x)}}{\exists x(\psi \to \varphi(x))}}{\exists x(\psi \to \varphi(x))}$$

[1]Because in this chapter we shall give a great number of extended derivations, for reasons of display we do not always annotate the deductions by the proof rules used. Moreover, when subdeductions are simple they will often be contracted.

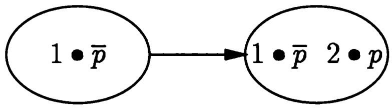

Figure 3.1: Intuitionistic counterexample to Plato's principle.

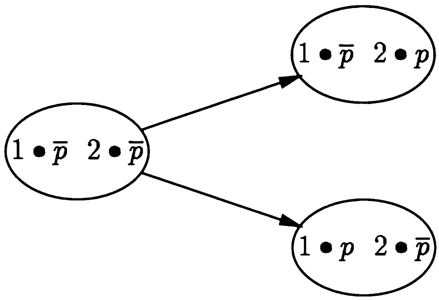

Figure 3.2: Model without a Skolem function for $\exists x Px \to Py$.

⊠

However, we know that $\not\vdash_{IPL} \exists v(\exists v\varphi \to \varphi)$, because of the completeness of IPL and the fact that there are Kripke models M of IPL with $M, \alpha \models \neg\exists x(\exists Px \to Px)$ (where α is a node in the Kripke model). Figure 3.1 gives such a counterexample. The model has just two nodes. In the first of these $\exists x Px$ does not hold, in the second node it does. The point is that the existential formula is made true by a *new* object which did not yet exist at the first node. So the only object at the first node, object 1, does *not* have the property that if some object acquires property P (at some accessible node), then 1 will acquire P. The conclusion must be that adding the epsilon rule is *not* a conservative extension of intuitionistic predicate logic.

If we think of this semantically, it is easy to see why the earlier expansion argument fails in the case of intuitionistic logic. Again, assume we have a formula ψ with $\not\vdash_{IPL} \psi$. Then there is a Kripke model K, an assignment s, with a node α at which ψ fails for s: $K, s, \alpha \not\models \psi$. Suppose we want to expand the model K with Skolem functions to interpret the epsilon terms. How should this be done?

Consider the example in Figure 3.2, and assume that we want to add a Skolem function for the formula $\exists x Px \to Py$. The list of parameters is empty in this case, so this should be a Skolem constant. It is easy to see that the formula is false in the initial node of the model and that it is made true by different objects in the two accessible nodes. So which object should one assign to $\epsilon y : (\exists x Px \to Py)$? It cannot be 1, for this is the wrong choice in one of the nodes, and it cannot be 2, for this is the wrong choice in another node. Van Dalen [Dal86] gives a formal proof of the fact that adding Skolem functions is not conservative over intuitionistic logic, based on an example by Smoryński.

This problem squares with the fact that IPL does not have prenex forms for formulas: it is impossible to rewrite an arbitrary IPL formula φ into the equivalent form $Q_1x_1 \cdots Q_nx_n\varphi'$, with Q_i equal to $\forall$ or $\exists$, and φ' quantifier free, as can be done for classical predicate logic. Prenex formulas are equivalent in turn to purely universal formulas with Skolem functions, so if prenex forms do exist, Skolemization is conservative.

So the addition of the epsilon rule to IPL is not conservative. As we shall see in Section 3.4.1, it does not give us classical predicate logic. But if we also add the extensionality principle of Section 2.4.1

$$\dfrac{\begin{matrix} \varphi(v_1, \ldots, v_n, v)\ [i] \\ \vdots \\ \psi(v_n, \ldots, v_n, v) \end{matrix} \qquad \begin{matrix} \psi(v_1, \ldots, v_n, v)\ [j] \\ \vdots \\ \varphi(v_1, \ldots, v_n, v) \end{matrix}}{\epsilon v\varphi = \epsilon v\psi}\ \epsilon\mathrm{I}_{i,j} \qquad \text{EXT}$$

we end up with classical logic (see [Bel93], [GM78]). It is convenient to take EXT in schematic form:

$$\forall v(\varphi \leftrightarrow \psi) \to (\epsilon v : \varphi = \epsilon v : \psi). \qquad \text{EXT}$$

On the assumption that there are two objects with $a \neq b$, adding ϵ plus EXT to IPL yields classical predicate logic.

3.12. PROPOSITION. *Assuming that there are two distinct objects, then the addition of EXT+ ϵ to IPL gives CPL.*

Proof: We shall show that under the assumption given, the law of excluded middle is IPL derivable from EXT. That is, $a \neq b, \text{EXT} \vdash_{IPL+\epsilon} \text{EM}$. Proposition 3.8 then gives the desired result.

1: $\exists y(y = a \vee \varphi)$	*given*
2: $\exists y(y = b \vee \varphi)$	*given*
3: $(\epsilon y : (y = a \vee \varphi)) = a \vee \varphi$	*1,ϵ-rule*
4: $(\epsilon y : (y = b \vee \varphi)) = b \vee \varphi$	*2,ϵ-rule*
5: $((\epsilon y : (y = a \vee \varphi)) = a \wedge (\epsilon y : (y = b \vee \varphi)) = b) \vee \varphi$	*3,4, distributivity*
6: $((\epsilon y : (y = a \vee \varphi)) \neq (\epsilon y : (y = b \vee \varphi))) \vee \varphi$	$5, a \neq b$
7: $\varphi \vdash_\epsilon \forall y((y = a \vee \varphi) \leftrightarrow (y = b \vee \varphi))$	*CPL*
8: $\varphi \vdash_\epsilon (\epsilon y : (y = a \vee \varphi)) = (\epsilon y : (y = b \vee \varphi))$	*7, EXT*
9: $(\epsilon y : (y = a \vee \varphi)) \neq (\epsilon y : (y = b \vee \varphi)) \vdash_\epsilon \neg\varphi$	*8, contraposition*

10: $\varphi \vee \neg\varphi$ *6,9*

⊠

Notice that only intuitionistically valid inference rules are used in this proof.

3.3.1 Plato's Principle and the ϵ-Rule

In this section will show that Plato's principle covers exactly the increase in derivational strength we get when we add the ϵ-rule to IPL.

3.13. THEOREM. *Let $\Sigma \cup \{\chi\}$ be a set of ϵ-free formulas. Then*

$$\Sigma \vdash_{IPL+\epsilon} \chi \iff \Sigma \vdash_{IPL+P\exists} \chi.$$

Proof: The direction from right to left: if $\Sigma \vdash_{IPL+P\exists} \chi$, then $\Sigma \vdash_{IPL} P\exists \rightarrow \chi$. But $\Sigma \vdash_{IPL+\epsilon} P\exists$ and so $\Sigma \vdash_{IPL+\epsilon} \chi$.

The proof of the left to right direction will take the form of a *conservativity proof*. Our strategy will be the following. We are going to show that, for Σ, χ ϵ-free, every ϵ-derivation of χ from Σ can be transformed into an IPL+P$\exists$ derivation. This we do by induction on the number n of ϵ-terms occurring in the ϵ-derivation of χ from Σ. If $n = 0$, then the derivation can have no application of the ϵ-rule, so χ follows already from Σ by pure IPL principles. If $n = m + 1$, then we are going to transform the derivation of χ from Σ in such a way that we have a new derivation of χ from Σ, but with all occurrences of one ϵ-term removed. Because any derivation uses only a finite number of ϵ-terms, this proves the theorem.

The following diagram captures our starting position for the removal of all occurrences of a selected ϵ-term. We have an ϵ-derivation $\mathcal{D}$ in which $m + 1$ different ϵ-terms are used.

$$\begin{array}{c} \Sigma \\ \mathcal{D}_1 \\ \Gamma \; \dfrac{\exists x\varphi}{\varphi[\epsilon x : \varphi/x]} \\ \mathcal{D}_2 \\ \psi \\ \vdots \\ \chi \end{array}$$

Here Σ is a set of the undischarged, ϵ-free, assumptions of the derivation $\mathcal{D}$, and χ is the ϵ-free conclusion of $\mathcal{D}$. A specific application of the ϵ-rule in this derivation is highlighted: $\exists x\varphi/\varphi[\epsilon x : \varphi/x]$. $\exists x\varphi$ is the conclusion of the subderivation $\mathcal{D}_1$ with assumptions in Σ. We assume that no applications of the ϵ-rule with proper term $\epsilon x : \varphi$ occurs in the subderivation $\mathcal{D}_1$. That is, we assume the highlighted application to be a *highest* one in $\mathcal{D}$. Because the assumptions in Σ and the conclusion χ do not contain any ϵ-terms, there must be a highest formula below $\varphi[\epsilon x : \varphi/x]$ in the derivation in which $\epsilon x : \varphi$ no longer occurs, and which does not depend on an assumption in which $\epsilon x : \varphi$ occurs: here this formula is denoted by ψ. The derivation $\mathcal{D}_2/\psi$ may use assumptions not in Σ: these are all contained in the set Γ (the elements of Γ will be discharged below ψ). So $\mathcal{D}_2$ is a derivation with assumptions in $\Gamma \cup \{\varphi[\epsilon x : \varphi/x]\}$ and conclusion ψ.

Now we are set to transform $\mathcal{D}$ to a derivation $\mathcal{D}'$ in which the term $\epsilon x : \varphi$ no longer occurs. Of course, Plato's principle will be used essentially. Consider the subderivation $\mathcal{D}_2$ with assumptions in $\Gamma \cup \{\varphi[\epsilon x\varphi/x]\}$ and conclusion ψ. We transform $\mathcal{D}_2$ to a derivation $\mathcal{D}_2'$ as follows.

$$\mathcal{D} \quad \begin{array}{c} \Gamma \\ \mathcal{D}_2 \\ \psi \end{array} \quad \Rightarrow \quad \mathcal{D}' \quad \dfrac{\exists x(\exists x\varphi \to \varphi) \qquad \begin{array}{c} \exists x\varphi \to \varphi[a/x]\ [1] \\ \mathcal{D}_2[a/\epsilon x : \varphi] \\ \psi \end{array}}{\psi}\ \exists\mathrm{E}_1$$

Here a is a fresh constant. Firstly, this a correct ($\exists$E) application. Because $\epsilon x : \varphi$ does not occur in Γ,ψ, or φ, a does not occur in $\Gamma[a/\epsilon x : \varphi]$, $\psi[a/\epsilon x\varphi]$ or $\varphi[a/\epsilon x : \varphi]$. Furthermore, $\mathcal{D}_2[a/\epsilon x : \varphi]$ is almost a correct derivation given the assumptions $\exists x\varphi \to \varphi[a/x]$. Every application of an inference rule other than the ϵ-rule, remains unaffected by the substitution of a for $\epsilon x : \varphi$. But applications of the ϵ-rule $\exists x\varphi/\varphi[\epsilon x : \varphi/x]$ in $\mathcal{D}_2$ become $\exists x\varphi/\varphi[a/x]$ in $\mathcal{D}_2[a/\epsilon x : \varphi]$, which are not correct proof steps. However, here is where the assumptions have their use. We we now transform $\mathcal{D}_2[a/\epsilon x : \varphi]$ to $\mathcal{D}_2'$ by reformulating every invalid $\exists x\varphi/\varphi[a/x]$ proof step as follows.

$$\mathcal{D}_2[a/\epsilon x : \varphi] \quad \dfrac{\exists x\varphi}{\varphi[a/x]} \quad \Rightarrow \quad \mathcal{D}_2' \quad \dfrac{\exists x\varphi \qquad \exists x\varphi \to \varphi[a/x]}{\varphi[a/x]}$$

Consequently, $\mathcal{D}_2'$ is still a correct derivation of ψ from assumptions in $\Gamma \cup \{\exists x\varphi \to \varphi[a/x]\}$. So the transformation of $\mathcal{D}_2$ gives a correct ($\exists$E) application with major premise $\exists x(\exists x\varphi \to \varphi)$, an instance of Plato's principle. Because *all* applications of the ϵ-rule in $\mathcal{D}$ with proper term $\epsilon x : \varphi$ lie in $\mathcal{D}_2$, by replacing $\mathcal{D}_2$ by $\mathcal{D}_2'$ we have removed all occurrences of the term $\epsilon x : \varphi$ from $\mathcal{D}$. Consequently, if $\mathcal{D}$ uses $m + 1$ ϵ-terms, then $\mathcal{D}[\mathcal{D}_2'/\mathcal{D}_2]$ uses m ϵ-terms. ⊠

Because the ϵ-rule derives Plato's principle with the help of ($\to$I) and ($\exists$I), and, conversely, Plato's principle allows us to transform any ϵ-derivation to an ϵ-free derivation, with the help of ($\to$E) and ($\exists$E), we can state the following corollary:

3.14. COROLLARY. *Let C be a logical system containing the rules* ($\exists$E),($\exists$I), ($\to$I), ($\to$E). *Then the following are equivalent:*

1. *the ϵ-rule is conservative over C,*
2. IP$\exists$ *is a (derived) rule of C,*
3. *P$\exists$ is a theorem of C.*

3.3.2 Interpretation of ϵ-Terms

Intuitionistically, the interpretation of ϵ-terms faces the problem that the addition of Skolem functions is not conservative over IPL. This situation changes in the extended system IPL + P$\exists$. First remark that if k is a model for intuitionistic predicate logic then $K \models \exists x(\exists \varphi(x, \overline{y}) \rightarrow \varphi(x, \overline{y}))$, an instance of Plato's principle, implies $K \Vdash \forall \overline{y} \exists x(\exists x \varphi(x, \overline{y}) \rightarrow \varphi(x, \overline{y}))$. Now we shall analyze the meaning of this schema on Kripke models according to the forcing definition. We can state the following proposition

3.15. PROPOSITION. *Let α be a node in an intuitionistic Kripke model K and suppose*

$$K, s, \alpha \Vdash \forall \overline{y} \exists x(\exists x \varphi(x, \overline{y}) \rightarrow \varphi(x, \overline{y})),$$

where $\overline{y}$ is the sequence of the n free variables in $\exists x \varphi$. Then for all $\beta \geq \alpha$ and all n-ary sequences $\overline{b}$ of elements form $dom(D(\beta))$ there is an $m \in dom(D(\beta))$ such that for all $\gamma \geq \beta$

1. $K, s(\overline{y}|\overline{b})(x|m), \gamma \Vdash \varphi(\overline{y}) \iff K, s(\overline{y}|\overline{b}), \gamma \Vdash \exists x \varphi(\overline{y})$,
2. $K, s(\overline{y}|\overline{b})(x|m), \gamma \Vdash \neg\varphi(\overline{y}) \iff K, s(\overline{y}|\overline{b}), \gamma \Vdash \forall x \neg\varphi(x, \overline{y})$.

Proof:
1) $\Rightarrow$, by the forcing definition.
1) $\Leftarrow$, because $K, s(\overline{y}|\overline{b}), \beta \Vdash \exists x(\exists x \varphi(x, \overline{y}) \rightarrow \varphi(x, \overline{y}))$, there is an $m \in dom(D(\beta))$ such that

$$K, s(\overline{y}|\overline{b})(x|m), \beta \Vdash \exists x \varphi(x, \overline{y}) \rightarrow \varphi(x, \overline{y}).$$

So if $K, s(\overline{y}|\overline{b}), \gamma \Vdash \exists x \varphi(x, \overline{y})$ for $\beta \leq \gamma$, then $K, s(\overline{y}|\overline{b})(x|m), \gamma \Vdash \varphi(x, \overline{y})$.
2) $\Rightarrow$, if for all $\gamma \geq \beta$: $K, s(\overline{y}|\overline{b}), (x|m), \gamma \nVdash \varphi(x, \overline{y})$, then for all $\gamma \geq \beta$:

$$K, s(\overline{y}|\overline{b}), \gamma \nVdash \exists x \varphi(x, \overline{y}).$$

Consequently $K, s(\overline{y}|\overline{b}), \beta \Vdash \neg\exists x \varphi(x, \overline{y})$.
2) $\Leftarrow$, by the forcing definition. ⊠

By this proposition, the principle P$\exists$ guarantees the existence of elements in the domain of any node in a Kripke model suitable for the interpretation of values of Skolem functions. And, in fact we have:

3.16. COROLLARY. *The addition of Skolem functions f for every existential formula $\exists x \varphi$ in a logic $L \supseteq IPL$ is conservative over L if and only if $\vdash_L \exists x(\exists x \varphi \rightarrow \varphi)$.*

Proof: For Σ a set of IPL formulas, let Σ^s be formed from Σ by adding function symbols $f_{\varphi(x,\overline{y})}$ to the language for every formula $\exists x\varphi(x,\overline{y})$, and adding the axioms $\forall\overline{y}(\exists x\varphi(x,\overline{y}) \to \varphi(f(\overline{y}),\overline{y}))$ to Σ.
$\Leftarrow$: We have to show that $\Sigma^s \vdash_{IPL+P\exists} \psi$ implies $\Sigma \vdash_{IPL+P\exists} \psi$ for ψ not containing any Skolem function symbol. Assume $\Sigma \nvdash_{IPL+P\exists} \psi$, where ψ has no Skolem functions. We shall show that $\Sigma^s \nvdash_{IPL+P\exists} \psi$. The assumption gives us, by completeness of IPL with respect to the class of Kripke models, a model K with a node α such that (1) $K \Vdash \forall\overline{y}\exists x(\exists x\varphi \to \varphi)$, (2) $K, \alpha \Vdash \Sigma$, and (3) $K, \alpha \nVdash \psi$. Because

$$K \Vdash \forall\overline{y}\exists x(\exists x\varphi \to \varphi),$$

we can define at each α a function $F^\alpha_{\varphi(x,\overline{y})} : dom(D(\alpha))^n \to dom(D(\alpha))$, such that

$$F^\alpha_{\varphi(x,\overline{y})}(\overline{d}) \in \{d' \in dom(D(\alpha)) \mid \mathcal{M}_C, \alpha, s(\overline{y}|\overline{d}) \Vdash \exists x\varphi(x,\overline{y}) \to \varphi(d',\overline{y})\}.$$

Because Plato's principle is valid on K, by Proposition 3.15 such a function is well-defined for all $\varphi(x,\overline{y})$. Now we extend the interpretation function 'Int' at every node $\alpha \in W$ to an interpretation 'Int^s' for function constants $f_{\varphi(x,\overline{y})}$ (for every formula $\varphi(x,\overline{y})$) in such a way that

1. $\text{Int}^s(D(\alpha))(f_{\varphi(x,\overline{y})}) = F^\alpha_{\varphi(x,\overline{y})}$ at every node $\alpha \in W$,
2. if $\alpha \leq \beta$, then $\text{Int}^s(D(\alpha))(f_{\varphi(x,\overline{y})}) = \text{Int}^s(D(\beta))(f_{\varphi(x,\overline{y})}) \restriction dom(D(\alpha))$.

This extension gives a model K^s. We then have for $\alpha \in W^s$,

$$\text{if } K, s, \alpha \Vdash \forall\overline{y}\exists x\varphi(x,\overline{y}), \text{ then } K, s, \alpha \Vdash \forall\overline{y}\varphi(f_{\varphi(x,\overline{y})}(\overline{y}),\overline{y}).$$

So, Int^s interprets Skolem functions in the appropriate way. Because ψ has no Skolem functions,

$$\text{if } K, s, \alpha \Vdash \Sigma \text{ and } K, s, \alpha \nVdash \psi, \text{ then } K^s, s, \alpha \Vdash \Sigma^s \text{ and } K^s, s, \alpha \nVdash \psi.$$

Consequently, $\Sigma^s \nvdash_{IPL+P\exists} \psi$.
$\Rightarrow$: We have to show that if the addition of Skolem functions is conservative over $L \supseteq IPL$, then L derives all instances of Plato's Principle. Suppose $\Sigma \nvdash_L P\exists$. By the canonical model construction we can create a Kripke model K for L, and a node α forcing Σ in K such that for some instance of Plato's principle, $K, s, \alpha \nVdash \forall\overline{y}\exists x(\exists x\varphi \to \varphi)$. Now, for α^s, α extended by an interpretation for Skolem functions, we have $K, s, \alpha^s \Vdash \Sigma^s$ and $K, s, \alpha^s \Vdash \forall\overline{y}(\exists x\varphi(x,\overline{y}) \to \varphi(f(\overline{y}),\overline{y})$. Here we may assume f to be the only Skolem function symbol occurring in the formula. Thus $K, s, \alpha^s \Vdash \forall\overline{y}\exists x(\exists x\varphi \to \varphi)$. But this is a formula without Skolem function symbols, and $K, s, \alpha \nVdash \forall\overline{y}\exists x(\exists x\varphi \to \varphi)$. By soundness of L for K, we then have $\Sigma^s \vdash \forall\overline{y}\exists x(\exists x\varphi \to \varphi)$ while $\Sigma \nvdash_L \forall\overline{y}\exists x(\exists x\varphi \to \varphi)$. Consequently, the addition of Skolem functions is not conservative over L. ⊠

$$\begin{tikzcd}
 & & CPL & & \\
 & & P\exists + P\forall \arrow[u] & & \\
P\exists \arrow[urr] & & EN + UN \arrow[u] & & P\forall \arrow[ull] \\
EN \arrow[u] \arrow[urr] & & & & UN \arrow[u] \arrow[ull] \\
 & & IPL \arrow[ull] \arrow[uu] \arrow[urr] & &
\end{tikzcd}$$

Figure 3.3: Inclusion diagram of the IPL extensions

3.17. DEFINITION. (Interpretation of ϵ-Terms) Let K be a Kripke model for IPL+P$\exists$ and let Int^s be an extension of the interpretation function Int of K over all Skolem function symbols. We now define the valuation function $V_{\alpha,s,\Phi}$ which is like $V_{\alpha,s}$ on all non ϵ-terms terms and for ϵ-term $\epsilon x : \varphi(x, \overline{y})$ is given by $V_{\alpha,s,\Phi}(\epsilon x : \varphi(x, \overline{y})) = \mathrm{Int}^s(D(\alpha))(f_{\varphi(x,\overline{y})})(s(\overline{y}))$. This creates a Kripke model $\mathcal{M}_\Phi$ which interprets ϵ-terms.

The following proposition follows straightforwardly.

3.18. PROPOSITION. *Let K be a Kripke model on which Plato's principle is valid. Then* $K_\Phi \Vdash \exists x\varphi \leftrightarrow \varphi[\epsilon x\varphi/x]$.

3.4 Intermediate Logics

The last section has shown that, for φ ϵ-free: $\vdash_{IPL+\epsilon} \varphi \iff \vdash_{IPL+P\exists} \varphi$. In the next section, the logic IPL+P$\exists$ will be the subject of investigation as an *intermediate* logic which is of interest in its own right. In fact, in the following sections we shall distinguish a *family* of intermediate logics.

A map of the logics that will be developed in the following sections is given in Figure 3.3. Here the arrows represent inclusion, the absence of arrows represents non-inclusion. As we shall show, all inclusions are proper.

The logics P$\exists$ (Section 3.4.1) and P$\forall$ (Section 3.4.2) are determined by the principles $\exists x(\exists x\varphi \to \varphi)$ and $\exists x(\varphi \to \forall x\varphi)$ respectively. They arise by addition to IPL of the epsilon and the tau rule respectively.

3.19. DEFINITION. (τ-Rule)

$$\frac{\varphi[\tau x : \varphi/x]}{\forall x\varphi}\forall I_\tau$$

The logics EN (Section 3.4.1) and UN (Section 3.4.2) are determined by $\neg\neg\exists x\varphi \rightarrow \exists x\neg\neg\varphi$ and $\forall x\neg\neg\varphi \rightarrow \neg\neg\forall x\varphi$ respectively. They arise in the addition of the rules

$$\epsilon\text{DN}\ \frac{\exists x\varphi}{\neg\neg\varphi[\epsilon x : \varphi/x]} \qquad \tau\text{DN}\frac{\varphi[\tau x : \varphi/x]}{\neg\neg\forall x\varphi}$$

to IPL. The rule ϵDN derives $\exists x(\exists x\varphi \rightarrow \neg\neg\varphi)$ straightforwardly. By Proposition 3.22 this will be shown to be equivalent to EN. The rule τDN derives UN. $\forall x\neg\neg\varphi$ gives $\neg\neg\varphi[\tau x : \varphi/x]$ by ($\forall$E). This gives, by τDN, $\neg\neg\neg\neg\forall x\varphi$ which is IPL equivalent to $\neg\neg\forall x\varphi$.

We shall not discuss the logic $EN + UN$. In Section 3.4.3 we shall discuss the logic P$\exists$ + $P\forall$. This is the only intermediate logic from Figure 3.3 which is complete (modulo decidable equality).

3.4.1 The Logic IPL+P$\exists$

Because Plato's principle is not derivable in IPL the logic IPL+P$\exists$ is stronger than IPL. That it is a proper intermediate logic (i.e., weaker than classical logic) is the import of the following proposition.

3.20. PROPOSITION. $m \not\vdash_{IPL+P\exists} \neg\forall x\varphi \rightarrow \exists x\neg\varphi$.

Proof: Let $K = \langle \mathbb{N}, \leq, \mathcal{O}\rangle$ be a Kripke frame where $\leq$ is the standard ordering on $\mathbb{N}$, and let the domain associated with every node be the set of natural numbers. It is easy to see that Plato's principle is valid on this model. Now let P be a monadic predicate and let the interpretation function for node n, Int_n, be defined by $\text{Int}_n(P) = \{m \mid m \leq n\}$. Then, for every $n \in M$ we have $K, n \Vdash \neg\forall x\varphi$, for every successor m of n has a successor l and a k in the domain of l such that $K, s(x|k), l \nVdash P(x)$, for at every node there are elements of the domain not (yet) in the extension of P. But there is no element k in the domain such that $K, s(x|k), n \Vdash \neg P(x)$, for every element of the domain will eventually end up in the extension of P. ⊠

Proof Theory

In this section we shall explore some of the derivational power of IPL+P$\exists$. We shall concentrate on IPL+P$\exists$ principles that are not IPL valid. First we derive a principle weaker than P$\exists$ and consider some of its consequences.

$$\neg\neg\exists x\neg\varphi \leftrightarrow \exists x\neg\varphi. \tag{DN$\exists$}$$

3.21. PROPOSITION. $\vdash_{IPL+\text{P}\exists}$ DN$\exists$.

Proof: We shall derive DN∃ in the IPL+ϵ calculus. Theorem 3.13 then tells us that DN∃ is a IPL+P∃ theorem. In IPL+(ϵ) the schema $\neg\exists x\neg\varphi \leftrightarrow \neg\neg\varphi(\epsilon x : \neg\varphi)$ is derivable. It results from the general schema

$$\vdash_{IPL+\epsilon} \exists x\psi \leftrightarrow \psi[\epsilon x : \psi/x]$$

by contraposition, with $\psi = \neg\varphi$. Contraposing once more gives $\neg\neg\exists x\neg\varphi \leftrightarrow \neg\neg\neg\varphi[\epsilon x\neg\varphi/x]$. But $\vdash_{IPL} \neg\varphi \leftrightarrow \neg\neg\neg\varphi$ and so, DN∃ follows. ⊠

This principle straightforwardly gives us *EN*

$$\neg\neg\exists x\varphi \leftrightarrow \exists x\neg\neg\varphi. \qquad \text{(EN)}$$

For,

$$\vdash_{IPL} \neg\exists x\varphi \leftrightarrow \neg\exists x\neg\neg\varphi.$$

So we get, by the contraposition of DN∃ and the contraposition of the above, that $\vdash_{IPL+\text{P}\exists}$ EN.

Now we shall consider some useful equivalents of this principle.

$$\exists x(\neg\varphi) \leftrightarrow \neg\forall x\neg(\neg\varphi), \qquad \text{(QN)}$$

$$\exists x(\exists x\varphi \rightarrow \neg\neg\varphi(x)), \qquad \text{(PN}\exists\text{)}$$

$$(\neg\exists\neg\varphi(x) \rightarrow \forall x\varphi(x)) \rightarrow (\neg\forall x\varphi(x) \rightarrow \exists x\neg\varphi(x)). \qquad \text{(MA)}$$

3.22. PROPOSITION. $\vdash_{IPL}$ *EN* $\leftrightarrow$ *QN* $\leftrightarrow$ PN∃ $\leftrightarrow$ *MA*.

Proof: See appendix to this chapter. ⊠

By means of Proposition 3.22, we can show the following.

3.23. PROPOSITION.

1. $\vdash_{IPL+PN\exists} \forall x(\varphi \vee \neg\varphi) \rightarrow (\exists x\varphi \vee \forall x\neg\varphi)$. (PM)
2. $\vdash_{IPL+PN\exists} \forall x(\varphi(x) \vee \psi) \rightarrow (\forall x\neg\neg\varphi(x) \vee \psi) \quad x \notin FV(\psi)$. (CDN)
3. $\vdash_{IPL+PN\exists} \forall x(\varphi \vee \neg\varphi) \rightarrow (\neg\neg\exists x\varphi \rightarrow \exists\varphi)$. (M∃)

Proof: See the appendix to this chapter. ⊠

Now we show that the logic $IPL + PN\exists$ is weaker than $IPL + P\exists$.

3.24. PROPOSITION. $\vdash_{IPL+P\exists}$ PN∃.

Proof: First we note that EN is derivable in IPL+P∃. This we have shown in Proposition 3.21. Also, $\vdash_{IPL}$ EN $\leftrightarrow$ PN∃, by Proposition 3.22. ⊠

3.25. PROPOSITION. PN∃ $\nvdash_{IPL}$ P∃.

Proof: Figure 3.4 gives a model forcing $\exists x(\exists xP \rightarrow \neg\neg P)$ in the first node. This node does not force $\exists x(\exists xP \rightarrow P)$. ⊠

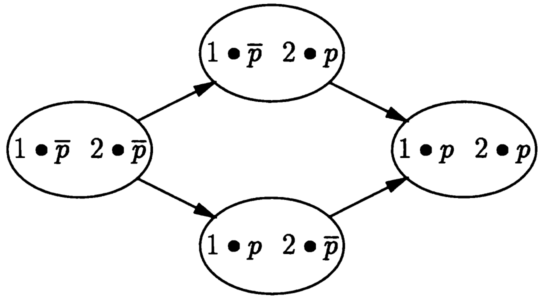

Figure 3.4: PN∃ does not derive P∃

Further Principles

In this section we shall show that IPL+P∃, modulo some conditions discussed below, derives two principles that define well-known classes of Kripke frames, namely, *directedness* and *linearity*. We shall proceed by proving two propositions expressing properties of *disjunction* in IPL+P∃. They will use a device, well-known in *Heyting Arithmetic*, to reduce disjunction to existential quantification (see, for instance, [TD88]). In Heyting Arithmetic, the formula $0 \neq 1$ is taken to be derivable, i.e., $\vdash_{HA} \neg(0 = 1)$ and 0 is taken to be a decidable element, i.e., $\vdash_{HA} \forall x(x = 0 \vee \neg(x = 0))$. Under these circumstances, disjunctions $\varphi \vee \psi$ can be introduced by definition through

$$\exists x((x = 0 \rightarrow \varphi) \wedge (\neg(x = 0) \rightarrow \psi)).$$

With only IPL rules for the existential quantifier, this can be shown to give exactly IPL disjunction. However, in IPL+P∃ the existential quantifier satisfies more properties than it does in IPL. By the definitional schema above, this carries over to disjunction.

3.26. PROPOSITION. *Let* $\Sigma = \{\forall x(x = a \vee \neg(x = a)), \neg(a = b)\}$. *Then*

$$\Sigma \vdash_{IPL+EN} \neg(\varphi \wedge \psi) \rightarrow (\neg\varphi \vee \neg\psi).$$

Proof: Let φ, ψ be arbitrary and set $A(x) = ((x = a \wedge \varphi) \vee (\neg(x = a) \wedge \psi)$. Notice that $\vdash_{IPL} \exists x \neg A(x) \leftrightarrow \exists x((x = a \rightarrow \neg\varphi) \wedge (x \neq a \rightarrow \neg\psi))$. Now we have

1: $\Sigma \vdash A(a) \leftrightarrow \varphi$	*by* $a = a$
2: $\Sigma \vdash A(b) \leftrightarrow \psi$	*by* $a \neq b$
3: $\Sigma, \neg\varphi \vdash \exists x \neg A(x)$	*contrap.1,*($\exists I$)
4: $\Sigma, \neg\psi \vdash \exists x \neg A(x)$	*contrap.2,*($\exists I$)
5: $\Sigma, \neg\exists x \neg A(x) \vdash \neg\neg\varphi \wedge \neg\neg\psi$	*contrapos.3,4,* ($\wedge I$)
6: $\Sigma, \neg\exists x \neg A(x) \vdash \neg\neg(\varphi \wedge \psi)$	*IPL*

7: $\Sigma, \neg(\varphi \wedge \psi) \vdash \neg\neg\exists x \neg A(x)$ *contrap.6,IPL*

8: $\Sigma, \neg(\varphi \wedge \psi) \vdash \exists x \neg A(x)$ *by EN*

9: $\Sigma, \neg(\varphi \wedge \psi) \vdash \neg\varphi \vee \neg\psi.$ $\forall x(x = a \vee \neg(x = a))$

⊠

So under the conditions mentioned in Proposition 3.26, all of De Morgan's laws hold.

3.27. COROLLARY. *Let Σ be as in Proposition 3.26, then*

1. $\Sigma \vdash_{IPL+EN} \neg\varphi \vee \neg\neg\varphi.$ (DIR)
2. $\Sigma \vdash_{IPL+EN} \neg\forall x \neg(\varphi \vee \psi) \rightarrow (\neg\forall x\varphi \vee \neg\forall x \neg\psi).$ (DIS)

Proof: DIR follows from Proposition 3.26 and the IPL theorem $\neg(\varphi \wedge \neg\varphi)$. (DIS) follows from the IPL theorem $\neg\forall x \neg(\varphi \vee \psi) \rightarrow \neg(\forall x \neg\varphi \wedge \forall x \neg\psi)$ and Proposition 3.26. ⊠

For the following proposition we finally need the full strength of P∃. Let LIN be the schema

$$(\xi \rightarrow (\psi \vee \chi)) \rightarrow ((\xi \rightarrow \psi) \vee (\xi \rightarrow \chi)). \tag{LIN}$$

Notice that the schema LIN constitutes a *propositional* version of Plato's principle P∃ (in the form IP∃) if we think of existential quantification as (infinite) disjunction. It is to show this similarity that we have opted for the LIN form of the linearity principle.

3.28. PROPOSITION. *The principle LIN is IPL equivalent to the standard* linearity *schema:*

$$(\varphi \rightarrow \psi) \vee (\psi \rightarrow \varphi) \tag{LIN$'$}$$

Proof: From LIN to LIN': We apply LIN to the IPL theorem $(\varphi \vee \psi) \rightarrow (\varphi \vee \psi)$. This gives us $((\varphi \vee \psi) \rightarrow \varphi) \vee ((\varphi \vee \psi) \rightarrow \psi)$. Now using the IPL theorems $\psi \rightarrow (\varphi \vee \psi)$ and $\varphi \rightarrow (\varphi \vee \psi)$ with the left and right disjunct respectively gives LIN'.
Form LIN' to LIN: Assume $\chi \rightarrow (\varphi \vee \psi)$. The consequent of this assumption with the left disjunct of LIN' (together with $\psi \rightarrow \psi$) gives $\chi \rightarrow \psi$. The consequent of the assumption and the right disjunct of LIN' (together with $\varphi \rightarrow \varphi$) gives $\chi \rightarrow \varphi$. Consequently LIN' derives $((\chi \rightarrow \psi) \vee (\chi \rightarrow \varphi))$ from $\chi \rightarrow (\varphi \vee \psi)$. ⊠

Now we show that the linearity schema is derivable in IPL+P∃ under the same conditions as those of Proposition 3.26.

3.29. PROPOSITION. $\{\exists x(x \neq a), \forall x(x = a \vee x \neq a)\} \vdash_{IPL+P\exists} (\xi \rightarrow (\psi \vee \chi) \rightarrow (\xi \rightarrow \psi) \vee (\xi \rightarrow \chi)).$

Proof: We set

$$A(x) = ((x = a \to \psi) \wedge (x \neq a \to \chi)).$$

We use this formula in two derivations which will be put together for the final result. Both derivations will use the same ($\exists$E) assumption $\exists x A(x) \to A(b)$ of an ($\exists$E) application with major premise $\exists x(\exists x A(x) \to A(x))$. This major premise lies outside both derivations. For the first derivation we use the fact that $\vdash_{IPL} A(a) \leftrightarrow \psi$.

$$\cfrac{\cfrac{\cfrac{\begin{array}{c}\xi\ (0)\\ \vdots\\ \vdots\\ \vdots\\ \psi\vee\chi\end{array} \quad \cfrac{\cfrac{\cfrac{\cfrac{[\psi](1)}{A(a)}}{\exists x A(x)} \quad \exists x A(x)\to A(b)\ (4)}{A(b)} \quad [b\neq a](3)}{\chi} \quad [\chi](2)}{\chi\ (-1,2)}}{\xi\to\chi\ (-0)}}{(b\neq a)\to(\xi\to\chi)\ (-3)}$$

Let's call this derivation I. It derives $b \neq a \to (\xi \to \chi)$ from assumption (4): $\exists x A(x) \to A(b)$. Here ($\vee$E) is closed while we still have an assumption, (3), in the proper term b of the global ($\exists$E) application. This assumption is only discharged after the ($\vee$E) closure.
For the second derivation, note that $\{\exists x(x \neq a), \chi\} \vdash_{IPL} \exists x A(x)$. This gives

$$\cfrac{\cfrac{\cfrac{\begin{array}{c}\xi\ (0)\\ \vdots\\ \vdots\\ \psi\vee\chi\end{array} \quad \cfrac{\cfrac{\cfrac{[\chi](1), \exists x(x\neq a)}{\exists x A(x)} \quad \exists x A(x)\to A(b)\ (4)}{A(b)} \quad [b = a](3)}{\psi} \quad [\psi](2)}{\psi\ (-1,2)}}{\xi\to\psi\ (-0)}}{(b = a)\to(\xi\to\psi)\ (-3)}$$

This we call derivation II. It derives $(b = a) \to (\xi \to \psi)$, assumption (4), and $\exists x(x \neq a)$. Putting I and II together, we get

$$\cfrac{\exists x(\exists x A(x)\to A(x)) \quad \cfrac{\cfrac{\cfrac{\exists x(x\neq a), [\exists x A(x)\to A(b)](4)}{I, II}}{(b = a\to(\xi\to\psi))\wedge(b\neq a\to(\xi\to\chi))}}{\exists x((x = a\to(\xi\to\psi))\wedge(x\neq a\to(\xi\to\chi)))}}{\exists x((x = a\to(\xi\to\psi))\wedge(x\neq a\to(((\xi\to\chi))\ (-4)}$$

But then $\exists x(x \neq a), \forall x(x = a \vee x \neq a) \vdash_{IPL+P\exists} \forall x((\xi \to \psi) \vee (\xi \to \chi))$, and because x does not occur in ξ, ψ or χ, we have $\exists x(x \neq a), \forall x(x = a \vee x \neq a) \vdash_{IPL+P\exists} (\xi \to \psi) \vee (\xi \to \chi)$. ⊠

Definability

The last section has left us with a plethora of IPL invalid principles derivable in IPL+P∃. This section will investigate the model theoretic counterpart. We shall determine the class of Kripke frames defined by this logic as well as some other intermediate logics.

We shall highlight some special classes of Kripke frames which will be seen to play a role in the models for IPL+P∃. These classes will be arranged according to the restrictions they put on the domain and on the structure of the accessibility relation.

3.30. DEFINITION. (Domain Principles) A Kripke frame $F = \langle W, \leq, \mathcal{O}\rangle$ has *constant domain* if $\mathcal{O}(\alpha) = \mathcal{O}(\nu)$ for every α, $\nu \in W$. Let A_c be the class of all Kripke frames with constant domain. Let A_1 be the class of Kripke frames with *singleton* domain, and A_f be the class of all Kripke frames with *finite* domain.

3.31. DEFINITION. (Structural Principles) A Kripke frame $F = \langle W \leq, \mathcal{O}\rangle$ has a *top* if there is an $\alpha \in W$ such that $\beta \leq \alpha$ for every $\beta \in W$. Let A_t be the class of all Kripke frames with a top.
A Kripke frame $F = \langle W \leq, \mathcal{O}\rangle$ is *well-ordered* if for every subset of $B \subseteq W$ there is an $\alpha \in B$ such that $\alpha \leq \beta$ for all $\beta \in B$. Let A_w be the class of all well-ordered Kripke frames.
A Kripke frame F is *linearly ordered* if for all $\alpha, \beta, \gamma \in W$ if $\alpha \leq \beta$ and $\alpha \leq \gamma$, then $\beta \leq \gamma$ or $\gamma \leq \beta$. Let A_l be the class of all linearly ordered Kripke frames.

We have seen that IPL+P∃ is a proper extension of IPL. The characteristic principle of the unrestricted ϵ-calculus,

$$\exists x(\exists x\varphi \rightarrow \varphi),$$

determines a class of Kripke frames, independent of any ϵ-interpretation, and the question is what this class looks like. We shall show that P∃ defines the class of frames with constant domain which either have a singleton domain or are linearly ordered with a finite domain or are well-ordered.

3.32. THEOREM. *The schema P∃ defines the class* $A_{P\exists} = A_c \cap (A_1 \cup (A_l \cap A_f) \cup A_w)$.

Proof: We first show that in every element of the class $A_{P\exists}$ Plato's principle is frame valid. That is,

$$F \in A_c \cap A_1 \text{ or } F \in A_c \cap A_w \text{ or } F \in A_c \cap A_l \cap A_f \Rightarrow F \Vdash \exists x(\exists x\varphi \rightarrow \varphi).$$

1. Suppose $F \in A_c \cap A_1$. It is evident that (P∃) holds on any Kripke model over F at every node.

2. Suppose $K \in A_c \cap A_w$, that is, K has constant domain and is well-ordered. Assume $K, s, \alpha \not\Vdash \exists x(\exists x\varphi \rightarrow \varphi)$ for some node α. We are going to arrive at a contradiction. Because $K, s, \alpha \not\Vdash \exists x(\exists x\varphi \rightarrow \varphi)$, we have for all $m \in dom(D(\alpha))$ a $\beta \geq \alpha$ such that $K, s, \beta \Vdash \exists x\varphi$ and $K, s(x|m), \beta \not\Vdash \varphi$. Because $K, s, \beta \Vdash \exists x\varphi$, there is an element n in $dom(D(\beta))$ such that

$$K, s(x|n), \beta \Vdash \varphi.$$

But, by constant domain, $n \in dom(D(\alpha))$ and

$$K, s(x|n), \alpha \not\Vdash \varphi.$$

Because P$\exists$ is not forced at α, there must be γ between α and β such that $K, s, \gamma \Vdash \exists x\varphi$ and $K, s(x|n), \gamma \not\Vdash \varphi$. So again, at γ we can find an element l, that statisfies φ at γ, but does not satisfy φ at some node δ between α and γ such that $K, s, \delta \Vdash \exists x\varphi$. This procedure can only be continued given an infinitely descending chain in K. But K is well-ordered. A contradiction. Consequently, it cannot be that $K, s, \alpha \not\Vdash \exists x(\exists x\varphi \rightarrow \varphi)$ for some $\alpha \in W$.
3. Suppose $F \in A_c \cap A_f \cap A_f$ and $K, s, \alpha \not\Vdash \exists x(\exists x\varphi \rightarrow \varphi)$ for some model K over F and some $\alpha \in W$. The proof that this leads to a contradiction proceeds exactly as above. In this case there may be an infinitely descending chain of nodes in a model K over F, but, because the common domain is finite, we cannot find the required infinite set of domain elements.

To show

$$F \Vdash \exists x(\exists x\varphi \rightarrow \varphi) \Rightarrow F \in A_c \cap A_1 \text{ or } F \in A_c \cap A_w \text{or } F \in A_c \cap A_l \cap A_f,$$

we proceed by contraposition. Under the assumption that F has nodes with more than one element in the domain, i.e., $F \notin A_1$. We shall show that if $F \notin A_c \cap A_w$ and $F \notin A_c \cap A_l \cap A_f$, then $F \not\Vdash \exists x(\exists x\varphi \rightarrow \varphi)$, i.e., there is model K over F such that $K \not\Vdash \exists x(\exists x\varphi \rightarrow \varphi)$. We first show that Plato's principle is not valid on F if $F \notin A_c$.

1. Let F have domains of varying cardinality: $F \notin A_c$. That is, there are nodes α, $\beta \in W$: $\alpha \leq \beta$, and a domain element m such that $m \in dom(D(\beta)) - dom(D(\alpha))$. Now, let P be a monadic predicate and fix a model K over F with an interpretation Int_β such that $Int_\beta(P) = \{m\}$. In this case $K, \alpha \not\Vdash \exists x(\exists x P(x) \rightarrow P(x))$ because for every element n in $dom(D(\alpha))$ there is a $\gamma \geq \alpha$ (namely β) such $K, \gamma \Vdash \exists x P$ and $K, \gamma \not\Vdash P(n)$. Consequently, $F \not\Vdash P\exists$. This situation is pictured in Figure 3.5.
2. Suppose F has constant domain, but is not well-ordered: $F \notin A_c \cap A_w$. This means that either F is not linearly ordered, or that F is linearly ordered but has an infinitely descending $\leq$-chain.

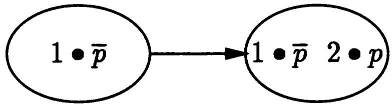

Figure 3.5: Non Constant Domain falsifies P∃

(a) Suppose F is not linearly ordered: $F \notin A_l \cap A_f$. So there are nodes α, β, γ in W such that $\alpha \leq \beta$ and $\alpha \leq \gamma$, but $\beta \not\leq \gamma$ and $\gamma \not\leq \beta$. Because F has constant domain with at least two elements, we can fix two subsets A, B of $dom(D(\alpha))$ such that $A \not\subseteq B$ and $B \not\subseteq A$. Now fix a model K over F, with the property $Int_\beta(P) = A - B$ and $Int_\gamma(P) = B - A$. So both $K, \beta \Vdash \exists x P$ and $K, \gamma \Vdash \exists x P$. But, because the intersection of $Int_\beta(P)$ and $Int_\gamma(P)$ is empty, there is no $m \in dom(D(\alpha))$ such that $K, \alpha \Vdash \exists x P \rightarrow P(m)$. Consequently, $K, \alpha \nVdash \exists x(\exists x P(x) \rightarrow P(x))$ and $F \nVdash P\exists$. This situation is graphically represented in Figure 3.6.

(b) Now suppose the frame F is linearly ordered with constant domain, but not well-ordered. So W is an infinite set and there must be a set of nodes $B \subseteq W$ without an $\leq$ minimum. By linearity, this must be an infinite set. Now suppose F has an infinite domain, that is, $F \notin A_l \cap A_f$. Now define a model K over F with an interpretation function Int such that for all $\beta \in B : Int_\beta(P) \neq \emptyset$ and $\bigcap_{\beta \in B} Int_\beta(P) = \emptyset$. That is, along the descending $\leq$ chain in B we take an ever decreasing non–empty subset of the domain as the interpretation of P. Now let $\alpha \in M - B$ be a node such that $\alpha \leq \beta$ and $\beta \neq \alpha$ implies $\beta \in B$. We then have $K, \alpha \nVdash \exists x(\exists x P(x) \rightarrow P(x))$, for all successors β of α force $\exists x \varphi$ but there is no element $m \in dom(D((\alpha))$ such that $\varphi(m)$

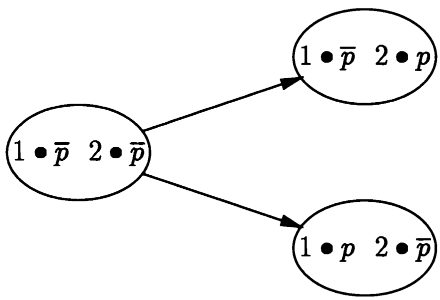

Figure 3.6: Non linearity falsifies P∃

is forced by all elements of B. Consequently $F \not\Vdash P\exists$. ⊠

Now we shall show that PN$\exists$ defines the class of frames with constant domain which have either a singleton domain, or are linearly ordered, or have a top.

3.33. PROPOSITION. *The schema* PN$\exists$, *that is,* $\exists x(\exists x\varphi \to \neg\neg\varphi)$ *defines the class*

$$A_{PN\exists} = A_c \cap (A_1 \cup A_t \cup A_l).$$

Proof: To show for a given frame F that

$$F \Vdash \text{PN}\exists \Rightarrow F \in A_c \cap (A_1 \cup A_t \cup A_l)$$

is left to the reader. Here we show the converse. Let's assume that the union of all domains of F has more than one element. We have to show that $F \notin A_c \cap A_l$ and $F \notin A_c \cap A_t$ implies $F \not\Vdash \exists x(\exists x\varphi \to \neg\neg\varphi)$. Suppose $F \in A_l \cup A_t$. The proof that $F \not\Vdash$PN$\exists$ if $F \notin A_c$ proceeds by exactly the same countermodel as the one in Theorem 3.32. Now suppose that $F \in A_c$ and $F \notin A_l \cup A_t$. So there are $\alpha, \beta, \gamma \in W$ such that $\alpha \leq \beta, \alpha \leq \gamma$ and there are two $\leq$-chains starting from β and γ which have no nodes in common. Because β and γ are unrelated, and the constant domain has at least two elements, we can define two disjoint subsets C, D over the domain, and define a model K over F with an interpretation Int of a monadic predicate letter P such that $Int_\delta(P) = C$ if $\beta \leq \delta$ and $Int_\delta(P) = D$ if $\gamma \leq \delta$. Thus $K, s(x|m)\beta \Vdash \exists x\varphi \to \neg\neg\varphi$ and $K, s(x|n)\gamma \Vdash \exists x\varphi \to \neg\neg\varphi$ for different m, n. But $K, \alpha \not\Vdash \exists x(\exists x\varphi \to \neg\neg\varphi)$. ⊠

Incompleteness of IPL+P$\exists$

In this section we shall show frame incompleteness for the logics $IPL + P\exists$ and $IPL + PN\exists$. Theorem 3.32 shows that Plato's principle defines the class of frames with constant domain which have either a singleton element or are well-ordered, or are linear with finite domain. Various of these classes of frames have been investigated with respect to definability and completeness. Here are some relevant principles.

$$\forall v(\varphi \vee \psi) \to (\varphi \vee \forall v\psi) \qquad v \notin FV(\varphi) \tag{CD}$$

$$\neg\varphi \vee \neg\neg\varphi \tag{DIR}$$

$$(\varphi \to \psi) \vee (\psi \to \varphi) \tag{LIN}$$

The following are well-known results.

3.34. PROPOSITION. (Completeness)

1. *The principle CD is complete for the class A_c.*
2. *The principle DIR is complete for the class A_t.*
3. *The principle LIN is complete for the class A_l.*

Discussion and proofs of these facts can be found in Van Dalen [Dal86] and Gabbay [Gab81].

In Propositions 3.26 and 3.29 we have seen that the principles of directedness and linearity are derivable in IPL+PN$\exists$ and IPL+$P\exists$ respectively, given the presence of a decidable element and two provably different elements. These conditions on the general derivability are understandable if we consider the definability results. On frames with singleton domain no structural constraints are put on $\leq$. The question of *completeness* of the logic IPL+P$\exists$ with respect to the class of frames it defines comes down to the question whether all formulas that are *valid on that class of frames* are indeed *derivable* in the logic. For the logic IPL+P$\exists$ we have the following answer.

3.35. THEOREM. *The logics IPL+P$\exists$ and IPL+PN$\exists$ are not complete for the class of frames they define.*

Proof: By algebraic means Umezawa (see [Ume59]) has shown that the constant domain principle, $\forall x(\varphi(x) \vee \psi) \rightarrow \forall x \varphi(x) \vee \psi$, cannot be derived from the schema

$$\exists x(\exists y \varphi \rightarrow \psi) \vee \exists y(\exists x \psi \rightarrow \varphi). \tag{P$\exists$+}$$

Plato's principle, P$\exists$, is a special instance of this schema (take $\varphi = \psi$). So CD cannot be derived from a schema stronger than P$\exists$. A fortiori, it cannot be derived by P$\exists$ or PN$\exists$. ⊠

It should be noted, that, given *linearity*, we can derive P$\exists$+ from P$\exists$. Linearity gives us $(\exists x \varphi \rightarrow \exists x \psi) \vee (\exists x \psi \rightarrow \exists x \varphi)$. The IP$\exists$ rule then gives us P$\exists$+.

3.4.2 The Logic IPL+P$\forall$

The logic IPL+P$\exists$ is incomplete. It is therefore natural to look for a strengthening of this logic which results in a complete logic. The logic IPL+P$\exists$ has been arrived at by analyzing the pure predicate logical substrate of IPL plus Hilberts ϵ-rule. In this section we shall find a complete extension of IPL+P$\exists$ in a second term rule introduced by Hilbert, the τ-rule.

The τ rule

Hilbert used $\tau x : \varphi$ for a term which denotes an arbitrary object which fails to satisfy φ, if there are such objects in the domain, and an arbitrary object *tout court* otherwise. So the term $\tau x : \varphi$ denotes an arbitrary *counterexample* to $\forall x \varphi$, if there is such a thing. Classically this means that $\tau x : \varphi$ is a shorthand for $\epsilon x : \neg\varphi$. In classical logic the epsilon rule then becomes equivalent to the tau rule:

$$\frac{\varphi[\tau v : \varphi / v]}{\forall v \varphi} \tag{τ}$$

But this is not the case in intuitionistic logic, for we do not have the full law of contraposition. Therefore, in the present set-up, the tau rule and the epsilon term rule have to be formulated as independent principles. The τ-rule resembles the ϵ-rule in the following aspect.

3.36. PROPOSITION. *The τ-rule is not conservative over IPL.*

Proof: This time a simple derivation gives the universal counterpart of Plato's principle:

$$\vdash_{IPL+\tau} \exists v(\varphi \to \forall v\varphi) \tag{P$\forall$}$$

which is not an IPL theorem. ⊠

So IPL+τ extends IPL by a fresh quantifier principle. P$\forall$ is equivalent to the $IP\forall$ rule

$$\frac{\forall v\varphi \to \psi}{\exists v(\varphi \to \psi)} \tag{IP$\forall$}$$

And, as will be shown, the addition of both P$\exists$ and P$\forall$ to IPL gives us a complete logic. We can prove a result corresponding to Theorem 3.13 for the τ–rule showing that P$\forall$ covers exactly the increase in strength resulting from the τ–rule.

3.37. THEOREM. *Let $\Sigma \cup \{\chi\}$ be a set of τ-free formulas. Then*

$$\Sigma \vdash_{IPL+\tau} \chi \iff \Sigma \vdash_{IPL+P\forall} \chi.$$

Proof: The proof of this theorem proceeds completely analogously to that of Theorem 3.13. Our starting position is a τ–derivation with $m+1$ τ–terms.

$$\begin{array}{c} \Sigma \\ \mathcal{D}_1 \\ \Gamma \ \dfrac{\varphi[\tau x : /\varphi]}{\forall x\varphi} \\ \mathcal{D}_2 \\ \psi \\ \vdots \\ \chi \end{array}$$

Again ψ is the first formula below $\forall x\varphi$ without $\tau x : \varphi$ and such that $\tau x : \varphi$ does not occur in any assumption on which ψ depends. The transformation to a derivation with only m τ–terms proceeds now by the P$\forall$ principle:

$$\begin{array}{c} \Gamma \\ \mathcal{D}_2 \\ \psi \end{array} \quad \Rightarrow \quad \dfrac{\exists x(\varphi(x) \to \forall x\varphi) \qquad \begin{array}{c} \varphi[a/x] \to \forall x\varphi \\ \mathcal{D}_2[a/\epsilon x : \varphi] \\ \psi \end{array}}{\psi}$$

3.38. COROLLARY. *Let C be a logical system containing the rules* ($\exists$E),($\exists$I), ($\rightarrow$I), ($\rightarrow$E). *Then the following are equivalent:*

1. *the τ-rule is conservative over C,*
2. *IP$\forall$ is a (derived) rule of C,*
3. *P$\forall$ is a theorem of C.*

Analogously to the case for IPL+P$\exists$ and ϵ-terms, every Kripke model validating the P$\forall$ principle can be supplied with an interpretation for τ-terms.

Proof Theory

We shall only mention the two most relevant theorems of IPL+P$\forall$ calculus.

3.39. PROPOSITION. *The principle DN$\forall$ is IPL+P$\forall$ derivable.*

$$\forall v \neg\neg\varphi \leftrightarrow \neg\neg\forall v\varphi. \tag{DN$\forall$}$$

Proof: This we show analogously to Proposition 3.21: we have

$$\vdash_{IPL+\tau} \neg\neg\forall x\neg\varphi \leftrightarrow \forall x\neg\varphi,$$

by contraposing the principle IPL+τ equivalence $\forall x\psi \leftrightarrow \psi[\tau x\psi/x]$ twice, with $\psi = \neg\varphi$. From this and the IPL theorem $\neg\neg\forall x\varphi \rightarrow \forall x\neg\neg\varphi$ we can derive DN$\forall$. ⊠

Again the principle DN$\forall$ is essentially weaker than P$\forall$.

3.40. PROPOSITION. $\vdash_{IPL+P\forall} \neg\forall x\varphi \rightarrow \exists x\neg\varphi$.

Proof: We may as well use the τ rule to show this and appeal to Theorem 3.37. By $\varphi[\tau x\varphi] \rightarrow \forall x\varphi$ we get by contraposition $\neg\forall x\varphi \rightarrow \neg\varphi[\tau x\varphi/x]$. This gives the desired result by existentially generalizing the consequent of the implication. ⊠

Consequently, IPL+P$\forall$ has all the classical interaction principles of quantifiers and negation.

Definability

We shall state a definability result for IPL+P$\forall$.

3.41. THEOREM. *IPL+ P$\forall$ defines the class $A_{P\forall}$ of frames with constant domain that either have singleton domain, or are conversely well-ordered, i.e., for every set B of nodes there is an $\alpha \in B$ such that $\beta \leq \alpha$ for all $\beta \in B$, or have a finite domain and are linearly ordered.*

Proof: First, if a Kripke frame F has constant domain and is conversely well-ordered, then P$\forall$ is valid. For suppose that $K, \alpha \not\Vdash \exists x(\varphi \rightarrow \forall x\varphi)$ were the case at some node α in a model K over F of the given class. That is, for all $m \in$

$dom(D(\alpha))$ there is a $\beta \geq \alpha : \beta \Vdash \varphi(m)$ and $K, \beta \nVdash \forall x\varphi$. Now, the set $X = \{\beta \mid K, \beta \nVdash \forall x\varphi\}$ has a maximum, call it t. Because $K, t \nVdash \forall x\varphi$ and $dom(D(t))$ has all the elements, there must be a domain element m at t such that $K, t \nVdash \varphi(m)$. But by assumption, for every m there is a node $\beta \geq \alpha$ such that $\beta \Vdash \varphi(m)$ and $K, \beta \nVdash \forall x\varphi$, i.e., $\beta \in X$, and so $\varphi(m)$ must hold at a node in X, and consequently at t. A contradiction.

And suppose a Kripke frame F has constant finite domain and is linear, but $K, \alpha \nVdash \exists x(\varphi \to \forall x\varphi)$ for some α and model K over F. That is, for all $m \in dom(D(\alpha))$ there is a $\beta \geq \alpha : K, \beta \Vdash \varphi(m)$ and $K, \beta \nVdash \forall x\varphi$. Now, the set $X = \{\beta \mid \exists m : K, \beta \Vdash \varphi(m) \;\&\; K, \beta \nVdash \forall x\varphi\}$ is linearly ordered. Because there are only a finite number of elements m in the global domain, there must be a γ in X, such $K, \gamma \Vdash \varphi(m)$ for all elements of the domain, consequently $K, \gamma \Vdash \forall x\varphi$. A contradiction.

For the converse, if F does not have constant domain, then the model of Figure 3.1 falsifies $\exists x(P(x) \to \forall x P(x))$ in the bottom node. So, suppose F has constant domain, but is not conversely well-ordered. That is, F has a set of nodes B such that for all nodes $\alpha \in B$ there is a node $\beta \in B$ such that $\alpha \not\geq \beta$. Consider such a set B: choose a node $\alpha \in B$ and a $\beta \in B$ such that $\alpha \not\geq \beta$. If $\alpha \not\leq \beta$, then we use α and β for the construction that follows; if $\alpha \leq \beta$, then we choose a $\gamma \in B$ such that $\beta \not\geq \gamma$. Again we stop if β and γ are unrelated and continue with γ if $\beta \leq \gamma$. In this way we either get two unrelated nodes in B or an infinite $\leq$ chain in B without a maximal element. In the case of two unrelated nodes α and β, we divide the common domain D in non–empty sets C and $D - C$ and let $Int_\alpha(P) = C$, $Int_\beta(P) = D - C$ for a monadic predicate P. Then on any node $\gamma \leq \alpha$, $\gamma \leq \beta$, P$\forall$ is falsified. Because for every element $m \in dom(D(\gamma))$ we can choose a successor such that $\varphi(m)$ holds there but $\forall x\varphi$ does not.

In case of an infinite $\leq$ chain, we can use the model of of Proposition 3.20 falsifying $\neg\forall x \to \exists x\neg\varphi$. Because this is a consequence of P$\forall$, this schema cannot hold.

Finally, suppose F has constant domain, is not conversely well-ordered. If F is linear but does not have a finite domain, then again the example of Proposition 3.20 can be used to show that P$\forall$ is not valid on F. If F has finite domain but is not linearly ordered, then we can again find two unrelated nodes in W over which we can define a valuation falsifying P$\forall$ as above. ⊠

Incompleteness

As is the case for the logic IPL+P$\exists$, the logic IPL+P$\forall$ is not canonical.

3.42. THEOREM. *The logics IPL+P$\forall$ and IPL+DN$\forall$ are not complete for the class of frames they define.*

Proof: Umezawa (see [Ume59]) has shown that the constant domain principle, $\forall x(\varphi(x) \vee \psi) \to \forall x\varphi(x) \vee \psi$, cannot be derived from the schema

$$\exists x(\varphi \to \forall y\psi) \vee \exists y(\psi \to \forall x\varphi). \tag{P$\forall$+}$$

P$\forall$ is a special instance of this schema (take $\varphi = \psi$). So CD cannot be derived from a schema stronger than P$\forall$. A fortiori, it cannot be derived by P$\forall$ or DN$\forall$. ⊠

3.4.3 The Logic IPL+P$\exists$+P$\forall$

The logic IPL+P$\forall$ is not contained in IPL+P$\exists$. This follows from Proposition 3.40 and Proposition 3.20. Conversely, the logic IPL+P$\forall$ is not contained in IPL+P$\exists$. This we show by constructing a model over a conversely well-founded Kripke frame with constant, non-singleton domain, with a node not forcing P$\exists$. This can be done analogously to the counterexamples constructed in the proof of Theorem 3.32. By Theorem 3.41 such a model will validate P$\forall$. So if we consider the logic IPL+P$\exists$+P$\forall$, we get a true extension of either of the component logics. An important consequence is: the *constant domain* principle is derivable.

3.43. PROPOSITION. *Given the conditions of Proposition 3.29 it holds that*

$$\vdash_{IPL+P\exists+P\forall} \forall x(\varphi(x) \vee \psi) \to (\forall x\varphi(x) \vee \psi).$$

Proof: Under the conditions of Proposition 3.29 we have $(\varphi \to \psi) \vee (\psi \to \varphi)$ in IPL+P$\exists$, an alternative formulation of linearity we have shown to be equivalent to LIN (Proposition 3.28). This we use in the form $(\forall x\varphi \to \psi) \vee (\psi \to \forall x\varphi)$. By $IP\forall$, a consequence of P$\forall$, and IPL this gives

$$\exists x(\varphi \to \psi) \vee \forall x(\psi \to \varphi). \tag{$*$}$$

Now we show

$$\vdash_{IPL} \exists x(\varphi \to \psi) \vee \forall x(\psi \to \varphi) \to \text{CD}.$$

Assume $(*)$ and the antecedent of CD, $\forall x(\varphi(x) \vee \psi)$. By ($\forall$E) this gives $\varphi(c) \vee \psi$ for some c. Now we apply ($\vee$E) with (1) $\varphi(c)$ and (2) ψ to both disjuncts of $(*)$, giving

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\forall x(\psi \to \varphi(x))}{\psi \to \varphi(c)} \quad \psi(2)}{\varphi(c)} \qquad \varphi(c)(2)}{\varphi(c)\ (-1,2)}}{\forall x\varphi}}{\forall x\varphi \vee \psi}$$

and

$$\dfrac{\exists x(\varphi \to \psi) \qquad \dfrac{\dfrac{\dfrac{\varphi(c)(1) \qquad \varphi(c) \to \psi\ (3)}{\psi} \qquad \psi(2)}{\psi(-1,2)}}{\forall x\varphi \vee \psi}}{\forall x\varphi \vee \psi\ (-3)}$$

Unlike its component logics, the logic IPL+P∃+P∀ can be shown to be weakly complete for the class $A_{P\forall} \cap A_{P\exists}$ it defines. Throughout we shall assume *decidable equality*:

$$\forall x \forall y(x = y \vee \neg(x = y)). \tag{DE}$$

We have not been able to do without this assumption. However we conjecture that it must be possible to prove weak completeness without this condition.

3.44. THEOREM. (Weak Completeness of IPL+P∃+P∀)

$$\vdash_{IPL+P\exists+P\forall} \varphi \iff \forall F \in A_{P\forall} \cap A_{P\exists} : F \models \varphi.$$

Proof: For soundness direction (left to right), note that both P∃ and P∀ are $A_{P\forall} \cap A_{P\exists}$ valid.

For completeness, right to left, we shall assume $\nvdash_{IPL+P\exists+P\forall} \varphi$. Now we are going to construct a Kripke model over a frame in the class $A_{P\forall} \cap A_{P\exists}$ with a node not forcing φ. Let a be a constant not occurring in φ. We shall distinguish two cases:

1. $\vdash_{IPL+P\exists+P\forall} \exists x \neg(x = a) \rightarrow \varphi$,
2. $\nvdash_{IPL+P\exists+P\forall} \exists x \neg(x = a) \rightarrow \varphi$.

For the first case we shall construct a model with constant singleton domain containing a node not forcing φ. For the second case we shall construct a finite linear Kripke model with constant domain containing such a node.

Case 1: $\vdash_{IPL+P\exists+P\forall} \exists x \neg(x = a) \rightarrow \varphi$ We first note that $\vdash_{IPL+P\exists+P\forall} \exists x \neg(x = a) \leftrightarrow \neg \forall x x = a$ for IPL+P∀ has all the classical interaction principles of quantifiers and negation. So $\vdash_{IPL+P\exists+P\forall} \neg \forall x x = a \rightarrow \varphi$. Because $\nvdash_{IPL+P\exists+P\forall} \varphi$ we have $\nvdash_{IPL+P\exists+P\forall} \neg \forall x x = a \vee \varphi$. So we can construct a prime theory Γ such that $\neg \forall x x = a, \varphi \notin \Gamma$.

Claim: $\Gamma \cup \{\forall x x = a\}$ *is consistent and* $\Gamma \cup \{\forall x x = a\} \nvdash_{IPL+P\exists+P\forall} \varphi$.

Proof: The consistency of $\Gamma \cup \{\forall x x = a\}$ follows immediately from the fact that $\neg \forall x x = a \notin \Gamma$. Now suppose $\Gamma \vdash_{IPL+P\exists+P\forall} \forall x x = a \rightarrow \varphi$. By IP∀ this implies $\Gamma \vdash_{IPL+P\exists+P\forall} \exists x(x = a \rightarrow \varphi)$. So there is constant b such that

$$b = a \rightarrow \varphi \in \Gamma.$$

On the other hand $\vdash_{IPL+P\exists+P\forall} \exists x \neg(x = a) \rightarrow \varphi$. By standard IPL this gives $\vdash_{IPL+P\exists+P\forall} \forall x(\neg(x = a) \rightarrow \varphi)$. Consequently,

$$\neg(b = a) \rightarrow \varphi \in \Gamma.$$

But $\Gamma \vdash_{IPL+P\exists+P\forall} \forall x(x = a \vee \neg(x = a))$, so $\Gamma \vdash_{IPL+P\exists+P\forall} \varphi$. This gives a contradiction with the assumption that $\Gamma \nvdash_{IPL+P\exists+P\forall} \varphi$. ⊠

Now we extend the consistent set $\Gamma \cup \{\forall xx = a\}$ not containing φ to a consistent prime theory Γ' not containing φ. The nodes in our model will consist of all prime extensions of Γ'. All these extensions will contain the formula $\forall xx = a$. So the resulting model K will have a constant singleton domain, i.e., $K \in A_{P\forall} \cap A_{P\exists}$ and node Γ' of this model will not force φ.

Case 2: $\nvdash_{IPL+P\exists+P\forall} \exists x\neg(x = a) \to \varphi$ In this case we can construct a linear model with infinite constant domain that is well-ordered and conversely well-ordered, with a node not forcing φ.

The proof proceeds in two stages. In the first stage we construct the (linear) IPL+P$\exists$+P$\forall$ canonical model K with an initial node Γ not forcing φ. In the second stage, we identify a *finite* linear submodel K_O of K with an initial node Γ_O having the same property. The frame F_O underlying K_O is an element of $A_{P\forall} \cap A_{P\exists}$. Our proof is then finished.

Let $\nvdash_{IPL+P\exists+P\forall} \varphi$ and let C be a countably infinite set of fresh individual constants. Because $\nvdash_{IPL+P\exists+P\forall} \exists x\neg(x = a) \to \varphi$ we can extend $\{\exists x\neg(x = a)\}$ to a P$\exists$+P$\forall$ prime theory Γ in language $\mathcal{L} \cup C$ such that $\Gamma \nVdash \varphi$. Let $W = \{\Gamma' \mid \Gamma' \text{ a } \mathcal{L} \cup C \text{ prime theory}, \Gamma \subseteq \Gamma'\}$. Notice that all theories in W are formulated in the language $\mathcal{L} \cup C$ of Γ. This is enough, for all elements of W have witnesses in C: if $\exists x\varphi \in \Gamma'$ for some $\Gamma' \in W$, then $\varphi[a/x] \in \Gamma'$ for some $a \in C$. This is the case because $\exists x(\exists x\varphi \to \varphi) \in \Gamma$, and so the *witnessing axioms* $\exists x\varphi \to \varphi[a/x]$ are in $\Gamma \subseteq \Gamma'$.

Because we have assumed decidable equality and Γ contains $\exists x\neg(x = a)$, by Proposition 3.29, Γ has all instances of $((\varphi \to \psi) \vee (\psi \to \varphi))$ and so $K = \langle W, \subseteq\rangle$ is linearly ordered (see [Dal86]). Now we set for all closed atomic formulas φ

$$K, \Gamma' \Vdash \varphi \iff \varphi \in \Gamma',$$

and get the forcing lemma.

3.45. LEMMA. (Forcing Lemma) *For all closed formulas φ, $K, \Gamma' \Vdash \varphi$ iff $\varphi \in \Gamma'$.*

Proof: The proof for the propositional connectives is completely standard. For the case of the existential quantifier it is standard to show that $\exists x\varphi \notin \Gamma'$ implies $K, \Gamma' \nVdash \exists x\varphi$. Now suppose $\exists x\varphi \in \Gamma'$. By the presence of the witness axioms in Γ' (and deductive closure) we know that $\varphi[a/x] \in \Gamma'$ for at least one term a. So, by induction hypothesis, $K, \Gamma' \Vdash \varphi[a/x]$, and consequently, $K, \Gamma' \Vdash \exists x\varphi$. Here is where the schema P$\exists$ does its work.

Now the universal quantifier. Here we shall use the schema P$\forall$. Suppose $\forall x\varphi \in \Gamma'$, then $\varphi[a/x] \in \Gamma''$ for all $a \in C$ and all $\Gamma'' \in W$ such that $\Gamma' \subseteq \Gamma''$. By induction hypothesis, for all $a \in C$ and all $\Gamma'' \in W : K, \Gamma'' \Vdash \varphi[a/x]$. Consequently $K, \Gamma' \Vdash \forall x\varphi$.

Now suppose $\forall x\varphi \notin \Gamma'$. We have to show that $K, \Gamma \nVdash \forall x\varphi$. For every universal formula $\forall x\varphi$, every prime theory contains $\varphi[a/x] \to \forall x\varphi$ for some a, by the presence of the schema $\exists x(\varphi \to \forall x\varphi)$. So, if $\forall x\varphi \notin \Gamma'$ then, by deductive closure of Γ', there must be at least one $a \in C$ such that $\varphi[a/x] \notin \Gamma'$. By induction hypothesis, $K, \Gamma' \not\Vdash \varphi[a/x]$ and so $K, \Gamma' \not\Vdash \forall x\varphi$. ⊠

3.46. REMARK. Notice that in the absence of P$\forall$ a constant 'domain' C, determined by the witnesses of the bottom node Γ, is not sufficient. If $\forall x\varphi \notin \Gamma'$ we may have to introduce a *new* constant a, i.e., $a \notin C$ to get a $\varphi[a/x]$ not forced by some node accessible to Γ. There is no guarantee that we can find such a 'witness' in C, for the set $\{\neg\forall x\varphi\} \cup \{\varphi[a/x] \mid a \in C\}$ is perfectly consistent in IPL+P$\exists$.

To make $\langle W, \subseteq\rangle$ well-ordered and conversely well-ordered, first a proposition.

3.47. PROPOSITION. *Let $B \subseteq W$, then $\bigcap B$ is a consistent prime $\mathcal{L} \cup C$ theory.*

Proof: The consistency and deductive closure of $\bigcap B$ follow straightforwardly: all elements of B are consistent and deductively closed. $\bigcap B$ is witnessing, because, if $\exists x\varphi \in \bigcap B$, then $\varphi[a/x] \in \bigcap B$ for some fixed a, by the presence of the witnessing axioms in Γ. Finally, $\bigcap B$ splits disjunctions, for if $\varphi \vee \psi \in \bigcap B$ but $\varphi, \psi \notin \bigcap B$, then there must be $\Gamma', \Gamma'' \in B$ such that $\varphi \notin \Gamma'$ and $\psi \notin \Gamma''$. But $\Gamma' \subseteq \Gamma''$ or $\Gamma'' \subseteq \Gamma'$. Suppose the former is the case. Then $\varphi \notin \Gamma'$ and $\varphi \vee \psi \in \Gamma'$ imply $\psi \in \Gamma'$, and so $\psi \in \Gamma''$. This gives a contradiction. The latter case gives a contradiction in the same way. ⊠

Now we construct a special kind of *filtration* of this model. Let O be the smallest set containing φ which is closed under Boolean subformulas such that, if $\exists x\varphi \in O$ ($\forall x\varphi \in O$), then O contains $\varphi[a/x]$ for some selected a such that $\exists x\varphi \to \varphi[a/x] \in \Gamma$ ($\varphi[a/x] \to \forall x\varphi \in \Gamma$). Notice that O is a finite set of formulas.

We set for $\Gamma', \Gamma'' \in W$:

$$\Gamma' \sim_O \Gamma'' \text{ if and only if } \forall\varphi \in O(\varphi \in \Gamma' \iff \varphi \in \Gamma'').$$

Now, in a standard filtration, the filtrate is constructed from $\sim_O$ equivalence classes of the elements of W. In the logic IPL+$P\forall$ + $P\exists$ we can go further, for, by Proposition 3.47 the intersection of such an equivalence class is itself a prime theory. We shall work with these intersections of equivalence classes. For $\Gamma' \in W$, we set $\Gamma'_O = \bigcap\{\Gamma'' \in W \mid \Gamma' \sim_O \Gamma''\}$.

Now we define $K_O = \langle W_O, \leq_O, D_O\rangle$, where

1. $W_O = \{\Gamma'_O \mid \Gamma \in W\}$,
2. $\leq_O = \leq \restriction (W_O \times W_O)$,
3. $D_O = D \restriction \mathcal{L}_O$.

Notice that $F_O = \langle W_O, \leq_O\rangle$ is a *finite* linear frame, so $F_O \in A_{P\exists} \cap A_{P\forall}$. Moreover, if $\Gamma' \leq \Gamma''$ then $\Gamma'_O \leq_O \Gamma''_O$ and if $\Gamma'_O \leq_O \Gamma''_O$, then there is a $\Gamma''' \sim_O \Gamma''$

such that $\Gamma' \leq \Gamma'''$. It remains to be shown that K_O is a Kripke model with a node Γ'_O not forcing φ.

3.48. PROPOSITION. *If we set for all closed* atomic *formulas* φ, $K_O, \Gamma'_O \Vdash \varphi \iff \varphi \in \Gamma'_O$, *then for all closed formulas* $\varphi \in O$,

$$K_O, \Gamma'_O \Vdash \varphi \iff \varphi \in \Gamma'_O.$$

Proof: The proof for the propositional connectives proceeds standardly. As an example, we shall treat implication.
Suppose $K_O, \Gamma_O \not\Vdash \varphi \rightarrow \psi$, for $\varphi \rightarrow \psi \in O$. So there is a $\Gamma'_O \geq \Gamma_O : K_O, \Gamma'_O \Vdash \varphi$ and $K_O, \Gamma'_O \not\Vdash \psi$. Because $\varphi, \psi \in O$, we have, by induction hypothesis, $\varphi \in \Gamma'_O$ and $\psi \notin \Gamma'_O$. But $\Gamma_O \subseteq \Gamma'_O$. So $\varphi \rightarrow \psi \notin \Gamma_O$.
Now suppose $\varphi \rightarrow \psi \notin \Gamma_O$. So there is a $\Gamma' \in W$ such that $\Gamma \subseteq \Gamma'$ and $\varphi \in \Gamma'$, $\psi \notin \Gamma'$. But $\varphi, \psi \in O$. So $\varphi \in \Gamma'_O$ and $\psi \notin \Gamma'_O$. By induction hypothesis, $K_O, \Gamma'_O \Vdash \varphi$, $K_O, \Gamma'_O \not\Vdash \psi$. But $\Gamma_O \leq_O \Gamma'_O$, so $K_O, \Gamma'_O \not\Vdash \varphi \rightarrow \psi$.
The cases for the quantifier follow the proof of Lemma 3.45 exactly. Here the fact that O contains $\exists x(\exists x\varphi \rightarrow \varphi)$- and $\exists x(\varphi \rightarrow \forall x\varphi)$-witnesses for $\exists x\varphi \in O$ and $\forall x\varphi \in O$ respectively, carries the induction hypothesis through. ⊠

3.49. COROLLARY. *For all closed formulas* $\varphi \in O$

$$K, \Gamma' \Vdash \varphi \iff K_O, \Gamma'_O \Vdash \varphi.$$

Proof: For all closed formulas φ in O and all $\Gamma \in W$ we have $\varphi \in \Gamma \iff \varphi \in \Gamma_O$. Proposition 3.48 and Lemma 3.45 then give the desired result. ⊠

Because $K, \Gamma \not\Vdash \varphi$ and $\varphi \in O$ we have the immediate consequence that $K_O, \Gamma_O \not\Vdash \varphi$. So we have created a model over a frame in $A_{P\forall} \cap A_{P\exists}$ with a node not forcing φ. ⊠

3.5 Conclusion

We have seen that the epsilon rule is not conservative over intuitionistic logic. Our investigations into this phenomenon has given us a number of intersting intermediate logics. The epsilon rule is conservative over the logic IPL+P∃, and the addition of Skolem functions over IPL+P∃ is conservative. However, this logic is frame incomplete. By weakening the principle P∃ to PN∃ we reached another incomplete intermediate logic. By considering logics arising from the addition of the τ-rule to IPL we have found the two universal counterparts P∀ and PN∀ of the logics above. These were also shown to be incomplete. Only by addition of both the ϵ-rule and the τ-rule, or alternatively, both the principles P∃ and P∀, did we achieve a complete logic.

A question, to be answered in the next chapter, is the following: how can we restrict the epsilon rule in such a way that we get a conservative extension over intuitionistic logic? This question belongs in the next chapter because it involves an awareness of *dependencies* that arise in a proof theoretic contexts. These dependencies can arise between *terms*, between *formulas* and between terms and formulas. In order to define an ϵ-rule that is conservative over IPL we have to investigate dependencies of the third kind: in the ϵ-rule an (existential) *formula* introduces an (epsilon) term. In unrestricted ϵ-derivations certain dependencies between formula and term will be seen to be broken. By respecting the 'relevant' dependencies, a conservative use of the ϵ-rule over IPL will be constructed.

3.6 Appendix

For EN, i.e., the principle $\neg\neg\exists x\varphi \to \exists x\neg\neg\varphi$ and the principles

$$\exists x(\neg\varphi) \leftrightarrow \neg\forall x\neg(\neg\varphi), \tag{QN}$$

$$\exists x(\exists x\varphi \to \neg\neg\varphi(x)), \tag{PN$\exists$}$$

$$(\neg\exists\neg\varphi(x) \to \forall x\varphi(x)) \to (\neg\forall x\varphi(x) \to \exists x\neg\varphi(x)), \tag{MA}$$

we have the propsosition:

3.50. PROPOSITION. $\vdash_{IPL} EN \leftrightarrow QN \leftrightarrow \mathrm{PN}\exists \leftrightarrow MA$.

Proof: For the proof of these equivalences we show EN $\vdash_{IPL}$ QN $\vdash_{IPL}$ PN$\exists$ $\vdash_{IPL}$ MA $\vdash_{IPL}$ EN.
From EN to QN: by $\neg\exists x\varphi \leftrightarrow \forall x\neg\varphi$ we get $\neg\neg\exists x\neg\varphi \leftrightarrow \neg\forall x\neg\neg\varphi$. Now EN (plus $\neg\varphi \leftrightarrow \neg\neg\neg\varphi$) gives the desired result. So, for *negated* φ we have the principle $\neg\forall x\varphi \to \exists x\neg\varphi$.
From QN to PN$\exists$:

$$\begin{array}{c}
\neg\neg\forall x\neg\neg(\exists x\varphi \wedge \neg\varphi) \\ \hline
\forall x\neg\neg(\exists x \wedge \neg\varphi) \\ \hline
\neg\neg(\exists x\varphi \wedge \neg\varphi(a)) \\ \hline
\neg\neg\exists x\varphi \wedge \neg\varphi(a) \\ \hline
\begin{array}{cc}
\vdots & \neg\varphi(a) \\ \hline
\neg\neg\exists x\varphi & \forall x\neg\varphi
\end{array} \\ \hline
\bot \\ \hline
\neg\neg\neg\forall x\neg\neg(\exists x\varphi \wedge \neg\varphi(x)) \\ \hline
\exists x\neg(\exists x\varphi \wedge \neg\varphi(x)) \\ \hline
\exists x(\exists x\varphi \to \neg\neg\varphi(x))
\end{array}$$

From IPN$\exists$ to MA: here we use the fact that $\exists x(\exists x\neg\varphi \to \neg\varphi)$ is an instance of $\exists x(\exists x\varphi \to \neg\neg\varphi)$ (given $\neg\varphi \leftrightarrow \neg\neg\neg\varphi$):

$$\dfrac{\exists x(\exists x\neg\varphi \to \neg\varphi) \qquad \dfrac{\dfrac{\dfrac{\dfrac{\dfrac{[\exists x\neg\varphi \to \neg\varphi(a)]}{\neg\neg\varphi(a) \to \neg\exists\neg\varphi} \quad [\neg\exists x\neg\varphi \to \forall x\varphi]}{\neg\neg\varphi(a) \to \forall x\varphi}}{\neg\forall x\varphi \to \neg\varphi(a)} \quad [\neg\forall x\varphi]}{\neg\varphi(a)}}{\exists x\neg\varphi}}{\dfrac{\exists x\neg\varphi}{\dfrac{\neg\forall x\varphi \to \exists x\neg\varphi}{(\neg\exists\neg\varphi(x) \to \forall x\varphi(x)) \to (\neg\forall x\varphi(x) \to \exists x\neg\varphi(x))}}}$$

From MA to EN: remark first that we can derive $\neg\forall x\neg\varphi$ from $\neg\neg\exists x\neg\neg\varphi$. $\exists x\neg\neg\varphi \to \neg\forall\neg x\varphi$ gives us $\neg\neg(\exists x\neg\neg\varphi \to \neg\forall x\neg\varphi)$ which leads to $\neg\neg\exists x\neg\neg\varphi \to \neg\neg\neg\forall x\neg\varphi$ with the desired result. Now consider

$$\dfrac{\begin{matrix}\neg\neg\exists x\neg\neg\varphi \\ \vdots \\ \vdots \\ \vdots \\ \neg\forall x\neg\varphi\end{matrix} \qquad \dfrac{\dfrac{\dfrac{\dfrac{\neg\neg\exists x\neg\neg\varphi \qquad [\neg\exists x\neg\neg\varphi]}{\bot}}{\forall x\neg\varphi}}{\neg\exists x\neg\neg\varphi \to \forall x\neg\varphi} \qquad \text{MA}}{\neg\forall x\neg\varphi(x) \to \exists x\neg\neg\varphi}}{\exists x\neg\neg\varphi}$$

⊠

3.51. PROPOSITION.

1. $\vdash_{IPL+PN\exists} \forall x(\varphi \vee \neg\varphi) \to (\exists x\varphi \vee \forall x\neg\varphi)$ (PM)
2. $\vdash_{IPL+PN\exists} \forall x(\varphi(x) \vee \psi) \to (\forall x\neg\neg\varphi(x) \vee \psi) \quad x \notin FV(\psi)$ (CDN)
3. $\vdash_{IPL+PN\exists} \forall x(\varphi \vee \neg\varphi) \to (\neg\neg\exists x\varphi \to \exists\varphi)$ (M∃)

Proof: The following IPL+P∃ derivation gives us the first theorem:

$$\dfrac{\exists x(\exists x\varphi \to \varphi) \qquad \dfrac{\dfrac{\forall x(\varphi \vee \neg\varphi)}{\varphi(a) \vee \neg\varphi(a)} \quad \dfrac{\dfrac{[\varphi(a)]}{\exists x\varphi}}{\exists x\varphi \vee \forall x\neg\varphi} \quad \dfrac{\dfrac{[\neg\varphi(a)] \qquad [\exists x\varphi \to \varphi(a)]}{\forall x\neg\varphi}}{\exists x\varphi \vee \forall x\neg\varphi}}{\exists x\varphi \vee \forall x\neg\varphi}}{\exists x\varphi \vee \forall x\neg\varphi}$$

For a derivation of the second theorem, assume $\forall x(\varphi(x) \vee \psi)$. This gives, by ($\forall$E), $\varphi(c) \vee \psi$ for some c. Now we apply the rule ($\vee$E) with assumptions (1) $\varphi(c)$ and (2) ψ:

$$
\begin{prooftree}
\AxiomC{$\exists x(\exists\neg\varphi \to \neg\varphi)$}
\AxiomC{$\varphi(c)\ (1)$}
\AxiomC{$[\neg\varphi(b)]\ (3)$}
\UnaryInfC{$\exists x\neg\varphi$}
\AxiomC{$[\exists x\neg\varphi \to \neg\varphi(c)]\ (4)$}
\BinaryInfC{$\neg\varphi(c)$}
\BinaryInfC{$\bot$}
\UnaryInfC{$\neg\neg\varphi(b)\ (-3)$}
\UnaryInfC{$\forall x\neg\neg\varphi$}
\UnaryInfC{$\forall x\neg\neg\varphi \vee \psi$}
\AxiomC{$\psi\ (2)$}
\UnaryInfC{$\forall x\neg\neg\varphi \vee \psi$}
\BinaryInfC{$\forall x\neg\neg\varphi \vee \psi\ (-1,2)$}
\BinaryInfC{$\forall x\neg\neg\varphi \vee \psi\ (-4)$}
\end{prooftree}
$$

For a proof of the third theorem, note that EN gives us $\forall x(\neg\neg\varphi \to \varphi) \to (\neg\neg\exists x\varphi \to \exists x\varphi)$ from the IPL theorem $\forall x(\neg\neg\varphi \to \varphi) \to (\exists x\neg\neg\varphi \to \exists x\varphi)$. Now $\forall x(\varphi \vee \neg\varphi) \to \forall x(\neg\neg\varphi \to \varphi)$ gives the desired result. ⊠

Chapter 4

Formula Dependencies

This chapter investigates dependencies arising in de course of natural deduction derivations by means of the epsilon calculus. The first section deals with the general phenomenon of dependence between formulas used as assumptions. A fundamental difference between assumption management in CPL and IPL will be observed. Based on the insights gathered, the second section will present the proof theory for a system of instantial logic in which reasoning about instances is maximally constrained, so as to obtain a conservative epsilon extension of intuitionistic logic. In a third section, we consider these issues from a semantic point of view, discussing various calculi that arise from expanding intuitionistic Kripke models with epsilon terms. These admit of Aczel-Thomason style completeness proofs. In our fourth and final section, we explore variants of Kripke semantics that should do the more complicated job of matching up with our original conservative instantial logic.

4.1 Dependence Management in Natural Deduction

The last chapter has left us with system of intuitionistic logic extended by the ϵ-rule, which is not conservative over intuitionistic logic. Indeed, we found a whole landscape of intermediate logics in this vein, that turned out to show some independent interest. Nevertheless, there remains the question what would be a natural conservative epsilon logic over an intuitionistic base. The search for such a system highlights some interesting features of natural deduction, which deserve independent study. In the intermediate logic IPL+ϵ, we have two alternative rules by means of which we can eliminate an existential quantifier: the rule ($\exists$E) and the ϵ-rule. In fact, the ϵ-rule alone will do, for every application of

($\exists$E) with major premise $\exists x\varphi$ and *assumption* $\varphi[a/x]$ in a derivation $\mathcal{D}$ can be replaced by an application of the ϵ-rule with premise $\exists x\varphi$ and *conclusion* $\varphi[\epsilon x : \varphi/x]$. By non-conservativity, the converse does not hold. This situation is caused by the different *dependence structures* of the two rules. In this section, we shall discuss the well-known notion of formulas *depending* on assumptions in a natural deduction framework. Eventually, this will suggest a restriction on derivations which results in a ϵ-calculus conservative over IPL.

There is a further spin-off of these considerations. The notion of dependence does not only apply to formulas in a derivation. The *terms* used in derivation also exhibit a dependence structure. These dependencies are witnessed by quantifier *interaction* principles. Here CPL and IPL do not differ, but the ϵ-calculus suggests various variations of the rules which affect the quantifier interaction principles. This will be the topic of a subsequent chapter.

4.1.1 Dependence on Assumptions

In its most straightforward form a formula occurrence ψ depends, in the natural deduction systems we discussed in chapter 2, intuitively on an assumption φ in a derivation $\mathcal{D}$ if φ is *used* in the derivation of ψ. In this sense, in the derivation

$$\dfrac{\begin{matrix}\varphi\ [i]\\ \vdots\\ \psi\end{matrix}}{\varphi \to \psi}\ {\to}\mathrm{I}_i$$

both the premise of the rule φ and the conclusion $\varphi \to \psi$ depend on the assumption φ. And in the derivation

$$\dfrac{\begin{matrix}\Gamma\\ \vdots\\ \exists x\varphi\end{matrix} \qquad \begin{matrix}\varphi[a/x]\ [i]\\ \vdots\\ \xi\end{matrix}}{\xi}\ \exists\mathrm{E}_i$$

both the conclusion ξ and the minor premise ξ use the assumptions in Γ. But in these derivations we can distinguish essential differences in the status of this dependence for the different formula occurrences. In the application of ($\to$I), the premise ψ depends on an assumption which has not been discharged, while at the conclusion this assumption has been discharged. It is used in the sense that it has been incorporated as a subformula of the conclusion and has disappeared as assumption that can be used in the derivation. In the application of the rule ($\exists$E) we again encounter this difference. But in this case, the assumption that is discharged does not occur as subformula in the conclusion. It is *used* in the derivation of ξ in a fundamentally different way than the assumption of the rule

($\to$I). Notice that, for a given derivation $\mathcal{D}$ it makes sense to speak of a formula depending in this way on an assumption, even though this assumption has been discharged previously in the derivation. In the proof tree constituting $\mathcal{D}$ both discharged and non-discharged assumptions are exposed on the leaves. However, if we want our notion of dependence to be less tied to individual derivations, then we should be interested in a notion of dependence which excludes dependence on discharged assumptions.

In order to define such a notion formally we shall introduce the notion of a *thread*.

4.1. DEFINITION. (Threads and Discharge) Let $\mathcal{D}$ be a derivation. A *thread* in $\mathcal{D}$ is a sequence $\varphi_1 \ldots \varphi_n$ of formula occurrences in $\mathcal{D}$ such that φ_1 is a top-formula in $\mathcal{D}$, φ_i lies immediately above φ_{i+1} for every $i < n$, and φ_n is the conclusion of $\mathcal{D}$. An *assumption occurrence* ψ in $\mathcal{D}$ can be *discharged at formula occurrence* φ if occurrences ψ and φ lie on an thread in $\mathcal{D}$ such that occurrence ψ has not been discharged between occurrences ψ and φ, and

1. if ψ is an ($\exists$E)-assumption of the form $\xi[a/x]$, then the proper term a does not occur in φ or in non discharged assumptions on a thread through φ,
2. if ψ, χ are the assumptions of some application of ($\vee$E), then the formula φ occurs on a thread starting with χ.

An assumption ψ will be called *active at formula occurrence* φ in $\mathcal{D}$, if ψ starts a thread through φ and at no formula occurrence above φ along that thread has ψ been discharged.

4.2. DEFINITION. (Dependence) A formula occurrence φ *depends* on assumption ψ in derivation $\mathcal{D}$, notation $\varphi \prec_{\mathcal{D}} \psi$, if ψ is active at φ in $\mathcal{D}$.

4.1.2 Dependence between Assumptions

From the perspective of dependence it makes sense to take a less structural point of view. Notice that a thread in a derivation contains only one assumption. But this does not exhaust the intuitive formula dependencies that arise in the course of a derivation. We also have dependencies *between assumptions*. For instance, in an application of ($\exists$E) in a derivation $\mathcal{D}$

$$\dfrac{\begin{array}{c}\Gamma\\ \vdots\\ \exists x\varphi\end{array} \qquad \begin{array}{c}\varphi[a/x]\ [i]\\ \vdots\\ \xi\end{array}}{\xi}\ \exists\mathrm{E}_i$$

we have a *main derivation* starting from assumptions in Γ and ending with $\exists x\varphi$, the major premise of the application, and a *side derivation* starting at $\varphi[a/x]$ and ending with ξ. Now, intuitively, the side derivation is subordinate to the main derivation: the *assumption* $\varphi[a/x]$ *depends* on the major premise $\exists x\varphi$, for only in the presence of this formula is the assumption that φ holds for some fresh term a a sound move. By the same argument, the *assumption* $\varphi[a/x]$ intuitively depends on all *assumptions* on which $\exists x\varphi$ depends, that is, on all elements of Γ. But this intuitive dependence of $\varphi[a/x]$ on $\exists x\varphi$ is not witnessed by the structure of the derivation, for there is no thread connecting $\exists x\varphi$ and $\varphi[a/x]$.

The same intuitive dependence occurs in the rule ($\vee$E)

$$\dfrac{\begin{array}{c}\Gamma\\ \vdots\\ \varphi \vee \psi\end{array} \qquad \begin{array}{c}\varphi\ [i]\\ \vdots\\ \xi\end{array} \qquad \begin{array}{c}\psi\ [j]\\ \vdots\\ \xi\end{array}}{\xi}\ \vee\mathrm{E}_{i,j}$$

Again, the assumptions φ and ψ intuitively depend on the major premise $\varphi \vee \psi$, and, as a consequence, on all assumptions in Γ. But there is no thread through $\varphi \vee \psi$ and φ or ψ. To capture the notion of assumptions depending on other assumptions we shall introduce a less structural notion of a thread in a derivation tree.

4.3. DEFINITION. (c-Threads) Let $\mathcal{D}$ be a derivation. A *c-thread* in $\mathcal{D}$ is a sequence $\varphi_1 \dots \varphi_n$ of formula occurrences in $\mathcal{D}$ such that φ_1 is a top-formula in $\mathcal{D}$, and for all $i \leq n$

1. φ_i lies immediately above φ_{i+1}, or
2. φ_i is the major premise of an application of ($\vee$E) or ($\exists$E) and φ_{i+1} is an assumption of that application,

and φ_n is the conclusion of $\mathcal{D}$.

So a c-thread treats the assumptions of ($\exists$E) and ($\vee$E)-applications as if they lie below the major premise of that application.

4.4. DEFINITION. (c-Dependence) Formula occurrence ψ *immediately c-depends* on assumption occurrence φ in derivation $\mathcal{D}$, notation $\psi \ll^c_{\mathcal{D}} \varphi$, if there is a c-thread $\varphi = \varphi_1, \dots, \varphi_n$ in $\mathcal{D}$ such that $\varphi_{i+1} = \psi$ for some $i < n$ and

1. if ψ is an *assumption occurrence*, then φ is active at φ_i;
2. if ψ is not an assumption occurrence, then φ is active at φ_{i+1}.

Formula occurrence ψ *c-depends* on assumption φ in derivation $\mathcal{D}$, notation $\psi \prec^c_{\mathcal{D}} \varphi$, if there is in $\mathcal{D}$ a sequence of immediate c-dependence steps leading from φ to ψ.

Assumption occurrence φ *supports* formula occurrence ψ in $\mathcal{D}$ if $\psi \prec^c_{\mathcal{D}} \varphi$ and either ψ is an assumption or $\psi \not\ll^c_{\mathcal{D}} \varphi$.

By this definition, the assumptions of ($\exists$E) and ($\vee$E)-applications c-depend on the assumptions active at the major premise.

4.5. PROPOSITION. *For any derivation $\mathcal{D}$ the relation $\prec^c_{\mathcal{D}}$ is transitive if and only if for any two assumption occurrences ψ and φ in $\mathcal{D}$ if φ precedes ψ on a c-thread t in $\mathcal{D}$, then the discharge of ψ precedes the discharge of φ on t.*

Proof: By definition, if $\varphi \prec^c_{\mathcal{D}} \psi$ and $\psi \prec^c_{\mathcal{D}} \chi$, then ψ and χ must be different *assumption* occurrences in $\mathcal{D}$. Now, there is a c-thread connecting φ and χ, so $\varphi \prec^c_{\mathcal{D}} \chi$ holds if and only if assumption χ is *active* at φ. By assumption ψ is active at φ. Consequently, $\varphi \prec^c_{\mathcal{D}} \chi$ if φ has not been discharged anywhere between ψ and φ on the c-thread connecting them. ⊠

So transitivity of $\prec^c_{\mathcal{D}}$ implies for assumptions on a c-thread: first in, last out. The set of assumptions along a c-thread behave at any point along that thread as a *stack*.

4.6. DEFINITION. For any formula occurrence φ in derivation $\mathcal{D}$, let the *stackset at* φ be the set $ST_{\mathcal{D}}(\varphi) = \{\psi \mid \varphi \prec^c_{\mathcal{D}} \psi\}$. Formula occurrence ψ is a *top element* in $ST_{\mathcal{D}}(\varphi)$ if $\psi \in ST_{\mathcal{D}}(\varphi)$ and there is no element $\chi \in ST_{\mathcal{D}}(\varphi)$ such that $\chi \prec^c_{\mathcal{D}} \psi$. Formula occurrence ψ is a *bottom element* in $ST_{\mathcal{D}}(\varphi)$ if $\psi \in ST_{\mathcal{D}}(\varphi)$ and there is no element $\chi \in ST_{\mathcal{D}}(\varphi)$ such that $\psi \prec^c_{\mathcal{D}} \chi$.

The set $ST_{\mathcal{D}}(\varphi)$ can be seen as a family of *stacks* of assumptions ordered by immediate c-dependence. Note that, if χ is a top element in $ST_{\mathcal{D}}(\varphi)$ (see Definition 4.2), then $\varphi \prec_{\mathcal{D}} \chi$, i.e., φ and χ are connected by a *thread*. The bottom elements of $ST_{\mathcal{D}}(\varphi)$ are either elements of the set of non-discharged assumptions of $\mathcal{D}$ or assumptions of the rule ($\rightarrow$I). Note that every stack in $ST_{\mathcal{D}}(\varphi)$ can have at most one assumption of ($\rightarrow$I).

The notion of a stackset is relative to that of a c-thread. Consequently, for other notions of a thread, other ways of traversing a proof tree, we get a correspondingly different notion of a stackset.

4.7. DEFINITION. (Stack-Discharge) An assumption occurrence ψ in derivation $\mathcal{D}$ may be *stack-discharged* at formula occurrence φ if ψ is a *top element* in $ST_{\mathcal{D}}(\varphi)$ and

1. if ψ is an ($\exists$E)-assumption of the form $\xi[a/x]$, then the proper term a does not occur in φ or in non discharged assumptions on a thread through φ,
2. if ψ, χ are the assumptions of some application of ($\vee$E), then the formula φ occurs on a thread starting with χ.

Stack-discharge of an assumption ψ at formula occurrence φ, means removal of ψ from $ST_{\mathcal{D}}(\varphi)$. This removal affects only the 'control' structure of the proof, that is, this action is not accompanied by the introduction or elimination of a logical operator. In the rules ($\vee$E) and ($\exists$E) this is witnessed by the fact that the proof transition from minor premise to conclusion does not change the formula involved. In ($\exists$E) the control structure of a derivation is changed at the conclusion of an application in that the proper term of the application is again available for general use.

4.8. PROPOSITION. (Stack Proposition) *For any derivation $\mathcal{D}$, if assumption occurrence φ is discharged at formula occurrence ψ, then φ is stack-discharged at ψ.*

Proof: by Proposition 4.5, we have to show that the relation $\prec^c_\mathcal{D}$ is transitive. So suppose $\varphi \prec^c_\mathcal{D} \psi$ and $\psi \prec^c_\mathcal{D} \chi$. We have to show that assumption χ has not been discharged anywhere between χ and φ along the c-thread connecting them. But the c-dependence of assumption ψ on assumption χ implies that ψ must be an assumption of an ($\exists$E) or ($\vee$E) application in $\mathcal{D}$ and the major premise of that derivation must c-depend on χ. So, no discharge of χ is allowed before ψ is discharged, because χ and ψ do not lie on a thread. ⊠

So in any derivation $\mathcal{D}$ if assumption ψ is introduced c-depending on assumption φ and φ *supports* ψ, then φ cannot be discharged before ψ.

4.9. EXAMPLE. Consider the derivation

$$\dfrac{\begin{matrix}\Gamma \\ \vdots \\ \exists x\varphi\end{matrix} \qquad \begin{matrix}\varphi[a/x]\ [i] \\ \vdots \\ \xi_1\end{matrix}}{\xi_2}\ \exists\mathrm{E}_i$$

where ξ_1 and ξ_2 denote different occurrences of the formula ξ. Let χ be an assumption in Γ. Then this derivation gives rise to the following thread and c-thread:

$$< \chi, \ldots, \exists x\varphi, \xi_2 > \qquad < \chi, \ldots, \exists x\varphi, \underbrace{\varphi[a/x], \ldots, \xi_1}, \xi_2 > .$$

On the left hand side we have a thread connecting χ and ξ, on the right hand side a c-thread. $ST_\mathcal{D}(\xi_1)$ contains $\langle \chi, \varphi[a/x] \rangle$ with top element $\varphi[a/x]$. At ξ_1 only $\varphi[a/x]$ may be discharged. The underbraced part of the c-thread consists of a *no-discharge zone* for assumption χ: $\varphi[a/x] \prec^c_\mathcal{D} \chi$ and for the penultimate occurrence of ξ we have $\xi \prec^c_\mathcal{D} \varphi[a/x]$ and, because χ is active at $\exists x\varphi$, we have $\xi \prec^c_\mathcal{D} \chi$. If we would allow discharge of χ in the underbraced part of the c-thread, then we could get $\xi \prec^c_\mathcal{D} \varphi[a/x]$ and $\varphi[a/x] \prec^c_\mathcal{D} \chi$, but $\xi \not\prec^c_\mathcal{D} \chi$. This destroys transitivity.

Notice that by Proposition 4.8 stack-discharge reduces to discharge simpliciter. If ψ is a top element of $ST_\mathcal{D}(\varphi)$, then φ and ψ are connected by a *thread.* But this reduction hides an important difference between CPL and IPL. This comes to light if we allow *permutations* of the stackset at some point in the derivation. In that case we allow discharge of assumptions at formula occurrences not thread-connected to those assumptions. The notion of c-dependence is defined with respect to assumptions of the rules ($\exists$E) and ($\vee$E), so they arise both in CPL and in IPL. But we shall show that it is only essential in IPL. In classical logic *permuting the stackset* is allowed at any point in the derivation.

That is, breach of the discharge order induced by c-dependence does not result in classically invalid principles, in intuitionistic logic it allows us to derive intuitionistically invalid principles. In IPL no permutations of the stackset are allowed. This will be discussed in the next two sections.

4.1.3 Classical Dependence Management

The notion of the *c-dependence* of a formula occurrence on an assumption occurrence has been defined relative to the notion of a c-thread. But if φ is an assumption occurrence starting a c-thread through formula occurrence ψ in derivation $\mathcal{D}$, then this does not entail that φ could be discharged at ψ nor that ψ is a conclusion of a subderivation of $\mathcal{D}$. In contrast with the notion of a *thread*, the notion of a c-thread does not control the dependence of a *conclusion* on an assumption, nor does it control *discharge*. This changes when we consider alternatives to the ($\exists$E) and ($\vee$E) rules. We have discussed the rules of existential instantiation (Section 2.5.3) and the *epsilon rule* (Section 2.3.8). Here is a representation of a derivation of ξ from Γ, Σ in the three systems:

$$\dfrac{\begin{array}{c}\Gamma\\ \vdots\\ \exists x\varphi\end{array}}{\begin{array}{c}\varphi[\epsilon x:\varphi/x]\\ \vdots\\ \xi_1\end{array}}\,\exists\mathrm{E}_\epsilon \qquad \dfrac{\begin{array}{c}\Gamma\\ \vdots\\ \exists x\varphi\end{array}}{\begin{array}{c}\varphi[a/x]\\ \vdots\\ \xi_1\end{array}}\,\exists\mathrm{E}\ a \prec AO(\varphi)$$

and

$$\dfrac{\begin{array}{c}\Gamma\\ \vdots\\ \exists x\varphi\end{array} \qquad \begin{array}{c}\varphi[a/x]\ [i]\\ \vdots\\ \xi_1\end{array}}{\xi_2}\,\exists\mathrm{E}_i$$

In the epsilon and the existential instantiation derivations, the side derivation of the ($\exists$E) application lies below the major premise, that is, the major premise of the application and the assumption lie on a *thread* in the instantiation frameworks. So in the alternatives to ($\exists$E), the notions of ψ c-depending on φ and ψ depending on φ coincide with respect to applications of the elimination rule for the existential quantifier. For χ an assumption in Γ, this gives

$$\underbrace{<\chi,\ldots,\exists x\varphi,\varphi[\epsilon x:\varphi/x],\ldots,\xi_1>}_{\text{thread}} \qquad <\underbrace{\chi,\ldots,\exists x\varphi,\varphi[a/x],\ldots,\xi_1>}_{\text{thread}}$$

and

$$\underbrace{<\chi,\ldots,\exists x\varphi,\underbrace{\varphi[a/x],\ldots,\xi_1},\xi_2>}_{\text{c- thread}}$$

Because discharge is controlled by threads, both alternatives to ($\exists$E) seem to allow *discharge* of an assumption which, in ($\exists$E) derivations amounts to discharge along c-threads. That is, in the ($\exists$E) derivation, the stackset $ST_{\mathcal{D}}(\xi_i)$ has $\varphi[a/x]$ as a top element. This element is lacking in the stackset at ξ_1 in the remaining two derivations. So *stack-discharge* will allow different actions in these examples.

Because both the instantiation rule and the ϵ-rule are conservative over classical but not over intuitionistic logic, this suggests that the notion of c-dependence captures an essential difference between CPL and IPL: in CPL c-dependence can be turned into dependence, in IPL these relations are essentially distinct. In CPL the *ordering* of the stackset of assumptions at a formula occurrence is not essential, in IPL it is.

4.10. DEFINITION. (c-Discharge) In a derivation $\mathcal{D}$ assumption occurrence ψ can be *c-discharged* at formula occurrence φ, if ψ and φ lie on an c-thread in $\mathcal{D}$ such that ψ has not been discharged between ψ and φ, and

1. if ψ is an ($\exists$E)-assumption of the form $\xi[a/x]$, then the proper term a does not occur in φ or in any non c-discharged assumption on a c-thread through φ,
2. if ψ, χ are the assumptions of some application of ($\vee$E), then the formula φ occurs on a c-thread starting with χ.

For $\Sigma \cup \{\varphi\}$ a set of formulas and C a proof system, $\Sigma \vdash^c_C \varphi$ means that φ can be derived from Σ in system C by using c-discharge.

Under c-discharge any of a number of assumption c-connected to a formula occurrence φ may be discharged at this occurrence. The c-dependence ordering becomes irrelevant in the ordering of discharge.

4.11. PROPOSITION. *Derivation $\mathcal{D}$ derives φ from Σ with c-discharge if and only if $\mathcal{D}$ derives φ from Σ with stack-discharge, where at any formula occurrence φ in $\mathcal{D}$, every stack in the stackset $ST_{\mathcal{D}}(\varphi)$ may be arbitrarily permuted.*

Proof: By c-discharge all assumptions lying on the same c-thread may be discharged (provided that the discharge conditions are satisfied) independent of their position on that c-thread. So c-discharge of an arbitrary non-top element of a stack in the stackset corresponds to stack-discharge where this element has been brought to the top of the stack. ⊠

4.12. EXAMPLE. Consider the derivation (†) which offends the classical discharge regime. The major premise of ($\exists$E) has been established along the thread starting at χ. Now this χ, an assumption on which the major premise depends, is discharged inside the application of ($\exists$E). That is, χ is discharged at a location not connected to that assumption by a thread. Notice that we mean proper discharge, not empty discharge. That is, the conclusion ψ no longer depends on χ. In this derivation we have $\varphi[a/x] \prec^c_{\mathcal{D}} \chi$, $\psi \prec^c_{\mathcal{D}} \varphi[a/x]$ but $\psi \not\prec^c_{\mathcal{D}} \chi$. The dependence between assumptions is not respected. We shall show that, classically, this disturbance of dependence does not lead to unsoundness.

$$\dfrac{\begin{array}{c}\chi\ [i]\\ \vdots \\ \exists x\varphi\end{array} \qquad \begin{array}{c}\varphi[a/x]\ [j]\\ \vdots\\ \dfrac{\alpha[a/x]}{\chi\to\alpha[a/x]}\to\mathrm{I}_i\\ \vdots\\ \psi\end{array}}{\psi}\exists\mathrm{E}_j \tag{$\dagger$}$$

4.13. PROPOSITION. *Let $\Sigma \cup \{\varphi\}$ be a set of formulas, then*

$$\Sigma \vdash_{CPL} \varphi \iff \Sigma \vdash^{c}_{CPL} \varphi.$$

Proof: Because c-discharge is a more liberal discipline than discharge, we only have to show that the right to left direction holds. This we have done if we can transform all c-threads in a derivation $\mathcal{D}$ with c-discharge into threads in a derivation $\mathcal{D}'$ with standard discharge. That is, we have to represent the ($\exists$E) and ($\vee$E)-*assumptions* as *conclusions* of their respective major premises. Classical logic has two principles allowing us to achieve just this

$$((\varphi \vee \psi) \to \varphi) \vee ((\varphi \vee \psi) \to \psi), \tag{P$\vee$}$$

$$\exists x(\exists x\varphi \to \varphi). \tag{P$\exists$}$$

The principle P$\vee$ is used to transform c-threads to threads in applications of ($\vee$E). This principle is a propositional version of Plato's principle and yet another form of the *linearity* schemas of the last chapter. We use it to transform an application of ($\vee$E)

$$\dfrac{\begin{array}{c}\Gamma\\ \vdots\\ \varphi\vee\psi\end{array} \quad \begin{array}{c}\varphi\ [i]\\ \vdots\\ \xi\end{array} \quad \begin{array}{c}\psi\ [j]\\ \vdots\\ \xi\end{array}}{\xi}\vee\mathrm{E}_{i,j}$$

in a derivation $\mathcal{D}$, into the following configuration:

$$\dfrac{\boxed{\mathrm{P}\vee} \qquad \begin{array}{c}\Gamma\\ \vdots\\ \dfrac{\varphi\vee\psi \quad \boxed{(\varphi\vee\psi)\to\varphi\ [i]}}{\varphi}\\ \vdots\\ \xi\end{array} \qquad \begin{array}{c}\Gamma\\ \vdots\\ \dfrac{\varphi\vee\psi \quad \boxed{(\varphi\vee\psi)\to\psi\ [j]}}{\psi}\\ \vdots\\ \xi\end{array}}{\xi}\vee\mathrm{E}_{i,j}$$

Here we use *standard* discharge and the rule (∨E) with the theorem P∨ as major premise. Obviously, P∨ can be derived with standard discharge. In the application of (∨E) the major premise $\varphi \vee \psi$ is connected to φ and ψ by a c-thread, while in the transformed derivation $\varphi \vee \psi$ is connected to φ and ψ by a thread. Consequently, an application of c-discharge with respect to (∨E) in a derivation $\mathcal{D}$ of φ from Σ can be changed to an application of standard discharge in a derivation $\mathcal{D}'$ of φ from Σ.

For applications of (∃E) we follow the same procedure, this time using the principle P∃. Again we transform a connection along a c-thread to a connection along a thread. That is, we transform

$$\dfrac{\begin{array}{c}\Gamma\\ \vdots\\ \exists x\varphi\end{array} \qquad\qquad \begin{array}{c}\varphi(a/x)\ [i]\\ \vdots\\ \xi\end{array}}{\xi}\ \exists \mathrm{E}_i$$

into

$$\dfrac{\boxed{\exists x(\exists x\varphi \to \varphi)} \qquad \begin{array}{c} \dfrac{\begin{array}{c}\Gamma\\ \vdots\\ \exists x\varphi\end{array}\quad \boxed{\exists x\varphi \to \varphi[a/x]\ [i]}}{\varphi[a/x]}\\ \vdots\\ \xi\end{array}}{\xi}\ \exists \mathrm{E}_i$$

Again we use standard (∃E) discharge in this derivation, with the theorem P∃ as the major premise. ⊠

Thus, as these transformations show, all assumptions on which the major premises of (∨E) and (∃E) applications are based may be discharged by (→I) at any point in the derivation. The assumptions of the major premise of an (∨E) application may be discharged at any point in the side derivations. Only the discharge order of (∃E) assumptions remains strict. This we can see in the last derivation above. In order to discharge $\exists x\varphi \to \varphi[a/x]$, the proper term should not occur in the conclusion ξ or in any assumption in Γ.

We have shown the following.

4.14. PROPOSITION. *Formula φ is CPL derivable from Σ if and only if φ is derivable from Σ under the stack-discharge regime with arbitrary permutations of the stackset.*

4.1.4 Intuitionistic Dependence Management

The principles P$\exists$, P$\vee$ (and $P\forall$) are classically derivable with standard discharge. But intuitionistically both are invalid rules, that is, not derivable with standard discharge. So intuitionistically, the notion of c-dependence arising from the natural deduction rules is essentially different from the notion of dependence: using c-discharge in derivations allows for the derivation of intuitionistically invalid principles. And the P$\vee$ and P$\exists$ principles characterize *precisely* intuitionistic logic with a c-discharge convention.

4.15. PROPOSITION. $\Sigma \vdash_{IPL+P\vee+P\exists} \varphi \iff \Sigma \vdash^{c}_{IPL} \varphi$.

Proof: From right to left follows because the rules P$\vee$ and P$\exists$ are derivable in IPL plus c-discharge. As an example we shall derive P$\exists$ with c-discharge.

$$\dfrac{\exists x\varphi\ [1] \qquad \dfrac{\dfrac{\varphi[a/x]\ [2]}{\exists x\varphi \to \varphi[a/x]}\to\mathrm{I}_{11}}{\exists x(\exists x\varphi \to \varphi)}\exists\mathrm{I}}{\exists x(\exists x\varphi \to \varphi)}\exists\mathrm{E}_2$$

From left to right follows because any application of c-discharge in a derivation $\mathcal{D}$ can be mimicked by an application of one of the IP rules. ⊠

A consequence of this proposition is the fact that the notion of a dependence *stack* is strict in IPL.

4.16. PROPOSITION. *Formula φ is IPL derivable from Σ if and only if Φ is derivable from Σ under the stack-discharge regime without permutations of the stackset.*

4.17. REMARK. In CPL we need the principle of double negation to derive P$\exists$. Here is a derivation (where we take some shortcuts):

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\forall x(\exists x\varphi \wedge \neg\varphi)\ (1)}{\exists x\varphi \wedge \neg\varphi[a/x]}}{\exists x\varphi} \qquad \dfrac{\varphi[b/x]\ (2) \qquad \dfrac{\dfrac{\forall x(\exists x\varphi \wedge \neg\varphi)\ (1)}{\exists x\varphi \wedge \neg\varphi[b/x]}}{\neg\varphi[b/x]}}{\bot\ (-2)}}{\bot}}{\neg\forall x(\exists x\varphi \wedge \neg\varphi)\ (-1)}}{\exists x\neg(\exists x\varphi \wedge \neg\varphi)}}{\exists x(\exists x\varphi \to \neg\neg\varphi)}}{\exists x(\exists x\varphi \to \varphi).}$$

The last three lines of the derivation use double negation (to get from $\neg\forall x$ to $\exists x\neg$, and to get from $\neg\neg\varphi$ to φ). If we compare this to the simple derivation of P$\exists$ with c-discharge, we notice that the principle of double negation hides the *dependency structure* which lies underneath the principle P$\exists$.

4.2 Conservative Epsilon Extensions of IPL

Derivations using the instantiation rules, or the epsilon rule cannot be shown to be intuitionistically sound by embedding them in intuitionistic derivations. This suggests the following question. How can we formulate a rule of existential instantiation and an epsilon rule sound for intuitionistic logic?

In order to develop conservative extensions of IPL with the epsilon rule we have to fix some conditions such an extension will have to satisfy. After all, it is no problem to device trivial conservative extensions: if we stipulate "do not apply any IPL rule to premises containing ϵ-terms", then the addition of the epsilon rule is conservative. To formulate our non-triviality condition, we shall take our lead from the the classical epsilon calculus of Chapter 2. There the epsilon rule can be seen as an *alternative* to the elimination rule for the existential quantifier. By *replacing* the rule ($\exists$E) by the ϵ-rule in CPL we do not reduce the set of ϵ-free theorems. This constraint we shall also invoke for ϵ-extensions of IPL. That is, we are after a conservative extension of IPL with the epsilon rule in which we can do without the rule ($\exists$E). Obviously the calculus arising by the above trivial stipulation does not satisfy this condition.

Before we formulate adequate restrictions on the calculus we shall discuss some proof-theoretic possibilities and their problems. In a later section, we shall also discuss these things semantically. To get at a conservative ϵ-extension, consider the derivation of the principle that creates the non-conservativity, Plato's principle.

$$\dfrac{\dfrac{\dfrac{\exists x\varphi\ [1]}{\varphi[\epsilon x:\varphi/x]}\ \exists\mathrm{E}_\epsilon}{\exists x\varphi\rightarrow\varphi[\epsilon x:\varphi]}\ {\rightarrow}\mathrm{I}_1}{\exists x(\exists x\varphi\rightarrow\varphi)}\ \exists\mathrm{I}$$

This derivation has to be ruled out as being incorrect. So we need restrictions which are tight enough to exclude this derivation, but which still allow us to derive all IPL theorems. In the above derivation, there are two proof steps at which restrictions can apply.

1. At the application of the rule ($\rightarrow$I) we can prevent the discharge of $\exists x\varphi$. This affects the deduction theorem. The intuitionistic epsilon calculus has the epsilon rule, but would not have $\exists x\varphi\rightarrow\varphi[\epsilon x:\varphi/x]$ as a theorem. The restriction suggested by our desideratum that ϵ-rule applications may be replaced by ($\exists$E) applications is the following. With respect to discharge, we have to treat the *conclusion* of the ϵ-rule, $\varphi[\epsilon x:\varphi/x]$ in the second line of the derivation, as an *assumption* of an ($\exists$E) application. We view $\epsilon x:\varphi$ as the *proper term* of that application. The discharge in the third line is then disallowed, because we may not discharge the major premise

of an ($\exists$E) application before we have discharged the assumption of that application. In the second line of the proof, the ϵ-term $\epsilon x : \varphi$ is *supported* by an assumption claiming $\exists x\varphi$. In the third line of the proof this is no longer the case. The analysis of the structure of ($\exists$E) of the last section, will give us the discharge restriction we need.

2. At the application of ($\exists$I) we can restrict abstraction over the ϵ-term. This course would then give us $\exists x\varphi \rightarrow \varphi[\epsilon x : \varphi/x]$ as a theorem, but not Plato's principle. This course is suggested by the analogy with modal predicate logic with non-rigid terms, and, semantically, this analogy is of course a natural one, given Gödels translation of intuitionistic logic into S.4. In this case we may restrict applications of ($\exists$I) to "non-modal" contexts, that is, to premises $\varphi[t/x]$, where φ is built from atomic propositions using only $\wedge, \vee$ and $\exists$.

Conservativity by Restricting Discharge

In this section we shall formulate a general restriction on discharge in the intuitionistic ϵ-calculus wich will guarantee a strong property of the calculus: in any derivation with conclusion and assumptions epsilon free, every application of the ϵ-rule can be replaced by an application of the rule ($\exists$E). Our calculus will resemble the one of Leivant [Lei73]. The main difference consists in the fact that we show conservativity for full IPL, while Leivant only discusses a disjunction and negation free fragment of this calculus. Solutions in the same vein have been developed by Celluci [Cel92] based on the sequent natural deduction calculus of Boričić [Bor85]. A solution of a different kind altogether is adopted by Mints [Min77]. There the ϵ-calculus is investigated in the format of Gentzens *sequent* calculus.

Our analysis of implicit dependence of the last section shows the main proof theoretic difference between the ϵ-rule and the rule ($\exists$E). Consider again the structure of the epsilon and the elimination rule.

$$
\begin{array}{c}
\Gamma \\
\vdots \\
\dfrac{\exists x\varphi}{\varphi[t/x]}\ \exists\text{E Inst.} \\
\left\{\begin{array}{c} \vdots \\ \gamma \\ \vdots \\ \xi \end{array}\right.
\end{array}
\qquad\qquad
\dfrac{\begin{array}{c}\Gamma \\ \vdots \\ \exists x\varphi\end{array} \qquad \left\{\begin{array}{c}\varphi[t/x]\ [i] \\ \vdots \\ \gamma \\ \vdots \\ \xi\end{array}\right.}{\xi}\ \exists\text{E}_i
$$

In the right hand derivation, no assumption from Γ may be discharged in the side derivation, for instance at γ. This situation we must make explicit in the left hand derivation, for the implicit dependence relation does not prevent this discharge. The derivation from $\varphi[a/x]$ to ξ must be explicitly declared a *no-discharge zone* for assumptions from Γ. This is easily done by copying the relevant ($\exists$E)-dependencies for the epsilon rule.

> In order to get a conservative IPL extension with the epsilon rule all we have to do is to put the conclusion of the $\varphi[t/x]$ of an application of the epsilon rule on top of $ST_{\mathcal{D}}(\varphi[t/x])$. That is, we have to treat it as an *assumption* with respect to dependence. This prevents discharge under the stack-discharge regime of any assumption on which $\exists x\varphi$ depends.
> The formula $\varphi[t/x]$ may be *removed* from the $ST_{\mathcal{D}}(\psi)$ if the proper term t does not occur in ψ or in any assumption χ such that $\psi \prec^{c}_{\mathcal{D}} \varphi$.

Thus we have precisely mimicked the discharge configuration of ($\exists$E) in derivations with the ϵ-rule.

4.18. DEFINITION. (I-Support) Let $IPL+I$ be the proof system of intuitionistic predicate logic plus the rule ϵ-rule. Let $\mathcal{D}$ be a derivation tree for $IPL+I$ with assumptions Σ, ψ. We say that assumption occurrence ψ *I-supports* formula occurrence φ in $\mathcal{D}$ *with respect to term* t, notation $\varphi \prec^{t}_{\mathcal{D}} \psi$, if there is an application of $(\exists E_\epsilon)$ in $\mathcal{D}$ with premise $\exists x\xi$ and conclusion $\xi[t/x]$ such that $\varphi \prec^{c}_{\mathcal{D}} \xi[t/x]$, $\exists x\xi \prec^{c}_{\mathcal{D}} \psi$, and the proper term t occurs in φ or in any assumption χ such that $\varphi \prec^{c}_{\mathcal{D}} \chi$. We say that assumption occurrence ψ *I-supports* formula occurrence φ in $\mathcal{D}$, notation $\varphi \prec^{I}_{\mathcal{D}} \psi$, if there is some term t such that $\varphi \prec^{t}_{\mathcal{D}} \psi$.

Now we stipulate:

4.19. DEFINITION. (Discharge Discipline) Let $IPL+I$ be an intuitionistic natural deduction system to which an instantiation rule is added. Let $\Sigma \cup \{\varphi\}$ be a set of formulas. Then $\Sigma \vdash^{s}_{i} \varphi$ if there is an $IPL+I$ derivation of φ from Σ where every application of ($\rightarrow$I) satisfies the following restriction:

$$\dfrac{\begin{array}{c}\varphi\ [i]\\ \vdots\\ \psi\end{array}}{\varphi \rightarrow \psi}\rightarrow\mathrm{I}_i$$

provided not $\varphi \prec^{I}_{\mathcal{D}} \psi$

The adequacy of this definition is shown by Proposition 4.25 which says, in fact, that the instantiation rule becomes *conservative* over IPL.

Properties of the Proof System

To get a feeling for the theorems of the ϵ-calculus restricted in this way, we shall discuss some peculiarities. Although we have $\exists x\varphi/\varphi[\epsilon x : \varphi/x]$ as a rule of inference, $\exists x\varphi \rightarrow \varphi[\epsilon x : \varphi/x]$ and $\exists x(\exists y\varphi \rightarrow \varphi)$ are not theorems in IPL+I, i.e., the deduction theorem does not hold in general in this system. However we do have the weaker $\{\exists x\varphi\} \vdash_i^I \exists x\varphi \rightarrow \varphi[\epsilon x : \varphi/x]$.

It is clear that, for instance, *contraposition* is not generally valid: we have $\exists x\varphi \vdash_i^I \varphi[\epsilon x : \varphi/x]$, but not $\neg\varphi[\epsilon x : \varphi/x] \vdash_i^I \neg\exists x\varphi$. We do have $\vdash_i^I \varphi[\epsilon x : \varphi/x] \rightarrow \exists x\varphi$ and $\vdash_i^I \neg\exists x\varphi \rightarrow \neg\varphi[\epsilon x : \varphi/x]$.

Whenever we consider an ϵ-deduction with ϵ-free assumptions and conclusion χ, then no formula will I-support one of the assumptions, and the restriction on discharge reduces to: no ψ to be discharged may I-support χ. However, when we also consider arbitrary ϵ-derivations with, for instance $\psi[\epsilon x : \varphi/x]$ among its premises, then assumption $\exists x\varphi$ cannot be discharged. Consequently we have

$$\{\exists x\varphi, \psi[\epsilon x : \varphi/x]\} \vdash_i^I \exists x(\varphi \wedge \psi)$$

and

$$\{\exists x\varphi\} \vdash_i^I \psi[\epsilon x : \varphi/x] \rightarrow \exists x(\varphi \wedge \psi)$$

but

$$\{\psi[\epsilon x : \varphi/x]\} \not\vdash_i^I \exists x\varphi \rightarrow \exists x(\varphi \wedge \psi).$$

The restrictions on discharge affect the notion of consistency. We have

$$\{\exists x\varphi), \neg\varphi[\epsilon x : \varphi/x]\} \vdash_i^I \bot.$$

But, having derived $\bot$ we cannot discharge $\exists x\varphi$. So $\{\neg\varphi[\epsilon x : \varphi/x]\} \not\vdash_i^I \neg\exists x\varphi$. As a consequence we now have

$$\{\neg\varphi[\epsilon x : \varphi/x], \neg\neg\exists x\varphi\} \not\vdash_i^I \bot.$$

For if $\{\neg\varphi[\epsilon x : \varphi/x], \neg\neg\exists x\varphi\}$ were inconsistent, then we could derive $\neg\exists x\varphi$ from $\neg\varphi[\epsilon x : \varphi/x]$ (discharge of $\neg\neg\exists x\varphi$ is allowed and $\neg\psi \leftrightarrow \neg\neg\neg\psi$ is an IPL-theorem). This implies that we can no longer conclude that $\Sigma \cup \{\varphi\}$ is consistent given the consistency of $\Sigma \cup \{\neg\neg\varphi\}$.

The general picture is the following. We cannot move an existential formula to the right of $\vdash_i^I$ if it leaves its proper term on the left unsupported. In fact, unsupported ϵ-terms on the left should not be interpreted as ϵ-terms at all: in order for an ϵ-term to get an *ϵ-interpretation*, its interpretation as a *witness*, it needs a licensing condition, the presence of an introducing formula. We are not allowed to reason: here we have $\psi[\epsilon x : \varphi/x]$, let's *assume* $\epsilon x : \varphi$ is a witness (i.e., let's assume $\exists x\varphi$) and consider the conclusions that follow as conclusions from $\psi[\epsilon x : \varphi/x]$. For, having once assumed $\exists x\varphi$ we cannot discharge it anymore. The following propositions justify this interpretation strictly.

4.20. DEFINITION. An ϵ-term $\epsilon x : \chi$ is *mute* in a derivation $\mathcal{D}$ if there is no ϵ-rule application in $\mathcal{D}$ with premise $\exists x\chi$. $Mu(\mathcal{D}) = \{\epsilon x : \varphi \in \mathcal{E}(\mathcal{D}) \mid \epsilon x\varphi \text{ is mute in } \mathcal{D}\}$

4.21. PROPOSITION. (Muteness) *For any ϵ-derivation $\mathcal{D}$ with conclusion ψ, and any constant a not occurring in $\mathcal{D}$, if $\epsilon x : \varphi \in Mu(\mathcal{D})$, then $\mathcal{D}[a/\epsilon x : \varphi]$ is an ϵ-derivation with conclusion $\psi[a/\epsilon x : \varphi]$.*

Proof: The truth of this proposition can be shown simply by checking all derivation rules. The only troublesome case, the ϵ-rule, is excluded by the muteness of $\epsilon x : \varphi$. ⊠

4.22. PROPOSITION. (Constant Proposition) *For every set of $\mathcal{L}$-sentences $\Sigma \cup \{\varphi\}$, if $\Sigma \nvdash^I_i \exists x\psi$, then $\Sigma \vdash^I_i \varphi$ if and only if $\Sigma[a/\epsilon x : \psi)] \vdash^I_i \varphi[a/\epsilon x : \psi)]$ for an arbitrary fresh constant a.*

Proof: Let $\epsilon x : \psi$ occur in $\Sigma \cup \{\varphi\}$ and $\Sigma \nvdash^I_i \exists x\psi$. There are two cases to consider.

1. The term $\epsilon x : \psi$ is not involved in derivation $\mathcal{D}$ of φ as a proper term of the (ϵ) rule. Then, by Proposition 4.21, $\epsilon x : \psi$ can be replaced in $\mathcal{D}$ by individual constant a while preserving derivationhood.
2. There is in derivation $\mathcal{D}$ of φ from Σ an application of the ϵ-rule in which $\epsilon x : \psi$ occurs as a proper term. Thus $\exists x\psi$ – or some ξ I-supporting $\exists x\psi$ – is discharged at some point in $\mathcal{D}$. By the restrictions on discharge, this implies that $\exists x\psi$ does not I-support φ. So, substituting a fresh individual constant a for $\epsilon x : \psi$ in $\mathcal{D}$ turns the application of the ϵ-rule, $\exists x\psi/\psi[\epsilon x\psi/x]$ into the pair $\exists x\psi$, $\psi[a/x]$, where the second formula is now the minor premis of $(\exists E)$ which can be discharged upon reaching the first ξ not I-supported by $\exists x\psi$ (such a ξ exists because φ is not I-supported by $\exists x\psi$). ⊠

It is interesting to remark that Lemma 4.22 does not hold for logics without the full restrictions of the intuitionistic ϵ-calculus. In such a logic there would be no guarantee that a derivation in which an existential formula has been discharged can be transformed into an ϵ-free one.

4.23. PROPOSITION. (Discharge Proposition) *Let $\Sigma \cup \{\varphi, \psi\}$ be a set of $\mathcal{L}$-formulas such that for all $\epsilon x : \xi \in \mathcal{E}(\Sigma, \psi)$*

$$\Sigma, \varphi \vdash^I_i \exists x\xi \Rightarrow \Sigma \vdash^I_i \exists x\xi.$$

Then

$$\Sigma, \varphi \vdash^I_i \psi \Rightarrow \Sigma \vdash^I_i \varphi \rightarrow \psi.$$

Proof: Let $\mathcal{D}$ be a derivation of ψ from Σ, φ. Suppose $\epsilon x : \xi$ occurs in ψ or Σ and there is an application of the ϵ-rule with $\exists x\xi$ as premise depending on φ. Now, by assumption, $\Sigma \vdash^I_i \exists x\xi$. Consequently, in the derivation $\mathcal{D}$ we can cut out all subderivations with conclusion $\exists x\xi$ depending on φ and replace them by a derivation of this formula from assumptions only in Σ. In this new derivation, φ does no longer support ψ and can be discharged. ⊠

4.24. REMARK. Notice that there is a definite notion of *binding* involved: an existential formula $\exists x\varphi(x) \in \Sigma$ 'binds' all occurrences of $\epsilon x : \varphi$ in Σ. If $\epsilon x : \varphi$ is 'free' in Σ, its identity is immaterial (it can be universally quantified).

Now the consistency of $\{\neg\varphi[\epsilon x : \varphi/x], \neg\neg\exists x\varphi\}$ can be interpreted by not viewing $\epsilon x : \varphi$ as the witness for $\exists x\varphi$, i.e., by considering $\{\neg\varphi[a/x], \neg\neg\exists x\varphi\}$. Even $\{\neg\varphi[a/x], \exists x\varphi\}$ can be consistent under this interpretation. However this set is only consistent if viewed as an extension of $\{\neg\varphi[\epsilon x : \varphi/x]\}$, for $\{\neg\varphi[\epsilon x : \varphi/x]\} \vdash_i^I \psi \iff \{\neg\varphi[a/x]\} \vdash \psi[a/x]$. It is inconsistent as an extension of $\{\exists x\varphi\}$! From $\exists x\varphi$ the negation of $\neg\varphi[\epsilon x : \varphi/x]$ is derivable, but the negation of $\exists x\varphi$ is not derivable from $\neg\varphi[\epsilon x : \varphi/x]$. The operations of replacing mute ϵ-terms in a theory by individual constants and extending the theory do not commute.

4.25. PROPOSITION. (Conservativity) *Let $IPL + I$ consist of the intuitionistic natural deduction system plus the ϵ-rule. Let $\Sigma \cup \{\varphi\}$ be a set of formulas free of proper terms of the instantiation rule. Then:* $\Sigma \vdash_i^I \varphi \iff \Sigma \vdash_{IPL} \varphi$.

Proof: We are going to proceed by transforming every application of the ϵ-rule in a derivation $\mathcal{D}$ of the specified form into an application of ($\exists$E), a rule present in IPL. The following figure contains the relevant information for this transformation process: we face an ϵ-derivation $\mathcal{D}$ with ϵ-free assumptions in Γ and ϵ-free conclusion χ.

$$
\mathcal{D} \quad \begin{array}{c} \Gamma \\ \vdots \\ \Sigma \quad \dfrac{\exists x\varphi}{\varphi[\epsilon x : \varphi/x]} \\ \mathcal{D}_1 \\ \psi \\ \vdots \\ \chi \end{array}
\qquad \Rightarrow : \mathcal{D}' \quad
\begin{array}{c} \dfrac{\begin{array}{c}\Gamma \\ \vdots \\ \exists x\varphi\end{array} \qquad \begin{array}{c}\varphi[a/\epsilon x : \varphi] \\ \mathcal{D}_1[a/\epsilon x : \varphi] \\ \psi[a/\epsilon x : \varphi]\end{array}}{\psi[a/\epsilon x : \varphi]} \\ \vdots \\ \chi \end{array}
$$

In the right hand derivation, a is some fresh constant, not occurring anywhere in $\mathcal{D}$. Let the highlighted application of the ϵ-rule be the *lowest* in $\mathcal{D}$ with the premise $\exists x\varphi$. That is, in $\mathcal{D}_1$ there occurs no application of the ϵ-rule with this premise. The formula ψ is here the first formula below $\varphi[\epsilon x : \varphi/x]$ in which $\epsilon x : \varphi$ does not occur and which does not lie on a thread starting with assumptions, active at ψ, in which this term occurs. Such a ψ must exist, because Γ and χ are ϵ-free. To see that we have a correct derivation, we have to check the following:

1. the constant a should not occur in any formula in $\Sigma[a/\epsilon x : \varphi]$, in $\psi[a/\epsilon x : \varphi]$ or in φ,
2. $\Sigma[a/\epsilon x : \varphi], \varphi[a/\epsilon x : \varphi]/\mathcal{D}_1[a/\epsilon x : \varphi]/\psi[a/\epsilon x\varphi]$ should be a correct deduction,

3. no assumption on which $\exists x\varphi$ depends has been discharged between $\varphi(\epsilon x : \varphi)$ and ψ, for in an application of ($\exists$E) the major premise, $\exists x\varphi$, can only be discharged after the minor one $\varphi(a)$.

We shall discuss the facts one by one. The first fact holds because a is fresh and neither ψ nor any element of Σ contains $\epsilon x : \varphi$. The formula ψ was chosen with these properties. The second demand is also satisfied. Because we have chosen a lowest application of the ϵ-rule with premise $\exists x\varphi$, in the subderivation

$$\dfrac{\dfrac{\Sigma, \varphi[\epsilon x : \varphi/x]}{\mathcal{D}_1}}{\psi}$$

of $\mathcal{D}$ there occurs no application of the ϵ-rule with proper term $\epsilon x : \varphi$. In such a derivation the term $\epsilon x : \varphi$ is mute and can be replaced by an ordinary (fresh) individual constant while preserving derivationhood. This follows by Proposition 4.21. So the first two demands are always satisfied, when we choose a lowest application of the ϵ-rule in a derivation. The third demand seems to exclude some ϵ-derivations from the transformation. But here our restriction on ($\to$I) comes into play. Discharge can be occasioned by either of the three rules ($\to$I), ($\vee$E), or ($\exists$E). We shall consider the cases one by one.

1. Suppose there is in $\mathcal{D}$ an assumption $\gamma \in \Gamma$ discharged by ($\to$I) at a formula occurrence α between $\varphi[\epsilon x : \varphi/x]$ and ψ. By the restriction on ($\to$I), this means that *not* $\gamma \prec^I_{\mathcal{D}} \alpha$. So α does not contain $\epsilon x : \varphi$ nor can it lie below an active assumption containing this term. But ψ was the highest such term below $\varphi[\epsilon x\varphi/x]$, so occurrence α must be equal to occurrence ψ.
2. In the notion of $IPL + I$ derivability no restrictions are placed on applications of ($\vee$E) or ($\exists$E). To show that this does not prevent conservativity, we shall argue that we can transform any application of one of these rules involving loss of breach of support into one in which support is respected. So suppose in a proof we have the following constellation of ($\vee$E)

$$\dfrac{\begin{array}{c}\Gamma\\ \vdots\\ \varphi\vee\psi\end{array} \quad \begin{array}{c}[\varphi]\\ \vdots\\ \xi\end{array} \quad \begin{array}{c}[\psi]\\ \vdots\\ \xi\end{array}}{\begin{array}{c}\xi\\ \vdots\\ \chi\end{array}}$$

where Γ, χ are ϵ-free, and suppose $\xi \prec^I_{\mathcal{D}} \varphi$ or $\xi \prec^I_{\mathcal{D}} \psi$. Then we cannot replace an application of the ϵ-rule in the derivation from φ to ξ or from ψ to ξ by an application of ($\exists$E). So this derivation does not satisfy the restriction imposed on ($\rightarrow$I). But we can always transform it to a derivation satisfying this restriction by

$$
\begin{array}{ccc}
\Gamma & [\varphi] & [\psi] \\
\vdots & \vdots & \vdots \\
\vdots & \xi & \xi \\
\cline{2-3}
\vdots & \xi & \xi \\
\vdots & \vdots & \vdots \\
\varphi \vee \psi & \eta & \eta \\
\hline
 & \eta & \\
 & \vdots & \\
 & \chi &
\end{array}
$$

Where η is the first formula occurrence below ξ not supported by φ or ψ. Such an η exists because χ is not supported by φ or ψ. Now, an application of the ϵ-rule in a derivation of ξ from φ or ψ can be replaced by ($\exists$E). There may seem to be a problem in this transformation, if, in the original derivation, some assumption from Γ is discharged between ξ and η. But this problem evaporates when one considers the possibilities.

(a) Discharges by applications of ($\vee$E) or ($\exists$E) are harmless, because the discharged formula does not, by the discharge, appear in ξ/η. Consequently, this subderivation stays correct and we can postpone discharge until we arrive at η.

(b) On the other hand, if discharge is induced by an application of ($\rightarrow$I), then, obviously, this discharge was allowed. For instance, if in the above derivation some $\alpha \in \Gamma$ is discharged by ($\rightarrow$I) at some formula δ between ξ and η, then *not* $\delta \prec^I_{\mathcal{D}} \alpha$. By the definition of dependence, this means that neither $\delta \prec^I_{\mathcal{D}} \varphi$ nor $\delta \prec^I_{\mathcal{D}} \psi$, for $\varphi \prec^I_{\mathcal{D}} \alpha$ and $\psi \prec^I_{\mathcal{D}} \alpha$ (there is no thread through ψ or φ and α), and, by transitivity, we would have $\delta \prec^I_{\mathcal{D}} \alpha$. But η was assumed to be the highest occurrence of a formula not supported by φ or ψ, so $\delta = \eta$.

3. Applications of ($\exists$E) can be transformed analogously to get correct discharge. ⊠

4.3 Kripke Models for Epsilon Terms

4.3.1 Semantic Strategies

In the preceding sections, our quest for a conservative intuitionistic epsilon calculus proceeded syntactically, through proof theory, with a heavy emphasis on dynamic dependence structures in proofs. But we can also think about this issue semantically. In Chapter 2, a guiding idea was the use of Skolem expansions for standard models. This semantic strategy forms another side of the coin. Proof-theoretically, we design some new calculus, and have to check that it does not generate undesired validities in the old language. But semantically, we start from a complete class of models for some underlying logic, and then expand these models with epsilon term interpretations. Evidently, the logic of the new model class does not extend the old logic in its original language, as no counter-examples are lost. But of course, now we have the problem of determining which proof theory describes the new logic completely.

Indeed, in semantics as in proof theory, the problem of 'conservative extension' is not unambiguous. More specifically, there are several ways of expanding Kripke models for intuitionistic logic with epsilon terms. A first strategy goes like the classical method of Chapter 2. We choose witnessing objects for $\epsilon x : \varphi$ in every node where an existential formula $\exists x \varphi$ is forced, while making arbitrary choices for $\epsilon x : \varphi$ in all other nodes. The striking feature of this semantics is that it is non-persistent. That is, formulas involving epsilon terms may change from true to false when going to nodes higher up, where the interpretation of their epsilon terms has shifted. A second strategy maintains persistence, by going partial: epsilon terms are only interpreted at those nodes where their supporting existential formula is forced. This is not enough: we also have to stick with the same choice of a witness higher up. Both strategies solve our problem. Both strategies pay a price: either non-persistence, or partiality. This multiplicity of solutions, and their logical side effects, seems a general feature of constructive epsilon calculi. For a third option, one could set up things in a partial logic with an existence predicate (see [Bla84], [RdL89]).

The fact that something has to go already emerged in our earlier proof-theoretic analysis. Even though we can maintain all intuitionistic 'object-level validities', we cannot maintain all general features of this logic. This will be clear by inspecting logical validities. On the first strategy, many principles even in the intuitionistic propositional base will become schematically invalid even though all their epsilon-free instances remain available. For, the validity of these principles often depends on semantic persistence. For instance, the conditional axiom $(\varphi \to (\psi \to \chi)) \to ((\varphi \to \psi) \to (\varphi \to \chi))$ remains unrestrictedly valid, but the conditional axiom $\varphi \to (\psi \to \varphi)$ drops out, as its validity presupposes persistence. More generally, like in the preceding section, the rule of conditionalization (or the deduction theorem) is only admissible for withdrawing persistent antecedents. And similar phenomena arise with other rules, such as universal generalization from premises not containing the quantified variable. There are solutions to these axiomatization problems, though. In partial logic (cf. Thijsse [Thi92], Jaspars [Jas94]), restricted versions of deduction theorems to persistent formulas only are well-known and Henkin-style completeness proofs can be adjusted to make do with these (see [Ben86b]). Such a strategy can also be pursued for the intuitionistic epsilon calculus. For instance, the rule of conditionalization would then only withdraw persistent antecedents, containing epsilon terms only in the scope of at least one implication or universal quantification. We shall not pursue this route here.

Instead, we shall explore the second, partial strategy in a little more detail. First, we define our models more precisely.

4.3.2 Partial Intuitionistic Epsilon Models

To construct an intuitionistic epsilon model we start with a standard Kripke model $K = \langle W, \leq, D\rangle$. In order to interpret epsilon terms at a node $\alpha \in W$ of this model we create for every node $\alpha \in W$ a *set of nodes* of the form $\{\langle \alpha, \Phi_\alpha\rangle \mid \Phi_\alpha \in C(\alpha)\}$, where $C(\alpha)$ is the set of all partial functions Φ_α from the set of epsilon terms and variable assignments in $dom(D(\alpha))$ satisfying

1. $\langle \epsilon x : \varphi, s\rangle \in dom(\Phi_\alpha)$ if and only if $K, \alpha, s, \Phi_\alpha \Vdash \exists x\varphi$,
2. if $\langle \epsilon x : \varphi, s\rangle \in dom(\Phi_\alpha)$, then $\Phi_\alpha(\langle \epsilon x : \varphi, s\rangle) \in \{m \in dom(D(\alpha)) \mid K, \alpha, s(m|x), \Phi_\alpha \Vdash \varphi\}$.

4.26. REMARK. To deal with the general case we actually need to associate with each node the *set* of all partial choice functions. For consider a Kripke model with an infinitely descending chain of nodes forcing $\exists x\varphi$ such that the set of φ-ers along this chain has an empty intersection. In that case, at no node can we fix a value for $\epsilon x : \varphi$ which takes into account the value assigned in all predecessors. At some $\leq$-predecessor $\epsilon x : \varphi$ must have been assigned a different value, thus precluding accessibility under $\leq_\epsilon$. If we want to preserve the structure of the model K for ϵ-free formulas, we need the guarantee that, for each node $\langle \alpha, \Phi_\alpha\rangle$ such that $\alpha \leq \beta$ there is a Φ_β such that $\langle \alpha, \Phi_\alpha\rangle \leq_\epsilon \langle \beta, \Phi_\beta\rangle$. Only by associating with a node the set of *all* partial choice functions there is such a guarantee. Now, given a node α forcing $\exists x\varphi$ we can always map $\epsilon x : \varphi$ to a domain element of the right sort by some function Φ_α, and any node β such that $\alpha \leq \beta$ can be supplied with a function Φ_β extending Φ_α.

Notice that, if we restrict ourselves to *well-founded* models the situation described above cannot occur. A well-founded model cannot yet be expanded with a single choice function at every node, for it may still occur that we have two minimal elements with a common $\leq$-successor: the choice functions defined at the minimal elements may assign different elements to the same ϵ-term, thus precluding accessibility of the common $\leq$-successor. But if we confine ourselves to well-founded Kripke models with a single bottom element, then we can do with a single choice function.

The *term valuation* V_{α,s,Φ_α} of the partial intuitionistic Kripke model is an extension of the valuation $V_{\alpha,s}$ to cover the interpretation of ϵ-terms. For atomic formulas $P(t_1, \ldots, t_n)$, we set: $K, \alpha, s, \Phi_\alpha \Vdash \varphi$ holds iff $\langle \epsilon x : \psi, s\rangle \in dom(\Phi_\alpha)$ for all ϵ-terms occurring in $P(t_1, \ldots, t_n)$ and

$$\langle V_{\alpha,s,\Phi_\alpha}(t_1), \ldots, V_{\alpha,s,\Phi_\alpha}(t_n)\rangle \in int(D(\alpha))(P).$$

The partial intuitionistic epsilon model K_ϵ is now constructed by setting:

1. $W_\epsilon = \{\langle \alpha, \Phi_\alpha\rangle \mid \alpha \in W, \Phi_\alpha \in C(\alpha)\}$
2. $\langle \alpha, \Phi_\alpha\rangle \leq_\epsilon \langle \beta\Phi_{\underline{e}ta}\rangle$ if $\alpha \leq \beta$ and for every variable assignment s, $V_{\alpha,s,\Phi_\alpha} \subseteq V_{\beta,s,\Phi_\beta}$
3. $D_\epsilon(\langle \alpha, \Phi_\alpha\rangle) = D(\alpha)$.

An epsilon term is interpreted by the valuation of terms of a node if and only if the corresponding existential formula is forced at that node. As a consequence, if $\langle \epsilon x : \varphi, \alpha\rangle \in dom(\Phi_\alpha)$ and $\epsilon x : \psi$ is a term occurring in φ, then $\langle \epsilon x : \psi, \alpha\rangle \in dom(\Phi_\alpha)$.

Notice the usual simultaneous recursion: the definition of the term valuation requires the definition of the forcing relation and vice versa.

4.27. LEMMA. (Expansion Lemma) *Every standard Kripke model can be expanded to an intuitionistic Kripke model.*

Proof: This follows by a simple extension of the argument analogous to the one given for classical models in Chapter 2, now for every node separatedly. We proceed in stages. We have a sequence of languages L^k, where L^0 is the language without ϵ-terms and L^{n+1} is L^n together with all ϵ-terms over L^n. The basic model $K^0_\epsilon = K$ interprets L^0. The model K^{n+1}_ϵ is created from K^n_ϵ by constructing the set $C^{n+1}(\alpha)$ for every node α: we add to every $\Phi_\alpha \in C^n(\alpha)$ all tuples $\langle\langle \epsilon x : \psi, s\rangle, m\rangle$ such that $K^n_\epsilon, s(m|x), \Phi_\alpha \Vdash \varphi$, thus creating from Φ_α a set of new partial choice functions, now interpreting L^{n+1}. The model K_ϵ is formed taking the union of the models created at all stages. ⊠

Digression: Modal Predicate Logic

It may be of help to note an analogy with modal predicate logic in what follows. (Cf. Hughes & Cresswell [HC84].) By the well-known Gödel Translation, intuitionistic predicate logic may be faithfully embedded inside quantified modal S4. In particular, intuitionistic implications $\varphi \to \psi$ become modalized implications $\Box(\varphi \to \psi)$ and universal quantifications $\forall x\varphi$ become modalized formulas $\Box\forall x\varphi$. (Here, existential quantifiers remain as they are, though, and so do conjunctions and disjunctions.) In this setting, epsilon terms play a role analogous to 'non-rigid individual constants' in modal predicate logic, whose interpretation is allowed to vary across different worlds. It is well-known that this necessitates changes in the base system, both in the semantics and the proof theory. The general reason is that these non-rigid constants can be 'captured' by modal operators bearing scope over them, which changes their behavior.

Semantically, this difference shows in the failure of the basic Substitution Lemma of standard first-order logic. It is no longer equivalent to say that $M, w, s \models \varphi[c/x]$ and $M, w, s(V_{M,w}(c)|x) \models \varphi$. For, inside φ, the substitution may relate the evaluation of c to other worlds than w. This difference also shows in a predicate-logical axiom, whose usual verification hinges on the Substitution Lemma. Existential Generalization is no longer valid. For instance, $\Box\varphi[c/x]$ can be true at a world without $\exists x\Box\varphi$ being true. No single object right now needs to witness the possibly different choices for c underlying the true modal statement. "A winner always wins", but there need not be anyone right here who always wins. In the usual modal completeness proofs, this problem is solved by restricting Existential Generalization to those statements which lack non-rigid constants in the scope of modalities. A similar solution is possible here. This modal analogy can probably be turned into an embedding argument extending the Gödel Translation, taking intuitionistic epsilon calculus into a suitable version of modal S4-style predicate logic with non-rigid constants satisfying suitable axioms. We leave this technical connection as an open question here.

4.3.3 Proof Calculus

In line with standard intuitionistic completeness proofs, we set up an axiomatic calculus matching the above semantics. Turning this calculus into a complete natural deduction formulation seems a matter of routine.

4.28. DEFINITION. (Minimal Intuitionistic Epsilon Calculus) The *minimal intuitionistic epsilon calculus* is given by the following proof rules.

1. The propositional rules of IPL.
2. The quantifier rules of IPL, but with the following restriction on Existential Generalization.

   In any application of ($\exists$I),

   $$\frac{\varphi[t/x]}{\exists x\varphi}$$

   if t is an epsilon term, then φ is constructed from only $\wedge, \vee$ and $\exists$.
3. The following two rules for epsilon terms:

$$\epsilon\text{R1}\ \frac{\exists x\varphi}{\exists x(\varphi \wedge \bigwedge_{i\leq n}(\psi_i[\epsilon x : \varphi/x] \leftrightarrow \psi_i))} \qquad \epsilon\text{R2}\ \frac{\psi[\epsilon x : \varphi/x]}{\exists x\varphi}$$

provided ψ is constructed from $\wedge, \exists$

4.29. EXAMPLE. (Some Derivations) We derive a few theorems of this calculus.

1. First of all, the epsilon schema

$$\vdash \exists x\varphi \to \varphi[\epsilon x : \varphi/x].$$

This we get by ϵR1 where we set $i = 1$ and $\psi_1 = \varphi$. This gives us $\exists x\varphi \to \exists x(\varphi \wedge (\varphi[\epsilon x : \varphi/x] \leftrightarrow \varphi))$, which gives us $\exists x\varphi \to \exists x(\varphi \wedge \varphi[\epsilon x : \varphi/x])$. This gives $\exists x\varphi \to \exists x\varphi[\epsilon x : \varphi/x]$. Because x does not occur free in $\varphi[\epsilon x : \varphi/x]$ the epsilon schema follows.

2. For ψ built from $\wedge$ and $\exists$

$$\psi[\epsilon x : \varphi/x] \to \exists(\psi \wedge \varphi).$$

By ϵR2 we have $\psi[\epsilon x : \varphi/x] \to \exists x\varphi$. Using now ϵR1, with $i = 1$ and $\psi = \psi_1$, this gives $\exists x(\varphi \wedge (\psi[\epsilon x : \varphi/x] \leftrightarrow \varphi))$ under assumption $\psi[\epsilon x : \varphi]$. By applying ($\exists$E) we get $\varphi[a/x] \wedge (\psi[\epsilon x : \varphi/x] \leftrightarrow \varphi[a/x])$ for some term a. Still under assumption $\psi[\epsilon x : \varphi/x]$ this gives $\varphi[a/x] \wedge \psi[a/x]$ and, consequently $\exists x(\varphi \wedge \psi)$.

3. For all occurrences of $\epsilon x : \varphi$

$$\exists x\varphi \wedge \psi[\epsilon x : \varphi/x] \to \exists x(\varphi \wedge \psi).$$

This we get by the same derivation as the previous theorem, but now we do not need ϵR2 to derive $\exists x\varphi$. We may use ($\wedge$E) instead, thus avoiding any restriction on the structure of ψ.

4. If ψ and χ are constructed only from $\wedge$ and $\exists$, and $\epsilon x : \varphi$ occurs in both ψ and ψ, then

$$(\varphi \vee \psi) \to \exists x\varphi.$$

This follows simply by using the rule ($\vee$E). Notice that if $\epsilon x : \varphi$ does not occur in both conjuncts, $\exists x\varphi$ does not follow. The restriction on rule ϵR2 cannot allow disjunction (even though this has no 'modal' character), because $P(\epsilon x : Qx)$ derives $P(\epsilon x : Qx) \vee R(\epsilon x : \varphi)$ for arbitrary φ. So $P(\epsilon x : Qx)$ would allow us to derive $\exists x\varphi$ for arbitrary φ if we allow the premise of rule ϵR2 to be constructed from $\wedge, \vee$ and $\exists$.

Without the restriction on Existential Generalization, we could derive Plato's Law (Chapter 3) from principle 1 above. Quod non, by the following Proposition.

4.30. PROPOSITION. (Soundness) *The above calculus is sound for partial intuitionistic epsilon models.*

Proof: By a simple inspection of the rules. We shall only consider the epsilon rules.

Suppose for some node α, $K_\epsilon, \alpha, s, \Phi_\alpha \Vdash \exists x\varphi$. Then $\epsilon x : \varphi$ is assigned a value at $\langle \alpha, \Phi_\alpha \rangle$ and this value is persistent over all $\leq_\epsilon$-accessible nodes. So there is some element m of the domain, namely $V_{\alpha,s,\Phi_\alpha}(\epsilon x : \varphi)$, such that at all accessible nodes, if this m satisfies both $\varphi(x)$ and $\psi(x)$, then the value of $\epsilon x : \varphi$ satisfies them both, and if $\epsilon x : \varphi$ satisfies $\psi(x)$ at such a node, then m satisfies both $\varphi(x)$ and $\psi(x)$. So $K_\epsilon, \alpha, s, \Phi_\alpha \Vdash \exists x(\varphi \wedge (\psi[\epsilon x : \varphi/x] \leftrightarrow \psi))$.

Suppose for some node α, $K_\epsilon, \alpha, s, \Phi_\alpha \Vdash \psi[\epsilon x : \varphi/x]$. If ψ is built only from $\wedge$ and $\exists$ then, by the forcing definition, this reduces to evaluation of atomic formulas containing $\epsilon x : \varphi$ at node $\langle \alpha, \Phi_\alpha \rangle$. But the fact that the formula is evaluated at all, means that V_{α,s,Φ_α} is defined over $\epsilon x : \varphi$ and this can only be if $\exists x \varphi$ is forced. ⊠

4.3.4 Completeness Proof

The completeness proof for this calculus follows a standard pattern. For convenience, we reproduce an outline of the Aczel-Thomason construction for basic intuitionistic logic, which is a Henkin-type modal argument, using *prime theories*, rather than the usual ultrafilters. (Cf. Troelstra & van Dalen [TD88].)

4.31. DEFINITION. (Prime Theories) A *prime theory* in a language L is a set Γ of L-sentences satisfying the following three demands¿

1. If $\Gamma \vdash \varphi$, then $\varphi \in \Gamma$.
2. If $\varphi \vee \psi \in \Gamma$, then $\varphi \in \Gamma$ or $\psi \in \Gamma$.
3. If $\exists x \varphi \in \Gamma$, then $\varphi(t) \in \Gamma$ for some individual constant c.

Such prime theories will be the nodes in the Henkin model. Now, given a set of closed L-formulas Σ and a closed L-formula ψ such that $\Sigma \nvdash \psi$ it is standard to show that there is a prime theory, in an expanded language, Γ extending Σ such that $\Gamma \nvdash \psi$. We fix an enumeration of all existential formulas and disjunctions in the expanded language and construct the desired Γ in stages where we show that, at stage Γ_i, if $\Gamma_i \vdash \exists x \varphi$ we can add $\varphi[c/x]$ to Γ_i for some fresh constant c and still have $\Gamma_i \cup \{\varphi[c/x]\} \nvdash \psi$. And if $\Gamma_i \vdash \chi \vee \xi$, we can add one of the disjuncts with the same result. In the first case the proof involves an application of ($\exists$E): because $\Gamma_i \vdash \exists x \varphi$ and c is a fresh constant, if $\Gamma_i \cup \{\varphi[c/x]\} \vdash \psi$, then we can discharge $\varphi[c/x]$ in an application of ($\exists$E), implying that we already have $\Gamma_i \vdash \psi$. In the second case, the proof goes in the same way, now using ($\vee$E). So in the standard set-up, if $\Sigma \nvdash \psi$ we can construct a prime theory Γ extending Σ such that $\Gamma \nvdash \psi$. This is a node in the Henkin model and the *Truth Lemma* then guarantees that we have a Kripke model with node Γ forcing Σ but not forcing ψ. Consequently, we have shown $\Sigma \not\models \psi$.

4.32. REMARK. In classical logic we can extend a consistent theory Σ by *witness axioms* $\exists x \varphi \rightarrow \varphi[c/x]$ in order to make it witnessing. This is not possible in the intuitionistic case. For, if we proceed to make prime theories by adding witness axioms, then every prime theory will contain all instances of $\exists x(\exists x \varphi \rightarrow \varphi)$ (by the rule ($\exists$I) and deductive closure). Consequently Plato's principle $\exists x(\exists x \varphi \rightarrow$

φ) will be universally valid on the Henkin model. But this is not intuitionistically valid.

Let us analyze this argument with a view towards our richer epsilon models. The propositional base steps are completely independent of epsilon expansions, and hence they go through for our richer language. (They do presuppose persistence, but we have taken care of this.) Quantifier steps are also independent but note that they presuppose the availability of rigid individual constants, in the argument for both the existential quantifier (witnesses were rigid constants) and the universal one (fresh rigid constants were needed to invoke the 'constant lemma').

Now we set up the Henkin model as follows.

The *nodes* of our model will be constructed from prime theories Γ satisfying: if $\Gamma \vdash \exists x\varphi$, then there is a constant c in L_Γ such that $\Gamma \vdash (\varphi[c/x] \wedge \bigwedge_{i\leq n}(\chi_i[\epsilon x : \varphi/x] \leftrightarrow \chi_i[c/x]))$ for all sets $\{\chi_i \in L_\Gamma \mid i \leq n\}$, $n < \omega$.

Let $\Gamma^{f.o.}$ be the restriction of prime theory Γ to ϵ-free formulas. Notice that $\Gamma^{f.o.}$ is a prime theory in $L_\Gamma^{f.o}$, the ϵ-free fragment of L_Γ. After all, if $\exists x\varphi$ has no ϵ-terms and $\Gamma^{f.o.} \vdash \exists x\varphi$, then $\varphi[c/x] \in \Gamma$ is without ϵ-terms for the constant c guaranteed to exist. So $\Gamma^{f.o.} \vdash \varphi[c/x]$. Consequently, with every prime theory Γ we can identify the set $[\Gamma] = \{\Delta \mid \Gamma^{f.o} = \Delta^{f.o}\}$. This set will constitute a node in the underlying *standard* Kripke model.

For each prime theory Γ, let Φ_Γ be a partial function from the set of *closed* ϵ-terms of L_Γ into the set of individual constants of L_Γ such that

1. $\epsilon x : \varphi \in dom(\Phi_\Gamma)$ if and only if $\Gamma \vdash \exists x\varphi$,
2. $\Phi_\Gamma(\epsilon x : \varphi) = c$ only if $\Gamma \vdash (\varphi[c/x] \wedge \bigwedge_{i\leq n}(\chi_i[\epsilon x : \varphi/x] \leftrightarrow \chi_i[c/x]))$ for all sets $\{\chi_i \in L_\Gamma \mid i \leq n\}$, $n < \omega$.

For $[\Gamma]$ a node in the standard Kripke model, the set of choice functions $C([\Gamma])$ over node $[\Gamma]$ will be the set $\{\Phi_\Delta \mid \Delta \in [\Gamma]\}$.

Now we define the Henkin model to be $K_\epsilon = \langle W_\epsilon, \leq_\epsilon, D_\epsilon\rangle$ where

1. $W_\epsilon = \{\langle[\Gamma], \Phi_\Delta\rangle \mid \Delta \in [\Gamma]\}$,
2. $\langle[\Gamma], \Phi_\Delta\rangle \leq_\epsilon \langle[\Gamma'], \Phi_{\Delta'}\rangle$ if $\Gamma^{f.o.} \subseteq \Gamma'^{f.o}$ and $\Phi_\Delta \subseteq \Phi_{\Delta'}$,
3. D_ϵ assigns to each $[\Gamma]$ the constant model of Γ.

If we now set for closed atomic formulas φ that $K_\epsilon, \langle[\Gamma], \Phi_\Delta\rangle \Vdash \varphi$ if and only if $\varphi \in \Delta$, then we can prove the *truth lemma* in the standard way.

4.33. LEMMA. (Truth Lemma) *For every node $\langle[\Gamma], \Phi_\Delta\rangle$ in the Henkin model K_ϵ and every closed formula φ in L_Γ:*

$$K_\epsilon, \langle[\Gamma], \Phi_\Delta\rangle \Vdash \varphi \iff \varphi \in \Delta.$$

Orthogonally to the truth lemma, we have to check that the model obtained is in our intended class of partial intuitionistic epsilon models. In one direction, this is clear. Whenever an existential statement holds, we have an epsilon witness, which will continue to work in all consistent witnessing splitting extensions. In the opposite direction, however, a supplementary argument is needed. We must also show that epsilon terms are interpreted at a world only when their corresponding existential 'supports' are true. This is the task of rule ϵR2. Whenever we have an atomic statement in a set containing an epsilon term, this will imply the latter's existential support.

Now we can prove the completeness theorem.

4.34. THEOREM. (Completeness Theorem) For all sets of formulas $\Sigma \cup \{\psi\}$ if $\Sigma \models \psi$, then $\Sigma \vdash \psi$.

Proof: We proceed as usual by contraposition. So suppose $\Sigma \not\vdash \psi$. Because Σ does not derive ψ we know Σ is consistent. Starting from a consistent Σ we construct a prime theory Γ extending Σ, satisfying $\Gamma \not\vdash \psi$, and a partial choice function Φ_Γ. This will give us a node in the Henkin model which, by the truth lemma, forces Σ without forcing ψ.

As usual we proceed in stages guided by an enumeration of the formulas in the language of Σ expanded by an infinite set of fresh individual constants. We set $\Gamma^0 = \Sigma$ and $\Phi_\Gamma = \emptyset$. So $\Gamma^0 \not\vdash \psi$. Now suppose we have constructed Γ^n such that $\Gamma^n \not\vdash \psi$ and Φ_{Γ^n}. Suppose the $n+1$'th element of our enumeration is the formula $\chi \vee \xi$, then we proceed standardly by adding one of the disjuncts. On the other hand, if this formula is of the form $\exists x\varphi$ and $\Gamma^n \vdash \exists x\varphi$. Then we set

1. $(\varphi[c/x] \wedge \bigwedge_{i\leq n}(\chi_i[\epsilon x : \varphi/x] \leftrightarrow \chi_i[c/x])) \in \Gamma^{n+1}$ for all sets $\{\chi_i \mid i \leq n\}$, $n < \omega$, and c some fresh constant.
2. $\langle \epsilon x : \varphi, c\rangle \in \Phi_{\Gamma^{n+1}}$ where c is the constant chosen in (1).

4.35. CLAIM. $\Gamma^{n+1} \not\vdash \psi$.

Proof: Suppose the contrary. That is, there is some set of χ_i, $i \leq k$, such that

$$\Gamma^n \cup \{\varphi[c/x] \wedge \bigwedge_{i\leq k}(\chi_i[\epsilon x : \varphi/x] \leftrightarrow \chi_i)\} \vdash \psi.$$

Notice that the constant c does not occur in Γ^n or ψ. Now, Γ^n derives $\exists x\varphi$, so by ϵR1, Γ^n derives $\exists x(\varphi \wedge \bigwedge_{i\leq k}(\chi_i[\epsilon x : \varphi/x] \leftrightarrow \chi_i))$. Consequently, because c does not occur in Γ^n or ψ, we can derive ψ from Γ^n by an application of ($\exists$E) (discharging $\varphi[c/x] \wedge \bigwedge_{i\leq k}(\chi_i[\epsilon x : \varphi/x] \leftrightarrow \chi_i)$). This contradicts our assumption, so $\Gamma^{n+1} \not\vdash \psi$. ⊠

4.3.5 Additional Principles

There may be something surprising about our epsilon calculi. We mean the absence of any rules manipulating the internal structure of epsilon terms.

Of course, one principle to this effect would be Extensionality. This was not used in the above intensional semantics (as we only require that $\Phi_\alpha(\langle \epsilon x : \varphi, s\rangle) \in$

$\{m \in dom(D(\alpha)) \mid K, \alpha, s(m|x), \Phi_\alpha \Vdash \varphi\}$ for $\langle \epsilon x : \varphi, s\rangle \in dom(\Phi_\alpha))$, but it could easily be added. But this is still close to nothing.

What we have in mind are rather Monotonicity and Distribution rules like the following:

$$\frac{\varphi(\epsilon x : (\psi \wedge \chi)}{\varphi(\epsilon x : \psi)} \qquad \frac{\varphi(\epsilon x : (\psi \vee \chi)}{\varphi(\epsilon x : \psi) \vee \varphi(\epsilon x : \chi)}$$

expressing intuitively "what holds for some $\psi \wedge \psi$ holds for some ψ" and "what holds for some $\psi \vee \chi$ holds for some ψ or some χ". These principles are not valid in our semantics, nor in any semantic we have considered in the previous chapters. Moreover, they are not derivable in the proof system of Section 2. This follows, because the theorems of this system form a subset of the theorems that are classically derivable (we only exclude some derivations) and we can use the completeness proof for the classical epsilon calculus to construct counterexamples. It is not quite clear how principles like Monotonicity and Distributivity could be validated, unless one adds a new aspect to our semantic modeling, namely, further correlations between choices of witnesses. We shall not pursue this here — but note that it will return in Chapter 6, when we discuss linguistic applications to indefinites. What we shall also find there is a delicate balance between 'epsilon logic' and 'representation'. Clearly, the fact that "A blonde cop fired" implies that "A cop fired", but it is not as obvious as might seem at first sight that this really expresses validity of the first principle mentioned above.

As a final observation, we point out that the above calculus, modest as it is, does have one bold feature which distinguishes it from its classical counterpart. In classical epsilon calculus, the following rule is clearly invalid, even for atoms:

$$\frac{P(\epsilon x : Q)}{\exists x(P \wedge Q)}$$

as it would trivialize our models. But here, due to our partial set up, we have managed to validate it. Thus, intuitionistic base logics can conservatively support stronger epsilon principles than their classical counterparts!

4.4 Conclusion

In this chapter we have shown an important difference between assumption management in CPL and IPL: CPL may permute its stack of assumptions freely, while IPL has to keep this stack intact at all times. Awareness of this difference has allowed us to construct a proof theory for intuitionistic epsilon logic which is conservative over IPL. However, semantically we have only shown completeness with respect to a less intuitive epsilon extension of IPL. The creation of an conservative epsilon extension is not a straightforward matter but involves strategic choices each giving its own logic. Instantial reasoning in intuitionistic logic is non-trivial.

4.5 Appendix

In Section 2, we have introduced a proof system for intuitionistic epsilon logic which is conservative over IPL. Below we try to construct a semantics for this system and discuss its completeness. Although we have not quite achieved this goal yet, we think the ideas involved are of some independent interest.

The essential problem of a Kripke semantics for the restricted ϵ-calculus consisted in the question how to treat *ϵ-terms* $\epsilon x : \varphi$ at nodes α such that α does not force $\exists x\varphi$. If we want to interpret the ϵ-term at α we cannot ***persistently*** assign a value, for $\exists x\varphi \vee \forall x\neg\varphi$ is not intuitionistically valid. At α we may choose an element to interpret $\epsilon x : \varphi$ which at some later node β may not satisfy φ although β forces $\exists x\varphi$. Again, we shall employ *Skolem expansions*, but this time, striking out in a direction somewhat different from that of Section 3, allowing non-persistent choices. More precisely, we shall define an interpretation of ϵ-terms over intuitionistic Kripke models by formulating the notion of a *local* Skolem function. The interpretation of an ϵ-term $\epsilon x : \varphi$ over a Kripke model will map this term, at each node, to a Skolem function F_φ over the model associated with that node, *if* $\exists x\varphi$ *is forced.* This value will be persistently assigned to that term. If $\exists x\varphi$ is *not* forced, then an arbitrary *non-persistent* value is assigned. When we then interpret sets of assumptions at a node, we shall restrict our notion of accessibility in such a way that only nodes which treat all ϵ-terms occurring in these assumption persistently are accessible.

The constant Lemma 4.22 tells us that theories using ϵ-terms which have not been 'introduced', may, proof theoretically, be treated as arbitrary constants or free variables. Especially the analogy with free variables is of interest. An ϵ-term occurring 'free' in a theory Σ may become 'bound' in extensions of Σ. That is, at a node α in a Kripke model not forcing $\exists x\varphi$, the ϵ-term $\epsilon x : \varphi$ will be interpreted as a free variable: it is assigned an arbitrary value. However, future nodes forcing $\exists x\varphi$ bind $\epsilon x\varphi$ in the sense that the value cannot be chosen arbitrarily. In these nodes the arbitrary value may have to be changed.

In a standard Kripke model $K = \langle W, \leq, D\rangle$ the valuation of terms at every node α is determined by the associated first-order model $D(\alpha)$. When ϵ-terms have to be interpreted this situation changes. The valuation of terms will depend on the *forcing definition* analogously to the situation in classical logic where the interpretation of ϵ-terms is interconnected with the *truth definition.* This means that we cannot refer the interpretation of ϵ-terms to the valuation function of the first-order model associated with nodes of the model, for it is the relation to other nodes in the model which determines the interpretation of these terms.

The Model

Let $K = \langle W, \leq, D\rangle$ be a standard Krike model. To interpret ϵ-terms we define in simultaneous induction on the dependence level of ϵ-terms at each node α

of K and every variable assignment s an intensional choice function Φ_α and a valuation function V_{α,s,Φ_β}. We set $\mathcal{L}^0$ the ϵ-free fragment of $\mathcal{L}$, and $\mathcal{L}^{n+1} = \mathcal{L}^n \cup \{\epsilon x : \varphi \mid \varphi \in \mathcal{L}^n\}$. Moreover $\mathcal{E}(\varphi)$ will denote the set of all ϵ-terms occurring in φ. For all $\alpha, \beta \in W$ and every variable assignment s:

1. $\Phi^0_\alpha = \emptyset$,
2. $V_{\alpha,s,\Phi^0_\beta} = V_{\alpha,s}$,
3. $K_{\Phi^0}, \langle \alpha, \Phi_\beta \rangle, s \Vdash \varphi \iff K, \alpha, s \Vdash \varphi$.

Now suppose we have defined Φ^n_α, $V_{\alpha,s,\Phi^n_\beta}$, and $K_{\Phi^n}, \langle \alpha, \Phi_\beta \rangle, s \Vdash \varphi$. For every $k < \omega$, let F^k be the set of all k-ary functions from $\bigcup_{\alpha \in W} dom(D(\alpha))$ in $\bigcup_{\alpha \in W} dom(D(\alpha))$ and let f^k be a k-ary function variable. Then for every $\alpha, \beta \in W$

1. $\Phi^n_\beta \subseteq \Phi^{n+1}_\beta$ and for every ϵ-term $\epsilon x : \varphi(x, \overline{y}) \in \mathcal{L}^{n+1}$ where $\overline{y} = \langle y_1 \ldots y_k \rangle$, if

   (a) $\mathcal{E}(\varphi) \subseteq dom(\Phi^n_\beta)$,
   (b) $K_{\Phi^n}, \langle \alpha, \Phi^n_\beta \rangle, s \Vdash \forall \overline{y} \exists x \varphi(x, \overline{y})$,

   then $\{\langle \epsilon x \varphi(x, \overline{y}), F \rangle\} \in \Phi^{n+1}_\beta$ where $K_{\Phi^n}, \langle \alpha, \Phi^n_\beta \rangle, s(f|F) \Vdash \forall \overline{y} \varphi(f^k(\overline{y}), \overline{y})$.
2. $V_{\alpha,s,\Phi^n_\beta} \subseteq V_{\alpha,s,\Phi^{n+1}_\beta}$ and for every $\mathcal{L}^{n+1}$-term t,

$$V_{\alpha,s,\Phi^{n+1}_\beta}(t) = \begin{cases} \Phi^{n+1}_\beta(\epsilon x \varphi(x, \overline{y})(s(\overline{y})) & t = \epsilon x \varphi(x, \overline{y}) \in dom(\Phi^{n+1}_\beta) \\ m \in dom(D(\alpha)) & t \notin dom(\Phi^{n+1}_\beta),\ t \text{ an } \epsilon\text{-term} \\ V_{\alpha,s}(t) & \text{otherwise} \end{cases}$$

3. $K_{\Phi^{n+1}} = \langle W_{\Phi^{n+1}}, \leq_{\Phi^{n+1}}, D \rangle$ where

   (a) $W_{\Phi^{n+1}} = \{\langle \alpha, \Phi^{n+1}_\beta \rangle \mid \alpha, \beta \in W, \beta \leq \alpha\}$,
   (b) $\leq_{\Phi^{n+1}} \subseteq W_\Phi \times W_\Phi$ such that $\langle \alpha, \Phi_\beta \rangle \leq^{n+1}_\Phi \langle \gamma, \Phi_\delta \rangle$ if

      (i) $\alpha \leq \gamma$,

      (ii) $\Phi_\beta \subseteq \Phi_\delta$,

      (iii) $V_{\alpha,s,\Phi_\beta} \upharpoonright (dom(\Phi_\beta) \cup \overline{dom(\Phi_\gamma)}) \subseteq V_{\gamma,s,\Phi_\delta} \upharpoonright (dom(\Phi_\beta) \cup \overline{dom(\Phi_\gamma)})$,
   (c) $D(\langle \alpha, \Phi_\beta \rangle) = D(\alpha)$.
4. For P an atomic formula in $\mathcal{L}^{n+1}$

$$K_{\Phi^{n+1}}, \langle \alpha, \Phi_\beta \rangle, s \Vdash Pt_1 \ldots t_n \iff \langle V_{\alpha,s,\Phi_\beta}(t_1), \ldots, V_{\alpha,s,\Phi_\beta}(t_n) \rangle \in int(D(\alpha))(P).$$

Now we set

$$K_\Phi = \langle \bigcup_{n<\omega} W_{\Phi^n}, \bigcup_{n<\omega} \leq_{\Phi^n}, D \rangle.$$

So nodes of the Kripke model K_Φ consist of pairs of standard nodes α and choice functions Φ_β. These choice functions may have been defined with respect to β

lying below α. We have no general persistence on these models. That is, if $K_\Phi, \pi, s \Vdash \varphi$, then it is possible that $K_\Phi, \pi', s \nVdash \varphi$ for some $\pi \leq_\Phi \pi'$. Notice that $\epsilon x : \varphi \in dom(\Phi_\alpha)$ if α forces $\forall \overline{y} \exists x \varphi$ and Φ_α is defined over all ϵ-terms occurring in φ. At a node with choice function Φ_α all ϵ-terms in the domain of this function have a persistent denotation. The ϵ-terms not in the domain of this function are treated persistently as long as they do not enter the domain of the choice function at an accessible node. That is, the statement

$$\langle \alpha, \Phi_\beta \rangle \leq_\Phi \langle \gamma, \Phi_\delta \rangle$$

holds if $\alpha \leq \beta$, and V_{γ,s,Φ_δ} assigns the same value to all terms in $dom(\Phi_\beta) \cup dom(\Phi_\gamma)$ as V_{α,s,Φ_β} does: an ϵ-term is persistently assigned a value, if it is an element of both $dom(\varphi_\beta)$ and $dom(\Phi_\delta)$ or of neither. So the two valuations may differ on the values they assign to terms in $dom(\Phi_\gamma) - dom(\Phi_\beta)$. Consequently we cannot have full persistence. A simple proposition describes the situation.

4.36. PROPOSITION. *Let $\langle \alpha, \Phi_\beta \rangle \leq_\Phi \langle \gamma, \Phi_\delta \rangle$ and suppose that $\mathcal{E}(\varphi) \cap dom(\Phi_\gamma) \subseteq dom(\Phi_\beta)$, then*

$$K, \langle \alpha, \Phi_\beta \rangle, s \Vdash \varphi \Rightarrow \begin{cases} (1) K, \langle \gamma, \Phi_\delta \rangle, s \Vdash \varphi \\ (2) K, \langle \gamma, \Phi_\gamma \rangle, s \Vdash \varphi \end{cases}$$

Proof: Consequence (1) holds by definition for atomic formulas. The general case then follows by induction on the standard forcing clauses. The second consequence follows, because $dom(\Phi_\gamma) - dom(\Phi_\delta) \subseteq dom(\Phi_\gamma) - dom(\Phi_\beta)$. ⊠

4.37. COROLLARY. *Assuming conditions 1–3 below, $\langle \alpha, \Phi_\beta \rangle \leq_\Phi \langle \gamma, \Phi_\delta \rangle$ implies $K, \langle \alpha, \Phi_\beta \rangle, s \Vdash \varphi \Rightarrow K, \langle \gamma, \Phi_\delta \rangle, s \Vdash \varphi$.*

1. *$\Phi_\beta = \Phi_\delta$. Evaluating with respect to a fixed choice function, gives us full persistence.*
2. *Φ_β is defined over all ϵ-terms occurring in φ. This implies that if for all $\epsilon x : \psi \in \mathcal{E}(\varphi)$ we have $K, \langle \alpha, \Phi_\alpha \rangle, s \Vdash \exists x \psi$, then $K, \langle \alpha, \Phi_\alpha \rangle, s \Vdash \varphi \Rightarrow K, \pi, s \Vdash \varphi$ for all π such that $\langle \alpha, \Phi_\alpha \rangle \leq_\Phi \pi$.*
3. *$\mathcal{E}(\varphi) = \emptyset$. This is a special case of 2. So for all ϵ-free formulas we have full persistence.*

The valuation functions associated with a node need not interpret the ϵ-terms 'correctly' at that node. That is, the following situation may typically occur for $\pi = \langle \alpha, \Phi_\alpha \rangle$, $\pi' = \langle \beta, \Phi_\alpha \rangle$ such that $\pi \leq_\Phi \pi'$:

1. $K_\Phi, \pi, s \nVdash \exists x \varphi$
2. $K_\Phi, \pi', s \Vdash \exists x \varphi$ and $K_\Phi, \pi', s \nVdash \varphi[\epsilon x : \varphi / x]$.

Because $\exists x \varphi$ is not forced at π with respect to s, the valuation function V_{β,s,Φ_α} assigns at β an element of $dom(D(\alpha))$ to $\epsilon x \varphi$. This element need not be contained in $\{d \in dom(D(\alpha)) \mid K_\Phi, \langle\beta, \Phi_\alpha\rangle, s(x|d) \Vdash \varphi\}$. The correct interpretation of ϵ-terms at β is only guaranteed on the *diagonal* node $\langle\beta, \Phi_\beta\rangle$. But this node is not *accessible* from $\langle\alpha, \Phi_\alpha\rangle$ if we interpret $\epsilon x : \varphi$ persistently. The situation at the node $\langle\beta, \Phi_\alpha\rangle$ shows that the *ϵ-axiom schema* is not generally valid on the models as we have defined them. In order to validate the *ϵ-rule* nevertheless, we shall define a notion of logical consequence which always evaluates formulas with respect to *diagonal nodes*. That is, with respect to nodes where the choice function is 'appropriate'.

Persistent Interpretation of Assumptions

Whenever we want to to interpret a set Σ of $\mathcal{L}$-formulas on a model K_φ we intend to evaluate it *persistently*. That is, we shall *stipulate* that all ϵ-terms in the set of assumptions have to be treated persistently. This we do as follows.

4.38. DEFINITION. (E-Persistent Kripke Models) For every set E of ϵ-terms and Kripke model K_Φ, we define an accessibility relation $\leq_\Phi^E$ and an *E-persistent* Kripke model K_Φ^E by

1. $\langle\alpha, \Phi_\beta\rangle \leq_\Phi^E \langle\gamma, \Phi_\delta\rangle$ if $\langle\alpha, \Phi_\beta\rangle \leq_\Phi \langle\gamma, \Phi_\delta\rangle$ and $E \cap dom(\Phi_\gamma) \subseteq dom(\Phi_\beta)$,
2. $K_\Phi^E = \langle W_\Phi, \leq_\Phi^E, D\rangle$.

So for a given formula φ, the model $K_\Phi^{\mathcal{E}(\varphi)}$ interprets φ persistently. Notice that, if $\mathcal{E}(\varphi) \subseteq dom(\Phi_\beta)$, then $\leq_\Phi^{\mathcal{E}(\varphi)} = \leq_\Phi$. In particular, the relation $\leq_\Phi^{\mathcal{E}(\varphi)}$ coincides with $\leq_\Phi$ for ϵ-free φ. On E-persistent Kripke models we define the forcing relation as follows.

4.39. DEFINITION. (E-Persistent Forcing) For E a set of ϵ-terms, we have the following forcing clauses:

1. $K_\Phi^E, \pi, s \Vdash \neg\varphi[g] \iff \forall \pi' \geq_\Phi^{E\cup\mathcal{E}(\varphi)} \pi : \ K_\Phi^{E\cup\mathcal{E}(\varphi)}, \pi', s \nVdash \varphi$,
2. $K_\Phi^E, \pi, s \Vdash \varphi \wedge \psi \iff K_\Phi^E, \pi, s \Vdash \varphi \ \&\ K_\Phi^E \pi, s \Vdash \psi$,
3. $K_\Phi^E, \pi, s \Vdash \varphi \vee \psi \iff K_\Phi^E, \pi, s \Vdash \varphi$ or $K_\Phi^E, \pi, s \Vdash \psi$,
4. $K_\Phi^E, \pi, s \Vdash \varphi \rightarrow \psi \iff$
   $$\forall \pi' \geq_\Phi^{E\cup\mathcal{E}(\varphi)} \pi : \ K_\Phi^{E\cup\mathcal{E}(\varphi)}, \pi', s \Vdash \varphi \Rightarrow K_\Phi^{E\cup\mathcal{E}(\varphi)}, \pi', s \Vdash \psi,$$
5. $K_\Phi^E, \pi, s \Vdash \forall x \psi \iff \forall \pi' \geq_\Phi^E \pi,\ \forall m \in dom(D(\pi')) : \ K_\Phi^E, \pi', s(m|x) \Vdash \varphi$,
6. $K_\Phi^E, \pi, s \Vdash \exists x \varphi \iff \exists m \in dom(D(\pi)) : \ K_\Phi^E, \pi, s(m|x) \Vdash \varphi$.

Notice especially the clauses for negation and implication. We evaluate an implication $\varphi \rightarrow \psi$, at π E-persistently by moving $E \cup \mathcal{E}(\varphi)$-persistent to some node π' forcing φ and then evaluating persistent with respect to the extended set $E \cup \mathcal{E}(\varphi)$. This is required for the notion of logical consequence we shall define: if we add a fresh assumption φ to our set of assumptions Σ, then this new assumptions must also be interpreted persistently.

The move from K_Φ to K_Φ^E changes the accessibility relation. The extend of this change is given by the following proposition.

4.40. PROPOSITION. *If* $\langle\alpha,\Phi_\beta\rangle \leq_\Phi \langle\gamma,\Phi_\delta\rangle$, *then* $\langle\gamma,\Phi_\delta\rangle$ *is* $\leq^E_\Phi$*-accessible to* $\langle\alpha,\Phi_\beta\rangle$ *if for all* $\epsilon x:\varphi \in E$, $K^E_\Phi\langle\gamma,\Phi_\gamma\rangle$, $\Vdash\exists x\varphi$ *implies* $K^E_\Phi,\langle\alpha,\Phi_\beta\rangle, s \Vdash\exists x\varphi$.

Proof: If for all $\epsilon x:\varphi \in E$, $K^E_\Phi,\langle\gamma,\Phi_\gamma\rangle \Vdash\exists x\varphi$ implies $K^E_\Phi,\langle\alpha,\Phi_\beta\rangle, s \Vdash\exists x\varphi$, then $E \cap (dom(\Phi_\gamma)) \subseteq dom(\Phi_\beta)$. ⊠

Now we define our notion of logical validity on with respect to models which treat the assumptions persistently, and interpret all ϵ-terms appropriate to the point of evaluation.

4.41. DEFINITION. (ϵ-Consequence) Let $\Sigma \cup \{\varphi\}$ be a set of $\mathcal{L}$-formulas. Formula φ is an *ϵ-consequence* of Σ, notation $\Sigma \Vdash_\epsilon \varphi$ if for all Kripke models $K = \langle W, \leq, D\rangle$, all $\alpha \in W$, all choice functions Φ and all variable assignments s

$$K^{\mathcal{E}(\Sigma)}_\Phi, \langle\alpha,\Phi_\alpha\rangle, s \Vdash\Sigma \Rightarrow K^{\mathcal{E}(\Sigma)}_\Phi, \langle\alpha,\Phi_\alpha\rangle, s \Vdash\varphi.$$

As was mentioned, for a proof of soundness and completeness of this semantics with respect to the prof notion $\vdash^I_i$, we refer the reader to work in progress. We shall conclude this section with two lemma's which are characteristic for this semantics.

4.42. LEMMA. (Epsilon Lemma) *For all existential formulas* $\exists x\varphi$

$$\exists x\varphi \Vdash_c \varphi[\epsilon x:\varphi/x].$$

Proof: Suppose $K^{\mathcal{E}(\exists x\varphi)}_\Phi, \langle\alpha,\Phi_\alpha\rangle, s \Vdash\exists x\varphi$. Because Φ_α is defined at α we have

$$V_{\alpha,s,\Phi_\alpha}(\epsilon x:\varphi) \in \{m \in dom(D(\alpha)) \mid K^{\mathcal{E}(\exists x\varphi)}_\Phi, \langle\alpha,\Phi_\alpha\rangle, s(x|m) \Vdash\varphi\}.$$

Consequently, $K^{\mathcal{E}(\exists x\varphi)}_\Phi, \langle\alpha,\Phi_\alpha\rangle, s \Vdash\varphi[\epsilon x:\varphi/x]$. ⊠

The next lemma should be compared to Proposition 4.23.

4.43. LEMMA. (ϵ-Deduction Lemma) *Let* $\Sigma\cup\{\varphi,\psi\}$ *be a set of* $\mathcal{L}$*-formulas such that for all* $\epsilon x:\xi \in \mathcal{E}(\Sigma,\psi)$

$$\Sigma \Vdash_\epsilon \exists x\xi \Rightarrow \Sigma,\varphi \Vdash_\epsilon \exists x\xi.$$

Then

$$\Sigma,\varphi \Vdash_\epsilon \psi \Rightarrow \Sigma \Vdash_\epsilon \varphi \rightarrow \psi.$$

Proof: Assume $\Sigma,\varphi \Vdash_\epsilon\psi$. That is

for all $\langle\alpha,\Phi_\alpha\rangle \in W_\Phi$, if $K^{\mathcal{E}(\Sigma,\varphi)}_\Phi\langle\alpha,\Phi_\alpha\rangle, s \Vdash\Sigma,\varphi$ then $K^{\mathcal{E}(\Sigma,\varphi)}_\Phi\langle\alpha,\Phi_\alpha\rangle, s \Vdash\psi$.

Now suppose for some node $\langle\alpha,\Phi_\alpha\rangle$, $K^\Sigma_\Phi, \langle\alpha,\Phi_\alpha\rangle, s \Vdash\Sigma$, and for some $\langle\beta,\Phi_\gamma\rangle$ such that $\langle\alpha,\Phi_\alpha\rangle \leq^{\mathcal{E}(\Sigma,\varphi)}_\Phi \langle\beta,\Phi_\gamma\rangle$ we have

$$K^{\mathcal{E}(\Sigma,\varphi)}_\Phi, \langle\beta,\Phi_\gamma\rangle, s \Vdash\varphi.$$

We need to show:

$$K_\Phi^{\mathcal{E}(\Sigma,\varphi)}, \langle \beta, \Phi_\gamma \rangle, s \Vdash \psi.$$

We proceed in two steps.

1. We first show that $K_\Phi^{\mathcal{E}(\Sigma,\varphi)}, \langle \beta, \Phi_\beta \rangle, s \Vdash \psi$. We have done this if we can show that $\langle \beta, \Phi_\gamma \rangle \leq_\Phi^{\mathcal{E}(\Sigma,\varphi)} \langle \beta, \Phi_\beta \rangle$. For this gives us, by persistence,

$$K_\Phi^{\mathcal{E}(\Sigma,\varphi)}, \langle \beta, \Phi_\beta \rangle, s \Vdash \Sigma, \varphi.$$

By assumption we then have $K_\Phi^{\mathcal{E}(\Sigma,\varphi)}, \langle \beta, \Phi_\beta \rangle, s \Vdash \psi$.
The accessibility of node $\langle \beta, \Phi_\beta \rangle$ follows by the condition of the lemma and Proposition 4.40. Suppose $K_\Phi^{\mathcal{E}(\Sigma,\varphi)}, \langle \beta, \Phi_\beta \rangle, s \Vdash \exists x \xi$, where $\epsilon x : \xi \in \mathcal{E}(\Sigma, \varphi)$. Because $\langle \beta, \Phi_\beta \rangle$ forces Σ, φ, the condition of the lemma guarantees that $K_\Phi^{\mathcal{E}(\Sigma)}, \langle \alpha, \Phi_\alpha \rangle, s \Vdash \exists x \xi$, because $K_\Phi^{\mathcal{E}(\Sigma)}, \langle \alpha, \Phi_\alpha \rangle, s \Vdash \Sigma$. But then $K_\Phi^{\mathcal{E}(\Sigma)}, \langle \beta, \Phi_\gamma \rangle, s \Vdash \exists x \xi$, and, by Proposition 4.40, $\langle \beta, \Phi_\gamma \rangle \leq_\Phi^{\mathcal{E}(\Sigma,\varphi)} \langle \beta, \Phi_\beta \rangle$.
2. Secondly we show that $K_\Phi^{\mathcal{E}(\Sigma,\varphi)}, \langle \beta, \Phi_\beta \rangle, s \Vdash \psi$ implies $K_\Phi^{\mathcal{E}(\Sigma,\varphi)}, \langle \beta, \Phi_\gamma \rangle, s \Vdash \psi$. This we do analogously to the previous proof step. Now the condition of the lemma guarantees that every $\exists x \xi$ forced at $\langle \beta, \Phi_\beta \rangle$, for $\epsilon x : \chi \in \mathcal{E}(\psi)$, is already forced at $\langle \beta, \Phi_\gamma \rangle$. Consequently V_{β,s,Φ_β} and V_{β,s,Φ_γ} agree on all terms in ψ. ⊠

The deduction lemma guarantees that we can conditionalize whenever all ϵ-terms in assumptions and conclusion retain their status of being defined or undefined.

4.44. EXAMPLE. We shall give some examples of situations in which the condition of Lemma 4.43 is not satisfied. For all cases of non-forcing we can find counterexamples.

$P(\epsilon x : Q), \exists x Q \Vdash_\epsilon \exists x (P \wedge Q)$	$P(\epsilon x : Q) \not\Vdash_\epsilon \exists x Q \to \exists x (P \wedge Q)$
$\exists x P, \neg P(\epsilon x : P) \Vdash_\epsilon \bot$	$\neg P(\epsilon x : P) \not\Vdash_\epsilon \exists x P \to \bot$
$\exists x P \Vdash_\epsilon P(\epsilon x : P)$	$\not\Vdash_\epsilon \exists x P \to P(\epsilon x : P)$

For the first non-equivalence, notice that $P(\epsilon x : Q), \exists x Q \Vdash_\epsilon \exists x Q$ while $P(\epsilon x : Q) \not\Vdash_c \exists x Q$ for $\epsilon x : Q \in \mathcal{E}(P(\epsilon x : Q), \exists x (P \wedge Q))$. So we cannot carry through the first proof step in the deduction lemma. By the same argument we have the second non-equivalence. For the third, note that $\exists x P \Vdash_\epsilon \exists x P$ while $\emptyset \not\Vdash_\epsilon \exists x P$ for $\epsilon x : P \in \mathcal{E}(P(\epsilon x : P))$. Consequently, we cannot carry through the second proof step of the deduction lemma.

Figure 4.1 gives a counterexample to the validity of $\exists x \varphi \to \varphi[\epsilon x : \varphi / x]$. There we see the domains of two nodes, $\langle \alpha, \Phi_\alpha \rangle$ and $\langle \beta, \Phi_\beta \rangle$, with $V_{\alpha,s,\Phi_\alpha}(\epsilon x : P) = \{1\}$. So $K_\Phi, \langle \beta, \Phi_\alpha \rangle, s \Vdash \exists x P$, but $K_\Phi, \langle \beta, \Phi_\alpha \rangle, s \not\Vdash P[\epsilon x : P/x]$. Moreover,

$$K_\Phi, \langle \alpha, \Phi_\alpha \rangle, s \Vdash \neg\neg \exists x P \wedge \neg P[\epsilon x : P/x].$$

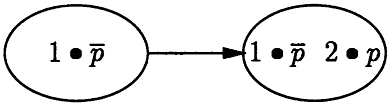

Figure 4.1: Counterexample to the ϵ-axiom schema

We can also consider node $\langle\beta, \Phi_\beta\rangle$, then we have at that node both $\exists x P$ and $P[\epsilon x : P/x]$, but $\langle\alpha, \Phi_\alpha\rangle \not\leq_\Phi \langle\beta, \Phi_\beta\rangle$.

Counterexamples witnessing the remaining two non-consequences can be constructed analogously. The trick is always to start at a node where the existential formula corresponding to the ϵ-term in the assumption is not forced. This gives us complete freedom to create a node falsifying the conclusion.

Chapter 5

Term Dependencies

This chapter explores dependencies between terms arising within derivations. It will present no definite results, but will discuss a variety of ways in which term dependencies can be made explicit. In the first sections we discuss some sources of dependence and mention connections between these sources. We shall deal mainly with derivation in natural deduction. The reasons for this choice are discussed in the third section. Our interest in term dependencies will lead us to define the notion of a choice process. This notion will be our guiding principle throughout the remainder of this chapter. In general, classical logic does not take term dependence seriously (an exception being, for instance, the proof of Herbrands Theorem). We shall investigate the substructure of term dependence in a number of logics which respect dependence in various ways. This will be the subject of the remaining sections of this chapter. We conclude with a brief discussion of a possible semantics for term dependencies in general.

5.1 Dependence as a Logical Parameter

Sequences of operators may express dependencies, as exemplified in the well-known logical phenomenon of scope. In particular, sequences of successive quantifiers may exhibit dependencies, a prime example being $\forall\exists$ combinations for functional dependencies between choices of objects. Some of these dependencies are expressible in standard logic, witness its account of operator scope ambiguities, or its use of Skolem functions. But intuitively, dependencies may even arise in quantifier combinations like $\exists x \exists y$ where the second object y may be chosen depending on the choice of the first object x. In classical logic, this phenomenon cannot be modeled, since this sequence is equivalent with $\exists y \exists x$. In the semantic literature on generalized quantifiers (De Mey [Mey90], Zimmermann [Zim93], Keenan & Westerstahl [KW95]), people have also been interested in the phenomenon of independence, which may occur despite the linear surface order of operators imposed by natural languages.

Various authors have stressed the central importance of the phenomenon of dependence. An early example is the semantic framework of Hintikka [Hin73], [HK83], whose game theory owes many of its more deviant aspects to a desire to model linguistic dependencies in discourse and reasoning. Game theory is indeed a natural mathematical paradigm for modeling dependencies, as we can deal with them in terms of 'prior information' available to, or hidden from, players at a certain stage of the game. A second example is the arbitrary object theory of Fine, discussed in our Chapter 2. This is essentially a semantic account of dependence, viewed as constraints on possible values for related objects. Finally, dependence has again emerged in the treatment of individual variables in so-called 'generalized assignment models' (Németi [Ném93]), where the absence of certain assignments (out of the total space DVAR) may force variable 'registers' to co-vary in their assigned objects. A related approach is the dependence semantics proposed in Alechina & van Benthem [BA93], following ideas by van Lambalgen [Lam91]. A semantics in a similar vein may be found in van der Does [Doe95].

Our preferred perspective here will be the phenomenon of dependence as it arises in proofs, in particular, in the natural deduction format. Occasionally, we shall also link up with more semantic approaches. Our key intuition in what follows is that of the ways in which we perform choice of witnesses.

5.2 Dependence in Proofs

5.2.1 Sources

Dependencies naturally arise in the course of derivations between propositions, between propositions and terms, and between terms themselves. We give some examples of all three.

Dependencies between propositions In Section 4.1.2 we have seen a variety of dependencies between assumptions and conclusions.

$$\begin{array}{c}\varphi\\ \vdots\\ \psi\end{array}$$

Here, in a derivation the formula occurrence ψ *depends* on assumption occurrence φ if *psi* lies on a thread starting with φ and φ has not been discharged. We have also discussed dependencies between *assumption* occurrences in a derivation involving the rules ($\vee$E) and ($\exists$E). Here we noticed that dependencies, although *defined* in terms of derivational structure common to both CPL and IPL, are nevertheless sensitive to the logical context.

Dependencies of terms on propositions In the intuitionistic epsilon calculus of Section 4.2 we have introduced the notion of an assumption *supporting* an epsilon term to achieve conservativity. An assumption occurrence φ supports an ϵ-term $\epsilon x : \psi$ in formula occurrence χ, or, alternatively, an ϵ-term $\epsilon x : \varphi$ *depends* on an assumption occurrence φ, if an application of the ϵ-rule with premise $\exists x\psi$ depends on φ, and χ depends on the conclusion of this application.

But even without epsilon terms, the notion of a term depending on a formula makes sense. Consider the formula $\varphi \rightarrow \exists x\psi$: That is, given the assumption φ we can conclude to the existence of an object satisfying $\psi(x)$

$$\begin{array}{c} \varphi \\ \vdots \\ \exists x\psi \end{array}$$

In terms of the Brouwer-Kolmogorov interpretation of IPL, any proof of φ can be turned into a proof of $\exists x\psi$, that is, it gives us an object satisfying ψ. So the existence of this object *depends* on a proof of φ. This example shows that dependencies arising in a proof theoretic context are sensitive to the presence of logical principles. For classically we have the IP$\exists$ principle (see Chapter 3).

$$\frac{\varphi \rightarrow \exists x\psi}{\exists x(\varphi \rightarrow \psi)}, \text{ provided } x \text{ does not occur free in } \varphi.$$

This principle states, in effect, that the dependence of the object satisfying $\psi(x)$ on the assumption φ is only apparent: by the conclusion such an object already exists without the assumption that φ holds. As we have seen in Chapter 3, the rule IP$\exists$ is not IPL valid.

Dependencies between Terms In proof theory term dependencies can be identified at various levels of magnification.

First, in the case of ϵ-terms, we can determine dependencies on the level of the *terms themselves.* The ϵ-term $\epsilon x : Rx(\epsilon y : Qyz)$ depends on the term $\epsilon y : Qyz$, because the latter is a subterm of the former. That is, $\epsilon x : Rx(\epsilon y : Qyz)$ is of the form $\epsilon x : Rxy[\epsilon y : Qyz/y]$. In semantic terms: the value assigned to $\epsilon x : Rx(\epsilon y : Qyz)$ *functionally* depends on the value assigned to $\epsilon y : Qyz$.

Next we can identify dependencies on the level *bare assertions*. An example of this is the dependence of the variable y on x in the formula

$$\forall x \exists y Rxy.$$

This dependence can be interpreted as a *functional* one by introducing the appropriate Skolem function.

$$\forall x Rxf(x)$$

In this formula the dependence of the variable y on x is restricted to the confines of this formula. That is, the structure of the variables does not show the dependence.

Finally, we can identify term-dependencies on the level of *assertions-in-context*. An instance of this can be found in the natural deduction framework. In such a framework the existential quantifier is eliminated by introducing a proper term. In a sequence of such eliminations there arises a sequence of proper terms each of which should be fresh to the derivation at that point. Consider the derivation of $\exists y \exists x Rxy$ from $\exists x \exists y Rxy$.

5.1. EXAMPLE.

$$\dfrac{\exists x \exists y Rxy \qquad \dfrac{\exists y Ray\ [1] \qquad \dfrac{\dfrac{Rab\ [2]}{\exists x Rxb\ (*)}\,\exists\mathrm{I}}{\exists y \exists x Rxy}\,\exists\mathrm{I}}{\exists y \exists x Rxy}\,\exists\mathrm{E}_2}{\exists y \exists x Rxy}\,\exists\mathrm{E}_1$$

In the starred line of the proof, the rule ($\exists$I) abstracts over the term a in Rab to give $\exists x Rxb$. In this situation the term b *depends on* a, not present in the formula itself. This dependence makes itself felt in two ways.

- We may not *eliminate* the existential quantifier from $\exists x Rxb$, at this point, by an application of ($\exists$E) with assumption Rab. This assumption would be undischargeable, for the proper term occurs in an assumption on which the major premise of the application depends (namely Rab itself). The term a is not yet released for general use. So we have to introduce a fresh constant in the elimination of the quantifier.
- We may not yet discharge assumption [1] which introduces the proper term because that term is still present in assumption [2] on which $\exists x Rxb$ depends.

5.2.2 Varieties of Dependence

We have identified dependencies between propositions, between propositions and terms and a trio of dependence relations between terms. But in many cases it is only a matter of perspective to which category we assign a certain kind of dependence. For instance, in *Correspondence Theory* dependence relations between propositions in *Intuitionistic Predicate Logic* (but also in classical predicate logic, see [Ben86b]) are translated to term dependencies. For instance, the propositional formula $\varphi \to \neg\neg\varphi$, in which the formula $\neg\neg\varphi$ depends on the formula φ, is roughly translated by the first-order formula

$$\forall\alpha\exists\beta((\alpha \leq \beta \wedge \varphi(\alpha)) \to \forall\gamma(\beta \leq \gamma \to \neg\varphi(\gamma))).$$

That is, in the translation we get a dependence between *terms.* Intuitively the term β depends on the term α.

But also proposition-term dependencies can be reduced to term-dependencies in this way. Consider again the formula $\varphi \to \exists x\psi$, which expresses a dependence of a term on a formula. In this case the intuitionistic translation gives the first order formula

$$\forall\alpha,\beta(\alpha \leq \beta \to (\varphi(\beta) \to (\exists m \in dom(\beta) \wedge \psi(m,\beta)))).$$

In this translation we observe a dependence between *terms* only, albeit terms of different sorts: the term m intuitively depends on the terms α and β.

Conversely, we have reductions of term dependencies arising in *assertions-in-context* to *formula dependencies.* The relation $\prec_\exists$ defined between $\exists$-proper terms in the last section derives from the order in which ($\exists$E) assumptions may be discharged. These assumptions are related by the relation $\prec^c$ of the last chapter. In fact, the two relations coincide. So this relation between terms in context can be reduced to a relation between assumptions in context.

But also the different varieties of dependencies between terms themselves can be related.

We have used the device of Skolem functions to make explicit the dependencies between variables in bare assertions. But in the second chapter we have interpreted ϵ-terms by Skolem functions. This suggests that dependencies in *bare assertions* can be reduced to dependencies between the *terms themselves* by using Skolem functions. But this is not immediate. Because Skolem functions give rise to functions depending only on the *free variables* of a formula, they lack the right sensitivity to correctly represent dependencies. For instance, both the terms $\epsilon x : Rx(\epsilon y : Qyz)$ and $\epsilon y : Qyz$ have only the variable z free. This leads to a Skolem representation of the formula $R(\epsilon x : Rx(\epsilon y : Qyz))(\epsilon y : Qyz)$ as

$$R(f_1(z))(f_2(z)),$$

thus losing the dependence of $f_1(z)$ on $f_2(z)$. In the literature on the epsilon calculus this is standardly solved by interpreting only ϵ-terms in matrix form by Skolem functions (for instance, [Tai65], [Min94]).

5.2. DEFINITION. (Matrix Forms) A *matrix* of an ϵ-term $\epsilon x : \varphi$ consists of a term $\epsilon x : \varphi'$, such that $\epsilon x : \varphi = \epsilon x : \varphi'[\epsilon x : \varphi_1/x_1, \dots, \epsilon x : \varphi_n/x_n]$ where $\epsilon x : \varphi_1, \dots, \epsilon x : \varphi_n$ are all proper subterms of $\epsilon x : \varphi$ that are proper subterms of no other subterm of $\epsilon x : \varphi$ (i.e., they are *maximal subterms*) and $x_1, \dots, x_n$ are n fresh and different variables.

So $\epsilon y : Qyz$ is itself a matrix, a matrix of $\epsilon x : Rx(\epsilon y : Qyz)$ is $\epsilon x : Rxy$. (Note that $\epsilon x : Rx(\epsilon y : Zxy)$ is a matrix: it is not of the form $\epsilon x : Rxz[\epsilon y : Zxy/z]$). Now interpret only matrices by Skolem functions. Epsilon terms in general are then interpreted by *compositions* of these functions. Thus, if $f_1(y)$ is the Skolem function interpreting the matrix $\epsilon x : Rxy$ and $f_2(z)$ is the Skolem function interpreting the matrix $\epsilon y : Qyz$, then $\epsilon x : Rx(\epsilon y : Qyz)$ is interpreted by $f_1(f_2(z))$. So the dependence structure is preserved.

Finally, in Section 5.3 we shall see that ϵ-terms give us a way to reduce term dependencies arising through assertions-in-context to dependencies purely between terms themselves. Thus, in the epsilon calculus all levels of magnification which give rise to term dependencies can be reduced to the level of relations between the terms themselves.

5.2.3 Proof Theoretic Formats

$$\forall\mathrm{I}\quad \frac{\varphi}{\forall x(\varphi[x/t])} \qquad\qquad \forall\mathrm{E}\quad \frac{\forall x\varphi}{\varphi[t/x]}$$

$$\exists\mathrm{I}\quad \frac{\varphi[t/x]}{\exists x\varphi} \qquad\qquad \exists\mathrm{E}\quad \frac{\exists x\varphi \qquad \begin{array}{c}\Sigma, [\varphi[t/x]]\\ \vdots \\ \xi\end{array}}{\xi}$$

In ($\forall$I), t should not occur free in Σ. In ($\exists$E), t should not occur free in Σ, φ or ξ.

Figure 5.1: Prawitz Rules for Quantifiers

Our general perspective on term-dependencies in a proof theoretic context will be that these arise in the course of a *choice process*. This viewpoint determines

1.	$\vdash \varphi[t/x] \rightarrow \exists x\varphi$	t free for x in φ
2.	$\vdash \exists x(\varphi \vee \psi) \leftrightarrow \exists x\varphi \vee \exists x\psi$	
3.	$\vdash \exists x\varphi \rightarrow \varphi$	x not free in φ

Figure 5.2: Quantifiers in an Axiomatic Set-Up

$$\mathrm{L}\exists\frac{\Gamma, \varphi[y/x] \Rightarrow \Delta}{\Gamma, \exists x\varphi \Rightarrow \Delta} \qquad \mathrm{R}\exists\frac{\Gamma \Rightarrow \varphi[t/z], \Delta}{\Gamma \Rightarrow \exists z\varphi\Delta}$$

$$\mathrm{L}\forall\frac{\Gamma, \varphi[t/z] \Rightarrow \Delta}{\Gamma, \forall z\varphi \Rightarrow \Delta} \qquad \mathrm{R}\forall\frac{\Gamma \Rightarrow \varphi[y/x], \Delta}{\Gamma \Rightarrow \forall x\varphi\Delta}$$

For y not free in Γ, Δ

Figure 5.3: Quantifier Rules in the Sequent Calculus

to a large extend our choice of proof theory. In Figure 5.1, 5.2, and 5.3 we have presented the quantifier rules in different proof theoretic set-ups. From a choice-perspective the natural deduction framework seems to be the eminently suitable for our investigations. In natural deduction, the rules for the quantifiers involve the introduction and elimination of *proper terms*, and we can locate the *moment of choice* at the application of the introduction rule.

An application of the rule ($\exists$E) with major premise $\exists x\varphi$ involves the introduction of some proper term a. Such an application represents a *choice* made to witness the existential premise: an element a is chosen to satisfy the *condition* $\varphi(x)$. An application of this rule is concluded by *discharging* the choice made. The conditions on discharge guarantee that the choice made has been *arbitrary*. In game-theoretic terms: we defend a statement $\exists x\varphi \rightarrow \psi$ by making an *arbitrary* choice to witness the existential formula and showing that ψ follows given this choice.

An application of the rule ($\exists$I) concludes to $\exists x\varphi$ from $\varphi[a/x]$. Here, we can defend the conclusion $\exists x\varphi$ by simply choosing the witness a for which $\varphi(x)$ apparently holds. (Notice that this only works if any term satisfying $\varphi(x)$ can in fact be chosen as a witness. If we consider choices as *resources*, this in fact need not be the case.)

Once we have chosen a value to witness an existential formula, we may use this witness to formulate the condition of subsequent choices. This imposes a natural *dependence* ordering on the choices made in the course of a derivation: the value chosen in one choice is used in the condition of a subsequent one. Consider again the derivation

$$
\cfrac{\exists x \exists y Rxy \qquad \cfrac{\exists y Ray\ [1] \qquad \cfrac{\cfrac{Rab\ [2]}{\exists x Rxb}\ \exists\mathrm{I}}{\exists y \exists x Rxy}\ \exists\mathrm{I}}{\exists y \exists x Rxy}\ \exists\mathrm{E}_2}{\exists y \exists x Rxy}\ \exists\mathrm{E}_1
$$

Here we can locate the points at which choices are made to witness the existential formulas. The value chosen for $\exists xyRay$, the term b, depends on the choice made for $\exists x\exists yRxy$, the term a. This is witnessed by the fact that the term a occurs in the *condition* of the choice which b has to satisfy. The following definition fixes this relation of dependence among $\exists$-proper terms precisely. The relation will be defined with respect to *pure* derivations, that is, the proper terms are local to the rules that introduce (eliminate) them. Every derivation can be brought to pure form ([Pra65]).

5.3. DEFINITION. (Dependence Ordering[1]) Let $\alpha_1, \ldots, \alpha_n$ be all applications of ($\exists$E) in a pure derivation Δ. Let for each $i \leq n$ a_i be the proper term of application α_i. Then a_i *immediately ($\exists$E)-depends* on a_j in Δ, notation $a_i \ll_\exists a_j$, if a_j occurs in the major premise of α_i .
We shall write $a_i \prec_\exists a_j$ if there is a finite sequence of immediately $\exists$–dependent terms relating a_i with a_j.

In pure derivations, the relations $\prec_\exists$ is transitive and irreflexive. The relation $\prec_\exists$ between $\exists$-proper terms mirrors exactly the order in which the assumptions introducing these terms can be discharged.

Compare the above derivation to a derivation of the same principle but now in the format of the sequent calculus:

$$
\cfrac{\cfrac{\cfrac{\cfrac{Rxy[a/x][b/y] \Rightarrow Rxy[a/x][b/y]}{Rxy[a/x][b/y] \Rightarrow \exists x Rxy[b/y]}}{Rxy[a/x][b/y] \Rightarrow \exists y \exists x Rxy}}{\exists y Rxy[a/x] \Rightarrow \exists y \exists x Rxy}}{\exists x \exists y Rxy \Rightarrow \exists y \exists x Rxy}
$$

[1] All quantifier rules can give rise to dependence orderings. Only the $\prec_\exists$ dependence relation has derivational content in the standard system. The relation $\ll_\exists$ reflects the *discharge ordering* of the assumptions connected to the ($\exists$E) applications: nested applications of ($\exists$E) result in dependence chains of proper terms. The $\prec_\exists$ relation is irreflexive and asymmetric for every derivation Δ. As there is nothing to discharge in the case of ($\exists$I), ($\forall$I) or ($\forall$E), the dependence ordering they induce is devoid of content in standard CPL and IPL. This is not the case in, for instance, *free logic*. There ($\forall$E) needs the assumption that the proper term *denotes*.

Here, there seem to be no points at which we can locate the *choice* of proper terms: term dependencies seem to have no natural interpretation in this derivation (although see Appendix I to this chapter). The same holds for a derivation of this principle in an axiomatic set-up. We shall spare the reader such a derivation.

5.2.4 Explicit Dependencies

The central role of dependence in the elimination rule for the existential quantifier can be highlighted by considering a natural deduction format which uses a rule for existential *instantiation*, instead of an elimination rule. In Section 2.5.3 we have introduced the Copi-Kalish system in its formulation by Fine [Fin85]. In this system we have the usual quantifier rules $(\forall\text{I}), (\forall\text{E}), (\exists\text{I})$ but the rule $(\exists\text{E})$ is replaced by

$$\frac{\exists x\varphi}{\varphi[a/x]}\ \exists\text{E}\ a \prec AO(\varphi)$$

This rule of existential *instantiation* explicitly introduces dependencies between the term introduced and the set $AO(\varphi)$ of proper terms upon which the derivation of φ depends. This set is defined recursively with $AO(\varphi) = \emptyset$ for φ occurring at a leaf node in a proof tree, and $AO(\varphi) = AO(\psi)$ if φ is the result of any application other than $\exists E$ or $\forall E$ with ψ as premise, and $AO(\varphi) := AO(\varphi) \cup \{a\}$ for φ the result of an application of $\exists E$ or $\forall E$ with φ as premise and a as proper term. In this calculus, the proper term introduced by $(\exists\text{E})$ should always be fresh to the derivation and the rule $(\forall\text{I})$ has the restriction that no term may depend on the proper term of this application. Here are two correct derivations in this calculus:

$$\dfrac{\dfrac{\dfrac{\dfrac{\exists x\forall y Rxy}{\forall y Ray}\ \exists\text{E}\ a \prec \emptyset}{Rab}\ \forall\text{E}}{\exists x Rxb}\ \exists\text{I}}{\forall y\exists x Rxy}\ \forall\text{I}
\qquad
\dfrac{\dfrac{\dfrac{\dfrac{\exists x\exists y Rxy}{\exists y Ray}\ \exists\text{E}\ a \prec \emptyset}{Rab}\ \exists\text{E}\ b \prec a}{\exists x Rxb}\ \exists\text{I}}{\exists y\exists x Rxy}\ \exists\text{I}$$

And here is an example of how the dependence relation is used in blocking the derivation of $\exists y\forall x Rxy$ from $\forall x\exists y Rxy$:

$$\dfrac{\dfrac{\dfrac{\forall x\exists y Rxy}{\exists y Rax}\ \forall\text{E}}{Rab}\ \exists\text{E}\ b \prec a}{\text{application of } \forall I \text{ to } a \text{ blocked because } b \prec a}$$

The switch from Prawitz-style natural deduction to natural deduction with an existential instantiation rule recording dependencies is one for convenience of presentation. For CPL, these calculi are interchangeable.

5.4. PROPOSITION. *Every derivation $\mathcal{D}$ of φ from Σ in the Copi-Kalish system where φ is free of proper terms can be embedded in a derivation using existential elimination. Moreover, the conditions on the use of the proper terms in $\mathcal{D}$ can be read off from the discharge restrictions on ($\exists E$) applications.*

Proof: We can consider every application of the existential instantiation rule as an application of *Modus Ponens* in standard natural deduction with 'hidden' major premise and assumption in the following way:

$$\dfrac{\begin{matrix}\Gamma\\ \vdots\\ \exists x\varphi\end{matrix}}{\begin{matrix}\varphi[a/x]\\ \vdots\\ \xi\end{matrix}}\,\exists\mathrm{E}\ a \prec AO(\varphi) \quad\Rightarrow\quad \dfrac{\boxed{P\exists}\qquad \dfrac{\begin{matrix}\Gamma\\ \vdots\\ \exists x\varphi\end{matrix}\qquad \boxed{\exists x\varphi \rightarrow \varphi[a/x]\ [i]}}{\begin{matrix}\varphi[a/x]\\ \vdots\\ \xi\end{matrix}}}{\xi}\,\exists\mathrm{E}_i$$

Here the classical *theorem* $P\exists$ is used as the major premise of an ($\exists$E) application. The assumption of this ($\exists$E)-application is the hidden assumption of an application of the instantiation rule. ⊠

In the left hand derivation we have $a \prec b$ if and only if in the right hand derivation there are two ($\exists$E) assumptions $\psi_1 = \exists x\varphi \rightarrow \varphi[b/x]$ and $\psi_2 = \exists y\chi[b/x] \rightarrow \chi[b/x, a/y]$ such that $\psi_1 \prec^c_{\mathcal{D}'} \psi_2$.

Given the fact that in the Copi-Kalish system the instantiation rule uses a 'hidden' ($\exists$E)-assumption, the conditions on the quantifier rules particular to this system make evident sense.

5.2.1. CONDITION. *The proper term of an application of $\forall I$ should not occur as an $\exists$-proper term in the proof tree of the premise.*

This holds because the $\exists$-proper term occurs in the hidden assumption. So by the standard condition on ($\forall$I), we may not universally generalize over it.

5.2.2. CONDITION. *The proper term t of an application of $\forall I$ should not occur in the conclusion or in any assumption on which the premise φ depends, nor should any term b with $b \prec t$.*

Again, if $b \prec t$, then t occurs in a hidden ($\exists$E)-assumption, so by the standard restriction on ($\forall$I), we may not generalize over it.

5.2.3. CONDITION. *The proper term of an application of $\exists E$ is fresh (i.e., it should not occur in the proof tree of the premise).*

This reduces to the standard condition that ($\exists$E)-proper terms should be fresh.

This representation of derivations in the Copi-Kalish system makes it clear that a proper conclusion of a derivation from constant-free assumptions should have no proper terms itself: when the conclusion still contains a proper term, the hidden minor premise corresponding to that term cannot be discharged. So the conclusion still follows only with a hidden assumption. Moreover, in a derivation $\mathcal{D}$, the dependence relation between terms, '$\prec$', used in the instantiation rule, directly reflects the order in which the hidden assumptions may be discharged in the ($\exists$E) representation. That is, it coincides with the relation $\prec_\exists$ we have introduced in the previous section.

An interesting alternative instantiation framework has been introduced by Quine [Qui52]. Again we take the formulation of K. Fine [Fin85]. This system is like the Copi-Kalish system, in having introducing dependencies when we instantiate existentially, but now also the rule ($\forall$I) introduces dependencies.

$$\frac{\varphi[a/x]}{\forall x\varphi}\,\forall I\ a \prec AO(\varphi)$$

The set $AO(\varphi)$ contains now both the $\exists$- and $\forall$-proper terms in φ. The dependence relation is addressed by both the $\exists E\ a \prec AO(\varphi)$ and the $\forall I\ a \prec AO(\varphi)$ rule. The interesting variation on the previous system is that by this rule, the proper term a becomes dependent on the terms in φ when it is *eliminated* from the proof.

In contrast to the system of Copi and Kalish, in the Quine system there are no *local* restrictions on the rules. Here the dependence relation is used to rule out, on a global level, some derivations as being correct. In a *correct* derivation the dependence relation satisfies the following conditions:

1. No proper term shall be instantial term twice, i.e., to two applications of of the same rule or of different rules.
2. $\prec$ must be antisymetric and irreflexive.

In this system there is no requirement that the proper terms of $\forall I\ a \prec AO(\varphi)$ applications should not occur in the assumptions on which the premise depends. In this, it differs from the standard rule ($\forall$I). This difference allows the following derivation.

$$\dfrac{\dfrac{\varphi[a/x]\ [1]}{\forall x\varphi}\,\forall\mathrm{I}\ a \prec AO(\varphi)}{\varphi[a/x] \rightarrow \forall x\varphi}\,\rightarrow\mathrm{I}_1$$

A derivation which extends this proof with the conclusion $\forall x(\varphi \rightarrow \forall x\varphi)$ by applying ($\forall$I) is not a correct one, by the restriction that $\prec$ must be irreflexive.

5.5. PROPOSITION. *Every derivation $\mathcal{D}$ of φ from Σ in the Quine system, where φ is free of proper terms, can be embedded in a derivation using existential elimination. Moreover, the conditions on the use of the proper terms in $\mathcal{D}$ can be read off from the discharge restrictions on ($\exists E$) applications.*

Proof: The rule $\forall I\ a \prec AO(\varphi)$ can be analyzed by means of the universal counterpart of Plato's principle:

$$\dfrac{\boxed{\exists x(\varphi[a/x] \rightarrow \forall x\varphi)} \qquad \begin{matrix} \dfrac{\begin{matrix}\Gamma \\ \vdots \\ \varphi[a/x]\end{matrix} \qquad \boxed{\varphi[a/x] \rightarrow \forall x\varphi\ [i]}}{\forall x\varphi} \\ \vdots \\ \xi \end{matrix}}{\xi}\ \exists E_i$$

Again, the conditions on the use of proper terms in the Quine system can be read off from the conditions on ($\exists$E) discharge. ⊠

5.2.5 Substructural Variation

Our eventual aim is the following. Modern proof theory is bringing to light various resource elements to reasoning, which have been neglected, or rather, set to implicit 'default values' in standard (predicate) logic. A prime example is of course linear logic, which brought out *occurrences*, and the importance of structural rules as choice points for their manipulation. We propose the same for the quantifier rules, with their associated dependencies between terms. As some predecessors in this field we mention Fine [Fin85], van Lambalgen [Lam91]. One gets a spectrum of decisions as to what depends on what. The purpose of this chapter is merely to look around in this landscape, which will throw light on the hidden dependence structure of the standard quantifiers and will provide a much richer setting for linguistic applications. We propose that the proper setting for a substructural investigation of quantification is in a theory of *choice processes.* In game-theoretical semantics an existentially quantified formula is intuitively analyzed in terms of a *friendly* choice of a witness, while the universal quantifier is intuitively analyzed in terms of a *hostile* choice of such a witness. This corresponds in the epsilon calculus to the interpretation of an ϵ-term $\epsilon x : \varphi$ by a choice function assigning an element satisfying $\varphi(x)$ (if there is such an element), and the interpretation of a τ-term $\tau x : \varphi$ by a choice function assigning an element satisfying $\neg\varphi(x)$ (if there is such a one). Now classical logic is characterized by a principle of *free choice.* In a natural deduction framework this is witnessed for the existential quantifier by the principle ($\exists$I):

$$\frac{\varphi[a/x]}{\exists x\varphi}$$

From the perspective of a choice interpretation this means that any object or term satisfying $\varphi(x)$ can be chosen as a witness for $\exists x\varphi$. Now if we consider choice processes as *resources* then this free choice principle characterizes only some specific class of processes.

By our choice interpretation, a formula $\exists x\varphi$ does not merely mean that $\varphi[a/x]$ holds for some arbitrary a, but that $\varphi(x)$ holds for some a that *we can choose.* By this reasoning $\varphi[a/x]$ is not enough to conclude to $\exists x\varphi$, unless a is a choice we can make to satisfy the condition $\epsilon x : \varphi$.

5.3 Epsilon Calculus as a Testing Ground

The epsilon calculus in its natural deduction formulation constitutes an excellent medium to test substructural interpretations of the existential quantifier in terms of choice processes. Its semantics has traditionally been formulated in terms of *choice* functions (see Sections 2.16 and 2.39) and its syntax expresses choice dependencies *explicitly* where standard natural deduction set-ups exhibit only dependencies-in context. Consider the following derivations of $\exists y\exists xRxy$ from $\exists x\exists yRxy$, the first one in the Copi-Kalish system, the second one in the epsilon calculus.

$$\dfrac{\dfrac{\dfrac{\dfrac{\exists x\exists yRxy}{\exists yRay}\,\exists\mathrm{E}\ a\prec\emptyset}{Rab}\,\exists\mathrm{E}\ b\prec a}{\exists xRxb}\,\exists\mathrm{I}}{\exists y\exists xRxy}\,\exists\mathrm{I}$$

$$\dfrac{\dfrac{\dfrac{\dfrac{\exists x\exists yR(x,y)}{\exists yR(\epsilon x:\exists yR(x,y),y)}\,\exists\mathrm{E}_\epsilon}{R(\epsilon x:\exists yR(x,y),\epsilon y:R(\epsilon x:\exists yR(x,y),y))}\,\exists\mathrm{E}_\epsilon}{\exists xR(x,\epsilon y:R(\epsilon x:\exists yR(x,y),y))}\,\exists\mathrm{I}}{\exists y\exists xR(x,y)}\,\exists\mathrm{I}$$

Observe the structural similarity between the derivations. The term correspondences are here

$$\underbrace{\epsilon x:\exists yR(x,y)}_{a} \qquad \underbrace{\epsilon y:R(\underbrace{\epsilon x:\exists yR(x,y)}_{a},y)}_{b}\,.$$

In the first derivation the dependence of the term b on a, arising in *assertions-in-context*, is recorded by annotations of the proof. In the ϵ-derivation, this dependence is recorded *syntactically* in the corresponding *ϵ-terms themselves*: the term $a = \epsilon x : \exists yR(x,y)$ occurs as a subterm of $b = \epsilon y : R(\epsilon x : \exists yR(x,y),y)$.

In an ϵ-derivation terms occurring on the surface level of a formula may have subterms that do not occur on the surface level of that formula, but on the surface of *assumptions* of that formula. Thus the "global" dependence relation $\prec_\exists$ of standard ($\exists$E)-derivations is "locally" witnessed in ϵ-derivations: at the level of terms occurring in a formula. The complete dependence structure of the terms occurring in the derivation can be determined at the level of bare-assertions. This means that, in ϵ-derivations, ($\exists$I) may abstract over terms in a formula φ that, in ($\exists$E)-derivations, occur outside of φ. For instance in the fourth line of the above ϵ-derivation, we could have abstracted over *both* occurrences of $\epsilon x : \exists y Rxy$. This possibility is lacking in the Copi-Kalish derivation. In Section 5.6 we shall consider extensions of the Copi-Kalish system which allow quantifiers in a derivation to have scope over the dependence relation, thus incorporating the extended quantificational possibilities of the ϵ-calculus into a standard framework.

Proper Terms and Epsilon-Terms

The relation between the *epsilon* terms occurring in the derivation in the epsilon language and the *proper* terms occurring in its twin in the Copi-Kalish system can be clarified by the notion of a choice process.

A proper term occurring in a Copi-Kalish derivation is always a proper term with respect to a specific application of an inference rule with specific premises. Because of this we can identify these terms by specifying (1) which of the two rules they have been introduced by, and (2) the formula they have been introduced to satisfy, given by the premise of the respective rule. For instance, the term a introduced by ($\exists$E) with premise $\exists x\varphi$ can be characterized by the following *term description*:

$$\langle\langle\epsilon, x, \varphi\rangle, a\rangle.$$

Here a is the term chosen, for the ($\exists$E) application with major premise $\exists x\varphi$. The *condition* of the choice is represented in the triple $\langle\epsilon, x, \varphi\rangle$. Here we see the components $\exists, x$, and φ of the premise of the application reflected. In $\langle\epsilon, x, \varphi\rangle$ the components identify

- ϵ the *kind of application*. In this case ($\exists$E). Other kinds of terms (for instance τ) may be used for other rule applications (for instance ($\forall$I)).
- x The *variable* to which the (value of the) proper term is to be bound in the formula φ (which may contain other free variables).
- φ The *condition* the (value of the) proper term has to satisfy with respect to the variable x in φ.

Note that, in the condition of a term description, other proper terms may occur. To avoid needless notation we shall write the triple $\langle\epsilon, x, \varphi\rangle$ as a (metalinguistic) term of the form $\epsilon x : \varphi$ and shall refer to tuples of the form $\langle \epsilon x : \varphi, a\rangle$ as *choice tuples*. Every proper term in a pure derivation is described by such a tuple. In the three components of an ϵ-term we have exhaustively described the *parameters* of choice, and choice strategies may differ with respect to their sensitivity to the structure of these parameters.

This suggests our interpretation of the relation between the derivation in the epsilon calculus and the one in the Copi-Kalish calculus.

> In the epsilon derivation we work with the *descriptions* of the choices.
> In the Copi-Kalish derivation we work with the *values* of these choices.

A *choice process* is now an ordered structure of choice tuples. We may accompany a natural deduction proof by specifying the sequence of choice tuples 'active' at any point. A tuple $\langle \epsilon x : \varphi, a\rangle$ can be seen as the choice of a witness to *defend* the existential statement $\exists x\varphi$. By the classical rules this tuple is introduced as an *assumption*, given $\exists x\varphi$ where a must be fresh to the derivation. We may *discharge* this choice at a formula occurrence not containing the choice value as long as that value does not occur in the condition of some previous choice tuple.

Notice that, given $\varphi[a/x]$, we may always *assume* the choice $\langle \epsilon x : \varphi, a\rangle$, but this assumption may not be dischargeable. For instance, if we conclude $Ray[a/y]$ from $\exists xRxx$ with choice $\langle \epsilon x : Rxx, a\rangle$, then assuming the choice $\langle \epsilon y : Ray, a\rangle$ we can derive $\exists yRxy$. But the value of the second choice occurs in the first and vice versa. So neither choice is dischargeable. (Moreover, the choice tuple $\langle \epsilon y : Ray, a\rangle$ could not have been introduced by applying ($\exists$E), for such an application always replaces *all* occurrences of the bound variable by the proper term.) What we need in this case are rules that *derive* new choice tuples from old ones.

The structure of a choice process can be formulated in terms of *dependence*.

5.6. DEFINITION. (Choice Dependence) We say that choice tuple $\langle \epsilon x : \varphi, a\rangle$ *immediately depends* on tuple $\langle \epsilon x : \psi, b\rangle$, notation $\langle \epsilon x : \varphi, a\rangle \ll \langle \epsilon x : \psi, b\rangle$, if b occurs in φ. We say that choice tuple $\langle \epsilon x : \varphi, a\rangle$ *depends* on tuple $\langle \epsilon x : \psi, b\rangle$, notation $\langle \epsilon x : \varphi, a\rangle \prec \langle \epsilon x : \psi, b\rangle$, if there is a finite sequence of immediate dependence steps connecting $\langle \epsilon x : \varphi, a\rangle$ to $\langle \epsilon x : \psi, b\rangle$.

Now the restriction on discharge of an assumed choice tuple is that it must be *arbitrary*: the value chosen must be fresh, it may not occur in an assumption nor in some other choice tuple. That is, the choice dependence relation associated with a derivation must be a *strict partial order*, and every value may occur only once.

So with respect to a choice process we can distinguish two orthogonal aspects. Firstly, there is the *internal structure* of the choice tuples involved. Secondly, there is its *dependence structure*: this is an abstract property, independent of the internal structure of the choice tuples. We shall see that the two aspects of choice processes have to be addressed independently to give us various quantifier principles.

The Fine Structure of Choice

The elimination of the existential quantifier introduces choice tuples (or term descriptions). Now it is natural to suggest that we may *introduce* an existential statement $\exists x\varphi$, given $\varphi[a/x]$, *only if the choice tuple* $\langle \epsilon x : \varphi, a\rangle$ *is available.* For we can defend $\exists x\varphi$, given $\varphi[a/x]$, only if a is in fact available as a choice. This entails that it is around the rule ($\exists$I), the introduction rule for the existential quantifier, that substructural logics for the existential quantifier center.

Given $\varphi[a/x]$, when can we choose a as a witness for $\exists x\varphi$? Classically, there is a simple principle: whenever we have $\varphi[a/x]$, we have the $\langle \epsilon x : \varphi, a\rangle$. *Every object* satisfying $\varphi(x)$ is a potential witness: the presence of $\varphi[a/x]$ is enough to defend $\exists x\varphi$. This means that the existence of a choice tuple is independent of the introduction by ($\exists$E).

If we want to take the notion of a choice process seriously, then we have to find restrictions on the set of available choices. We shall take the viewpoint here that all non-dischargeable choice tuples must eventually derive from choices made in the elimination of the existential quantifier. That is, the rule ($\exists$E) introduces choices. What we require are rules that tell us how these choices propagate.

In the quantifier free epsilon calculus (see Section 2.3.3), this freedom of choice has an especially clear formulation. Notice that, in the epsilon calculus, the requirement that $\langle \epsilon x : \varphi, a\rangle$ has to be present in order to conclude $\exists x\varphi$ from $\varphi[a/x]$, comes down to the requirement that we need $\varphi[\epsilon x : \varphi/x]$ in order to conclude $\exists x\varphi$. Now in this calculus we introduced the existential quantifier by definition.

5.7. DEFINITION. (Defined Existential Quantifier)

$$\mathcal{E}x\varphi \equiv_{df} \varphi[\epsilon x : \varphi/x]$$

and checked whether the quantifier '$\mathcal{E}$' could be interpreted as the real thing, '$\exists$'. This was the case in the presence of the *epsilon term rule*,

$$\frac{\varphi[a/x]}{\varphi[\epsilon x : \varphi/x]}$$

which, by the definition of '$\mathcal{E}$', again expresses the the standard rule ($\exists$I). Now, as we stated, substructural versions of proof rules for the existential quantifier in a Prawitz set-up, will center around the rule ($\exists$I). Analogously, in the epsilon calculus, they will revolve around the epsilon term rule. If we keep the definition of the quantifier symbol '$\mathcal{E}$' as given, then variations of the epsilon term rule will give us a variety of existential quantifiers.

5.4 Benchmark Problems

In this section we shall consider a number of *benchmark* principles which illustrate various orthogonal properties of the existential quantifier. This particular set has turned out to be significant in *Modal State Semantics* [Ben94b], [Ben94a], [ANvB94], [Ném93], and the substructural framework of [AvL95] (a brief discussion of the central definitions of these semantics is given in Appendix II to this chapter). In this semantics the fine structure of first-order quantifiers is investigated within the framework of modal logic. As is the case in our present *choice* perspective, the truth of an existential statement $\exists x\varphi$ is taken to imply more than the mere existence of a domain element satisfying $\varphi(x)$. In modal state semantics this surplus of meaning is taken to be the fact that the domain element has to stand in a specific relation to *the values of all other variables.* In the framework of [AvL95], this surplus means that this domain element has to be related in a specific way to *the parameters of* φ. In the present set-up this surplus is interpreted as the availability of the domain element at a certain state of a choice process. As we shall see, these two perspectives cut up the list of benchmark problems in different ways. This has to do with our special interest in the notion of *dependence.*

1. $\exists x(\varphi \vee \psi) \rightarrow \exists x\varphi \vee \exists x\psi$,
2. $\exists x\varphi \vee \exists x\psi \rightarrow \exists x(\varphi \vee \psi)$,
3. $\exists x Px \rightarrow \exists y Py$,
4. $\exists x Py \rightarrow Py$,
5. $\exists x Rxx \rightarrow \exists y \exists x Rxy$,
6. $\exists x \forall y Rxy \rightarrow \forall y \exists x Rxy$,
7. $\exists x \exists y Rxy \rightarrow \exists y \exists x Rxy$.

The first two principles determine the interaction of the existential quantifier with the Booleans.

In the minimal calculi we shall consider, these will be derivable. We are interested in the substructure of quantifier interactions. In the interaction of the existential quantifier with the Boolean connectives no dependencies are involved that are of interest to us. From a choice perspective on existential quantification: if we have chosen a witness for $\varphi \vee \psi$, then, ipso facto, we have chosen one for φ or for ψ. In terms of choice tuples, principle (1) expresses the following. Suppose we have chosen a value a as a witness for $\exists x(\varphi \vee \psi)$. That is we have the choice

$$\langle \epsilon x : (\varphi \vee \psi), a \rangle.$$

Then we may use the choice $\langle \epsilon x : \varphi, a \rangle$ as choice tuple, given $\varphi[a/x]$. Principle (2) expresses that, if we have made choice $\langle \epsilon x : \varphi, a \rangle$, then we may use $\langle \epsilon x : (\varphi \vee \psi), a \rangle$, given $\varphi[a/x] \vee \psi[a/x]$. These rules allow the standard derivations of

(1) and (2):

$$\dfrac{\exists x\varphi \vee \exists x\psi \qquad \dfrac{\dfrac{\dfrac{\exists x\varphi\ [i]}{\varphi[a/x]}}{\varphi[a/x] \vee \psi[a/x]}}{\exists x(\varphi \vee \psi)} \qquad \dfrac{\dfrac{\dfrac{\exists x\psi\ [j]}{\psi[b/x]}}{\varphi[b/x] \vee \psi[b/x]}}{\exists x(\varphi \vee \psi)}}{\exists x(\varphi \vee \psi)}\ \vee\mathrm{E}_{i,j}$$

$$\dfrac{\dfrac{\exists x(\varphi \vee \psi)}{\varphi[a/x] \vee \psi[a/x]} \qquad \dfrac{\dfrac{\varphi[a/x]\ [i]}{\exists x\varphi}}{\exists x\varphi \vee \exists\psi} \qquad \dfrac{\dfrac{\psi[a/x]\ [j]}{\exists x\psi}}{\exists x\varphi \vee \exists x\psi}}{\exists x\varphi \vee \exists x\psi}\ \vee\mathrm{E}_{i,j}$$

Notice that the principles required for these derivations are sensitive to the *internal structure* of the choice conditions, but do not involve dependence structure.

Principles (3) and (4) deal with variable management. In terms of a choice interpretation these principles address the *variable* occurring in the choice tuple $\langle\langle\epsilon, \mathrm{x}, \varphi\rangle, a\rangle$. By principle (3), the identity of the variable we choose a value for, is immaterial. That is, having the description $\langle\epsilon x : Px, a\rangle$ we may use $\langle\epsilon y : Py, a\rangle$ for every variable y. The variable occurring in the condition of the choice is merely a placeholder. So the proper choice condition is in fact a class of conditions that are identical up to the identity of the bound variable.

By principle (4) if we choose a value for a variable not involved in the condition, then this does not affect the truth of the predicate. This means that the action of choice itself has no logical content. Notice that in the description $\langle\epsilon y : Px, a\rangle$ there is no condition for a to satisfy. That is, a is a free choice.

Principle (5) deals with the property of *deidentification* or *weakening* of the existential quantifier ([Sán91]). By this principle we can access the *locations* present in a predicate symbols: if two locations of a predicate are occupied, even if it is by the same term, then we can abstract over the locations separately. In the natural deduction framework this property derives from an aspect of the rule ($\exists$I) in which it is unique among the quantifier rules. It is the only rule in which the substitution box occurs in the premise (see Figure 5.1). By this property it is the only quantifier rule which allows deidentification. If we would change the rule ($\exists$I) to

$$\frac{\varphi}{\exists x(\varphi[x/t])}$$

i.e., if we move the substitution box to the conclusion, then our calculus would no longer be able to address predicate *locations*. That is, the same term occurring at different locations in a predicate can not be treated differently. Identities between arguments of a predicate are preserved in the course of a derivation.

Note that from a choice perspective of the existential quantifier, this principle means that we can choose the same object twice. Under some choice regimes this may be disallowed. This principle is not sensitive to the internal structure of the choice condition, but now dependence is involved essentially. This can be seen if we consider the choices occurring in a derivation of $\exists x \exists y Rxy$ from $\exists x Rxx$. We introduce $\langle \epsilon x : Rxx, a\rangle$ by assumption and require $\langle \epsilon y : Ray, a\rangle$. Here we need a choice tuple in which the value chosen occurs itself in the condition of that choice (the description depends on itself, according to definition 5.6). Principles (6) and (7) involve quantifier *interaction.* Again the internal structure of choice tuples is not involved, only their dependence structure. Principle (6) states that any independent choice satisfying some universal property can be construed as functionally depending on this property. Principle (7) expresses the fact that we can choose witnesses for existentially bound variables in any order. In a derivation of (7) we shall produce the description $\langle \epsilon y : Ray, b\rangle$ and shall need $\langle \epsilon x : Rxb, a\rangle$. The value b depends on the choice-value a and vice versa. This gives to a circular dependence relation. Again, under some choice regime this may be excluded.

5.5 Dependence Sensitive Prawitz Calculi

In this section we shall start taking term dependencies seriously by formulating inference rules which refer to the dependence relation $\prec_\exists$. We shall leave the propositional rules untouched as well as the rules for the universal quantifier. That is, we shall concentrate on the dependence structure of choice processes and disregard their internal structure. As a consequence, in these calculi, only benchmark principles (5) and (7) will be affected.

Rules for the Booleans: Standard.

Rules for the universal quantifier:

$$\frac{\varphi}{\forall x(\varphi[x/t])}\,\forall\mathrm{E} \qquad\qquad \frac{\forall x\varphi}{\varphi[t/x]}\,\forall\mathrm{E}$$

provided t does not occur in an assumption to φ, and for no $a \in AO(\varphi) : t \prec_\exists a$

In the standard calculus, the dependencies introduced in a derivation by applications of the rule ($\exists$E) play no role at all at applications of ($\exists$I): they do not *restrict* applications of ($\exists$I). Here is where the restricted calculi we shall introduce differ from the standard calculus. In the last two lines of example 5.1 assumptions are discharged. These are actions belonging to an *elimination* rule.

That is, the property of *normal* derivations — like this one — that introductions follow eliminations holds only up to *segments* ([Pra65]), in this case, the segment consisting of copies of $\exists y \exists x Rxy$. From the perspective of introduction and elimination, it would be more natural to let the elimination rule eliminate the quantifier and introduce the assumption, while the introduction rule would introduce the quantifier and eliminate the assumption.

As we have noticed, from a perspective of choices as *resources* it is the rule ($\exists$I) that seems too liberal: it assumes that from the fact that $\varphi[a/x]$ holds alone it is permissible to conclude $\exists x\varphi$. This disregards the requirement that we must be able to choose the term a as a witness for this existential formula. A dependence sensitive proof calculus arises when the proof rule ($\exists$I) refers to the relation $\prec_\exists$ built by the rule ($\exists$E). Here, we shall discuss two such calculi which use the dependence relation in fundamentally different ways. The first calculus takes a *local* perspective. We restrict applications of the quantifier rule ($\exists$I) to those satisfying certain conditions (the Copi-Kalish calculus is an example of this approach). The second one takes a *global* perspective. We can apply quantifier rule ($\exists$I) without any but the usual restrictions, but accept only derivations with dependence relations of the right form. For an arbitrary property P of dependence relations, $\mathcal{D}$ is said to be a P-derivation if the dependence relation associated with $\mathcal{D}$ has property P. Now we can state that only P-derivations will be considered correct (Quine's calculus of Section 5.2.4 is an example of this approach).

5.5.1 Calculus I: Local Restrictions on Rules

The first calculus we shall consider is of the local kind. It is an adaptation of the Copi-Kalish system in which the dependence introduced by the rule ($\exists$E) is consulted in applying the rule ($\exists$I).

Rules for the existential quantifier:

$$\frac{\exists x\varphi}{\varphi[a/x]}\ \exists\mathrm{E}\ \ a \prec_\exists AO(\varphi) \qquad\qquad \frac{\varphi[a/x]}{\exists x\varphi}\ \exists\mathrm{I}_I$$

provided there is no
$b \in AO(\varphi) : b \prec_\exists a$

In this calculus, a formula of the form $\varphi[t/x]$ is no longer sufficient to conclude $\exists x\varphi$. The mere fact that t is a φ-er is not sufficient to conclude that there exists some φ-er. We also need that the term t satisfying φ stands in the proper dependence relation to the parameters in φ.

5.5.2 Calculus II: Global Constraints on Proofs

In the second calculus we shall consider, there are no restrictions on the quantifier rules. In the Copi-Kalish system, we annotate the rule (∃I) with an instruction to extend the dependence relation. Moreover, we add a general constraint on correct derivations.

Rules for the existential quantifier:

$$\frac{\exists x\varphi}{\varphi[a/x]}\,\exists\mathrm{E}\ a \prec_{\exists} AO(\varphi) \qquad\qquad \frac{\varphi[a/x]}{\exists x\varphi}\,\exists\mathrm{I}_{II}\ a \prec AO(\varphi)$$

5.5.1. CONDITION. *A derivation is correct if the dependence relation at the conclusion is a* strict partial order.

In this calculus both rules add their proper terms to the dependence relation. Notice that if the term a already depends on the elements of $AO(\varphi)$ (by having been introduced by $(\exists E_I)$), then the rule $(\exists E_{II})$ does not change the dependence relation.

5.5.3 Discussion of Benchmarks

We shall compare the standard Copi-Kalish system with the two extended calculi we have introduced. First permutation of quantifiers. In the standard, unrestricted calculus this has the following form.

$$\cfrac{\cfrac{\cfrac{\cfrac{\exists x\exists yRxy}{\exists yRay}\,\exists\mathrm{E}\ a \prec \emptyset}{Rab}\,\exists\mathrm{E}\ b \prec a}{\exists xRxb}\,\exists\mathrm{I}}{\exists y\exists xRxy}\,\exists\mathrm{I}$$

In the restricted calculi we get the following:

$$\cfrac{\cfrac{\cfrac{\exists x\exists yRxy}{\exists yRay}\,\exists\mathrm{E}\ a \prec \emptyset}{Rab}\,\exists\mathrm{E}\ b \prec a}{\exists xRxb\ (*)}\,\exists\mathrm{I}_I \qquad\qquad \cfrac{\cfrac{\cfrac{\cfrac{\exists x\exists yRxy}{\exists yRay}\,\exists\mathrm{E}\ a \prec \emptyset}{Rab}\,\exists\mathrm{E}\ b \prec a}{\exists xRxb}\,\exists\mathrm{I}_{II}\ a \prec b}{\exists y\exists xRxy}\,\exists\mathrm{I}_{II}$$

The left hand derivation is one in the $\exists I_I$ calculus. Here permutation is blocked by the proviso on $\exists I_I$. Because $b \prec a$, we are not allowed to abstract over a before we have abstracted over b: the order in which we eliminate proper terms should respect the *stack* of terms created by ($\exists E_I$). The right hand derivation is one in the $\exists I_{II}$ calculus. Here, nothing is blocked but we end up with a dependence relation which is circular. This derivation is then ruled out by the global restriction on dependence relations of Calculus II. So, on the permutation of existential quantifiers the two calculi agree: they both exclude this. Notice that by our definition of $AO(\varphi)$ permutation of a prefix of existential quantifiers is allowed if the bound variables do not actually occur in the matrix of the formula.

$$\dfrac{\dfrac{\dfrac{\exists x \exists y Ruv}{\exists y Ruv}\ \exists\mathrm{E}\ a \prec \emptyset}{Ruv}\ \exists\mathrm{E}\ b \prec \emptyset}{\exists x Ruv}\ \exists\mathrm{I}_I$$

Because the variable x does not occur in Ruv, $AO(\exists y Ruv) = \emptyset$.

Now consider weakening of the existential quantifier.

$$\dfrac{\dfrac{\dfrac{\exists x Rxx}{Raa}\ \exists\mathrm{E}\ a \prec \emptyset}{\exists y Rya}\ \exists\mathrm{I}_I}{\exists x \exists y Ryx}\ \exists\mathrm{I}_I \qquad \dfrac{\dfrac{\dfrac{\exists x Rxx}{Raa}\ \exists\mathrm{E}\ a \prec \emptyset}{\exists y Rya}\ \exists\mathrm{I}_{II}\ a \prec a}{\exists x \exists y Ryx}\ \exists\mathrm{I}_{II}\ a \prec a$$

The left hand derivation is correct according to the rules of Calculus I. The right hand derivation is not correct in Calculus II, for we end up with a dependence relation with a reflexive point. So Calculus II is more strict than Calculus I.

Notice that Calculus II actually consists of a family of Calculi when we vary the constraint on the eventual dependence relation. For instance, to block permutation we only have to require non-circularity of the relation, and to block weakening we only have to require irreflexivity.

What do these rules express from a choice perspective? The rule ($\exists I_I$) puts the following restriction on the availability of choice tuples: If we have $\varphi[a/x]$, then we may use the tuple $\langle \epsilon x : \varphi, a \rangle$, that is a is available to satisfy the condition $\varphi(x)$, if a does not occur in any *other* choice tuple depending on $\langle \epsilon x : \varphi, a \rangle$. So permutation is excluded because we need the tuple $\langle \epsilon x : Rxb, a \rangle$ on which the tuple $\langle \epsilon y : Ray, b \rangle$ depends. Notice that we are allowed to use the tuple $\langle \epsilon y : Rya, a \rangle$ in the derivation of weakening, because there occurs no *other* tuple depending on it in the derivation. On the other hand, Calculus II states: if we have $\varphi[a/x]$, then we may use the tuple $\langle \epsilon x : \varphi, a \rangle$ if a does not occur in any choice tuple depending on $\langle \epsilon x : \varphi, a \rangle$. Now $\langle \epsilon y : Rya, a \rangle$ may not be used, for it depends on itself.

5.8. REMARK. In the ϵ-calculus we get Calculus I by stipulating: ($\exists$I) may abstract over terms that have only surface occurrences. Calculus II we get by stipulating that (∃I) always abstracts over *all* such occurrences.

5.6 Extended Dependence Language

5.6.1 Quantifying in Dependence Structures

In this section we shall take a closer look at the interaction between quantification and dependence, by considering the extended quantificational capabilities of the epsilon calculus. We shall introduce an extended first-order language and a proof theory which mirrors the expressivity of the epsilon calculus with respect to dependence. Of course, there is more to epsilon terms than a dependence structure. They also have *logical* content, but already at the level of dependence can we gain interesting insights into hidden dependence structure of first-order logic proper.

In an ϵ-derivation the dependence between proper terms in a Prawitz style derivation is exhibited on the formula level. This means that in ϵ-derivations, the rule (∃I) may abstract over terms in a formula φ that, in (∃E)-derivations, occur outside of φ. Proof theoretically, the ϵ-calculus can express more quantificational patterns, than the standard calculus can. However, the *conservativity* of the epsilon calculus over classical first-order logic entails that quantifying in the dependence relation as will occur in the following examples of $\exists E\ \{a \prec AO(\varphi)\}$ derivations, does not increase the derivational strength of the calculus as long as we are interested in proper-term free conclusions from proper term-free assumptions.

Consider the following derivations.

$$\dfrac{\dfrac{\exists x \exists y Rxy}{\exists y Ray}\ \exists\mathrm{E}\ a \prec \emptyset}{Rab}\ \exists\mathrm{E}\ b \prec a \qquad \dfrac{\dfrac{\exists x \exists y R(x,y)}{\exists y R(\epsilon x : \exists y R(x,y), y)}\ \exists\mathrm{E}_\epsilon}{R(\epsilon x : \exists y R(x,y), \epsilon y : R(\epsilon x : \exists y R(x,y), y))}\ \exists\mathrm{E}_\epsilon$$

In the ϵ-derivation we may continue by an application of (∃I) abstracting over occurrences of $\epsilon x : \exists y R(x,y)$ in three ways:

1. We can only abstract over the surface occurrence of the term $\epsilon x : \exists y R(x,y)$.

$$\dfrac{R(\overbrace{\epsilon x : \exists y R(x,y)}, \epsilon y : R(\epsilon x : \exists y R(x,y), y))}{\exists x R(x, \epsilon y : R(\epsilon x : \exists y R(x,y), y))}\ \exists\mathrm{I}$$

This we have used in the derivation of $\exists y \exists x R(x,y)$ from $\exists x \exists y R(x,y)$.

2. We can also abstract over *both* occurrences of $\epsilon x : \exists y R(x,y)$.

$$\dfrac{R(\overbrace{\epsilon x : \exists y R(x,y)}, \epsilon y : R(\overbrace{\epsilon x : \exists y R(x,y)}, y))}{\exists x R(x, \epsilon y : R(x,y))}\ \exists \mathrm{I}$$

Now, this quantificational pattern, can be mimicked on the $(\exists E)$ side by abstracting both over the surface occurrence of the term a and over its occurrence in the dependence relation.

$$\dfrac{\dfrac{\exists x \exists y Rxy}{\exists y Ray}\ \exists \mathrm{E}\ a \prec \emptyset}{Rab}\ \exists \mathrm{E}\ b \prec a \quad \Rightarrow \quad \dfrac{\dfrac{\dfrac{\exists x \exists y Rxy}{\exists y Ray}\ \exists \mathrm{E}\ a \prec \emptyset}{Rab}\ \exists \mathrm{E}\ b \prec a}{\exists x Rxb}\ \exists \mathrm{I}_I\ b \prec x$$

3. Finally, we may abstract only over the embedded occurrence of $\epsilon x : \exists y R(x,y)$.

$$\dfrac{R(\epsilon x : \exists y R(x,y), \epsilon y : R(\overbrace{\epsilon x : \exists y R(x,y)}, y))}{\exists x R(\epsilon x : \exists y R(x,y), \epsilon y : R(x,y))}\ \exists \mathrm{I}$$

This corresponds to a $(\exists E_I)$ derivation where we *only* abstract over the occurrence of a in the dependence relation.

$$\dfrac{\dfrac{\exists x \exists y Rxy}{\exists y Ray}\ \exists \mathrm{E}\ a \prec \emptyset}{Rab}\ \exists \mathrm{E}\ b \prec a \quad \Rightarrow \quad \dfrac{\dfrac{\dfrac{\exists x \exists y Rxy}{\exists y Ray}\ \exists \mathrm{E}\ a \prec \emptyset}{Rab}\ \exists \mathrm{E}\ b \prec a}{\exists x Rab}\ \exists \mathrm{I}_I\ b \prec x$$

In the last two derivation, the 'local' quantifier $\exists x$ in the conclusion of the ϵ-derivation amounts to a 'global' existential quantifier in the derivation in the Copi-Kalish Calculus. The quantifier can bind in the dependence structure.

Quantification into the dependence structure, and substitution, will have to take into account the fact that dependence of a term a on a term b involves 'hidden' binding structure. Consider the following incorrect derivation of $\exists x \forall y Rxy$ from $\forall x \exists y Rxy$.

$$\dfrac{\dfrac{\dfrac{\dfrac{\forall x \exists y Rxy}{\exists y Ray}\ \forall \mathrm{I}}{Rab}\ \exists \mathrm{E}\ b \prec a}{\forall x Rxb}\ \forall \mathrm{I}\ b \prec x}{\exists y \forall x Rxy\ (*)}\ \exists \mathrm{I}_I\ y \prec x \qquad \dfrac{\dfrac{\dfrac{\dfrac{\forall x \exists y R(x,y)}{\exists y R(a,y)}\ \forall \mathrm{E}}{R(a, \epsilon y : R(a,y))}\ \exists \mathrm{E}_\epsilon}{\forall x R(x, \epsilon y : R(x,y))}\ \forall \mathrm{I}}{\exists y \forall x R(x,y)\ (*)}\ \exists \mathrm{I}$$

In the ϵ-derivation on the right hand side the last proof step is prevented by the fact that $\forall x R(x, \epsilon y : R(x,y))$ is not of the form $\forall x R(x,y)[\epsilon y : R(x,y)/y]$. That is, $\epsilon y : R(x,y)$ is not free for y in $\forall x R(x,y)$. In the left hand derivation the (functional) dependence of b on the variable x is witnessed by the fact that $b \prec x$.

5.6.2 An Explicit Language for Dependencies

In order to get the quantificational patterns of the epsilon calculus represented in the Copi-Kalish Calculus, we shall introduce explicitly a dependence structure as a syntactic object associated with the atomic formulas of the language. This gives us a way to manipulate dependence structure *explicitly* and to bring into formulas dependence structure 'from the outside', unrelated to the quantificational pattern of the formula. First we define the notion of a dependence structure on $\mathcal{L}$- terms

5.9. DEFINITION. (Dependence Structures for $\mathcal{L}$) A *dependence structure* R for language $\mathcal{L}$ is a tuple $\langle D_R, \ll_R \rangle$ where D_R is a finite set of $\mathcal{L}$-terms and $\ll_R$ is a binary relation on $\mathcal{L}$-terms. The term t *depends on term* t' with respect to R, notation $t \prec_R t'$, if there is a sequence $t_1 \ll_R t_2, \ldots, t_{n-1} \ll_R t_n$ such that $t = t_1$, $t' = t_n$. The set $\mathcal{R}_{\mathcal{L}}$ is the set of all dependence structures for $\mathcal{L}$.

In a dependence structure R, the set D_R contains a class of terms, and $\ll_R$ specifies dependencies between terms. Notice that D_R need not be the basis of the relation $\ll_R$, in fact, the relation $\ll_R$ may be infinite. In our calculus the set D_R will consist of existentially bound variables and the relation $\ll_R$ will record dependencies in which these terms are involved.

5.10. DEFINITION. (Dependence Language) Let $\mathcal{L}$ be a first-order language. The non-logical vocabulary of $\mathcal{L}^R$ consists of the $\mathcal{L}$ vocabulary plus an infinite set PT of new individual constants $a_1, a_2, \ldots$, the *proper terms* of $\mathcal{L}^R$. Let $R \in \mathcal{R}_{\mathcal{L}}$ be a dependence structure for $\mathcal{L}$. The language $\mathcal{L}^R$ is given by

terms $t ::= a \mid c \mid v \mid f(t_1 \cdots t_n)$.

formulas $\varphi ::= \bot : R \mid Pt_1 \cdots t_n : R \mid t_1 = t_2 : R \mid \neg\varphi \mid (\varphi \wedge \psi) \mid (\varphi \vee \psi) \mid (\varphi_1 \rightarrow \varphi_2) \mid \forall v r_v(\varphi) \mid \exists v d_v(\varphi)$.

The interaction of quantifiers and dependence structures is mediated by the two functions, r_x and d_x which are defined as follows:

1. if $\varphi : R$ is an atomic $\mathcal{L}^R$-formula, then
   (a) $d_x(\varphi : R) = \varphi : \langle D_R \cup \{x\}, \ll_R \cup \{\langle y, x\rangle \mid y \in D_R\} \cup \{\langle x, a\rangle \mid$ a a closed proper term in $\varphi\}\rangle$;

(b) $r_x(\varphi : R) = \varphi : \langle D_R, \ll_R \cup \{\langle y, x\rangle \mid y \in D_R\}\rangle$

2. d_x and r_x commute with the Boolean connectives and the quantifiers.

So, at the atomic level, d_x adds an element to the domain of the dependence structure and adds tuples to the relation $\ll$. This is used to compose dependencies resulting from existential quantification. The function r_x adds only tuples to the dependence relation and leaves the domain untouched. This is used for universal quantification. In general, wide scope of an existential quantifier means low in the dependence relation. Universally quantified terms end up on the bottom of the relation: they depend on nothing.

In this language atomic formulas have an explicit dependence structure attached to them. Quantifiers introduce and bind variables in the $\mathcal{L}$-formula part as well as the dependence part of an $\mathcal{L}^R$ formula.

Notice that, like in the epsilon calculus, atomic formulas may come with an arbitrary dependence structure already attached.

The *proper terms* of $\mathcal{L}^R$ will be treated differently from ordinary $\mathcal{L}$-constants. They will be used in derivations as proper terms of quantifier rules. In the application of these rules they will enter into dependencies.

The notion of substituting a term for a variable in an $\mathcal{L}^R$ formula has to be adapted to the presence of the dependence structures. In particular, if a proper term a depends in φ on a variable, this variable should be treated as if it occurs in the term a.

5.11. Definition. (Substitution) Let φ be an $\mathcal{L}^R$ formula. Term t is *free for variable* x in φ if

1. there is no variable y such that $x \prec_R y$ for some dependence structure R in φ,
2. no occurrence of x in φ lies within the scope of a quantifier binding a variable free in t.

The result of *substituting term t for free variable* x in φ, notation $\varphi[t/x]$, consists of the formula φ with all occurrences of x in φ replaced by t, if t is free for x in φ. A proper term a is *closed* in R if there is no variable x such that $a \prec_R x$.

In the language $\mathcal{L}^R$ we can associated with every $\mathcal{L}$ formula φ a set of of formulas by coupling the atomic formulas of φ to arbitrary dependence structures. But each formula φ of $\mathcal{L}$ has its dependence *eigen*-structure in $\mathcal{L}^R$, the structure expressing precisely the quantificational pattern of φ. This is the structure we shall introduce in the next section.

5.6.3 From First-Order Logic to Dependence Logic

Now, to get *the* dependence of an $\mathcal{L}$ formula, we have to define the translation function as follows.

5.12. DEFINITION. (Embedding $\mathcal{L}$ in $\mathcal{L}^R$) The *embedding* of $\mathcal{L}$ in $\mathcal{L}^R$ is the function $^* \in \mathcal{L}^{R^{\mathcal{L}}}$ satisfying

1. $\varphi^* = \varphi : \langle\emptyset, \emptyset\rangle$ for atomic φ,
2. * commutes with the Boolean connectives,
3. $(\exists x\varphi)^* = \exists x d_x(\varphi^*)$,
4. $(\forall x\varphi)^* = \forall x r_x(\varphi^*)$.

5.13. EXAMPLE. We shall give some examples of embeddings. In our system, the domain of a dependence structure contains only existentially quantified variables. Dependencies are only constructed with respect to elements of this domain. So universal quantifiers among themselves do not create dependencies:

$$(\forall x\forall yQxy)^* = \forall x\forall yQxy : \langle\emptyset, \emptyset\rangle.$$

Existential quantifiers create full dependence structures:

$$(\exists x\exists yQxy)^* = \exists x\exists yQxy : \langle\{x, y\}, y \ll x\rangle.$$

The interaction of universal and existential quantifier is regulated by the different ways they construct dependencies:

$$(\forall x\exists yQxy)^* = \forall x\exists yQxy : \langle\{y\}, y \ll x\rangle.$$

The existential quantifier enters the variable y in the domain. The universal quantifier constructs a dependence with respect to this variable, but does not introduce its variable in the domain. Reversing the quantifiers gives

$$(\exists x\forall yQxy)^* = \exists x\forall yQxy : \langle\{x\}, \emptyset\rangle.$$

Now the universal quantifier does not create a dependence, for it works on a domain that is (still) empty. Consequently, the existential quantifier only enters its variable in the domain.

In the presence of a proper term a the existential quantifier creates dependencies, but the universal one does not:

$$(\exists xQxa)^* = \exists xQxa : \langle\{x\}, x \ll a\rangle \qquad (\forall xQxa)^* = \forall xQxa : \langle\emptyset, \emptyset\rangle.$$

The embedding of complex formulas is straightforward as the translation only applies to atomic formulas.

$$(\exists x(Px \to \exists yQxy))^* = \exists x(Px : \langle\{x\}, \emptyset\rangle \to \exists yQxy : \langle\{x, y\}, y \ll x\rangle).$$

When we are dealing with formulas containing only existential quantifiers we shall disregard the domain and represent only the relational information. For instance, we can represent the above embedding as

$$(\exists x(Px \to \exists yQxy))^* = \exists x(Px : \emptyset \to \exists yQxy : y \ll x).$$

The embedding $*$ has the following invariant.

5.14. PROPOSITION. *For every $\mathcal{L}$-formula φ every dependence structure R occurring in φ^* is transitive.*

Proof: by induction on the complexity of φ. The AOproposition holds trivialiter for quantifier free formulas. For the quantifiers, consider the functions d_x and r_x. Let φ be transitive by induction hypothesis. Suppose $d_x(\varphi)$ or $r_x(\varphi)$ adds $\langle y, x\rangle$ to some R in φ where $\langle t, y\rangle \in \ll_R$. Then $t \in D_R$. For, by the construction of the d and r functions, whenever $\langle t', t''\rangle \in \ll_R$, then $t' \in D_R$. Consequently, $\langle t, x\rangle \in R$ by the definition of the d and r functions. ⊠

The language $\mathcal{L}^R$ thus has both dependence structures and quantification. It seems that one of both is superfluous. After all, the formula

$$Rxy : \langle\{x, y\}, y \ll x\rangle$$

can be uniquely identified as $\exists x \exists y Rxy$. But the same holds for the epsilon calculus. There every first-order formula (implicit dependencies) has a quantifier free equivalent (explicit dependencies). But the converse is not the case: not every epsilon formula has a first-order equivalent. The same holds for $\mathcal{L}^R$ formulas. The epsilon rule and the rule ($\exists$I) function in the epsilon calculus to bring us from epsilon formulas to new epsilon formulas.

$$\dfrac{\dfrac{R(\epsilon x : \varphi, \epsilon x : \psi)}{\exists x R(x, \epsilon x : \psi)}\,\exists\mathrm{I}}{R(\epsilon x : R(x, \epsilon x : \psi), \epsilon x : \psi)}\,\exists\mathrm{E}_\epsilon$$

Notice that an application of ($\exists I$) followed by an application of ($\exists E_\epsilon$) does not bring us back in the old situation. Thus the quantifier rules serve to eliminate and introduce dependence structures. This will be made explicit in the proof system for $\mathcal{L}^R$.

5.7 Extended Proof System

The proof calculus for $\mathcal{L}^R$ we shall introduce syntactically records dependencies as they arise in the course of a derivation. The dependencies are made explicit in the language. The proof calculus for $\mathcal{L}^R$ is standard for the Boolean connectives.

5.15. DEFINITION. (Quantifier rules for $\mathcal{L}^R$) For a an $\mathcal{L}^R$ proper term

$$\dfrac{\exists x \varphi}{\varphi[a/x]}\,\exists\mathrm{E} \qquad\qquad \dfrac{\varphi[a/x]}{\exists x d_x(\varphi)}\,\exists\mathrm{I}_I$$

In the existential elimination rule, the proper term a must be fresh to the derivation. In the existential introduction rule the proper term a should be free for x in φ.

For t an $\mathcal{L}^R$ term

$$\frac{\varphi}{\forall x(\varphi[x/t])}\,\forall\mathrm{I} \qquad\qquad \frac{\forall x\varphi}{(r_x(\varphi))[t/x]}\,\forall\mathrm{E}$$

In applications of the introduction rule for the universal quantifier the term t should not occur in any assumption on which the premise depends or in any domain of a dependence structure occurring in the premise.

With these proof rules we have extended the variable occurrences that can be bound by an existential quantifier. There is no choice involved in applications of $(\exists\mathrm{E}_I)$

$$\frac{\exists x\exists yRxy : \{y \ll x\}}{\exists yRay : \{y \ll a\}}\,\exists\mathrm{E}$$

All occurrences of the variable x are involved in the application of the elimination rule. But applications of $(\exists\mathrm{I}_I)$ now involve choices not present in the standard calculus.

$$\frac{\exists yRxy[a/x] : \{y \ll a\}}{\exists x\exists yRxy : \{y \ll a, y \ll x\}}\,\exists\mathrm{I}_I \qquad \frac{\exists yRxy[a/x] : \{y \ll x\}[a/x]}{\exists x\exists yRxy : \{y \ll x\}}\,\exists\mathrm{I}_I$$

$$\frac{\exists yRay : \{y \ll x\}[a/x]}{\exists x\exists yRay : \{y \ll x\}}\,\exists\mathrm{I}_I$$

Here we have three applications of $(\exists\mathrm{I}_I)$ with premise $\exists yRay : \{y \ll a\}$.

1. The first application abstracts only on the formula side, only over the *surface* occurrence of the term a. Such applications can be seen as $(\exists\mathrm{I})$ applications in the standard, dependence-free, calculus. The above application leaves us with an occurrence of a in the dependence structure. The meaning of this fact will be discussed below.
2. The second $(\exists\mathrm{I}_I)$ application above abstracts over both occurrences of the term a. Notice that this brings us back to $(\exists x\exists yRxy)^*$, the premise of the $(\exists\mathrm{E}_I)$ application with conclusion $\exists yRay : \{y \ll a\}$.
3. The third $(\exists\mathrm{I}_I)$ application above abstracts only over the dependence structure. From the perspective of the formula side, this application gives a *vacuous* quantification, but in the dependence calculus conclusion and premise of this application are not interderivable.

5.7.1 Benchmarks Once More

Weakening of the Existential Quantifier In this calculus, it is the rule ($\exists \mathrm{I}_I$) that creates dependencies. Consider the derivation of existential quantifier weakening.

$$\dfrac{\dfrac{\dfrac{\exists x Rxx : \langle \{x\}, \emptyset \rangle}{Rxa[a/x] : \langle \{a\}, \emptyset \rangle}\,\exists \mathrm{E}}{\exists x Rxy[a/y] : \langle \{x, y\}[a/y], \{x \ll y\}[a/y] \rangle}\,\exists \mathrm{I}_I}{\exists y \exists x Rxy : \langle \{x, y\}, \{x \ll y\} \rangle}\,\exists \mathrm{I}_I$$

Here we witness a splitting of D_R. This corresponds intuitively to what happens in 'deidentification'.

Universal-Existential Interaction

$$\dfrac{\dfrac{\dfrac{\dfrac{\exists x \forall y Rxy : \langle \{x\}, \emptyset \rangle}{\forall y Ray : \langle \{a\}, \emptyset \rangle}\,\exists \mathrm{E}}{Rxb[a/x] : \langle \{x\}, \emptyset \rangle [a/x]}\,\forall \mathrm{E}}{\exists x Rxb : \langle \{x\}, x \ll b \rangle}\,\exists \mathrm{I}_I}{\forall y \exists x Rxy : \langle \{x\}, x \ll y \rangle}\,\forall \mathrm{I}$$

Notice that in the above proofs we derive $(\exists y \exists x Rxy)^*$ from $(\exists x Rxx)^*$ and $(\forall y \exists x Rxy)^*$ from $(\exists x \forall y Rxy)^*$. This situation we shall identify as a restriction on *correct* derivations.

5.16. DEFINITION. (Independent derivations) For $\Sigma \cup \{\varphi\}$ a set of $\mathcal{L}$-formulas, Σ ***d*-derives** φ, notation $\Sigma \vdash^d \varphi$, if $\{\psi^* \mid \psi \in \Sigma\} \vdash \varphi^*$.

The conclusions of independent derivations show no trace of their derivational history. Notice that an independent derivation cannot have a conclusion containing a proper term. A typical case of a correct derivation in this sense is the following.

$$\dfrac{\dfrac{\dfrac{\dfrac{\exists x \exists y Rxy : \{y \ll x\}}{\exists y Ray : \{y \ll a\}}\,\exists \mathrm{E}}{Rab : \{b \ll a\}}\,\exists \mathrm{E}}{\exists y Ray : \{y \ll a\}}\,\exists \mathrm{I}_I}{\exists x \exists y Rxy : \{y \ll x\}}\,\exists \mathrm{I}_I$$

Here, we have a derivation $\mathcal{D}$ such that $R_{\mathcal{D}} = R_{\exists y \exists x Rxy}$. At the conclusion, the dependence structure is given by the translation function $*$.

Not all derivations give derivation independent conclusions. Our next example will exemplify this situation.

Permutation of Existential Quantifiers Consider a derivation of $\exists y \exists x Rxy$ from $\exists x \exists y Rxy$:

$$
\cfrac{\cfrac{\cfrac{\cfrac{\exists x\exists y Rxy : \{y \ll x\}}{\exists y Ray : \{y \ll a\}}\ \exists\mathrm{E}}{Rxb[a/x] : \{b \ll a\}}\ \exists\mathrm{E}}{\exists x Rxy[b/y] : \{b \ll a, x \ll b, x \ll a\}}\ \exists\mathrm{I}_I}{\exists y \exists x Rxy : \{b \ll a, x \ll b, x \ll a, x \ll y, y \ll b, y \ll a\}}\ \exists\mathrm{I}_I
$$

In this derivation we have left out the domain of the dependence relation for reasons of display. The dependence structure $\langle\{x, y\}, y \ll x\rangle$, implicit in the premise of the derivation, is turned into an explicit structure $b \ll a$, accompanying the formula Rab. This again leads to the structure $\psi = \exists y \exists x Rxy : \{b \ll a, y \ll b, x \ll y\}$ (where, for clarity, we leave out tuples that can be computed by transitivity, which is justified by Proposition 5.17). Now, the formula $(\exists y \exists x Rxy)^*$ has dependence structure $\langle\{x, y\}, x \ll y\rangle$, the reverse of the relation at the assumption. Notice that $(\exists y \exists x Rxy)^* \neq \psi$. Rearranging the dependence structure at ψ, we get the structure

$$x \ll y \ll b \ll a.$$

At the conclusion there is a residu of dependence structure not accounted for by $(\exists y \exists x Rxy)^*$. That is, the conclusion of this derivation still carries information about its derivational history. However, the calculus allows us to continue the derivation in order to reach a derivation independent conclusion, by abstracting only over the dependence part.

$$
\cfrac{\cfrac{\exists y \exists x Rxy : \{x \ll y \ll z \ll a\}\}[b/z]}{\exists z \exists y \exists x Rxy : \{x \ll y \ll z \ll u\}[a/u]}\ \exists\mathrm{I}_I}{\exists u \exists z \exists y \exists x Rxy : \{x \ll y \ll z \ll u\}}\ \exists\mathrm{I}_I
$$

Now $(\exists u \exists z \exists y \exists x Rxy)^* = \exists u \exists z \exists y \exists x Rxy : \{x \ll y \ll z \ll u\}$. The dependence structure at the conclusion of the derivation is independent of its derivation. In the process, there arise two existential quantifiers that are vacuous with respect to the formula part. This can be understood by comparing the complete derivation

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\exists x\exists yRxy : \{y \ll x\}}{\exists yRay : \{y \ll a\}}\,\exists\mathrm{E}}{Rxb[a/x] : \{b \ll a\}}\,\exists\mathrm{E}}{\exists xRxy[b/y] : \{x \ll b \ll a\}}\,\exists\mathrm{I}_I}{\boxed{\exists y\exists xRxy} : \{x \ll y \ll z \ll a\}[b/z]}\,\exists\mathrm{I}_I}{\boxed{\exists z\exists y\exists xRxy} : \{x \ll y \ll z \ll u\}[a/u]}\,\exists\mathrm{I}_I}{\boxed{\exists u\exists z\exists y\exists xRxy} : \{x \ll y \ll z \ll u\}}\,\exists\mathrm{I}_I$$

to the situation in a standard ($\exists$E) derivation

$$\dfrac{\exists x\exists yRxy \qquad \dfrac{\exists yRay\ [1] \qquad \dfrac{\dfrac{Rxb[a/x]\ [2]}{\exists xRxy[b/y]}\,\exists\mathrm{I}}{\boxed{\exists y\exists xRxy}}\,\exists\mathrm{I}}{\boxed{\exists y\exists xRxy}}\,\exists\mathrm{E}_2}{\boxed{\exists y\exists xRxy}}\,\exists\mathrm{E}_1$$

The 'vacuous' quantifications in the dependence derivation correspond to the two discharge actions performed in the ($\exists$E) derivation. In the latter derivation the discharge actions give rise to an $\exists y\exists xR(x,y)$ *segment* of length 2. In the former derivation we get a sequence of two 'vacuous' quantifications. The highest occurrence of $\exists y\exists xRxy$ in the ($\exists$E) derivation is not yet a conclusion of a derivation with premise $\exists x\exists yRxy$. There are still discharge actions to be performed. In the Copi-Kalish system of existential introduction these discharge actions disappear, consequently no segments arise (they do arise, however, if we embed this Copi-Kalish derivation in a standard derivation, as we have done in Section 5.2.4). Adding dependencies to this system brings the discharge actions again to the surface. This time in the form of vacuous quantification. Note that we have to abstract over the occurrences of b before we can abstract over a: in the ($\exists$E) derivation we must discharge the assumption introducing b before we can discharge the one introducing a. Abstracting first over a would leave us in a situation where abstraction over b is no longer possible:

$$\dfrac{\exists y\exists xRxy : \{x \ll y \ll b \ll u\}[a/u]}{\exists u\exists y\exists xRxy : \{x \ll y \ll \underbrace{b \ll u}\}}\,\exists\mathrm{I}_I$$

Now $\{x \ll y \ll b \ll u\} \neq \{x \ll y \ll z \ll u\}[b/z]$. For the same reason we cannot permute quantifiers by concluding from $Rab : \langle\{b\}, b \ll a\rangle$ to

$\exists xRxb : \langle\{x, b\}, b \ll x\rangle$. At this point we cannot abstract over b.

In all of the above derivations, the dependence structure is a transitive relation at every step of the derivation. This is not a coincidence.

5.17. PROPOSITION. *Let $\mathcal{D}$ be a derivation from premise in Σ such that every $\psi \in \Sigma$ is of the form φ^* for some $\mathcal{L}$-formula φ. Then at every formula occurrence χ in $\mathcal{D}$ every dependence structure occurring in χ will be transitive.*

Proof: by induction on the length of the derivation. ⊠

Vacuous Quantification With respect to dependencies, quantification that is vacuous over the formula part of a formula-dependence structure pair now acquires meaning. Standardly Pa and $\exists xPa$ are interderivable. In the dependence calculus they loose this proof theoretic equivalence.

$$\frac{\exists xPa : \{x \ll a\}}{Pa : \{b \ll a\}}\,\exists\text{E} \qquad\qquad \frac{Pa[b/x] : \emptyset}{\exists xPa : \{x \ll a\}}\,\exists\text{I}_I$$

Moreover if we continue the left hand derivation

$$\cfrac{\cfrac{\exists xPa : \{x \ll a\}}{Pa : \{b \ll a\}}\,\exists\text{E}}{\exists xPx : b \ll x}\,\exists\text{I}_I$$

we have arrived at a situation where we can no longer quantify over b (it depends on the bound variable x). Only by *reversing* the ($\exists\text{E}_I$) application, i.e., applying ($\exists$I) with respect to x, we can achieve a derivation independent conclusion. Here we see that quantification into hidden structure (resulting from a derivational history) can express 'states' of the derivation that are beyond the reach of a standard Prawitz calculus.

Liberal Substitution and Quasi-Logical Form Up until now we have $\{a \prec y\} \neq \{x \prec y\}[a/x]$ to preserve functional dependencies. But the dependence structure can be seen as exactly describing the functional dependencies among terms. So suppose we maintain for $\mathcal{L}^R$ the standard notion for a term being free for a variable in a formula: term t is *free for variable* x in φ if no occurrence of x in φ lies within the scope of a quantifier binding a variable free in t. Then the following derivation is allowed

$$\cfrac{\cfrac{\cfrac{\cfrac{\forall x\exists yRxy : \langle\{y\}, y \ll x\rangle}{\exists yRay : \langle\{y\}, y \ll a\rangle}\,\forall\text{I}}{Rab : \langle\{b\}, b \ll a\rangle}\,\exists\text{E}}{\forall xRxy[b/y] : \langle\{y\}, y \ll x\rangle[b/y]}\,\forall\text{I}}{\exists y\forall xRxy : \langle\{y\}, y \ll x\rangle}\,\exists\text{I}_I$$

Compare the conclusion of this derivation to

$$(\exists y \forall x Rxy)^* = \exists y \forall x Rxy : \langle \{y\}, \emptyset \rangle.$$

The conclusion of the derivation is associated with a dependence structure which does not match that of the embedding. In this case the dependence structure exhibits a dependence of y on x in spite of the quantifier prefix. That is, the dependence structure shows that the existentially quantified variable y should be interpreted as a *function* f_y applied to values of x. Because x does not occur in D_R we know that f_y is such that for all values t of x, $Rtf_y(t)$ holds.

We revisite the permutation of existential quantifiers, now with free substitution.

$$\dfrac{\dfrac{\dfrac{\dfrac{\exists x \exists y Rxy : \langle \{x, y\}, \{y \ll x\} \rangle}{\exists y Ray : \langle \{a, y\}, \{y \ll a\} \rangle} \exists \mathrm{E}}{Rxb[a/x] : \langle \{a, b\}, \{b \ll x\}[a/x] \rangle} \exists \mathrm{E}}{\exists x Rxy[b/y] : \langle \{x, b\}, \{y \ll x\}[b/y] \rangle} \exists \mathrm{I}_I}{\exists y \exists x Rxy : \langle \{x, y\}, \{y \ll x\} \rangle} \exists \mathrm{I}_I$$

Again the conclusion of the derivation does not match the embedding, for

$$(\exists y \exists x Rxy)^* = \exists y \exists x Rxy : \{x \ll y\}.$$

According to the dependence structure, the bound variable y in the conclusion should be interpreted as the value of a function f_y with argument x. In this case the structure shows an '$\exists\exists$' pattern: f_y is such that there is a value t of x for which $Rf(t)t$ holds.

5.8 Possible Semantics

Semantics of Statements or Semantics of Proofs? In constructing a semantics for dependence one is immediately confronted with the following fact. Dependence between terms in a proof does not reside in the *denotations* of these terms. Dependence arises in the way denotations of terms are *chosen* or *constructed* within a proof. This fits a *dynamic* or *representational* view of dependence. But as dependence is not inherent in term denotation, what kind of a semantics can we expect? It seems that the optimal semantics would be one which is closely linked to the structure of derivations. This suggests that the true semantics for dependence would be one in the vein of the Curry-Howard semantics for categorial proofs, rather than a Tarskian Semantics. Be this as it may, in this section we shall pursue a less than optimal semantics of the Tarskian kind in the form of the *arbitrary object models* of Section 2.5.

5.8.1 Arbitrary Object Semantics

Arbitrary object semantics can be closely tied to derivations through the following observation. In a natural deduction treatment of existential information, if we arrive at a formula of the form $\exists x\varphi$, we choose a fresh term a, an instance of a φ–er, and continue reasoning with $\varphi[a/x]$. Having arrived at some formula no longer specific to the instance chosen, a formula without the term a, we may take this to be a conclusion of the general information that $\exists x\varphi$. The semantics of this process can best be seen as the possibility to *expand* any model for $\exists x\varphi$ with an interpretation for a *fresh* term a such that $\varphi[a/x]$ holds on the expanded model. Any formula in the unextended language which holds on the expansion, holds on the original model. This lies behind the soundness of the elimination rule for the existential quantifier.

Whenever we have *nestings* of existential quantifiers, the elimination rule will give us a strict partial order of instances where the choice of one instance determines the possible subsequent choices. In semantic terms, nested existential quantifiers are interpreted by *sequences* of expansions of a model over fresh terms. This is the connection between the order structure of choices and their *arbitrariness.*

Arbitrary object models incorporate this structure of sequences of expansion in their domain of arbitrary objects, structured by a dependence relation. We shall repeat the definition of an arbitrary object model of Section 2.5.

5.18. DEFINITION. (Arbitrary Object Models) The quadruple

$$\mathcal{M} = \langle M, A, <, V \rangle$$

is an *arbitrary object model* if M is a first-order model, A a domain of arbitrary objects disjoint from $dom(M)$, V a set of partial functions from A in $dom(M)$, and $<$ a conversely well-founded binary relation on A.

Recall that the set V has to satisfy some closure properties. Furthermore, recall that for an arbitrary object model $\mathcal{M}$ over first-order model M and variable assignment s, the relation $\mathcal{M}, s \models \varphi$ is interpreted as: *for all* $v \in V$, $M, s, v \models \varphi$.

We shall interpret our choice processes in arbitrary object models. In a choice process, an ordered structure of choices, we can distinguish two fundamentally different aspects. Firstly, there is the pure dependence structure of a choice process. This has been proof theoretically investigated in the last section. From this perspective we disregard the conditions of the choices completely, only the *order* in which these choices are made matters. Secondly, there are the choice *conditions* involved in the process: the reasons for the choices in the first place. This difference cani, in fact, be interpreted in the distinction between *arbitrary object models* which interpret arbitrary 'somethings' without going into the nature of this 'something', and models for epsilon terms, where the internal structure of the term determines (in part) its denotation. Our definition of an arbitrary object model *realizing* a choice process will incorporate both aspects.

5.19. DEFINITION. (Choice Processes) A *choice process* is a non-empty set of choice tuples $\langle \epsilon x : \varphi, a\rangle$, where $\epsilon x : \varphi$ is a closed ϵ-term and a an individual constant. A choice process C is *arbitrary* if the choice dependence relation restricted to C is a strict partial order and no term t occurs as the value of different ϵ-terms in C.

5.20. DEFINITION. Arbitrary object model $\mathcal{M}$ is *appropriate for* choice process C if for all $\langle \epsilon x : \varphi, a\rangle \in C$ and all $c \in A$: $a < c$ iff for some term b in φ, $c = b$ or $b < c$. Model $\mathcal{M}$ *realizes* choice process C if it is appropriate for C, and for all tuples $\langle \epsilon x : \varphi, a\rangle \in C$ and all $v \in V$: $v(a) = m \iff M, s(m|x), v \models \exists x\varphi \rightarrow \varphi$.

Notice that the notion of an appropriate model is the right one for the explicit dependence logic of Section 5.2.4. There the elements of a dependence relation have no logical content. Thus the perspective of choice *tuples* does not make much sense: only the values of choices, not the choice *conditions*, occur in that logic. There is much detail to this structure, but that will not be the topic of this section. Here we are especially interested in realizability, as this notion deals with choice conditions as well as with choice values.

5.21. PROPOSITION. *If choice process C is arbitrary, then we can expand any first order model M to an* arbitrary object model $\mathcal{M}$ *realizing C.*

Proof: If choice process C is arbitrary, then any first order model can be expanded to an arbitrary object model appropriate for C (for the simple proof of this part of the Proposition we refer the reader to Fine [Fin85], Section I.7.). Now this appropriate model can be supplied with *intensional choice functions* to interpret the ϵ-terms (see Chapter 2, Section 2.3.3). These choice functions can be used to define V, the set of value assignments of the model. ⊠

In the tuples in C that are realized in a model $\mathcal{M}$ the ϵ-terms are really interpreted as ϵ-terms should be. This suggests our basic definition for the substructural existential quantifier.

5.22. DEFINITION. For $\mathcal{M}$ an arbitrary object model and s a variable assignment, the relation $\mathcal{M}, s \models \langle \epsilon x : \varphi, a\rangle$ holds if $\mathcal{M}$ realizes $\langle \epsilon x : \varphi, a\rangle$.

5.23. DEFINITION. (Primitive Existential Quantifier) Let $\mathcal{M}$ be an arbitrary object model. Then

$$\mathcal{M}, s \models \mathcal{E}x\varphi \iff \mathcal{M}, s \models \langle \epsilon x : \varphi, a\rangle \ \& \ \mathcal{M}, s \models \varphi[a/x].$$

(This definition should be compared with the minimal existential quantifier of *Modal State Semantics*, [Ben94b], [Ben94a] and the one from the van Lambalgen-Alechina framework [AvL95], see Appendix II.) So $\mathcal{M}, s \models \varphi[a/x]$ is not sufficient to to conclude $\mathcal{E}x\varphi$. As yet, this does not give us much of a logic for the existential quantifier. In fact, why would we call this an existential quantifier at all? Well, at least we have $\mathcal{M}, s \models \mathcal{E}x\varphi \Rightarrow \mathcal{M}, s \models \exists x\varphi$. But this does not yet distinguish it from the universal quantifier. However, by definition, if $\langle \epsilon x : \varphi, a\rangle$ is realized on $\mathcal{M}$ but $\mathcal{M}, s \models \mathcal{E}x\varphi$ does not hold, then $\mathcal{M}, s \models \forall x\neg\varphi$. This does not square with the universal quantifier. To determine the quantifier $\mathcal{E}$ further, what we need are *closure properties* on the set of realizable choices by turning ever more internal structure of the choice conditions, the ϵ-terms, into choice parameters.

Here is a way to do this. Suppose $\mathcal{M}, s \models \mathcal{E}x(\varphi \vee \psi)$, where $\langle \epsilon x : (\varphi \vee \psi), a\rangle \in C$, and suppose $\mathcal{M}, s \models \varphi[a/x]$. In that case we want to have $\mathcal{M}, s \models \mathcal{E}x\varphi$ to get one of the typical properties of the existential quantifier. To achieve this we need a closure property, because in arbitrary C we need not find the right choice tuple. We can state our requirement:

If $\langle \epsilon x : (\varphi \vee \psi), a\rangle \in C$ and $\mathcal{M}, s \models \varphi[a/x]$, then $\mathcal{M}, s \models \varphi[b/x]$ for the unique b such that $\langle \epsilon x : \varphi, b\rangle \in C$.

Let $\overline{\in} \subseteq C \times C$ be a binary relation on a set of choice tuples C. We shall denote C ordered by $\overline{\in}$ by $C_{\overline{\in}}$.

5.24. DEFINITION. *An arbitrary object model $\mathcal{M}$ realizes $C_{\overline{\in}}$ if it realizes C and $\langle \epsilon x : \varphi, a\rangle \overline{\in} \langle \epsilon x : \psi, b\rangle$ implies $VR_a \subseteq VR_b$.*

(See Section 2.5 for a definition of the *value range* VR_a.)

5.25. DEFINITION. For any set of L-assumptions Σ and choice process C,

$$\Sigma, C_{\overline{\in}} \models \langle \epsilon x : \varphi, a\rangle$$

if for all models $\mathcal{M}$ for Σ realizing $C_{\overline{\in}}$ there is a $\langle \epsilon x : \varphi, b\rangle \in C_{\overline{\in}}$ such that $\langle \epsilon x : \psi, a\rangle \overline{\in} \langle \epsilon x : \varphi, b\rangle$ for the unique $\epsilon x : \psi$ such that $\langle \epsilon x : \psi, a\rangle \in C$.

The relation '$\overline{\in}$' is a partial order which can be strengthened to an equivalence relation, or to one including even the complete dependence structure of terms.

How does this semantics connect to the proof theory of the preceding sections? With every derivation $\mathcal{D}$ with conclusion φ and assumptions Σ in the Copi-Kalish Calculus we can associate a choice process C, a set of *assumptions* of the form $\langle \epsilon x : \varphi, a\rangle$ for every application of $(\exists E_I)$ in $\mathcal{D}$ with premise $\exists x\varphi$ and proper term a. This gives a realizable process. (See Fine [Fin85] pp.111-112, where this is done in terms of *definitional systems* which are notational variants of our choice processes.) The set C of assumptions is *dischargeable* at the conclusion φ if and only if C is realizable on any model for Σ.

Thus we can interpret the rule $(\exists E)$ as being of the form

$$\frac{\exists x\varphi \quad \langle \epsilon x : \varphi, a\rangle}{\varphi[a/x]}$$

where $\langle \epsilon x : \varphi, a \rangle$ is an assumption to be discharged. For a more extensive discussion of this point of view, see [MV95].

Now if we go substructural and require choice tuples also for the rule ($\exists$I), then the associated choice process C may no longer be realizable on any model for the assumptions Σ of $\mathcal{D}$. An example being the conclusion of $\exists x Rxa$ from premise Raa. This requires the tuple $\langle \epsilon x : Rxa, a \rangle$.

Thus we can interpret the rule ($\exists$I) as being of the form

$$\frac{\varphi[a/x] \quad \langle \epsilon x : \varphi, a \rangle}{\exists x \varphi}$$

Again $\langle \epsilon x : \varphi, a \rangle$ is an assumption to be discharged.

In order to maintain realizable choice processes (i.e., dischargeable assignments) we need principles which derive choice tuples from assumptions in Σ and choice tuples in C. Here are the principles that allow us to deal with the benchmark problems:

1. From $\langle \epsilon x : \varphi, a \rangle$ and $\varphi[a/x] \vee \psi[a/x]$ conclude $\langle \epsilon x : (\varphi \vee \psi), a \rangle$.
   From $\langle \epsilon x : (\varphi \vee \psi), a \rangle$ and $\varphi[a/x]$ conclude $\langle \epsilon x : \varphi, a \rangle$.
2. From $\langle \epsilon x : \varphi, a \rangle$ conclude $\langle \epsilon y : (\varphi[y/x]), a \rangle$.
3. From $\mathcal{E} x \varphi[a/y]$ and $\langle \epsilon x : \varphi[a/y], a \rangle$ conclude $\langle \epsilon x : (\varphi[x/y]), a \rangle$.
   From $\mathcal{E} x (\varphi[y/x])$ and $\langle \epsilon x : (\varphi[x/y]), a \rangle$ conclude $\langle \epsilon x : (\varphi[a/y]), a \rangle$.
4. From $\langle \epsilon x : \varphi[b/y], a \rangle$ and $\langle \epsilon x : \psi, b \rangle$ conclude $\langle \epsilon x : \varphi[\epsilon x : \psi/y], a \rangle$.
5. From $\langle \epsilon x : \varphi[\epsilon y : \varphi/y], a \rangle$ conclude $\langle \epsilon x : \mathcal{E} y \varphi, a \rangle$.
   From $\langle \epsilon x : \mathcal{E} y \varphi, a \rangle$ conclude $\langle \epsilon x : \varphi[\epsilon y : \varphi/y], a \rangle$.
6. From $\forall x \varphi$ and $\langle \tau x : \varphi[\epsilon y : \varphi/y], a \rangle$ conclude $\langle \epsilon x : \varphi[\epsilon y : \varphi/y], a \rangle$.
7. From $\mathcal{E} x \forall y \varphi$ and $\langle \epsilon x : \forall y \varphi, a \rangle$ conclude $\langle \epsilon x : \varphi[b/y], a \rangle$ for all b.

These principles correspond to various properties of $\overline{\in}$ as defined above. By rule (1) we have distribution of the existential quantifier. These rules together turn $\overline{\in}$ into an equivalence relation (for substructural rules dealing with *monotonicity* we refer the reader to the next chapter, Section 6.10 and [MV93]). By the second rule we may change the bound variable arbitrarily. Under this rule, also the complete dependence structure is preserved (not only the value range). Rule (3) expresses subordination principles dealing with weakening and permutation of the existential quantifier. Rule (4) expresses a necessary substitution property. It does not put any constraint on realizing models. Rule (5) gives the epsilon equivalences. Rule (6) and (7) deal with the interaction of $\mathcal{E}$ with the 'real' universal quantifier.

To conclude we shall show the application of these principles in the case of weakening and permutation.

First weakening of existential quantifier. This is essentially derivable by rules under (3):

$$\dfrac{\dfrac{\mathcal{E}xR(x,x) \qquad \langle \epsilon xR(x,x),a\rangle\ (1)}{R(a,a)} \qquad \dfrac{\dfrac{(1)}{\langle \epsilon xR(x,a),a\rangle}}{}}{\mathcal{E}xR(x,a)} \qquad \begin{array}{c}(1),\langle \epsilon yR(y,y),a\rangle \\ \vdots \\ \langle \epsilon y\mathcal{E}xR(x,y),a\rangle\end{array}$$

$$\dfrac{\dfrac{\mathcal{E}xR(x,a) \qquad \langle \epsilon y\mathcal{E}xR(x,y),a\rangle}{\mathcal{E}y\mathcal{E}xR(x,y)}}{\mathcal{E}y\mathcal{E}xR(x,y)\ (-1)}$$

The necessary choice tuple can be derived: we get $\langle \epsilon xR(x,a),a\rangle$ from $\langle \epsilon xR(x,x),a\rangle$ by (3) straightforwardly, given the premise $\langle \epsilon y\exists xR(x,y),a\rangle$. Here is the derivation:

$$\dfrac{\dfrac{\dfrac{\dfrac{\langle \epsilon yR(y,y),a\rangle}{\langle \epsilon yR(a,y),a\rangle} \qquad \langle \epsilon xR(x,a),a\rangle}{\langle \epsilon yR(\epsilon xR(x,a),y),a\rangle}}{\langle \epsilon yR(\epsilon xR(x,y),y),a\rangle}}{\langle \epsilon y\mathcal{E}xR(x,y),a\rangle}$$

$\langle \epsilon yR(y,y),a\rangle$ follows from $\langle \epsilon xR(x,x),a\rangle$ by Rule 2.

Now for permuation of existential quantifiers. We derive $\mathcal{E}x\mathcal{E}yR(x,y) \rightarrow \mathcal{E}y\mathcal{E}xR(x,y)$ as follows: let $\mathcal{D}$ be the derivation

$$\dfrac{\dfrac{\mathcal{E}x\mathcal{E}yR(x,y) \quad \langle \epsilon x\mathcal{E}yR(x,y),a\rangle\ (1)}{\mathcal{E}yR(a,y)} \qquad [\langle \epsilon yR(a,y),b\rangle]\ (2)}{R(a,b)}$$

Then this gives our result:

$$\dfrac{\dfrac{\dfrac{\dfrac{\begin{array}{c}\mathcal{D}\\ \\ R(a,b)\end{array} \quad \begin{array}{c}(1,2)\\ \\ \langle \epsilon xR(x,b),a\rangle\ (3)\end{array}}{\mathcal{E}xR(x,b)} \qquad \begin{array}{c}(1,2)\\ \vdots \\ \langle \epsilon y\mathcal{E}xR(x,y),b\rangle\ (4)\end{array}}{\mathcal{E}y\mathcal{E}xR(x,y)}}{\mathcal{E}y\mathcal{E}xR(x,y)\ (-2)}}{\mathcal{E}y\mathcal{E}xR(x,y)\ (-1)}$$

Without any but the propositional rules, $\mathcal{E}y\mathcal{E}xR(x,y)$ follows only from

$$\{\mathcal{E}x\mathcal{E}yR(x,y),(1),(2),(3),(4)\},$$

for none of these assumptions can be discharged at the conclusion. However, the subordination rules together with the substitution rules and the epsilon equivalences, allow us to derive permutation of existential quantifiers, by deriving (3) and (4) from (1) and (2). Here is the derivation:

$$\dfrac{\dfrac{\dfrac{\dfrac{\langle\epsilon x\mathcal{E}yR(x,y),a\rangle\ (1)}{\langle\epsilon xR(x,\epsilon yR(x,y)),a\rangle}}{\langle\epsilon xR(x,\epsilon yR(a,y)),a\rangle} \quad \langle\epsilon yR(a,y),b\rangle\ (2)}{\langle\epsilon xR(x,b),a\rangle\ (3)} \quad \langle\epsilon yR(a,y),b\rangle\ (2)}{\dfrac{\dfrac{\langle\epsilon yR(\epsilon xR(x,b),y),b\rangle}{\langle\epsilon yR(\epsilon xR(x,y),y),b\rangle}}{\langle\epsilon y\mathcal{E}xR(x,y),b\rangle\ (4)}}$$

5.9 Links to Linguistic Applications

In this and the previous chapter we have made a plea for treating the notion of 'dependence' as a core concept of logic, and we have seen that important aspects of it can be isolated and studied in connection with the analysis of proof structure.

Dependencies show up in linguistics in a number of ways. 'Dependence' or 'coherence' is what is holding a string of words together and makes it into a sentence, or what glues a sequence of sentences together to make up a text. Indeed, the field of linguistics can be defined as the study of dependence or coherence in sequences of sound patterns, strings of words or sequences of sentences. Indeed, the spelling out of dependencies can be viewed as charting out specific kinds of coherence.

Anaphoric reference resolution for pronouns can be viewed as creating dependence links between pronouns and their antecedents. Scoping resolution can be viewed as spelling out the dependencies between terms representing the quantifiers. There is a close connection with instantial logic here, for epsilon terms are a direct kin of the *qterms* (unscoped terms in underspecified logical form representations) used in Alshawi c.s. [EA91]. Resolution of underspecified plural references (as collective, distributive, cumulative, etcetera) can be viewed as charting the internal dependencies within the reference set of a plural expression (see Carpenter [Car94] for a proof theoretic account of this). Reference resolution for underspecified relations (e.g., the *ownership* relations expressed by possessive pronouns) and of other forms of underspecification may use *coordinated choices* which also create dependencies (see Alshawi c.s. [EA91], Van Deemter [Dee91]). Some of the connections between dependencies in instantial logic and dependencies in linguistics (in connection with anaphoric linking, plurality and genericity) are taken up in the next chapter.

5.10 Conclusion

We have discussed a variety of proof systems each of which makes dependence explicit in its own way. We have reached no definite results about the relation between these systems, nor about the most suitable semantics. However it is clear that there is a rich area to be explored beneath the standard quantifiers and 'dependence' is the notion by which this area can be charted. Throughout this chapter we have referred to the notion of a choice process as the one tying the different aspects of dependence together. It is an open question how this notion is connected to some quite different logical frameworks like [BA93], [Ben94b] and [AvL95] which also deal with the substructure of quantifiers. This connection is the topic of work in progress.

5.11 Appendix I

In this chapter, we have mainly dealt with proofs in a natural deduction set-up. In this appendix we shall briefly discuss a way in which the choice perspective on substructural quantifiers can be incorporated into the sequent calculus.[2] Here, the anchor point for structural rules seems to lie in the interaction between substitution and quantification. We shall suggest a way how substitution can be incorporated into the sequent calculus. We shall make substitution explicit by rules moving the substitution boxes on top of the sequent arrow. Attached to formulas, the substitution boxes are metalinguistic devices (they are not part of the formula syntax), but on top of the sequent arrow, they belong to the logic proper. We consider the set S of finite sequences of substitution boxes $\langle[x_1/y_1],\ldots,[x_n/y_n]\rangle$. The empty sequence, 0, is included in this set. If $s \in S$ then the variable x occurs *free* in s if there is variable y such that $[x/y]$ occurs in s and there is no $[z/x]$ in s following $[x/y]$ in s.

Axioms Axioms are all sequents of the form $\Delta \overset{0}{\Rightarrow} \Delta$. The valid sequents are all sequents of the form $\Delta \overset{0}{\Rightarrow} \Gamma$ that can be derived form axioms by means of the standard rules for the Boolean connectives, and the following special rules.

Substitution Rules For s a finite sequence of substitution boxes we have the following 'left' and 'right' rules:

$$\frac{\Delta \overset{s\cdot[t/y]}{\Rightarrow} \Gamma}{\Delta[t/y] \overset{s}{\Rightarrow} \Gamma[t/y]} \qquad \frac{\Delta[t/y] \overset{s}{\Rightarrow} \Gamma[t/y]}{\Delta \overset{s\cdot[t/y]}{\Rightarrow} \Gamma}$$

[2]For a different approach to substructural quantifiers in the setting of the sequent calculus, see [AvL95].

Quantifier rules For y not free in Γ, Δ or s

$$\text{L}\exists \frac{\Gamma, \varphi \overset{s\cdot[y/x]}{\Rightarrow} \Delta}{\Gamma, \exists x\varphi \overset{s\cdot[y/x]}{\Rightarrow} \Delta} \qquad \text{R}\exists \frac{\Gamma \overset{s\cdot[t/z]}{\Rightarrow} \varphi, \Delta}{\Gamma \overset{s\cdot[t/z]}{\Rightarrow} \exists z\varphi, \Delta}$$

$$\text{L}\forall \frac{\Gamma, \varphi \overset{s\cdot[t/z]}{\Rightarrow} \Delta}{\Gamma, \forall x\varphi \overset{s\cdot[t/z]}{\Rightarrow} \Delta} \qquad \text{R}\forall \frac{\Gamma \overset{s\cdot[y/x]}{\Rightarrow} \varphi, \Delta}{\Gamma \overset{s\cdot[y/x]}{\Rightarrow} \forall x\varphi\Delta}$$

This can be considered as the *minimal* system without structural rules for substitution. We can derive weakening of the existential quantifier.

5.26. EXAMPLE. (Weakening)

$$\cfrac{\cfrac{\cfrac{\cfrac{\cfrac{\cfrac{\cfrac{\cfrac{Rxx[a/y][a/x] \Rightarrow Rxy[a/y][a/x]}{Rxx[a/y] \overset{[a/x]}{\Rightarrow} Rxy[a/y]}}{Rxx \overset{[a/x][a/y]}{\Rightarrow} Rxy}}{Rxx \overset{[a/x][a/y]}{\Rightarrow} \exists yRxy}}{Rxx[a/y] \overset{[a/x]}{\Rightarrow} \exists yRxy[a/y]}}{Rxx \overset{[a/x]}{\Rightarrow} \exists x\exists yRxy}}{\exists xRxx \overset{[a/x]}{\Rightarrow} \exists x\exists yRxy}}{\exists xRxx[a/x] \Rightarrow \exists x\exists yRxy[a/x]}}{\exists xRxx \Rightarrow \exists x\exists yRxy}$$

But we cannot derive permutation of existential quantifiers. Here is a possible structural rule to solve this problem.

Possible Structural Rule

$$\frac{\Gamma \overset{s\cdot[a/x][b/y]}{\Rightarrow} \Delta}{\Gamma \overset{s\cdot[b/y][a/x]}{\Rightarrow} \Delta}$$

And here is a derivation using this rule.

5.27. EXAMPLE. (Permutation)

$$
\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{Rxy[a/x][b/y] \Rightarrow Rxy[a/x][b/y]}{Rxy \overset{[b/y][a/x]}{\Rightarrow} Rxy}}{Rxy \overset{[b/y][a/x]}{\Rightarrow} \exists x Rxy}}{Rxy[a/x] \overset{[b/y]}{\Rightarrow} \exists x Rxy[a/x]}}{Rxy[a/x] \overset{[b/y]}{\Rightarrow} \exists y \exists x Rxy}}{Rxy[a/x] \overset{[b/y]}{\Rightarrow} \exists y \exists x Rxy[a/x]}}{Rxy \overset{[b/y][a/x]}{\Rightarrow} \exists y \exists x Rxy}}{Rxy \overset{[a/x][b/y]}{\Rightarrow} \exists y \exists x Rxy \ (*)}}{\exists y Rxy \overset{[a/x][b/y]}{\Rightarrow} \exists y \exists x Rxy}}{\exists x \exists y Rxy \overset{[a/x][b/y]}{\Rightarrow} \exists y \exists x Rxy}}{\exists x \exists y Rxy[b/y] \overset{[a/x]}{\Rightarrow} \exists y \exists x Rxy[b/y]}}{\exists x \exists y Rxy[a/x] \Rightarrow \exists y \exists x Rxy[a/x]}}{\exists x \exists y Rxy \Rightarrow \exists y \exists x Rxy}
$$

In the starred line of the derivation, the structural rule is used.

5.12 Appendix II

In this section we shall give the basic definition of the existential quantifier in *model state semantics* and in the van Lambalgen-Alechina framework. These definitions give the 'minimal existential quantifier' and should be compared to our Definition 5.23.

5.28. DEFINITION. (Existential Quantifier in Modal State Semantics) A model in modal state semantics consists of a pair $\langle M, \mathcal{R} = \{R_x \mid x \in VAR\}\rangle$, where M is a first-order model and for every variable x, R_x is a binary relation on the set of variable assignments over M. The truth definition for the Booleans is standard. For the existential quantifier the definition is as follows.

$$M, \mathcal{R}, s \models \mathcal{E}x\varphi \iff \text{for some } s' : R_x ss' \;\&\; M, \mathcal{R}, s' \models \varphi.$$

The relations R_x constitute an abstract version of the usual relation $\equiv_x$ of being-an-x-variant-of. Universal validities of this logic are all classical Boolean propositional laws, the modal distribution axiom $\mathcal{E}x(\varphi \vee \psi) \leftrightarrow \mathcal{E}x\varphi \vee \mathcal{E}x\psi$, the rule of modal necessitation, if $\models \varphi$ then $\models \neg\mathcal{E}x\neg\varphi$, and the definition of $\forall x\varphi$ as $\neg\mathcal{E}x\neg\varphi$.

More fine grained logics arise by requiring that the accessibility relations R_x satisfy further conditions (for more information, see [Ben94b] and [ANvB94]).

In the generalized quantification framework of van Lambalgen-Alechina the minimal existential quantifier is again defined with respect to a first-order model and a relation. This time, however, this is a relation of arbitrary arity on the domain of the model.

5.29. DEFINITION. (Existential Quantifier in van Lambalgen-Alechina Logic) In the logic of van Lambalgen-Alechina, a proof theoretic investigation is undertaken of the substructure of quantification by varying the structural rule of *substitution*. A general semantics for the resulting logics is supplied by extensions of first-order models M with a relation R of arbitrary arity on the domain of M. The truth definition for the Booleans is standard on these models. For $\mathcal{E}x\varphi(x,\overline{y})$ an existential formula where $\overline{y}$ contains all free variables, the truth definition is as follows:

$$M, R, s \models \mathcal{E}x\varphi(x,\overline{y})$$

iff

$$\text{for some } s', s \equiv_x s' : (M, R, s' \models R(x,\overline{y}) \Rightarrow M, R, s' \models \varphi(x,\overline{y})).$$

In the interpretation of the existential quantifier the domain element chosen for the existentially bound variable must stand in the R-relation to the free variables in the formula. Various proof theoretic systems are shown to be sound and complete with respect to models of this kind, where the relation R satisfies a variety of constraints. For more information, we refer the reader to [BA93], [Lam91] and [AvL95].

It is an open question, the subject of work in progress, how these semantics relate to the one proposed in this chapter in terms of choice processes.

Chapter 6

Epsilon Terms in Natural Language Analysis

In this Chapter we explore applications of instantial logic to the problem of representing anaphoric links in natural language semantics, to the representation of plural noun phrases, and to the representation of generic noun phrases (generic uses of bare plurals). First, we explore how far classical epsilon logic, as it was defined in Chapter 2, gets us. It will turn out that this allows us to deal unbounded anaphoric linking and donkey pronouns in universal and conditional contexts. We then extend the discussion to cases of plural anaphora. We give a simple representation language for the singular/plural distinction (distributive cases of plurality only). A logic with singular and plural epsilon terms is proposed. At the end of the chapter, we briefly consider the problem of interpreting generics. We shall hint at the possibility of using instantial logic as one of the building blocks for a theory of genericity.

6.1 Introduction

Three applications of instantial logic to natural language semantics will be explored in this chapter:

1. an application to the mechanism of anaphoric linking, where we shall sketch the rudiments of a theory of pronoun resolution on the basis of instantial logic,
2. an application to the representation of plurality,
3. an application to the representation and interpretation of genericity.

The aim of this chapter is to demonstrate that instantial logic carries a promise for each of these application areas. Further work along the lines sketched below will have to prove, in each case, that using instantial logic as a tool will enable us to build theories that can compete with the best semantic theories in these areas that are currently available.

For the case of anaphoric linking, it is obvious that a theory which focuses on the representation of indefinite (and definite) terms can be brought to bear to problems of anaphora. This is along the general lines of the tradition inaugurated by Kamp [Kam81] and Heim [Hei82], where representations of indefinite and definite noun phrases turn out to have a different status from those of quantified noun phrases in that they introduce discourse referents or are linked to such referents. Instead of discourse referents, instantial logic has epsilon terms. Instead of anaphoric reference resolution by means of establishing links to discourse referents, instantial logic could proceed by finding a suitable epsilon term in the preceding context, or constructing such an epsilon term from the context, and using that term to resolve the pronoun meaning.

This method of interpreting pronouns by finding suitable descriptions in the antecedent context and using these descriptions to flesh out the pronoun meaning also has a respectable history. Both Geach [Gea80] and Evans [Eva80] propose it for the resolution of what they call 'pronouns of laziness' and 'E-type pronouns', respectively.

Heim, in a recent article, has drawn attention to the similarities and differences between an E-type approach and a Discourse Representation (or File Change Semantics) approach to anaphora. According to her, sentences like (6.1) are crucial for making a theoretical choice between the two approaches.

6.1. *When a bishop meets another bishop he blesses him.*

This kind of example provides a case against an E-type approach, she maintains, because there need not be a unique description of any of the two bishops. The example is constructed in such a way that both clergymen are interchangeable. (See also Kadmon [Kad87] for this issue of uniqueness of reference.)

It would seem that the problem is caused by the fact that the translation Heim proposes uses definite rather than indefinite description. There are two bishops satisfying the description *bishop who blesses another bishop.* Instantial logic suggests we use epsilon terms, with a dependence between them to keep the two bishops apart. Using indefinite descriptions is unacceptable to Heim because the 'familiarity theory of definiteness' she takes as her starting point. What instantial logic does is provide us with 'familiar indefinites' in the form of epsilon terms created from context. Thus, instantial logic urges us to replace the 'familiarity theory of definiteness' with a 'committed choice theory of definites', so to speak.

Instantial logic, when taken by itself, is not a theory of pronominal reference resolution. Nor is it a theory of plural reference or a theory of generic interpretation. Rather, as we hope to demonstrate in this chapter, it provides a sound basis for building such theories. We shall show the following:

1. By allowing instantial terms in the representation of natural language meaning we can formulate natural conditions for processing these representations to generate suitable antecedents for pronoun resolution.
2. There is a natural extension of instantial logic to plural terms; we can use this to deal with plural pronoun resolution.
3. The interpretation of generic terms can make use of a natural modification of epsilon logic which results from incorporating certain principles of non-monotonic reasoning.

However, in this chapter we shall *not* work out these sketches into full-fledged theories of anaphora resolution, plural reference or genericity. The aim of the chapter is merely to show how instantial logic points the way and suggests avenues for each of these areas.

6.2 Noun Phrases, Pronouns and ϵ-Terms

In the epsilon calculus first-order formulas can be reduced to quantifier-free form by the introduction of epsilon- and tau-terms. These terms can be used in the meaning representation of sentences of natural language to represent a variety of noun phrases.

They incorporate information about the context in which they occur. Let's look at some options this calculus gives us for the translation of the following simple sentence.

6.2. *A man loved a woman.*

A first-order translation could be either (6.3) or (6.4).

6.3. $\exists x(Mx \wedge \exists y(Wy \wedge Lxy))$.

6.4. $\exists y(Wy \wedge \exists x(Mx \wedge Lxy))$.

If we reduce (6.3) to pure term form, with $a = \epsilon x : (Mx \wedge \exists y(Wy \wedge Lxy))$ we get, by applying the epsilon rule:

6.5. $Ma \wedge W(\epsilon y : (Wy \wedge Lay)) \wedge La(\epsilon y : (Wy \wedge Lay))$.

We observe that the term corresponding to 'a man' has the shape $\epsilon x : (Mx \wedge \exists y(Wy \wedge Lxy))$ and the term for 'a woman' the shape $\epsilon y : (Wy \wedge Lay)$. The latter term *depends* on the former. Thus we generate in fact terms corresponding to the NPs 'a man such that there is a woman he loves' and a 'woman loved by a man such that there is a woman he loves'. But (6.3) is equivalent in classical logic to (6.4). By applying the ϵ-rule to (6.4) we get the terms $a = \epsilon y : (Wy \wedge \exists x(Mx \wedge Lxy))$ and $\epsilon x(Mx \wedge Lxa)$ in the following reduction:

6.6. $Wa \wedge M(\epsilon x : (Mx \wedge Lxa)) \wedge L(\epsilon x : (Mx \wedge Lya))a$.

This time the term corresponding to 'a man' depends on the term corresponding to 'a woman': the dependencies are reversed.

These translations are derived from the first-order formulas (6.3) and (6.4) by applying the ϵ-rule. But the ϵ-calculus gives us the possibility to translate (6.2) in ways that have no first-order equivalent. Consider formula (6.7).

6.7. $\exists x Mx \wedge (\exists y Wy \wedge L(\epsilon x : Mx)(\epsilon y : Wy))$.

In this formula there are no dependencies between the terms corresponding to 'a man' and 'a woman'. Notice that the existential quantifiers of the subformulas $\exists x Mx$ and $\exists y Wy$ in (6.7) do not have scope over the subformula $L(\epsilon x : Mx)(\epsilon y : Wy)$. Nevertheless, we claim the following.

6.1. CLAIM. *Formula* (6.7) *has exactly the same ϵ-free consequences as formulas* (6.3) *and* (6.4). *Formula* (6.7) *cannot be derived from* (6.3) *or* (6.4).

Proof: We could refer to Chapter 2, Proposition 2.28 for a proof, but we shall spell out the simple details. For the first statement, note that (6.7) derives both (6.3) and (6.4) above: by using the equivalence of $\exists x\varphi$ and $\varphi[\epsilon x : \varphi/x]$, we can first derive

$$\exists x Mx \wedge \exists y(Wy \wedge L(\epsilon x : Mx)y)$$

from (6.7) and then (6.3). So if (6.3) derives ψ, then (6.7) derives ψ. On the other hand, suppose an ϵ-free formula ψ cannot be derived from (6.3). By first-order completeness, we can find a model M for (6.3) falsifying ψ. Now, the essential insight is that any model for (6.3) can be supplied with a choice function Φ to verify (6.7). Consequently, because ψ has no ϵ-symbols, M, Φ is a model for (6.7) falsifying ψ. By completeness of the ϵ-calculus, ψ cannot be derived from (6.7).

For the second statement take a model-choice function pair M, Φ on which (6.3) is true such that Φ assigns to $\epsilon x : Mx$ and $\epsilon y : Wy$ a pair of elements not in the 'love' relation. Such a pair can be found in a suitably 'realistic' model M where not every man loves every woman. By completeness of the epsilon calculus it then follows that (6.3) does not derive (6.7). ⊠

Consequently, the epsilon calculus gives us a representational medium which is richer than first-order logic proper. For instance, here is a quite different formula with the same ϵ-free consequences as (6.3) and (6.4).

Note that when we claim in this chapter that a certain formula φ has the same ϵ-free consequences as some ϵ-free formula ψ, then the proof of this claim generally involves two insights. Firstly, φ derives ψ and, secondly, any model for ψ can be supplied with a choice function verifying φ.

We have not yet exhausted the expressive possibilities of the ϵ-calculus with respect to translating sentence (6.2). Here is a translation of a completely different nature.

6.8. $\exists x Mx \wedge (W(\epsilon y : L(\epsilon x : Mx)y)) \wedge \exists y L(\epsilon x : Mx)y$.

Again, this formula has the same ϵ-free consequences as (6.3). This formula we can roughly render as 'A man loves someone and this someone loved by that man is a woman'. Note that in formula (6.7) the predicate 'woman' occurs under the scope of an ϵ-symbol, it is part of the *identifier* of an ϵ-term. In formula (6.8) the same predicate is used as a condition on the term $\epsilon y : L(\epsilon x : Mx)y$. These two options are always there when we translate sentences of natural language into formulas of the epsilon calculus.

Now consider the following universal-existential sentence.

6.9. *Every man loved a woman.*

A first-order translation could be either of the following two:

6.10. $\forall x(Mx \rightarrow \exists y(Wy \wedge Lxy))$.

6.11. $\exists y(Wy \wedge \forall x(Mx \rightarrow Lxy))$.

This time, the formulas are not logically equivalent. For $a = \tau x : (Mx \rightarrow \exists y(Wy \wedge Lxy))$ the pure term form of (6.10) is (6.12):

6.12. $Ma \rightarrow (W(\epsilon y : (Wy \wedge Lay)) \wedge La(\epsilon y : (Wy \wedge Lay)))$.

We see that the term corresponding to the NP 'a woman' depends on the term for the NP 'every man'. Again the terms translate in fact more complex NPs. Again we can get epsilon formulas which have the same first-order consequences as (6.10), but which are not derivable from it. Here are two of them:

6.13. $\forall x(Mx \rightarrow (\exists y(Wy \wedge Lay)))$,

6.14. $\exists x Mx \rightarrow (\exists y Lay \wedge W(\epsilon y : Lay))$.

We may conclude from these examples that representations with epsilon terms provide us with intriguing new options. In the next section we shall further explore these possibilities.

6.3 Pronouns and Epsilon Terms

In this section we shall illustrate the uses of ϵ- and τ-terms for the treatment of intersentential and intrasentential *pronouns*.[1] We shall investigate the distribution and interpretation of pronouns under the restriction that *only* pronouns can be rendered as ϵ- or τ-terms. This we shall do by extending the semantic part of a standard sample grammar with semantic representation of pronoun NPs. This grammar is taken from Jan van Eijck's lecture notes [Eij95]. Every syntax rule has a semantic counterpart to specify how the meaning representation of the

[1] For a different use of ϵ-terms for the treatment of E-type pronouns, see [Nea90].

whole is built from the meaning representations of the components. X is always used for the meaning of the whole, and X_n refers to the meaning representation of the n-th daughter.

S	::=	**NP VP**	X	::=	$(X_1 X_2)$
NP	::=	*Mary*	X	::=	$(\lambda P.(Pm))$
NP	::=	*John*	X	::=	$(\lambda P.(Pj))$
NP	::=	**DET CN**	X	::=	$(X_1 X_2)$
NP	::=	**DET RCN**	X	::=	$(X_1 X_2)$
DET	::=	*every*	X	::=	$(\lambda P.(\lambda Q.\forall x((Px) \to (Qx))))$
DET	::=	*some*	X	::=	$(\lambda P.(\lambda Q.\exists x((Px) \land (Qx))))$
DET	::=	*no*	X	::=	$(\lambda P.(\lambda Q.\forall x((Px) \to \neg(Qx))))$
DET	::=	*the*	X	::=	$(\lambda P.(\lambda Q.\exists x(\forall y((Py) \leftrightarrow x = y) \land (Qx))))$
CN	::=	*man*	X	::=	$(\lambda x.(Mx))$
CN	::=	*woman*	X	::=	$(\lambda x.(Wx))$
CN	::=	*boy*	X	::=	$(\lambda x.(Bx))$
RCN	::=	**CN** *that* **VP**	X	::=	$(\lambda x.((X_1\ x) \land (X_3\ x)))$
RCN	::=	**CN** *who* **NP TV**	X	::=	$(\lambda x.((X_1\ x) \land (X_3(\lambda y.((X_4\ y)x)))))$
VP	::=	*walked*	X	::=	$(\lambda x.(Lx))$
VP	::=	**TV NP**	X	::=	$(\lambda x.(X_2\ (\lambda y.((X_1\ y)x))))$
TV	::=	*loved*	X	::=	$(\lambda x.(\lambda y((Lx)y)))$

We shall often use examples with words not incorporated in the lexicon of this fragment. We shall use however no new syntactic types or constructions.

To this grammar we add rules for the construction of texts, and a general pronoun rule. We shall introduce a set $\mathcal{N}$ of *ν-terms* of the form ν_i. These will function as *placeholders* for ϵ- and τ-terms. Sometimes we shall denote ν-terms as $\nu v : \varphi$. In that case they stand for terms of the form $\epsilon v : \varphi$ or $\tau v : \varphi$. If we write ν-terms as $\nu x : \varphi$, that is with a concrete variable 'x', they will stand for the terms $\epsilon x : \varphi$ or $\tau x : \varphi$. For φ an $\mathcal{L}$ formula, the set $\mathcal{N}(\varphi)$ will consist of all ν-terms occurring in φ.

S	::=	*if* **S S**	X	::=	$X_1 \to X_3$
T	::=	**S**	X	::=	X
T	::=	**T . S**	X	::=	$(X_1 \land X_2)$
NP	::=	*he$_i$*	X	::=	$(\lambda P.(P\nu_i))$
DET	::=	*his$_i$*	X	::=	$(\lambda P.(\lambda Q.\exists x((Px) \land poss(x, \nu_i) \land (Qx))))$

For good measure, we also add a rule for reflexive pronouns (but we shall not bother to spell out the feature constraints for gender agreement).

VP	::=	**TV** *himself*	X	::=	$(\lambda x.((X_1\ x)x))$
VP	::=	**TV** *herself*	X	::=	$(\lambda x.((X_1\ x)x))$

In this example grammar (non-reflexive) pronoun occurrences are translated as

ν-terms in the meaning representation of sentences and texts before anaphoric reference resolution. We shall assume that every surface pronoun gets assigned a ν-term with a different index. During the process of pronoun resolution, these schematic terms have to be *instantiated* to concrete ϵ- or τ-terms. We shall call a formula of our logical language $\mathcal{L}$ *schematic* if it contains ν-terms. If an $\mathcal{L}$-formula is not schematic, we call it *pure*. In this section we shall consider principles to *instantiate* schematic terms. That is, we shall consider principles that reduce schematic formulas to pure formulas. We shall formulate these principles in terms of a general *instantiation relation*. Taken together, the pronoun principles we shall propose are meant as a rudimentary theory of pronoun resolution. The principles are meant to illustrate how pronoun resolution relates to and builds on concepts from instantial logic.

6.2. DEFINITION. (Instantiation Relations) For every schematic term t, an *instantiation relation* R_t is a binary relation on $\mathcal{L}$ satisfying the constraint

$$\text{if } \varphi R_t \psi, \text{ then } \mathcal{N}(\psi) = \mathcal{N}(\varphi) - \{t\}.$$

That is, if $\varphi R_t \psi$ holds, then ψ has at most the schematic terms of φ, and if t is a schematic term of φ then t does not occur in ψ. Notice that a formula φ is pure if and only if for all schematic terms t, $\varphi R_t \varphi$. By further specifying the relations R_t we determine the *admissible* instantiations of schematic ν-terms. In this section we shall explore instantiation principles which define a specific reduction relation R_t. Most will be of the form "$\varphi R_t \psi$ holds if $\psi = \varphi[e/t]$", where e is some ϵ- or τ-term. So instantiation along the R_t relation does not change logical form. It consists of a *substitution* operation. In Section 6.6.2 we shall briefly consider instantiations which also change logical structure.

In many of the cases we shall consider, we can instantiate a ν-term already in a partial reduct of a sentence. But we shall have no stipulation that we must instantiate at the earliest possible occasion. In fact, our strategy will be that we try to instantiate the terms only in the finished product.

A general requirement on the use of pronouns must be formulated without further ado.

Pronoun Principle I For φ to be the representation of a sentence, it must be that $\varphi R_t \varphi$ for all schematic terms t.

This just fixes what we mean by 'meaning representation', namely: representation in which all pronouns have been linked to an appropriate antecedent. In other words, φ does only count as a meaning representation for a sentence if every ν-term is instantiated.

6.3.1 Intersentential Donkey Pronouns

Instantial logic looks promising as a representation medium for anaphoric linking. We shall start by considering how it fares with the typical intersentential *donkey* pronouns. But as we shall see, in the present set-up, there are no essential differences between these pronouns and intrasentential ones.

6.15. *Some man walked. He talked.*

By applying the rules of our grammar we get the representation:

6.16. $\exists x(Mx \wedge Wx) \wedge T(\nu)$.

Now we have to instantiate the schematic term. The most obvious idea is to use the descriptive content of the antecedent noun phrase (*man* in this case) as material to construct an epsilon terms from. This would give us the following translation:

6.17. $\exists x(Mx \wedge Wx) \wedge T(\epsilon x : Mx)$.

Unfortunately, this is too naive. The problem is that nothing forces the interpretation of $(\epsilon x : Mx)$ to be a walking man. Some sitting male talker could make this true. Notice that (6.17) does not entail (6.18).

6.18. $\exists x(Mx \wedge Wx \wedge Tx)$.

Still, this formula should be true given the truth of (6.15) (on the assumption that the pronoun *he* is linked to *some man*).

The moral seems to be that we need the whole subformula inside the scope of the indefinite to 'load' the epsilon term. Only then we can be sure that we have 'picked up the reference' to the appropriate antecedent. In the example case, this would get us the following:

6.19. $\exists x(Mx \wedge Wx) \wedge T(\epsilon x : Mx \wedge Wx)$.

The pronoun is linked not just to an antecedent noun phrase, but to the whole antecedent phrase, so to speak. The pronoun is translated as an epsilon term, with descriptive content given by the existential formula which translates the 'antecedent sentence'. We note that formula (6.19), but not formula (6.17), entails $\exists x(Mx \wedge Wx \wedge Tx)$. Here is a derivation:

$$\dfrac{\dfrac{T(\epsilon x : (Mx \wedge Wx)) \qquad \dfrac{\exists x(Mx \wedge Wx)}{M(\epsilon x : (Mx \wedge Wx)) \wedge W(\epsilon x : (Mx \wedge Wx))}\,\exists\mathrm{E}_\epsilon}{T(\epsilon x : (Mx \wedge Wx)) \wedge M(\epsilon x : (Mx \wedge Wx)) \wedge M(\epsilon x : Mx \wedge Wx))}\,\wedge\mathrm{I}}{\exists x(Mx \wedge Tx \wedge Wx)}\,\exists\mathrm{I}$$

So, in the epsilon calculus the typical *extension of scope* of dynamic frameworks is present for the existential quantifier, but only *if we link a pronoun to the whole antecedent phrase.* The notion of an *antecedent* phrase is not to be interpreted here in terms of the sequential order of the formula: in both $\exists x\varphi \wedge \psi[\epsilon x : \varphi/x]$ and $\psi[\epsilon x : \varphi/x] \wedge \exists x\psi$ the term $\epsilon x : \varphi$ is supported by the existential formula. That is, all choice functions in a model for either sentence will map the term $\epsilon x : \varphi$ to an element satisfying φ.

We shall formulate a notion of *Extended Scope* of one formula over another to capture the notion of an antecedent. This notion of scope is motivated by the relation between a formula $\exists x\varphi$, an *introducer* of the term $\epsilon x : \varphi$, and of formula in which the schematic term ν may be instantiated as $\epsilon x : \varphi$. The truth of a *conjunction* $\exists x\varphi \wedge \psi[\nu/x]$ guarantees that the term $\epsilon x : \varphi$ has indeed an interpretation restricted by the set of φ-ers. The truth of an implication $\exists x\varphi \rightarrow \psi[\nu/x]$ guarantees that, if the antecedent is true, then an instantiation of ν in the consequent as $\epsilon x : \varphi$ is indeed a φ-er. Moreover, in $\exists x Rx(\nu)$ we can never instantiate ν to a term introduced by the existential quantifier in which scope it lies: no $\exists x\varphi$ can contain the term $\epsilon x : \varphi$.

6.3. DEFINITION. (E-Scope) The notion *E-scope* is governed by the following principles.

1. In $\varphi \wedge \psi$ the formula occurrence φ (ψ) has E-scope over occurrence $\varphi \wedge \psi$.
2. If occurrence φ has E-scope over $\psi\#\chi$, for $\# \in \{\wedge, \vee, \rightarrow, \leftrightarrow\}$, then it has E-scope over ψ, unless it is identical to occurrence φ, and E-scope over χ, unless it is identical to occurrence φ.
3. In $\varphi \rightarrow \psi$, the formula occurrence φ has E-scope over occurrence ψ.
4. If formula occurrence φ has E-scope over occurrence ψ and occurrence ψ has E-scope over occurrence χ, then occurrence φ has E-scope over occurrence χ.

So in $\psi[\epsilon x : \varphi/x] \wedge (\exists x\varphi \wedge \chi)$ and in $(\chi \wedge \exists x\varphi) \rightarrow \psi[\epsilon x : \varphi/x]$ the term $\epsilon x : \varphi$ occurs in the E-scope of the formula $\exists x\varphi$, but in $(\exists x\varphi \rightarrow \chi) \wedge \psi[\epsilon x : \varphi/x]$ it does not. The notion of E-scope for implications can be derived if we use the fact that $(\varphi \rightarrow \psi) \leftrightarrow (\varphi \rightarrow (\varphi \wedge \psi))$.

Now we can temptatively formulate a pronoun principle:

> For any pronoun translation $\lambda P : P(\nu_i)$, the term ν_i can only be instantiated to the term $\epsilon x : \varphi$ if ν_i occurs in the E-scope of $\exists x\varphi$.

In terms of the instantiation relation, this principle can be formulated as:

> If all occurrences of schematic term t in ψ lie in the E-scope of $\exists x\chi$, then $\psi R_t \psi[\epsilon x : \chi/t]$.

We shall modify this principle later on to get our Pronoun Principle II.

By our tentative principle, (6.19) is a correct instantiation of (6.15). It explains why the following is not correct.

6.20. (*) *No man walked. He talked.*

This has the translation:

6.21. $\neg\exists x(Mx \wedge Wx) \wedge T(\nu_i)$.

The ν-term does not find an introducer for $\epsilon x : (Mx \wedge Wx)$ in the right E-scope relation. Semantically, this makes intuitive sense: after all, it is claimed that there is no walking man. But a semantic motivation hides the structural part of our principle. For instance, also in (6.22) the existence of a man is claimed.

6.22. *A man who had a brother loved* him.

But still we cannot take 'him' to refer to that man. The translation shows why:

6.23. $\exists x(Mx \land \exists yBRxy \land Lx(\nu))$.

Here ν may be instantiated as $\epsilon y : BRxy$. We have an introducer for this term in E-scope. But there is no introducer for the man. To express that the man mentioned in the subject is the object of 'love' we must use (6.24).

6.24. *A man who had a brother loved* himself.

Now we shall consider some more complex examples.

6.25. *A man loved a woman. He kissed her.*

Its translation:

6.26. $\exists x(Mx \land \exists y(Wy \land Lxy)) \land K(\nu_1)(\nu_2)$.

We have an introducer for $a = \epsilon x : (Mx \land \exists y(Wy \land Lxy))$. This can be the instantiation of ν_1. But the E-scope of the subformula $\exists y(Wy \land Lxy)$ does not extend over the the last conjunct. This is as it should be, because the term $\epsilon y : (Wy \land Lxy)$ has a free variable which would not be bound in this conjunct. To get an instantiation for ν_2, we note that $\exists x(Mx \land \exists y(Wy \land Lxy))$ derives the logically equivalent $Ma \land \exists y(Wy \land Lay)$ by an application of the ϵ-rule. Now, taking the latter formula as the antecedent, we have a new introducer, for the last conjunct does lie in the E-scope of $\exists x(Wy \land Lay)$. So ν_2 can be instantiated as $\epsilon y : (Wy \land Lay)$. Notice that in this way *dependencies* arise between the instantiated terms. By allowing derivations on the antecedent, we can get the translation

6.27. $\exists x(Mx \land \exists y(Wy \land Lxy)) \land K(\epsilon x : (Mx \land \exists y(Wy \land Lxy)), \epsilon y : (Wy \land Lay))$

where $a = \epsilon x : (Mx \land \exists y(Dy \land Own(x, y))$. What we need is the principle: if a subformula of φ lies in the E-scope of $\exists x\psi$, then it lies in the E-scope of $\varphi[\epsilon x : \varphi/x]$. We shall take the following general formulation to be our second pronoun principle.

Pronoun Principle II $\varphi R_t \varphi[\epsilon x : \psi/t]$ holds if all occurrences of the schematic term t occur in the E-scope of the subformulas $\psi_1, \ldots, \psi_n$ of φ and

$$\{\psi_1, \ldots, \psi_n\} \vdash \exists x\psi.$$

We shall call the formula $\exists x\varphi$ the *introducer* of the term $\epsilon x : \varphi$, and we shall call the *support set* of a subformula ψ of φ, the smallest set Σ of subformulas of φ such that ψ lies in the E-scope of all elements of Σ, and Σ derives the introducer of all terms in ψ.

Principle II allows us to instantiate (6.26) as (6.27). But there are more instantiations possible. Notice that in formula (6.26) the conjunct $\exists x(Mx \wedge \exists y(Wy \wedge Lxy))$ classically derives $\exists y(Wy \wedge \exists x(Mx \wedge Lxy))$. So if we allow arbitrary derivations to supply antecedents, we can also get the following formula as an instantiation of (6.26).

6.28. $\exists x(Mx \wedge \exists y(Wy \wedge Lxy)) \wedge K(\epsilon x : (Mx \wedge \exists y(Wy \wedge Lxy)), \epsilon y : (Wy \wedge \exists x(Mx \wedge Lxy)))$.

An important restriction is still present in Pronoun Principle II: we may only use proof theory to get *antecedents* of a pronoun. That is, we only consider derivations of an existential formula $\exists x\varphi$ with assumption in a set defined by *E-scope*. For instance,

6.29. (*)*A man loved* him.

is ungrammatical with 'him' referring to the man. If we consider the translation

6.30. $\exists x(Mx \wedge L(x, \nu))$,

then this ungrammaticality shows by the fact the ν-term does not lie in the E-scope of any existential quantifier connected to 'man'. We may use the ϵ-rule to get $Ma \wedge La(\nu)$ with $a = \epsilon x : (Mx \wedge Lx(\nu))$ and then derive $\exists x Mx$ from Ma. But then we do not have an application of Principle II. In the application of the ϵ-rule the entire formula was involved. That is $\exists x Mx$ cannot be derived *solely* from formulas which have $Lx(\nu)$ in its E-scope. For the same reasons Principle II does not allow us to construct 'a man' as an antecedent in (6.22). As a third example, consider a formula of the form $\forall x(\varphi \rightarrow \psi[\nu/y])$ in which the variable x does not occur free in $\psi[\nu/y]$. This formula allows us to derive $\exists x\varphi \rightarrow \psi[\nu/y]$. But still, in the original formula, $\exists x\varphi$ cannot be derived from subformulas which have $\psi[\nu/y]$ in their E-scope. In the section on donkey pronouns in universal contexts, this situation will occur repeatedly.

However, when we have a representation of a two sentence text, then all derivations using the representation of the first sentence may be used to get antecedents for pronouns occurring in the second sentence. As a final example, consider:

6.31. (*) *A man loved no woman. He kissed her.*

We can find a candidate antecedent for *he*, namely $\exists x(Mx \wedge \neg\exists y(Wy \wedge Lxy))$, but we cannot find an appropriate candidate antecedent for *her*, as the preceding discourse does not imply the existence of any women at all.

6.3.2 Donkey Pronouns in Universal and Conditional Contexts

In this section we shall address the issue of universal readings of pronouns. Some of the most well-known examples are of this kind.

6.32. *Every man who loves a woman kisses her.*

We translate this by our grammar into:

6.33. $\forall x((Mx \wedge \exists y(Wy \wedge Lxy)) \rightarrow Kx(\epsilon y : Wy \wedge Lxy))$.

Here the ϵ-term in the consequent lies in the E-scope of its introducer. So we have an admissible instantiation. Moreover, Principle II does not allow us to get construct an antecedent from the phrase 'every man': $\exists xMx$ cannot be derived solely form formulas which have $Kx(\epsilon y : (Wy \wedge Lxy))$ in their E-scope. So Principle II gives the right results concerning the *distribution* of pronouns. But does this give the right semantics for (6.32)? The formula expresses that every woman-loving man kisses at least one of the women he loves. But in general (6.25) is taken to mean that every woman-loving man kisses all the women he loves. That is, the ϵ-term in the consequent should have universal force. And Principle II does not give us this reading. However, our set-up till now works all right for examples where this universality is not claimed.

6.34. *Every man who has a dime puts* it *in the meter.*

Here we have the translation

6.35. $\forall x((Mx \wedge \exists y(Dy \wedge Own(x,y)) \rightarrow PiM(x, \epsilon y : (Dy \wedge Own(x,y)))$.

The pronoun 'it' has been instantiated as $\epsilon y : (Dy \wedge Own(x,y))$ with free variable x. It lies in the E-scope of the introducer $\exists y(Dy \wedge Own(x,y))$.

The lack of universal readings crops up if we consider the typical Donkey equivalences for universal sentences. Sentence (6.34) has the same meaning as the following conditional sentence.

6.36. *If a man has a dime* he *puts* it *in the meter.*

Consider the following translation

6.37. $\exists x(Mx \wedge \exists y(Dy \wedge Own(x,y))) \rightarrow PiM(\nu_1, \nu_2)$.

We note that again Principle II gets the distribution right. In contrast to (6.35), for this sentence we can construct an antecedent for both 'he' and for 'it'. But (6.37) does not give the right meaning to (6.36). It does not imply that *every* dime-owning man puts some dime in the meter. The same problem occurs in the translation of

6.38. *If a man loves a woman he kisses her.*

Here we get the translation

6.39. $\exists x \exists y(Mx \wedge Wy \wedge Lxy) \rightarrow K(\epsilon x : \exists y(Mx \wedge Wy \wedge Lxy))(\epsilon y : (Ma \wedge Wy \wedge Lay)))$,

where $a = \epsilon x : \exists y(Mx \wedge Wy \wedge Lxy)$. In this case both ϵ-terms must be interpreted with universal force to get the right interpretation.

The lack of universal readings is not a consequence of the semantics of the ϵ-calculus. This calculus allows us to formulate a wide variety of quantificational patterns by instantiating the schematic terms to the right ϵ-terms. It solely derives from the lack of instantiating principles in the set-up up to now. To account for universal interpretations of pronouns in conditional contexts, we shall use τ-terms in a judicious way.

To illustrate the solution we shall propose, consider the formula $\exists xA \to B(\nu)$, the general form of the conditional sentences of interest. Suppose we instantiate ν by the term $\tau x : (A \to B)$. This gives $\exists xA \to B(\tau x : (A \to B))$. Because the variable x does not occur freely in the consequent of the implication, this is equivalent to $\forall x(A \to B(\tau x : (A \to B)))$. From this we can derive, by ($\forall$E), the formula $A(\tau x : (A \to B)) \to B(\tau x : (A \to B))$ and so $\forall x(A \to B)$, by the characteristic τ-principle. Consequently, under this instantiation we get universal readings of conditionals.

Notice that

- $(\exists xPx \wedge Q\epsilon x : Px)$ implies $\exists x(Px \wedge Qx)$,
- $(\exists xPx \to Q\tau x : (Px \to Qx))$ implies $\forall x(Px \to Qx)$.

In both cases the reverse implication does not hold. To see this for the second implication, consider a model M where $\forall x(Px \to Qx)$ holds, such that there are objects that are neither P nor Q. Let the choice function Φ map $\tau x : (Px \to Qx)$ to such an element. Then a counterexample to $\forall x(Px \to Q\tau(Px \to Qx))$ is supplied by any element which is P.

We formulate our pronoun principle for universal and conditional sentences. Informally, this principle states "If a schematic term ν in the consequent $\psi[\nu/x]$ of an implication lies in the E-scope of $\exists x\varphi$ in the antecedent, then ν can be instantiated as $\tau x : (\varphi \to \psi)$."

Pronoun Principle III The instantiation relation $\chi R_t \chi[\tau x : (\xi \to \psi[x/t])/t]$ holds if $\chi = \varphi \to \psi$, φ lies in the E-scope of $\exists x\xi$ and all occurrences of term t lie in ψ.

Now we use Principle III to get translations.

6.40. *Every farmer who has a donkey beats* it.

By lambda conversions this gives

6.41. $\forall x((Fx \wedge \exists y(Dy \wedge O(x,y))) \to B(x,\nu))$

with a schematic term. By Principle III we may instantiate 'it' by $\tau : y(Dy \wedge O(x,y) \to B(x,y))$, giving

6.42. $\forall x((Fx \wedge \exists y(Dy \wedge O(x,y))) \to B(x, \tau : y(Dy \wedge O(x,y) \to B(x,y))))$.

Because the variable y does not occur freely in the consequent, this is equivalent to the first-order sentence

6.43. $\forall x \forall y((Fx \wedge (Dy \wedge Oxy)) \to Bxy)$.

We end this section with a remark about freedom to instantiate universally.

6.4. REMARK. The logical background to Principle III is constituted by the soundness of the following rule

$$\frac{\begin{array}{c}\Sigma, \exists x\chi \ [i] \\ \vdots \\ \psi[\epsilon x : \chi/x]\end{array}}{\forall x(\chi \to \psi)} \to\mathrm{I}_i$$

$$\text{provided } \epsilon x : \chi \notin \mathcal{E}(\Sigma)$$

That is, $\vdash \exists x\chi \to \psi[\epsilon x : \chi/x]$ implies $\vdash \forall x(\chi \to \psi)$. For if $\forall x(\chi \to \psi)$ is falsified on some model M, s, G_Φ, then we can find there an element in $|\varphi(x) \wedge \neg\psi(x)|^{s,G_\Phi}$. Such an element can be taken as the V_{M,s,G_Φ} value of $\epsilon x : \chi$ (for $\epsilon x : \chi$ is not restricted by Σ). Consequently we have a model for $\Sigma, \exists x\chi$ falsifying $\varphi[\epsilon x : \chi/x]$.

If we are interested in the analysis of natural language, Hilbert's definition of τ-terms need not be sacrosanct. If we translate 'every man walked' by $\forall x(Mx \to Wx)$ then, classically, we can conclude $\exists x(Mx \to Wx)$. But we cannot conclude $\exists x(Mx \wedge Wx)$. That is, in the semantics of 'every man walks' there is no guarantee that there is a walking man at all. This is unsatisfactory from a linguistic point of view. By changing our definition of τ-terms slightly, we can remedy this.

6.5. DEFINITION. (Variant τ-term Interpretation) If χ is implicational formula of the form $\varphi \to \psi$, then the interpretation of $\tau x : \chi$ is given by

$$V_{M,s,G_\Phi}(\tau x : \chi) = \begin{cases} \in |\varphi(x) \wedge \neg\psi(x)|^{s,G_\Phi} & \text{if } |\varphi(x) \wedge \neg\psi(x)|^{s,G_\Phi} \neq \emptyset \\ \in |\varphi(x)|^{s,G_\Phi} & \text{if } |\varphi(x)|^{s,G_\Phi} \neq \emptyset \\ \in dom(M) & \text{otherwise} \end{cases}$$

If χ is not of implicational form, then the τ-term corresponding to χ has the form $\tau x : (\top \to \varphi)$.

For non-implicational formulas this interpretation reduces to the standard one. For implicational formulas, $\tau x : (\varphi \to \psi)$ is assigned an element in φ if such an element exists. Now

$$\vdash \forall x(\varphi \to \psi) \to \forall x(\varphi \to \psi[\tau x : (\varphi \to \psi)/x]),$$

and so the following rule is sound:

$$\frac{\begin{array}{c}\Sigma, \varphi \ [i] \\ \vdots \\ \psi[\epsilon x : \chi/x]\end{array}}{\varphi \to \psi^*} \to\mathrm{I}_i \qquad \text{where } \psi^* = \begin{cases} \psi[\tau x : (\chi \to \psi)/x] & \text{If } \varphi = \exists x\chi \text{ and} \\ & \epsilon x : \chi \notin \mathcal{E}(\Sigma) \\ \psi & \text{otherwise} \end{cases}$$

With this remark about freedom to instantiate universally we end the section.

6.4 Truth-Conditional Semantics and Incrementality

The difference between

6.44. *Some man walked and talked.*

and

6.45. *Some man walked. He talked.*

comes down to the difference between the formulas

6.46. $\exists x(Mx \wedge Wx \wedge Tx)$

and

6.47. $\exists x(Mx \wedge Wx) \wedge T(\epsilon x : Mx \wedge Wx)$

respectively. In the second translation an ϵ-term is *introduced* and this term is used as a term in a predicate. These representations have exactly the same *first-order consequences*: every model for the first formula can be supplied with a choice function to verify the second formula, and every model-choice function pair for the second formula is a model for the first. So if we consider the truth-functional interpretation of a sentence or text to be given by a first-order, ϵ-free formula, then the difference between the two formulas above is not a truth functional one. The intuitive interpretation of this difference can best be seen, if we continue text (6.45) as follows:

6.48. *Some man walked. He talked. Then he smiled.*

This can be translated as

6.49. $\exists x(Mx \wedge Wx) \wedge T(\epsilon x : Mx \wedge Wx) \wedge S(\epsilon x : Mx \wedge Wx)$

and as

6.50. $\exists x(Mx \wedge Wx) \wedge T(\epsilon x : Mx \wedge Wx) \wedge S(\epsilon x : Mx \wedge Wx \wedge Tx)$.

Both formulas derive $\exists x(Mx \wedge Wx \wedge Tx \wedge Sx)$ and again they cannot be distinguished by their first-order consequences. But we claim that only the second formula corresponds to (6.48). The first is in fact a translation of

6.51. *Some man walked. He talked and smiled.*

The second, but not the first, translation exhibits the sequential order present in the text. This order is not evident from the sequential order of the logical *string.* It can be found in the relation between the *support sets* of the subformulas in the string. The term $\epsilon x : (Mx \wedge Wx)$ only requires the subformula $\exists x(Mx \wedge Wx)$. But the introducer of $\epsilon x : (Mx \wedge Wx \wedge Tx)$ needs the accumulated text translation $\exists x(Mx \wedge Wx) \wedge T(\epsilon x : (Mx \wedge Tx))$ for its derivation. In the interpretation of the second occurrence of 'he' all accumulated information is incorporated. We can give this accumulation principle an official status by formulating the following simple pronoun principle.

Pronoun Principle IV Every occurrence of a pronoun in the text must have its own introducer in the semantic representation.

In other words: two pronoun occurrences cannot have the same introducer. This is made more specific in the following rule.

Pronoun Principle V If $\xi[\nu/x]$ occurs in the E-scope of $\psi[\epsilon x : \chi]$, then it may be instantiated as $\xi[\epsilon x : (\chi \wedge \psi)/x]$. If $\xi[\nu/x]$ occurs in the consequent of an implication, in the E-scope of $\psi[\epsilon x : \chi]$ occurring in the antecedent, then it may be instantiated as $\xi[\tau x : (\chi \wedge \psi \rightarrow \xi)/x]$.

So now, to be instantiated, any pronoun requires the present of a 'previous' ν-term in the translation. Again, the soundness of this instantiation schema can be seen from the following facts. By the grammar, all ν-terms, apart from pronouns, that occur in translations lie within the scope of an introducer. This does not change with respect to the previous grammar. This implies that for Principle V, we can always *derive* an introducer from a formula in the E-scope of which the instantiation of the pronoun lies. For instance, $\exists x\varphi \wedge \psi[\epsilon x : \varphi/x]$ derives $\exists x(\varphi \wedge \psi)$, an introducer for $\epsilon x : (\varphi \wedge \psi)$.

By our definition of E-scope this resembles the standard notion of accessibility which roughly states that a pronoun and its antecedent cannot be arguments of the same predicate. For instance, in

6.52. *A man loved* him.

the pronoun 'him' cannot be anaphorically linked to 'a man'. In the translation $\exists x Mx \wedge L(\epsilon x : Mx)(\nu)$ this follows because we need a formula with an ϵ-term with $L(\epsilon x : Mx)(\nu)$ in its E-scope. But there is no such formula. According to Principle IV we may not use $\exists x Mx$ again.

As long as we use the same ϵ-term, we are dealing with a complex predicate, in DRT terms, with a complex *condition*. This will be the translation of one occurrence of a pronoun. *Updating* the ϵ-term $\epsilon x : (Mx \wedge Wx)$ to $\epsilon x : (Mx \wedge Wx \wedge Tx)$ corresponds to the use of a new pronoun.

Again formulas (6.49) and (6.50) have the same first-order consequences as $\exists x(Mx \wedge Wx \wedge Tx \wedge Sx)$. In this sense the meaning supplied by (6.49) and (6.50) coincides with the one DRT or DPL assigns to this sentence. However if we consider the extended interpretation including ϵ-terms, then, intuitively, there is a problem with (6.50). For a model for this formula need not assign $\epsilon x : (Mx \wedge Wx)$ and $\epsilon x : (Mx \wedge Wx \wedge Tx)$ the same value. The choice made for both terms need not be coordinated. And it is this coordination of choices, we claim, that gives pronouns their flavor of being like some kind of definite descriptions.

6.53. *Some man walked. He talked. Then* the *man that walked and talked smiled.*

We cannot use our semantical representation of 'the' to translate the pronoun, for this would lead to models where only *one* man walked and talked (see Van Eijck [Eij85]). We shall interpret the intuition of 'definiteness' surrounding the use of pronouns by taking it to mean that the antecedent of the pronoun has led to a *choice* of some walking and talking man, and taht the pronoun in 'Then *he* smiled' refers to *that choice*.

6.6. REMARK. The flavor of definiteness associated with *singular* pronouns corresponds to the flavor of 'exhaustiveness' surrounding *plural* pronouns. In the text

6.54. *Some congressmen admired Kennedy.* They *were of Irish descent.*

the interpretation of the pronoun 'they' involves the fact that *all* congressmen that admired Kennedy where of Irish descent. Again it seems undesirable to translate the pronoun by the semantic representation of 'every', for we may continue this text with "Some other congressmen admired Kennedy. And *they* were Catholics". In the case of plural pronouns we shall take exhaustiveness to mean that *all choices* made for Kennedy-admiring congressmen where in fact of Irish descent.

For this interpretation of definiteness and exhaustiveness to work we need the assurance that, in (6.50) for instance, the ϵ-terms get assigned the same value.

It is clear that every model for the $\exists x(Mx \wedge Wx \wedge Tx \wedge Sx)$ has a choice function assigning the terms in (6.50) the same value. We shall now show that we can define the class of *coordinated choice* models by a proof rule. In the formulation of this rule we use an aspect of the ϵ-calculus we have quite neglected up to now. Because ϵ-terms have formulas as subparts, we can formulate inference rules referring to these formulas. The following rules are prime examples.

6.7. DEFINITION. (Coordinated Choice Rules)

$$\frac{\exists x\varphi \wedge \psi[\epsilon x : \varphi/x] \quad \chi[\epsilon x : \varphi/x]}{\chi[\epsilon x : (\varphi \wedge \psi)/x]} \qquad \frac{\exists x\varphi \wedge \psi[\epsilon x : \varphi/x] \quad \chi[\epsilon x : (\varphi \wedge \psi)/x]}{\chi[\epsilon x : \varphi/x]}$$

These proof rules carefully express our requirements[2]. If we have identity in the language, then these principles entail *extensionality* (Section 6.10). If we do not have identity, then they express the indistinguishability of $\epsilon x : \varphi$ and $\epsilon x : (\varphi \wedge \psi)$ on any model where $\exists x\varphi \wedge \psi[\epsilon x : \varphi/x]$ holds.

[2]For our present requirements we could have used

$$\frac{\exists x\varphi \wedge \psi[\epsilon x : \varphi/x]}{\chi[\epsilon x : \varphi/x] \leftrightarrow \chi[\epsilon x : (\varphi \wedge \psi)/x]}$$

But the formulation we have chosen is adapted to the *truth-to-truth* interpretation we shall use in Section 6.7 for the treatment of plurals.

6.5 Subject Predicate Form

The above argument about incrementality suggests a basic change in our grammar. Observe that the term $\epsilon x : Mx$, corresponding to 'a man', the *subject* of the sentence "Some man walked", does not appear in the translation

$$\exists x(Mx \wedge Wx) \wedge T(\epsilon x : Mx \wedge Wx)$$

of 'Some man walked. *He* talked.' Here the translation of the occurrence of 'he' as *that man* suggests that we have already made a choice to interpret 'some man'. That is, it suggest that 'Some man walked' has the form

6.55. $\exists x Mx \wedge W(\epsilon x : Mx)$.

Here an ϵ-term occurs in the representation that is not the instantiation of a pronoun. It is introduced by the subject of the sentence. The subject of a sentence is generally acknowledged to have a special status, compared, for instance, to the *object* of a sentence. This difference can be made explicit by the epsilon calculus. Consider again

6.56. *A man loved a woman.*

with the following translation:

6.57. $\exists x Mx \wedge (\exists y L(\epsilon x : Mx)y \wedge W(\epsilon y : L(\epsilon x : Mx)y))$.

Here the subject of the sentence introduces an ϵ-term into the verb phrase. The object of the sentence, 'a woman', does not enter in the form of an ϵ-term, but as a condition on the term introduced by the (representation of the) transitive verb. Notice that (6.57) expresses the ϵ-free truth-conditions of $\exists x(Mx \wedge \exists y(Wy \wedge Lxy))$.

We shall now give a modified version of our grammar, where epsilon and tau terms get introduced already in the determiner translations. This modification is for illustrative purposes only, namely to show the direction of further elaborations of the system. We shall not spell out the modifications in the rules for transitive verb-phrases that we need to get the subject-object distinction that we want, but we shall just assume that our grammar gives us what we want in the way we want it.

$$\begin{array}{lllll}
\textbf{DET} & ::= & \textit{every} & X & ::= & (\lambda P.(\lambda Q.\forall x((Px) \rightarrow (Q\tau x : ((Px) \rightarrow (Qx)))))) \\
\textbf{DET} & ::= & \textit{some} & X & ::= & (\lambda P.(\lambda Q.\exists x(Px) \wedge (Q\epsilon x : (Px)))) \\
\textbf{DET} & ::= & \textit{no} & X & ::= & (\lambda P.(\lambda Q.\forall x((Px) \rightarrow \neg(Q\tau x : ((Px) \rightarrow \neg(Qx)))))) \\
\textbf{DET} & ::= & \textit{the} & X & ::= & (\lambda P.(\lambda Q.\exists x \forall y((Py) \leftrightarrow x = y) \wedge \\
 & & & & & \qquad (Q\epsilon x : (\forall y((Py) \leftrightarrow x = y)))))
\end{array}$$

Now, ϵ- and τ-terms are already introduced in the representation by the rules of our grammar. Notice that $\exists x Px \wedge Q\epsilon x : Px$ derives $\exists x(Px \wedge Qx)$, our previous representation (but not vice versa), and that $\forall x(Px \to Q\tau(Px \to Qx))$ derives $\forall x(Px \to Qx)$ (again not vice versa).

With this modified grammar "Some man walked. He talked" has the translation (after reference resolution)

6.58. $\exists xMx \wedge W(\epsilon x : Mx) \wedge T(\epsilon x : (Mx \wedge Wx))$

and we have the definite reading: 'Some man walked. *That walking man* talked.' where *that* refers to the choice made to interpret $\epsilon x : Mx$. Thus, as was mentioned in the introduction, instantial logic gives us a theory of definiteness as committed or coordinated choice.

All Pronoun Principles can remain as they are. Consider again the example

6.59. *Every man who has a dime puts* it *in the meter.*

Now we have the translation

$$\forall x((Mx \wedge \exists yDy \wedge Own(x, \epsilon y : Dy)) \to PiM(x, \nu)).$$

By Principle IV the pronoun 'it' can been instantiated as $\epsilon y : (Dy \wedge Own(x, y))$ with free variable x. It lies in the E-scope of the introducer $Own(x, \epsilon y : Dy)$, and $\exists x(Mx \wedge \exists y(Dy \wedge Own(x, y))$ is *derivable* from the antecedent of the implication. Notice that we get the same ϵ-term as under the previous translation (formula (6.35)).

6.6 Harder Cases

In this section we shall discuss a slightly more problematic case (and admit that the grammar still needs further modification). Here we concentrate on the terms we *want* to generate to instantiate pronouns. Consider the sentence

6.60. *A man loved a woman who hates* him.

Up to now, we are not able to use the subject as an introducer for an ϵ-term. But here is a possible solution. When we make the subject-object distinction suggested in Section 6.5 we can get the following translation.

6.61. $\exists xMx \wedge (\exists yL(\epsilon x : Mx)y \wedge (W(\epsilon y : L(\epsilon x : Mx)y) \wedge H(\epsilon y : L(\epsilon x : Mx)y)\nu)))$.

Now, according to pronoun Principle IV (and according to Principle II) we have the formula $\exists x(Mx \wedge \exists yLxy)$ as a possible introducer for $\epsilon x : (Mx \wedge \exists yLxy)$ (for it is derivable from $\{\exists xMx, \exists yL(\epsilon x : Mx)y\}$). By the coordinated choice rule we may replace $\epsilon x : (Mx \wedge \exists yLxy)$ by $\epsilon x : Mx$ giving

6.62. $\exists xMx \wedge \exists yL(\epsilon x : Mx)y \wedge (W(\epsilon y : L(\epsilon x : Mx)y) \wedge H(\epsilon y : L(\epsilon x : Mx)y)\epsilon x : Mx)$.

And this formula has the same first-order consequences as

6.63. $\exists x(Mx \wedge \exists y(Lxy \wedge (Wy \wedge Hyx)))$.

Still, note that the modified rules do not handle the following example.

6.64. *A man who loved a woman who hated* him *smiled sadly.*

The problem here is that the pronoun should get linked to a noun phrase of which it itself constitutes a part. We have to admit defeat here, at least provisionally, but we comfort ourselves with the thought that problematic cases like this did already lead to discussion and controversy in the early days of Montague grammar (see Janssen [Jan86]).

6.6.1 Bach-Peters Sentences

Famous 'tough cases' are constituted by so-called *Bach-Peters sentences* [Gea80] like the following.

6.65. *A woman who really loved him forgave a man who didn't care about her.*

We get a translation

6.66. $\exists x(Wx \wedge Lx(\nu_1)) \wedge (\exists y(My \wedge \neg Cy(\nu_2)) \wedge F(a,b))$

where

$$a = \epsilon x : (Wx \wedge Lx(\nu_1)),$$
$$b = \epsilon y : (My \wedge \neg Cy(\nu_2)).$$

The object is now to use the introducers $\exists x(Wx \wedge Lx(\nu_1))$ and $\exists y(My \wedge \neg Cy(\nu_2))$ to get instantiations for the schematic terms ν_2 and ν_1 respectively. The problem here is that in using these introducers we do not get rid of schematic terms. For substitution in a and b of the terms thus introduced gives

$$a \mapsto \epsilon x : (Wx \wedge Lx(\epsilon y : (My \wedge \neg Cy(\nu_2)))),$$
$$b \mapsto \epsilon y : (My \wedge \neg Cy(\epsilon x : (Wx \wedge Lx(\nu_1)))).$$

Moreover, the representation of the pronoun 'him', i.e., ν_1, is to be coreferential with the term $b(\nu_1)$ after substitution and, mutatis mutandis, the same holds for the denotation of 'her', i.e., ν_2, and $a(\nu_2)$. But then the schematic terms ν_1 and ν_2 reappear in the very terms we use as instantiations. To get a correct representation what we need are the instantiations:

$$\nu_2 \mapsto \epsilon x : (Wx \wedge Lx(\epsilon y : (My \wedge \neg Cyx))),$$
$$\nu_1 \mapsto \epsilon y : (My \wedge \neg Cy(\epsilon x : (Wx \wedge Lxy))).$$

Here ν_2 is instantiated by a term denoting a woman loving a man who does not care about *that same woman*, and ν_1 is instantiated by a term denoting a man not caring about a woman who loves *that same man*.

Notice, that we can get the required terms if we replace the ν-terms in a and b by the variable bound by the outermost ϵ-symbol. This we can achieve by addition of the following *Subordination Principle* (see Section 5.8.1 and ([MV95]).

Subordination Principle If $\psi(\nu)$ occurs in the E-scope of $\exists x\varphi$ and ν may be instantiated in ψ by $\epsilon x : \varphi$, then it may be instantiated by $\epsilon x : (\varphi[x/\nu])$.[3]

The soundness of this principle can be argued for as follows. Suppose $\exists x Rx(\nu)$ holds on a model M, Φ, and ν is instantiated as $\epsilon x : Rx(\nu)$. Thus ν is assigned a value in $\{m \in dom(M) \mid M, \Phi \models Rm(\nu)\}$. But that means that ν is assigned a value in $\{m \in dom(M) \mid M, \Phi \models Rmm\}$. That is, ν may be instantiated by $\epsilon x : Rxx$. Notice that it is essential that $\exists x Rx(\nu)$ holds on the model. For if it would not hold, then the term $\epsilon x : Rx(\nu)$ could be assigned an arbitrary value, say the value of ν. But it may still be that $M, \Phi \models \exists x Rxx$, so the values assigned to $\epsilon x : Rxx$ are *not* arbitrary and, in particular, this term cannot be assigned the value of ν. So if $\exists x Rx(\nu)$ does not hold on M, Φ, then the fact that ν is instantiated by $\epsilon x : Rx(\nu)$ does not mean that it can be instantiated by $\epsilon x : Rxx$.

Armed with this principle we can deal adequately with our Bach-Peters example. But again, we have to admit that further principles have to be assumed for the handling of universal Bach-Peters sentences such as (6.67).

6.67. *Every woman who really loves him will forgive a man even if he does not care about her.*

6.6.2 Modal Subordination

A particularly tough nut to crack has been the phenomenon called *modal subordination.*

6.68. *Every man came forward.* He *accepted his award and thanked the committee.*

Here the pronoun 'he' needs an antecedent that cannot have E-scope beyond the boundaries of the implication. In this case we are after a reduction of the form

6.69. $\forall x(Mx \to CF(\tau x : (Mx \to CFx)) \wedge AW(\nu)$

to something like

6.70. $\forall x(Mx \to (CF \wedge AW)[\tau x : (Mx \to (CFx \wedge AWx)/x])$.

How to get this instantiation? Here, for the first time, the process of instantiation appears to require more than substitution. We can limit the restructuring of logical form as follows.

[3] In [MV95] also a converse principle is discussed which can be paraphrased here as

> If $\psi(\nu)$ occurs in the E-scope of $\exists x(\varphi[x/\nu])$ and ν may be instantiated in ψ by $\epsilon x : (\varphi[x/\nu])$, then it may be instantiated by $\epsilon x : \varphi$.

This gives us the possibility to *create* circular instantiations $\nu_1 \mapsto b(\nu_2)$ and $\nu_2 \mapsto a(\nu_1)$ like the ones we used in the above example, given the right existential formulas. If $\exists x Rx(\epsilon y : Qxy)$ has $\psi(\nu_1)$ in its E-scope, ν_1 is instantiated by $\epsilon x : Rx(\epsilon y : Qxy)$, and ν_2 by $\epsilon y : Qx(\nu_1)$, then by this principle, we may instantiate ν_1 by $\epsilon x : Rx(\epsilon y : Q(\nu_1)y)$. So we in fact instantiate ν_1 by $\epsilon x : Rx(\nu_2)$: a circular instantiation.

6.71. $\forall x(Mx \rightarrow CF(\tau x : (Mx \rightarrow CFx))) \wedge \forall x(Mx \rightarrow AW(\tau x : (Mx \rightarrow (CFx \wedge AWx))))$.

This formula derives 6.70, for the first conjunct derives $\forall x(Mx \rightarrow CFx)$ and consequently $M(\tau x : (Mx \rightarrow (CFx \wedge AWx)) \rightarrow CF(\tau x : (Mx \rightarrow (CFx \wedge AWx))$. The rest follows by standard first-order logic. Notice that we only have to restructure the formula containing the schematic term. We do not have to break up the previous context. We shall formulate our proposal and leave it at that.

Pronoun Principle V If $\varphi[\tau x : (\varphi \rightarrow \psi)/x] \rightarrow \psi[\tau x : (\varphi \rightarrow \psi)/x]$ can be derived in the E-scope of $\chi(\nu)$, then $\chi(\nu)$ may be instantiated as $\forall x(\varphi \rightarrow \chi[\tau x : (\varphi \rightarrow \psi \wedge \chi)/\nu])$.

Notice that by this proposal only the subformula containing the pronoun needs to be restructured. Further discussion of modal subordination is beyond the scope of the present chapter.

6.7 Plural E-type Pronouns

Many famous problem cases of anaphora involve plural pronouns.

6.72. *A man and a boy walked in. They smiled.*

Here the plural pronoun 'they' has as antecedent the summation of the singular NPs 'a man' and 'a boy'.

6.73. *John borrowed some books. Mary read them.*

This sentence should mean that Mary read all the books that John borrowed. That is, the plural pronoun 'them' should get an *exhaustive* reading.

6.74. *Few women admire John, but they are very beautiful.*

This should mean that all women who admire John are very beautiful.

6.75. *Few books mention John, and they are very hard to find.*

Again, this should mean that all books that mention John are hard to find.

To discuss the meanings of examples involving plural pronouns in a formal way we need at least a rudimentary logical treatment of the singular/plural distinction. We shall use the possibility of the ϵ-calculus to interpret ϵ-terms by *sets* of choice functions. Let there be for every predicate Q of the language a second predicate Q^p, a *plural* version of Q. The interpretation function of the models interprets Q^p *exactly* like Q. It is only within the scope of an ϵ-symbol that the superfix p will do some work. The ϵ-terms in the unextended language we shall call *singular* terms. If the formula φ contains any predicate symbol superfixed with p, then the term $\epsilon x : \varphi$ will be called a *plural* term. We shall take τ-terms $\tau x : \varphi$ to be abbreviations of $\epsilon x : \neg\varphi$. For any formula φ in the extended language, we let $S(\varphi)$ be the formula we get from φ by stripping all predicate symbols of the superfix 'p'.

We interpret this language by *sets* of choice functions.

6.8. DEFINITION. (Adequate Sets of Choice Functions) A set of choice functions

$$G_\Phi \subseteq \{\Phi' \mid \Phi' \text{ a choice function over } M\}$$

is called *adequate* for M if G_Φ is a non-empty set, $\Phi \in G_\Phi$, and for all variable assignments s it holds that

1. for all singular ϵ-terms $(\epsilon x : \varphi)$: $V_{M,s,G_\Phi}(\epsilon x : \varphi) = \{V_{M,s,G_\Phi}(\epsilon x : \varphi)\}$,
2. for all *plural* terms $(\epsilon x : \varphi)$: $V_{M,s,G_\Phi}(\epsilon x : \varphi) = \{V_{M,s,G_{\Phi'}}(\epsilon x : S(\varphi)) \mid \Phi' \in G_\Phi\}$.

On the standard singular terms the set of choice functions G_Φ behaves as the choice function Φ. On plural terms this set behaves 'generically'.

Now we set

$$M, s, G_\Phi \models \varphi \iff \forall \Phi' \in G_\Phi : M, s, G_\Phi \models S(\varphi).$$

Notice that $\exists x\varphi \leftrightarrow \varphi[\epsilon x : \varphi/x]$ and $\forall x\varphi \leftrightarrow \varphi[\epsilon x : \neg\varphi/x]$ are universally valid also under the generic interpretation.

We add to our grammar entries for plural common nouns, of the form

$$\mathbf{CN} \ ::= \ \mathit{men} \qquad X \ ::= \ (\lambda x.(M^p x)).$$

Now

6.76. *Some men loved some women*

can be translated as

6.77. $\exists x M^p x \wedge \exists y W^p y \wedge L(\epsilon x : M^p x)(\epsilon y : W^p y)$.

Notice that this translation does not entail that all men in the denotation of $\epsilon x : M^p x$ loved all women in the denotation of $\epsilon y : W^p y$.

Now we shall consider some typical E-type phenomena concerning plurals. A famous example

6.78. *Some congressmen admire Kennedy.* They *are of Irish descent.*

Here it seems that *all* congressmen who admire Kennedy are of Irish descent. That is, in the translation

6.79. $\exists x C^p x \wedge AK(\epsilon x : C^p x) \wedge ID(\epsilon x : (C^p x \wedge AKx))$

we want an *exhaustive* interpretation of $\epsilon x : (Cx \wedge AKx)^p$. Because we may continue this text with "Some other congressmen admire Kennedy and *they* are catholics". We shall take exhaustiveness to mean here that all elements of the set assigned to $\epsilon x : (C^p x \wedge AKx)$ are indeed of Irish descent. This is the case if we consider only models for the *coordinated choice* principle. By this principle, any model for $\exists x P \wedge Q(\epsilon x : Px) \wedge R(\epsilon x : (Px \wedge Qx))$ will assign the same element to $\epsilon x : Px$ and $\epsilon x : (Px \wedge Qx)$. In the case of singular terms this takes care of the continuity of reference of the term $\epsilon x : Px$. In the case of plural terms, this guarantees that all values assigned to $\epsilon x : Px$ will satisfy R.

Now the use of plural pronouns can be an extension of the singular case.

6.80. *Some boys walked.* They *talked.*

This translates simply to

6.81. $\exists x B^p x \wedge W(\epsilon x : B^p x) \wedge T(\epsilon x : (B^p x \wedge Wx))$

with a plural interpretation. Here we can use the same Pronoun Principles as for singular terms.

In cases where we have to construct an antecedent for a plural pronoun out of several singular noun phrases, we need some further machinery. In order to interpret sentences like

6.82. *Some man and some woman walked in.* They *sat down.*

we must be able to gather singleton terms into plural antecedents.

6.9. DEFINITION. (Summation Operator) The *summation operator* '$\sqcup$' is a binary operation on the set ϵ-terms giving new ϵ-terms such that

$$V_{M,s,G_\Phi}(\epsilon x : \varphi \sqcup \epsilon y : \psi) = V_{M,s,G_\Phi}(\epsilon x : \varphi) \cup V_{M,s,G_\Phi}(\epsilon y : \psi)$$

holds.

The term operator '$\sqcup$' is governed by the following proof rules.

6.10. DEFINITION. (Summation Rules)

$$\frac{\varphi[\epsilon x : \psi/x] \wedge \varphi[\epsilon y : \chi/x]}{\varphi[\epsilon x : \psi \sqcup \epsilon y : \chi/x]} \qquad \frac{\varphi[\epsilon x : \psi \sqcup \epsilon y : \chi/x]}{\varphi[\epsilon x : \psi/x] \wedge \varphi[\epsilon y : \chi/x]}$$

In the logic of plural terms all proof rules have to be interpreted *truth-to-truth* (see Section 2.3.4). That is, an inference $\varphi_1, \ldots, \varphi_n/\psi$ is valid on a model M, s, if for all adequate sets G_Φ: if $\varphi_1, \ldots, \varphi_n$ are true on M, s, G_Φ, then ψ is true on M, s, G_Φ. Recall that this interpretation disallows conditionalization: from the validity of φ/ψ we cannot conclude to the validity $\varphi \rightarrow \psi$.

Notice that $M, s, G_\Phi \models \neg\varphi[\epsilon x : \psi \sqcup \epsilon y : \chi/x]$ holds if no element of G_Φ verifies $\varphi[\epsilon x : \psi/x]$ or $\varphi[\epsilon y : \chi/x]$. This is as it should be. The sentence "I did not meet John and Mary" is true if I met neither of them. Now we can summate singular existentials into an antecedent for a plural pronoun.

Assuming that we have a grammar to deal with coordinated NPs like 'some men and some women', a sentence like

6.83. *Some men and some women walked in.*

can now be translated as

6.84. $\exists x M^p x \wedge \exists y W^p y \wedge Wa(\epsilon x : M^p x \sqcup \epsilon y : W^p y)$.

From this we can conclude $\exists x M^p x \wedge W(\epsilon x : M^p x)$ and $\exists y W^p y \wedge W(\epsilon y : W^p y)$. That is, plural NPs like 'some men and some women' get a *distributive* reading. The sentence

6.85. *Some men and all women walked in.*

will have the meaning representation

6.86. $\exists x M^p x \wedge \forall y(W^p y \rightarrow Wa(\epsilon x : M^p x \sqcup \epsilon y : \neg(W^p y \wedge \neg Wa(y))))$.

Now we may use plural pronouns with a summation of singular NPs as an antecedent.

6.87. *A man hated a woman.* They *agreed not to meet.*

Here we need the following simple principle.

Plural Pronoun Principle If the schematic term ν in $\psi[\nu/x]$ can be instantiated to $\epsilon x : \varphi$ and $\epsilon x : \psi$, then it can be instantiated to $\epsilon x : \varphi \sqcup \epsilon x : \psi$.

This principle allows the translation

6.88. $\exists x Mx \wedge \exists y Wy \wedge H(\epsilon x : M, \epsilon y : W) \wedge AGR(\epsilon x : (M \wedge H(x, \epsilon y : W)) \sqcup \epsilon y : (Wy \wedge H(\epsilon x : M, y)))$.

These remarks about plurals should be enough to give the flavor of the instantial logic treatment of ordinary plural noun phrases and of the way in which references of plural pronouns can be resolved. (A treatment along slightly different lines is proposed by van den Berg [Ber95].)

6.8 Bare Plurals

In the area of plurals, *bare plurals* take a special place. A plural CN like 'lions' can occur quantified — 'Some lions have manes' — or 'bare' — 'Lions have manes' (see [Car77, Car91]). In the second sentence the plural is to be interpreted generically. That is, the second sentence must be interpreted as some qualified sort of *universal* statement. Of course, the main question is here: qualified in what way? The occurrence of the CN 'lions' in the first sentence does not cause any problems, it can be interpreted standardly as an existentially quantified phrase. Now ϵ-terms seem eminently suitable to treat bare plurals. As Carlson [Car77] has remarked, on the one hand, *syntactically* bare plurals seem to behave as proper names; they can occur in the same contexts, with the same behavior as proper names. This behavior lead Carlson to the conclusion that bare plurals, if generically used, should be interpreted as names of *kinds*. On the other hand, *semantically* they seem to be belong to the family of quantifier expressions and, indeed, the general analysis of bare plurals in the most extensive study at present [KPC+92] interprets generic expressions as belonging to the family *generalized quantifier* expressions. In ϵ-terms the quantifier expressions and proper names, that is, *terms* of the language, seem to come together.

Not all occurrences of bare plurals are to be interpreted generically. A plural like 'cigars' can occur generically or existentially in a sentence and the difference can be quite subtle. Consider

6.89. *John hates cigars.*

versus

6.90. *John smokes cigars.*

The occurrence of 'cigars' in the first sentence is a generic one but in the second sentence it is not. The argument here is by monotonicity. If John *smokes* cuban cigars, then he smokes cigars. So in this context the bare plural has the typical upward monotonicity behavior of indefinites ([Ben86a]). But the fact that John *hates* cuban cigars does not entail that John hates cigars. In this context the bare plural has the monotonicity behavior of proper names or universally quantified expressions.

From a logical point of view the interpretation of bare plurals faces two related problems. First of all, in a logical semantics we are interested in the *truth conditions* of sentences. Here one is prone to confuse two facts: the specification of truth conditions and truth in in the 'real world'. Consider the sentences "Lions have manes" and "Lions are male". Both sentences are generic, both express generalizations about lions, or the kind 'lion'. The first is a 'true' statement, the second a 'false' one. Now, when we are doing model theoretical semantics, it is not our task to explain why the first sentence is true *in the real world* (itrw) and the second sentence false itrw. After all, a world where, normally, lions are male is not a *logical* impossibility. This much is obvious, as semanticists it is our task to specify truth-conditions, and if this can be done for generic sentences at all, then it can be done for the first as well as for the second sentence. The confusion about the task of semantics is evident in Krifka's et al. treatment of Declerck's *Relevant Quantification.* Declerck [Dec91] analyzes generic quantification in terms of universal quantification and a monadic relation **R** restricting the domain of quantification to the 'relevant' individuals. Our proposal will take this intuition from Declerck. A sentence like

6.91. *Whales give birth to live young.*

is analyzed as (6.92).

6.92. $\forall x(whale(x) \wedge \mathbf{R}(x) \rightarrow gives\ birth\ to\ live\ young(x))$.

Of course, this is problematic. For the appropriate relation **R** must come out of the blue. Consider "Lions have manes" and "Lions suckle their young". In the first case the relevant set has to be the set of male lions, in the second case the set of female lions. Then how do we analyze the (true) generic "Lions have manes and suckle their young"? However, every semantics that has been suggested faces this problem, so we want to focus on a different point, namely, the following objection of Krifka et al. ([KPC+92], Ch. 2, p. 32) to this semantics.

> "One obvious problem with this approach is that this principle, as it stands, can easily justify all kinds of generic sentences ... since it is easy to find restrictions which would make any quantification come true. For example, the analysis could make
>
> (79) Whales are sick
>
> be a true generic, since we can take **R** to be the predicate **sick**, hence to restrict the quantification to sick whales".

The problem here seems not so much that we can trivialize this proposal when no restrictions on possible relations **R** are imposed, but that it is not true that whales are sick, so we should not allow relations **R** that force this to be true (after all, this semantics can be trivialized also for true generics).

But why is there a problem here? Indeed, any generic can be justified in this way, just as any particular sentence "Mary swam across the Atlantic ocean","John held his breath for ten months" can be true on the right model. A semantic analysis of these sentences should result in a class of models verifying the sentences; whether these models fit the real world is not our concern. Even if we find a way to restrict the **R** relation so as to exclude triviality, still it must be possible to make "any quantification come true", for even "whales are sick" can be a true generic in some model.

It is important to emphasize this point, because the above restriction of quantification occurs in every proposal, most notably in the Modal Semantics [Del87] and that in terms of Situation Theory [GL88]. Given the right 'ordering source', we can make "whales are sick" true in the modal set-up, and given the right 'back-ground' situation the same generic can be made true in the situation theoretical framework.

There remains however the problem of *preservation of truth* by logical reasoning. A sound reasoning pattern allowing us to infer B from A will guarantee the truth of B itrw if A is true itrw. So a pattern like: from "Lions have manes" and "If a lion has a mane, then it is male" conclude "Lions are male", should not be allowed[4].

[4]Most non-monotonic logics allow this however: if $\varphi(x)$ is non-monotonically derivable and

The fact that a proper semantics should be able to deal with true as well as with untrue generic statements constitutes a major problem for an interpretation of generics in a non-monotonic framework. In a *data base* we shall not add a statement like "birds have rubber wings". So, the problem of false generics does not occur there. For instance, in default logic, "Lions have manes" is interpreted as: if we have a lion and it is consistent to assume it has manes, then conclude it has manes. So far, so good. But if it is consistent to assume that this lion has manes, then it is consistent to assume it is male, so "Lions are male" is also accounted for. In data base logic this does not constitute a problem because we just don't add a statement like "Lions are male" to our base, but a semantics of natural language will have to deal with untrue generics. The problem is of course that we cannot deal with falsehood in terms of standard *counter examples*, for these exist also for true generics.

6.9 Generics Explained in Terms of Relevant Instances

Because of the wide variety of data the treatment of generic statements as 'universal statements with exceptions' seems to be misguided. If we only concentrate on a sentence like "Birds fly", then indeed most birds fly. But in "Rats bother people" and "Turtles live to be very old", the exception seems to be the rule: very few rats ever bother people and most turtles do not survive their first day. Every 'proportion' of individuals can occur in the interpretation of some generic statement, from the full set of individuals (dogs are mammals) to the majority (birds fly), a minority (turtles live to be very old) down to the empty set (dogs are widespread). In this sense the number, or proportion, of individuals verifying a sentence seems an irrelevant attribute for the genericity of a sentence.

In this section we are going to treat bare plurals as plural ϵ-terms which denote sets of *relevant* instances. This notion of relevancy is independent of the *cardinality* of the (non-empty) set of instances. Thus we exclude sentences like "Dogs are common" from our analysis. That is, we shall only deal with so-called *derived kind predication* [KPC$^+$92]. These sets of *relevant* elements, assigned to a plural ϵ-term $\epsilon x : \varphi$, we shall represent, relative to a specific set G_Φ of choice functions, by $|\epsilon x : \varphi|^{s,G_\Phi}$.

6.11. DEFINITION. (Value Functions) For a model M and variable assignment s, The *value function* $|\cdot|^{s,G_\Phi}$ is a mapping from ϵ-terms and formulas to subsets of M defined as follows

$\forall x(\varphi(x) \rightarrow \psi(x))$ holds classically, then $\psi(x)$ is derivable. This shows that extensional treatments in terms of majorities or probabilities are impotent here.

1. $|\epsilon x : \varphi|^{s,G_\Phi} = V_{M,s,G_\Phi}(\epsilon x : \varphi)$,
2. $|\varphi(x)|^{s,G_\Phi} = \{m \in dom(M) \mid M, s, G_\Phi \models \varphi\}$.

Here the set $|\epsilon x : \varphi|^{s,G_\Phi}$ determines the set of *relevant* elements of $|\varphi(x)|^{s,G_\Phi}$ relative to G_Φ. The next section will characterize logics with respect to relations between these sets.

6.10 Extended ϵ-Calculi

Whenever we have $M, s, G_\Phi \models \psi[\epsilon x : \varphi/x]$, we can identify a conditional, $\varphi(x) \succeq\psi(x)$, relating the *identifier* $\varphi(x)$ of $\epsilon x : \varphi$ with the formula $\psi(x)$. In the standard ϵ-calculus no such conditional is universally valid. For instance, $\mathcal{M} \models \neg\varphi(\epsilon x : \varphi)$ is not excluded (in fact, this is the case if and only if $\mathcal{M} \models \forall x \neg\varphi(x)$), and neither is $\mathcal{M} \models \psi(\epsilon x \varphi) \wedge \neg\psi(\epsilon x : (\varphi \wedge \varphi))$. Indeed, it is up to us to create conditionals, to create a logic, by extending the ϵ-calculus with deduction rules. For instance, the rule

$$\frac{\psi(\epsilon x : \varphi)}{\psi(\epsilon x : (\varphi \wedge \varphi))}$$

would exclude the second of the situations above. These new rules extend a system of classical deduction rules, so the interaction with a classical (proof theoretic) environment is straightforward. Moreover, models for these extended ϵ-calculi are classes of ϵ-models in which the sets of choice functions G_Φ satisfy special properties. Finally, every model for the ϵ-free fragment of a set Σ of $\mathcal{L}$-sentences can be supplied with a set of value assignments G_Φ satisfying these special demands such that the result is a model for Σ. So, given completeness, the resulting calculus is conservative over the ϵ-free fragment.

We shall introduce some notions to formulate the variety of semantics for the ϵ-calculus. The various semantics will be formulated in terms of relations between the sets $|\epsilon x : \varphi|^{s,G_\Phi}$ and $|\varphi(x)|^{s,G_\Phi}$. We have two families of subsets of M.

1. The family of *s, G_Φ-truth sets* will be denoted by $D_{s,G_\Phi}(M) = \{|\varphi(x)|^{s,G_\Phi} \mid \varphi(x) \in \mathcal{L}\}$.
2. The family of *s, G_Φ-value sets* will be denoted by $D^\epsilon_{s,G_\Phi}(M) = \{|\epsilon x : \varphi(x)|^{s,G_\Phi} \mid \varphi(x) \in \mathcal{L}\}$.
3. The mapping $i \subseteq D_{s,G_\Phi}(M) \times D^\epsilon_{s,G_\Phi}(M)$ is defined by $i = \{\langle |\varphi(x)|^{s,G_\Phi}, |\epsilon x : \varphi(x)|^{s,G_\Phi}\rangle \mid \varphi(x) \in \mathcal{L}\}$.

The sets $D_{s,G_\Phi}(M)$ and $D^\epsilon_{s,G_\Phi}(M)$ consist respectively of all *definable* sets on $\mathcal{M}$ and all *value ranges* of ϵ-terms. Note that these ranges need not be definable in the standard sense and that the mapping i need not be functional.

The various semantics we shall consider will be formulated in terms of different relations between sets of these two domains. We shall have some minimal demands on these relations. First of all, we want to tie in with a recent proposal from the linguistics literature to interpret generic statements as shaped by some *generalized quantifier*, i.e., 'Birds fly' has the logical form $Q_\epsilon x(\mathrm{Bird}(x))(\mathrm{Fly}(x))$ for some quantifier $Q_\epsilon x$ to be specified. Such a quantifier $Q_\epsilon x(\psi(x))(\varphi(x))$ can be seen as a notational variant of a conditional assertion $\psi(x) \mathrel{|\!\!\approx} \varphi(x)$ in the sense of [KLM90]. Standardly, a generalized quantifier is interpreted as some relation between (definable) subsets of the domain:

$$\langle \mathcal{M}, v\rangle \models Q_\epsilon x(\mathrm{Bird}(x))(\mathrm{Fly}(x))[g] \iff |\mathrm{Bird}(x)|^{s,G_\Phi} R_\epsilon |\mathrm{Fly}(x)|^{s,G_\Phi}.$$

In our case we fix the interpretation:

6.12. Definition. (Generalized quantifier interpretation) The binary relation R_ϵ on $D_{s,G_\Phi}(M)$ is defined by

$$|\psi(x)|^{s,G_\Phi} R_\epsilon |\varphi(x)|^{s,G_\Phi} \iff |\epsilon x\psi(x)|^{s,G_\Phi} \subseteq |\varphi(x)|^{s,G_\Phi}.$$

For this interpretation to work we need invariance of R_ϵ under logically equivalence: R_ϵ must be a binary relation on $D_{s,G_\Phi}(M)$. In other words i must be a *functional* relation on $\mathcal{P}(M)$. As a consequence, logically equivalent formulas $\psi(x)$ and $\varphi(x)$ must give rise to identical sets $|\epsilon x : \psi(x)|$ and $|\epsilon x : \varphi(x)|$.[5]

Moreover, if we set $E(\varphi(x))_{s,G_\Phi} = \{\psi(x) \mid |\epsilon x : \varphi(x)|^{s,G_\Phi} = |\epsilon x : \psi(x)|^{s,G_\Phi}\}$ then, by our definition of R_ϵ, we have: if $|\psi(x)|^{s,G_\Phi} R_\epsilon |\varphi(x)|^{s,G_\Phi}$ and $\chi(x) \in E(\psi(x))_{s,G_\Phi}$, then $|\chi(x)|^{s,G_\Phi} R_\epsilon |\varphi(x)|^{s,G_\Phi}$.

6.10.1 Standard Models

To get a logic of ϵ-terms started at all, we shall have to demand that all classical consequences of the *identifier* $\varphi(x)$ of an ϵ-term $\epsilon x : \varphi$ are properties of the object denoted by this term. This is the case if and only if $\mathcal{M} \models \exists x\varphi(x)$. For then we have $\mathcal{M} \models \varphi[\epsilon x : \varphi/x]$ and, consequently, if $\mathcal{M} \models \forall x(\varphi(x) \to \psi(x))$ then $\mathcal{M} \models \psi[\epsilon x : \varphi/x]$. So only the terms in $A^+ = \{\epsilon x\varphi \mid \mathcal{M} \models \exists x\varphi\}$ can be involved in logical relations on $\mathcal{M}$. Notice that this implies *reflexivity* of the quantifier $Q_\epsilon x$.

6.13. Definition. For every ϵ-model $\mathcal{M}$: $\mathrm{CON}_\mathcal{M} = \{\varphi(x) \mathrel{|\!\!\approx} \psi(x) \mid \langle \mathcal{M}, v\rangle \models \exists x\varphi(x) \wedge \psi(\epsilon x : \varphi)\}$.

6.14. Proposition. *On every ϵ-model $\mathcal{M}$, the set $CON_\mathcal{M}$ is closed under conjunction of consequents (AND), and weakening of consequents (RW). Moreover, for every $\varphi(x)$ occurring in some conditional statement in $CON_\mathcal{M}$ we have $\varphi(x) \mathrel{|\!\!\approx} \varphi(x) \in CON_\mathcal{M}$ (REFL).*

[5] Notice that this need not imply that $V_{M,s,G_\Phi}(\epsilon x : \psi) = V_{M,s,G_\Phi}(\epsilon x : \varphi)$ for individual Φ.

To give an idea of the possibilities of the extended ϵ-calculi we shall make a selection of three systems from [KLM90] adapted to our purpose.

6.15. DEFINITION. (Extended Calculi) 1. The *standard ϵ-calculus*, E, consists of a classical proof system plus the (ϵ) rule

"from $\exists x\varphi(x)$ infer $\varphi(\epsilon x : \varphi)$";

2. the *cumulative ϵ-calculus*, CE, consists of E plus the EQUIV rule

"from $\exists x\varphi$, $\psi(\epsilon x : \varphi)$ and $\varphi(\epsilon x : \psi)$ infer $\chi(\epsilon x : \varphi)$ if and only if $\chi(\epsilon x : \psi)$";

3. the *preferential ϵ-calculus*, PE, consists of CE plus the OR rule

" from $\exists x\varphi$, $\psi(\epsilon x : \varphi)$ and $\psi(\epsilon x : \chi)$ infer $\psi(\epsilon x : (\varphi \vee \chi))$";

4. the *monotonic ϵ-calculus*, ME, consists of PE plus the MON rule

"from $\exists x\varphi$, $\forall x(\varphi(x) \rightarrow \psi(x))$ and $\chi(\epsilon x : \psi)$ infer $\chi(\epsilon x : \varphi)$".

6.16. DEFINITION. (Extended models)

1. A *cumulative ϵ-model* is an ϵ-model satisfying for all variable assignments s and G_Φ

$|\epsilon x : \varphi(x)|^{s,G_\Phi} \subseteq |\psi(x)|^{s,G_\Phi} \subseteq |\varphi(x)|^{s,G_\Phi} \Rightarrow |\epsilon x : \varphi(x)|^{s,G_\Phi} = |\epsilon x : \psi(x)|^{s,G_\Phi}$.

2. A *preferential ϵ-model* is a cumulative ϵ-model satisfying for all variable assignments s and G_Φ

$|\epsilon x : (\varphi(x) \vee \psi(x))|^{s,G_\Phi} \subseteq |\epsilon x : \varphi(x)|^{s,G_\Phi} \cup |\epsilon x : \psi(x)|^{s,G_\Phi}$.

3. A *monotonic ϵ-model* is a preferential ϵ-model satisfying for all variable assignments s and G_Φ

$|\varphi(x)|^{s,G_\Phi} \subseteq |\psi(x)|^{s,G_\Phi} \Rightarrow |\epsilon x\varphi(x)|^{s,G_\Phi} \subseteq |\epsilon x : \psi(x)|^{s,G_\Phi}$.

We shall discuss the logics and their models.

6.10.2 Cumulative Models

The cumulative ϵ-calculus consists of the ϵ-calculus together with the rule

$$\text{EQUIV} \qquad \frac{\varphi(\epsilon x : \psi) \wedge \psi(\epsilon x : \varphi) \qquad\qquad \chi(\epsilon x\varphi)}{\chi(\epsilon x : \psi)}$$

This rule and the ones we shall formulate on the following pages must be interpreted *truth-to-truth* on the ϵ-models. For our purposes, the main consequence of the EQUIV rule is the fact that ϵ-terms with logically equivalent identifiers have the same derivable properties. So the mapping i is functional on all models satisfying EQUIV. In the cumulative system the familiar rules of *cumulative monotonicity* (CM) and *cautious cut* (CC) are derivable *for* $\epsilon x : \varphi \in A^+$.

$$\text{CM} \quad \frac{\psi(\epsilon x : \varphi) \quad \chi(\epsilon x : \varphi)}{\chi(\epsilon x : (\varphi \wedge \psi))} \qquad\qquad \text{CC} \quad \frac{\psi(\epsilon x : \varphi) \quad \chi(\epsilon x : (\varphi \wedge \psi))}{\chi(\epsilon x : \varphi)}$$

A *cumulative* ϵ-model satisfies: for every $E(\varphi(x))_{s,G_\Phi}$ and every $\psi(x)$, $\chi(x) \in E(\varphi(x))_{s,G_\Phi}$, if $|\psi(x)|^{s,G_\Phi} \subseteq |\xi(x)|^{s,G_\Phi} \subseteq |\chi(x)|^{s,G_\Phi}$ then $\xi(x) \in E(\varphi(x))_{s,G_\Phi}$. So in a cumulative model every $E(\varphi(x))_{s,G_\Phi}$ is a $\subseteq$-convex set with non-empty intersection.[6] By reference to convexity it is easy to see that if the mapping i satisfies cumulativity for all formulas $\varphi(x)$, and variable assignments s, then setting $i(\bigcup_{\Phi' \in G_\Phi} |\varphi(x)|^{s,G_\Phi}) = \bigcup_{\Phi' \in G_\Phi} i(|\varphi(x)|^{s,G_\Phi})$ lifts i to a cumulative mapping on $|\varphi(x)|^s$. This allows us to interpret $\epsilon x : \varphi$ on M with variable assignment s by $i(|\varphi(x)|^s) = |\epsilon x : \varphi|^s$.

6.17. PROPOSITION. *The cumulative ϵ-calculus is sound and complete with respect to the class of cumulative ϵ-models.*

6.18. PROPOSITION. *A set CON_M is closed under EQUIV if and only if $\mathcal{M}$ is a cumulative model.*

Minimal Entailment

A "minimal entailment" interpretation arises in a simple way. EQUIV allows us to define equivalence classes on A^+ by stipulating: $E(\epsilon x : \varphi)_{\mathcal{M}} = \{\epsilon x : \psi \mid \mathcal{M} \models \varphi(\epsilon x \psi) \wedge \psi(\epsilon x : \varphi)\}$. Let $[A^+] = \{E(\epsilon x : \varphi)_{\mathcal{M}} \mid \epsilon x : \varphi \in A^+\}$. Now define a binary relation R^c on $[A^+]$ as follows.

$$E(\epsilon x : \varphi)_{\mathcal{M}} R^c E(\epsilon x : \psi)_{\mathcal{M}} \iff \exists \chi \in E(\varphi(x))_{\mathcal{M}} : \ |\epsilon x : \psi| \subseteq |\chi(x)|.$$

R^c is a reflexive and antisymetric ordering on $[A^+]$ and

> $\mathcal{M} \models \psi(\epsilon x : \varphi)$ iff for all $\epsilon x : \chi$ R^c-minimal in $|\varphi(x)|$ we have $\mathcal{M} \models \psi(\epsilon x \chi)$.

This holds trivially because $|\varphi(x)|$ has exactly one minimum (modulo equivalence), namely $|\epsilon x \varphi|$.

6.10.3 Preferential Models

The *preferential* ϵ-calculus consists of the cumulative proof system together with the rule *OR*.

$$\text{OR} \qquad \frac{\varphi(\epsilon x : \psi) \qquad\qquad \varphi(\epsilon x : \chi)}{\varphi(\epsilon x : (\psi \vee \chi)}$$

Actually, CC is derivable from OR and CM. Consequently, equivalence relations between ϵ-terms can be defined in the OR+CM calculus. The following are derived rules of this calculus.

$$\frac{\varphi(\epsilon x : (\psi \wedge \chi))}{\psi(\epsilon x : \chi) \to \varphi(\epsilon x : \chi)} \qquad\qquad \frac{\varphi(\epsilon x : (\psi \wedge \chi)) \wedge \varphi(\epsilon x(\psi \wedge \neg\chi))}{\varphi(\epsilon x : \psi)}$$

[6] The intersection of $E(\varphi(x))_{s,G_\Phi}$ need not be an element of $E(\varphi(x))_{s,G_\Phi}$: $|\epsilon x \varphi|^{s,G_\Phi}$ need not be definable.

A preferential ϵ-model is a cumulative one in which for every $\varphi(x)$, if $\psi(x), \chi(x) \in E(\varphi(x))_{s,G_\Phi}$ then $\psi(x) \vee \chi(x) \in E(\varphi(x))_{s,G_\Phi}$. So here $E(\varphi(x))_{s,G_\Phi}$ consists of a $\subseteq$-convex set with non-empty intersection which is closed under finite unions. Consequently $E(\varphi(x))_{s,G_\Phi}$ is a convex set with a unique $\subseteq$-minimum, namely $\bigcap E(\varphi(x))_{s,G_\Phi}$, and a unique maximum, namely $\bigcup E(\varphi(x))_{s,G_\Phi}$. Neither of these need to be elements of $E(\varphi(x))_{s,G_\Phi}$, but if $\psi(x) \in E(\varphi(x))_{s,G_\Phi}$ is such a minimum or maximum, then $|\epsilon\psi(x)|^{s,G_\Phi} = |\psi(x)|^{s,G_\Phi}$, i.e., the set $|\epsilon\psi(x)|^{s,G_\Phi}$ is a *fixed* point of the mapping i.

6.19. PROPOSITION. *The preferential ϵ-calculus is sound and complete with respect to the class of all preferential ϵ-models.*

6.20. PROPOSITION. *A cumulative set $CON_{\mathcal{M}}$ is closed under OR if and only if $\mathcal{M}$ is a preferential ϵ-model.*

Again we may consider an minimal entailment interpretation. This time we define the ordering on $[A^+]$ by

$$E(\epsilon x : \varphi)_{\mathcal{M}} R^p E(\epsilon x : \psi)_{\mathcal{M}} \iff |\epsilon x : (\varphi(x) \vee \psi(x))| \subseteq |\varphi(x)|.$$

Now R^p is a partial order on $[A^+]$ and we have: $\mathcal{M} \models \varphi(\epsilon x : \psi)$ if and only if for all $E(\epsilon x \chi)_{\mathcal{M}}$ that are R^p-minimal in $|\varphi(x)|$ we have $\mathcal{M} \models \varphi(\epsilon x : \chi)$.

6.10.4 Monotonic Models

The clearest statement of the monotonicity of the monotonic ϵ-calculus, comes from the derivability of

$$\frac{\varphi(\epsilon x : \psi)}{\varphi(\epsilon x : (\psi \wedge \chi))} \qquad \frac{\varphi(\epsilon x : (\psi \vee \chi)}{\varphi(\epsilon x : \psi) \wedge \varphi(\epsilon x : \chi)}$$

On a monotonic ϵ-model the i-function respects Boolean structure. I.e., $|\epsilon x : (\varphi \vee \psi)| = |\epsilon x : \varphi| \cup |\epsilon x : \psi|$ and $|\epsilon x : (\varphi \wedge \psi)| = |\epsilon x : \varphi| \cap |\epsilon x : \psi|$.

A special kind of monotonic models, the *full* models are given by the rule

$$\frac{\psi(\epsilon x : \varphi)}{\forall x(\varphi(x) \to \psi(x))}$$

which forces $|\epsilon x : \varphi| = |\varphi|$ for all formulas φ.

6.21. PROPOSITION. *The monotonic ϵ-calculus is sound and complete with respect to the class of monotonic ϵ-models.*

6.22. PROPOSITION. *A preferential set $CON_{\mathcal{M}}$ is closed under monotonicity if and only if $\mathcal{M}$ is a monotonic ϵ-model.*

6.10.5 Sensitive Generic Semantics

Now that we have introduced logics to work with *relevant* extensions of plural ϵ-terms, we conclude with a brief sketch of how to apply this to concrete linguistic examples. In concrete applications we have to determine how a particular set of relevant instances is chosen. It is immediately clear that this set cannot be fixed for every context. Consider again the examples 'Lions have manes' and 'Lions breast feed their young'. In the first case the relevant lions are the male ones. In the second case only the female lions are relevant. In this case, it seems to be the *predicate* we apply to the term which determines relevancy. That is, the predicate 'have manes' maps the set of all choice functions G to some set G_Φ^{manes}, while the predicate 'breast feed their young' maps G to some other subset. The following definition shows how this simple strategy can be pursued.

6.23. DEFINITION. Let G be the set of all choice functions adequate for M. Every (complex) predicate Q of the language determines a set $G_\Phi^Q \subseteq G_\Phi$. We set $[Q] = \{P \mid G^Q = G^P\}$. Furthermore, let $M, s \models^+ Q[\epsilon x : \varphi/x]$ if $M, s, G^Q \models Q[\epsilon x : \varphi/x]$.

In '$\models^+$ the interpretation of an ϵ-term is influenced by the predicate it occurs in. Notice immediately the following consequence: the fact that $M, s, G_\Phi \models^+ Q[\epsilon x : \varphi/x]$ and that $M, s \models^+ \forall x(Qx \rightarrow Px)$ does not imply that $M, s \models^+ P[\epsilon x : \varphi/x]$. This does hold for stable fragments of the language. We call a fragment of language $\mathcal{L}$ *stable* if $[Q] = [P]$ for all predicate symbols Q and P occurring in the fragment.

6.24. DEFINITION. (Generic τ-Terms) If χ is implicational formula of the form $\varphi \rightarrow \psi$, then the interpretation of $\tau x : \chi$ is given by:

$$V_{M,s,G_\Phi}(\tau x : \chi) = \begin{cases} \in |\epsilon x : \varphi|^{s,G_\Phi^\psi} \wedge \overline{|\psi(x)|}^{s,G_\Phi^\psi} & \text{if } |\epsilon x : \varphi|^{s,G_\Phi^\psi} \wedge \overline{|\psi(x)|}^{s,G_\Phi^\psi} \neq \emptyset \\ \in |\epsilon x : \varphi|^{s,G_\Phi^\psi} & \text{if } |\varphi(x)|^{s,G_\Phi^\psi} \neq \emptyset \\ \in dom(M) & \text{otherwise.} \end{cases}$$

So, if we render 'lions' as $\lambda P.(P(\tau x : (Lions(x) \rightarrow P))$, then 'Lions are male' is an untrue, or incorrect, generic statement, because $\tau x : (Lions(x) \rightarrow Male(x))$ is mapped to a *relevant* counterexample: The predicate 'male' does not select a relevant subset of the form $|\epsilon x : Lion(x)|^{s,G_\Phi^{male}}$.

These definitions illustrate how naturally instantial logic extends to a theory of genericity. For further discussion and clarification of this extension we refer to Meyer Viol and Santos [MVS93] and Meyer Viol [MV93].

6.11 Conclusion

We have demonstrated the use of instantial logic as a stepping stone for a theory of anaphoric linking and anaphoric reference resolution, for a theory of plurality, and for a logic of genericity. To be sure, all three of these applications could and should be worked out further before a fruitful comparison with competing theories in natural language semantics is possible. Still, we hope to have demonstrated that the epsilon and tau terms of instantial logic can be used to shed new light on a number of key issues in the semantics of natural language.

One further application of instantial logic in the semantics of natural language which comes to mind is the making of a dynamic turn in instantial logic. Instantial logic suggests the following straightforward variation on Groenendijk and Stokhof's [GS91] Dynamic Predicate Logic (DPL): where DPL considers dynamic changes to variable assignment functions, dynamic instantial logic considers dynamic changing or extension of term valuation functions. In dynamic instantial logic, an epsilon term $\epsilon x : \varphi$ is interpreted dynamically as an extension of the domain of the current term valuation function with a new value d satisfying φ. Thus, dynamic instantial logic can be viewed as an extension of DPL with complex individual terms and suitable dynamic interpretations for those. Something along these lines is explored in Van Eijck [Eij94]. One nice feature of dynamic instantial logic is that the reference markers of Kamp [Kam81] or the file cards of Heim [Hei82] get natural interpretations as terms. Further connections between discourse representation theory and instantial logic are explored in Meyer Viol [MV92].

Bibliography

[ANvB94] H. Andréka, I. Németi, and J. van Benthem. Back and forth between modal logic and classical logic. Technical report, ILLC, University of Amsterdam, 1994.

[Ass57] G. Asser. Theorie der logischen Auswahlfunktionen. *Zeitschrift für mathematische Logik und Grundlagen der Mathematik*, 3:30–68, 1957.

[AvL95] N. Alechina and M. van Lambalgen. Generalized quantification as substructural logic. Technical report, ILLC, University of Amsterdam, 1995.

[BA93] J. van Benthem and N. Alechina. Modal quantification over structured domains. Technical report, ILLC, University of Amsterdam, 1993.

[Bar74] E.M. Barth. *The logic of articles in traditional philosophy.* Reidel, Dordrecht, 1974.

[Bar79] J. Barwise. On branching quantifiers in english. *Journal of Philosophical Logic*, 8:47–80, 1979.

[Bel93] J.L. Bell. Hilbert's ϵ-operator and classical logic. *Journal of Philosophical Logic*, 22:1–18, 1993.

[Ben86a] J. van Benthem. *Essays in Logical Semantics.* Reidel, Dordrecht, 1986.

[Ben86b] J. van Benthem. Partiality and nonmonotonicity in classical logic. *Logique et Analyse*, 114:225–247, 1986.

[Ben94a] J. van Benthem. Modal foundations for predicate logic. Technical Report CSLI-94-191, ILLC, University of Amsterdam, 1994.

[Ben94b] J. van Benthem. A new world underneath standard logic: Cylindric algebra, modality and quantification. In K. Apt, L. Schrijver, and N. Temme, editors, *From Universal Morphisms to Megabytes*, pages 179–186. Stichting Mathematisch Centrum, Amsterdam, 1994.

[Ber95] M. van den Berg. *The Internal Structure of Discourse.* PhD thesis, University of Amsterdam, 1995.

[Bet69] E.W. Beth. *Moderne Logica.* van Gorcum & Comp., N.V., 1969.

[Bla84] S. Blamey. Partial logic. In D. Gabbay and F. Guenthner, editors, *Handbook of Philosophical Logic, III*, pages 1–70. Reidel, Dordrecht, 1984.

[Bor85] B.R. Boročić. On sequence-conclusion natural deduction systems. *Journal of Philosophical Logic*, 14:359–377, 1985.

[Car77] G.N. Carlson. *Reference to Kinds in English.* PhD thesis, University of California, 1977.

[Car91] G.N. Carlson. Natural kinds and common nouns. In A. von Stechow and D. Wunderlich, editors, *Semantik (HSK)*, pages 370–398. Walter de Gruyter, Berlin, 1991.

[Car94] B. Carpenter. Distribution, collection, and quantification: A type-logical account of plurality. Technical report, OTS, University of Utrecht, 1994.

[Cel92] C. Celluci. Existential instantiation and normalization in sequent natural deduction. *Annals of Pure and Applied Logic*, 58:111–148, 1992.

[Dal83] D. van Dalen. *Logic and Structure (Second Edition).* Springer, Berlin, 1983.

[Dal86] D. van Dalen. Intuitionistic logic. In D. Gabbay and F. Guenthner, editors, *Handbook of Philosophical Logic, Volume III*, pages 225–339. Reidel, Dordrecht, 1986.

[Dec91] R. Declerck. The origins of genericity. *Linguistics*, 29:79–101, 1991.

[Dee91] K. van Deemter. *On the Composition of Meaning.* PhD thesis, University of Amsterdam, 1991.

[Del87] J. Delgrande. A semantics for defaults using conditional logic. *Artificial Intelligence*, 1987.

[Doe94] H.C. Doets. Wijzer in wiskunde. Unpublished Lecture Notes (in Dutch), ILLC, Amsterdam, 1994.

[Doe95] J. van der Does. Dependency logics. Unpublished article, University of Amsterdam, 1995.

[EA91] J. van Eijck and H. Alshawi. Logical forms. In H. Alshawi, editor, *The Core Language Engine*, pages 11–40. MIT press, Cambridge Massachusetts, 1991.

[Eij85] J. van Eijck. *Aspects of Quantification in Natural Language.* PhD thesis, University of Groningen, Groningen, 1985.

[Eij94] J. van Eijck. A natural term language. In K. Apt, L. Schrijver, and N. Temme, editors, *From Universal Morphisms to Megabytes*, pages 287–300. Stichting Mathematisch Centrum, Amsterdam, 1994.

[Eij95] J. van Eijck. Formal forays into language. Lecture Notes, University of Utrecht, Spring, 1995.

[End72] H.B. Enderton. *A Mathematical Introduction to Logic*. Academic Press, New York, 1972.

[Eva80] G. Evans. Pronouns. *Linguistic Inquiry*, 11:337–362, 1980.

[Fin85] K. Fine. *Reasoning With Arbitrary Objects*. Blackwell, Oxford, 1985.

[Gab81] D. Gabbay. *Semantical Investigations in Heyting's intuitionistic logic*. Reidel, Dordrecht, 1981.

[Gea80] P.T. Geach. *Reference and Generality: An Examination of Some Medieval and Modern Theories*. Cornell University Press, Ithaca, 1962 (Third revised edition: 1980).

[Gen34] G. Gentzen. Untersuchungen ueber das logische Schliessen. *Math. Zeitschrift*, 39:176–210, 405–431, 1934.

[GL88] C. Gerstner-Link. *Über Generizität. Generische Nominalphrasen in singulären Aussagen und Generische Aussagen*. PhD thesis, University of Munich, 1988.

[GM78] N. Goodman and J. Myhill. Choice implies excluded middle. *Zeitschrift für Mathematische Logik und Grundlagen der Mathematik*, 24:461, 1978.

[GS91] J. Groenendijk and M. Stokhof. Dynamic predicate logic. *Linguistics and Philosophy*, 14:39–100, 1991.

[HB39] D. Hilbert and P. Bernays. *Grundlagen der Mathematik*. Springer, Berlin, 1939. Second edition: Berlin 1970.

[HC84] G.E. Hughes and M.J. Cresswell. *A Companion to Modal Logic*. Methuen, London, 1984.

[Hei82] I. Heim. *The Semantics of Definite and Indefinite Noun Phrases*. PhD thesis, University of Massachusetts, Amherst, 1982.

[Hin73] J. Hintikka. *Logic, Language Games and Information*. Clarendon, 1973.

[HK83] J. Hintikka and J. Kulas. *The Game of Language*. Reidel, 1983.

[Hod93] W. Hodges. *Model Theory*. Cambridge University Press, 1993.

[Jan86] T.M.V. Janssen. *Foundations and Applications of Montague Grammar*. CWI Tract 19. CWI, Amsterdam, 1986.

[Jas94] J. Jaspars. *Calculi for Constructive Communication*. PhD thesis, Catholic University of Brabant, 1994.

[Kad87] N. Kadmon. *On Unique and Non-Unique Reference and Asymmetric Quantification*. PhD thesis, University of Massachusetts at Amherst, 1987.

[Kam81] H. Kamp. A theory of truth and semantic representation. In J. Groenendijk et al., editors, *Formal Methods in the Study of Language*. Mathematisch Centrum, Amsterdam, 1981.

[KLM90] S. Kraus, D. Lehman, and M. Magidor. Nonmonotonic reasoning, preferential models and cumulative logics. *Artificial Intelligence*, 44:167–207, 1990.

[Kol67] A.N. Kolmogorov. On the principle of excluded middle. In J. van Heijenoord, editor, *From Frege to Gödel*, pages 414–437. Harvard University Press, 1967.

[KPC+92] M. Krifka, J. Pelletier, G. Carlson, A. ter Meulen, G. Link, and G. Chierchia. Genericity: an introduction, version 5 (final). Unpublished manuscript, 1992.

[KW95] E.L. Keenan and D. Westerstahl. Generalized quantifiers in linguistics and logic. In J. van Benthem and A. ter Meulen, editors, *Handbook of Logic and Language*. 1995.

[Lam91] M. van Lambalgen. Natural deduction for generalized quantifiers. In J. van der Does and J. van Eijck, editors, *Generalized Quantifier Theory and Applications*, pages 143–154. Dutch Network for Language, Logic and Information, Amsterdam, 1991.

[Lei69] A.C. Leisenring. *Mathematical Logic and Hilbert's ϵ-symbol*. MacDonald, London, 1969.

[Lei73] D. Leivant. Existential instantiation in a system of natural deduction for intuitionistic arithmetic. Technical report, Stichting Mathematisch Centrum, Amsterdam, 1973.

[Lew70] D. Lewis. General semantics. *Synthese*, 22:18–67, 1970.

[Mey90] S. de Mey. *Determiner Logic or the Grammar of the NP*. PhD thesis, University of Groningen, 1990.

[Min77] G. Mints. Heyting predicate calculus with epsilon symbol. *J. Soviet Math.*, 8:317–323, 1977.

[Min94] G. Mints. An epsilon substitution method for the predicate logic. manuscript, CSLI, April, 1994.

[MV92] W. Meyer Viol. Partial objects and DRT. In P. Dekker and M. Stokhof, editors, *Proceedings of the Eighth Amsterdam Colloquium*, pages 381–402. ILLC, Amsterdam, 1992.

[MV93] W. Meyer Viol. Non–monotonic logic in the epsilon calculus. In *Proceedings of the Dutch/German Workshop on Nonmonotonic Reasoning Techniques and Their Applications*. RWTH Aachen, 1993.

[MV95] W. Meyer Viol. A proof-theoretic treatment of assignments. *Bulletin of the IGPL*, 3 (forthcoming), 1995.

[MVS93] W. Meyer Viol and D. Santos. The generic spell. Technical report, INESC, Lisbon, 1993.

[Nea90] S. Neale. *Descriptions*. Bradford Books, MIT Press, 1990.

[Ném93] I. Németi. Decidability of weakened versions of first-order logic. Technical report, Fifth European Summer School in Logic, Language and Information, Lisbon, 1993. Lecture Notes Workshop on Algebraization of Logics.

[Pra65] D. Prawitz. *Natural Deduction: A Proof-theoretical Study*. Almquist and Wiksell, Stockholm, 1965.

[Qui52] W.V.O. Quine. *Methods of Logic*. Routledge and Kegan Paul, London, 1952.

[RdL89] G.R. Renardel de Lavalette. Interpolation in fragments of intuitionistic propositional logic. *Journal of Symbolic Logic*, 54:1419–1430, 1989.

[Sán91] V. Sánchez. *Studies on Natural Logic and Categorial Grammar*. PhD thesis, University of Amsterdam, 1991.

[Str86] D.J. Struik, editor. *A Source Book in Mathematics, 1200–1800*. Princeton University Press, 1986.

[Tai65] W. Tait. The substitution method. *Journal of Symbolic Logic*, 30:175–192, 1965.

[TD88] A.S. Troelstra and D. van Dalen. *Constructivism in Mathematics / An Introduction Volume I*. North-Holland, Amsterdam, 1988.

[Thi92] E. Thijsse. *Partial Logic and Knowledge Representation*. Ph.D. thesis, Catholic University of Brabant, 1992.

[Ume59] T. Umezawa. On logics intermediate between intuitionistic and classical predicate logic. *Journal of Symbolic Logic*, 24:141–153, 1959.

[Zim93] E. Zimmerman. Scopeless quantifiers and operators. *Journal of Philosophical Logic*, 22:545–561, 1993.

[ZT87] J.I. Zucker and R.S. Tragesser. The adequacy problem for inferential logic. *Journal of Philosophical Logic*, 7:501–516, 1987.

[Zuc87] J.I. Zucker. The adequacy problem for classical logic. *Journal of Philosophical Logic*, 7:517–535, 1987.

Samenvatting

Instantiële Logica

Dit proefschrift heeft als thema logisch redeneren met zogenaamde willekeurige objecten. Zulke objecten worden geïntroduceerd in een logische of wiskundige redenering als een instantiëring van een algemeen begrip. De instantiëring is *willekeurig* als redeneringen met behulp daarvan tot conclusies leiden die geldigheid hebben voor het algemene begrip. Een klassiek voorbeeld uit de wiskunde is het bewijs dat de som van de hoeken van een driehoek 180 graden bedraagt. Hiertoe voert men een constructie uit op een *conrete* driehoek (op papier of op het bord). Het ligt in de aard van die constructie dat hetgeen men toont voor deze instantiëring van het begrip 'driehoek' geldt voor driehoeken in het algemeen.

Na een korte kennismaking met het onderwerp in hoofdstuk 1 van het proefschrift worden in hoofdstuk 2 de componenten van instantiële logica geïntroduceerd die in het proefschrift aan een onderzoek worden onderworpen. Dit zijn het bewijssysteem van ***natuurlijke deductie***, de ***epsilon calculus*** en de theorie der *willekeurige objecten*. In natuurlijke deductie wordt de betekenis van de logische kwantoren volledig bepaald door zogenaamde introductie en eliminatie regels. Met name de existentiële kwantor is interessant vanuit dit perpectief. Een existentiële kwantor $\exists x\varphi(x)$ — het algemene begrip — wordt geëlimineerd door een term te kiezen, zeg t, en de deductie voort te zetten met $\varphi(t)$, dat wil zeggen, met de instantiëring van het algemene begrip. Leidt vervolgens de deductie tot een conclusie ψ en kan de term t beschouwd worden als zijnde willekeurig, dan is ψ een conclusie van $\exists x\varphi$.

In de epsilon calculus krijgen de termen die gekozen worden bij de eliminatie van een existentiële kwantor syntactische structuur. Deze structuur verraadt de reden waarom de term geïntroduceerd is. De calculus wordt bepaald door de *epsilon regel*: uit $\exists x\varphi(x)$ concludeer $\varphi(\epsilon x : \varphi)$. Hier is de term $\epsilon x : \varphi$ een zogenaamde epsilon term. De interne structuur van dergelijke termen maakt afhankelijkheden syntactisch expliciet die impliciet blijven in de epsilon-vrije natuurlijke deductie. Deze afhankelijkheden zijn het onderwerp van de hoofdstukken 4 en 5.

De theorie der willekeurige objecten geeft een semantiek waarin willekeurige en concrete objecten naast elkaar bestaan. Zij verschaft *generisch redeneren* een gezonde logische basis. In dit hoofdstuk worden natuurlijke deductie, epsilon termen en willekeurige objecten behandeld binnen het kader van de klassieke, eerste-orde, logica.

In hoofdstuk 3 wordt de epsilon calculus onderzocht in het kader van de *intuitionistische logica*. Als toevoeging aan de klassieke logica levert de epsilon regel niets nieuws: zolang we ons beperken tot epsilon-vrije formules is hetgeen afleidbaar is met behulp van de epsilon regel ook afleidbaar zonder deze regel. Toevoeging van de epsilon regel levert een *konservatieve uitbreiding* van de klassieke eerste-order logica. Deze situatie verandert wanneer we de epsilon regel toevoegen aan de intuitionstische logica. Dit hoofdstuk onderwerpt de logica die aldus ontstaat aan een bewijstheoretisch en een semantisch onderzoek. Dit resulteert in een zestal zogenaamd 'intermediate logics' die, op één na, alle 'frame onvolledig' zijn.

Hoofdstuk 4 behandelt de vraag hoe de epsilon regel aangepast moet worden opdat een konservatieve uitbreiding ontstaat van de intuitionistische logica. De beantwoording van deze vraag leidt tot een diepergaande analyse van *afhankelijkheden* tussen formules en termen die optreden in een natuurlijk deductie systeem waarin assumpties gemaakt en weer ingetrokken kunnen worden. Het voorkomen van een formule van de vorm $\exists x\varphi(x)$ in een afleiding leidt bijvoorbeeld tot de keuze van een term t en de introductie van de formule $\varphi(t)$. In dit geval *hangt* de term t *af* van de formula $\exists x\varphi$. Het blijkt dat er essentiële verschillen zijn tussen de boekhouding van assumpties in klassieke en in intuitionistische logica. In klassieke logica mag de assumptie $\exists x\varphi$ ingetrokken worden terwijl de term t nog aanwezig is als 'getuige', in intuitionistische logica daarentegen is dit niet toegestaan.

In hoofdstuk 5 wordt het onderzoek van afhankelijkheden in natuurlijke deductie bewijzen voortgezet. De nadruk ligt hier op afhankelijkheden tussen *termen* en het begrip van een *keuze proces* wordt geïntroduceerd om deze term afhankelijkheden te interpreteren. In dit hoofdstuk wordt de fijnstructuur onderzocht van de regels die de existentiële kwantor introduceren en elimineren. Dit leidt tot variaties op deze regels die zwakkere existentiële kwantoren bepalen dan de standaard kwantor. Bovendien wordt de epsilon calculus gebruikt voor de onwikkeling van een logische taal waarin term afhankelijkheden expliciet worden bijgehouden. Dit hoofdzakelijk bewijstheoretische hoofdstuk wordt afgesloten met een korte bespreking van een mogelijke semantiek.

Hoofdstuk 6 bevat een drietal toepassingen van de epsilon calculus op de analyse van de natuurlijke taal. De eerste toepassing betreft de semantiek van persoonlijke voornaamwoorden. Er worden algemene principes voorgesteld voor de representatie van voornaamwoorden als epsilon termen in de logsiche vorm van een zin of tekst. Deze principes beïnvloeden zowel de *distributie* als de *betekenis* van deze woorden. De zogenaamde 'donkey zinnen' komen aan de orde, maar ook 'Bach-Peters zinnen' and het verschijnsel van 'modale subordinatie'.

De tweede toepassing behandelt meervoudige naamwoordsgroepen. De 'generische' interpretatie van epsilon termen die in hoofdstuk 2 is besproken wordt gebruikt voor een verrassend eenvoudige semantiek van deze woordgroepen.

De derde toepassing bestaat uit een semantische verkenning van het netelige gebied van generisch taalgebruik. Hier worden in het kort enkele bestaande theorieën besproken en wordt een keuze gemaakt voor een analyse van generisch gebruikte meervouden in termen van "relevante" instanties. Een semantiek wordt besproken die nauw aansluit bij recente theorieën over niet-monotoon redeneren.

Curriculum Vitae

Op 17 juni 1954 ben ik geboren te Den Haag. Mijn jeugd heb ik doorgebracht in Maastricht waar ik het Henric van Veldeke college heb bezocht. In 1973 heb ik daar het diploma gymnasium β behaald. Na een werelreis van een jaar heb ik me in 1974 ingeschreven aan the Centrale Interfaculteit van de Rijks Universiteit Groningen voor de studie Wijsbegeerte. In 1976 heb ik me bovendien ingeschreven aan de Faculteit Psychologie van dezelfde universiteit. Het jaar 1978 heb ik wederom rondtrekkend doorgebracht. In 1983 heb ik in het kader van de studie Psychologie een jaar door gebracht aan het Max Planck Institut für Psycholinguistik te Nijmegen. In 1986 heb ik, onder leiding van professor J. van Benthem, de studie Wijsbegeerte cum laude voltooid met een doctoraal scriptie over Temporele Logica. In 1987 heb ik een 8000 DM in de wacht gesleept door de EWG kwis te winnen op de Duitse televisie. Mijn onvrede met de huidige Kognitieve Psychologie heeft zijn weerslag gevonden in het boek "De Oorsprong van Gedrag" waaraan in 1989 de Wolters-Noordhoff Academieprijs is verleend. Het jaar 1990 heb ik doorgebracht met tijdelijke dienstverbanden aan de Faculteit Wiskunde en Informatica van de Universiteit van Amsterdam. In 1991 ben ik voor een periode van 4 jaar als AIO in dienst getreden bij het Onderzoeksinstituut voor Taal en Spraak (OTS) van de Letterenfaculteit van de Universiteit Utrecht.

Wilfried Meyer Viol

Titles in the ILLC Dissertation Series:

Transsentential Meditations; Ups and downs in dynamic semantics
Paul Dekker
ILLC Dissertation series 1993-1

Resource Bounded Reductions
Harry Buhrman
ILLC Dissertation series 1993-2

Efficient Metamathematics
Rineke Verbrugge
ILLC Dissertation series 1993-3

Extending Modal Logic
Maarten de Rijke
ILLC Dissertation series 1993-4

Studied Flexibility
Herman Hendriks
ILLC Dissertation series 1993-5

Aspects of Algorithms and Complexity
John Tromp
ILLC Dissertation series 1993-6

The Noble Art of Linear Decorating
Harold Schellinx
ILLC Dissertation series 1994-1

Generating Uniform User-Interfaces for Interactive Programming Environments
Jan Willem Cornelis Koorn
ILLC Dissertation series 1994-2

Process Theory and Equation Solving
Nicoline Johanna Drost
ILLC Dissertation series 1994-3

Calculi for Constructive Communication, a Study of the Dynamics of Partial States
Jan Jaspars
ILLC Dissertation series 1994-4

Executable Language Definitions, Case Studies and Origin Tracking Techniques
Arie van Deursen
ILLC Dissertation series 1994-5

Chapters on Bounded Arithmetic & on Provability Logic
Domenico Zambella
ILLC Dissertation series 1994-6

Adventures in Diagonalizable Algebras
V. Yu. Shavrukov
ILLC Dissertation series 1994-7

Learnable Classes of Categorial Grammars
Makoto Kanazawa
ILLC Dissertation series 1994-8

Clocks, Trees and Stars in Process Theory
Wan Fokkink
ILLC Dissertation series 1994-9

Logics for Agents with Bounded Rationality
Zhisheng Huang
ILLC Dissertation series 1994-10

On Modular Algebraic Prototol Specification
Jacob Brunekreef
ILLC Dissertation series 1995-1

Investigating Bounded Contraction
Andreja Prijatelj
ILLC Dissertation series 1995-2

Algebraic Relativization and Arrow Logic
Maarten Marx
ILLC Dissertation series 1995-3

Study on the Formal Semantics of Pictures
Dejuan Wang
ILLC Dissertation series 1995-4

Generation of Program Analysis Tools
Frank Tip
ILLC Dissertation series 1995-5

Verification Techniques for Elementary Data Types and Retransmission Protocols
Jos van Wamel
ILLC Dissertation series 1995-6

Transformation and Analysis of (Constraint) Logic Programs
Sandro Etalle
ILLC Dissertation series 1995-7

Frames and Labels. A Modal Analysis of Categorial Inference
Natasha Kurtonina
ILLC Dissertation series 1995-8

Tools for PSF
G.J. Veltink
ILLC Dissertation series 1995-9

(To be announced)
Giovanna Ceparello
ILLC Dissertation series 1995-10

Instantial Logic. An Investigation into Reasoning with Instances
W.P.M. Meyer Viol
ILLC Dissertation series 1995-11