On the formulae-as-types correspondence for classical logic

Charles Alexander Stewart

Worcester College

Revised version of thesis, July 2000

Oxford University Computing Laboratory
Programming Research Group

On the formulae-as-types correspondence for classical logic

Charles Stewart

Worcester College

Submitted for the degree of Doctor of Philosophy
Trinity Term 1999

Abstract

The Curry–Howard correspondence states the equivalence between the constructions implicit in intuitionistic logic and those described in the simply-typed lambda-calculus. It is an insight of great importance in theoretical computer science, and is fundamental in modern approaches to constructive type theory. The possibility of a similar formulae-as-types correspondence for classical logic looks to be a seminal development in this area, but whilst promising results have been achieved, there does not appear to be much agreement of what is at stake in claiming that such a correspondence exists. Consequently much work in this area suffers from several weaknesses; in particular the status of the new rules needed to describe the distinctively classical inferences is unclear.

We show how to situate the formulae-as-types correspondence within the proof-theoretic account of logical semantics arising from the work of Michael Dummett and Dag Prawitz, and demonstrate that the admissibility of Prawitz's inversion principle, which we argue should be strengthened, is essential to the good behaviour of intuitionistic logic.

By regarding the rules which determine the deductive strength of classical logic as *structural* rules, as opposed to the logical rules associated with specific logical connectives, we extend Prawitz's inversion principle to classical propositional logic, formulated in a theory of Parigot's lambda-mu calculus with eta expansions.

We then provide a classical analogue of a subsystem of Martin-Löf's type theory corresponding to Peano Arithmetic and show its soundness, appealing to an extension of Tait's reducibility method. Our treatment is the first treatment of induction in classical arithmetic that truly falls under the aegis of the formulae-as-types correspondence, as it is the first that is consistent with the intensional reading of propositional equality.

Contents

Introduction **4**

1 Logic and proof theory **29**
1.1 Modern Logic . . . 29
Logical formalism . . . 31
1.2 Natural Deduction . . . 34
The syntax of natural deduction . . . 34
The inversion principle . . . 38
1.3 Normalisation . . . 42
Properties of the calculus . . . 50
Normal form theorem . . . 52
1.4 The Curry–Howard correspondence . . . 53
The simple theory of constructions . . . 54
Some consequences of the correspondence . . . 60
1.5 Recursion . . . 66

2 Intuitionistic type theory **73**
2.1 Type dependency . . . 73
2.2 The calculus ITT . . . 76
2.3 Type soundness . . . 85
2.4 Representability . . . 89
2.5 Consistency . . . 96
2.6 Conversion theory . . . 101
Head normal form theorem . . . 104
2.7 Logical and constructive harmony . . . 105
2.8 A note about semantics . . . 111

3 Classical proofs I: propositions **113**
3.1 Classical strength reasoning . . . 113
3.2 Classical Natural Deduction . . . 120
3.3 Reduction and the inversion principle . . . 126
3.4 Recursion . . . 139
3.5 An alternative formulation . . . 153

4 Classical proofs II: arithmetic **157**
4.1 The naïve theory 157
Failure of the theory NTT 160
Reflections upon the unsound theory 161
4.2 The calculus CTT 164
4.3 Type soundness 165
4.4 Conversion theory 166
Canonical form theorem 168
4.5 Representability 170
4.6 Consistency and harmony 173

Conclusions **175**

A A summary of $\mathrm{PRA}^{\omega}_{\mu}$ and CTT **185**
A.1 The calculus $\mathrm{PRA}^{\omega}_{\mu}$ 185
A.2 The calculus CTT 188

Bibliography **193**

Index **206**

Introduction

This work grows out of recent developments in the proof theory of classical logic that suggest that one may provide a *formulae–as–types* correspondence for classical logic along the lines of the Curry–Howard correspondence for intuitionistic logic which has proved to be such a fruitful vein of ideas in theoretical computer science. The principal aims of this thesis are to make clear what is at stake in advancing such a claim, to examine and reformulate the existing correspondence for classical propositional logic based on Parigot's lambda-mu calculus, and to apply these ideas in making an extension of this correspondence to a treatment of the classical predicate calculus and arithmetic[1].

The picture presented here developed slowly, and had its origins in my earlier work upon providing an operational semantics for μPCF[2], a simple functional programming language with the additional syntactic features of Parigot's lambda-mu calculus, employed to model a theory of control after that of Mathias Felleisen[3]. It was my original aim to write a thesis that extended these ideas to a treatment that rivalled the operational account given by Andrew Pitts for PCF, perhaps similar to that given by Ian Stark for the nu calculus, a treatment of imperative state based upon ML[4].

However all of my attempts to develop such an account ended with results which I found profoundly unsatisfactory. A central task of such an account would be to provide a systematic account of the operational behaviour of the calculus, together with methods of reasoning about its properties, such as the coinductive techniques Andrew Pitts derived from the semantical account of logical relations for the lazy lambda calculus. Whilst progress was made towards such a definition, this was by a process owing more to educated guesswork than to analysis according to well-founded principles, a situation I found difficult to square with the clarity purported to flow from the formulae–as–types correspondence with classical logic.

[1]See Parigot's 'Classical proofs as programs' [ParigotM:clapp]. We shall discuss Parigot's work on treating the classical predicate calculus later in this introduction.

[2]This was joint work carried out with my supervisor, Luke Ong, attempting to provide operational semantics along the lines of that argued for by Andrew Pitts. See his *Operationally-Based Theories of Program Equivalence* [PittsAM:opebtp]. Our joint paper *A Curry–Howard foundation for functional computation with control* [OngCHL:curhff] appeared in January 1997.

[3]In 'The revised report on the syntactic theories of sequential control and state' [FelleisenM:revrst].

[4]See his *Names and Higher-Order Functions* [StarkI:namhof].

I came to the conclusion that there is something misleading about the discussion of the formulae–as–types correspondence in the literature. Often identified with the slogan 'proofs as programs', one sees in the literature claims that, because a term calculus has been developed 'in the formulae–as–types tradition', the term calculus is in correspondence with some logic. 'Which logic?' one might ask. No better answer is or could be given than the 'logic' that is obtained from considering the terms of each type as 'proofs' of the respective 'propositions'[5].

Something has gone wrong when we start to regard such vacuous sleight of hand as illuminating. In fact, the formulae–as–types correspondence gives no license to regard an arbitrary term calculus as a theory of proofs; rather it gives a congruence between formal structures proposed by two quite different subject matters: namely a particular kind of attempt to explain where it is that logic derives its meaning, and a particular theory of typed computations. What is deep about the formulae–as–types correspondence is that it forms a bridge between notions related to truth and logical justification on the one hand, and data type and effective procedure on the other. Existing work on the 'classical formulae–as–types correspondence' fails to answer the question: why is this term calculus of any *logical* interest?

I argue in this work that the best vantage point from which to understand the significance of the formulae–as–types correspondence lies in Michael Dummett's writings on the semantics of logic and language. As we will see in the following paragraphs, this takes us some distance from the original question we wish to answer into the realms of a philosophical controversy, but I believe that this is a necessary diversion: until we are clear in our conceptual underpinnings we are in persistent danger of being led astray in our reasoning.

The internal justification of logic

Dummett's original ideas on the semantics of logic and language centre upon his response to the dominant, truth-conditional account of semantics. This account has its origins in Frege's logicism, and has received its clearest formulation in the work of Alfred Tarski, which we shall briefly examine.

Tarski's logical work is best known for his definition of logical consequence and his semantic conception of truth. His definition of logical consequence in turn depends upon the notion of valuation: we relate the class of expressions we wish to explain to a denotational interpretation by means of an evaluation function, where

[5]David Lewis famously criticised a similarly vacuous system of semantics in linguistics, namely the Generative Semantics of Katz and Fodor:

> We can know the Markerese translation of an English sentence without knowing the first thing about the meaning of the English sentence: namely, the conditions under which it would be true. Semantics with no treatment of truth conditions is not semantics. *General Semantics* [LewisD:gens].

A treatment of truth, if perhaps not truth-conditions, must be at the heart of semantics.

the value of the function at an expression is determined according to the formal structure of that expression, and in particular the denotation of a sentence is its truth-value.

Each evaluation function will be associated with a consequence relation, and we are especially concerned with *logical consequence* which is defined to be the minimal consequence relation that holds under all possible[6] valuations. The logical connectives are those expression formers whose behaviour is not dependent upon the details of any particular interpretation, that is to say they are the connectives whose meaning can be specified by truth functions.

We may call Tarski's approach an exemplar of *external semantics*: we obtain our meaning for the formal language by relating it to an independently meaningful target language. In general we may characterise the external approach to justification by these two characteristics: it regards the semantics of a language in terms of valuation, and this valuation is conceived of extensionally. Such an extensional, evaluative approach to the relationship between meaning and truth is the hallmark of a *truth-conditional* theory of meaning.

Given a good picture of what kinds of things are to count as domains of valuation and evaluation functions, the Tarskian approach can achieve striking successes. The most notable success is the ability to interpret classical mathematics in terms of the cumulative hierarchy of set theory: all of the diverse parts of classical mathematics can be embedded in set theory in such a way that all mathematical entities can be interpreted as elements of a suitable set, and that all true theorems about these elements can be interpreted as true statements in set theory. This important achievement has been characterised as the *foundational completeness* of set theory for classical mathematics[7].

What could be wrong with this account? It is intuitive, economical and flexible. It is precisely these strengths that would give this account, if wrong, such a great power to mislead. Let me try to sow the seeds of doubt about this picture, in terms of its paradigmatic application to first-order arithmetic.

1. The interpretation of quantification over an infinite domain such as the natural numbers is necessarily based upon some such device as the following: to an expression $\forall x^{\mathbb{N}}.\phi(x)$, we determine the truth value by means of the set $\{[\![\phi(a)]\!] \mid a \in \mathbb{N}\}$, where '$[\![-]\!]$' is our evaluation function. If this set contains the value 'false' then the proposition is false, otherwise it is true.

2. It appears as though this gives us a determinate truth value for all sentences of arithmetic: all closed unquantified sentences of arithmetic are decidable, and an expression with a single quantifier, say $\forall x^{\mathbb{N}}.\phi(x)$, is true just when there is no number n for which $\phi(n)$ is false. This process can be applied recursively to determine the truth-value of sentences all the way up the arithmetic hierarchy.

[6] In the sense of 'logical possibility'.

[7] This phrase is due to Harvey Friedman.

3. This informal intuition outlined above can be considered as a kind of pseudo-categoricity result: Tarski's very reasonable model-theoretic interpretation of the quantifiers, together with our grasp of the truth of the closed arithmetic sentences, constitute our grasp of the standard model of arithmetic. It may seem as if we have decided the matter of the truth-values of sentences of arithmetic; certainly this is close to the picture I learnt when attending undergraduate lectures on model theory.

4. The above picture may be reinforced if we were to suppose that we must make the following either-or choice: *either* we embrace the standard model, *or* we embrace a non-standard model. The non-standard models include such monstrosities as numbers that cannot be reached by counting in a finite number of steps. Thus many have been led by our picture of evaluative semantics into regarding the pseudo-categoricity picture as a kind of completeness result. But this is not so: we only have fixed things *subject to having supposed our interpretation of set theory to be similarly fixed*. Indeed, any formal axiomatisation set theory is necessarily only an incomplete guide to the content of the infinite collections we interpret quantification in terms of: for instance any given primitive recursive formulation of set theory will determine formulae in first-order arithmetic whose truth-value is not determined under our interpretation, using Gödel's incompleteness theorem.

5. Let us emphasise the point: the set theoretic picture attempts to settle the interpretation of arithmetic in terms of axioms best understood as being about which infinite sets there are. But which is the right account? Can we accept Zermelo-Frankel set theory? Which large cardinal axioms are valid? What criteria are appropriate to criticise further hypothesised axioms? There is no formal 'statute of limitations' to assure us that the truth-under-interpretation of a given arithmetic formulae will not depend upon our answers to these kinds of question.

Perhaps we should change our picture to one that does not mislead us in this way? A minimum requirement we might place upon such a theory is that it does not artificially extend the meaning of the concepts of the theory we wish to understand with concepts properly lying beyond that theory. In particular we see that we must reject the presumption of the association of truth values to expressions, since to interpret quantification this way leads us to concepts to do with infinite sets.

How are we to achieve this? One kind of answer is presented by the idea of an *internal semantics*. To motivate this let us first examine the evaluative (but not purely external) account of Frege, and its criticism by Wittgenstein.

Frege's semantical account Frege was driven to provide a semantics for language not because of any great interest he had in the functioning of natural language,

but to provide a conceptual justification for his new logical system. Aristotle's syllogistic system derives its power to convince from his analysis of the structure and meaning of language. So when Frege overturns the Aristotelian analysis of judgement into subject and predicate, he must put something new in its place.

I do not wish to write a treatise on Frege's theory, so let me just briefly summarise the main points in bullet-point fashion. The numbered points correspond to propositions of Frege, the following bullet points are further comments either elaborating or clarifying those points.

1. Antipsychologism: our account of the meaning of logical and arithmetic concepts should make no reference to subjective states.

2. Against Kant's idea that analytic concepts are immediate, Frege motivates a notion of *indirect* analytic truth via the notion of proof.

    - Eg. 135664+37863=173527 is not immediate, but it can be proven by means of a complex demonstration in terms of a formal theory whose axioms are immediate. This is Frege's project of the Grundgesetze.

3. To express not only what things are true and what things are not, but also show how propositions may be informative, we must give an account of how it is we grasp truth.

    - 'a = b' is true just when the expressions 'a' and 'b' denote the same object, so our account must contain an account of denotation;
    - But an account of meaning which only makes use of denotation cannot explain how it is that 'a=b' is informative, for when 'a' and 'b' coincide in denotation, there is no denotational distinction between 'a=b' and the trivial, uninformative 'a=a'.

4. In the case of proper names, their reference (Frege uses "Bedeutung"in the German original) is the object itself whilst their sense ("Sinn") is how we pick out the object, ie. it is the object's mode of presentation in language.

    - Sense, unlike the idea or impression we may have of an object, is itself objective: the same sense may be possessed by many people, and it is what is grasped by any competent language user.
    - Reference is what language is about. Whilst sense may fail to pick out a referent, it is ordinarily the aim of sense to do so.

5. To explain how it is possible for many different senses to pick out the same denotation, Frege introduces his analysis of incomplete expressions and concepts. Both 'There are 500 men here' and 'There are four companies here' may be correct descriptions of the same state of affairs;

- Distinct senses pick out different ways we may analyse the denotation.
- This idea of the analysis of a sentence is the key to Frege's solution of the problem of multiple generality[8] in the Begriffschrift.
- In the above example, the incomplete expressions 'There are — men here' and 'There are — companies here' (where '—' marks the gap in the expression that may be filled by other expressions) denote distinct *concepts* applicable to the situation, fulfilled by different numbers. In other words, a concept is what Frege takes to be the denotation of an incomplete expression.
- Also the above examples may be seen to relate distinct concepts to truth-values: 'How many men are there here? 500.' or 'How many companies are there here? Four.' Here the distinct concepts are completed by numbers to yield true sentences.

6. Context principle: "the meaning of a word is not to be sought in isolation, but in the context of a sentence." In its application to number words:

   - Number expressions could not be simple referents to collections of 'units', ie. contentless individuals[9]: since the meaning of 'one' does not change from instance to instance, the unit referred to by 'one' is always the same, but in the denotation of 'two' we would have to distinguish between the two units referred to. Thus our abstract units would have contradictory properties, at once identical and distinct.
   - Numbers are not properties of individuals either: 'wise' is such a property and one may move from 'Solon was wise' and 'Thales was wise' to 'Solon and Thales were wise', but since one may not move from 'Solon was one' and 'Thales was one' to 'Solon and Thales were one', 'one' does not behave like a property of individuals.
   - Instead numbers are meaningful in the context of completing sortal concepts[10] into complete judgements, as in the above examples of 'There are — men here' and 'There are — companies here'. The numbers may are picked out by such expressions as 'The number of men that are here'.

[8]The problem of multiple generality is the failure of Aristotle's syllogistic to explain such inferences as 'If there is a cat of whom all mice are afraid, then all mice find some cat fearsome.'

[9]This is Euclid's view in the *Elements* (book VII) [Euclid:ele]:

Definition 1. *A unit is that by virtue of which each of the things that exist is called one.*

Definition 2. *A number is a multitude composed of units.*

Frege does not attack Euclid directly, but instead his target in the *Grundlagen* is Schröder's formulation of a derivative view in *Vorlesungen ber die Algebra der Logik* [SchroederE:voruda].

[10]Sortal concepts are concepts whose objects possess a well-behaved notion of identity, like 'is an animal' but unlike 'is water'.

Wittgenstein's break with evaluative semantics In his *Philosophical Investigations* [WittgensteinL:phii], Wittgenstein attacks in the most general terms theories that attempt to understand the semantics of language in terms of reference. Wittgenstein characterises such an approach as being one that subscribes to the 'Augustinian picture of meaning': after quoting a telling passage of Augustine from his Confessions[11], Wittgenstein goes on to observe "These words, it seems to me, give us a particular picture of the essence of human language. It is this: the individual words in language name objects –sentences are combinations of such names.– In this picture of language we find the roots of the following idea: every word has a meaning. This meaning is correlated with the word. It is the object for which the word stands."

Wittgenstein holds that this picture has the power to confuse because it leads us to make analogies between the meaning of words where language does succeed in being 'about' reference, in simple examples such as 'apple' or 'chair', to quite different units of meaning such as whole sentences, number words or abstract nouns such as 'goodness'.

It is hard to deny that Frege is guided by this Augustinian picture of meaning: in 'Über Sinn und Bedeutung' [FregeG:ubesub] he begins by making a number of observations about sense and denotation in the case of proper names, and then goes on to apply these observations to draw conclusions about the meaning of whole sentences and incomplete expressions. Let us examine Wittgenstein's rejection of Frege's picture, and what he seeks to put in its place. As before, only the numbered paragraphs ascribe views to Wittgenstein, the rest is further commentary.

1. Wittgenstein particularly targets Frege's adherence to the determinacy of sense, ie. the principle that every concept determines exactly which objects fall under it and which do not. Against this Wittgenstein holds that we cannot know how in advance how to apply our concepts in unexpected situations.

    - In the Grundlagen, Frege gives a revealing example: 'The same number belongs to the concept "inhabitant of Germany at the beginning of the year 1883, Berlin time" throughout eternity.' Throughout eternity? Even *before* 1883? And regardless of the numerous borderline cases between visitors, guests and inhabitants?
    - Certainly it is possible to offer refinements of our words that render their application in particular situations determinate (indeed Frege motivates his Begriffschrift in just such a way), but such a fixing would not correspond to the original sense of the word: we use words in practice before we discover these difficult cases, and so there is no sense in which offered refinements belong to the shared competence that Frege held to determine the sense of the disputed utterance.

[11] *Confessions* [Augustine:con]. Wittgenstein is generally held to have chose Augustine, rather than a particular semanticist, because he wanted to attack not the details of this or that theory, but the root, pretheoretical intuition that leads to referential theories of language.

- Frege's insistence upon this condition is presumably a consequence of his attachment to the objectivity of meaning. But difficult borderline cases do not normally threaten that objectivity: as Wittgenstein observes, we may be certain that a riverbank is overgrown with plants even though we have not yet decided which micro-organisms are plants and which are not.

2. Wittgenstein is as insistent about antipsychologism and the public nature of meaning as Frege; however he locates the source of the objectivity of language not in existence of an independent realm of meaning objects, but in our capacity to correctly determine the scope and application of a rule, and to recognise the correct application of these rules.

   - There is no need, therefore, to insist that a concept fixes a determinate extension on this view, it is enough that we have the capacity to recognise that it applies to one case, fails to apply in another.

3. Wittgenstein understands our grasp of language as being composed a bit like a jigsaw puzzle out of language games which are our practice-based understanding of how to apply words in certain paradigmatic situations. Just as the significance of the knight in the game of chess is not the piece of wood out of which the piece is made, but rather the powers of that piece in game play, so too the meaning of a word is correlated with the role that word plays in the language game, rather than any reference that may be ascribed to it.

   - 'How many cats are there in the garden?' we might ask a child. 'One... two... three... four... There are four cats in the garden!' would be an appropriate reply. Here we have an example of a question which asks for the number falling under a particular concept, and an algorithm (ie. recognising instances of a concept and counting them) which can be employed to fulfil the request.
   - This kind of language game can be seen as a paradigm situation establishing the meaning of number: to learn this language game is to learn what kind of role a number may play in our language and how it is possible to construct and recognise correct uses of number in assertions.
   - The language game makes use of capabilities we have: if we cannot move from the question to a grasp of the concept contained in the question then one is at a loss as to what has been asked for. Similarly if one grasps the concept, but cannot keep track of which cats one has so far counted, then one cannot apply this concept to come up with a tally.

Dummett's synthesis As I have presented him above, it may seem as if Wittgenstein is attacking particular weaknesses of Frege's account. This is not the case at

all: Wittgenstein takes himself to be demolishing the idea of anything that may be regarded as a systematic theory of meaning.

It is Dummett who saw that Wittgenstein's criticisms could be applied in a more targeted way, maintaining a broadly Fregean account of semantics, whilst rejecting Frege's evaluative semantics by moving to an account of meaning that emphasises sense over denotation. The definitive statement of this position is presented in his *Frege: philosophy of language* [DummettMAE:frepl].

1. Broad adherence to Fregean doctrines about meaning, subject to movement of the central locus of our semantical account from denotation to sense.

   - It is still the normal intent of words to denote. However language functions because of our *grasp* of those denotations, and so it is in terms of this grasp that our semantic theory should be formulated.
   - Furthermore it may be the case that significant parts of our language are not about an independently constituted reality. The existence of the objects that language appears to be about is frequently a matter of philosophical controversy, controversies Dummett characterises in terms of 'realism vs. anti-realism'. Ethics is often offered as an example, so, significantly to us here, is mathematics.

2. 'Meaning is use'

   - This slogan is derived from Wittgenstein: "For a large class of cases – though not for all – in which we employ the word 'meaning' it can be defined thus: the meaning of a word is its use in the language." Here Wittgenstein's qualification is taken to exempt such cases as 'These clouds mean rain'. Also, we employ meaning to indicate our intentions, such as in 'When I said "Slow down!" I meant that you were driving too fast for me.' So adherence to this slogan may be taken to mean that there is no more to the semantics of a word than may be seen in its employment.
   - Wittgenstein emphasised the motley of things we accomplish with our language, asking questions, complaining, expressing hopes, making declarations, etc. He is sometimes seen as arguing that this diversity renders the kind of systematic theory Frege was after impossible. Against this, proponents of speech act theory[12] that it is possible to accommodate this diversity by making a distinction in our grasp of an expression between its sense (that part of its meaning capable of being judged true), its force (the speech act effected by the expression) and its tone (which concerns

[12] Austin and Searle are recognised as the originators of speech act theory. See *How to do Things with Words* [AustinJL:howtww] and *Speech Acts* [SearleJ:spea]. Austin uses a different terminology to Dummett; he talks of the locutionary, illocutionary and perlocutionary acts associated with an utterance. I follow Dummett's terminology of sense, force and tone, where force is roughly equivalent to what Austin means by the illocutionary act.

the indirect connotations of a word, for example between 'policeman' and 'cop'). The plausibility of this analysis lies in our recognition of a common component of meaning shared between such expressions as 'The door is closed.', 'Is the door closed?' and 'Please close the door.'

- Assertions are kind of the speech act most characteristically used to relate expressions to truth-values, and so it is in relation to our usage of assertions that Dummett explains sense. Dummett argues that this sense cannot be explained in terms of truth-conditions, however: due to the phenomena of *verification transcendence*[13] use necessarily diverges from truth-conditions. So our theory cannot be given in terms of truth-conditions whilst remaining correlated with use.

3. Semantic molecularism.

   - The creative dimension of language, namely our ability to devise and understand sentences we have never seen before, and, equally importantly, to be in agreement with other language users in so doing, suggests that out grasp of language is reducible to our grasp of a finite number of simpler components.
   - This *compositionality* is sometimes expressed by the following principle[14] "The sense of a sentence is composed out of the senses of its parts".
   - The principle is applied by Dummett in an argument against a strong form of conventionalism, which says that there is no more to meaning than the arbitrary linguistic practices that people happen to follow. A consequence of this view is that if the meaning of a word lies in its use, then no established use can be incorrect. Against this, compositionality introduces the possibility of specifying possible incoherent linguistic practices (we will examine Prior's 'tonk' example in chapter one): as a consequence there must be more to meaning than convention alone.

The above remarks give a general picture of the compositional, antipsychologistic and use-based approach to semantics that lies behind the body of this work. However they do not by themselves explain how we are to provide an alternative formalisation of arithmetic. We need to move from general reflections about the nature of meaning, to particular judgements about how to capture meaning that can be applied in our metamathematical programme[15].

In *The Logical basis of Metaphysics* [DummettMAE:logbm], Dummett outlines his *verificationist* programme, which locates the meaning of a proposition with its *assertion conditions*, ie. grounds upon which it may be validly asserted. In this work

[13]Verification transcendence occurs when correct application of the truth condition of a sentence cannot be precisely determined by the users of the language. Dummett argues compellingly that any language which is sufficiently expressive must contain verification transcendent sentences.

[14]'Gedankengefüge' [FregeG:ged]. This is often called Frege's principle.

[15]These Dummett calls the 'programmatic aspects' of our semantics. I prefer the Scholastic term 'middle axioms' for such additional hypotheses.

I shall reject this proposal in favour of what I take to be a more general and robust idea, which is to locate the meaning of a proposition with its inferential role. I shall explain why I consider this to provide us with a better foundation after explaining how it is to be applied.

The plausibility of ascribing the meaning of a proposition to its inferential role depends upon a crucial observation due to Belnap, namely that a certain *harmony* must exist in that inferential role, between the valid grounds for making assertions and the consequences that flow from those assertions. If harmony fails to prevail in our system of inferences, than the practice of making assertions associated with those inferences will be incoherent, and so will fail to constitute a useful practice at all. This assignment of meaning to inferential role, subject to the requirements of harmony, I shall call *logical formalism*.

The next stage is to apply this intuition about meaning in the context of a formal system. We will work within Gentzen's calculus of natural deduction. Within this system Belnap's requirements can be satisfied by appeal to Prawitz's inversion principle and his proof of normalisation.

We have thus described a three-tiered system of the internalist semantics of logic:

1. The first and most general tier consists of the broad theses about meaning described in our discussion of Frege, Wittgenstein and Dummett;

2. Logical formalism constitutes our second tier, that is the location of meaning in inferential role subject to Belnap's criterion of harmony.

3. The most specific tier is our collection of techniques, derived from the writings of Dag Prawitz, used to justify the claim of harmony in the context of a set of rules given in a natural deduction formulation.

To the verificationist account of Michael Dummett, the third tier is again founded upon Prawitz's inversion principle, but the principle assumes a quite different kind of importance: namely in order to explain the whole of inferential role in terms of assertion conditions, the verificationist is committed to believe that elimination rules arise as a kind of inverse[16] to introduction rules. I believe this commitment puts verificationism at a justificatory disadvantage compared to the more relaxed inferential role semantics that I have adopted:

1. If this faith in the possibility of deriving elimination rules from introduction rules is justified in general, then verificationism and logical formalism coincide, since each account implies the truth of the premisses of the other.

2. If it is not possible in certain cases to obtain the elimination rule from the introduction rule, as I believe is the case in arithmetic, then there may be several possible elimination rules, none of which has any obvious grounds to be

[16]Hence Prawitz's choice of terminology.

preferred by the verificationist. This the verificationist is unable to systematically justify even inference.

3. Furthermore there appear to be difficulties facing the specification of certain logical connectives (eg. '$\supset$') purely in terms of assertion conditions.

This, then, is our justificatory framework. How does it relate to the formulae–as–types correspondence?

We find in the course of this work that we need to apply the formulae–as–types in two main places. Firstly, the account of arithmetic that we give in chapter two is the intensional variant of the intuitionistic type theory of Martin-Löf, which is founded upon the formulae–as–types correspondence. Our justification of that theory depends upon that correspondence in an apparently ineliminable way: in particular our demonstration of logical harmony in the case of the rules for equality depends upon an appeal to a new notion of constructive harmony which itself depends upon the formulae–as–types correspondence.

Secondly our formulation of the inversion principle for classical propositional logic depends upon the computational theory due to Griffin, and the refined theory of reductions of Parigot. Without this computational reading of classical proofs, there could be no internalist justification of classical logic and arithmetic.

Statement of achievement

The central idea of the thesis is that the so-called Curry–Howard correspondence for classical logic can be used to give a semantic basis for classical logic that accords with the three-tiered account of semantics that we have described, and that is built by analogy with the intuitionistic case.

To this end, we develop the semantic basis for intuitionistic logic in order to present a yardstick by which our achievements for classical logic can be developed. Chapters one and three present the accounts for intuitionistic and classical propositional logic, whilst chapters two and four treat the intuitionistic and classical accounts of arithmetic. By pairing the chapters in this way, and by following so far as is possible the same order of presentation of results, it is possible to judge the success of the classical analogs to their intuitionistic forbears.

The first chapter is responsible for introducing the key ideas of the semantic account. The idea that logical harmony may be used as a basis for a formal justification of logic, and that Prawitz's inversion principle shows how this idea may be applied to natural deduction, is present in the writings of Michael Dummett, but the further attempt to relate this account to the formulae-as-types correspondence is new[17].

The second chapter presents an exposition of Martin-Löf's intensional type theory, restricted to the arithmetic fragment (that is, using the type formers Π, Σ, $\mathbb{N}$

[17]New, but not surprising though: the related attempt to justify the meaning of intuitionistic logic by Martin-Löf, *is* explicitly related to the formulae-as-types correspondence.

and $=$). I give a justification of the semantic soundness of this type theory along the logical formalist lines given in chapter one; in the discussion we see that formulation of the inversion principle does not extend satisfactorily to the $\mathbb{N}$ and $=$ connectives, and so we reformulate the principle, relating it to a closer analysis of harmony than we provided in chapter one, and then use it to justify harmony for these connectives. This internal justification of Martin-Löf's type theory is distinct from Martin-Löf's own approach: I provide a sketch of an alternative semantics for the theory, and compare it to that of Martin-Löf. I shall describe an important weakness of Martin-Löf's account later in this introduction.

Chapter three applies the treatment of chapter one to classical propositional logic by appealing to an alternative notion of construction provided in Parigot's lambda-mu calculus. Parigot's work by itself does not by itself go far enough to secure the internal semantic coherence of classical logic, as the account of chapter one secures for intuitionistic logic, and so we begin our account by grounding the operations of the calculus in terms of elementary contraraieties, and showing how the inversion principle can be extended to the classical case.

It should be emphasised that the treatment of classical logic is not intended to convince someone who so far accepts only the intuitionistic rules that classical reasoning is valid: I make no attempt to argue that the new rules relating to contraraieties are acceptable to an intuitionist, as they patently are not. Instead the concern is to show that classical logic is perfectly intelligible according to the same desiderata used to justify intuitionistic logic, thus undermining claims that classical logic is incoherent, or incompatible with a constructive reading of the meaning of $\supset$ and $\wedge$.

Chapter four, the final chapter, demonstrates the power of the logical formalist semantics of meaning by providing an internal semantics for first-order Peano Arithmetic. Applying the constructions of chapter three to Martin-Löf's type theory, which we use as our semantic framework for interpreting proofs of Peano Arithmetic, is not straightforward: the formulation of the type theory needs to be fine-tuned to avoid difficulties with the characterisation of 'induction as recursion' in the presence of Parigot's operator. A close analog of the head normal form theorem can be proven, used to justify the semantics of the theory.

Chapter one: Logic and proof theory

Chapter one begins with an exposition of the logical formalist picture, introducing the problem of the justification of logic. We present a very brief history of the development of logic, principally in order to motivate some observations on Frege's achievement and also to introduce the problem of the justification of logic. We give a brief discussion on the exchange between Arthur Prior and Nuel Belnap on the possibility of a rules-based justification of logic.

In the next two sections we introduce the natural deduction calculus and its proof theoretic justification due to Prawitz as a fulfilment of Belnap's desiderata.

Our treatment broadly follows that of Prawitz[18], with two main differences. Firstly, we give a rather different account of the inversion principle, which is formulated with two clauses so as to fulfil both parts of Belnap's requirement of harmony, and furthermore our formulation is *local*: unlike in Prawitz's formulation, satisfaction of this formulation does not depend upon strong normalisation (the global property) obtaining in the whole calculus.

Secondly, we give a proof of a strong normalisation property. Whereas Prawitz's normalisation proof only entails weak normalisation, because only contractions associated with reducing head redexes (that is redexes on the principal path) need lead to reductions in Prawitz's ordinal measure, under our more subtle ordinal measure we present here, all reductions are decreasing. To my knowledge no such direct proof appears in the literature: the advantage of my proof is that the reduction measure is explicitly determined rather than being left implicit.

The results proven follow the treatment of Prawitz: we show the principal path theorem, the subformula property, and we show how consistency and Belnap's conservativity requirement follow from the subformula property.

In the following section we introduce the Curry–Howard correspondence, the formulae–as–types correspondence for intuitionistic logic. The technical details of the correspondence are well-known; however it is a defect of most introductions to the correspondence that appear in the literature that they do not give a clear idea as to what precisely is accomplished by the correspondence. In the light of our account we emphasise two points, firstly that the correspondence gives a precise articulation to the claim that intuitionistic logic is constructive, and secondly it gives us a theory of identity for proofs, a useful addition to the tool-box that we are developing to support our logical formalist account.

The last section of this chapter develops a brief treatment of recursion, allowing us to introduce Gödel's PRA$^{\omega}$[19], and outline the Girard–Tait proof of strong normalisation for it. As our theory involves eta expansions, we need a slight syntactic innovation to make our Girard-style proof go through, and we introduce a *relative normalisation* argument for these eta expansions to this end.

Chapter two: Intuitionistic type theory

We introduce *Martin-Löf's type theory* as, in the first place, a natural extension of the account developed for the intuitionistic propositional calculus to the predicate case in accordance with the Brouwer–Heyting–Kolmogorov account of the meaning of the intuitionistic connectives.

I had hoped not to need to write a full introduction to Martin-Löf's type theory, but I have not encountered a treatment of his calculus which is satisfactory for my purposes. Most treatments neglect to explain how we can apply cut-elimination results to prove the consistency of the calculus, but instead give semantic accounts,

[18] In *Natural Deduction* [PrawitzD:natd].

[19] Better known as system T, we follow the treatment of Feferman in referring to it as PRA$^{\omega}$.

or more common, no treatment of this at all[20]. Some treatments do explain this, but are not interested in the local dimension of the proof-theoretic justification of the rules. Indeed surprisingly few accounts even attempt to relate the provability strength of Martin-Löf's type theory to better known formal systems. It is a surprising lacuna in the literature that it is not possible to unhesitatingly recommend a text as an introduction to such an influential theory.

As a consequence, though this chapter assumes some familiarity with carrying out derivations in type theory, the meta-theory is presented from scratch.

We develop an abstract syntax, simpler than that presented in the work of Nordström, Petersson and Smith[21], which provides a theory of incomplete expressions and definitions for our system. Our principal reason for developing our simple account is that it is explicitly based upon substitution, and so unlike the former work, where the authors are able to infer eta-conversion from the conversions of the abstract syntax, we do not get any properties in the equational theory 'for free'.

We also base our calculus upon a system of telescopes, after de Bruijn, rather than the natural deduction resembling account of Martin-Löf. This arose from my own confusion when first studying the system about the precise role of type variables in the theory, which is made explicit in the account with telescopes. It also has the advantage that it is closer in form to those type theories intended for use in automated theorem provers. Otherwise, however, our treatment is standard, though restricted to quite a small fragment without universes.

We introduce the calculus proper, with brief explanations of the interpretations of the type formers, and then give a meta-theoretic investigation of the calculus, divided into four parts. First we provide some elementary adequacy results for our formal syntax, results which though not substantive are needed for the calculus to be meaningful. In the next section we consider representability, showing how it is possible to interpret equational logic, primitive recursive arithmetic and Heyting Arithmetic in the theory.

In the following two sections we consider meta-theoretic properties of the calculus: principally consistency, a correspondence with the system PRA^{ω} by which we can prove strong normalisation, and the canonical form theorem. The proof of the canonical form theorem depends upon our proof of the head normal form theorem, which provides the basis for our logical formalist justification of the type theory.

Our penultimate section is devoted to showing that logical harmony prevails in the calculus. The introduction of types such as $\mathbb{N}$ and $\mathbb{N} \Rightarrow \mathbb{N}$, which can only be understood as constructive sets, introduces a need for what we call an auxiliary account of *constructive harmony*. Furthermore, the justification of harmony in the case of propositional equality depends upon a reanalysis of harmony, and a new formal technique, derived from an idea of Prawitz. We also note that the term

[20]The treatment of $\mathbf{ML}_0^i$ in chapter 11 of *Constructivism in Mathematics* [TroelstraAS:conmi] is close to my perspective here, but the theory differs: the authors do not maintain a distinction between convertive and conceptual equality.

[21]*Programming in Martin-Löf's Type Theory* [NordstromB:promlt].

former $\mathbb{N}$ cannot justify the original harmony requirement, due to the phenomenon of the extensibility of arithmetic, and so we show a weaker formal property in this case.

In the last section, I shall say a little by way of comparison of the semantics given there with that of Per Martin-Löf. The meaning-theoretic interpretation given by Martin-Löf for each type is given in terms of the *head-normal forms* at each type. For example f is an element of $\Pi x^A.B(x)$ (under any assumptions) if it gives a method for transforming any head normal form $a \in A$ into a head-normal form $b \in B(a)$.

A consequence of this interpretation is that Martin-Löf's semantics violates compositionality: if A is an infinite type then the meaning of $\Pi x^A.B(x)$ depends upon the infinite range of particular instances $B(a)$. Martin-Löf argues[22] that this shows 'Frege's dictum' (ie. the claim that the sense of a proposition is composed out of the sense of its parts[23]) fails in the case of function space. I rather take it to show that Martin-Löf has failed to provide an account of the sense of $\Pi x^A.B(x)$.

The difficulty is one that attends any attempt to specify the meaning of functions space in terms of the introduction rules alone: Dummett is lead into just this difficulty in his 'Philosophical Basis of Intuitionistic Logic' [DummettMAE:phibil], where he tries to justify the meaning of universal quantification in terms of Brouwer's 'fully-analysed derivations'. We can avoid the difficulty in our two factor semantics, by appealing to the elimination rules for A in formulating the inference conditions for $\Pi x \in A.B(x)$, and to the introduction rules for A in providing our account of the consequences of $\Pi x \in A.B(x)$. It is worth emphasising that the two-factor semantics I give in the last section depends upon the account of logical harmony: it allows us to separate assertoric content from hypothetical contribution, and so rescues the account from the vacuity that Dummett notes threatens attempts to define content in terms of inference rules.

Review of the 'classical formulae–as–types' literature

Before we move onto considering the contents of chapter three, it will be valuable to survey the key achievements and defects of the work carried out on the formulae–as–types correspondence for classical logic.

The attempt to provide constructive characterisations of classical proofs date back to the independent work of Kolmogorov, Gödel and Gentzen[24] on the double negation translations, which establish a simple characterisation of classical provability in terms of intuitionistic provability by means of a translation on formulae. Andrei Kolmogorov at first believed that his translation established a constructive

[22] In the introduction to 'Constructive mathematics and computer programming' [MartinLofP:conmcp].

[23] A dictum not occurring anywhere in Frege's writings, but essentially the requirement of compositionality we have insisted upon already.

[24] Kolmogorov's 'On the principle of the excluded middle' [KolmogorovAN:priem] is the earliest account with those of Gödel and Gentzen following in the early 30s.

basis for classical logic, whilst vindicating the consistency of intuitionistic logic. Whilst it certainly achieves the latter aim, the former does not, due to the failure of the translation to satisfy the disjunction and existence properties.

The double negation was later used together with functional interpretation in a proof by Kreisel[25] to show an important conservativity argument, namely that for formulae of arithmetic complexity up to Π^0_2, first-order Peano Arithmetic and its intuitionistic analogue, Heyting Arithmetic, coincide in provability. An elementary translation on formulae was provided by Harvey Friedman establishing the same result[26] by more general means.

I described these results as the 'prehistory' of our work here because, though their results are seminal, and provide bounds upon the kind of account we want to develop, they are not concerned with providing direct interpretations of classical proofs.

Interest in this possibility was sparked by Girard's observation[27] that it was possible to give a denotation of cut-free proofs of classical logic by translation into his linear logic. The weakness of the translation that prevents it providing a full semantics of proofs is that certain terms are expected to serve as either of '!' or '?' modality. Girard later presented[28] a calculus that overcame this difficulty: at its heart was a distinction between formulae on the basis of an inferred polarity, and controlling which cuts are available to a formula based upon this polarity. We shall say a little more about Girard's result shortly.

Another result of major influence upon this paper arose quite independently. Mathias Felleisen proposed a systematisation of existing accounts of control, such as continuations and exceptions, in functional programming languages by means of an extension to the lambda calculus by special control operators. The reduction rules for these operators are obtained by means of a special internal interpretation called 'continuation passing style', or cps-, translations.

Cps-translations were used by Gordon Plotkin to encode alternative reduction strategies in the lambda calculus, so that the target derivation is an a sense 'strategy independent': the sequence of reductions performed by the reduction strategy upon the source term are made available only in that same sequence in the translated term. Felleisen observed that the control intuition lying behind the cps-translation allows a natural representation of the control operators, and so allows us to provide a theoretic perspective on this most operational of behaviours.

Felleisen's work was principally aimed at providing interpretations for control operators used in LISP, and so he was not concerned with the types of these operations. Timothy Griffin's observation was that if one attempted to infer types for one of Felleisen's operators, the '$\mathcal{C}$' operator, then one obtained the type $\neg\neg A \supset A$, suggesting that it could be used as a computational interpretation of classical proofs.

[25]'Mathematical significance of consistency proofs' [KreiselG:matscp].

[26]The A-translation, in 'Classically and intuitionistically provably recursive functions' [FriedmanH:claipr].

[27]In 'Linear logic' [GirardJY:linl].

[28]In 'A new constructive logic: classical logic' [GirardJY:newclc].

Furthermore the cps-translation of the calculus $\lambda + \mathcal{C}$ is type sound, yielding strong normalisation for the system.

Griffin's result was investigated in depth by Chetan Murthy in his PhD thesis, who attempted to apply the ideas to extracting constructive content from classical proofs. Murthy observed a close link between the double negation translations we described and the cps-translations of the $\lambda + \mathcal{C}$ calculus: different reduction strategies corresponded to different translations on formulae. Murthy successfully derived an automatic witness extraction algorithm for Higman's lemma (a non-trivial property concerning well-foundedness of a special kind of ordinal).

The success of work carried out in this vein lies in the fact that we possess both a clear grasp of what a 'classical' computation is, by its relationship with a theory of computation of independent interest, and also of what kind of thing we are attempting to obtain, namely terms-as-values in the intuitionistic formulae-as-types correspondence.

However work carried out in this vein suffers from a serious flaw, quite apart from the technical difficulties that dogged early results[29], namely that the account of classical proofs exists in a parasitic relationship with the underlying intuitionistic logic, deriving its meaning through a mediating interpretation by the double-negation translation.

Thus, for example, in Rehof and Sorensen's $\lambda_\triangle$ calculus, though the idea of an independently meaningful theory of classical proofs is clearly beginning to emerge, there is no canonical form theorem, and so no guarantee that a proof of some formula ϕ actually corresponds to any reasonable semantics we might specify for 'ϕ'.

The interpretation of classical proof theory into linear logic due to Girard, which we mentioned earlier, can be seen to be a candidate for such a theory. A difficulty with such an account is one that attends linear logic in general: linear logic is a revisionist theory, and one which, unlike intuitionistic logic, lacks a solid philosophical underpinning. If we hope to cast light upon classical logic from the logical formalist point of view articulated earlier, then one must provide a justification of the meaning of linear logic connectives on the logical formalist account, and what is obtained from them. Providing such a justification for linear logic is certainly a worthwhile achievement, but it seems to me that great difficulties face such a treatment.

We saw how the achievements of Griffin's calculus flowed from its close relationship with intuitionistic logic: we might for these reasons expect that treatments of classical proofs based upon extending existing accounts for intuitionistic logic to be the most promising. Distinguished amongst such treatments is Parigot's lambda-mu calculus.

[29] Griffin's original calculus suffered from both the failure of subject reduction and a conversion theory that depended upon the choice of reduction strategy. Felleisen's later work, 'The revised report on the syntactic theories of sequential control and state' [FelleisenM:revrst], corrected the difficulty with the type assignment system, and building on this, the Rehof and Sorensen developed a formulation of the $\mathcal{C}$ operator in 'The $\lambda_\triangle$ calculus' [RehofJ:lamdc] with a Church–Rosser reduction strategy.

Whilst the lambda-mu calculus possesses a close relationship with the control intuition[30], it actually arose out of Parigot's earlier work on Free Deduction, which can be seen as an attempt to find a formal representation of proofs with a character intermediate between natural deduction and sequent calculus.

Parigot's treatment of the propositional case is formally quite similar to our presentation here, and we discuss our differences in chapter three, though he is concerned only to prove existence and uniqueness of normal forms for the $\supset$ connective, and does not provide an account of eta equality. Parigot goes on to consider[31] the nature of inhabitants of the $\mathbb{N}$ type, and the existence of deviant inhabitants if we do not consider terms of the form $\mathsf{succ}(\mu\alpha.e)$ to be redexes.

Parigot rightly concludes that we cannot accommodate all of these deviant numbers into a correct theory of arithmetic. He presents a treatment of arithmetic in a polymorphic predicate calculus setting, based on two measures for reining in the bad behaviour of these terms.

Parigot's first measure is to use an analogue of Krivine's output operator to design terms which, when applied to deviant terms, reduce to canonical forms, and when applied to the usual inhabitants behave like the identity.

Parigot's second measure is a consequence of the logical framework he uses, based upon the calculus AF_2 of Krivine[32]. AF_2 is a calculus with both second-order quantification over types, permitting the definition of inductive data-types as described by Girard[33], and first-order quantification over typed individuals. The combination is a very expressive syntax for formulae.

The difficulty we describe in chapter four arises due to the potential bad behaviour of the new type formers of lambda-mu when they occur as terms in formulae. Classical AF_2 avoids this difficulty by distinguishing between the terms that may occur in the formula language and the terms that are the proofs of formulae. The restriction to intuitionistic terms in the former eliminates those terms which Parigot calls paradoxical, that is with inconsistent behaviours, though it does not ensure that all terms have canonical normal forms. The existence of output operators then allows us to extract the constructive meaning of these terms.

The logical framework brings with it some disadvantages, however. Firstly, we lose the identification of terms with proofs which is one of the cornerstones of the logical formalist justification of Martin-Löf's type theory. Secondly is the perhaps surprising observation that for AF_2, as for all logics of strength underneath the calculus of constructions, it is not possible to prove the principle of induction[34].

[30]First exposed in Phillippe de Groote's 'A CPS-translation of the lambda-mu calculus' [deGrooteP:cpstlm].

[31]See his 'Classical proofs as programs' [ParigotM:clapp].

[32]See his *Lambda-calculus, types and models* [KrivineJL:lamctm]. This calculus is based upon Daniel Leivant's ideas in 'Reasoning about functional programs and complexity classes associated with type disciplines' [LeivantD:reaafp].

[33]In *Proofs and Types* [GirardJY:prot].

[34]Thierry Coquand shows the unprovability of induction for the calculus of constructions to be a consequence of cut-elimination in 'Metamathematical investigations of a theory of constructions' [CoquandT:metitc], by observing that no normal form can inhabit the proposition expressing induc-

This is surprising since AF_2 is an extension of System F, which was formulated by Girard as a kind of functional interpretation[35] for the usual second-order logic: cut-elimination for System F is equivalent to the consistency of second-order arithmetic, and the interpretation of the principle of induction is just the type of the natural numbers[36].

However, whilst in a sense it is correct to say that system F captures the computational aspect of induction, the correspondence is not quite exact, and so one still needs to explicitly formulate the notion of recursion to fully capture the content of induction.

In consequence, it seems more promising to stick with Martin-Löf's approach to quantifiers and induction. However if we abandon Parigot's second measure, the first no longer gives us our guarantee of constructive content: to paradoxical terms we may find different output operators which yield several different numerals, and so we can no longer pretend that these operators are a kind of identity function. Furthermore, for an arbitrary predicate $\phi(n)$ (where n ranges over the natural numbers), we cannot in general associate to each discriminator $x \in \mathbb{N} \mapsto D(x) \in \mathbb{N}$ an associated function $D_\phi : \phi(n) \mapsto \phi(D(n))$, since such a pair of functions would allow us to directly give a constructive interpretation of the axiom of choice, contradicting the demonstration in section 4.1 that there can be no such interpretation.

Chapter three. Classical proofs I: propositions

We begin this section with a discussion of some arguments against the possibility of a successful logical formalist account of classical logic, which we argue are unconvincing because we may regard the strength of classical logic arising independently of our account of the rules associated with the individual connectives. We develop the claim that classical logic differs from intuitionistic logic due to the presence of stronger forms of structural rules in classical logic. These arguments are of crucial importance to the aim of this work: it is here that we introduce and argue for the idea that the classical logical rules should be seen as prior to the ordinary justification of logical connectives.

The task is to apply this insight to the development of a theory of classical logic to rival the account given in chapter one for intuitionistic logic. We argue that the additional term formers of lambda-mu fulfil the role of the right structural rules

tion. The same demonstration works in the presence of the lambda-mu calculus.

[35]Girard gives a summary of the interpretation in the last chapter of *Proofs and Types* [GirardJY:prot].

[36]In fact the paradox can be explained as follows: to define the provably recursive functions associated with a given proof theory one requires only an *iterator*, which can be expressed in Calculus of Constructions-like theories: one needs the stronger principle of induction only when one wants to show that these functions satisfy particular properties. See Splawski and Urzyczyn 'Type fixpoints: Iteration vs. Recursion' [UrzyczynP:typfiv] for an illuminating discussion; further to their discussion it is interesting to note that recursion can be defined in terms of iteration in the presence of the strong sum type former of Martin-Löf's type theory.

in the sequent calculus. We also show how it is possible to provide an intuitive foundation for these new term operators in terms of the notion of contraraiety.

Introducing these term operators disturbs the account of logical harmony we developed in chapter one. Calling derivations whose last rule is an introduction rule *natural* derivations, and other derivations *structural* derivations, we see that the *contraversion* operator provides a new *structural* introduction rule, and our account of the inversion rule must be extended to accommodate it. We provide this extension by means of a new clause 'contravertive decomposition' which must be shown at each type.

The new decomposition lemma allows us to see that the structural introductions are equivalent to natural derivations, this equivalence being captured by a form of conversion unique to the classical theory, which we call zeta conversion. Our treatment here differs from that of Luke Ong's formulation[37] in that we consider these zeta rules in the context of eta expansions, which allows us to avoid the need for the special conversions '$\bot 1 - \bot 3$'. We then prove the strong normalisation and Church–Rosser properties for this calculus.

To prove the principal path lemma we need to introduce a notion of path similar to that Prawitz introduced for his treatment of disjunction and existence, but in which segments need not consist of formulae directly above one another. With this notion we can prove the principal path lemma, and so prove that the subformula property for normal derivations, and consequently that the theory is consistent.

We obtain an interesting corollary to the subformula property[38], namely the identification of a logic that stands in the same relation to classical logic as minimal logic stands in relation to intuitionistic logic. We call this *classical minimal logic*, and we show two properties: firstly that it is equivalent in provability strength to the theory obtained by adding the axiom scheme $((A \supset B) \supset A) \supset A$ (Peirce's law) to minimal logic, and secondly that full classical logic is a conservative extension of classical minimal logic obtained by adding the axiom scheme $\bot \supset A$.

In the next section we consider the analogous system of recursion to that of section 1.5, the calculus $\mathrm{PRA}^{\omega}_{\mu}$. The principal result of this section is also the most substantial theorem of the thesis: the proof of strong normalisation for $\mathrm{PRA}^{\omega}_{\mu}$. The calculus cannot be coded into calculi such as Parigot's second order lambda-mu[39], due to the presence of what we call 'wild' conversions which may increase the number of occurrences of cut-redexes in the proof of any logical complexity. The proof presented here depends upon two main innovations, firstly the extension of the Tait reducibility predicate to a new notion, called part reducibility, which permits inductive reasoning over the complexity of open mu names, which is necessary to show that the set of strongly normalising terms is closed under the contra-

[37]See his 'A semantic view of classical proofs: type-theoretic, categorical, denotational characterizations' [OngCHL:semvcp].

[38]As an abuse of terminology, we shall use the term 'subformula property' to apply both to the property of derivations, and to the theorem that all normal forms have this property.

[39]Parigot proves strong normalisation for this calculus in 'Proofs of strong normalisation for second order classical natural deduction' [ParigotM:prosnf].

version term former. It is hoped that the technical device used may help to provide a serviceable notion of logical relation for the lambda-mu calculus.

The second innovation in the proof is the notion of full reducibility, which in turn depends upon an idea of closure under a special class of substitutions. We show that full reducibility is encompassed by a still stronger notion of reducibility, called ramified reducibility, which we can show is closed under structural induction.

In the last section we briefly examine a rather different basis for constructing a proof theory of the classical propositional calculus, which substitutes the idea of Gentzen's elimination rule by the idea of dual introduction rules which give the grounds for rejecting a proposition of a given type. We show the calculus to be in a pleasing correspondence with the account given earlier in the chapter, where the negative introduction rules are in a precise sense direct analogues of the elimination rules of natural deduction.

Although I develop this account only at an informal level, I hope that this account can be used to provide a new and arguably better justification of classical logic than so far presented. Michael Dummett argues[40] against attempts to develop multi-levelled theories of propositional meaning, where one has an account of negation in the meta-theory and one in the object theory, as violating Occam's principle. As my account makes use of just such a device, it may be desirable to reformulate the role of the meta-theory. In the novel formulation the semantic account of negation appears in a mutually recursive formulation, where the rejective rules attached to each type are determining the *consequences* associated with each connective.

The disadvantage of the novel formulation is that the logical rules are no longer the same as those of intuitionistic logic, and so the point about the separation of the logical and structural rules is better made with the lambda-mu style formulation. Further the account involves a major amendment to the structure of familiar derivations, and it is perhaps implausible to call it a natural deduction formulation.

Chapter four. Classical proofs II: arithmetic

The last chapter develops an extension of the theory described in chapter two to a *classical type theory* by employing the classical structural rules of chapter three.

We begin by presenting a 'naïve' extension of the theory ITT with a new class of structural rules as an exercise in demonstrating the pitfalls that attend an attempt to provide such a theory by ad-hoc methods. We give a short analysis of various approaches that fail.

There is little difficulty about the choice of language, it is the choice of rules that is troublesome. We need to formulate conversion rules which satisfy *local soundness*, that is, they must satisfy subject reduction, the inversion principle[41] and the

[40] In *Frege: philosophy of language* [DummettMAE:frepl].

[41] Recall that, unlike Prawitz, we give a local formulation of the inversion principle.

diamond property, and they must satisfy *global soundness* which here is the canonical form theorem. Perhaps surprisingly, strong normalisation presents no difficulties at all: all reasonable candidates for conversion rules can easily be coded into the system $\mathrm{PRA}^{\omega}_{\mu}$.

On the other hand, formulating rules which satisfy just local soundness is really quite difficult: we cannot admit the usual formulation of conversion for the natural numbers together with its zeta rules without sacrificing the diamond property, and so the identification of a *left recursive* formulation of the rules is one of the two key insights making the account we give possible.

The other insight is that it is adequate for our purposes to restrict zeta conversion for the wild term formers to distribute only through *vacuous* contraversions (that is, where the μ abstraction constructor does not bind any free variables). Without this restriction, it is impossible to show the subformula property holds for the wild zeta conversions, due to type dependency.

The main difficulty facing the proof of the canonical form theorem is the constructibility of what Parigot's paradoxical inhabitants at various types, such as

$$\mu\alpha^{\mathbb{N}}.[\alpha]\mathsf{succ}(\mu\beta^{\mathbb{N}}[\alpha]\underline{0})$$

which, in the right recursive form with zeta reduction, can be shown in some contexts to behave like $\underline{0}$ and in other contexts like a successor, leading to the derivability of $0 =_{\mathbb{N}} 1$ true.

Given the soundness proof it is not difficult to show that it satisfies many of the same properties as ITT, together with our prize, the validity of the principle of the excluded middle. Unfortunately it is not possible to provide a sound treatment of the strong sum type former, for reasons which we discuss in section 4.1, and so we lose the axiom of choice. The weak existential is available to us, through the definition $\exists x \in A.\phi(x) \mathrel{\hat{=}} \neg\forall x \in A.\neg\phi(x)$.

Our last result in this section is to show that we cannot give a direct set-theoretic semantics for this calculus, due to a simple counter-example showing that equality at higher types is not extensional. We end the chapter with a brief discussion of logical and constructive harmony, noting a significant caveat that follows from our modified treatment of reduction.

Reading this work

The main body of the work is organised into four chapters, the first two concerned with the intuitionistic propositional calculus and intuitionistic type theory respectively, and the last two concerned with their classical variants. We have developed the two accounts in parallel as far as possible to make it easier to see where the differences lie, a point I hope will be of especial convenience to the reader in the account of classical type theory. Also, it is not necessary to read chapter 2 before reading chapter 3.

I have also included a brief, critical summary of the approach to semantics of logic and arithmetic in the concluions, together with some discussion of possible further work.

I have referred to the formulae-as-types correspondence as the Curry–Howard correspondence only in the context of the intuitionistic case, and not in the context of Martin-Löf's type theory. The main intuitionistic systems of this thesis are: the calculus NJ of section 1.2, the simply-typed lambda calculus (which we refer to simply as λ) the system PRA$^{\omega}$(which is the same as Gödel's system T), and the theory ITT given in chapter 2, which is a subsystem of Martin-Löf's type theory adequate for intuitionistic arithmetic, and equivalent in expressivity to Heyting Arithmetic.

The other four main systems are the classical analogues of the above: the natural deduction calculus and the lambda-mu calculus (or $\lambda\mu$) of section CND of section 3.2, the system PRA$^{\omega}_{\mu}$ of section 3.4, and the system CTT of chapter 4. We summarise the theories lambda-mu and CTT in the appendix.

We provide an index of definitions also, in which the page references of the above can be found, under the heading 'calculus', together with the other systems mentioned in the thesis: PRA, HA, LK, $\lambda + \mathcal{A}$, $\lambda + \mathcal{P}$, $\lambda + \mathcal{C}$, $\lambda\mu^*$, $\lambda\mu^{**}$, PA and NTT.

I have tried to keep the thesis as self-contained as possible, so that the work will appeal to the widest audience. An elementary grasp of intuitionistic logic, basic proof theory and the simply-typed lambda calculus is assumed pretty much throughout, and chapters two and four assume familiarity with the use of dependent type theory. Familiarity with the contents of Girard, Lafont and Taylor's *Proofs and Types* would also be an asset.

Chapter 1

Logic and proof theory

1.1 Modern Logic

Logic is the study of arguments that are correct in virtue of their form. It is typically divided up into three parts: identifying argument forms as they occur in natural language, justifying the validity of these argument forms and providing the symbolic analysis of correct arguments.

When we say that an argument is valid due to its form, we mean that it conforms to a scheme of correct arguments with many instances. For example the two arguments:

> Since all mothers are female and some horses are mothers, then some horses are female.

and

> Since all logicians are clever and some people are logicians, then some people are clever.

share the same form, which is given by the syllogism Darii in Aristotle's *Prior Analytics*. The logical form of the argument, then, is what essentially equivalent arguments have in common, and the analysis of this form allows us to present the argument in a symbolic scheme, with symbols replacing the interchangeable parts of the argument.[1]

For example we might render the logical form of the above two arguments in the scheme:

$$\frac{\text{Some As are B} \qquad \text{All Bs are C}}{\text{Some As are C}}$$

[1]This is not to say that it is always easy to determine the logical form of such an argument. For example, in the following argument:

> Since my brother is married, I have a sister–in–law.

the form of the argument is not explicit, and requires elucidation of the meanings of the words to make clear. Such analysis is, perhaps misleadingly, called informal logic.

Aristotelian logic traditionally placed little stress upon the distinction between utterance and logical form, a neglect no doubt due the conceptual simplicity of his scheme. All the propositions of his system are one of four categorical judgements relating a subject to a predicate, and all of the arguments are composed of 14 syllogisms[2] which take two judgements as premisses and have a single conclusion. Though it had its challengers, it is fair to say that it was clearly superior to its rivals for over two millennia.

However Aristotle's system was far from complete. For example the following argument:

> (Syl-C) If there is some cat whom all mice fear, then each mouse is afraid of some cat.

cannot be derived in Aristotle's system: although both premiss and conclusion take the form of a categorical judgement, the predicate forms cannot be formulated in a manner that expresses their intuitive relationship. The Scholastics called this the problem of *muliple generality*.

It was not until Frege wrote his *Begriffsschrift* that the notion of logical form received a treatment that could capture such complex schemes. Frege's revolutionary achievement was founded upon two principal formal innovations:

1. Frege placed the central emphasis of logic upon the declarative sentence, the class of expressions of language that we may take to be true or false. In so doing he overturned the traditional perspective based upon the Aristotelian analysis of subject and predicate.

   One of the distinctive features of this arrangement is that predicates are not considered to be first-class bearers of meaning, but are regarded as *incomplete expressions*, which become complete when their gaps are filled by object denoting expressions or *proper names*. The important feature of this alternative arrangement is that there are many incomplete expressions occurring within any given sentence, and this allows us to analyse the formal content of a proposition in many distinct ways. For example 'Zero is less than succ(zero)' may be decomposed into, amongst others:

   - ϕ(zero) where $\phi(x)$ is 'x is less than succ(zero)';
   - ϕ(succ(zero)) where $\phi(x)$ is 'zero is less than x';
   - ϕ(zero) where $\phi(x)$ is 'x is less than succ(x)'.

2. He gave a novel treatment of categoricals where 'All' and 'Some' are formalised by quantifiers, constructs taking an incomplete expression (where we indicate the gap with a free variable) and yielding a complete expression.

[2] In fact, most expositions of Aristotle's system give 24 syllogisms, though the 24 syllogism theory allows the same inferences as the original 14 syllogism theory.

This introduces the ability to define the scope of categorical judgements, allowing us to represent (Syl-C) as:

$$(\exists x.\text{Cat}(x) \wedge (\forall y.\text{Mouse}(y) \supset y \text{ fears } x)) \\ \supset (\forall y.\text{Mouse}(y) \supset (\exists x.\text{Cat}(x) \wedge y \text{ fears } x))$$

which is a theorem of Frege's calculus[3].

But though Frege's system constitutes a dramatic step forward in the expressive power of logic, it is attended by new difficulties. Whilst Aristotle's syllogistic calculus was grounded in his theory of categories, making it clear what it is that makes valid judgements true, Frege's more sophisticated judgements would need a new system of justification.

Frege explained the validity of his *Begriffsschrift* in terms of his profoundly influential theory of concept and object, articulating the meaning of each expression in terms of sense and denotation, and determining the denotations of complex expressions by evaluating their parts according to their function–argument structure. His account depended, however, upon a strong form of conceptual realism, leading Frege to found his theory of the *Grundgesetze* upon a principle of unrestricted comprehension, which made Russell's famous paradox possible.

Evaluative approaches to semantics were not abandoned by mathematical logicians however: the most widely accepted account of the meaning of logic is due to Alfred Tarski, where the denotation of a sentence is given by its truth condition, and its truth-value determined by means of the *material adequacy* condition.

All such truth-conditional accounts are open to a deep criticism: they suppose that predicates precisely divide their domain into those which are true and those which are false. As Michael Dummett has convincingly argued, such *determinacy of sense* cannot be made explicit in our theories of meaning for those expressions. Such theories of meaning, he argues, make significant metaphysical assumptions.

We are faced with a problem: if we wish to eliminate metaphysical controversy from our theory of logic, how do we know which are the correct laws of logic? Whilst a system such as Tarski's can be used to provide a justification, if we have doubts about the theory, it follows that we will have doubts about the justification. We may hope that we are able to provide a justification which does not depend upon such contentious assumptions.

Logical formalism

If we abandon truth-conditions as the criteria of logical truth, then how are we to distinguish valid arguments from fallacies? What is to be the intuitive foundation of our conception of logic?

We may observe that we are furnished with explicit evidence of the meaning of logical truth in the rules that govern the practice of logical inference. Perhaps it is the practice of inference which determines the content of logical truth.

[3] In fact Frege only gives rules for universal quantification, and defines $\exists x.\phi$ by $\neg\forall x.\neg\phi$.

Such a claim would be a reversal of the usual priority existing between provability and truth: instead of the notion of truth and the meaning of the logical constants being used to justify the practices of inference, we have the notion of truth and the meaning of propositions being given by deductive practice.

This position was argued for by Gerhard Gentzen in his paper 'Investigations into Logical Deduction' [GentzenG:invld]. In his formulation of logic, Gentzen separates the rules that concern each logical connective into those which determine the valid grounds which we may have for asserting a proposition governed by a logical connective, and those which permit us to draw conclusions from a hypothesis involving that connective. In a natural deduction setting these rules are called respectively the introduction and elimination rules, whilst in the sequent calculus they are the right and left rules. Gentzen's investigations were concerned with the possibility of 'reading off' a meaning from the form of the rules; such a position might be called *logical formalism*[4].

This appears to raise the possibility that what governs logical truth is simply the conventional acceptance of a collection of rules. But, as Arthur Prior noted[5], there must be more to deductive practice than mere conventional acceptance, because such a position could give no compelling grounds for the rejection of any particular set of inference rules. In particular, we could have a binary connective *tonk* for which 'A tonk B' could be asserted whenever 'A or B' can (that is, it shares the same introduction rules as disjunction), and whose formal consequences are the same as those of 'A and B' (it shares the same elimination rules as conjunction).

But in such a system the following derivation is valid:

$$\dfrac{\dfrac{A}{A \text{ tonk } B}}{B}$$

We cannot accept this because it implies that one proposition is true whenever any other is, a conclusion that clearly defies any reasonable conception of logical validity. Therefore there must be more to the meaning of logical truth in general, and the meaning of the logical connectives in particular, than the description of the logical rules that govern them.

Nuel Belnap[6] resisted the conclusion which Prior arrived at, that his example had refuted the possibility of a project like Gentzen's succeeding on the basis of the form of the rules alone. Instead, he argued, what was wrong with Prior's connective 'tonk' was that a natural harmony that ought to exist between the introduction and elimination rules was absent, because it is possible to draw conclusions from a proof involving the connective 'tonk' that were unrelated to the premisses.

[4]Michael Dummett described Wittgenstein's position as 'logical formalism' in *The Logical Basis of Metaphysics* [DummettMAE:logbm], but the context suggests that he does not intend the phrase in the wider sense to which I apply it in this work. It is not to be confused with the kind of formalism that Brouwer accused Hilbert of.

[5]In 'The runabout inference ticket', [PriorA:runit].

[6]See his paper 'Tonk, Plonk and Plink' [BelnapND:tonpp].

Belnap further argued that it is possible to set down formal requirements which ensure the immunity of the logical formalist project from objections like Prior's. His first requirement, *conservativity* is that the addition of the new connective and its attendant rules should not affect the truth status of statements which do not involve the connective. The second requirement, which we call *logical harmony*, is that there should be a duality between the rules permitting us to make assertions of statements governed by a connective and the rules which permit us to draw conclusions from such statements, ensuring that the two kinds of rule match each other in strength. Finally the logical theory should not support distinctions between the meaning of connectives if they share the same proof conditions, this is his requirement of *uniqueness*[7].

Let us examine more closely this duality between the introduction and elimination rules required by logical harmony. Consider the introduction and elimination rules governing conjunction:

$$\frac{A \quad B}{A \wedge B} \wedge\mathcal{I} \qquad \frac{A \wedge B}{A} \wedge\mathcal{E}1 \qquad \frac{A \wedge B}{B} \wedge\mathcal{E}2$$

The requirement of logical harmony can be divided into two complementary properties: we say that *analytic harmony* prevails in the theory if the elimination rules are at least as strong as the introduction rules, and that *synthetic harmony* prevails if the introduction rules are at least as strong as the elimination rules. Our work is to articulate this sense, and for conjunction, these requirements admit a particularly simple formulation. If analytic harmony does not hold then it is possible to draw conclusions not governed by a connective in the presence of the rules that cannot be drawn in their absence, such as occurs with Prior's tonk. If the synthetic harmony does not hold then the conclusions we may validly draw from a proposition may be affected by the presence of other, well-behaved connectives[8].

We show that analytic harmony is satisfied by observing that each elimination rule has as its conclusion a formula that forms the content of one of the premisses to the introduction rule, and so the consequences of the elimination rule are contained in the premisses of the introduction rule. Synthetic harmony follows from observing that every premiss of the introduction rule is the conclusion to some elimination rule, and so all of the grounds for asserting $A \wedge B$ proposition are recoverable from the possible consequences of holding it.

In general though, the logical rules governing any particular connective may not admit such a simple test of acceptability, as they may involve the discharge of hypotheses, or they may involve formulae that are not directly related to the formula being introduced or eliminated. In these cases, the formalisation of the duality requirements will have to look at more of the derivations that may justify the premisses than is necessary in the case of conjunction. We shall re-examine these more complex formulations when we come to consider Prawitz's inversion principle later in the section.

[7]The uniqueness requirement will not be mentioned again until the conclusions.

[8]We shall discuss this kind of failure in the last section of chapter two.

1.2 Natural Deduction

To proceed with our investigation, we shall introduce a formalisation of a simple theory of logic, corresponding to the minimal logic of Prawitz. Before we examine the particular rules of this logical system, we will introduce some general terminology for talking about proofs.

The syntax of natural deduction

A natural deduction system is a proof calculus in which we do not introduce logical tautologies as axioms, as in Hilbert's system, but instead we deduce theorems by applying rules of inference related to the logical connectives. To begin a derivation we may introduce any formula as a hypothesis in an assertion. Since all of the rules have a single conclusion, derivations in a natural deduction system may take a tree like form where there is a single conclusion at the root of the tree.

Before we examine the particular formulation that we will study in the remainder of this section, let us provide in general terms a description of natural deduction. There are four kinds of entity that make up a natural deduction proof: the formulae, the rules of inference, the derivations, and the judgements of the logical calculus. The proof calculi that we will examine in the remainder of this work will form refinements and extensions of these basic concepts.

Formulae

The formulae of a natural deduction system are defined schematically, and the possible formulae are determined by the connectives of the logical system. Each connective is associated with an *arity* which determines how it is used to generate new formulae.

A *formula* is either a *schematic letter*, which is displayed as A, B, C or A′ or A_i and so on, or it is built out of an n-ary connective $\otimes$, and n formulae $A_1, \ldots, A_n$ to give the new formula $\otimes(A_1, \ldots, A_n)$. The collection of formulae is generated inductively by these two rules. If a formulae is not a schematic letter, but is built from other formulae and the connective $\otimes$, then we say that it is *governed by* $\otimes$. We may supplement the calculus with symbols to denote particular propositions: we shall call these *primitive letters*, and the collection of schematic and primitive letters are the *atomic formulae*.

The *degree* of a formula ϕ is given by the function:

$$\mathsf{deg}(\phi) \mathrel{\hat{=}} \begin{cases} 1, & \text{if } \phi \text{ is atomic} \\ 1 + \mathsf{max}\{\mathsf{deg}(A_i) \mid 1 \leq i \leq n\}, & \text{if } \phi \equiv \otimes(A_1, \ldots, A_n) \end{cases}$$

The *subformulae* of a formula ϕ is defined to be a set of formulae determined

recursively as follows:

$$\mathsf{sf}(\phi) \mathrel{\hat{=}} \begin{cases} \{\phi\}, & \text{if } \phi \text{ is atomic} \\ \{\phi\} \cup \bigcup\{\mathsf{sf}(A_i) \mid 1 \leq i \leq n\}, & \text{if } \phi \equiv \otimes(A_1, \ldots, A_n) \end{cases}$$

The *eigenformulae* of a formula are the schematic letters occurring amongst its subformulae. If the schematic letters $X_1, \ldots, X_n$ are eigenformulae of the formula A, then we say that the result of systematically replacing the occurrences in A of formulae X_i by new formulae B_i is an *instance of* A. Observe that any instance of a formula governed by $\otimes$ is itself governed by $\otimes$.

Rules

Each connective $\otimes$ is associated with a collection of rules, which we call the *rules governing* $\otimes$. These rules are divided into introduction rules and elimination rules, and are defined schematically.

An *inference rule* consists of a single conclusion and zero or more premisses, the number of premisses being the *arity* of the rule. The rule is displayed by a horizontal line with the premisses occurring above the line, the conclusion occurring beneath, and we may indicate the name of the rule to the right of the line.

A conclusion always consists simply of a formula, whilst premisses may be of one of two forms. A *simple premiss* is a premiss which consists simply of a formula, called the *required formula* whilst a *hypothetical premiss* is one in which two formulae occur, the required formula and the *discharged formula*. A hypothetical premiss is usually displayed in the definition of a rule as follows:

$$\begin{array}{c} A \\ \vdots \\ B \end{array}$$

where A is the discharged formula and B is the required formula.

The *eigenformulae* of a rule are the collection of the eigenformulae occurring in its conclusion and premisses, with the eigenformulae of each premiss being the union of the eigenformula of the one or two constituent formulae. The *instances of the rule* are defined analogously. For example, in the $\supset\mathcal{I}$ rule:

$$\frac{\begin{array}{c} X \\ \vdots \\ Y \end{array}}{X \supset Y}$$

there are two eigenformulae, X and Y, which are the union of the eigenformulae of the three formulae in the rule. By replacing X by $U \supset V$, we obtain a rule instance:

$$\frac{\begin{array}{c} U \supset V \\ \vdots \\ Y \end{array}}{(U \supset V) \supset Y}$$

with three eigenformulae, Y, U and V.

All logical rules have a special identified premiss, the *principal premiss* (though in the case of introduction rules there may be several principal premisses).

For the introduction rules governed by $\otimes$, the formula appearing at the conclusion is governed by $\otimes$, and all of the premisses are principal premisses.

For the elimination rules governed by $\otimes$, the principal premiss is the left-most premiss, which will be governed by $\otimes$. The other premisses are called auxiliary premisses. If the eigenformulae of the whole rule are the eigenformulae of the principal premiss, then we call it a *direct elimination rule*, otherwise we call it an *indirect elimination rule*.

Derivations

A derivation is built up recursively from assertions by applying inference rules. Each derivation is associated with an identified formula called the *conclusion*, and a set of formulae called the *open assumptions*. Much of the complexity of what follows is due to the fact that we use a sophisticated discharge convention involving assumption packets. Our reason for not using the more usual crude discharge convention will become clear in the next section.

A *derivation* consists of either:

1. An assertion of a formula, which consists of the formula asserted and a label indicating the *assumption packet* to which it belongs, where each assumption packet is associated with the unique formula which it may label. The conclusion of the assertion is A and the open assumptions are $\{A\}$. The derivation is displayed as the single formula A with the label of the assumption packet indicated as a superscript;

2. Or it is an application of a rule $\mathcal{R}$. An application of $\mathcal{R}$, where $\mathcal{R}$ is of arity n consists of an instance of $\mathcal{R}$ together with n admissible derivations $d_1, \ldots, d_n$, the *direct subderivations*, where derivation d_i is *admissible* if its conclusion is the same as the required formula of the i-th premiss of the rule instance.

   The conclusion of the resulting derivation is the conclusion of the rule instance, whilst the open assumptions are the union of the contributed assumptions from each of the direct subderivations, where the *contributed assumptions* of the i-th direct subderivation are the open assumptions of that subderivation if the i-th premiss is simple. If the i-th premiss is hypothetical, with discharge formula A, then we must choose an assumption packet x associated with A. If all of the formula occurrences of A which are not discharged are labelled x, then the contributed assumptions are the open assumptions less A, otherwise they are the all of the open assumptions. All of the assertions labelled x are discharged.

   The derivation is displayed as a rule labelled with rule $\mathcal{R}$ and the labels of all of the assumption packets discharged to the right, and the conclusion be-

neath the line. The direct subderivations are written above the line with their conclusions where the premisses would appear in the ordinary rule. We write square brackets about any assertions that are discharged by the rule.

The collection of derivations is generated recursively by the application of the above two cases. If the derivation is obtained by an application of a rule R, then we say that it is *governed by R*, and that R is the *last rule* of the derivation. In practice we usually omit the name of the rule from derivations. In the case of an indirect rule, the auxiliary premiss is called the *cut formula* of the rule.

We define the *subderivation*s recursively to be the singleton set consisting of the derivation itself if the derivation is an assertion, or the union of this set with the subderivations of the direct subderivations if the derivation if the result of a rule application. The *formula occurrences* of a derivation to be the set of conclusions of its subderivations. The *leaves* of a derivation are the subset of formulae occurrences which are the conclusion of an assertion or a 0-ary rule application.

The *height* of a derivation d is given by the function:

$$\mathsf{ht}(d) \mathrel{\hat{=}} \begin{cases} 1, & \text{if } d \text{ arises by assertion} \\ 1 + \mathsf{max}\{\mathsf{ht}(d_i) \mid 1 \leq i \leq n\}, & \text{if } d \text{ arises by an } n\text{-ary rule and has} \\ & \text{immediate subderivations } d_1, \ldots, d_n \end{cases}$$

We talk about the collection of derivations relative to a logical system. This means that all subformulae of the formula occurrences in the derivation are either atomic, or are governed by one of the connectives in the logical system. Thus all of the rules that occur in the derivation are rules that govern these occurrences.

A *stands immediately above* B in d if there is some subderivation of d with conclusion B and one of whose direct subderivations has conclusion A. If A stands immediately above B, and the rule of which B is the conclusion and A the premiss is $\mathcal{R}$, then we say that $\mathcal{R}$ *relates A to B*.

A *thread* of d is a sequence of formulae $A_1, \ldots, A_n$ which are all formulae occurrences of d, and where each A_i stands immediately above A_{i+1}. A thread is an *initial thread* if its first formula is a leaf and a *terminal thread* if its last formula is a conclusion, whilst it is a *spanning thread* if it is both initial and terminal.

Observe that all of the formulae occurrences in d which arise as the result of an assertion are leaves. Each such leaf A is either an open assumption of d, or it is discharged by the last rule of some subderivation d' of d. We say that d' is the *discharge*subderivation associated with the label of A.

Judgements

A *sequent* is a pair consisting of a set of formulae Γ and a single formula A. A sequent is displayed $\Gamma \vdash A$, and we call A the conclusion of the sequent and Γ the premisses of the sequent. If we wish to indicate that we are considering some particular logical system Λ, then we indicate so by adding a subscript to the turnstile so: $\Gamma \vdash_\Lambda A$.

Let d be a derivation with conclusion A and open assumptions Γ in the logical system Λ. Then we say that for any set of formulae Γ' with $\Gamma \subseteq \Gamma'$ that d *justifies* the sequent $\Gamma' \vdash_\Lambda A$. Conversely we say that $\Gamma \vdash_\Lambda A$ is *derivable* if in the logical system Λ there is such a derivation whose open assumptions occur in Γ.

The inversion principle

The propositional calculus of minimal logic, NJ which we shall be principally concerned with in this chapter is the system involving only the connectives of implication and conjunction, in order to avoid the syntactic complexities associated with disjunction. We shall discuss disjunction and its attendant complications in the conclusion.

Our first step is to examine the outstanding formal requirements that the rules governing each connective must satisfy in order to be considered acceptable in a natural deduction calculus. These requirements correspond to the three conditions advanced by Belnap as adequacy conditions for systems of logical reasoning, and they constitute the formal grounds for claiming that a formal system satisfies Belnap's criteria.

The first requirement is the duality condition that must hold between introduction and elimination rules. As noted earlier, it is quite easy to formalise this requirement for conjunction as a direct relationship that must exist between the premisses of the introduction rule and the conclusions of the elimination rules. However, the existence of discharge formulae complicates the matter for the other logical rules, and we require a more sophisticated condition for these connectives.

First, let us consider the case of implication:

$$\dfrac{\begin{array}{c}[A]^x \\ \vdots \\ B\end{array}}{A \supset B}\,(x)\supset\mathcal{I} \qquad \dfrac{A \supset B \quad A}{B}\supset\mathcal{E}$$

We can see that the required formula of the premiss of the introduction rule matches the conclusion of the elimination rule, and that the discharged formula of the premiss of the introduction rule matches the subsidiary premiss of the elimination rule. But how are we to take these observations to a demonstration of duality?

In the case of disjunction, the problem appears even more acute. Gentzen gave the introduction rules:

$$\dfrac{A}{A \vee B}\vee\mathcal{I}1 \qquad \dfrac{B}{A \vee B}\vee\mathcal{I}2$$

and the elimination rule:

$$\dfrac{A \vee B \quad \begin{array}{c}[A]^x \\ \vdots \\ C\end{array} \quad \begin{array}{c}[B]^y \\ \vdots \\ C\end{array}}{C}\,(x)(y)\vee\mathcal{E}$$

and at an intuitive level it is hard to see how to take the match between the formula to a compelling account of duality. How is the role of the formula C to be explained? How do we explain the relationship between the principal premiss of the elimination rule and the discharged formulae of the auxiliary premisses?

The solution was provided by Dag Prawitz in his seminal work *Natural Deduction* [PrawitzD:natd]: we see the introduction rules governing each connective as determining a natural shape of a derivation whose conclusion is governed by that connective. We then say that any derivation of the connective implicitly contains a derivation of this natural shape; we can tabulate this account so:

> A derivation of $A \wedge B$ implicitly contains a derivation of A and a derivation of B;
> A derivation of $A \supset B$ implicitly contains a derivation of B under assumption A;
> A derivation of $A \vee B$ either implicitly contains a derivation of A or it implicitly contains a derivation of B.

As it stands, this relation of 'implicit containment' is a little unclear. We may distinguish between two cases: we say that d implicitly contains d_0 in a local sense if d_0 is a subderivation of d, whilst if we justify the claim by reference to a more indirect property we say that the containment is in a global sense.

In the case of a derivation whose last rule is an introduction rule we see that above properties hold in a local sense, since the immediate subderivations satisfy the condition.

It is the task of our semantics on the logical formalist account to provide an explicit justification for the proof theory, which we provide in two stages. Firstly, we justify the claim that logical harmony exists between an introduction rule and an elimination rule with an appeal to the local containment. Secondly, we show how it is possible to apply this duality to obtain from any given derivation governed by a logical rule, a derivation justifying the same judgement whose form clearly satisfies the above intuitive scheme, a scheme sufficient to demonstrate conservativity.

The first part is demonstrated by appeal to Prawitz's inversion principle. In its original formulation it only shows that the elimination rules are no stronger than the introduction rules (which we call the eliminative part), the method of the converse direction (which we call the decompositional part) is introduced in a later paper of Prawitz's[9], although he does not remark upon its significance there. Also the property as formulated by Prawitz depends upon the global form of containment; it may be significant that the local form is sufficient.

DEFINITION 1 (INVERSION PRINCIPLE)

1. (Eliminative part) Let d be an instance of a derivation whose conclusion rule is the elimination rule governing the connective $\otimes$, and whose principal premiss is the conclusion of an introduction rule. Then it is possible to assemble

[9] 'Ideas and Results in Proof Theory' [PrawitzD:iderpt]

from the immediate subderivations of the introduction rule, and from the auxiliary subderivations, a derivation justifying the conclusion without introducing any new rules.

2. (Decomposition part) Let A be a formula governed by $\otimes$. Then there is a derivation justifying the judgement $A \vdash A$ which is of the natural form associated with $\otimes$.

To show that all of the connectives satisfy this inversion principle, it is first necessary to provide a lemma:

LEMMA 2 (SUBSTITUTION LEMMA)
Let d justify $\Gamma \vdash A$ and let d' justify $\Gamma, A \vdash B$. Then there is a derivation $d'[A \backslash d]$ which justifies $\Gamma \vdash B$

PROOF If $A \in \Gamma$, then the conclusion is immediately true for $d'[A \backslash d] \hat{=} d'$. Otherwise, let $\vec{x}$ be the assumption packets of d' associated with with the formula A. Since the derivations are generated freely from the rules of the system, the derivation obtained from d' by replacing each assertion A^{x_i} by the derivation d is a derivation of d'. It is clear that its conclusions B, and that all of its open assumptions must be in Γ. $\square$

Let A be a formula occurrence of a derivation d. We say that the *derivation above* A in d is the subderivation of d whose conclusion is the formula occurrence A. We also say that the *derivation below* A in d is the derivation obtained from d by replacing the subderivation above A with a fresh assumption of A.

We are now in a position to show that the rules satisfy the inversion principle:

$\wedge$ 1. (Elimination) Let d be a derivation whose last rule is a conjunctive elimination rule and which satisfies the form required by the inversion principle. Then there are derivations d_0, d_1 such that either

$$d \text{ is } \dfrac{\dfrac{\begin{matrix}\vdots\, d_0 \\ A\end{matrix} \quad \begin{matrix}\vdots\, d_1 \\ B\end{matrix}}{A \wedge B}}{A} \quad \text{or } d \text{ is } \dfrac{\dfrac{\begin{matrix}\vdots\, d_0 \\ A\end{matrix} \quad \begin{matrix}\vdots\, d_1 \\ B\end{matrix}}{A \wedge B}}{B}$$

In the first case, the required derivation id d_0, in the second it is d_1.

2. (Decomposition) $\dfrac{\dfrac{A \wedge B^x}{A} \quad \dfrac{A \wedge B^x}{B}}{A \wedge B}$ justifies $A \wedge B \vdash A \wedge B$.

$\supset$ 1. (Elimination) Let d be a derivation whose last rule is $\supset \mathcal{E}$ and is of the form

required by the inversion principle. Then d is

$$\dfrac{\dfrac{\begin{array}{c}[A]^x\\ \vdots\, d'\\ B\end{array}}{A \supset B}\,(x) \quad \begin{array}{c}\vdots\, d''\\ A\end{array}}{B}$$

and so the required derivation is $d'[A \backslash d'']$.

2. (Decomposition) $\dfrac{\dfrac{A \supset B^x \quad [A]^y}{B}}{A \supset B}\,(y)$ justifies $A \supset B \vdash A \supset B$.

We can form new derivations from old by means of the figures associated with the elimination and decomposition parts of the inversion principle. Such new derivations are called rewrites of the old derivation, and they are held in some sense to be contained within the old derivation. We shall first define the rewrites associated with eliminations first as they are of a simpler form than the decompositions.

DEFINITION 3

1. A *maxima* in a derivation is a formula occurrence of that derivation which is both the conclusion of an introduction rule and the principal premiss of an elimination rule.

2. A *maxima elimination* is a rewrite of a derivation. If A is a maxima of d with principal connective $\otimes$, and B is the formula occurrence immediately beneath A, then the rewritten derivation is obtained from d by replacing the subderivation above B with its simplified counterpart obtained from the justification of the eliminative part for $\otimes$.

3. Let d' arise from d by the application of a rewrite. We say that d is the *antecedent derivation* with respect to the rewrite and that d' is the *residual derivation*. From the definition of maxima elimination for $\wedge$ it is possible to associate each formula occurrence of the antecedent derivation with each formula occurrence of the residual derivation, which we call its *antecedent formula occurrence*. We want this association to be uniquely determined, so in the case of an elimination of a maxima $A \supset B$, where the substituted formula B could be associated with several different formula occurrences, we identify the auxiliary premiss of the elimination rule to be the antecedent formula occurrence.

4. Let d admit two maxima eliminations, at the maxima A and A'. If in the derivation obtained by eliminating the maxima A the maxima A' is not the antecedent formula occurrence of any maxima in the residual derivation, or vice versa for A when A' is eliminated, then we say that A and A' are *overlapping redexes*.

We sometimes call formulae removed by a maxima eliminations the *detours* of the antecedent derivation.

The decomposition part shows how we may guarantee that the natural form of derivations governed by a connective can always be obtained. If d has a conclusion A governed by a connective $\otimes$, then by applying the substitution lemma to d and the derivation given in the justification of the decomposition part of the inversion principle, we obtain a rule justifying the same judgement whose last rule is an introduction rule governed by $\otimes$.

The first of Belnap's requirements, conservativity, is not guaranteed by the inversion principle alone. To see this, observe that it is possible to conceive that there might be a derivation d of $\vdash A \wedge B$ with the property that the last rule of d is an introduction, but there is no derivation d_0 justifying $\vdash A$ for which $A \wedge B$ is not a formula occurrence of d_0. This is to say that while the purely local justification suffices for the second requirement, we need more to justify his first requirement of conservativity.

To obtain this, what we additionally wish to know is that there is an ideal derivation whose last rule is an introduction rule, and that its immediate subderivations are of a logically simpler form, providing the global sense to the relation of containment.

1.3 Normalisation

Such a definition of an ideal property and a method of obtaining one is provided by Gentzen's normalisation theorem. The normalisation theorem states that we may associate with each derivation a derivation justifying the same conclusion, and with derivation wide structural properties. From this, conservativity follows as an easy corollary.

DEFINITION 4 A derivation is a *simple normal form* if it has no maxima.

To show that to each derivation we may associate one in normal form, we prove that any sequence of maxima eliminations applied to a derivation must be finite. Prawitz proved this theorem by associating each derivation with an ordinal under ω^ω, and then finding a reduction strategy which associates every term which is not a normal form with a term with a strictly lower measure.

Prawitz's proof is perfectly satisfactory, however it is desirable to prove a stronger result, namely that the result of *any* maxima elimination is a derivation that is in some sense simpler. The former property is called *weak normalisation*, the strategy-independent property is called *strong normalisation*.

There are two main means by which strong normalisation is proven in the literature. Firstly, it is possible to use Tait's method, based upon reducibility predicates, and a version of this proof is outlined in section 1.5. Or secondly, there is Gandy's method[10], where we provide an internal 'coding' of proofs into other proofs which

[10] In 'Proof of strong normalisation' [GandyRO:prosn].

forces all maxima of the original to be reduced, and so allows us to deduce strong normalisation from weak normalisation.

Unfortunately both of these proofs have drawbacks. The reducibility predicate for implication is defined by means of an impredicative quantification, and so the proof is not finitistically acceptable, which is undesirable given our goal of justifying logic without major metaphysical assumptions. Gandy's method *is* finitistically acceptable; however it is also very indirect, and it does not cast much light on the nature of the normalisation process.

We shall here go through an original alternative, an extension of Prawitz's method by an additional measure on redexes, that guarantees that the overall measure on derivations reduces at each step. For the sake of motivation, we shall first examine two complicating matters.

Firstly the application of a maxima elimination may result in the creation of new maxima that did not occur in the antecedent derivation. For example, consider the following derivation of $A, B, A \supset C, B \supset D \vdash D$:

$$\dfrac{\dfrac{\dfrac{\dfrac{A \supset C^c \quad \dfrac{[A \wedge B]^x}{A}}{C} \quad \dfrac{B \supset D^d \quad \dfrac{[A \wedge B]^x}{B}}{D}}{C \wedge D}}{A \wedge B \supset C \wedge D}\,(x) \quad \dfrac{A^a \quad B^b}{A \wedge B}}{C \wedge D}}{D}$$

There is only one maxima in this derivation, an occurrence of $A \wedge B \supset C \wedge D$, but applying the maxima elimination creates three new maxima, two occurrences of $A \wedge B$ and one of $C \wedge D$.

Secondly, applying a maxima elimination to a maxima governed by $\supset$ may cause a single antecedent maxima to have several residuals, if the assumption packet bound by the introduction rule contains several occurrences, and the maxima occurs in the substituted derivation.

Each of these kinds of difficulty may be avoided by finding a measure of the complexity of a derivation which is guaranteed to reduce under any maxima elimination. We provide such a measure, which we call the complexity of a formulae occurrence, and obtain the measure on the whole derivation as the sum of the complexity of the maxima of that derivation.

The first complication is easy to sort out. Inspection of the maxima eliminations reveals that its application can only create new maxima of degree strictly less than that which was eliminated. The second complication is much harder to resolve, as reflection makes clear that the measure must depend upon the context in which the formula occurs. Indeed Prawitz avoided this complication by considering only those $\supset$ maxima eliminations where all of the maxima in the subsidiary

subderivation of the $\supset\mathcal{E}$ rule are of lesser degree than the eliminated maxima. We shall, however, press ahead to find a satisfactory measure that reduces under all eliminations.

The key change in the situation of a formula occurrence which is in the subsidiary subderivation of a $\supset$ maxima elimination is that after substitution it is no longer above that $\supset\mathcal{E}$ rule. However, we cannot obtain our measure by simply counting the number of $\supset\mathcal{E}$ rules which occur beneath the formula, as there may be several such rules occurring in the derivation above the $\supset\mathcal{I}$ rule. Before we consider how to handle this difficulty, let us make a few useful definitions.

DEFINITION 5

1. A *junction* is an elimination rule of a connective for which the maxima elimination involves a substitution. The *major premisses* of a junction are those premisses that do not occur as the conclusion of the substituted subderivation, and the *minor premisses* are those that do. In the current context a junction is an $\supset\mathcal{E}$ rule, whose major premiss is the principal premiss, and whose minor premiss is the subsidiary premiss.

2. A *branch* is a thread in a derivation whose first formula is a leaf, whose last formula is either the conclusion of the derivation, or is a minor premiss of a junction, and in which no formula is the conclusion of a junction whose predecessor is the minor premiss. A *main branch* is a branch whose last formula is the concluding formula of the derivation, and a *minor branch* is one whose last formula is the minor premiss of a junction.

3. The *order* of a branch is the number of minor premisses to junctions which occur in the terminal thread starting at the last formula of the branch. The order of a formula occurrence is the order of the branch upon which it occurs. The order of a derivation is the maximum of all the orders of its branches.

By considering the nature of the terminal thread beneath each formula occurrence, it should be clear that each leaf of a derivation determines a unique branch, and that every branch and formula occurrence is associated with a unique order.

DEFINITION 6 An *assumption binding* is a derivation whose last rule binds an assumption packet. In this proof system the only such rule is $\supset\mathcal{I}$.

Consider the general form of the $\supset$ maxima:

$$\dfrac{\dfrac{\begin{array}{c}[A]^x\\ \vdots\, d_0\\ B\end{array}}{A\supset B}\,(x) \qquad \begin{array}{c}\vdots\, d_1\\ A\end{array}}{B}$$

Our strategy is to formulate a measure dually upon both assumption bindings and the junction, which we shall call its depth, such that the depth of any assumption binding is less than that of the junction, and where the depth of the assumption binding records the maximal complexity which arises in the bound assumptions A^x due to their relative position within the assumption binding.

There are two chief difficulties facing an attempt to give a definition of depth. The first is that the parts of the definition are mutually dependent. The depth of an assumption binding depends upon the junctions that occur within it, whilst the depth of a junction depends upon all the assumption bindings whose conclusion might appear as the principal premiss. The second difficulty is that new assumption bindings may appear above a junction as a result of substitutive reductions which occur beneath it, and so the depth of the junction must depend upon both the derivation above the main premiss and beneath the conclusion. Thus the depth of a junction cannot be defined by a simple structural induction.

As a result, the definition of depth and the proof of its soundness are rather intricate, and so the reader is advised to take care in following the argument.

Our first step is to distinguish between a measure applied to a subderivation in context and the same measure applied to a subderivation considered separately as a derivation.

DEFINITION 7

1. The *depth* of a formula occurrence is zero if its order is zero, and it is the sum of the depths of the junctions occurring beneath it for a formula occurrence on a minor branch;

2. The depth of an assumption binding is always calculated for the assumption binding independently. It is the maximum of the depths of all the assumptions bound by the last rule;

3. The depth of a junction is one plus the maximum of the depths of all the relevant assumption bindings. For an $\supset\mathcal{E}$ rule with principal premiss $A \supset B$, the relevant assumption bindings are all those whose conclusion is $A \supset B$ and whose conclusion does not occur on the terminal thread beneath the conclusion of the $\supset\mathcal{E}$ rule.

EXAMPLE 8 Let d be

$$\dfrac{\dfrac{\dfrac{[A \supset A]^f \quad \dfrac{[A \supset A]^f \quad A^x}{A}}{A}}{(A \supset A) \supset A}\,(f) \qquad \dfrac{\dfrac{A \supset A^g \quad \dfrac{A \supset A^g \quad A^y}{A}}{A}}{A \supset A}\,(y)}{A}$$

The depth of the assumption binding which binds y is 3, and so the depth of each of the two junctions whose principal premiss is an assertion $A \supset A$ is 4. Consequently the depth of the assertion A^x is 8. The reader is invited to verify that the normal form obtained by applying three rewrites to this derivation has a single assertion in the assumption packet x whose depth is 6.

LEMMA 9 For any derivation d:

1. There is a unique depth associated with every formula occurrence;
2. Every formula occurrence in a given branch shares the same depth, and that depth is always at least as large as the order of the branch.

Observe that we only state the lemma in terms of the depth of formula occurrences, as the other two notions of depth are easily calculated from these.

PROOF

1. The first part is proven by induction on the height of derivations.
   (a) $\mathsf{height}(d) = 1$: As the only formula occurrence in the derivation is on the main branch, it has depth 0.
   (b) $\mathsf{height}(d) = n$ and $n > 1$: First we notice that all assumption bindings in d of height less than n have determinate depth by the induction hypothesis.

   Since the depth of each formula occurrence is determined by the depth of the junctions in d, if we can show that any junction in d has a depth, then we are done.

   The depth of a junction is determined by the depth of the subderivations of d whose conclusion is not beneath the conclusion of the junction. But since the last formula of the derivation is beneath this conclusion, all of these subderivations have height less than n, and so as we have noted have determinate depth.
2. Elementary.

□

DEFINITION 10

1. The *extended polynomials*, XP, are the ordinals[11] closed under the following operations:

$$0 \in XP$$
$$e + 1 \in XP \text{ when } e \in XP$$
$$e \cdot \omega \in XP \text{ when } e \in XP$$

[11]This treatment is taken from Muhammad Ali McBeth's *Combinatorial number theory* [McBethMA:comnt].

where addition and multiplication are understood in the usual polynomial sense[12];

2. Let $e, e' \in XP$. $e < e'$ if there exists $f \in XP$ such that $e' = e + f + 1$;

3. A *descending chain* of XP is a sequence $\langle e_i \rangle$ of ordinals of XP such that for each i, $e_{i+1} < e_i$. We do not require that the sequence to be finite;

4. The *measure* of a maximal formula occurrence is the sum of its degree and its depth. The *simple measure* of a derivation d is given by the following summation:
$$\sum_{v \in V} \omega^{m(v)}$$
where V is the collection of maxima in d, and for each $v \in V$, $m(v)$ is the measure of v.

The following properties of XP are quite elementary to establish, with the exception of the descending chain condition, which is equivalent to the consistency of finitary mathematics[13].

PROPOSITION 11

1. XP is closed under the operations of sum and product;
2. If $e, e' \in XP$ then either $e < e'$, $e = e'$ or $e > e'$;
3. The simple measure of every derivation of NJ is in XP;
4. All strictly descending chains in XP are finite.

We are now in a position to prove the normal form theorem:

DEFINITION 12 A maxima elimination chain, or just *elimination chain* is a sequence, finite or infinite, of derivations $\langle d_i \rangle$ where each d_{i+1} is obtained from d_i by eliminating a maximal formula. Call such a chain *full* if it is finite and its last derivation is a simple normal form.

LEMMA 13 Let d be a derivation, and let A be a maxima of d. Then the derivation obtained from d by applying the maxima elimination to A has measure strictly less than that of d.

PROOF We need to consider separately maxima eliminations governed by $\wedge$ and $\supset$:

[12]These operations correspond to Gentzen's natural sum and product, *not* Cantor's sum and product; consequently they are commutative.

[13]This follows from the fact that the XP consists of the ordinals under ω^ω, and this ordinal characterises the consistency strength of PRA, a formalism of the finitist mathematics of Hilbert. Consequently due to Gödel's incompleteness result it is impossible to prove the well-foundedness of ω^ωusing only finitist methods.

1. Without loss of generality, let the elimination rule be $\wedge\mathcal{E}1$. Then d has the general form:

$$\begin{array}{c} \begin{array}{cc} \vdots\, d_0 & \vdots\, d_1 \\ A & B \end{array} \\ \hline A \wedge B \\ \hline \begin{array}{c} A \\ \vdots\, d' \\ C \end{array} \end{array}$$

   To establish the conclusion it suffices to show that

   (a) The depth of each formula occurrence in the antecedent derivation is no more than its matching occurrence, if it has one, in the residual derivation;
   (b) No residual formula occurrence is a maxima in the residual derivation that is not either a maxima in the antecedent, or a subformula of the eliminated maxima.

   The last condition suffices since the natural sum of all the strict subformula of the eliminated maxima is less than the depth of the maxima, and it follows by inspection.

   There are two cases where after applying the maxima elimination to $A \wedge B$ the depth of a residual formula might differ from its antecedent:

   (a) If there is an assumption binding in d whose conclusion is on the terminal thread beneath the formula A displayed above, and whose deepest bound assumptions are in d_1. The depth of such a binding will reduce, as may junctions and formula occurrences which depend upon it;
   (b) A junction in d_0 or d' may depend upon assumption bindings occurring in d_1, and the depth of such a junction may reduce.

   In all these cases, the depth may reduce or stay the same, and as there are no substituted residuals, we are done.

2. Let the elimination rule be $\supset\mathcal{E}$ and let the antecedent derivation of the maxima elimination have the form:

$$\begin{array}{c} \begin{array}{cc} \begin{array}{c} [A]^x \\ \vdots\, d_0 \\ B \end{array} & \vdots\, d_1 \\ \hline A \supset B & A \end{array} \\ \hline \begin{array}{c} B \\ \vdots\, d' \\ C \end{array} \end{array}$$

and the residual derivation be

$$\begin{array}{c} \vdots\, d_1 \\ A \\ \vdots\, d_0 \\ B \\ \vdots\, d' \\ C \end{array}$$

and let $m+1$ be the depth of the eliminated junction. We call the residual formula occurrences of d_1 *indirect residuals* to indicate that they are substituted, and the other residuals direct. Our result here is complicated by the fact that there may be many indirect residuals matching each antecedent formula occurrence. We must establish:

(a) The depth of each direct residual is no more than its antecedent;
(b) The depth of each indirect residual is strictly less than its antecedent;
(c) No residual formula occurrence is a maxima in the residual derivation that is not either a maxima in the antecedent, or a subformula of the eliminated maxima.

A taste of the logical complications of this matter are given by the fact that we must establish:

(a) i. The depth of the assumption packets of d_0 and d_1 do not change;
    ii. The depth of assumption packets strictly contained in d' (ie. those which do not have their concluding formula on the terminal thread beneath the eliminated maxima) do not change;
(b) The depth of junctions of d_0 and d_1 do not increase;
(c) i. The depth of other assumption packets do not increase;
    ii. The depth of junctions in d' do not increase;
(d) i. Formulae occurrences of d_1 strictly reduce in depth;
    ii. Other formulae occurrences do not increase in depth.

and these propositions can only be established in this order. To complete the lemma:

(a) All of the assumption packets under this case do not change from antecedent to residual, and so their depth remains unchanged.
(b) The junctions of d_0 and d_1 depend only on the assumption packets in the previous part, since the assumption packets that we have not shown to have determinate depth are on the terminal thread beneath the junctions under consideration.
(c) Since junctions on the terminal thread beneath B may depend upon assumption packets further up this thread, we need to establish this result by induction over the length of the thread. But the induction is quite straightforward.

(d) i. If the formulae of d_1 have no residuals then this case is vacuously true. Otherwise it is the case that the depth of the eliminated junction is strictly greater than the differences between the depth of the occurrence of B in the residual derivation and each of the residuals A. Thus the depth of the substituted formulae strictly decreases from antecedent to residual.

ii. The branch structure is conserved from antecedent to residual, and each of the junctions has non-increasing depth.

□

THEOREM 14 (SIMPLE NORMAL FORM)
Let d justify $\Gamma \vdash A$. Then there is a full maxima elimination chain whose first derivation is d.

PROOF It is an easy corollary of the previous lemma that maxima elimination chains induce decreasing chains under the ordinal ω^ω, and so by well-foundedness all such chains are finite. We observe that a maxima elimination chain is full iff it is not a strictly initial subsequence of another chain. □

In fact, as we noted, we have succeeded in proving strong normalisation.

Properties of the calculus

An important consequence of the normal form theorem is the principal path lemma. For logical systems only involving direct elimination rules, the definition of path has a particularly simple form, and indeed in the current system it exactly coincides with the notion of branch.

DEFINITION 15

1. A thread in a derivation is a *principal thread* if every formula occurrence in it except the last is a principal premiss to a rule.

2. A principal thread is an *analytic thread* if every one of its formula occurrences except the last is the principal premiss of an elimination rule, and it is a *synthetic thread* if every formulae occurrence except the first is the conclusion of an introduction rule.

3. A *principal path* is a thread that is both initial and principal, and whose last formula occurrence is either the subsidiary premiss of an elimination rule, or is the concluding formula.

PROPOSITION 16 (PRINCIPAL PATH LEMMA)
If d is a normal derivation justifying $\Gamma \vdash A$, and which only contains rules governed by $\wedge$, and $\supset$, then any principal path in d is obtained by catenating an analytic and a synthetic thread.

PROOF Observe that in any triple of consecutive formulae in a principal path, $\langle A_i, A_{i+1}, A_{i+2}\rangle$, if A_{i+1} is the conclusion of an introduction and the premiss of an elimination, then it is a maximal formula. Thus by transitivity, the conclusion of an introduction rule is not a maxima only if either it is the final formula in the path, or it is a premiss of an introduction rule, and so all the introduction rules appear in a synthetic thread. By a dual argument, all the elimination rules appear in an analytic thread. □

PROPOSITION 17 (THE SUBFORMULA PROPERTY)
If d is a normal derivation justifying $\Gamma \vdash A$, then every formula occurrence of d is a subformula either of A or of some formula in Γ.

PROOF Every formula occurrence occurs in some principal path. By the principal path lemma it occurs either in the analytic part, in which case it is the subformula of the first formula of the path, or it occurs in the synthetic part, and is a subformula of the last formula of the path. In each of these cases, the subformula relationship arises from the transitivity of the subformula relationship existing between the principal premiss and conclusion that exists in introduction and elimination rules.

To complete the proof we must show that the end-formulae of any branch are subformulae of $\Gamma \cup \{A\}$. We achieve this by induction on the order of the branches in d. First we must prove a lemma:

LEMMA 18 Let A be a formula occurrence of order i which arises by an assertion. If A is discharged by an instance of $\supset \mathcal{E}$, then the conclusion of the rule is a formula occurrence in a branch of order less than or equal to i.

To see this, consider the assumption packet which binds A. Since the conclusion of the assumption packet is on the terminal thread beneath A, the difference between the order of these formula occurrences is positive, and is equal to the number of minor premisses of junctions occurring between the two formulae in this thread.

Now we seek to establish by induction the following property: for all branches of order i, the end-formulae of the branch are subformulae of $\Gamma \cup \{A\}$.

When $i = 0$, the branch is a main branch. If the first formula of this branch is discharged, then it is a subformula of the conclusion of the assumption packet that binds it. Since this conclusion is on the synthetic part of the thread, it is a subformula of the conclusion. Otherwise, if the assumption is open, it is a subformula of Γ. In either case the last formula is trivially a subformula of A.

When $i > 0$, it is the order of a minor branch. The last formula is the subsidiary premiss of an $\supset \mathcal{E}$ rule, and so it is a subformula of a formula of the principal premiss. Since the order of this formula occurrence is less than i, we may apply the induction hypothesis to obtain the conclusion.

If the first formula is not discharged, then it occurs among the open assumptions. If it is discharged, then there are two cases:

1. If it is discharged in this branch, then it is a subformula of the last formula of the branch, and so is a subformula of $\Gamma \cup \{A\}$;
2. If it is discharged in another branch, then it is a subformula of the last formula of that branch, and so we have by the induction hypothesis that it is a subformula of $\Gamma \cup \{A\}$.

□

COROLLARY 19 (CONSISTENCY)
There is no derivation d which justifies $\vdash X$ for any schematic letter X.

PROOF For a contradiction, we shall assume that d is such a derivation, and by the simple normal form theorem, we may assume that it is in a simple normal form. Since X is atomic, it cannot be the conclusion of an introduction rule, and since there are no open assumptions, it cannot be obtained by an assertion. Finally, it cannot be the conclusion of an elimination rule because by the subformula property the principal premiss of this rule would have to be a subformula of X. □

Normal form theorem

Finally, we shall justify the global form of the containment relation, where we say that d_0 is implicitly contained in d if d' is obtained from d by some sequence of the following rewrites, and d_0 is contained in d in the local sense.

DEFINITION 20

1. A formula occurrence is *minimal* if one of the following four cases holds:
   (a) It is the only formula occurrence on a path;
   (b) It is the first formula occurrence on a path, and it is the principal premiss of an introduction;
   (c) It is the last formula on a path and the conclusion of an elimination;
   (d) It is the conclusion of an elimination and the principal premiss of an introduction.
2. A *normal form* is a simple normal form all of whose minima are atomic.
3. A *minima decomposition* may be applied to any minima in a formula that is not atomic. If A is a non-atomic minima in d, then the minima decomposition is obtained by two applications of the substitution lemma as follows: if d' and $d''(x)$ are the subderivations of d above and below A, and f is the derivation justifying $A \vdash A$ in the decomposition part of the inversion principle for the connective governing A, then $d''(f[A \backslash d'])$.
4. A *rewrite chain* is a sequence, finite or infinite, of derivations $\langle d_i \rangle$ where each d_{i+1} is obtained from d_i by applying a rewrite, that is either a maxima elimination of a minima decomposition. It is *full* if its last element is a normal form.

THEOREM 21 (NORMAL FORM THEOREM)
There is a full rewrite chain starting from every derivation of minimal logic.

We shall not prove this theorem here, but instead we observe that it is a corollary of the normal form theorem proven in section 3.3, and also of the Tait method we show in section 1.5.

1.4 The Curry–Howard correspondence

The results of the previous section constitute strong support for Belnap's defence of logical formalism. By arguing that we can justify the claim of logical harmony by showing that the rules associated to each connective satisfy the inversion principle, and show that use of these rules do not affect our use of the other parts of our language, by means of the subformula property, we have shown how Belnap's criteria can be shown to obtain in the context of a formalised proof theory.

We might wish for more however. Prawitz's early writings suggest that if one derivation can be obtained from another by the rewrites we have described, then they should be regarded as in some sense the same. Can we make this idea precise, by means of a theory of identity for proofs?

We can give an affirmative response to this question by appealing to an observation of W. A. Howard[14], that proofs in intuitionistic logic may, under a suitable (bijective and homeomorphic) interpretation, be seen as constructions, and share their theory of identity.

The questions are resolved as follows: the notion of construction can be seen as the semantic content of the derivations of Prawitz's system, this notion being the inhabitant of a constructive type. The sameness of interpretations is established in two steps: firstly the definition of reduction can be seen to conserve the intuitive content of elementhood, and secondly we show that in the equivalence class of terms denoting an element, there is a unique canonical term. This canonical term corresponds to the normal form in natural deduction.

We shall begin our account by developing formally the simple theory of constructions.

The simple theory of constructions

We shall introduce a theory of constructions formulated as a typed lambda calculus, that is a theory whose basic syntax is derived from Church's lambda calculus, but in which terms are unambiguously associated with a type. The theory will

[14]In 'The formulae-as-types notion of construction' [HowardWA:fortnc]. H. B. Curry noted that application in the lambda calculus could be seen as an analogue to the rule *Modus Ponens* of logic, and W. W. Tait noted the similarity between normalisation in the lambda calculus and cut-elimination in Gentzen's calculi; see *Combinatory logic* [CurryHB:coml] and 'Intensional interpretations of functions of finite type I' [TaitWW:intiff]. In view of Tait's important contribution, perhaps the 'Curry–Howard–Tait' correspondence would have been a better label, but we shall follow the established usage.

have two type formers corresponding to Cartesian product and function space. We shall refer to the calculus as the theory λ.

The *types* of the theory fall into two kinds. There are the atomic types, which may include types drawn from a collection of primitive types, but will also include schematic letters in order to permit a simple form of polymorphism. The complex types are then freely generated over the set of atomic types by the binary type formers $\times$ and $\Rightarrow$.

The terms of the theory are those given by a type inference scheme whose general form resembles that we described for natural deduction. Instead of derivations, however we are concerned with type inference, and the nodes of our inference trees are type judgements whose content we shall describe in a moment. All terms are generated by a term grammar, the elements of which we call *term candidates*:

$$\begin{aligned} s ::= {} & x \\ & \mid \lambda x : A.s \mid \mathsf{ev}(s, s) \\ & \mid \langle s, s \rangle \mid \mathsf{outl}(s) \mid \mathsf{outr}(s) \end{aligned}$$

where x is a variable letter, and A is a type.

We allows a little syntactic sugar to make expressions more readable. We let ev admit more than two arguments by currying, so that

$$\mathsf{ev}(s, t_1, \ldots, t_n) \mathrel{\hat{=}} \mathsf{ev}(\, \mathsf{ev}(s, t_1) \ldots, t_n)$$

We also allow type information to appear as superscripts in variable binders, so that $\lambda x^A.s$ means $\lambda x : A.s$.

A type judgement is of the form $\Gamma \vdash s : A$ where s is a term candidate, A is a type, and Γ is a *telescope*[15], defined to be a finite set of pairs $x_i : A_i$ (where each x_i is a variable name, and each A_i is a type), which satisfies the property that each variable is distinct. The set of variables in a telescope Γ is called its *domain*, or $\mathsf{dom}(\Gamma)$, and the set of types its *range*. As with judgements in natural deduction, we must distinguish between the syntactic form of a type judgement, and the judgement qua judgement, which means that it is the conclusion of some valid type inference.

The type inferences, are generated in a fashion similar to those of natural deduction, except that we additionally have schematic variables representing telescopes, for which substitution is defined pointwise, and schematic letters representing terms. The rule basis is given as follows:

1. Variable rule:

$$\frac{x : A \in \Gamma}{\Gamma \vdash x : A}\,\mathrm{Var}$$

[15]Telescopes are more usually known as variable contexts, but we follow de Bruijn's usage to avoid confusion with term contexts.

2. Product rules:

$$\frac{\Gamma \vdash s : A \quad \Gamma \vdash t : B}{\Gamma \vdash \langle s, t\rangle : A \times B}$$

$$\frac{\Gamma \vdash s : A \times B}{\Gamma \vdash \mathsf{outl}(s) : A} \qquad \frac{\Gamma \vdash s : A \times B}{\Gamma \vdash \mathsf{outr}(s) : B}$$

3. Function space rules:

$$\frac{\Gamma, x : A \vdash s : B}{\Gamma \vdash \lambda x : A.s : A \Rightarrow B} \qquad \frac{\Gamma \vdash s : A \Rightarrow B \quad \Gamma \vdash t : A}{\Gamma \vdash \mathsf{ev}(s, t) : B}$$

We shall first be concerned to prove that this type inference system is well-behaved.

DEFINITION 22

1. The *subterms* of a term candidate are a set of term candidates given by the following function:

$$\begin{aligned}
\mathsf{st}(x) &\triangleq \{x\} \\
\mathsf{st}(\lambda x : A.s) &\triangleq \{\lambda x : A.s\} \cup \mathsf{st}(s) \\
\mathsf{st}(\mathsf{ev}(s,t)) &\triangleq \{\mathsf{ev}(s,t)\} \cup \mathsf{st}(s) \cup \mathsf{st}(t) \\
\mathsf{st}(\langle s,t\rangle) &\triangleq \{\langle s,t\rangle\} \cup \mathsf{st}(s) \cup \mathsf{st}(t) \\
\mathsf{st}(\mathsf{outl}(s)) &\triangleq \{\mathsf{outl}(s)\} \cup \mathsf{st}(s) \\
\mathsf{st}(\mathsf{outr}(s)) &\triangleq \{\mathsf{outr}(s)\} \cup \mathsf{st}(s)
\end{aligned}$$

2. The *free variables* of a term candidate are a set of variables given by the following function:

$$\begin{aligned}
\mathsf{FV}(x) &\triangleq \{x\} \\
\mathsf{FV}(\lambda x : A.s) &\triangleq \mathsf{FV}(s) - \{x\} \\
\mathsf{FV}(\mathsf{ev}(s,t)) &\triangleq \mathsf{FV}(s) \cup \mathsf{FV}(t) \\
\mathsf{FV}(\langle s,t\rangle) &\triangleq \mathsf{FV}(s) \cup \mathsf{FV}(t) \\
\mathsf{FV}(\mathsf{outl}(s)) &\triangleq \mathsf{FV}(s) \\
\mathsf{FV}(\mathsf{outr}(s)) &\triangleq \mathsf{FV}(s)
\end{aligned}$$

PROPOSITION 23

1. If s is a term, then every $t \in \mathsf{st}(s)$ is a term.
2. If $\Gamma \vdash s : A$, then for every $v \in \mathsf{FV}(s)$, v is in the domain of Γ.
3. If $\Gamma \vdash s : A$, then there is a unique type inference justifying this conclusion.
4. It is decidable whether a term candidate is a term or not.

PROOF The first three parts can be shown by a simple structural induction over a type inference justifying $\Gamma \vdash s : A$. The last part is a consequence of the Hindley-Milner type inference algorithm[16]. Indeed the type inference algorithm is more general than required, as we need not have type annotations on the λ term former. □

The theory of constructions also requires an account of when one term represents the same construction as another. Before we engage in this account, we shall insist upon a variable naming convention. The convention we shall use is weaker than the well-known Barendregt variable convention[17], but also is easier to formalise.

CONVENTION 24 (BOUND IDENTIFIER CONVENTION)
Suppose $\Gamma \vdash s : A$. Then the identifier associated with each bound variable of s is distinct from that of all the variables in the domain of the telescope Γ. Furthermore for any subterm $s' \in \mathsf{st}(s)$ there is a telescope $\Gamma, \Gamma' \vdash s' : A'$; the above condition applies to these subterms also. The inference rules of λ can only be used to justify terms satisfying this convention.

To define the substitution operation, it is necessary to appeal to alpha conversion, or the principle that if $\Gamma \vdash \lambda x^B.t : B \Rightarrow A$ then for $y \notin \mathsf{dom}(\Gamma)$, $\lambda x^B.t$ is considered to be identical to the term $\lambda y^B.t'$ where t' is obtained from t by replacing all occurrences of x by y.

DEFINITION 25 The *substitution operation* $s[x := t]$ is defined for any variable x and term candidates s, t. Then $s[x := t]$ is obtained from s by replacing every occurrence of x in s by t', where t' is obtained by repeated applications of alpha conversion so that the bound variables of t' are distinct from any bound variables of s.

REMARK 26 Note that this substitution operation is not capture free; its correctness depends upon the convention to ensure that the free variables of t' are distinct from the bound variables of s.

LEMMA 27 (WEAKENING LEMMA) If $\Gamma \vdash s : A$, then for any Γ' with $\mathsf{dom}\Gamma \cap \mathsf{dom}(\Gamma') = \emptyset$, $\Gamma, \Gamma' \vdash s : A$.

PROOF We see by inspection that we may apply the weakening lemma to the conclusion of each rule if we may apply it to each of its premisses, and so the lemma follows by structural induction. □

PROPOSITION 28 If $\Gamma \vdash s : A$ and $\Gamma, x : A \vdash t : B$ then $\Gamma \vdash t[x := s] : B$.

PROOF Replace every type inference leaf justifying $\Gamma, x : A, \Gamma' \vdash x : A$ in the type inference of $\Gamma, x : A \vdash t : A$ with the Γ'-weakening of s, $\Gamma, \Gamma' \vdash s : A$, and remove all occurrences of the variable–type pair $x : A$ from the remainder of the tree. A simple induction validates that this is a correct type inference. □

[16]See the chapter on type assignment in *Introduction to combinators and lambda calculus* [HindleyJR:intclc], or Barendregt's 'Typed lambda calculi' [BarendregtH:lamcwt] for details.

[17]See *The lambda calculus* [BarendregtH:lamcss].

DEFINITION 29

1. A *beta redex* is a term with one of the forms $\mathsf{ev}(\lambda x : A.s, t)$, $\mathsf{outl}(\langle s, t\rangle)$, or $\mathsf{outr}(\langle s, t\rangle)$.

2. A *beta conversion* is a relation holding between the set of redexes and the set of term candidates defined:

$$\begin{aligned}\mathsf{outl}(\langle s, t\rangle) &\to^c_\beta s\\ \mathsf{outr}(\langle s, t\rangle) &\to^c_\beta t\\ \mathsf{ev}(\lambda x : A.s, t) &\to^c_\beta s[x := t]\end{aligned}$$

3. A *redex–contractum pair* is a pair $\langle r, c\rangle$ such that $r \to^c c$.

PROPOSITION 30 (SUBJECT REDUCTION)
If $\Gamma \vdash s : A$ and $s \to^c_\beta t$, then $\Gamma \vdash t : A$.

PROOF We obtain the result by examining the general form of a type inference tree of each redex, and showing that we can build a type inference tree of the contractum from it. □

DEFINITION 31

1. An (n-ary) *simple definition* is a judgement form displayed

$$\Gamma \vdash f \mathrel{\hat{=}} s : A$$

It indicates firstly that $\Gamma \vdash s : A$, and secondly that occurrences of the identifier f are to be understood as shorthand for the term s, in which the variables of Γ may occur;

2. An (n-ary) *parameterised definition* is a judgement form displayed

$$\Gamma \vdash f(x_1 : B_1, \ldots, x_n : B_n) \mathrel{\hat{=}} s : A$$

where for each i, $x_i \notin \mathsf{dom}(\Gamma)$. It indicates that firstly, $\Gamma, x_1 : B_1, \ldots, x_n : B_n \vdash s : A$, and that each x_i occurs precisely once in s, and secondly that occurrences of the form $f(t_1, \ldots, t_n)$, where for each i we are to take $\Gamma \vdash t_i : B_i$ as shorthand for the term $s[x_1 := t_1, \ldots, x_n := t_n]$;

3. An (n-ary) *context*[18] is a judgement form displayed

$$\Gamma \vdash f((\Delta_1)x_1 : B_1, \ldots, (\Delta_n)x_n : B_n) \mathrel{\hat{=}} s : A$$

It indicates firstly that $\Gamma \vdash f(x_1 : B_1, \ldots, x_n : B_n) \mathrel{\hat{=}} s : A$ is a parameterised definition and also that each judgement $\Gamma, \Delta_i \vdash x_i : B_i$ occurs once in the type inference of $\Gamma \vdash s : A$. Secondly the restriction on the t_i in the instantiations of $f(t_1, \ldots, t_n)$ are relaxed to $\Gamma, \Delta_i \vdash t_i : B_i$, or equivalently, the Δ_i indicate variables of t_i that may be captured upon substitution;

[18] This leads to the same set of contexts as the usual definition of term context. I give the more complex definition to indicate which terms may be validly substituted into the context, and because of its closeness to the treatment of de Bruijn.

4. An *n-ary term decomposition* of a term s with $\Gamma \vdash s : A$ consists of an n-ary context f and terms $t_1, \ldots, t_n$ such that $s \equiv f(t_1, \ldots, t_n)$.

EXAMPLE 32

1. $f : A \Rightarrow A \vdash F \triangleq \lambda x^A.\mathsf{ev}(f, \mathsf{ev}(f, \mathsf{ev}(f, x))) : A \Rightarrow A$ is a simple definition;
2. $\vdash d(x : A) \triangleq \langle x, x \rangle : A \times A$ is a 1-ary parameterised definition;
3. The term F defined in part 1 admits the following term decompositions, amongst others:
   (a) $f : A \Rightarrow A \vdash F_0((x^A)y : A) \triangleq \lambda x^A.\mathsf{ev}(f, y) : A \Rightarrow A$ in $F_0(\mathsf{ev}(f, \mathsf{ev}(f, x)))$;
   (b) $f : A \Rightarrow A \vdash F_1((x^A)y : A \Rightarrow A, (x^A)z : A) \triangleq \lambda x^A.\mathsf{ev}(y, \mathsf{ev}(f, z))) : A \Rightarrow A$ in $F_0(f, \mathsf{ev}(f, x))$.

DEFINITION 33

1. The *compatible closure* of a relation $\phi(-, -)$ between term candidates is defined to be $\phi^*(-, -)$ as follows: $\phi^*(s, t)$ is true when there are term decompositions $s \equiv M[s']$ and $t \equiv M[t']$ such that $\phi(s', t')$;
2. $\rightarrow^1_\beta$ is defined to be the compatible closure of $\rightarrow^c_\beta$, and $\rightarrow^*_\beta$ is defined to be the reflexive, transitive closure of $\rightarrow^1_\beta$. $=_\beta$ is the reflexive, symmetric and transitive closure of $\rightarrow^1_\beta$.

We are now in a position to state the Curry–Howard correspondence formally[19]. There are three limbs to the correspondence: the bijection between formulae and types, between derivations and terms, and between normalisation and reduction.

We must first presuppose an equivalence between the primitive formulae and primitive types. We shall avoid a need to presuppose a correspondence in the other two limbs of the correspondence, by restricting our attention to derivations not containing any reductions we have not explicitly introduced.

DEFINITION 34 We define a mapping $[\![-]\!]$ from formulae to terms:

$$\begin{aligned} [\![A]\!] &\triangleq A, \text{ if } A \text{ is atomic} \\ [\![A \wedge B]\!] &\triangleq [\![A]\!] \times [\![B]\!] \\ [\![A \supset B]\!] &\triangleq [\![A]\!] \Rightarrow [\![B]\!] \end{aligned}$$

PROPOSITION 35 $[\![-]\!]$ induces a bijection between formulae and types.

The second limb relies upon a correspondence between the term formers and the logical rules. We observe that term formers fall into two kinds, *constructors* which allow us to introduce terms govern by a type, and *extractors* which allow us to recover information from an element of such a type. There is then a formal

[19] This is not exactly the same as stated by W. A. Howard in his paper, as we have 'cleaned up' the theory of constructions somewhat, and we use Prawitz's system rather than Gentzen's.

relationship between the constructors of a type former, and the introduction rule of the related logical connective, and between the extractors and the elimination rules. We shall also relate the identifiers of assumption packets with variables.

DEFINITION 36 Let d justify $\Gamma \vdash A$. We associate a term $[\![d]\!]$ with d as follows:

1. Assertion. If the derivation consists of an assertion, then the term is x, where x is the identifier associated with the assertion A.

2. Conjunction. In case d is

$$\dfrac{\begin{array}{c}\vdots\, d_0\\ A\end{array} \quad \begin{array}{c}\vdots\, d_1\\ B\end{array}}{A \wedge B}\wedge\mathcal{I} \quad \text{or} \quad \dfrac{\begin{array}{c}\vdots\, d'\\ A \wedge B\end{array}}{A}\wedge\mathcal{E}1 \quad \text{or} \quad \dfrac{\begin{array}{c}\vdots\, d'\\ A \wedge B\end{array}}{B}\wedge\mathcal{E}2$$

then $[\![d]\!]$ is either $\langle [\![d_0]\!], [\![d_1]\!] \rangle$, $\mathsf{outl}([\![d']\!])$, or $\mathsf{outr}([\![d']\!])$ respectively.

3. Implication. In case d is

$$\dfrac{\begin{array}{c}[A]^x\\ \vdots\, d'\\ B\end{array}}{A \supset B}(x)\supset\mathcal{I} \quad \text{or} \quad \dfrac{\begin{array}{c}\vdots\, d_0\\ A \supset B\end{array} \quad \begin{array}{c}\vdots\, d_1\\ A\end{array}}{B}\supset\mathcal{E}$$

then $[\![d]\!]$ is either $\lambda x : A.[\![d']\!]$ or $\mathsf{ev}([\![d_0]\!], [\![d_1]\!])$.

The expression of this leg of the correspondence is complicated somewhat by the disparity between the form of judgements usual in natural deduction, which do not explicitly mention assumption packets, and those in the theory of constructions, which do. We have only a surjection from the latter form of judgement to the former, however this does not affect the bijectivity of the correspondence between terms and derivations.

PROPOSITION 37

1. Let $\Gamma \vdash s : A$, and let Γ^* be the range of Γ. Then there is a derivation d such that $s \equiv [\![d]\!]$, and d justifies $\Gamma^* \vdash A$;

2. $[\![-]\!]$ induces a bijection between closed derivations and terms.

We need to restrict the above bijection to closed terms, since free variables of terms of λ carry no type information.

The final leg of the correspondence rests upon recognising that in each redex, the type associated with the constructor subterm of the redex corresponds to a maxima in the derivation corresponding to the redex.

PROPOSITION 38

1. Let d be a derivation such that $[\![d]\!]$ is a redex. Then the last rule of d is an elimination, and the last rule of the principal subderivation is an introduction.

2. Let s be a term, and let $s \to^1_\beta t$. Let d, d' be derivations corresponding to s, t. Then d' arises from d by a maxima elimination.

We also note that a maxima elimination chain induces a relation which is the reflexive transitive closure of that induce by maxima elimination, and so is in a bijective correspondence with that of beta reduction. Similarly there is a bijection between normal forms and irreducible terms.

Some consequences of the correspondence

The first benefit which derives from the correspondence is that it makes precise the claim that intuitionistic logic is constructive. As we shall see shortly, Gödel was dissatisfied with the account of constructivity of the Brouwer–Heyting–Kolmogorov (or BHK) interpretation of the intuitionistic connectives as it appealed to the abstract notion of proof. Let us examine this correspondence for the connectives $\wedge$ and $\supset$:

1. A proof of $A \wedge B$ consists of a proof of A and a proof of B;
2. A proof of $A \supset B$ consists of a method of transforming proofs of A to proofs of B.

We can make this interpretation precise using the Curry–Howard correspondence, for 'a proof of A and a proof of B' has a natural correspondence with 'element of the Cartesian product of the type of proofs of A and the type of proofs of B', and similarly between 'method of transforming' and 'function space'. Indeed, if we define the function Prf from propositions to the type of proofs justifying the argument, we obtain:

1. $p \in \mathsf{Prf}(A \wedge B)$ when $p \in \mathsf{Prf}(A) \times \mathsf{Prf}(B)$, and
2. $p \in \mathsf{Prf}(A \supset B)$ when $p \in \mathsf{Prf}(A) \Rightarrow \mathsf{Prf}(B)$;

which, given extensionality, yields $\mathsf{Prf}(A) = [\![A]\!]$. Furthermore, we obtain a vindication of the maxima decomposition rewrites, as they are equivalent to the notion of beta equality, and as we shall see the minima decomposition rules have their analogue as well.

The articulation of the precise sense in which intuitionistic logic is constructive was an important goal of many proof theorists around the time of Hilbert's programme and during the years following. To see that it is not a trivial problem consider the following observation[20]:

DEFINITION 39 Let D be an inhabited universe, and let $R_0, \ldots, R_n$ be atomic relations defined upon D. Let ϕ be any sentence built up from primitive formulae $\phi(\vec{d})$ using the logical connectives $\perp$, $\supset$ and $\forall$. Define the set $P(\phi)$ for each ϕ, interpreting the right hand comprehensions in classical ZF set theory.

[20] Adapted from exercise 1.3.4 of Troelstra and van Dalen's *Constructivism in Mathematics* [TroelstraAS:conmi].

1. $P(R_i(\vec{d}) \mathrel{\hat{=}} \begin{cases} \{*\} & \text{if } R_i(\vec{d} \text{ is true.} \\ \emptyset & \text{otherwise} \end{cases}$
2. $P(\bot) \mathrel{\hat{=}} \emptyset$
3. $P(\phi \supset \psi) \mathrel{\hat{=}} \{f : P(\psi) \Rightarrow P(\phi)\}$
4. $P(\forall x.\phi(x)) \mathrel{\hat{=}} \{f : D \Rightarrow P^* \mid \forall d \in D.f(d) \in P(\phi(d))\}$ where $P^* \mathrel{\hat{=}} \bigcup\{P(\phi(d)) \mid d \in D\}$.

PROPOSITION 40 If ϕ is provable in classical predicate calculus, then $P(\phi)$ is inhabited.

Consequently a classical notion of 'construction' immediately yields an interpretation of the BHK correspondence consistent with classical logic. So we must make precise what we mean by construction to give the BHK interpretation more than platitudinous import. Three such formulations have become cornerstones of modern proof theory:

1. Gentzen's cut-elimination procedure and the technique of ordinal assignment;
2. Gödel's functional, or 'Dialectica', interpretation;
3. Kleene's recursive realizability.

It is possible to regard the Curry–Howard correspondence as bridging the gap between cut-elimination and functional interpretation (we shall not be further concerned with realisability). In the first place, the proof-theoretic side is very close as we have explained[21], and in the second place, there is a close relationship between the target calculi of Gödel's functional interpretation and the calculi in formulae-as-types correspondence with various formalisations of intuitionistic proof theories. Let us examine Gödel's technique a little more closely.

Gödel's functional interpretation[22] is a reduction of intuitionistic arithmetic to the propositional theory PRA$^\omega$ whose atomic propositions are equations between the terms of the system, known as the finite-type functionals. To each proof of $\vdash_{\text{HA}} \phi(\vec{x})$, there is an associated unquantified formula $\phi_D(\vec{x}, \vec{y}, \vec{z})$ and terms $\vec{t}$ of PRA$^\omega$ such that $\vdash_{\text{PRA}^\omega} \phi_D(\vec{x}, \vec{t}, \vec{z})$, where the variables $\vec{x}$ range over integers and $\vec{z}$ range over terms of given finite type.

Gödel presented his functional interpretation as a natural extension of finitary metamathematics capable of proving the consistency of Heyting Arithmetic[23]

[21] However, the sequent calculus is preferred by a majority of proof-theorists, and its correspondence with natural deduction is not exact. See Kleene's 'Permutability of inferences in Gentzen's calculi LJ and LK' [KleeneSC:perigc] and Ungar's *Normalisation, cut-elimination and the theory of proofs* [UngarAM:norcet].

[22] 'On a hitherto unutilised extension of the finitary standpoint' [GoedelK:hituef].

[23] Gödel's aims in presenting his functional interpretation, and the evolution of his theory are traced in Anne Troelstra's excellent introduction to Gödel's paper in the *Collected Works* [GoedelK:colw].

(which has equivalent consistency strength to Peano Arithmetic by the double-negation translation). He was dissatisfied with the Brouwer–Heyting–Kolmogorov account of the constructivity of proofs due to its dependence upon the abstract idea of proof. Gödel's technique has the advantages, summarised by Solomon Feferman[24] as follows:

1. It recasts the logical superstructure into operations explicitly defined in a computationally meaningful context;
2. It verifies the intuitionistic admissibility of the axiom of choice;
3. It shows how induction can be characterised in terms of recursion, allowing us to obtain a perspicuous characterisation of the provable recursive functions.

We see immediately from the results of this section that the Curry–Howard correspondence answers the first concern. Furthermore, when we come to treat existence and induction in the context of Martin-Löf's intuitionistic type theory, which depends heavily upon the formulae-as-types correspondence, we shall see that the other two advantages also accrue. We might also add that the Curry–Howard is more accessible, due to its directness.

The correspondence does not entirely supersede the Dialectica correspondence, however. Not all proof theories easily admit formulae-as-types correspondences, especially theories based upon classical logic, and the functional interpretation is often more revealing when it is applied to the results of the double-negation translation. Furthermore, the feature which makes functional interpretation opaque to the student, the role of the formula Φ_D, makes possible the revealing 'no counterexample' interpretation of classical logic[25], and Troelstra remarks that functional interpretation was originally conceived by Gödel's desire to show that $\neg\neg\forall x.(\phi(x) \vee \neg\phi(x))$ was unprovable in intuitionistic logic.

In the remainder of this section we shall be concerned with two issues, firstly to show how the theory of reduction for λ gives us a decidable theory of identity of proofs, and the presentation of a strengthened account of the semantic content of a derivation.

DEFINITION 41

1. A relation on terms $\phi(-,-)$ is *Church–Rosser* if for all terms s, t, t' with $\phi^*(s,t)$ and $\phi^*(s,t')$, there is a term u for which $\phi^*(t,u) \wedge \phi^*(t',u)$, where ϕ^* is the transitive closure on ϕ;

2. A relation on terms $\phi(-,-)$ satisfies the *diamond property* if for all terms s, t, t' with $\phi(s,t)$ and $\phi(s,t')$, there is a term u for which $\phi(t,u) \wedge \phi(t',u)$;

THEOREM 42 The one-step reduction relation $\rightarrow^1_\beta$ is strongly normalising, and the full reduction relation $\rightarrow^*_\beta$ is Church–Rosser.

[24]See his 'Gödel's Dialectica interpretation and its two-way stretch' [FefermanS:goedii].

[25]Georg Kreisel, 'A survey of proof theory' [KreiselG:surpt].

PROOF Strong normalisation is a direct consequence of the finiteness of maxima rewrite chains. We shall briefly sketch an outline proof of the Church–Rosser property; the details may be filled in by consulting as standard reference[26].

In the presence of strong normalisation, the Church–Rosser property is equivalent to the diamond property. The diamond property holds if for all pairs of formula occurrences A, A' of a given term s considered as a derivation, if A, A' are maxima or minima, then the two associated rewritten terms can in turn be rewritten to a shared form. If A, A' are non-overlapping, then this common form can be obtained by rewriting all the residuals of the redex not contracted. Thus proof of Church–Rosser reduces to proof of the diamond property for overlapping redexes, which are commonly known as *critical pairs*.

Since all redexes are determined by the rule to which they are the conclusion and the rule to which they are the premiss (or the fact that they are the last rule), it follows that the overlapping redexes are all vertically adjacent in the proof tree, and so it is quite easy to enumerate them, a process whose content is captured by the Knuth–Bendix algorithm. In fact there are no critical pairs for the beta theory, and so we are done. □

COROLLARY 43

1. Every term has a beta unique normal form.
2. The relation $=_\beta$ is decidable.

We have so far only considered the relationship between maxima elimination and beta reduction. However the full theory of proof equivalence that we considered in the last section also involves minima decompositions.

The equivalence on terms induced by allowing minima decompositions is represented in λ by eta equality, a form of reduction introduced to model intuitions about extensionality. However eta conversions are usually conceived of as running in the opposite direction to minima decompositions, ie. they are reductions, which eliminate dual pairs of rule applications. There are two good reasons for this: firstly, we cannot obtain the definition of one-step reduction from the definition of the eta conversions directly by compatible closure without sacrificing strong normalisation, and secondly, we will also need typing constraints on the definition of eta conversions to preserve subject reduction if reduction is to run the other way, as a variable might be governed by any connective, or none at all.

We have chosen the reverse direction, known as eta-expansion, for two reasons: firstly, it leads to a more natural statement of the semantic content of a proof, and secondly, it leads to a better behaved reduction relation for the calculus when we introduce extensions later in this work.

DEFINITION 44 The *eta conversions* are defined as follows:

1. For $\Gamma \vdash s : A \times B$:

$$s \to_\eta^c \langle \mathsf{outl}(s), \mathsf{outr}(s) \rangle$$

[26]See, for example, *Introduction to combinators and lambda calculus* [HindleyJR:intclc].

2. For $\Gamma \vdash s : A \Rightarrow B$:

$$s \to_\eta^c \lambda x : A.\mathsf{ev}(s, x)$$

We will obtain the definition of one-step reduction by direct analogy with the definition of minima, which we repeat:

1. It is the only formula occurrence on a path;
2. It is the first formula occurrence on a path, and it is the principal premiss of an introduction;
3. It is the last formula on a path and the conclusion of an elimination;
4. It is the conclusion of an elimination and the principal premiss of an introduction.

Equivalently, a formula is not a minima if it is the conclusion of an introduction, or the principal premiss of an elimination rule.

DEFINITION 45 Let $\Gamma \vdash s : A$, and let $f((\Gamma')x : B)$ and t be a term decomposition of s. Then

1. t is a $\times\eta$ redex of s if $B \equiv C \times C'$, the last rule of t is not an introduction, and the assertion $C \times C'^x$ is not the premiss an $\wedge\mathcal{E}$ rule (or equivalently neither $\mathsf{outl}(x)$ nor $\mathsf{outr}(x)$ are subterms of $f(x)$). We write $f(t) \to_\eta^1 f(\langle \mathsf{outl}(t), \mathsf{outr}(t)\rangle)$;

2. t is a $\Rightarrow\eta$ redex of s if $B \equiv C \Rightarrow C'$, and t is not a lambda abstraction, and the assertion $C \times C'^x$ is not the premiss to the $\supset\mathcal{E}$ rule (or equivalently $\mathsf{ev}(x, u)$ is not a subterm of $f(x)$ for any u). We write $f(t) \to_\eta^1 f(\lambda x^C.\mathsf{ev}(t, x))$.

$\to_\eta^*$ is defined to be the reflexive transitive closure of $\to_\eta^1$, and $\to_{\beta\eta}^1$ be defined to be the union of $\to_\beta^1$ and $\to_\eta^1$. $\to_{\beta\eta}^*$ is defined to be the reflexive transitive closure of $\to_{\beta\eta}^1$.

PROPOSITION 46 $\to_{\beta\eta}^1$ is strongly normalising, and $\to_{\beta\eta}^*$ is Church–Rosser.

PROOF Strong normalisation follows from finiteness of elimination chains, of which we noted we still owe a proof, to be given in section 3.3. The Church–Rosser property follows from the diamond property, which follows from the diamond property restricted to critical pairs. There are two critical pairs: they take the form $\mathsf{ev}(\lambda x^A.s, t)$ where t is either a $\Rightarrow\eta$ redex or a $\times\eta$ redex. It is quite elementary to confirm that these share a reduced form. □

COROLLARY 47

1. Every term has a unique beta-eta normal form.
2. $=_{\beta\eta}$ is decidable.

Finally let us consider an extension of the principal path lemma that we can formulate in λ.

DEFINITION 48

1. Let $\Gamma \vdash s : A$. We say that s is a *recursor term* if it is normal form and A is atomic.
2. The context $I((\Delta_1)x_1 : B_1, \ldots, (\Delta_n)x_n : B_n)$ and the terms $t_1, \ldots, t_n$ form a *constructor–recursor term decomposition* of s if:
   (a) I consists only of introduction rules;
   (b) Each $\Gamma, \Delta_i \vdash t_i : B_i$ is a recursor term; and
   (c) $I(t_1, \ldots, t_n)$ is the normal form of s.

PROPOSITION 49 Every term has a unique constructor–recursor decomposition.

PROOF Since every term has a unique normal form, we may without loss of generality consider only these. Let s be such a normal form, with $\Gamma \vdash s : A$, and let $A_1, \ldots, A_n$ be its minimal formula occurrences of order 0. Observe that by the principal path lemma, for no i, j is A_i above A_j. Therefore we may define $M(t_1, \ldots, t_n)$ to be the characteristic formula decomposition of $A_1, \ldots, A_n$. Since each A_i is atomic, each t_i is an extractor. It only remains to show that $M((\Delta_1)x_1, \ldots, (\Delta_n)x_n)$ consists only of introduction rules.

Firstly we should note that all formula occurrences of M correspond to formula occurrences in s of order 0. There are two main cases to consider:

1. The last rule of s is an elimination rule, or an assertion. Then the conclusion of s is atomic, since otherwise it would admit a minima decomposition. Therefore s is an extractor and M is single-ended and consists of a single variable;
2. The last rule of s is an introduction. Then the terminal thread beneath each A must consist only of the conclusion of introductions, since every formula occurrence of order 0 is on one of the principal paths of s.

□

The constructor–recursor decomposition of a term can be seen as a kind of analog of the principal path lemma for natural deduction. However, whilst the threads in a constructor are all introduction threads, the structure of the extractors are more complex in two respects:

1. An extractor t_i may depend upon new assumptions, ie. they are derivations of $\Gamma, \Delta_i \vdash B_i$;

2. Recursors need not only contain elimination rules and assumptions, but may also contain introductions in the subsidiary subderivations of $\supset\mathcal{E}$ rules.

This complexity is due to the fact that the semantic content of a proposition $A \supset B$ depends upon the way in which we may draw conclusions from A as well as the way in which we may assert B, introducing a kind of polarity reversal into the justifications.

1.5 Recursion

The simple theory of constructions may be considered to be a theory of computation as it is a fragment of the lambda calculus. However it lacks the power to express such typical computations such as the evaluation of arithmetic expressions[27].

It is possible to extend the existing system by adding a type to represent the natural numbers, complete with formation rules and a recursion operator, together with an equational theory based upon conversion. The system we obtain we call *PRA*$^{\omega}$[28], a system capable of defining all provably total functions of Peano Arithmetic, and rather more powerful than primitive recursion.

We allow a new, primitive, type $\mathbb{N}$, and term formers $\mathsf{zero}, \mathsf{succ}(-)$. The type inferences of this system are given as follows[29]:

$$\frac{}{\Gamma \vdash \mathsf{zero} : \mathbb{N}} \qquad \frac{\Gamma \vdash n : \mathbb{N}}{\Gamma \vdash \mathsf{succ}(n) : \mathbb{N}}$$

$$\frac{\Gamma \vdash s : \mathbb{N} \quad \Gamma \vdash a : A \quad \Gamma \vdash f : \mathbb{N} \Rightarrow (A \Rightarrow A)}{\mathsf{Rec}(s, a, f) : A}$$

For convenience we also define the closed terms of $\mathbb{N}$ $\underline{0} \hat{=} \mathsf{zero}$, and for each k $\underline{k+1} \hat{=} \mathsf{succ}(\underline{k})$.

We shall describe the recursor as an elimination rule whose first premiss is the principal premiss and whose other premisses are auxiliary premisses, although the rules just defined are not of the same status with respect to logical harmony as the rules for $\supset$ and $\wedge$. Recall that the formula A in the definition above is called a *cut formula*

There is a difficulty in the definition of the theory: that there is no obviously superior form of conversion for the calculus. There are three principal alternatives, whose beta conversions we give as follows, the relations $\rightarrow^1_{\mathbb{N}\beta}$ and $\rightarrow^*_{\mathbb{N}\beta}$ arising by

[27]One *can* encode some arithmetic in the simply-typed lambda calculus by using the Church coding of the natural numbers as the type $X \Rightarrow (X \Rightarrow X) \Rightarrow X$, but one can only code the extended polynomials in this treatment, as was shown by Schwichtenberg [SchwichtenbergH:defflk], so it is impossible to represent, for example, such typically primitive recursive functions as Gödel's beta function, or even the test for whether a number is odd.

[28]It is essentially Gödel's system T. We call it PRA$^{\omega}$ after Solomon Feferman's treatment in 'Theories of finite-type related to mathematical practice' [FefermanS:theftr].

[29]Recall that $\mathsf{ev}(f, a, b)$ is shorthand for $\mathsf{ev}(\mathsf{ev}(f, a), b)$.

compatible closure and transitivity in the usual way (there is no eta convertibility for this type):

1. Strict recursion.

$$\begin{aligned}\mathsf{Rec}(\underline{0}, a, f) &\to^c_\beta a\\ \mathsf{Rec}(\underline{k+1}, a, f) &\to^c_\beta \mathsf{Rec}(\underline{k}, \mathsf{ev}(f, \underline{0}, a), f)\end{aligned}$$

2. Left recursion.

$$\begin{aligned}\mathsf{Rec}(\underline{0}, a, f) &\to^c_\beta a\\ \mathsf{Rec}(\mathsf{succ}(s), a, f) &\to^c_\beta \mathsf{Rec}(s, \mathsf{ev}(f, \underline{0}, a), f)\end{aligned}$$

   It is called *left recursion* by analogy with the **foldl** combinator of Bird and Wadler[30];

3. Right recursion.

$$\begin{aligned}\mathsf{Rec}(\underline{0}, a, f) &\to^c_\beta a\\ \mathsf{Rec}(\mathsf{succ}(s), a, f) &\to^c_\beta \mathsf{ev}(f, s, \mathsf{Rec}(s, a, f))\end{aligned}$$

   Dually, *right recursion* is by analogy with Bird and Wadler's **foldr** combinator.

For terms of the form $\mathsf{Rec}(\underline{k}, a, f)$, the three schemes above give the same normal form. However, when the first argument to the recursion operator is not of the form k, the three schemes can give different normal forms, as in the case of $\mathsf{Rec}(\mathsf{succ}(x), \underline{0}, \lambda m^{\mathbb{N}}.\lambda n^{\mathbb{N}}.\underline{1})$.

Thus the equational theory of PRA$^\omega$ comes in three different flavours, with Gödel's original treatment corresponding to the right-recursive form. PRA$^\omega$ has a sub-theory PRA which is obtained by insisting that the only permissible cut-formula for the recursor rule is $\mathbb{N}$.

PRA is a very expressive language for computation: as Gödel showed it is possible to encode a decision procedure for the validity of proofs of almost all existing proof theories, and almost all useful computations can be coded in it. We give a few simple examples:

1. Addition.

$$\mathsf{add}(x, y) \doteq \mathsf{Rec}(x, y, \lambda i^{\mathbb{N}}.\lambda z^{\mathbb{N}}.\mathsf{succ}(z))$$

[30] See their *An Introduction to Functional Programming* [BirdR:intfp].

In the right recursive form, the conversion theory proves the following equations (where m, n are free variables, and the equality sign denotes beta-eta conversion), showing that addition satisfies the intended meaning:

$$\begin{aligned}\mathsf{add}(\underline{0}, n) &= n,\\ \mathsf{add}(\mathsf{succ}(m), n) &= \mathsf{succ}(\mathsf{add}(m, n)).\end{aligned}$$

2. Multiplication.

$$\mathsf{mul}(x, y) \mathrel{\hat{=}} \mathsf{Rec}(x, \underline{0}, \lambda i^{\mathbb{N}}.\lambda z^{\mathbb{N}}.\mathsf{add}(y, z));$$

3. Predecessor.

$$\mathsf{pred}(x) \mathrel{\hat{=}} \mathsf{Rec}(x, \underline{0}, \lambda i^{\mathbb{N}}.\lambda z^{\mathbb{N}}.i);$$

4. Subtraction.

$$\mathsf{subtr}(x, y) \mathrel{\hat{=}} \mathsf{Rec}(x, y, \lambda i^{\mathbb{N}}.\lambda z^{\mathbb{N}}.\mathsf{pred}(z));$$

5. Maximum.

$$\mathsf{max}(x, y) \mathrel{\hat{=}} \mathsf{add}(\mathsf{subtr}(x, y), x)$$

Equations characterising the other functions can be found easily enough, and shown to be validated by PRA.

EXERCISE 50 PRA formulated with left-recursion doesn't satisfy the second of the above equations characterising addition. The reader should find an equation that it *does* satisfy[31]

One can define very large functions in PRA. Exponentiation can be defined by using the same iterative idea to multiplication that was used to define multiplication in terms of addition. A function quadration can be obtained from exponentiation in the same fashion, where $\mathsf{quadrate}(x, y)$ is a tower of exponentials of x y occurrences of x high.

Nonetheless one can define functions in PRA$^{\omega}$ that are faster-growing than any in PRA: given the following definition of Ackermann's function:

$$\begin{aligned}\mathsf{Ack}(n, x, y) \mathrel{\hat{=}} \mathsf{ev}(F(\mathsf{pred}(n), x), y)\\ \text{where } F(n, x) \mathrel{\hat{=}} \mathsf{Rec}(n,\\ \lambda y^{\mathbb{N}}.\mathsf{add}(x, y),\\ \lambda i^{\mathbb{N}}.\lambda z^{\mathbb{N}\Rightarrow\mathbb{N}}.\lambda y^{\mathbb{N}}.\mathsf{Rec}(y, \underline{1}, \lambda j^{\mathbb{N}}.\lambda u^{\mathbb{N}}.\mathsf{ev}(z, u)))\end{aligned}$$

we have that the function $\lambda n^{\mathbb{N}}.\mathsf{Ack}(n, 3, 2)$ provides just such a dominating function[32].

[31] *Hint:* you will need an auxiliary definition.

[32] See Muhammed Ali McBeth's *Combinatorial Number Theory* [McBethMA:comnt] for an illuminating discussion of this classical result of Ackermann's.

The key to this expressive power lies in our ability to use lambda abstraction to apply a function to itself variably many times by use of higher-type recursion. This has the consequence, however, that cut-formulae of higher-type cannot be re-expressed without implication, and so they are *ineliminable detours*.

Let us now turn to the proof of the soundness of the conversion theory of PRA$^\omega$. We shall not attempt to provide a direct, finitary proof of strong normalisation for PRA$^\omega$ as we did for NJ since, although it is not too difficult to provide such a proof for weak normalisation by using the ordinal assignment scheme of Gentzen, it is much harder to find such an assignment that is reducing under all reductions[33].

So instead we shall appeal to the impredicative method of Tait[34]. Girard has introduced some refinements to this method, and we shall follow the presentation of the proof in chapters 6 and 7 of his *Proofs and Types* [GirardJY:prot].

As this reference is easily obtainable, I shall only outline the skeleton of the proof for comparison with proof we shall give in section 3.4. One point worth emphasising is that their proof does not consider eta expansions, and so we must extend the result which we achieve by an appeal to our relative normalisation lemma. It would not be too difficult to amend the Girard–Tait proof, however, the chief complication would be that we must prove that variables at all types are reducible.

We also note that the choice of recursive form makes little difference to the proof, only slightly affecting our argument when we consider the commutation of eta reductions through $\mathbb{N}\beta$ reductions and when we come to consider Rec term former in the final induction. Consequently we shall assume the right-recursive form throughout this section: when we come to section 3.4 we will be concerned with the left-recursive form.

LEMMA 51 (CHAIN BOUND LEMMA)
If the term s is strongly normalising, then there is an associated numeral $\mathsf{rcb}(s)$ (the *rewrite chain bound* of s), which is the length of the longest reduction sequence starting at s.

PROOF The set of rewrite chains from s form a tree, whose nodes are determined by the common initial subsequences of these chains. Since every term has only finitely many rewrites, the tree is finitely branching, and since we have strong normalisation for s, it has no infinite threads. By König's lemma the tree is finite, and so it has a finite depth, which is the bound rcb(s).

We note that the only property of the reduction system we have appealed to is that there are only finitely many rewrites possible from each term, a property that applies to all the systems of this thesis. □

PROPOSITION 52

1. $\rightarrow^1_\eta$ is strongly normalising.

[33]It is not too difficult to formulate such a measure if we appeal to the 'tree' ordinal Γ_0, but this ordinal is much bigger than Gentzen's ordinal ϵ_0. I would be interested to hear if anyone has such a strategy-independent assignment under ϵ_0.

[34]See his 'Intensional interpretations of functionals of finite type' [TaitWW:intiff].

2. If $s \to^1_\beta t$ and $t \to^1_\eta u$, then there are terms t', u' such that $s \to^1_\eta t'$ and $t' \to^+_\beta u'$, where $u \to^*_\eta u'$;

3. If $s \to^1_\beta t$ and s possesses no eta redexes, then neither does t;

4. If $s \to^1_\beta t \to^*_\eta u$ and u possesses no eta minima, then there is a term t' such that $s \to^*_\eta t' \to^+_\beta u$.

PROOF Part one follows by providing a measure on derivations that reduces with any eta reduction. Let $\Sigma_{v \in V}\, \omega^{\mathsf{deg}(v)}$ be our measure, where V is the set of minimal formulae occurrences of the derivation, with the *degree* $\mathsf{deg}(v)$ defined as in section 1.2, and with $\mathsf{deg}(\mathbb{N}) \mathrel{\hat{=}} 0$. Inspection shows that this measure has the desired property.

Part two is the main property needed for the corollary, and is shown by case analysis. There are six cases to consider, with no overlapping redexes:

1. $(\times\eta)$

   (a) $(\times\beta)$ If the residual of the beta redex is a subterm of the eta redex, then exchanging the orders of the reduction will create two beta reductions where there were one, but $u \equiv u'$. If neither residual or if the eta minima occurs in the beta redex shares a subterm we can commute the reductions exactly;

   (b) $(\supset\beta)$ Except if the eta minima occurs in the beta redex this case is the same as before. However, in this case, the eta minima occurs as part of the substituted term, and so when we commute the terms, there may be many residuals of the contracted redex, where we contracted only one before. In this case, however, we can contract these other eta redexes, obtaining $u \to^+ u'$;

   (c) $(\mathbb{N}\beta)$ As in the *implies* case, there may be many eta redexes corresponding to the one contracted in the original if the eta redex occurs in the third argument of the recursor: it makes no difference whether we adopt the left or right recursive form. We handle this case in the same way;

2. $(\Rightarrow\eta)$ There are three cases to consider here, all of which are similar to that for $\times\eta$, except that we preserve the number of beta reductions when we commute the eta reduction.

Part three follows by a simple case analysis on reductions. Part four follows easily from parts one to three. □

COROLLARY 53 (RELATIVE NORMALISATION)
If $\to^1_\beta$ is strongly normalising, then so is $\to^1_{\beta\eta}$.

PROOF We shall show how, given any finite rewrite chain of $\to^1_{\beta\eta}$, we can associate a finite rewrite sequence of $\to^1_\beta$. Thus, if there is an infinite rewrite chain of

$\rightarrow^1_{\beta\eta}$ starting from some term s, then because there must be only a finite number of reductions from each $\rightarrow^1_\beta$ to the next (due to part one of the above lemma), we can find a rewrite chain of $\rightarrow^1_\beta$ exceeding any given finite bound. But by the chain bound lemma this must mean that s is not strongly normalising.

Now to show how to obtain the associated rewrite chain. We shall restrict our attention to chains ending in an eta normal form, since any chain may be extended to such by part one above. The rewrite chain $s \rightarrow^*_{\beta\eta} v$ may be divided into three sections so:

$$s \rightarrow^*_{\beta\eta} t \rightarrow^*_\eta u \rightarrow^*_\beta v$$

since the last two segments may be empty, and without loss of generality we may consider that the reduction yielding t is a beta reduction.

We associate to each such sequence a new rewrite chain:

$$s \rightarrow^*_{\beta\eta} t_1 \rightarrow^*_\eta u_1 \rightarrow^*_\beta v$$

obtained by commuting across the last beta reduction of the first segment to the third segment. We see by repeating this procedure that the number of beta reductions in the first segment reduces by one each time, and so this sequence must terminate, and we also see that there are at least as many beta reductions in the obtained rewrite chain as in the one before, so this measure is non-decreasing. Thus the third segment contains no fewer beta reductions than were present in the original rewrite chain. □

DEFINITION 54

1. Let $\Gamma \vdash s : A$. We define the predicate $\mathsf{Red}_A(s)$ by structural induction on the type A:
   (a) If A if atomic (that is, it is a schematic letter or the type $\mathbb{N}$), then $\mathsf{Red}_A(s)$ iff s is strongly normalising;
   (b) If $A \equiv B \times B'$ then $\mathsf{Red}_A(s)$ iff $\mathsf{Red}_B(s)$ and $\mathsf{Red}_{B'}(s)$;
   (c) If $A \equiv B \Rightarrow B'$ then $\mathsf{Red}_A(s)$ iff for all terms t such that $\Gamma \vdash t : B$ and $\mathsf{Red}_B(t)$ then $\mathsf{Red}_{B'}(\mathsf{ev}(s,t))$. Note that we are quantifying impredicatively over the set of terms for which $\mathsf{Red}_B(t)$ in defining $\mathsf{Red}_A(s)$ here.
2. A term is *canonical* if it has one of the following forms: $\langle s, s' \rangle$, $\lambda x^A.s$, or $\underline{k}$. It is *non-canonical* if it is not canonical. Girard calls non-canonical terms *neutral* terms.

PROPOSITION 55 Let $\Gamma \vdash s : A$.

1. If $\mathsf{Red}_A(s)$ then s is strongly normalising;
2. If $\mathsf{Red}_A(s)$, and $s \rightarrow^*_\beta s'$, then $\mathsf{Red}_A(s')$;
3. If s is non-canonical, and if for every term s' with $s \rightarrow^1_\beta s'$ we have $\mathsf{Red}_A(s')$, then $\mathsf{Red}_A(s)$.

PROOF We establish these properties for all types A by structural induction on A. The case of $A \equiv \mathbb{N}$ is the same as any other atomic type. It is worth noting that the last case is complicated somewhat by the presence of eta reduction, which is not present in the theory considered by Girard, Lafont and Taylor. □

PROPOSITION 56

1. For any numeral $\underline{k}$, $\mathsf{Red}_{\mathbb{N}}(\underline{k})$;
2. If $\Gamma \vdash \langle s, s' \rangle : A \times A'$ and s, s' are reducible, then so is $\langle s, s' \rangle$;
3. If $\Gamma, x : A \vdash s : B$ and for all terms t such that $\Gamma \vdash t : A$ and $\mathsf{Red}_A(t)$ we have that $s[x := t]$ is reducible.

PROOF Part 1 is immediate. Parts 2 and 3 follow the treatment of Girard, Lafont and Taylor exactly. □

THEOREM 57 All terms are reducible.

COROLLARY 58 The calculus PRA$^\omega$ is strongly normalising.

PROOF We proceed by structural induction on terms. We need a slightly devious induction hypothesis: we show inductively that under arbitrary substitutions of free variables for reducible terms, the term is reducible. □

THEOREM 59 The calculus PRA$^\omega$ is Church–Rosser.

PROOF The only new critical pairs for this calculus are eta redexes on the auxiliary premisses of the new elimination rule, which we resolve in the same manner as for the $\Rightarrow \beta$, and so the Church–Rosser property extends quite painlessly from λ to these calculi. Again the left-recursive form presents no complications. □

Chapter 2

Intuitionistic type theory

2.1 Equality and type dependency

We may summarise the achievements of the theory developed in the last chapter by saying that we have a semantics of intuitionistic propositional logic which satisfies the following desiderata:

1. (Naturality) We can formulate proofs in the logic which represent reasoning naturally, and whose logical rules admit justification by internal methods;

2. (Constructivity) We have an account of the formal semantic content of proofs which makes explicit the concepts required in the justification of the logical rules, and which depends only upon computationally effective mathematics;

3. (Epistemic) We have an account of what it is to know a proposition that is in harmony with the system of justifications used to show naturality.

Naturality is satisfied by Gentzen's NJ calculus along with the Belnap–Prawitz account of logical harmony. Constructivity is satisfied by the lambda calculus and the Curry–Howard correspondence. Finally, the epistemic criteria is satisfied by the correspondence between the lambda calculus and the BHK interpretation of the propositional connectives.

Satisfying though this correspondence is, it falls short of capturing the full strength of Frege's calculus, as it has no account either of equality statements or of quantification. To introduce this, we shall reverse the order of the above account, beginning at the epistemic criteria and ending at a presentation of our formalisation of proof theory, and indeed this process mirrors the historical development.

So we begin by an examination of the BHK interpretation of the quantifiers:

> A proof of $\forall x \in X.A(x)$ consists of a function from elements x of the species X to proofs of $A(x)$.
>
> A proof of $\exists x \in X.A(x)$ consists of an element x_0 of the species X and a proof of $A(x_0)$.

Naïvely, we may observe that the constructor in each case is respectively function space and Cartesian product. However if we attempt to apply this intuition directly using the theory of section 1.4, we obtain:

$$p \in \mathsf{Prf}(\forall x \in X.A(x)) \text{ when } p \in \mathsf{El(X)} \Rightarrow \mathsf{Prf}(A(?))$$
$$p \in \mathsf{Prf}(\exists x \in X.A(x)) \text{ when } p \in \mathsf{El(X)} \times \mathsf{Prf}(A(?))$$

where the '?' indicates a hole in the proposition A(x) which arises when x becomes unbound.

The key to handling this conceptual dependency of the second argument upon the first lies in the account of indexed families which first appears in a mature form in the work of Dana Scott[1].

The essential idea is that a type $A(x)$ where x is a free variable of type X, is modeled by a family of types $\{A_x \mid x \in X\}$. If x_0 is a particular element of X, then $A(x_0) \equiv A_{x_0}$. We may then generalise the type formers $\Rightarrow$, $\times$ as follows:

> The *dependent function space* from X to $Y(x)$ is given by the type former $\Pi x \in X.Y(x)$ and consists of functions f such that for each $x_0 \in X$, $f(x_0) \in Y_{x_0}$.
>
> The *dependent Cartesian product* of X and $Y(x)$ is given by the type former $\Sigma x \in X.Y(x)$ and consists of pairs of $x_0 \in X$ and $y_0 \in Y_{x_0}$.

We may then model the type of proofs of quantifiers as follows:

$$p \in \mathsf{Prf}(\forall x \in X.A(x)) \text{ when } p \in \Pi x \in X.A(x)$$
$$p \in \mathsf{Prf}(\exists x \in X.A(x)) \text{ when } p \in \Sigma x \in X.A(x)$$

Attractive though Scott's paper is, it lacks equality as a propositional form, and so it possesses no concept corresponding to a 'term representing the proof of $a = b$', and so judgements involving equality must be held as falling outside the bounds of what we may call a Curry–Howard style correspondence.

There are two approaches to providing a constructive, or intensional, treatment of equality. It is possible to use Leibniz's rule to encode propositional equality using the device of Russell and Whitehead[2], which is the approach used in the calculus of constructions. However, such an approach depends upon the availability of impredicative universes. This spoils the kind of fine-grained analysis of harmony we wish to conduct, since impredicative quantification is a very non-conservative

[1]See his 'Constructive Validity' [ScottDS:conv]. The idea that one can model the predicate calculus inside indexed families appears to have been discovered independently by F. W. Lawvere, unpublished [LawvereFW:eletcs]. Also dependent types have been formalised around this time by De Bruijn in 'The mathematical language AUTOMATH' [DeBruijnNG:matlai], but not in a higher order sense.

[2]Naturally, the *Principia Mathematica*, [RussellB:prim]. Peter Aczel provides an enlightening description of this encoding in 'The Russell–Prawitz modality', [AczelP:ruspm].

extension of our conceptual framework, and defining equality in terms of impredicative quantification ensures that we cannot isolate its contribution to our framework from that of it.

The alternative is to provide an exact constructive characterisation of equality directly. Normal presentations of equational logic do not provide their rules in a form suitable for our techniques, since they do not separate the introduction and elimination rules for equality statements. Giving such a characterisation is perhaps the main achievement of Martin-Löf's intuitionistic type theory. It distinguishes between three kinds of equality: convertive equality, which is defined by conversion rules analogous to those of the simple theory of constructions; conceptual equality, which captures the intuitively validated notion of equality between terms, and propositional equality, which is a type corresponding to the proposition 'objects s and t are equal'.

This aspect effectively forces itself upon us as the right choice for extending the story we have developed in chapter one to predicate calculus: nowhere else is equality decomposed into its

Martin-Löf also introduces a significant new analysis of what it is to make judgements, and in particular, what it is to make the judgement 'A type', whose importance we shall discuss later. Also his later formulations of his intuitionistic type theory[3] possess a high degree of formal polish. Like the work of Frege, the writings of Martin-Löf are distinguised by the careful thought applied to identifying the concepual significance of the most elementary operations of his calculus.

It is beyond the scope of this work to provide a thorough treatment of Martin-Löf's type theory. Instead our efforts here shall be directed at understanding what Martin-Löf's theory has to say about the extension of the Curry–Howard correspondence to higher-order predicate logic.

The calculus we shall develop we call 'ITT' (for *intuitionistic type theory*). It is essentially a subset of the intensional theory of Martin-Löf[4] sharing the properties:

1. It is a polymorphic system, in the simple sense of Hindley and Milner. Our system of judgements allow us two kinds of basic judgement, and analogously two kinds of assumption. In the terminology of De Bruijn's AUTOMATH we allow degree 1 judgements, introducing type judgements, and degree 2 judgements, introducing term judgements. Our degree 1 assumptions are of the form 'X type' where X is a schematic letter, introducing a simple schematic polymorphism;

2. The theory is predicative. Quantifiers may only range over types that are

[3]The intensional formulation as given in 'Programming in Martin-Löf's Type Theory' [NordstromB:promlt], and the slightly earlier extensional formulation as given in 'Constructive mathematics and computer programming' [MartinLofP:conmcp] and *Intuitionistic Type Theory* [MartinLofP:intttt].

[4]We refer to the theory of 'Constructive mathematics and computer programming' [MartinLofP:conmcp] as the intensional theory, due to the property related to equality, which was not possessed by his earlier theories.

explicitly provided;

3. We support the strong existential quantifier of Bill Howard in 'The formulae-as-types notion of control' [HowardWA:fortnc]. As Howard observed, this allows us to derive the axiom of choice;

4. We support intensional equality. This allows us to give a semantics via a Curry–Howard like correspondence.

However, our account differs from the standard accounts of Martin-Löf's type theory in some respects:

1. The most significant difference is that we do not make use of Martin-Löf's meaning theory for his calculus, due to the fact that his system violates compositionality as discussed in the introduction.

   Our discussion of the adequacy of the framework is composed of three stages: we show that the calculus validates, through natural translations, the expected meaning of the arithmetic rules. Then we show the consistency of the framework by translation into PRA^{ω} (from which we obtain an 'incomplete' completeness result: the Dialectica interpretation tells us that HA and ITT define the same recursive functions). Finally we show that in the case of three of the rules, the expected harmony between the introduction and elimination rules can be given; for the last, that is $\mathbb{N}$ type former used to formalise the natural numbers, we identify the cause of the weakness and provide an appropriate caveat to the applicability of Belnap's conservativity criterion. We shall discuss the significance of this caveat in the conclusions.

   In the last section of the chapter we provide a preliminary sketch of how an alternative semantics might run, by providing an identification of sense with 'two factors' (ie. their decomposition into assertoric and hypothetical contribution) allows our theory of sense to retain compositionality. A critical discussion of this sketch, which falls short of providing all that we should require of a systematic semantics, follows in the conclusions.

2. We give a single calculus for the type theory. Presentations of Martin-Löf's type theory usuall follow one of three different styles. The first is to present the axiomatically, as Martin-Löf's own writings did, with no meta-theoretic demonstration of their adequacy. The other two approaches[5] are in the first case to give a higher order abstract syntax for the calculus, or to present the calculus in terms of an intermediate logical framework. Both of these approaches introduce an additional, non neutral, intermediary into our framework, which disrupts our attempt to give a framework independent treatment of logical harmony.

[5]Both are given in *Programming in Martin-Löf's Type Theory* [NordstromB:promlt].

Whilst not using an independent intermediary in the construction of the calculus, it is helpful to introduce a condensed notation for definitions[6]. This is neutral with respect to the framework, as it is defined explicitly in terms of substitution operations, and is similar to the abstract syntax used in part one of *Programming in Martin-Löf's Type Theory*.

Finally I should note that this is not the only 'disintermediated' presentation of type theory. Healfdene Goguen's PhD thesis[7] presents an alternative approach, of which I was not aware of when I began work on this chapter. A significant difference between his approach and that takes here, is that he introduces conversions explicitly and then synthesises definitional equality, whilst here, we introduce conversions implicitly as definitional equality judgements, and relate it to our account of conversion in chapter one through a synthetic decomposition in section 2.6.

3. Martin-Löf's type theory normally is formulated without use of eta conversions. Since eta expansion proved important in chapter one to our justification of synthetic harmony, we give the conversion here.

   That we can do so at all in our eta-expansion–based meta theory, is due to our implicit treatment of conversion. Neil Ghani has shown[8] that a type theory that introduces the natural conversion rule:

$$\frac{A \to^1 A'}{\lambda x^A.s \to^1 \lambda x^{A'}.s}$$

   will violate strong normalisation. Without this rule, it is not clear that Martin-Löf's type congruence conditions can be satisfied.

4. We have a strictly limited range of types sufficient for arithmetic, covering dependent function space, dependent Cartesian product, equality types and natural numbers. We do not give rules for disjunction, well-orderings, enumeration types or universes, that is to say we give a treatment of the *arithmetic* fragment of the intensional type theory.

5. We give a sequent-style presentation, whose assumptions are explicitly given as they are in the type inference system of section 1.3. This is not of semantic significance, but it allows a closer formal similarity to other comparable formal frameworks such as the Edinburgh LF and the calculus of constructions. As far as is convenient, we shall follow the conventions of De Bruijn's AUTOMATH; our sequents accordingly are called *telescopes*.

6. Also not of semantic significance, we give a systematic form for recursors at each type former $\otimes$, written $\mathrm{R}^{\otimes}(\ldots)$. This is, I think, due to Martin Hoffman[9].

[6] In the *Substitutions and definitions* subsection of the next section.
[7] *A Typed Operational Semantics for Type Theory* [GoguenH:phdthesis].
[8] 'Eta-expansions in dependent type theory' [GhaniN:etaedt].
[9] See 'The syntax and semantics of dependent types' [HofmannM:synsdt].

2.2 The calculus ITT

Formulae and Terms

The valid formulae and terms are those that can be derived from the rules of the calculus that follow. But it is at least possible to define a grammar for *formula candidates* and *term candidates*:

$$
\begin{aligned}
T ::= & X \mid \Pi x \in T.T \mid \Sigma x \in T.T \mid s =_T s \mid \mathbb{N} \\
s ::= & x \\
& \mid \lambda x : T.s \mid \mathsf{ev}(s, s) \\
& \mid \langle s, s \rangle \mid \mathrm{R}^{\Sigma}(s, (x : T, x : T)s) \\
& \mid \mathsf{refl}^{T}(s) \mid \mathrm{R}^{=}(s, (x : T)s) \\
& \mid \mathsf{zero} \mid \mathsf{succ}(s) \mid \mathrm{R}^{\mathbb{N}}(s, s, (x : T, x : T)s) \\
& \text{where } X \text{ ranges over schematic letters and } x \text{ ranges over variable letters}
\end{aligned}
$$

We define the free variables and subterms of term candidates in the usual way. We also define a special function subj from each type candidate to a set of term candidates called the *subjects* of the type inductively as follows:

$$
\begin{aligned}
\mathsf{subj}(\Pi x \in A.B) = \mathsf{subj}(\Sigma x \in A.B) &= \mathsf{subj}(A) \cup \mathsf{subj}(B) \\
\mathsf{subj}(X) = \mathsf{subj}(\mathbb{N}) &= \emptyset \\
\mathsf{subj}(s =_A t) &= \{s, t\}
\end{aligned}
$$

We are also interested in the free variable occurrences of type, which are not simply the free variables of the subjects, as the Π, Σ type formers act as binders analogously to λ. The variables occurring in brackets in the arguments to the recursors, that is the term formers of the form R^{T}, also behave as binders.

In practice we may omit the types of bound variables if the type may be inferred from the context, and we may write '$x : A$' as x^{A}. Furthermore, in term binders, if the bound variable does not occur in the term, we may indicate this by '$(-)$'.

Telescopes

A *type assumption*, or degree 1 assumption, is an expression 'X type' for a schematic letter X. A *term assumption*, or degree 2 assumption, is an expression '$x : A$' where x is a variable and A is a type candidate.

The valid telescopes are those that can be validly derived from the rules of the calculus, but as with terms and types we may introduce *telescope candidates* Γ, where each *assumption* is specified by the grammar $\mathcal{A}$:

$$
\begin{aligned}
\Gamma ::= & \cdot \mid \Gamma, \mathcal{A} \\
\mathcal{A} ::= & x : T \mid X \text{ type}
\end{aligned}
$$

The valid telescopes also possess a number of properties which we describe in propositions 60 and **??** in section 2.3.

Note that telescope candidates are ordered sequences, unlike the telescopes of the lambda calculus. We may define the *domain* of a given telescope, a collection of variable and schematic letters, recursively:

$$\begin{aligned}\mathsf{dom}(\cdot) &\triangleq \emptyset\\ \mathsf{dom}(\Gamma, x : A) &\triangleq \{x\} \cup \mathsf{dom}(\Gamma)\\ \mathsf{dom}(\Gamma, X\ \ \mathsf{type}) &\triangleq \{X\} \cup \mathsf{dom}(\Gamma)\end{aligned}$$

and the *range of a telescope* analogously:

$$\begin{aligned}\mathsf{ran}(\cdot) &\triangleq \emptyset\\ \mathsf{ran}(\Gamma, x : A) &\triangleq \{A\} \cup \mathsf{ran}(\Gamma)\\ \mathsf{ran}(\Gamma, X\ \ \mathsf{type}) &\triangleq \mathsf{ran}(\Gamma)\end{aligned}$$

Judgements and equality

There are two core judgements of the calculus, consisting firstly of the judgement that a type candidate is a type, and secondly the judgement that a term candidate is a term. These judgements are written 'A type' and '$s : A$' respectively.

Martin-Löf has written a lucid exposition of the meanings of these judgements, relating them to his categorical treatment of knowledge (see 'On the meanings of the logical constants and the justifications of the logical laws' [MartinLofP:mealcj]). Briefly we may summarise his treatment, moving from truth as the key concept to constructibility. To say that we know 'A type' is to say that we know what methods may be used to create constructions in A, and dually that we know how to make use of constructions in A in forming new constructions. To say '$s : A$' is to say that s is a construction, and that it is constructed in accordance with the methods appropriate to A.

We also wish to have judgements under assumptions, and it is for this reasons that we formalise de Bruijn's notion of a telescope. The significance of de Bruijn's system of formalisations is that it is the foundation to his analysis of the content of such mathematical commonplaces as definitions, theorems and axioms, and is the key to their automated representation. So far as possible in this work I have tried to follow a 'de Bruijn'-like approach to definitions and the syntax of judgements.

We thus have an auxiliary form of judgement 'Γ tel' which says that the telescope candidate Γ is a telescope. We then use telescopes in the formalisation of *hypothetical judgements*: $\Gamma \vdash \mathcal{J}$, where $\mathcal{J}$ ranges over core judgements. As an example to illustrate hypothetical judgements, we may derive the following judgement in the rules of ITT:

$$X\ \ \mathsf{type}, x : X, y : X \vdash x =_X y\ \ \mathsf{type}$$

This conditional judgement indicates that if we are shown what it is to be a construction of type X, and given two elements of that type, then we know what it is to show that the two elements are equal.

Central to Martin-Löf's type theory also is his treatment of equality, and in this work we shall distinguish between four kinds of equality.

1. Definitional equality: this is the notion of equality at work 'behind the scenes' of the calculus. It is concerned with substitutions and parameterised definitions. Significantly, we do not subsume alpha-conversion under definitional equality, for two reasons. Firstly, we do not provide a higher-order abstract syntax for the calculus, which is usually how alpha-conversion is handled in practice, and secondly, in the presence of type conversions it is not immediately obvious that alpha-conversion is computationally trivial.

2. Convertive equality: this expresses the equality on types and terms that coincides with the compatible, symmetric and transitive closure of the rewrite relation. We reverse the usual priority between the two judgements: we introduce the equality judgement and then move to the rewrite relation at the end of the chapter;

3. Inhabitation of the equality type: this holds when it is possible to find a term inhabiting the type whose meaning is the equality holding between two terms. This is much stronger than convertive equality;

4. Conceptual equality: this is the semantic conception of sameness of terms that guides the definition of the type formers. To Martin-Löf, part of knowing what a construction of a particular type is, is to know when two constructions inhabiting that type are equal. It is intended for inhabitation of the equality type to coincide with conceptual equality.

We shall require judgement forms to handle these notions of equality. Definitional equality is handled by the auxiliary judgement '$s \equiv s'$', which is used in case analysis of terms. Convertive equality is given by three judgements, two matching core judgements and one for telescopes: '$\Gamma \vdash A = A'$ type', '$\Gamma \vdash s = s' : A'$' and '$\Gamma = \Gamma'$ tel'. The equality type can be handled by the core judgement on terms. Finally we shall discuss conceptual equality in section 2.7.

Substitutions and definitions

The theory ITT has two substitution operators, one which replaces type variables by type candidates and one which replaces assertion variables by term candidates. These are '$X := A$' and '$x := s$' respectively, and little more needs to be said about them, except to comment on bound variables. Firstly, the terms R^{Σ} , $\mathrm{R}^{=}$ and $\mathrm{R}^{\mathbb{N}}$ bind variables, secondly that substitution must make use of alpha-conversion (to be defined shortly) in order to respect the bound identifier convention, and thirdly we should note that free type variables may occur in the types associated with bound variables.

We may build a more sophisticated theory of substitutions upon these two operation, the *general substitutions*, which may be seen as maps between telescopes. We introduce a new auxiliary judgement to express this relation: '$\sigma :: \Gamma \to \Gamma'$.

The general substitutions inhabit a grammar, whose elements are the *substitution candidates*:

$$\begin{aligned}\sigma ::= & \mathsf{wk}^{\Gamma}_{\Delta} \mid \mathsf{exp}_{\Delta}(\sigma) \mid \sigma, \sigma \\ & \mid X_{\Gamma} := A \mid x_{\Gamma} := s\end{aligned}$$

We associate with the substitution candidates a mapping on type candidates and term candidates as follows:

$$\begin{array}{rcrcl} A[\mathsf{wk}^{\Gamma}_{\Delta}] & \equiv & A & \qquad & s[\mathsf{wk}^{\Gamma}_{\Delta}] \equiv s \\ A[\mathsf{exp}_{\Delta}(\sigma)] & \equiv & A[\sigma] & & s[\mathsf{exp}_{\Delta}(\sigma)] \equiv s[\sigma] \\ A[\sigma_1, \sigma_2] & \equiv & (A[\sigma_1])[\sigma_2] & & s[\sigma_1, \sigma_2] \equiv (s[\sigma_1])[\sigma_2] \end{array}$$

with the remaining two cases being the familiar operations, the subscripted 'Γ' not affecting their behaviour. We shall also define a substitution operation on hypotheses $\mathcal{A}[\sigma]$, where this is defined only if σ does not act upon the variable $\mathsf{dom}(\mathcal{A})$:

$$\begin{aligned}(X\ \mathsf{type})[\sigma] &\equiv X\ \mathsf{type} \\ (x : A)[\sigma] &\equiv x : (A[\sigma])\end{aligned}$$

We extend the substitution on assumptions to substitution on telescope candidates in the usual way.

The valid substitution judgements are obtained using the following rules:

1. Weakening

$$\frac{\Gamma, \Delta\ \mathsf{tel}}{\mathsf{wk}^{\Gamma}_{\Delta} :: \Gamma \to \Gamma, \Delta}$$

2. Expansion

$$\frac{\sigma :: \Gamma \to \Gamma' \quad \Gamma, \Delta\ \mathsf{tel}}{\mathsf{exp}_{\Delta}(\sigma) :: \Gamma, \Delta \to \Gamma', (\Delta[\sigma])}$$

3. Composition

$$\frac{\sigma_1 :: \Gamma \to \Gamma' \quad \sigma_2 :: \Gamma' \to \Gamma''}{\sigma_1, \sigma_2 :: \Gamma \to \Gamma''}$$

4. Substitution operators

$$\frac{\Gamma \vdash A\ \mathsf{type} \quad X \notin \mathsf{dom}(\Gamma)}{X_{\Gamma} := A :: \Gamma, X\ \mathsf{type} \to \Gamma}$$

$$\frac{\Gamma \vdash s : A \quad x \notin \mathsf{dom}(\Gamma)}{X_{\Gamma} := s :: \Gamma, x : A \to \Gamma}$$

In practice we usually omit the expansion substitution former, the brackets from composition (it is associative), and the subscript from the substitution operators. The interested reader will have no difficulty in obtaining equivalences between various general substitutions and the existence of canonical forms for them.

We shall also use parameterised definitions and contexts for ITT and their definition is analogous to that given in definition 30 of section 1.4. We shall find it convenient to use implicit parameterised definitions in presenting logical rules, which lead a double life: in addition to specifying which derivations may be 'glued' into place, they also specify parameterised definitions used in the other parts of the rule.

EXAMPLE 60
The $\Sigma - \mathrm{c}$ rules is defined as follows:

$$\frac{\Gamma \vdash \langle s, s' \rangle : \Sigma x^A.B \quad \Gamma \vdash C(z : \Sigma x^A.B)\ \mathsf{type} \quad \Gamma, x : A, y : B \vdash t : C(\langle x, y \rangle)}{\Gamma \vdash \mathrm{R}^{\Sigma}(\langle s, s' \rangle, (x^A, y^B)t) = t[y := s', x := s] : C(\langle s, s' \rangle)}\ \Sigma - \mathrm{c}$$

The second premiss may be equivalently rendered '$\Gamma, z : \Sigma x^A.B \vdash C\ \mathsf{type}$'; it also indicated that the expression '$C(\langle x, y \rangle)$' in the third premiss should be understood as '$C[z := \langle x, y \rangle]$', and similarly the expression '$C(\langle s, s' \rangle)$' in the conclusion should be understood as '$C[z := \langle s, s' \rangle]$'.

Structural rules

Many of the rules of the calculus do not have substantial logical content, and so we distinguish these rules, which we call the structural rules, from the rules which are explicitly related to the connectives.

As there are quite a large number of structural rules (21 in all), we organise them into five classes. All but five of the structural rules are concerned with ensuring that convertive equality is well-behaved.

To make the presentation slightly more readable we shall use schematic letters ranging over parameterised definitions in typing judgements. If we write $\Gamma \vdash C(x : A, y : B)\ \mathsf{type}$, then we consider C to range over types in which the variables x and y occur, and instances of $C(s, t)$ are to be interpreted by substitution.

Class 1. Formation of telescopes Recall that . is the empty telescope, and Γ ranges over telescope candidates:

$$\frac{}{\cdot\ \mathsf{tel}}\ \mathrm{tel} - \mathrm{emp} \qquad \frac{}{\cdot = \cdot\ \mathsf{tel}}\ \mathrm{tel} - \mathrm{emp} - \mathrm{eq}$$

$$\frac{\Gamma\ \mathsf{tel} \quad X \notin \mathsf{dom}(\Gamma)}{\Gamma, X\ \mathsf{type\ tel}}\ \mathrm{tel} - \mathrm{ty} \qquad \frac{\Gamma = \Gamma'\ \mathsf{tel} \quad X \notin \mathsf{dom}(\Gamma)}{\Gamma, X\ \mathsf{type\ tel} = \Gamma', X\ \mathsf{type\ tel}}\ \mathrm{tel} - \mathrm{ty} - \mathrm{eq}$$

$$\frac{\Gamma \vdash A\ \mathsf{type} \quad x \notin \mathsf{dom}(\Gamma)}{\Gamma, x : A\ \mathsf{tel}}\ \mathrm{tel} - \mathrm{ass}$$

$$\frac{\Gamma = \Gamma' \ \mathsf{tel} \quad \Gamma \vdash A = A' \ \mathsf{type} \quad x \notin \mathsf{dom}(\Gamma)}{\Gamma, x : A = \Gamma', x : A' \ \mathsf{tel}} \ \mathrm{tel-ass-eq}$$

Class 2. Introduction of assumptions

$$\frac{\Gamma, X \ \mathsf{type}, \Gamma' \ \mathsf{tel}}{\Gamma, X \ \mathsf{type}, \Gamma' \vdash X \ \mathsf{type}} \ \mathrm{hyp-ty} \qquad \frac{\Gamma, x : A, \Gamma' \ \mathsf{tel}}{\Gamma, x : A, \Gamma' \vdash x : A} \ \mathrm{hyp-ass}$$

Class 3. Convertive equality is an equivalence relation

$$\frac{\Gamma \vdash A \ \mathsf{type}}{\Gamma \vdash A = A \ \mathsf{type}} \ \mathrm{eq-ty-r} \qquad \frac{\Gamma \vdash A = B \ \mathsf{type} \quad \Gamma \vdash A = C \ \mathsf{type}}{\Gamma \vdash B = C \ \mathsf{type}} \ \mathrm{eq-ty-st}$$

$$\frac{\Gamma \vdash s : A}{\Gamma \vdash \Gamma \vdash s = s : A} \ \mathrm{eq-tm-r} \qquad \frac{\Gamma \vdash s = t : A \quad \Gamma \vdash s = u : A}{\Gamma \vdash t = u : A} \ \mathrm{eq-tm-st}$$

Class 4. Convertive compatibility

$$\frac{\Gamma \vdash s : A \quad \Gamma \vdash A = A' \ \mathsf{type}}{\Gamma \vdash s : A'} \ \mathrm{ty-conv}$$

$$\frac{\Gamma \vdash s = s' : A \quad \Gamma, x : A \vdash B \ \mathsf{type}}{\Gamma \vdash B[x := s] = B[x := s'] \ \mathsf{type}} \ \mathrm{tm-ty-conv}$$

$$\frac{\Gamma \vdash s = s' : A \quad \Gamma, x : A \vdash t : B}{\Gamma \vdash t[x := s] = t[x := s'] : B[x := s]} \ \mathrm{tm-tm-conv}$$

Class 5. Alpha conversion and type congruence

$$\frac{\Gamma \vdash A = A' \ \mathsf{type} \quad \Gamma, x : A \vdash B = B' \ \mathsf{type} \quad y \notin \mathsf{dom}(\Gamma)}{\Gamma \vdash \Pi x^A.B = \Pi y^{A'}.B'[x := y] \ \mathsf{type}} \ \mathrm{alpha-ty-\Pi}$$

$$\frac{\Gamma \vdash A = A' \ \mathsf{type} \quad \Gamma, x : A \vdash s : B \quad y \notin \mathsf{dom}(\Gamma)}{\Gamma \vdash \lambda x^A.s = \lambda y^{A'}.s[x := y] : \Pi x^A.B} \ \mathrm{alpha-\Pi}$$

$$\frac{\Gamma \vdash A = A' \ \mathsf{type} \quad \Gamma, x : A \vdash B = B' \ \mathsf{type} \quad y \notin \mathsf{dom}(\Gamma)}{\Gamma \vdash \Sigma x^A.B = \Sigma y^{A'}.B'[x := y] \ \mathsf{type}} \ \mathrm{alpha-ty-\Sigma}$$

$$\frac{\Gamma \vdash A = A' \ \mathsf{type} \quad \Gamma, x : A \vdash B(x) = B'(x) \ \mathsf{type} \quad \Gamma, x : A, y : B(x) \vdash s : C(\langle x, y \rangle) \quad u, v \notin \mathsf{dom}(\Gamma)}{\Gamma, z : \Sigma x^A.B(x) \vdash \mathrm{R}^{\Sigma}(z, (x^A, y^{B(x)})s) = \mathrm{R}^{\Sigma}(z, (u^{A'}, v^{B'(u)})s[v := y, u := x]) : C(z)} \ \mathrm{alpha-\Sigma}$$

$$\frac{\Gamma \vdash A = A' \ \mathsf{type} \quad \Gamma \vdash \mathrm{R}^{=}(p, (a^A)f) : C \quad x \notin \mathsf{dom}(\Gamma)}{\Gamma \vdash \mathrm{R}^{=}(p, (x^{A'})f[a := x]) : C} \ \mathrm{alpha-=}$$

$$\frac{\Gamma, i : \mathbb{N} \vdash C(i) = C'(i)\ \mathsf{type} \quad \Gamma \vdash \mathrm{R}^{\mathbb{N}}(s, a, (i^{\mathbb{N}}, z^{C(i)})f) : C(s) \quad j, w \notin \mathsf{dom}(\Gamma)}{\Gamma \vdash \mathrm{R}^{\mathbb{N}}(s, a, (i^{\mathbb{N}}, z^{C}(i))f) = \mathrm{R}^{\mathbb{N}}(s, a, (j^{\mathbb{N}}, w^{C'(j)})f[i := j, z := w]) : C(s)}\ \mathrm{alpha} - \mathbb{N}$$

Logical and arithmetic rules

In the subsystem of Martin-Löf's calculus that we shall be using in this section, there are four type formers, $\Pi(-,-), \Sigma(-,-), =_T$ and $\mathbb{N}$. Each type former has four kinds of rule:

1. Formation rules. These describe when we are permitted to infer 'A type' for a formula A governed by the connective $\otimes$. The rule is written $\otimes - \mathrm{f}$;

2. Introduction and elimination rules. These describe the conditions under which we can form terms governed by either a constructor or recursor of the respective type. They are annotated $\otimes - \mathrm{i}$ and $\otimes - \mathrm{e}$ respectively;

3. Conversion rules describe when we may form the equality judgements describing convertive equality at the given type. They are annotated $\otimes - c$.

We recall the discussion of parameterised definitions which are used in the elimination and conversion rules of the Σ, $\mathbb{N}$ and $=_A$ type formers.

Π type former

$$\frac{\Gamma \vdash A\ \mathsf{type} \quad \Gamma, x : A \vdash B\ \mathsf{type}}{\Gamma \vdash \Pi x^A.B\ \mathsf{type}}\ \Pi - \mathrm{f}$$

$$\frac{\Gamma, x : A \vdash s : B}{\Gamma \vdash \lambda x^A.s : \Pi x^A.B}\ \Pi - \mathrm{i}$$

$$\frac{\Gamma \vdash f : \Pi x^A.B \quad \Gamma \vdash t : A}{\Gamma \vdash \mathsf{ev}(f, t) : B[x := t]}\ \Pi - \mathrm{e}$$

$$\frac{\Gamma, x : A \vdash s : B \quad \Gamma \vdash t : A}{\Gamma \vdash \mathsf{ev}(\lambda x^A.s, t) = s[x := t] : B[x := t]}\ \Pi - \mathrm{c}$$

$$\frac{\Gamma \vdash s : \Pi x^A.B \quad y \notin \mathsf{dom}(\Gamma)}{\Gamma \vdash s = \lambda y^A.\mathsf{ev}(s, y) : \Pi x^A.B}\ \Pi - \mathrm{c} - \mathrm{eta}$$

We have discussed the motivation for the Π type former as a generalisation of the $\supset$ connective. We note that type dependency affects the form of the rules barely at all: the principal difference between this account and that of $\supset$ is the need for the $\Pi - f$ rule.

The Π type former governs the only collection of rules whose elimination rule is not of the standard recursor form. This is because the usual 'funsplit' form depends

upon[10] having higher-order assumptions, and we only have a first-order abstract syntax.

Σ type former

$$\frac{\Gamma \vdash A \ \mathsf{type} \quad \Gamma, x : A \vdash B \ \mathsf{type}}{\Gamma \vdash \Sigma x^A.B \ \mathsf{type}} \Sigma - \mathrm{f}$$

$$\frac{\Gamma \vdash B(x : A) \ \mathsf{type} \quad \Gamma \vdash s : A \quad \Gamma \vdash t : B(s)}{\Gamma \vdash \langle s, t \rangle : \Sigma x^A.B(x)} \Sigma - \mathbf{i}$$

$$\frac{\Gamma \vdash s : \Sigma x^A.B \quad \Gamma \vdash C(z : \Sigma x^A.B) \quad \Gamma, x : A, y : B \vdash t : C(\langle x, y \rangle)}{\Gamma \vdash \mathrm{R}^{\Sigma}(s, (x^A, y^B)t) : C(s)} \Sigma - \mathrm{e}$$

$$\frac{\Gamma \vdash \langle s, s' \rangle : \Sigma x^A.B \quad \Gamma \vdash C(z : \Sigma x^A.B) \quad \Gamma, x : A, y : B \vdash t : C(\langle x, y \rangle)}{\Gamma \vdash \mathrm{R}^{\Sigma}(\langle s, s' \rangle, (x^A, y^B)t) = t[y := s', x := s] : C(\langle s, s' \rangle)} \Sigma - \mathrm{c}$$

$$\frac{\Gamma \vdash s : \Sigma x^A.B}{\Gamma \vdash s = \langle \mathrm{R}^{\Sigma}(s, (x^A, y^B)x), \mathrm{R}^{\Sigma}(s, (x^A, y^B)y) \rangle : \Sigma x^A.B} \Sigma - \mathrm{c} - \mathrm{eta}$$

We note that it is not in general possible to infer the type of a term constructed by the $\Sigma - \mathrm{i}$ rule, due to the existence of many incomplete types $B(x : A$. For example, suppose we know that $\Gamma \vdash a : A$ and $\Gamma \vdash d : \phi(a, a)$. Then it is possible to infer either $\Gamma \vdash \langle a, d \rangle : \exists x^A.\phi(x, x)$ or $\Gamma \vdash \langle a, d \rangle : \exists x^A.\phi(x, a)$.

We also note that it is possible to give a quite different form to the elimination and conversion rules, which in the presence of the Π type former is equivalent to the current formulation:

$$\frac{\Gamma \vdash s : \Sigma x^A.B}{\Gamma \vdash \mathsf{outl}(s) : A} \Sigma - \mathrm{e} - \mathsf{outl}$$

$$\frac{\Gamma \vdash s : \Sigma x^A.B}{\Gamma \vdash \mathsf{outr}(s) : B(\mathsf{outl}(s))} \Sigma - \mathrm{e} - \mathsf{outr}$$

$$\frac{\Gamma \vdash a : A \quad \Gamma \vdash b : B(a)}{\Gamma \vdash \mathsf{outl}(\langle a, b \rangle) = a : A} \Sigma - \mathrm{c} - \mathsf{outl}$$

$$\frac{\Gamma \vdash a : A \quad \Gamma \vdash b : B(a)}{\Gamma \vdash \mathsf{outr}(\langle a, b \rangle) = b : B(a)} \Sigma - \mathrm{c} - \mathsf{outr}$$

$$\frac{\Gamma \vdash s : \Sigma x^A.B}{\Gamma \vdash s = \langle \mathsf{outl}(s), \mathsf{outr}(s) \rangle} \Sigma - \mathrm{c} - \mathrm{eta}'$$

[10] In the usual abstract syntax presented in modern expositions. It is used in *Programming in Martin-Löf's type theory* [NordstromB:promlt].

The equivalence is shown by means of dual encodings.

1. If $\Gamma \vdash s : \Sigma x^A.B$
   (a) Let $\mathsf{outl}(s) \triangleq \mathrm{R}^{\Sigma}(s, (x^A, y^B)x)$
   (b) and $\mathsf{outr}(s) \triangleq \mathrm{R}^{\Sigma}(s, (x^A, y^B)y)$.
2. $\mathrm{R}^{\Sigma}(s, (x^A, y^B)t) \triangleq \mathsf{ev}(\lambda x^A.\lambda y^B.t, \mathsf{outl}(s), \mathsf{outr}(s))$.

The reader can easily verify that the definitions have the correct typings, and the conversion rules associated with each form are induced by the conversion rules of the other form, with appeal to the conversions of Π in the second case.

$\mathbb{N}$ type former

$$\frac{\Gamma\ \mathsf{tel}}{\Gamma \vdash \mathbb{N}\ \mathsf{type}}\ \mathbb{N}-\mathrm{f}$$

$$\frac{\Gamma\ \mathsf{tel}}{\Gamma \vdash \mathsf{zero} : \mathbb{N}}\ \mathbb{N}-\mathrm{i}(0) \qquad \frac{\Gamma \vdash n : \mathbb{N}}{\Gamma \vdash \mathsf{succ}(n) : \mathbb{N}}\ \mathbb{N}-\mathrm{i}(+)$$

$$\frac{\Gamma \vdash s : \mathbb{N} \quad \Gamma \vdash C(n : \mathbb{N})\ \mathsf{type} \quad \Gamma \vdash a : C(\mathsf{zero}) \quad \Gamma, m : \mathbb{N}, x : C(m) \vdash t : C(\mathsf{succ}(m))}{\mathrm{R}^{\mathbb{N}}(s, a, (m^{\mathbb{N}}, x^{C(m)})t) : C(s)}\ \mathbb{N}-\mathrm{e}$$

$$\frac{\Gamma \vdash C(n : \mathbb{N})\ \mathsf{type} \quad \Gamma \vdash a : C(\mathsf{zero}) \quad \Gamma, m : \mathbb{N}, x : C(m) \vdash t : C(\mathsf{succ}(m))}{\mathrm{R}^{\mathbb{N}}(\mathsf{zero}, a, (m^{\mathbb{N}}, x^{C(m)})t) = a : C(\mathsf{zero})}\ \mathbb{N}-\mathrm{c}-0$$

$$\frac{\Gamma \vdash s : \mathbb{N} \quad \Gamma \vdash C(n : \mathbb{N})\ \mathsf{type} \quad \Gamma \vdash a : C(\mathsf{zero}) \quad \Gamma, m : \mathbb{N}, x : C(m) \vdash t : C(\mathsf{succ}(m))}{\mathrm{R}^{\mathbb{N}}(\mathsf{succ}(s), a, (m^{\mathbb{N}}, x^{C(m)})t) = t[x := \mathrm{R}^{\mathbb{N}}(s, a, (m, x)t), m := s] : C(\mathsf{succ}(s))}\ \mathbb{N}-\mathrm{c}-\mathsf{succ}$$

We have already encountered a close relative of these rules in the system PRA^{ω}. It repays careful attention to the form of the elimination rule $\mathbb{N} - e$ to make sure that we see how it corresponds to proof by induction, and in particular how the third argument captures the idea of the induction step.

As we mentioned in section 1.4, it is one of the achievements of modern constructive logic that it shows how the explicit formal content of proof by induction can be understood as recursion.

$=_A$ type former The $=_A$ type, or *propositional equality* type, is subscripted to avoid confusion with the symbol indicating definitional equality. This type former is the only type former in this system that actually introduces type dependency.

$$\frac{\Gamma \vdash s : A \quad \Gamma \vdash t : A}{\Gamma \vdash s =_A t\ \mathsf{type}} = -\mathrm{f}$$

$$\frac{\Gamma \vdash s = t : A}{\Gamma \vdash \mathsf{refl}(s) : s =_A t} = -\mathrm{i}$$

$$\frac{\Gamma \vdash p : s =_A t \quad \Gamma \vdash C(x : A, y : A, z : x =_A y)\ \mathsf{type} \quad \Gamma, x : A \vdash d : C(x, x, \mathsf{refl}(x))}{\Gamma \vdash \mathrm{R}^{=}(p, (x^A)d) : C(s, t, p)} = -\mathrm{e}$$

$$\frac{\Gamma \vdash s = t : A \quad \Gamma \vdash C(x : A, y : A, z : x =_A y)\ \mathsf{type} \quad \Gamma, x : A \vdash d : C(x, x, \mathsf{refl}(x))}{\Gamma \vdash \mathrm{R}^{=}(\mathsf{refl}(s), (x^A)d) = d[x := s] : C(s, t, \mathsf{refl}(s))} = -\mathrm{c}$$

There are two points to be aware of. Firstly, the decidability of the judgement $\Gamma \vdash \mathsf{refl}(s) : t =_A t'$ is in general equivalent to the decidability of the conversion theory, due to the rule $\mathrm{tm} - \mathrm{tm} - \mathrm{conv}$. Secondly, the syntax of the recursor as it stands is type ambiguous, since there may be many posible incomplete types $C(x : A, y : A, z : x =_A y)$ that may match an instance $C(s, t, \mathsf{refl}(s))$. Thus to obtain decidability of the term judgement, we would need to add a third type-valued clause to the recursor, which we omit only for the sake of brevity.

2.3 Type soundness

The calculus ITT has a much more complex syntax than that of λ, and we correspondingly find providing an account of its soundness to be a more complex affair.

The principal complication is the complex relationship that exists between the classes of judgement. Our account of the theory of λ proceeded in a linear fashion: first we introduced the types, then we introduced an account of the well-formed terms inhabiting these types, and finally we introduced an account of equality as a relation on terms defined in terms of a reduction semantics. In ITT we have to introduce all of these ideas simultaneously, as the set of well-formed types depends upon the set of well-formed terms, due to the $= -\mathrm{f}$ rule, and the set of well-formed terms depends upon the valid conversions, due to the $= -\mathrm{i}$ rule.

Consequently we cannot at the outset say that it decidable whether a given judgement is valid or not. Thus in this section we must make do with weaker properties, deferring issues of decidability until we have a more fully fleshed-out calculus.

PROPOSITION 61

1. If Γ, Γ' tel then Γ tel;
2. If $\Gamma, \mathcal{A}, \Gamma'$ tel then $\mathsf{dom}(\mathcal{A}) \notin \mathsf{dom}(\Gamma) \cup \mathsf{dom}(\Gamma')$;
3. (Weakening) If $\Gamma, \Gamma'' \vdash \mathcal{J}$ and $\Gamma, \Gamma', \Gamma''$ tel then $\Gamma, \Gamma', \Gamma'' \vdash \mathcal{J}$;
4. (Exchange) If $\Gamma, \mathcal{A}, \mathcal{A}', \Gamma' \vdash \mathcal{J}$, and $\Gamma, \mathcal{A}'$ tel then $\Gamma, \mathcal{A}', \mathcal{A}, \Gamma' \vdash \mathcal{J}$.
5. If $\Gamma \vdash s : A$ or $\Gamma \vdash s = s' : A$ then s satisfies the bound identifier convention;

PROOF The first four parts can be established by easy inductions on the length of telescopes. The last follows from a simple structural induction on term forming inferences. □

LEMMA 62 (SUBTERM LEMMA)
Let d justify $\Gamma \vdash s : A$, and let $s' \in \mathsf{st}(s)$. Then there is a telescope candidate Γ' and a type candidate A such that $\Gamma, \Gamma' \vdash s' : A'$ is the conclusion of some sub-inference of d.

PROOF This is established by transitivity of the subterm relation and a simple structural induction on subderivations of d. □

LEMMA 63 (SUBSTITUTION LEMMA)
Let $\Gamma \vdash \mathcal{J}$ where $\mathcal{J}$ is a core judgement or an equational judgement. If $\sigma :: \Gamma \to \Gamma'$ then $\Gamma' \vdash \mathcal{J}[\sigma]$.

PROOF Reflection upon the definition of a general substitution shows that the substitution lemma follows for all substitutions if it follows for all substitutions of one of the two forms:

$$\sigma = \mathsf{exp}_{\Delta}(X_{\Gamma} := A)$$
$$\sigma = \mathsf{exp}_{\Delta}(x_{\Gamma} := s)$$

Each of these two cases can be established by a structural induction over the type A and the term s respectively. □

The next proposition takes a perhaps slightly surprising form; the explanation is that if there is a conversion which does not satisfy the subject reduction property, then it will be formulated in the calculus using one of the $\otimes - \mathrm{c}$ rules, and consequently the term on the right hand side of the equality sign will not be typable. The intuition is that the equality judgement is at least as strong as the beta equality relation induced by the conversion rules and compatible closure: that this is so you should see follows from the substitution. Consequently a counter-exmple to subject reduction yields a counter-example to this proposition. The proof of confluence in section 2.6 will make this intuition rigorous, and improve the rleation between the equality judgement and beta equality from 'if' to 'if and only if'.

PROPOSITION 64 (SUBJECT REDUCTION)

1. If $\Gamma \vdash A = A'$ type then $\Gamma \vdash A$ type and $\Gamma \vdash A'$ type;
2. If $\Gamma \vdash s = s' : A$ then $\Gamma \vdash s : A$ and $\Gamma \vdash s' : A$.

PROOF Because the proofs of parts 1, 2 and 3 will depend upon parts 2, 3 and 1 respectively, we must establish them in a single induction. The induction will be a structural induction on the derivation justifying the judgement forming the conditional of each part; the induction hypothesis will be that all subderivations whose conclusion matches the form of one of the conditions, then the whole clause applies.

In each part, the proof will be established by a case analysis of the last rule of the derivation. For the first part, if the last rule is:

1. $\mathrm{eq} - \mathrm{ty} - \mathrm{r}$: immediate in each case.

2. $\mathrm{eq-ty-st}$: applying the induction hypothesis to each premiss in turn yields the dsired conclusion in each case.

3. $\mathrm{tm-ty-conv}$: apply 63(2) to the first premiss. The two results follow from applying the substitution lemma in each case.

4. $\mathrm{alpha-conv}$: For $\mathrm{alpha-ty-\Pi}$, the conclusion is given y an application of a $\Pi-\mathrm{f}$ whose first premiss is obtained by the induction hypothesis (this part), and whose second premiss is identical. The case of $\mathrm{alpha-ty-\Sigma}$ is analogous.

And for the second part:

1. $\mathrm{eq-tm-r}$, $\mathrm{eq-tm-st}$: exactly analogous to $\mathrm{eq-ty-r}$ and $\mathrm{eq-ty-st}$ respectively.

2. $\mathrm{tm-tm-conv}$:

   (a) To establish $\Gamma \vdash t[x := s] : B[x := s]$ the induction hypothesis may be applied to the first premiss, from which our conclusion follows by the substitution lemma.

   (b) To establish $\Gamma \vdash t[x := s] : B[x := s]$, we can apply the above argument to obtain $\Gamma \vdash t[x := s'] : B[x := s]$. By applying the induction hypothesis (third part) to the given premiss $\Gamma, x : A \vdash t : B$ we obtain $\Gamma, x : A \vdash B$ type, which is the missing premiss to $tm - ty - conv$, giving us $\Gamma \vdash B[x := s] = B[x := s']$. The goal then follows by $\mathrm{ty-conv}$.

3. The justification of subject reduction for inferences ending in an alpha conversion follow the same general form, and so we shall only consider the case of $\mathrm{alpha-\Pi}$. Consider the general form:

$$\frac{\Gamma \vdash A = A' \ \mathsf{type} \quad \Gamma, x : A \vdash s : B \quad y \notin \mathsf{dom}(\Gamma)}{\Gamma \vdash \lambda x^A.s = \lambda y^{A'}.s[x := y] : \Pi x : A.B}$$

   The judgement $\Gamma \vdash \lambda x^A.s : \Pi x : A.B$ is an immediate consequence of the second premiss, and so it remains to show that $\Gamma \vdash \lambda y^{A'}.s[x := y] : \Pi x : A.B$, which can be easily inferred from $\Gamma, y : A' \vdash \lambda y^{A'}.s[x := y] : B[x := y]$. We produce an inference justifying this conclusion by replacing every occurrence of $\Gamma, x : A, \Gamma' \vdash x : A$ by the inference:

$$\frac{\Gamma, y : A', \Gamma' \vdash y : A' \quad \dfrac{\Gamma \vdash A = A' \ \mathsf{type}}{\Gamma, y : A', \Gamma' \vdash A = A' \ \mathsf{type}}}{\Gamma, y : A', \Gamma' \vdash y : A} \ \mathrm{ty-conv}$$

   which uses the weakening lemma proved earlier. The correctness of this justification follows by an elementary induction.

4. Subject reduction is justified for equational judgements arising from the logical conversion rules by a similar rearrangement of derivations to that used to justify the admissibility of the rewrites for natural deduction.

Lastly we consider the cases that can justify a judgement of the form $\Gamma \vdash s : A$:

1. $\mathrm{hyp} - \mathrm{ass}$: From 60(1) we obtain the validity of the judgement $\Gamma, x : A$ tel. But the only rule that can justify this judgement is $\mathrm{tel} - \mathrm{ass}$ to which $\Gamma \vdash A$ type is a premiss.

2. $\mathrm{ty} - \mathrm{conv}$: follows by applying the induction hypothesis (first part) to the second premiss.

3. An introduction rule: follows by the induction hypothesis and the type formation rule.

4. An elimination rule: these all follow the same scheme, so we just treat the Π rule.

   (a) $\Pi - \mathrm{e}$: apply the induction hypothesis to the first premiss, to obtain a derivation whose last rule is $\Pi - \mathrm{f}$. Apply the substitution lemma to the second premiss of the rule together with the second premiss of the elimination rule to obtain the conclusion.

   (b) $\Sigma - \mathrm{e}, = -\mathrm{e}$: these follow the same scheme, ie. apply the induction hypothesis to the first premiss to obtain a derivation whose last rule is the formation rule; the conclusion follows from as an instance of the substitution lemma making use of the premisses of this formation rule and of the original elimination rule.

   (c) $\mathbb{N} - \mathrm{e}$: We apply the substitution rule directly to the first and second premisses.

□

PROPOSITION 65

1. If $\Gamma \vdash A$ type and B is a subformula of A, then for some telescope candidate Γ' $\Gamma, \Gamma' \vdash B$ type.

2. If $\Gamma = \Gamma'$ tel then Γ tel and Γ' tel.

3. If $\Gamma \vdash \mathcal{J}$ and $\Gamma = \Gamma'$ tel then $\Gamma' \vdash \mathcal{J}$.

PROOF The first part is shown in the same way as lemma 60. The second part is shown by a simple structural induction; we need to appeal to subject reduction in the case of the two $\mathrm{alpha} - \mathrm{ty} - *$ rules. Part three follows by the kind of simple in induction we used to prove the parts of proposition 60.

Part four is syntactically tricky to establish, but essentially follows from the observation that from the justification of the judgement $\Gamma = \Gamma'$ tel we can recover a number of type equalities, and by judicious use of $\mathrm{ty} - \mathrm{conv}$ together with an induction on the length of telescopes we may transform an inference of $\Gamma \vdash \mathcal{J}$ into one justifying $\Gamma' \vdash \mathcal{J}$. □

2.4 Representability

It is the aim of this section to show that ITT is capable of expressing several familiar constructive theories, namely minimal logic with equality, the systems PRA and PRA$^\omega$ of section 1.4, and Heyting arithmetic, and also to show how Martin-Löf's expressive type structure can be used to make the representation of formulae simpler than usual first-order logic. Before we do, let us consider a little more the identification of formulae with types.

The judgements set, prop and true In λ it is one of the consequences of the Curry–Howard correspondence that the formal distinction between proofs of some proposition, and constructions inhabiting some type, becomes almost invisible and so it may be considered either a formulation of minimal logic, or the foundations of the 'logic-free' PRA$^\omega$. With ITT the situation is a little more complex, since we have both types which can only be considered to contain constructions (such as $\mathbb{N}$ and $\Pi x \in \mathbb{N}.\mathbb{N}$) which we call sets, and types that can only be considered to be propositions (ie. the various equality types).

Without wishing to compromise the *underlying* identification of sets and propositions, let us define two kinds of *synthetic judgements*, that is judgements whose meaning derives directly from that of another judgement. These are A set and A prop, and they are defined as follows.

'A set', 'A prop' judgements

The sets and the propositions are generated inductively by the following rules:

$$\frac{\Gamma \vdash A \text{ set} \quad \Gamma \vdash B \text{ set} \quad x \notin \text{dom}(\Gamma)}{\Gamma \vdash \Pi x^A.B \text{ set}} \Pi - \text{set} - \text{f} - \text{alt}$$

$$\frac{\Gamma \vdash A \text{ set} \quad \Gamma \vdash B \text{ set} \quad x \notin \text{dom}(\Gamma)}{\Gamma \vdash \Sigma x^A.B \text{ set}} \Sigma - \text{set} - \text{f} - \text{alt}$$

$$\frac{\Gamma \text{ tel}}{\Gamma \vdash \mathbb{N} \text{ set}} \mathbb{N} - \text{set} - \text{f}$$

If $\Pi x^A.B$ set or $\Sigma x^A.B$ set can be derived by these two rules, then we may use the type synonyms $A \Rightarrow B \mathrel{\hat{=}} \Pi x^A.B$ and $A \times B \mathrel{\hat{=}} \Sigma x^A.B$, observing that there can be no type dependency in types A for which $\Gamma \vdash A$ set. Furthermore, all sets are inhabited.

and

$$\frac{\Gamma \text{ tel} \quad X \text{ type} \in \Gamma}{\Gamma \vdash X \text{ prop}} \text{hyp} - \text{prop} - \text{f}$$

$$\frac{\Gamma \vdash A \text{ set} \quad \Gamma, x : A \vdash B \text{ prop}}{\Gamma \vdash \Pi x^A.B \text{ prop}} \Pi - \text{prop} - \text{f}$$

$$\frac{\Gamma \vdash A \ \mathsf{set} \quad \Gamma, x : A \vdash B \ \mathsf{prop}}{\Gamma \vdash \Sigma x^A.B \ \mathsf{prop}} \Sigma - \mathrm{prop} - \mathrm{f}$$

$$\frac{\Gamma \vdash s =_A t \ \mathsf{type}}{\Gamma \vdash s =_A t \ \mathsf{prop}} = -\mathrm{prop} - \mathrm{f}$$

If $\Pi x^A.B$ prop or $\Sigma x^A.B$ prop can be derived by the above rules, then we may use the type synonyms $\forall x^A.B \triangleq \Pi x^A.B$ and $\exists x^A.B \triangleq \Sigma x^A.B$.

The above inference rules are not adequate to model conjunction and implication and their analogues in set, and so we also admit the rules:

$$\frac{\Gamma \vdash A \ \mathsf{prop} \quad \Gamma \vdash B \ \mathsf{prop} \quad x \notin \mathsf{dom}(\Gamma)}{\Gamma \vdash \Pi x^A.B \ \mathsf{prop}} \Pi - \mathrm{prop} - \mathrm{f} - \mathrm{alt}$$

$$\frac{\Gamma \vdash A \ \mathsf{prop} \quad \Gamma \vdash B \ \mathsf{prop} \quad x \notin \mathsf{dom}(\Gamma)}{\Gamma \vdash \Sigma x^A.B \ \mathsf{prop}} \Sigma - \mathrm{prop} - \mathrm{f} - \mathrm{alt}$$

If $\Pi x^A.B$ prop or $\Sigma x^A.B$ prop can be derived by these two rules, then we may use the type synonyms $A \supset B \triangleq \Pi x^A.B$ and $A \wedge B \triangleq \Sigma x^A.B$, for fresh x. Clearly, if $\Gamma \vdash A$ set or $\Gamma \vdash A$ prop then $\Gamma \vdash A$ type

Finally we introduce the synthetic judgement 'A true', which is introduced by the rule:

$$\frac{\Gamma \vdash s : A \quad \Gamma \vdash A \ \mathsf{prop}}{\Gamma \vdash A \ \mathsf{true}}$$

Equational logic Now let us turn to the representation of minimal logic with equality in ITT. For the sake of simplicity, let us assume that there is only one kind, no relations other than equality, no constants (they can be handled using free variables or 0-ary function letters), and where the only logical connectives are $\supset$ and $\forall x.-$. The terms are then the variables, and the saturated expressions of the form $f(s_1, \ldots, s_n)$ where the s_i are other terms.

The axiom schemes of equational logic are then given, assuming the usual conventions regarding association of '$\supset$' and the function FV to be defined:

1. $\phi \supset \psi \supset \phi$;
2. $(\phi \supset (\psi \supset \chi)) \supset (\phi \supset \psi) \supset \phi \supset \chi$;
3. If $x \notin \mathsf{FV}(\phi)$ then $(\phi \supset \psi) \supset \phi \supset \forall x.\psi$;
4. If s is any term, then $\forall x.\phi \supset \phi[x := s]$;
5. $x = x$;
6. $x = z \supset y = z \supset x = y$;

where ϕ, ψ, etc. range over the formulae of equational logic. There is a single rule, modus ponens.

The only difficulty facing the interpretation of this theory is deciding upon how to formulate the translation. We shall choose a set to serve as the type of terms, say $\mathbb{N}$, and we shall translate function letters as free variables of type $\mathbb{N} \supset \ldots \supset \mathbb{N}$.

DEFINITION 66 Define the translation $[\![\cdot]\!]$ on the terms and formulae of equational logic as follows:

$$\begin{aligned}
[\![x]\!] &\hat{=} x \\
[\![f(s_1, \ldots, s_n)]\!] &\hat{=} \mathsf{ev}(f, [\![s_1]\!], \ldots, [\![s_n]\!]) \\
[\![s = t]\!] &\hat{=} [\![s]\!] =_{\mathbb{N}} [\![t]\!] \\
[\![\phi \supset \psi]\!] &\hat{=} [\![\phi]\!] \supset [\![\psi]\!] \\
[\![\forall x.\phi]\!] &\hat{=} \forall x^{\mathbb{N}}.[\![\phi]\!]
\end{aligned}$$

We must first show that the type $=_A$ satisfies some properties:

PROPOSITION 67

1. (a) If $\Gamma \vdash s : A$ then $\Gamma \vdash s =_A s$ true;
   (b) If $\Gamma \vdash s =_A t$ true then $\Gamma \vdash t =_A s$ true;
   (c) If $\Gamma \vdash s =_A t$ true and $\Gamma \vdash t =_A u$ true then $\Gamma \vdash s =_A u$ true;
2. If $\Gamma \vdash B(x : A)$ prop, then $\Gamma, p : s =_A t \vdash B(s) \supset B(t)$ true;
3. If $\Gamma \vdash f : \Pi x^A.B(x)$ then $\Gamma \vdash s =_A t \supset \mathsf{ev}(f, s) =_{B(s)} \mathsf{ev}(f, t)$ true.

PROOF

1. (a) By reflexivity of definitional equality and $= -$i;
   (b) We have $\Gamma \vdash p : s =_A t$. Let $C(x, y) \hat{=} y =_A x$. Then we may derive

$$\frac{\Gamma \vdash p : s =_A t \quad \Gamma, x : A \vdash \mathsf{refl}(x) : C(x, x)}{\Gamma \vdash \mathrm{R}^{=}(p, (x)\mathsf{refl}(x)) : C(s, t)}$$

   and we are done.

   (c) We have $\Gamma \vdash p_0 : s =_A t$ and $\Gamma \vdash p_1 : t =_A u$. Let $C(x, y) \hat{=} y =_A u \supset x =_A u$. Then $x : A \vdash \lambda y^{x =_A u}.y : C(x, x)$. And so we have

$$\frac{\dfrac{\Gamma \vdash p_0 : s =_A t \quad \Gamma, x : A \vdash \lambda y^{x =_A u}.y : C(x, x)}{\Gamma \vdash \mathrm{R}^{=}(p_0, (x)\lambda y^{x =_A u}.y) : C(s, t) \equiv t =_A u \supset s =_A u} \quad \Gamma \vdash p_1 : t =_A u}{\Gamma \vdash \mathsf{ev}(\mathrm{R}^{=}(p_0, (x)\lambda y.y), p_1) : s =_A u}$$

2. We have $\Gamma \vdash B(x : A)$ prop. Let $C(x, y) \hat{=} B(x) \supset B(y)$. The $\Gamma, x : A \vdash \lambda z^{B(x)}.z : C(x, x)$. So by the hyp $-$ d2 and $= -$e we have $\Gamma, p : s =_A t \vdash \mathrm{R}^{=}(p, (x : A)\lambda z^{B(x)}.z : b(s) \supset B(t)$.

3. Let $C(x, y) \hat{=} x =_A y \supset \mathsf{ev}(f, x) =_A \mathsf{ev}(f, y)$. Then $\Gamma, x : A \vdash \lambda z^{x =_A x}.\mathsf{refl}(\mathsf{ev}(f, x)) : C(x, x)$. So we apply $= -$e and $\Pi -$ e to obtain $\lambda p^{s =_A t}.\mathsf{ev}(\mathrm{R}^{=}(p, (x), \lambda z.\mathsf{refl}(\mathsf{ev}(f, z))), p) : s =_A t \supset \mathsf{ev}(f, s) =_B \mathsf{ev}(f, t)$.

□

COROLLARY 68 If ϕ is a theorem of equational logic, then $\Gamma \vdash_{\text{ITT}} [\![\phi]\!]\ \mathsf{true}$, where Γ consists only of enough set-valued assumptions to bind the variables and function letters of ϕ. Furthermore it is a theorem of the subsystem of ITT with only the Π and $=$ type formers, and the rule $\mathbb{N} - \mathrm{f}$.

PROOF First we must show that if s is a term of equational logic, then $\Gamma \vdash_{\text{ITT}} [\![s]\!] : \mathbb{N}$. For each instance of the axioms of equational logic, it is easy to verify that the translation of the axiom is a theorem of ITT, given the above proposition and currying. Since modus ponens and term formation can be translated using $\Pi - e$, we are done. □

The above result can easily be extended to handle existential quantification and conjunction, given the usual axiomatisations[11] by translation into the Σ type.

PRA$^\omega$ and PRA The translation from PRA$^\omega$ is even easier if we restrict to the theory without the $\times$ connective: we only need to replace the subterms of the form $\mathsf{Rec}(s, a, f)$ by the subterm $\mathrm{R}^{\mathbb{N}}(\bar{s}, \bar{a}, (i^{\mathbb{N}}, z^A)\mathsf{ev}(\bar{f}, i, z)$ for appropriate type A. The types of PRA$^\omega$ are exactly the sets of ITT.

PROPOSITION 69 If $s =_{\beta\eta} t$ is a theorem of PRA$^\omega$, then $\Gamma \vdash_{\text{ITT}} \bar{s} = \bar{t} : A$, where Γ consists only of enough set-valued assumptions to bind the variables of s and t and A is some appropriate set. Furthermore it is a theorem of the subsystem of ITT with only the Π and $\mathbb{N}$ type formers.

With a little ingenuity in reformulating PRA, we can give simpler form for the target system of the sub-theory PRA: it maps onto the sub-theory of ITT with just the $\mathbb{N}$ type former. The details of the translation are left to the interested reader.

Heyting Arithmetic The last theory we shall consider is the first-order formulation of Heyting Arithmetic. This theory is an extension of equational logic. It has a constant 0, an unary function letter s, two binary function letters $\times$, $+$, an additional connective $\neg$, and new axioms:

1. $\neg\phi \supset \phi \supset \psi$;
2. $0 + x = x$;
3. $\mathsf{s}(x) + y = \mathsf{s}(x + y)$;
4. $x + y = y + x$;
5. $0 \times 0 = 0$;

[11] The weak existential rule can be formulated:

1. $\phi[x := s] \supset \exists x.\phi$;
2. $\exists x.\phi \supset ((\forall x.phi) \supset \psi) \supset \psi$.

6. $\mathsf{s}(x) \times y = (x \times y) + x$;
7. $x \times y = y \times x$;
8. $\neg 0 = 1$
9. $\mathsf{s}(x) = \mathsf{s}(y) \supset x = y$;
10. If $i \notin \mathsf{FV}(\phi)$ then $\phi[x := 0] \supset \forall i.(\phi[x := i] \supset \phi[x := \mathsf{s}(i)]) \supset \forall x.\phi$.

We would like to define a translation first by providing a denotation for $+$, $\times$ (it will also be convenient to define a predecessor function):

DEFINITION 70
1. $\mathsf{pred}(m) \mathrel{\hat{=}} \mathrm{R}^{\mathbb{N}}(m, \underline{0}, (i^{\mathbb{N}}, -)i)$;
2. $\mathsf{add}(m, n) \mathrel{\hat{=}} \mathrm{R}^{\mathbb{N}}(m, n, (-, z^{\mathbb{N}})\mathsf{succ}(z))$;
3. $\mathsf{mul}(m, n) \mathrel{\hat{=}} \mathrm{R}^{\mathbb{N}}(m, \underline{0}, (-, z^{\mathbb{N}})\mathsf{add}(z, n))$.

The reader is invited to confirm the following:

EXERCISE 71
1. (a) $\vdash \forall m^{\mathbb{N}}.\mathsf{pred}(\mathsf{succ}(m)) =_{\mathbb{N}} m$ true;
   (b) $\vdash \mathsf{pred}(\underline{0}) =_{\mathbb{N}} \underline{0}$ true;
2. (a) $\vdash \forall m^{\mathbb{N}}.\mathsf{add}(\underline{0}, m) =_{\mathbb{N}} m$ true;
   (b) $\vdash \forall m^{\mathbb{N}}.\forall n^{\mathbb{N}}.\mathsf{add}(\mathsf{succ}(m), n) \mathrel{\hat{=}} \mathsf{succ}(\mathsf{add}(m, n))$ true;
   (c) $\vdash \forall m^{\mathbb{N}}.\forall n^{\mathbb{N}}.\mathsf{add}(m, n) \mathrel{\hat{=}} (\mathsf{add}(n, m))$ true;
3. (a) $\vdash \forall m^{\mathbb{N}}.\mathsf{mul}(\underline{0}, m) =_{\mathbb{N}} \underline{0}$ true;
   (b) $\vdash \forall m^{\mathbb{N}}.\forall n^{\mathbb{N}}.\mathsf{mul}(\mathsf{succ}(m), n) \mathrel{\hat{=}} \mathsf{add}((\mathsf{mul}(m, n), n)$ true;
   (c) $\vdash \forall m^{\mathbb{N}}.\forall n^{\mathbb{N}}.\mathsf{mul}(m, n) \mathrel{\hat{=}} (\mathsf{mul}(n, m))$ true;

We shall then try to derive Peano's rules:

PROPOSITION 72
1. $\mathsf{zero} : \mathbb{N}$;
2. If $\Gamma \vdash n : \mathbb{N}$ then $\Gamma \vdash \mathsf{succ}(n) : \mathbb{N}$;
3. $m : \mathbb{N}, n : \mathbb{N} \vdash \mathsf{succ}(m) =_{\mathbb{N}} \mathsf{succ}(n) \supset m =_{\mathbb{N}} n$ true;
4. Suppose $\Gamma \vdash P(n : \mathbb{N})$ prop. If $\Gamma \vdash P(0)$ true and $\Gamma \vdash \forall n : \mathbb{N}.P(n) \supset P(\mathsf{succ}(n))$, then $\Gamma \vdash \forall n : \mathbb{N}.P(n)$.

PROOF 1,2 and 4 are easy consequences of the inference rules $\mathbb{N}{-}i{-}0$, $\mathbb{N}{-}i{-}\mathsf{succ}$ and $\mathbb{N}{-}e$. To show 3, we must apply the last part of proposition 67 with the function pred. □

Unfortunately we cannot prove Peano's fourth axiom directly, that $\neg 0 = 1$, or as we would like to render it X **type** $\vdash 0 =_{\mathbb{N}} 1 \supset X$ **true**. This can be shown by Jan Smith's 'proof irrelevance' semantics[12]. We define a map ϕ from the collection of types to the two point set $\{\mathsf{tt}, \mathsf{ff}\}$ as follows:

$$\begin{aligned}
\phi(X) &\hat{=} \mathsf{ff}, \text{ if } X \text{ is a schematic letter} \\
\phi(\mathbb{N}) &\hat{=} \mathsf{tt} \\
\phi(s =_A t) &\hat{=} \mathsf{tt} \\
\phi(\Pi x^A.B) &\hat{=} \phi(A) \underset{\mathbb{B}}{\supset} \phi(B) \\
\phi(\Sigma x^A.B) &\hat{=} \phi(A) \underset{\mathbb{B}}{\wedge} \phi(B)
\end{aligned}$$

where $\supset_{\mathbb{B}}$ and $\wedge_{\mathbb{B}}$ are the boolean functions corresponding to implication and conjunction.

PROPOSITION 73

1. X **type** $\not\vdash s : X$ for any term s;
2. Let $\Gamma \vdash s : A$, and let Γ consist of m type assumptions X_1 **type**, ..., X_n **type** and n variable assumptions $x_1 : B_1, \ldots, x_n : B_n$. Then

$$(\phi(B_1) \underset{\mathbb{B}}{\wedge} \ldots \underset{\mathbb{B}}{\wedge} \phi(B_n)) \underset{\mathbb{B}}{\supset} A = \mathsf{tt}$$

PROOF The two parts are proven by simultaneous induction over the height of type inferences. For derivations of height 1, the first part follows from the fact that the only inference rule without premisses that matches the type is the variable rule, which doesn't math the telescope, whilst the second part follows by verifying the conclusion of all zero premiss rules.

For the inductive part, we show the first part by observing that the last rule must be an elimination rule. But by inspecting the principal premiss and applying the induction hypothesis, we can see that the principal premiss is uninhabited (in the case of $= -$e we also need proposition 60).

The second part follows by observing for all the remaining rules that the conclusion satisfies the property whenever all the premisses do. Note that the second part must be proved alongside the first to deal with the $\Pi -$ e, where the auxiliary premiss is of schematic type. □

COROLLARY 74 X **type** $\not\vdash 0 =_{\mathbb{N}} 1 \supset X$ **true**

The solution is to *define* negation to be $\neg A \hat{=} A \supset 0 =_{\mathbb{N}} 1$, but then the first of the new axioms becomes problematic.

LEMMA 75 If $\Gamma \vdash \phi$ **prop** and ϕ contains no schematic letters, and the only equality types occurring in ϕ are over $\mathbb{N}$, then $\Gamma \vdash 0 =_{\mathbb{N}} 1 \supset \phi$ **true**.

[12] 'On the independence of Peano's fourth axiom from Martin-Löf's type theory' [SmithJM:indpfa]. We must slightly alter his theorem to make sense of typing assumptions.

PROOF The key step is to provide a proof of $p : 0 =_{\mathbb{N}} 1 \vdash \forall m.m =_{\mathbb{N}} 0$ true using $\mathbb{N}-\mathrm{e}$. One can then easily produce a witness to any $\Gamma \vdash s =_{\mathbb{N}} t$ type, and so the lemma proceeds by a simple induction on the relevant types. □

DEFINITION 76 We extend the earlier translation from equational logic to ITT as follows:

$$\begin{aligned}
[\![k]\!] &\triangleq \underline{k} \\
[\![\mathsf{s}(s)]\!] &\triangleq \mathsf{succ}([\![s]\!]) \\
[\![s+t]\!] &\triangleq \mathsf{add}([\![s]\!],[\![t]\!]) \\
[\![s \times t]\!] &\triangleq \mathsf{mul}([\![s]\!],[\![t]\!]) \\
[\![\neg\phi]\!] &\triangleq [\![\phi]\!] \supset 0 =_{\mathbb{N}} 1
\end{aligned}$$

It is then quite easy to prove the following theorem.

PROPOSITION 77 If ϕ is a theorem of HA, then $\Gamma \vdash_{\mathrm{ITT}} [\![\phi]\!]$ true, where Γ consists only of enough set-valued assumptions to bind the variables and function letters of ϕ.

Strong existential Finally, we may make a few remarks about the Σ type former. It has three principal uses: to represent existential quantification, to encode relations as types, and it is used in the coding of signatures, subsets and quotient types. We have not really got space to discuss this last point; the interested reader is directed to Martin Hofmann's PhD thesis *Extensional concepts in intensional type theory* [HofmannM:extcit].

The representation of existential quantification is quite straightforward; it is a simple matter of syntactic sugar. It also comes with a bonus: because the Σ type former corresponds to strong existence in the terminology of W. A. Howard [HowardWA:fortnc]), which means that it is possible to take left and right projections of the type $\Sigma x^A.B$[13], we can express the axiom of choice for this calculus, and for any calculus whose existence operator is defined according to the BHK account of the logical connectives. The proof of this is shown by the construction, that if $\Gamma \vdash s : \forall x^A.\exists y^B.C$ type, we have

$$\Gamma \vdash \left\langle \lambda x_0^A.\mathsf{outl}(\mathsf{ev}(s,x')), \lambda x^A.\mathsf{outr}(\mathsf{ev}(s,x)) \right\rangle : \exists f \in (\Pi x_0^A.B).\forall x^A.C[y := \mathsf{ev}(f,x)]$$

The coding of relations as types is made much easier by the strength of the Σ type former. Whilst it is quite easy to code standard arithmetic functions such as $x <_{\mathbb{N}} y$ as primitive recursive functions, eg. $x < y \Leftrightarrow f_<(x,y) =_{\mathbb{N}} \underline{0}$, the codings tend to be somewhat opaque, and it is not possible to encode undecidable relations in this way. A more natural encoding of the less than or equal to relation is by means of a defined type $\mathsf{LQ}(x : \mathbb{N}, y : \mathbb{N}) \triangleq \exists i.x + i =_{\mathbb{N}} y$. The reader is invited to verify that LQ has the desired properties.

[13]That is, we can define the constants outl and outr, as we discussed when we introduced the inference rules for Σ.

EXERCISE 78

1. $\vdash \forall m^{\mathbb{N}}.\mathsf{LQ}(\underline{0}, m)$ true;
2. $\vdash \forall m^{\mathbb{N}}, n^{\mathbb{N}}.\mathsf{LQ}(m, n) \supset \mathsf{LQ}(\mathsf{succ}(m), \mathsf{succ}(n))$ true;
3. $\vdash \forall l^{\mathbb{N}}, m^{\mathbb{N}}, n^{\mathbb{N}}.\mathsf{LQ}(l, m) \supset \mathsf{LQ}(m, n) \supset \mathsf{LQ}(l, n)$ true;
4. $\vdash \forall m^{\mathbb{N}}, n^{\mathbb{N}}.\mathsf{LQ}(\mathsf{succ}(m), \mathsf{succ}(n)) \supset \mathsf{LQ}(m, n)$ true;
5. $\vdash \forall m^{\mathbb{N}}, n^{\mathbb{N}}.\mathsf{LQ}(m, n) \supset \mathsf{LQ}(n, m) \supset m =_{\mathbb{N}} n$ true.

2.5 Consistency

We wish to draw attention to a number of claims that might be held to establish consistency, which we arrange in increasing order of strength:

1. There are uninhabited types;
2. Arithmetic is sound (ie. $\nvdash \underline{0} =_{\mathbb{N}} \underline{1}$ true);
3. We can provide a static semantics for the theory (ie. we can provide a modeling function for the calculus that identifies definitionally equal terms);
4. The conversion theory of the calculus is sound (ie. we have a reduction relation for the calculus that decides definitional equality, and possesses a good notion of normal form adequate to show reflection of the equality types).

It should be clear that each proposition is a consequence of the one following[14]. It is also the case that the first two do not demand the subject reduction property to make sense, whilst the last two do.

The first proposition follows from the proof irrelevance semantics given in proposition 73. We show the second proposition as a corollary to the third.

Our static semantics proceeds by a direct map from types and terms to sets, one in which the interpretation of Π-types is the usual pointwise function space. Despite their simplicity, such direct semantics are not common in the literature, presumably due to the fact that they cannot be extended to impredicative quantification, and I know of none for Martin-Löf's type theory.

The intuition lying behind the model is that if $\vdash s : A$, then $[\![s : A]\!]$ and $[\![A \ \mathsf{type}]\!]$ are defined, and $[\![s : A]\!] \in [\![A \ \mathsf{type}]\!]$. This allows a much more finely grained account of the type theory than that on the 'proof-irrelevance' account; in this account we can deduce that certain types are uninhabited in a much wider set of instances.

The account is complicated by the presence of open terms and types. If Γ is an n-ary telescope and $\Gamma \vdash s : A$ then our denotations are given $[\![\Gamma \vdash s : A]\!]$ and $[\![\Gamma \vdash A \ \mathsf{type}]\!]$. These are functions whose domain is an n-tuple of sets and range is a set. The above inhabitation condition cannot be generalised to arbitrary functions, since $[\![0 =_{\mathbb{N}} 1 \ \mathsf{type}]\!] = \emptyset$ but $p : 0 =_{\mathbb{N}} 1 \vdash p : 0 =_{\mathbb{N}} 1$.

We begin by presenting some standard functions on sets.

[14] A reductive semantics gives rise to a static semantics in ZFC if we can find a suitable ordinal bound to define a normalisation function.

DEFINITION 79

1. (a) Let $\mathsf{pair}(x, y) \triangleq \{\{x\}, \{x, y\}\}$;
   (b) Let $\mathsf{fst}(x) \triangleq \{a \in \bigcup x \mid \forall b \in x.a \in b\}$, and let $\mathsf{snd}(x) \triangleq \{a \in \bigcup x \mid \forall b \in x.b \subseteq \{\mathsf{fst}(x), a\}\}$;
   (c) Let $\mathsf{cart}(x, y) \triangleq \{\mathsf{pair}(a, b) \mid a \in x \wedge b \in y\}$.

2. (a) Let $\mathsf{Rel}(x) \triangleq \forall z \in x.\exists a.\exists b.z = \mathsf{pair}(a, b)$. Let $\mathsf{dom}(x) \triangleq \{\mathsf{fst}(a) \mid a \in x\}$ and $\mathsf{ran}(x) \triangleq \{\mathsf{snd}(a) \mid a \in x\}$;
   (b) Let $\mathsf{Func}(x) \triangleq \mathsf{Rel}(x) \wedge \forall y \in \mathsf{dom}(x).\exists b \in \mathsf{ran}(x).\forall c \in \mathsf{ran}(x).\mathsf{pair}(a, b) = x \wedge \mathsf{pair}(a, c) = x \supset b = c$;
   (c) Let

$$\mathsf{funsp}(x, y) \triangleq \{g \in \mathbb{P}(\mathsf{cart}(x, y)) \mid \mathsf{Func}(g) \wedge \mathsf{dom}(g) = x\}$$
$$\mathsf{eval}(x, y) \triangleq \begin{cases} \bigcup\{b \in \mathsf{ran}(x) \mid \mathsf{pair}(y, b) \in x\}, & \text{if } \mathsf{Func}(x) \\ \emptyset, & \text{otherwise} \end{cases}$$

3. (a) Let ω be the set of Von Neumann natural numbers, and let $\mathsf{zr} \triangleq \emptyset$ and $\mathsf{su}(x) \triangleq x \cup \{x\}$. Let num be the function from the natural numbers to their Von Neumann representation.
   (b) Let

$$\begin{aligned} \mathsf{rec}(a, f) \triangleq \bigcup\{g \in \mathsf{funsp}(\omega, A) \mid \mathsf{pair}(\mathsf{zr}, a) \in g \\ \wedge \forall z \in g.\mathsf{pair}(\mathsf{su}(\mathsf{fst}(z)), \mathsf{eval}(f, z) \in g\}) \\ \text{where } A \triangleq \{a\} \cup \{\mathsf{snd}(x) \mid x \in \mathsf{ran}(f)\} \end{aligned}$$

PROPOSITION 80

1. $\forall x.\forall y.\mathsf{fst}(\mathsf{pair}(x, y)) = x \wedge \mathsf{snd}(\mathsf{pair}(x, y)) = y$.

2. Let A and B be non-empty sets, and let $f \in \mathsf{funsp}(A, B)$. Then for all $a \in A$, $\mathsf{eval}(f, a) \in B$.

3. Let A be a non-empty set, and let $a \in A$ and $f \in \mathsf{funsp}(\mathsf{cart}(\omega, A), A)$. Also let

$$v(n) \triangleq \begin{cases} a, & \text{if } n = 0 \\ f(\mathsf{pair}(\mathsf{num}(n-1), v(n-1))), & \text{if } n > 0 \end{cases}$$

   Then we have

   (a) For each $n \in \mathbb{N}$, $\mathsf{pair}(\mathsf{num}(n), v(n)) \in \mathsf{rec}(a, f)$;
   (b) $\mathsf{rec}(a, f) \in \mathsf{funsp}(\omega, A)$.

PROOF SKETCH Parts one and two are standard in the literature. The first part of part three is proven by an easy induction. For the second part, observe that set manipulation gives $\mathsf{rec}(a, f) \in \mathbb{P}(\mathsf{cart}(\omega, A))$, and so it satisfies the predicate Rel. It remains to show that it is a total function. Clearly the previous induction establishes that the set $\{g \in \mathsf{funsp}(\omega, A) \mid \mathsf{pair}(\mathsf{zr}, a) \in g \wedge \forall z \in g.\mathsf{pair}(\mathsf{su}(\mathsf{fst}(z)), \mathsf{eval}(f, z) \in g\})$ contains an inhabitant. We can also see that there is a unique such inhabitant of this set, since an easy induction shows that the total functions in this set cannot disagree at any value. Therefore $\mathsf{rec}(a, f)$ is a total function. □

We are now in a position to provide the modelling functions $[\![\cdot]\!]$. We provide this by associating with each rule of the calculus

$$\frac{\Gamma_1 \vdash \mathcal{A}_1 \quad \ldots \quad \Gamma_n \vdash \mathcal{A}_n}{\Gamma' \vdash \mathcal{A}'},$$

(where each of $\mathcal{A}_i, \mathcal{A}'$ are core judgements) a conditional definition:

> Given $F_i \hat{=} [\![\Gamma_i \vdash \mathcal{A}_i]\!]$ for each $1 \leq i \leq n$, the denotation $\Gamma' \vdash \mathcal{A}'$ is given by $F \hat{=} \ldots$.

So by substitution, a complete type inference of a judgement gives a definition of the set-valued function as desired. We provide the auxiliary function F to allow explicit reference to parameters.

DEFINITION 81

1. All of the structural rules are essentially trivial, since the equality judgement is interpreted by set equivalence, and we do not interpret telescopes directly.

2. Π-type former:

   (a) $\Pi - \mathrm{f}$: Given

   $$F_1(\vec{z}, x) = [\![\Gamma, x : A \vdash B \ \mathsf{type}]\!]$$

   we let

   $$G_0(\vec{z}) \hat{=} [\![\Gamma \vdash A \ \mathsf{type}]\!]$$
   $$G_1(\vec{z}) \hat{=} \bigcup \{F_1(\vec{z}, x) \mid x \in G_0(\vec{z})\}$$

   Then the denotation of $[\![\Gamma \vdash \Pi x^A.B \ \mathsf{type}]\!]$ is

   $$\{g \in \mathsf{funsp}(G_0(\vec{z}), G_1(\vec{z})) \mid \forall a \in \mathsf{dom}(g).\mathsf{eval}(g, a) \in F_1(\vec{z}, a)\}$$

   (b) $\Pi - \mathrm{i}$: Given

   $$F_1(\vec{z}, x) = [\![\Gamma, x^A \vdash s : B]\!]$$

   we give the denotation of

   $$[\![\Gamma \vdash \lambda x^A.s : \Pi x^A.B]\!] \hat{=} \{\mathsf{pair}(a, b) \mid a \in [\![\Gamma \vdash A \ \mathsf{type}]\!](\vec{z}) \wedge b = F_1(\vec{z}, a)$$

(c) $\Pi - \mathrm{e}$: Given

$$\begin{aligned} F_1(\vec{z}, x) =& [\![\Gamma \vdash s : \Pi x^A.B]\!] \\ F_2(\vec{z}) =& [\![\Gamma \vdash t : A]\!] \end{aligned}$$

we give the denotation of $[\![\Gamma \vdash \mathsf{ev}(s,t) : B[x := A]]\!]$ by

$$F(\vec{z}) \hat{=} \hat{=} \mathsf{eval}(F_1(\vec{z}), F_2(\vec{z}))$$

3. Σ-type former:

(a) $\Sigma - \mathrm{f}$: Given

$$F_1(\vec{z}, x) = [\![\Gamma, x : A \vdash B \ \mathsf{type}]\!]$$

then the denotation of $[\![\Gamma \vdash \Sigma x^A.B \ \mathsf{type}]\!]$ is given by

$$F'(\vec{z}) \hat{=} \{\mathsf{pair}(a,b) \mid a \in [\![\Gamma \vdash A \ \mathsf{type}]\!](\vec{z}) \wedge b \in F_1(\vec{z}, a)\}$$

(b) $\Sigma - \mathrm{i}$: Given

$$\begin{aligned} F_1(\vec{z}) =& [\![\Gamma \vdash s : A]\!] \\ F_2(\vec{z}) =& [\![\Gamma \vdash t : B[x := s]]\!] \end{aligned}$$

then the denotation of $[\![\Gamma \vdash \langle s, t\rangle : \Sigma x^A.B]\!]$ is given by

$$F'(\vec{z}) \hat{=} \mathsf{pair}(F_1(\vec{z}, F_2(\vec{z})))$$

(c) $\Sigma - \mathrm{e}$: Given

$$\begin{aligned} F_1(\vec{z}) =& [\![\Gamma \vdash s : \Sigma x^A.B \ \mathsf{type}]\!] \\ F_2(\vec{z}, x, y) =& [\![\Gamma, x : A, y : B \vdash t : C(\langle(\rangle x, y))]\!] \end{aligned}$$

the denotation $[\![\Gamma \vdash \mathrm{R}^{\Sigma}(s, (x^A, y^B)t) : C(s)]\!]$ is given by

$$F'(\vec{z}) \hat{=} F_2(\vec{z}, \mathsf{fst}(F_1(\vec{z})), \mathsf{snd}(F_1(\vec{z})))$$

4. $=$-type former:

(a) $= -\mathrm{f}$: Given $F_1(\vec{z}) = [\![\Gamma \vdash s : A]\!]$ and $F_2(\vec{z}) = [\![\Gamma \vdash t : A]\!]$ the denotation $[\![\Gamma \vdash s =_A t \ \mathsf{type}]\!]$ is given by

$$F'(\vec{z}) \hat{=} \begin{cases} \{\emptyset\}, & \text{if } F_1(\vec{z}) = F_2(\vec{Z}) \\ \emptyset, & \text{otherwise.} \end{cases}$$

(b) $= -$i: The denotation $[\![\Gamma \vdash \mathsf{refl}(s) : s =_A t]\!]$ is given by $F'(\vec{z}) \hat{=} \emptyset$.

(c) $= -$e: Given

$$\begin{aligned} F_1(\vec{z}) =& [\![\Gamma \vdash p : s =_A t]\!] \\ F_2(\vec{z}, x) =& [\![\Gamma, x : A \vdash u : C(x, x, \mathsf{refl}(x))]\!] \end{aligned}$$

let $G(\vec{z}) \hat{=} [\![\Gamma \vdash A\ \mathsf{type}]\!]$ to obtain the denotation

$$[\![\Gamma \vdash \mathrm{R}^{=}(p, (x^A)u) : C(s, t, p)]\!](\vec{z}) \hat{=} F_2(\vec{z}, G(\vec{z}))$$

5. $\mathbb{N}$:

(a) $\mathbb{N} -$ f: $[\![\Gamma \vdash \mathbb{N}\ \mathsf{type}]\!]$ is given by $F'(\vec{z}) \hat{=} \omega$, where ω is the Cantor ordinal;

(b) $\mathbb{N} -$ f: $[\![\Gamma \vdash \mathsf{zero} : \mathbb{N}]\!]$ is given by $F'(\vec{z}) \hat{=} \mathsf{zr}$;

(c) $\mathbb{N} -$ i(succ): Given $F_1(\vec{z}) = [\![\Gamma \vdash n : \mathbb{N}]\!]$ the denotation $[\![\Gamma \vdash \mathsf{succ}(n) : \mathbb{N}]\!]$ is given by $F'(\vec{z}) \hat{=} \mathsf{su}(F_1(\vec{z}))$;

(d) $\mathbb{N} -$ e: Given

$$\begin{aligned} F_1(\vec{z}) =& [\![\Gamma \vdash s : \mathbb{N}]\!] \\ F_2(\vec{z}, m : \mathbb{N}) =& [\![\Gamma \vdash C\ \mathsf{type}]\!] \\ F_3(\vec{z}) =& [\![\Gamma \vdash a : C[n := \mathsf{zero}]]\!] \\ F_4(\vec{z}, n, z) =& [\![\Gamma, n : \mathbb{N}, z : C[m := n] \vdash f : c[m := \mathsf{succ}(n)]]\!] \end{aligned}$$

the denotation $[\![\Gamma \vdash \mathrm{R}^{\mathbb{N}}(s, a, (n^{\mathbb{N}}, z^{C[m:=n]})f : C[m := s])]\!]$ is given

$$\begin{aligned} F'(\vec{z}) \hat{=}\ & \mathsf{eval}(\mathsf{rec}(F_3(\vec{z}), f), F_1(\vec{z})) \\ \text{where } f \hat{=}\ & \{\mathsf{pair}(\mathsf{pair}(n, c), v) \mid n \in \omega \wedge c \in F_2(\vec{z}, n) \wedge v \in F_4(\vec{z}, n, c)\} \end{aligned}$$

DEFINITION 82 σ is a *valuation of* Γ for a telescope $\Gamma \equiv \mathcal{A}_1, \ldots, \mathcal{A}_n$, if σ is an n-tuple of sets $S_1, \ldots, S_n$ satisfying the following proposition:

For each i, if $\mathcal{A}_i \equiv x_i : A_i$ then $S_i \in [\![\mathcal{A}_1, \ldots \mathcal{A}_{i-1} \vdash A_i\ \mathsf{type}]\!](S_1, \ldots, S_{i-1})$.

The set of valuations of Γ is given by $\mathsf{val}(\Gamma)$.

PROPOSITION 83

1. If $\Gamma \vdash A = B\ \mathsf{type}$, then $[\![\Gamma \vdash A\ \mathsf{type}]\!] = [\![\Gamma \vdash B\ \mathsf{type}]\!]$
2. Let $\Gamma \vdash s : A$. Then $\forall \sigma \in \mathsf{val}(\Gamma).[\![\Gamma \vdash s : A]\!](\sigma) \in [\![\Gamma \vdash A\ \mathsf{type}]\!](\sigma)$;
3. Let $\Gamma \vdash s = t : A$. Then $\forall \sigma \in \mathsf{val}(\Gamma).[\![\Gamma \vdash s : A]\!](\sigma) = [\![\Gamma \vdash t : A]\!](\sigma)$;
4. Let $\Gamma \vdash p : s =_A t$. Then $\forall \sigma \in \mathsf{val}(\Gamma).[\![\Gamma \vdash s : A]\!](\sigma) = [\![\Gamma \vdash t : A]\!](\sigma)$;

PROOF We must prove this by a simultaneous induction, since parts are interdependent. We proceed by structural induction on complete type inferences. □

COROLLARY 84 $\not\vdash 0 =_{\mathbb{N}} 1$ true.

PROOF For a contradiction, assume $\vdash p : 0 =_{\mathbb{N}} 1$ true. Then $[\![\vdash p : 0 =_{\mathbb{N}} 1]\!] \in \emptyset$ by the previous proposition. □

2.6 Conversion theory

We have two main aims in this section: firstly, we wish to show that the judgement forms expressing convertive equality are decidable, and secondly we wish to prove the head normal form theorem, which will serve as our formal foundation for discussing issues in the semantics of the calculus.

DEFINITION 85

1. We write $\Gamma \vdash s \to^c s' : A$, or that s converts to s', if there is an inference of the judgement $\Gamma \vdash s = s' : A$ whose last rule is a logical conversion rule;

2. We write $\Gamma \vdash s =_\alpha s' : A$, for *alpha equivalence*, if there is an inference of the judgement $\Gamma \vdash s = s' : A$ in which none of the logical conversion rules appear (ie. which only contains structural conversion rules);

3. We write $\Gamma \vdash s \to^1 s' : A$ if there is a context-redex decomposition $s \equiv f(r)$ where $\Gamma \vdash f((\Delta)x : B) : A$, $\Gamma, \Delta \vdash r \to^c c : B$ and $s' \equiv f(c)$, which, in the case of an eta conversion we require an additional criterion:

   If the conversion arises by the $\Pi - c - eta$ rule, then the condition that r is not a lambda abstraction and the variable x in $f(x)$ is not the principal premiss of a $\Pi - e$ rule must obtain;

   Similarly if the conversion arises by the $\Sigma - c - eta$ rule, then the condition that r is not a pair and the variable x in $f(x)$ is not the principal premiss of a $\Sigma - e$ rule must obtain.

4. We write $\Gamma \vdash s \to^* s' : A$ if either $s \equiv s'$ or there is a sequence of terms $s_1, \ldots, s_n$ such that $s_1 \equiv s$, $s_n \equiv s'$ and for $1 \leq i < n$, $\Gamma \vdash s_i \to^1 s_{i+1} : A$.

It is elementary to establish the following proposition.

PROPOSITION 86

1. If $\Gamma \vdash s \to^* s' : A$ then $\Gamma \vdash s = s' : A$;
2. $\Gamma \vdash s \to^c s' : A$ iff there is an inference of the judgement $\Gamma \vdash s = s' : A$ in which none of the class 3, 4 or 5 structural rules appear.

We prove decidability firstly by showing that the relation '$\to^1$' is strongly normalising and Church–Rosser, and then by showing that for $\Gamma \vdash s : A$ and $\Gamma \vdash t : A$, we have $\Gamma \vdash s = t : A$ iff the normal forms of s, t are alpha equivalent.

DEFINITION 87 We define a translation of judgements of ITT into judgements of $\lambda\mathbb{N}_r$, called the *dependency forgetting translation*:

1. Firstly there is the translation on types:

$$\begin{aligned}
[\![X]\!] &\hat{=} X, \quad [\![\mathbb{N}]\!] \hat{=} \mathbb{N} \\
[\![\Pi x : A.B]\!] &\hat{=} [\![A]\!] \Rightarrow [\![B]\!] \\
[\![\Sigma x : A.B]\!] &\hat{=} [\![A]\!] \times [\![B]\!] \\
[\![s =_A t]\!] &\hat{=} [\![A]\!]
\end{aligned}$$

2. Then the translation on hypotheses:

$$\begin{aligned}
[\![X \ \mathsf{type}]\!] &\hat{=} \emptyset \\
[\![x : A]\!] &\hat{=} \{x : [\![A]\!]\}
\end{aligned}$$

3. and then the translation on terms:

$$\begin{aligned}
[\![x]\!] &\hat{=} x \\
[\![\lambda x^A.s]\!] &\hat{=} \lambda x^{[\![A]\!]}.[\![s]\!] \\
[\![\mathsf{ev}(s,t)]\!] &\hat{=} \mathsf{ev}([\![s]\!],[\![t]\!]) \\
[\![\langle s,t\rangle]\!] &\hat{=} \langle [\![s]\!],[\![t]\!]\rangle \\
[\![\mathrm{R}^{\Sigma}(s,(x^A,y^B)t)]\!] &\hat{=} \mathsf{ev}(\lambda x^{[\![A]\!]}.\lambda y^{[\![B]\!]}.[\![t]\!],\mathsf{outr}([\![s]\!]),\mathsf{outl}([\![s]\!])) \\
[\![\mathsf{refl}(s)]\!] &\hat{=} \mathsf{refl}([\![s]\!]) \\
[\![\mathrm{R}^{=}(s,(x^A)t)]\!] &\hat{=} \mathsf{ev}(\lambda x^{[\![A]\!]}.[\![t]\!],[\![s]\!]) \\
[\![\mathsf{zero}]\!] &\hat{=} \mathsf{zero} \quad [\![\mathsf{succ}(n)]\!] \hat{=} \mathsf{succ}([\![n]\!]) \\
\left[\!\!\left[\mathrm{R}^{\mathbb{N}}(s,a,(i^{\mathbb{N}}z^{C(i)})f)\right]\!\!\right] &\hat{=} \mathsf{outr}(\mathsf{Rec}([\![s]\!],\langle 0,[\![a]\!]\rangle,(w^{\mathbb{N}\times[\![C]\!]})\langle \mathsf{succ}(\mathsf{outl}(u)),f'\rangle)) \\
&\quad \text{where} f' \equiv [\![f]\!][z := \mathsf{outr}(w), i := \mathsf{outl}(u)]
\end{aligned}$$

PROPOSITION 88

1. If $\Gamma \vdash_{\mathrm{ITT}} s : A$ then $\bigcup\{[\![\mathcal{A}]\!] \mid \mathcal{A} \in \Gamma\} \vdash_{\lambda\mathbb{N}_r} [\![s]\!] : [\![A]\!]$;
2. If $\Gamma \vdash_{\mathrm{ITT}} s \to^1 s' : A$ then $[\![s]\!] \to^+ [\![s']\!]$ in $\lambda\mathbb{N}_r$.

PROOF We must first establish that if $\Gamma \vdash A = A'$ type, then $[\![A]\!] \cong [\![A']\!]$, which follows by a simple induction on the structure of types. Then the first part follows from an elementary induction on the structure of the term s, and the second follows by a case analysis on the logical conversions, and in the case of the $\Pi - \mathrm{c} - \mathrm{eta}$ and $\Sigma - \mathrm{c} - \mathrm{eta}$ rule the condition specified for eta rules must apply. The additional condition in the second part ensures that every reduction of ITT maps to a non-zero number of reductions of $\lambda\mathbb{N}_r$. □

COROLLARY 89 The relation $\rightarrow^1$ is strongly normalising.

PROPOSITION 90 (THE DIAMOND PROPERTY)
If $\Gamma \vdash s \rightarrow^1 t_1 : A$ and $\Gamma \vdash s \rightarrow^1 t_2 : A$ then there is a term u such that $\Gamma \vdash t_1 \rightarrow^* u : A$ and $\Gamma \vdash t_2 \rightarrow^* u : A$.

PROOF This follows by observing that the diamonds used to complete the critical pairs of $\lambda\mathbb{N}_r$ exist in ITT whenever the critical pair is the image of a term of ITT. □

COROLLARY 91 The relation $\rightarrow^1$ is Church–Rosser.

LEMMA 92

1. If $\Gamma \vdash s =_\alpha s' : A$ and $\Gamma \vdash s \rightarrow^1 t : A$, then there is a term t' such that $\Gamma \vdash s' \rightarrow^1 t' : A$ and $\Gamma \vdash t =_\alpha t' : A$;
2. Alpha equivalence is decidable;

PROOF Part 1. is shown, for each logical conversion, by induction on the structure of justifications of $\Gamma \vdash s =_\alpha s' : A$.

We describe the decision procedure for part 2. as follows. Consider $\Gamma \vdash s : A$. To each variable binding x^A of s we associate a depth, which is the number of variable bindings under which the binding of x appears. We replace every occurrence of the variable x as a binding and as an assertion, by the identifier v_k where k is the depth of x. A simple induction shows that the term $\overline{s}$ which arises by applying this replacement to all bound variables admits the typing $\Gamma \vdash \overline{s} : A$.

If $\Gamma \vdash s : A$ and $\Gamma \vdash t : A$, we say that s and t are *alpha compatible* if the terms $\overline{s}$ and $\overline{t}$ are identical, or differ only in the types attached to bound variables.

It remains to show that alpha compatibility coincides with alpha equivalence. We can show that $\Gamma \vdash \overline{s} =_\alpha \overline{t}$ is provable if the above condition holds, and similarly we can construct proofs of $\Gamma \vdash s =_\alpha \overline{s} : A$ and $\Gamma \vdash t =_\alpha \overline{t} : A$, whilst the reverse direction is shown by a simple induction on proofs of $\Gamma \vdash s =_\alpha t$. □

THEOREM 93 The equality relation on terms is decidable.

PROOF For two terms $\Gamma \vdash s : A$ and $\Gamma \vdash t : A$, if their normal forms are alpha equivalent, then clearly $\Gamma \vdash s = t : A$. It remains to show that the converse is true, which establishes that terms are equal iff their normal forms are alpha equivalent.

We show by induction on the inference of $\Gamma \vdash s = t : A$ that there are terms s', t' such that $\Gamma \vdash s \rightarrow^* s' : A$, $\Gamma \vdash t \rightarrow^* t' : A$ and $\Gamma \vdash s' =_\alpha t' : A$, for which we will require the first part of the previous lemma. We can continue to apply the lemma to reductions of s' or t' and obtain this separation, until the two terms are normal forms. □

Head normal form theorem

Finally we shall outline an important syntactic tool of the conversion theory, that of *head normal forms*, and show how it can be used to demonstrate consistency. In the context of Martin-Löf's type theory we are interested in the analog of what are known as *weak head normal forms*, that is, where the process of head reduction is completed if the term is either canonical or has a head variable; the theory of strong head normal forms only ends when the term has a head variable. We shall have nothing further to say about strong head normal forms.

DEFINITION 94

1. s is in *canonical form* if it has one of the following forms:

$$\begin{aligned} s &\equiv \lambda x^A.s' \\ s &\equiv \langle t_0, t_1 \rangle \\ s &\equiv \mathsf{refl}(s') \\ s &\equiv \underline{k}, \text{ for some numeral } k; \end{aligned}$$

2. The *head subterms* of a term s is a subset of $\mathsf{st}(s)$ consisting of:
   (a) Just the term itself if it is a canonical form or a variable;
   (b) The term itself together with the head subterms of s' if s matches one of $\mathsf{ev}(s', t)$, $\mathrm{R}^{\mathbb{N}}(s', t, (i, z)u)$ or $\mathrm{R}^{=}(s', (x)t)$;
   (c) The term itself together with the head subterms of s' in case $s \equiv \mathsf{succ}(s')$ and s' is not a canonical form.

   Note that the head subterms are all distinct, and as occurrences of s are arranged from left to right in the term, rather like Russian dolls;

3. We call a term *head reducible* if one of its head subterms is a redex. We say that a term is in *head normal form*, or *HNF*, if it is not head reducible;

4. If there is a variable amongst the head subterms of a term, we call this its *head variable*.

LEMMA 95 If a term has a head variable, then this variable is free.

PROOF Write out the term as a natural deduction tree in the style of section 1.3. Consider the principal path; observe that all of the subderivations with conclusions on the principal path either correspond to head subterms, or they are subterms of the rightmost head subterm, which is a canonical form.

Observe that the only variable binder that can occur on the principal path is a lambda abstraction, which corresponds to a canonical form, and that the topmost formula of the principal path must arise either by $\mathrm{hyp} - \mathrm{ass}$, or the rule $\mathbb{N} - \mathrm{t} - \underline{0}$. This latter case must form part of a canonical form, and so the rightmost head subterm must be either a variable, and thus the head variable, or a canonical form. Furthermore if the rightmost head subterm is a variable, it cannot be bound, since it does not form part of a lambda abstraction. □

THEOREM 96 (HEAD NORMAL FORM THEOREM)
If a term is in head normal form, then it is either a canonical term or it has a head variable.

PROOF For a contradiction, suppose this is not so. Then one of the following cases must hold:

1. There is a term whose rightmost head subterm is not either a head variable or a canonical form;

2. There is a term which is not a canonical form, has no head redexes, but whose rightmost subterm is a canonical form.

The first case cannot hold, since all terms are either variables, canonical forms or match one of the five cases in the definition above of terms whose principal subterm is a head subterm.

We see that the second case cannot hold, since if the rightmost subterm is a canonical form, then either it is the only head subterm, or it is part of a beta redex, since the next rightmost subterm must be an elimination rule. (The successor of a canonical form must itself be a canonical form if it is well-typed). □

COROLLARY 97 (CANONICAL FORM THEOREM)
All closed normal forms of ITT are canonical forms.

Finally let us note that the canonical form theorem allows us another means of deducing the unprovability of $\vdash 0 =_{\mathbb{N}} 1$ true.

2.7 Logical and constructive harmony

Neither our representation results, nor our consistency results are by themselves sufficient to establish that the theory ITT determines the meaning of its formulae; for this we must appeal to a theory based upon the roles of its assertions in a way similar to that our account of logical harmony did for the intuitionistic propositional logic in chapter one.

There are several new issues that we must deal with in providing our logical formalist justification of type theory. The first is that here we are not just dealing with types that are to be interpreted as propositions capable of being asserted, but also types that are to be interpreted as constructive sets, such as $\mathbb{N}$ and $\mathbb{N} \Rightarrow \mathbb{N}$.

A crucial insight of Martin-Löf is that the formulae-as-types correspondence allows us to blur the distinction between types–as–propositions and types–as–sets, letting us treat the justification of each class in the same way. Indeed it is desirable to do so: a type constructor such as Π may be the principal constructor of either a type–as–proposition or a type–as–set, and so it is natural and economical

to provide it with a single account of justification[15].

There are ways in which the two kinds of account are different. For one, we are dealing with different kinds of concepts. The claim that logical harmony vindicates the idea that the meaning of the logical connectives is determined by the associated rules depended upon our analysis of the practice of assertion, and this dependence seems to be fundamental to the logical formalist approach. To what are we appealing in the case of types–as–sets? Certainly not assertion in the simple sense. But we still can maintain, as noted by Martin-Löf, that asserting that a construction has been made is a judgement just as much as asserting that a proposition has been proven. So, just as the meaning of a proposition on the logical formalist proposition lies in when it is appropriate to assert it and how it may be used as a hypothesis validating other assertions, so too the meaning of a set lies in when it is possible to construct an inhabitant of it, and how such an inhabitant may be used as an ingredient of another construction. And so analogous to the need for harmony in our assertoric practice, there is a need for harmony in the practice of constructing mathematical objects.

This leads us to our second difference: when we come to our formulation of the inversion principle for constructions, it is not enough that we know how to *associate* a new term with the detour removed; we must know that the two derivations are *identical*. Nonetheless as we noted in section 1.4, the formulae–as–types correspondence provides us with a notion of identity on proofs, and so there is no difficulty in uniformly strengthening our description of the inversion principle with a requirement that the two derivation forms justifying the sequent are equivalent.

Π and Σ type formers The easiest cases are those of the Π and Σ type formers, whose justification is a direct analogue of that for the $\supset$ and $\wedge$ connectives. The terms justifying the eliminative and decompositional clauses are identical to the term representations of the respective proofs justifying the analogous clauses for NJ.

We have a stronger requirement, however. We must also justify that the detour bearing terms are identical to their associated terms, and consequently we need to decide which of the notions of equality available is the right one. I shall not argue this point at length: I choose the inhabitation of propositional equality as the natural notion of equality on terms (as opposed to the stricter convertive equality), but note that one may be concerned that there is a circularity involved in justifying harmony by means of a concept whose coherence depends upon demonstrating harmony. I shall discuss this point in a little more depth in the conclusions.

For the Π type former, the additional equality constraint is justified for the eliminative clause by the $\Pi - \mathrm{c}$ rule and equality reflection, and for the decompositional

[15]Furthermore, I believe there is a deeper reason behind this coincidence: there is an intuitive dependence of our ideas about mathematical propositions upon our grasp of mathematical algorithms, a relationship that has its philosophical roots in Wittgenstein's discussion of rule-following, and its mathematical roots in the Curry–Howard correspondence.

clause by the $\Pi - \mathrm{c} - \mathrm{eta}$ rule and equality reflection. For the Σ type former the eliminative clause is justified similarly, but we note that the expected decompositional clause is not a convertive equality: we note however that it is quite simple to prove inhabitation of propositional equality.

$=_A$ type former and synthetic harmony The situation with the justification of propositional equality is a little less straightforward. Whilst the rule $= -c$ satisfies the eliminative clause of the inversion principle, the decompositional clause cannot so be satisfied.

It does not follow from this that synthetic harmony fails: we have argued that the inversion principle provides sufficient grounds to make the claim of logical (and by analogy constructive) harmony, but nowhere have we claimed that it is necessary. Recall from our brief discussion of the matter at the end of section 1.1 the grounds for insisting that logical harmony prevails in our justificatory practice: they were firstly to provide the strongest logical rules consistent with that practice, and secondly to ensure the coherence of that practice.

Let us now examine the incoherence potentially arising from the failure of synthetic harmony. An illustrative example is that a common formulation of quantum logic in natural deduction exhibits a special kind of failure of conservativity. The formulation principally differs in its rules for disjunction[16], which follow the usual introduction rules:

$$\frac{A}{A \oplus B}\oplus\mathcal{I} \qquad \frac{B}{A \oplus B}\oplus\mathcal{I}$$

but in the elimination rule we attach an additional constraint:

$$\frac{\begin{array}{c}\vdots\, d\\ A \oplus B\end{array} \quad \begin{array}{c}[A]^x\\ \vdots\, d'\\ C\end{array} \quad \begin{array}{c}[B]^y\\ \vdots\, d''\\ C\end{array}}{C}\oplus\mathcal{E}$$

where we insist that only the assumptions appearing in the above proof tree may figure in the derivations d' and d'', ie. there are no side-assumptions in the auxiliary premisses.

It is easy enough to show that analytic harmony is satisfied for the above connective: the eliminative clause to the inversion principle is unaffected by the weakening of the eliminative rule. We can demonstrate that conservativity fails however: in the presence of of just the usual rule for conjunction it is not possible to derive the distributivity law $A \wedge (B \oplus C) \vdash (A \oplus B) \wedge (A \oplus C)$ of disjunction over conjunction, but in the presence of the usual law of disjunction it is possible to derive this law *even though the connective features nowhere in the law*. Note that this non-conservativity is of a different kind to the kind exhibited by Prior's tonk connective: tonk permitted us to derive consequences in which the connective is

[16]The significance of these rules is that they block the derivation of the distributive law of disjunction.

not present, whereas quantum disjunction allows the provability relation to turn on what connectives, formally unconnected with the judgement at hand, may be present. We distinguish the former kind as failure of *downwards conservativity* to be contrasted with the latter more subtle failure of *upwards conservativity*.

Now we have identified the danger, can we find an alternative basis for justifying synthetic harmony? Since we wish to avoid the formal incoherence arising from upwards non-conservativity, we must begin with an examination of the way in which such examples violate the principal path property, and so conservativity.

For derivations containing an indirect elimination, a path is allowed to move from the principal subderivation to the auxiliary subderivations in the way prescribed for each rule, and then from the auxiliary subderivation to the conclusion. So for disjunction:

$$\dfrac{\begin{array}{c}\vdots\, d\\ A\vee B\end{array}\quad \begin{array}{c}[A]^x\\ \vdots\, d'\\ C\end{array}\quad \begin{array}{c}[B]^y\\ \vdots\, d''\\ C\end{array}}{\begin{array}{c}C\\ \vdots\, d'''\\ D\end{array}}\vee\mathcal{E}$$

the principal paths take one of the following two forms:

- the catenation of the paths d, d' and d''';
- the catenation of the paths d, d'' and d''';

As Prawitz observed, the formula 'C' may occur as part of a detour, with the introduction rule whose conclusion is C appearing as the last rule of an auxiliary premiss, and the matching elimination rule (whose principal premiss is C) appearing beneath the conclusion. In this case the $\vee\mathcal{E}$ rule gets in the way of the detour elimination. If the connective satisfies the strong form of the decomposition clause these detours can still be eliminated, since the principal premiss of the elimination rule can always be converted to an eliminative redex, and then reduced, thus 'unblocking' the detour.

Prawitz's came up with a general solution to eliminate detours even where the strong form of the elimination clause does not apply: it is possible to prove the subformula property if it is possible to show that one can systematically apply *commuting conversions* to permute the elimination rule up into the auxiliary premiss: this method succeeds even where we can show only the weak form of the decompositional clause. The general form of such a commuting conversion is as follows:

$$\dfrac{\dfrac{\begin{array}{c}\vdots\, d\\ A\vee B\end{array}\quad \begin{array}{c}[A]^x\\ \vdots\, d'\\ C\end{array}\quad \begin{array}{c}[B]^y\\ \vdots\, d''\\ C\end{array}}{C}\vee\mathcal{E}}{D}\,r \quad\to^*\quad \dfrac{\begin{array}{c}\vdots\, d\\ A\vee B\end{array}\quad \begin{array}{c}[A]^x\\ \vdots\, r\\ D\end{array}\quad \begin{array}{c}[B]^y\\ \vdots\, r\\ D\end{array}}{D}\vee\mathcal{E}$$

where r is any elimination rule of our proof theory, and C is the principal premiss of r. We note that in the case of quantum disjunction that we cannot apply the commutation of the elimination rule for ordinary disjunction in the indirect proof of the distributive law.

Can this observation be used to give a formal requirement strong enough to justify synthetic harmony, but weak enough to prove harmony for equality? Here is our candidate formulation:

DEFINITION 98 (INVERSION PRINCIPLE)

Permutative decomposition A binary connective $\otimes$ with an indirect elimination rule satisfies *permutative decomposition* in the context of a theory Λ if it satisfies the following four properties.

Weak decomposition $x : A \otimes B \vdash d : A \otimes B$ for some term d[17];

Identity I $x : A \otimes B \vdash x =_{A \otimes B} d$ true;

Commutation For each connective $\oplus$ of the system Λ, it is possible to commute the elimination rule $\oplus\mathcal{E}$ up through the elimination rule $\otimes\mathcal{E}$, if the conclusion of the latter is the principal premiss of the former;

Identity II Each of the commutative conversions of the previous condition express identities in the theory.

Note that this clause differs from that of the original formulation in an important regard: it is proved relative to a *particular collection of logical rules*, rather than once and for all as in the formulation of the original inversion principle. Let us examine the individual clauses in particular. We have discussed the commutative conversion requirement and how it is needed to avoid the upwards non-conservativity we observed with quantum disjunction. By itself the commutation requirement is not sufficient to ensure synthetic harmony: the obviously defective connective $\wedge^*$ whose rules are given

$$\frac{A \quad B}{A \wedge^* B} \qquad \frac{\begin{array}{c}\vdots\\ A \wedge^* B\end{array} \quad \begin{array}{c}A\\ \vdots\\ C\end{array}}{C}$$

satisfies the commutation condition; consequently we see that we need the weak decomposition condition to ensure the elimination rules match the introduction rules in strength. Finally the identities are needed to ensure that the principle yields harmony for constructions as well as for assertions.

EXERCISE 99 The reader is invited to satisfy himself that the above principle is adequate to show harmony for the usual natural deduction formulation of disjunction, provided suitable conversions in the reduction theory.

[17]Note we have dropped the requirement of the last rule of d being an introduction rule.

Now let us apply our new formulation to the $=_A$ type former. We note that, although the type former generally resembles an indirect rule, as described in section 2.2, it does not quite match the description there, as the conclusion of the elimination rule differs from its auxiliary premiss in its eigenformulae. This variation renders the Prawitzian upwards commutation of elimination rules impossible in the case of the $\Pi - \mathrm{e}$ rule.

All is not lost however: because the $= -\mathrm{e}$ rule has just one auxiliary premiss, it is possible to commute introduction rules downwards to obtain precisely the same shortening of detours as Prawitz's method.

PROPOSITION 100 The rules attached to the propositional equality type former satisfy the permutative decompositional clause in the fragment of ITT restricted to the Π, $=_A$ and $\mathbb{N}$ type formers, subject to the amendment of the convertive sub-clause to downwards commutation of introduction rules.

PROOF All clauses are shown quite elementarily. □

Unfortunately we *are* forced to abandon the claim of synthetic harmony in the presence of the strong existence operator: both upwards and downwards commutation are unprovable for this clause. We shall not conclude directly that synthetic harmony fails for propositional equality as a result, since there is another difficult issue related to equality and strong existence whose resolution might cast light upon this difficulty. We shall return to this matter in the conclusion.

The failure of synthetic harmony for $\mathbb{N}$ As for the equality type, it is possible to satisfy the eliminative clause of the inversion principle but not the strong form of the decompositional clause. While the weak form of the decompositional clause (together with the respective identity condition) can be satisfied, we cannot hope to extend this to a justification of synthetic harmony as upwards conservativity fails for this connective.

PROPOSITION 101

1. There are functions of type $\mathbb{N} \Rightarrow \mathbb{N}$ definable in ITT only with use of rules connected to the Π type former;

2. There are quantifier free formulae $\phi(n)$ such that $n : \mathbb{N} \vdash \phi(n)$ true is provable only in the presence of the Π type former.

PROOF SKETCH The first part is justified by noting that all functions of type $\mathbb{N} \Rightarrow \mathbb{N}$ that do not use term formers associated with the Π type former are definable in PRA, and so Ackermann's function is not amongst them, though it may be defined in ITT. The second part is obtained by coding up the totality of Ackermann's function using the Kleene T-predicate: this formula is provable in Heyting Arithmetic, and so ITT, but not using only Σ^0_1-IND. □

The conclusion we must draw from this is that, since the meaning of an arithmetic proposition determines what truth-value it may have, the meaning of propositions involving integers are not fixed just by the rules associated with type formers occurring in the constitution of the proposition.

This is a weak kind of holism, and it is an inevitable consequence of the *extensibility of arithmetical concepts* that is a direct conclusion of the results of Gödel and Turing: no effective proof theory can capture all of the truths of arithmetic, and no set of constructions can describe all of the arithmetic functions. Our immediate conclusion is that we have to relax the demands of harmony we make for non-logical concepts. This raises a number of important issues, which I will briefly touch upon in the conclusions.

2.8 A note about semantics

Martin-Löf defines the semantics for his type theory in terms of the canonical forms at each type, and shows the soundness of his theory by means of the head-normal form theory. This means that Martin-Löf's semantics is a *verificationist* theory, since it holds that the meaning of each type is determined by the introduction rules for that type. Such a semantics suffers from a particular defect: the semantics for types $\Pi x \in A.B(x)$ is not compositional, since it is not determined by our account of meaning for its subformula, but instead is a *substitutional* theory that interprets the meaning of this type in terms of the potentially infinite conjunction of the meanings of each $B(a)$ for each $a \in A$.

It is possible to give a semantics that does not suffer this defect by giving a *two factor* semantics which associates the meaning of each type former with on the one hand, its *assertoric content* (or *constructive content* in the case of collection types), and on the other hand its *hypothetical contribution* (or *constructive contribution*), both of which are to be determined in terms of the respective subformulae. The assertoric content determines what grounds suffice to make a given assertion (or similarly, constructive content determines what constructions are necessary to produce an element of that type), whilst hypothetical contribution determines the contribution that may be drawn from a given hypothesis or argument in providing assertoric or constructive content. Let us make a sketch of how these ideas might be used in providing a semantics in practice, by examining the cases of the Π type former and propositional equality in detail.

Π types may have either assertoric or constructive content, depending upon whether they are viewed as propositions or sets. The assertoric/constructive content of $\Pi x \in A.B(x)$ depends upon the hypothetical contribution of A and the assertoric/constructive content of $B(x)$, where x specifies the hypothetical contribution provided by A. This content is specified exactly by the introduction rule for Π: the innovation is that we may interpret the assumption directly with our other two factor.

The hypothetical contribution of $\Pi x \in A.B(x)$ is given by a means of trans-

forming the content s of assertions/constructive sets of type A into hypothetical contributions of type $B(s)$. Again this is specified precisely by the elimination rule.

For the propositional equality type $s =_A t$, which may be considered to be atomic and always an assertion, we have that the assertoric content is determined by the conceptual equality relation associated with the type A: the grounds for making an assertion of type $s =_A t$ are having a demonstration that s and t are conceptually equal. This formulation is strictly weaker than the introduction rule, since conceptual equality is a coarser equivalence than convertive equality. This is necessary due to the fact that it may be possible to assert $s =_A t$ in the presence of hypotheses when s and t are not equivalent under conversion.

We may use the head normal form theorem to derive the following refinement of assertoric content: an assertion consists either of the introduction rule for propositional equality, or of a $\mathbb{N}$ or $=_A$ elimination rule whose principal premise is a free variable. Consequently the available hypothetical contribution determines the possible assertoric content of a propositional equality, and in the case where no hypothetical contribution is available, then the normal form theorem assures us that conceptual equality coincides with convertive equality.

The conceptual contribution of $=_A$ is rather easier to describe. The contribution of the formulae $s =_A t$ is to allow us to recast any type ϕ into a new type ϕ' obtained from ϕ by replacing some number of occurrences of s in ϕ by t and vica versa.

The meanings of the type formers Σ and $\mathbb{N}$ are not difficult to provide, though in the case of $\mathbb{N}$ we note the caveat that the assertoric content of a construction of $\mathbb{N}$ is given by a function whose return type is $\mathbb{N}$, and due to the failure of synthetic harmony, we cannot circumscribe the possible methods that may be used in generating this function, except relative to a particular formal system.

Chapter 3

Classical proofs I: propositions

3.1 Classical strength reasoning

It is our aim in this chapter to introduce an extension to the theory developed in chapter 1 that is capable of representing the full theory of classical propositional logic, and which retains many of the desirable properties of the system we have developed for minimal logic. However it will be useful to begin by discussing forerunners of our account, to highlight the pitfalls that confront a theory of classical logic.

Our first step is to consider the system proposed by Gentzen. Gentzen's original system of natural deduction came in two parts: firstly he developed an account of negation in the intuitionistic part, called NJ, and then he obtained an extension to classical logic, NK, by adding an axiom scheme corresponding to the principle of the excluded middle.

The rules for negation are defined using an auxiliary zero-place logical connective[1], called absurdity, as follows:

$$\dfrac{\begin{array}{c}[A]^x\\ \vdots\\ \bot\end{array}}{\neg A}\,(x)\neg\mathcal{I} \qquad \dfrac{\neg A \quad A}{\bot}\,\neg\mathcal{E} \qquad \dfrac{\bot}{A}\,\bot\mathcal{E}$$

and the principle of the excluded middle is given by the axiom scheme:

$$\dfrac{}{A \vee \neg A}\,\text{PEM}$$

The complexity of introducing rules, such as $\neg\mathcal{E}$ above, which involve not one but two operators is alleviated by the fact that the above definition of negation is

[1]Gentzen is ambivalent about whether it is to be considered a zero-place connective, or as a propositional constant denoting falsehood, a distinction which will influence how we are to treat its justification.

equivalent to the synthetic definition within the system: $\neg A \triangleq A \supset \bot$, using only the rule $\supset \mathcal{E}$.

Gentzen's strategy unfortunately has severe drawbacks. The most major drawback is that the rule PEM is an introduction rule that does not exist in any kind of duality with the corresponding elimination rule (ie. for disjunction). Consequently, we lose the good properties of the calculus, such as the subformula property, and the principal path lemma. The alternative he mentions, that of the stability rule:

$$\frac{\neg\neg A}{A}$$

also has this drawback. Another drawback is that the $\bot\mathcal{E}$ rule only satisfies half of the condition for logical harmony, the analytic part, of Belnap's criteria (which it satisfies vacuously), it does not satisfy synthetic harmony, suggesting that the elimination rule is stronger than the introduction rule.[2]

These difficulties lead to a claim by some meaning-theory oriented constructivists that classical logic is incoherent[3]. In brief, the argument runs:

1. We have two contested accounts of logic, the constructivist account whose principles only contain the theory of minimal logic so far developed, and the classical account which contains the rules $\neg\mathcal{I}$, $\neg\mathcal{E}$ and $\bot\mathcal{E}$ as well.

2. The two accounts share the same sub-language of atomic formulae and the connectives $\wedge$ and $\supset$, but they differ in the judgements derivable. For example $\neg\neg X \vdash X$ is a valid judgement of classical logic, but it cannot be derived in intuitionistic logic.

3. Belnap's criteria include a requirement of conservativity: since in general we cannot expect derivations involving subformulae whose principal connective is negation to be eliminable by any account of normalisation, the classical rules involving negation are not conservative over derivations not involving negation. Thus they fail the requirements we demanded of any logical connective, and so they should not be so considered.

Reflection upon this argument shows that there is a weakness in it: we have only shown non-conservativity for a particular formulation of classical logic. The argument will, however, be effective against all attempts to introduce classical logic by using either special classical strength rules, or by the addition of axiom schemes.

From the natural deduction perspective, it does not seem that there is much alternative to providing classical strength provability by changing the rules governing the connectives. However, from the perspective of the sequent calculus,

[2]We also note that it is not possible to formulate a right logical rule for $\bot$ in the sequent calculus.

[3]This claim is expressed clearly in Dummett's 'The philosophical basis of intuitionistic logic' [DummettMAE:phibil], and again on page 290 of his *The Logical Basis of Metaphysics* [DummettMAE:logbm].

another formalisation of logic due to Gentzen, it becomes possible to attribute the strength of classical logic to another source.

We shall briefly present the sequent calculus in an informal manner. Gentzen's calculus LK has as its basic judgement sequents of form $\Gamma \vdash \Gamma'$ where Γ, Γ' are sequences of formulae. We may start a derivation with a zero premiss rule, called hyp whose schema is $X \vdash X$, and we have rules governed by logical connectives, which come in two dual varieties, left and right logical rules. The rules for $\supset$ are given:

$$\frac{\Gamma \vdash A, \Gamma' \quad \Delta, B \vdash \Delta'}{\Gamma, \Delta, A \supset B \vdash \Gamma', \Delta'} \supset \mathcal{L} \qquad \frac{\Gamma, A \vdash B, \Gamma'}{\Gamma \vdash A \supset B, \Gamma'} \supset \mathcal{R}$$

In addition, we require explicit structural rules:

$$\frac{\Gamma_0, A, B, \Gamma_1 \vdash \Gamma'}{\Gamma_0, B, A, \Gamma_1 \vdash \Gamma'} \mathrm{l-exch} \qquad \frac{\Gamma \vdash \Gamma'_0, A, B, \Gamma'_1}{\Gamma \vdash \Gamma'_0, A, B, \Gamma'_1} \mathrm{r-exch}$$

$$\frac{\Gamma \vdash \Gamma'}{\Gamma, A \vdash \Gamma'} \mathrm{l-wk} \qquad \frac{\Gamma \vdash \Gamma'}{\Gamma \vdash A, \Gamma'} \mathrm{r-wk}$$

$$\frac{\Gamma, A, A \vdash \Gamma'}{\Gamma, A \vdash \Gamma'} \mathrm{l-con} \qquad \frac{\Gamma \vdash A, A, \Gamma'}{\Gamma \vdash A, \Gamma'} \mathrm{r-con}$$

And finally we have the cut rule:

$$\frac{\Gamma \vdash A, \Gamma \quad \Delta, A \vdash \Delta'}{\Gamma, \Delta \vdash \Gamma', \Delta'} \mathrm{cut}$$

Derivations are built up inductively by gluing together subderivations using the rule applications arising from the above rules by instantiating propositional formulae in the usual manner, with Γ, Δ, etc. being schematic letters standing for lists of formulae. Intuitively we interpret a sequent of the form $A_1, \ldots, A_n \vdash B_1, \ldots, B_m$ as being provable when the formula $(A_1 \wedge \ldots \wedge A_n) \supset (B_1 \vee \ldots B_m)$ is.

It is beyond the scope of this brief exposition to give a precise account of the relationship of the sequent calculus to natural deduction, but it will be useful to point out a few surface analogies. The left logical rules play a role analogous to the elimination rules, and the right logical rules to introduction rules. 'hyp' plays a role analogous to assertion, and the left structural rules are related to the conventions governing assumption packets. The cut rule guarantees compositionality, and Belnapian duality is shown for the system by means of two meta-theorems, the first showing that derivations involving cut can be rewritten as derivations without cut (the cut-elimination theorem), and secondly that instances of 'hyp' where the concerned formula is not atomic can be rewritten as subderivations involving only applications of 'hyp' only using atomic formulae.

The sequent calculus has two advantages over natural deduction from the point of view of logical analysis: cut-elimination is more obviously related to duality than

maxima elimination, and analysis of the logical complexity of derivations is more straightforward than for natural deduction. Both of these advantages derive from the fact that in a certain sense, elimination segments are 'upside-down' compared to its equivalent in the sequent calculus, e.g. if there is a hypothesis of type $A \supset B$ in a cut-free derivation of the sequent calculus, it should be obtained by a left rule which may appear at the bottom of the derivation, whereas in natural deduction it appears upside down. Thus the left-rules of the sequent calculus create logical structure in the hypotheses rather than destroy it, as the elimination rules do in natural deduction. Nonetheless, the account of proof-equality in natural deduction is generally considered to be superior, especially in the light of the Curry–Howard correspondence, which for the purposes of this work is a decisive advantage. The interested reader is directed towards A. M. Ungar's excellent *Normalization, cut-elimination and the theory of proofs* [UngarAM:norcet].

So far, we have not discussed the right structural rules, which have no equivalent in the theory developed in section 1.2. It is these rules which introduce classical strength provability, as the following derivation of Peirce's law indicates:

$$\dfrac{\dfrac{\dfrac{\dfrac{}{A \vdash A}}{A \vdash B, A}\,\mathrm{r-wk}}{\vdash A \supset B, A} \qquad \dfrac{}{A \vdash A}}{\dfrac{(A \supset B) \supset A \vdash A, A}{(A \supset B) \supset A \vdash A}\,\mathrm{r-con}}$$

The strength of minimal logic can be recovered by eliminating the right structural rules, a restriction which enforces the invariant that valid judgements of minimal logic all have precisely one formula on the right-hand side. LJ is the calculus which arises from this restriction.

The possibility of providing an account of classical logic based upon structural rules undermines the previous argument for the formal unacceptability of classical logic, but it is possible to provide another, namely that classical logic cannot support an account of the semantic content of a proof. If true, this claim would undermine the use of the inversion principle in section 1.2, as its use in proving Belnapian duality depended upon being able to describe one derivation as equivalent to another.

Although this claim is not quite as deadly as the previous one (the proof of conservativity, in particular, does not depend upon any such conception of 'sameness of proof'), it would, if upheld, support the claim that intuitionistic logic has a significant advantage in formal clarity of meaning over classical logic. Let us now examine the grounds for the claim.

The argument advanced in support of the claim seeks to show that any notion of proof equality conserved by normalisation or cut-elimination must be degenerate, that is, if both d_0 and d_1 justify $\Gamma \vdash \Gamma'$, then d_0 and d_1 are equal.

The clearest argument of this form is due to Yves Lafont[4]. Let d_0, d_1 justify $\Gamma \vdash \Gamma'$. Then there is the following derivation justifying $\Gamma \vdash \Gamma'$:

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{}{B \vdash B}}{B \vdash C, B}\,\mathrm{r-wk} \quad \dfrac{\begin{array}{c}\vdots\, d_0\\ \Gamma \vdash \Gamma'\end{array}}{\Gamma, B \vdash \Gamma'}\,\mathrm{l-wk}}{\Gamma \vdash B \supset C, \Gamma'} \quad \dfrac{\begin{array}{c}\vdots\, d_1\\ \Gamma \vdash \Gamma'\end{array}}{\Gamma, B \supset C \vdash \Gamma'}\,\mathrm{l-wk}}{\Gamma, \Gamma \vdash \Gamma', \Gamma'}\,\mathrm{cut}}{\Gamma \vdash \Gamma'}\,\mathrm{l,r-con*}$$

Since the right formula eliminated by each cut arises by weakening, and presupposing a structural rule elimination if we have a weakening of a formula followed by a its contraction, we may obtain by applying either to the upper cut d_0 or to the lower cut d_1. Since equality is preserved by eliminating either cut, by transitivity we have that d_0 and d_1 are equal.

Another pessimistic result casting doubt upon the possibility of a good account of the semantic content of classical proofs is Joyal's theorem[5], which shows that a Cartesian closed category, a general model for the theory of constructions described in section 1.3, becomes degenerate if we allow a natural semantic equivalence between proofs of $\Gamma \vdash A$ and proofs of $\Gamma \vdash \neg\neg A$.

Again, arguments of this form are vulnerable to the criticism that they apply only to particular formalisations of classical logic, and it is the central claim of this work that we can give an account of the semantic content of classical proofs that avoids this class of difficulty.

Before we give this account, let us examine two issues. The first issue is to classify the strength of some usual theories of logic, and the second is to investigate possible formalisations for the negation and absurdity rule.

Whilst classical strength provability is normally introduced into natural deduction in the presence of the absurdity rule, our derivation of Peirce's law before shows that it need not be. We can strengthen the case for an 'absurdity independent' formulation of classical logic by observing that we can only avoid the force of the first (conservativity based) argument against classical logic in such a calculus, since otherwise arguments using classical reasoning are not conservative over the fragment without absurdity. Since the usual codification of classical strength provability via the stability rule is not absurdity independent, we shall wish to examine its relationship to classical strength laws which are.

DEFINITION 102

1. A *term constant* is a pair consisting of an identifier, possibly subscripted by

[4]Appendix B of *Proofs and Types* [GirardJY:prot].

[5]Joyal never published his result, but a proof due to Peter Freyd is given in Lambek and Scott's *Introduction to higher-order categorical logic*, chapter I.8 [LambekJ:inthoc]

parameters, and a formula, whose eigenformulae are parameters of the identifier.

2. All of the following are constant terms:

$$\begin{aligned} \mathcal{A}_A &: \bot \supset A \\ \mathcal{C}_A &: ((A \supset \bot) \supset \bot) \supset A \\ \mathcal{P}_{A,B} &: ((A \supset B) \supset A) \supset A \end{aligned}$$

   An *instantiation* of a constant term is a substitution of a propositional formula for the schematic letters occurring in the type and in the parameter.

3. An *axiomatic system* over a logical calculus consists of a number of constants together with the rules that govern them, and a number of constant terms. The valid derivations of an axiomatic system may have instantiations of constant terms as the leaves of derivations. If Λ is a logical system, and $\{C_1, \ldots, C_n\}$ is a set of constant terms, then $\Lambda + C_1 + \ldots + C_n$ is the axiomatic system obtained by adding the constant terms to the theory Λ.

4. The *theory of an axiomatic system* is the set of formulae in the axiomatic system which are justified by a closed derivation. If Λ is an axiomatic system, then $\mathsf{Th}(\Lambda)$ is its theory.

The theory of $\lambda + \mathcal{A}$ correspond to the usual propositional formulation of *intuitionistic logic*, and we christen the theory of $\lambda + \mathcal{P}$, *classical minimal logic*.

PROPOSITION 103

1. (a) $\mathsf{Th}(\lambda + \mathcal{A}) \subseteq \mathsf{Th}(\lambda + \mathcal{C})$
   (b) $\mathsf{Th}(\lambda + \mathcal{P}) \subseteq \mathsf{Th}(\lambda + \mathcal{C})$
   (c) $\mathsf{Th}(\lambda + \mathcal{C}) \subseteq \mathsf{Th}(\lambda + \mathcal{A} + \mathcal{P})$
2. (a) $\bot \supset X \notin \mathsf{Th}(\lambda + \mathcal{P})$
   (b) $((X \supset Y) \supset X) \supset X \notin \mathsf{Th}(\lambda + \mathcal{A})$

We leave the proof of the first part as an exercise to the reader, whilst the second part is a corollary of a result appearing later in the chapter.

Finally, we are interested in providing an account of absurdity and negation that are as far as possible compatible with Belnap's criteria. Following the discussion about introducing classical strength via the connectives, this means using intuitionistically acceptable rules.

Gentzen's rules for $\bot$, and the synthetic definition of $\neg$ in terms of $\supset$ and $\bot$, are usually considered to be intuitionistically acceptable, however as we noted earlier they do not satisfy Belnapian duality. In response to this argument, Michael Dummett proposes an infinitary introduction rule for $\bot$, where $\{A \ldots \Omega\}$ are all the formulae of the calculus:

$$\frac{A \ \ldots \ \Omega}{\bot} \bot\mathcal{I}$$

This allows us to justify the eliminative clause of duality for $\perp$ as follows:

$$\dfrac{\begin{array}{c}\vdots\, d_A\\ A\end{array} \quad \ldots \quad \begin{array}{c}\vdots\, d_U\\ U\end{array} \quad \ldots \quad \begin{array}{c}\vdots\, d_\Omega\\ \Omega\end{array}}{\dfrac{\perp}{U}}$$

which admits a reduction to d_U, and the decompositional clause:

$$\dfrac{\dfrac{\perp}{A} \quad \ldots \quad \dfrac{\perp}{\Omega}}{\perp}$$

However the use of infinitary rules necessitates some unstated assumptions about the semantics of proofs, which the requirement of explicitness in developing a meta-theory requires that we articulate, such as the existence of infinite sets of formulae, if we are to deduce the validity of induction from the ω-rule.

The unstated assumption in Dummett's rule may be articulated by making use of schematic variables. We interpret the introduction rule:

$$\dfrac{X}{\perp}\perp\mathcal{I}$$

where we must insist upon some *eigenvariable conditions*: firstly the only permissible instantiations of this rule are substitutions for X of other schematic variables, and secondly we only allow application of the rule where the principal premiss is not an eigen-parameter of any open assumption of the principal subderivation (without this second requirement, one could prove $X \vdash Y$).

We shall not use this reformulation for two reasons. Firstly, there is the author's prejudice that eigenvariable conditions have no place in the formulation of a natural deduction calculus, and secondly the intuitive subformula-hood relationship that exists between the principal premiss of an introduction rule and its conclusion, which we discussed briefly in section 1.2, fails to hold for this rule.

Instead we shall again appeal to algebra for our characterisation of absurdity. The $\supset\mathcal{E}$ rule already considered gives rise to the same theory as the rule:

$$\dfrac{A \vee \perp}{A}$$

in the presence of the usual rules governing disjunction. Thus $\perp$ may be considered not as a zero-place logical connective, but as a special propositional constant whose meaning is the unit of disjunction. Dually the truth constant may be understood as the unit of conjunction.

Since classical strength provability arises from permitting structural rules on the right, and in a sequent we understand multiple formulae on the right disjunctively, the algebraic characterisation leads to the following characterisation of $\perp$

and $\top$ in the sequent calculus:

$$\frac{\Gamma \vdash \bot, \Gamma'}{\Gamma \vdash \Gamma'} \bot - \text{const} \qquad \frac{\Gamma, \top \vdash \Gamma'}{\Gamma \vdash \Gamma'} \top - \text{const}$$

These formulations have the advantage that they do not explicitly depend upon disjunction and conjunction for their meaning; they may easily be verified to correspond to the usual rules by application of cut. We illustrate the use of the rule in the following example:

$$\dfrac{\dfrac{\dfrac{\dfrac{A \vdash A}{A \vdash \bot, A}\,\text{wk}}{\vdash A \supset \bot, A} \qquad \bot \vdash \bot}{(A \supset \bot) \supset \bot \vdash A, \bot}}{(A \supset \bot) \supset \bot \vdash A} \bot - \text{const}$$

Finally we may see a third reason for distinguishing the rule governing $\bot$ from the rules governing connectives: when we add $\bot - \text{const}$ to minimal logic, we violate the invariant we observed earlier, that there must be exactly one formula on the right hand side. In the case that there are no formulae to the right of the turnstile, we may permit the weakening of a single formula, recovering the strength of the $\bot\mathcal{E}$ rule. Nonetheless, this suggests that derivations with $\bot\mathcal{E}$ may have different properties when compared with those without.

3.2 Classical Natural Deduction

We have resolved to introduce classical strength provability into our system of natural deduction by means not tied to the rules governing the logical connectives, but by means relating to what we shall call 'book-keeping' rules. Moreover we have a model for this kind of characterisation in the sequent calculus where classical reasoning arises through permitting structural rules on the right-hand side of the sequent.

We shall achieve this by adapting Parigot's lambda-mu calculus. First, we will describe the calculus we shall use, and then briefly discuss the differences between it and Parigot's system. We shall call the natural deduction form of the calculus CND, and the system in which we formulate the lambda-mu calculus, $\lambda\mu$.

The first, and most fundamental, step in our calculus is that we provide syntactic support for concepts relating to the basic logical idea of contraraiety. Instead of a natural deduction derivation consisting purely of propositional formulae, formula occurrences in a derivation will also carry an indication of the propositional attitude of the formula, in that we allow formulae to be affirmed or rejected.

This distinction is captured in the analysis of a formula into two parts, the *attitude* of the formula, which indicates whether it is to be understood as being affirmed or rejected, and the *bare formula* which corresponds to the subject of the

attitude, ie. it is the same as the idea of formula in NJ. We indicate affirmations by 'A', where A is the bare formula, and rejections by '$\bullet A$'[6].

To accommodate formulae which occur as denials into our theory, we shall require new *structural* rules, of which there are three:

1. In addition to introducing affirmations as hypotheses by the assertion rule, we permit the introduction of rejections as assumptions by means of the rule of *denial*. As with assertions, denials belong to assumption packets, which in this calculus are tied to propositional attitudes as well as bare formulae. We annotate rejective assumption packets with Greek letters α, β, γ to distinguish them from assertions;

2. If we have derivations d, r with conclusions A and $\bullet A$ respectively, then we may form a new derivation which lacks a conclusion, called a *contradiction*, by means of the rule of *antithesis*. We indicate such a rule by a forward slash as follows:

$$\frac{\begin{array}{c}\vdots\, r\\ \bullet A\end{array} \quad \begin{array}{c}\vdots\, d\\ A\end{array}}{}\,/$$

3. If we can derive a contradiction in the presence of a rejective assumption packet $\alpha : \bullet A$, then we may discharge this assumption to obtain a deduction of A by means of the rule of *contraversion*. If $\alpha \in \mathsf{FV}(e)$, where e is a contradiction, the this is represented as follows:

$$\frac{\begin{array}{c}[\bullet A]^{\alpha}\\ \vdots\end{array}}{A}\,\mu\alpha$$

Observe that an application of the antithesis rule followed by a contraversion involves only a single horizontal rule line. We indicate this by combining both annotations into $\mu\alpha/$. We also indicate bound rejective assumptions by enclosing them in square brackets, just as we do with affirmative assumption packets.

Negative assumptions are handled using precisely the same conventions as positive assumptions. Denials introduce open rejective assumptions, and contraversions bind rejective assumptions. As with NJ, it is quite possible to bind empty packets of denials. There are no rules so far developed other than contraversion which may bind denials.

[6]In an earlier draft of this work, I had affirmations displayed '$\circ A$' to emphasise the symmetry between assertions and denials. However the additional annotations are cumbersome, and since this calculus is not symmetric, the inclusion is perhaps somewhat misleading. See the conclusions for a more symmetric formulation of classical natural deduction.

The following derivation of Peirce's law ($(A \supset B) \supset A \vdash A$) should help to illustrate these new ideas:

$$\dfrac{[\bullet A]^{\alpha} \qquad \dfrac{(A \supset B) \supset A^{y} \qquad \dfrac{\dfrac{[\bullet A]^{\alpha} \quad [A]^{x}}{B}\,\mu\beta/}{A \supset B}\,(x)}{A}}{A}\,\mu\alpha/$$

Note that the rule instantiation of contraversion binding $\beta : \bullet B$ is an empty assumption packet.

Finally we will also need to distinguish between three kinds of derivation, to cope with the fact that we may have an affirmative or rejective conclusion, or none at all; we will also wish to have three kinds of judgement which these derivations justify. The judgements of classical natural are distinguished from those of natural deduction by the subscripting of a 'c' to the turnstile as follows: $\Gamma \vdash_c A$.

1. A judgement whose formula to the right of the sequent has an affirmative propositional attitude is called an *affirmative judgement*. A derivation whose conclusion is an affirmation is called a *deduction*.

2. A judgement whose formula to the right of the sequent has rejective propositional attitude is called a *rejective judgement*. A derivation whose conclusion is a rejection is called a *refutation*.

3. A judgement which lacks a formula to the right of the sequent is called an *empty judgement*, and is written '$\Gamma \vdash_c$'. As we noted before a derivation without a conclusion is called a contradiction.

Just as the derivation of natural deduction may be put into a correspondence with the terms of the lambda calculus, so too the derivations of classical natural deduction may be put into a correspondence with terms of an extension of the lambda calculus, called the *lambda-mu calculus*. This correspondence we shall call the *classical Curry–Howard correspondence*. We shall introduce the correspondence as we did for natural deduction, beginning with the definition of term candidates.

DEFINITION 104

1. The grammar of terms are divided into three parts, where the elements of s are *affirmative terms*, those of e are *empty terms* and those of r are *rejective terms*. α ranges over assumption packets carrying denials.

$$\begin{aligned} s &::= \ldots \mid \mu\alpha : A.e \\ e &::= [r]s \\ r &::= \alpha \end{aligned}$$

2. Variable contexts may consist of either assertions or denials, ie. they are finite sets whose elements are either of the form $x : A$ or $\alpha : \bullet A$.

3. We have three kinds of type judgement where we had one before:

$$\Gamma \vdash s : A$$
$$\Gamma \vdash r : \bullet A$$
$$\Gamma \vdash e \ \mathsf{empty}$$

If any of the above judgements apply to a term candidate, then it is a term. We write $\Gamma \vdash \mathcal{J}(s)$ to indicate that under assumptions Γ, the terms s can be assigned an affirmative type or rejective type, or it can be judged empty.

4. We have three new type inference rules:

(a) Denial. If $\alpha : \bullet A \in \Gamma$ then

$$\Gamma \vdash \alpha : \bullet A$$

(b) Antithesis.

$$\frac{\Gamma \vdash r : \bullet A \quad \Gamma \vdash s : A}{\Gamma \vdash [r]s \ \mathsf{empty}}$$

(c) Contraversion.

$$\frac{\Gamma, \alpha : \bullet A \vdash e \ \mathsf{empty}}{\Gamma \vdash \mu\alpha^A.e : A}$$

Note that we omit the '$\bullet$' on the mu abstraction, since the abstraction binds only rejective variables. We shall repeat our proofs of the coherence of our type system.

DEFINITION 105

1. The *subterms* of a term candidate are a set of term candidates given by the following function:

$$\begin{aligned}
\mathsf{st}(x) &\triangleq \{x\} \\
\mathsf{st}(\lambda x : A.s) &\triangleq \{\lambda x : A.s\} \cup \mathsf{st}(s) \\
\mathsf{st}(\mathsf{ev}(s,t)) &\triangleq \{\mathsf{ev}(s,t)\} \cup \mathsf{st}(s) \cup \mathsf{st}(t) \\
\mathsf{st}(\langle s,t\rangle) &\triangleq \{\langle s,t\rangle\} \cup \mathsf{st}(s) \cup \mathsf{st}(t) \\
\mathsf{st}(\mathsf{outl}(s)) &\triangleq \{\mathsf{outl}(s)\} \cup \mathsf{st}(s) \\
\mathsf{st}(\mathsf{outr}(s)) &\triangleq \{\mathsf{outr}(s)\} \cup \mathsf{st}(s) \\
\mathsf{st}(\mu\alpha : A.e) &\triangleq \{\mu\alpha : A.e\} \cup \mathsf{st}(e) \\
\mathsf{st}([r]s) &\triangleq \{[r]s\} \cup \mathsf{st}(r) \cup \mathsf{st}(s) \\
\mathsf{st}(\alpha) &\triangleq \{\alpha\}
\end{aligned}$$

The *affirmative subterms* of a term candidate are given by $\mathsf{st}_\circ(s)$ which is the subset of $\mathsf{st}(s)$ restricted to affirmative term candidates. Analogously we have $\mathsf{st}_\bullet(s)$ for the *rejective subterms* and $\mathsf{st}_\epsilon(s)$ for the *empty subterms*.

2. The free variables $\mathsf{FV}(s)$ of a term candidate are defined by analogy with the definition in section 1.3, where we apply the same naming convention to mu abstractions. We have $\mathsf{FV}_\circ(s)$ to be the subset of $\mathsf{FV}(s)$ restricted to the variables arising by assertion, and $\mathsf{FV}_\bullet(s)$ to be the subset of $\mathsf{FV}(s)$ restricted to the variables arising by denial.

PROPOSITION 106

1. If s is a term, of any attitude, then every $t \in \mathsf{st}(s)$ is a term also;
2. If $\Gamma \vdash_{\lambda\mu} \mathcal{J}(s)$, then $\mathsf{FV}(s) \subseteq \Gamma$;
3. It is decidable whether a term candidate is a term or not.

PROOF As with the analogous proposition of section 1.3, parts 1 and 2 are elementary. Part three can be shown by providing a Hindley–Milner style type unification algorithm, which the author provides in an appendix. □

We are now in a position to provide the first two pillars of a Curry–Howard style correspondence for classical natural deduction:

1.

$$\begin{aligned} [\![A \wedge B]\!] &\hat{=} [\![A]\!] \times [\![B]\!] \\ [\![A \supset B]\!] &\hat{=} [\![A]\!] \Rightarrow [\![B]\!] \\ [\![\bullet A]\!] &\hat{=} \bullet [\![A]\!] \end{aligned}$$

2. (a) $[\![\bullet A^\alpha]\!] \hat{=} \alpha$

   (b) $\left[\!\!\left[\dfrac{\begin{matrix}\vdots\, r \\ \bullet A\end{matrix} \quad \begin{matrix}\vdots\, d \\ A\end{matrix}}{} \right]\!\!\right] \hat{=} [[\![r]\!]][\![d]\!]$

   (c) $\left[\!\!\left[\begin{matrix}[\bullet A]^\alpha \\ \vdots\, e \\ \overline{A}\ \mu\alpha/\end{matrix} \right]\!\!\right] \hat{=} \mu\alpha^A.[\![e]\!]$

PROPOSITION 107

1. If d is a derivation of CND, then $[\![d]\!]$ is a term;
2. (a) If d justifies $\Gamma \vdash A$, then there is a variable context Γ' with $\Gamma = \mathsf{dom}(\Gamma')$ such that $\Gamma' \vdash [\![d]\!] : A$;
   (b) If r justifies $\Gamma \vdash \bullet A$, then there is a variable context Γ' with $\Gamma = \mathsf{dom}(\Gamma')$ such that $\Gamma' \vdash [\![r]\!] : \bullet A$;
   (c) If e justifies $\Gamma \vdash$, then there is a variable context Γ' with $\Gamma = \mathsf{dom}(\Gamma')$ such that $\Gamma' \vdash [\![e]\!]\ \mathsf{empty}$.

PROOF SKETCH The key to proving this proposition lies in establishing an inductive mapping Hyp taking derivations to a variable context associating the open assumptions of a derivation with the proposition. □

The account we have just given is technically quite close to that of Parigot's. The main differences are that our account does not involve any explicit treatment of the type $\perp$, but instead treats empty terms as if they do not possess a type at all, and secondly we give an account of eta equality. There is another difference: Parigot's account allows us simply to 'disappear' rejective assumptions of type $\perp$, whilst we introduce a special rejective term former of type $\bullet\perp$. A consequence of this is that the theory without this special term former can be coded in the lambda calculus with Peirce's law.

PROPOSITION 108 $\mathsf{Th}(\lambda\mu) = \mathsf{Th}(\lambda + \mathcal{P})$

PROOF By translation of affirmative terms of $\lambda\mu$ into terms of $\lambda + \mathcal{P}$ and vice versa. We shall only consider the subsystem of λ with the single connective $\supset$; conjunction is quite elementary.

DEFINITION 109

1. The forward translation is a mapping of terms of $\lambda + \mathcal{P}$ onto $\lambda\mu$:

$$\begin{aligned}
[\![x]\!]_{\mathcal{P}} &\hat{=} x \\
[\![\lambda x^A.s]\!]_{\mathcal{P}} &\hat{=} \lambda x^A.[\![s]\!]_{\mathcal{P}} \\
[\![\mathsf{ev}(s,t)]\!]_{\mathcal{P}} &\hat{=} \mathsf{ev}([\![s]\!]_{\mathcal{P}}, [\![t]\!]_{\mathcal{P}}) \\
[\![\mathcal{P}_{A,B}]\!]_{\mathcal{P}} &\hat{=} \lambda h^{(A\Rightarrow B)\Rightarrow A}.\mu\alpha^A.[\alpha]\mathsf{ev}(h, \lambda x^A.\mu\beta^B.[\alpha]x)
\end{aligned}$$

2. We say a sequence of formulae σ is a *rejective covering* of a term s, if for each $\mu\alpha^A.e \in \mathsf{st}_\circ(s)$, A is an element of σ, and the sequence is at least two elements long. It is clear that we can find a rejective covering for any term.

3. If $\sigma \equiv \langle A_1, A_2, \ldots, A_n\rangle$, then $\bigwedge\sigma \hat{=} (A_1 \wedge (A_2 \ldots A_n))$, and $\mathsf{proj}_\sigma^{A_i} \hat{=} \mathsf{outl}(\mathsf{outr}^i(s))$. proj_σ^A is undefined if $A \notin \sigma$.

4. The reverse translation maps open rejective variables α onto affirmative variables α^*:

$$\begin{aligned}
[\![x]\!]_\mu^\sigma &\hat{=} x \\
[\![\lambda x : A.s]\!]_\mu^\sigma &\hat{=} \lambda x : A.[\![s]\!]_\mu^\sigma \\
[\![\mathsf{ev}(s,t)]\!]_\mu^\sigma &\hat{=} \mathsf{ev}([\![s]\!]_\mu^\sigma, [\![t]\!]_\mu^\sigma) \\
[\![[\alpha]s]\!]_\mu^\sigma &\hat{=} \mathsf{ev}(\alpha^*, [\![s]\!]_\mu^\sigma) \\
[\![\mu\alpha^A.e]\!]_\mu^\sigma &\hat{=} \mathsf{ev}(\mathcal{P}_{A,\bigwedge\sigma}, \lambda\alpha^* : A \Rightarrow \bigwedge\sigma.\mathsf{proj}_\sigma^A([\![e]\!]_\mu^\sigma))
\end{aligned}$$

LEMMA 110

1. If σ is a rejective covering of s, then $[\![(s)_\mu^\sigma]\!]$ is well-defined;
2. If $\Gamma \vdash_{\lambda+\mathcal{P}} s : A$ then $\Gamma \vdash_{\lambda\mu} [\![s]\!]_{\mathcal{P}} : A$;

3. (a) If $\Gamma \vdash_{\lambda\mu} s : A$, Γ is positive, and σ is a rejective covering of s, then $\Gamma \vdash_{\lambda+\mathcal{P}} [\![s]\!]^{\sigma}_{\mu} : A$.
   (b) If $\Gamma \vdash_{\lambda\mu} e$ empty, Γ is positive, and σ is a rejective covering of e, then $\Gamma \vdash_{\lambda+\mathcal{P}} [\![e]\!]^{\sigma}_{\mu} : \bigwedge \sigma$.

The proposition follows directly from the lemma, so it remains only to check the parts of the lemma. The first two parts are quite elementary, and so it remains only to inspect the last part, which requires an induction hypothesis that covers the cases where there are open denials. We observe that translating each open denial $\alpha : \bullet A$ to an open assertion $\alpha^* : A \Rightarrow \bigwedge \sigma$ suffices. □

We may obtain the use of the absurdity law by allowing a new kind of rejection, ie. we allow rejections to be introduced by means other than assumption. For the sake of symmetry, we shall also introduce a new type and term constructor to allow the tautology type.

We permit new CND rules:

$$\frac{}{\bullet\bot} \qquad \frac{}{\top}$$

and new term formers:

$$\begin{aligned} s &::= \ldots \mid * \\ r &::= \ldots \mid \mathcal{D}_{\bot} \end{aligned}$$

and new type inference rules:

$$\frac{}{\Gamma \vdash \mathcal{D}_{\bot} : \bullet\bot} \qquad \frac{}{\Gamma \vdash * : \top}$$

It is easy to verify that we conserve the properties of the Curry–Howard style correspondence so far derived. We can also see that the absurdity rule may be defined synthetically within this system:

$$\frac{\Gamma \vdash s : \bot \quad \alpha \text{ fresh}}{\Gamma \vdash \mu\alpha : A.[\mathcal{D}_{\bot}]s : A}$$

REMARK 111 We may also make an informal comment upon the relationship of this system to the system LK of the sequent calculus. If Γ and Δ are sets of affirmative formulae, then we may say that $\Gamma \vdash_{LK} \Delta$ iff $\Gamma, \bullet\Delta \vdash_{\lambda\mu}$. The details of the demonstration are left to the interested reader; for a hint one models the derivation of $A \vdash_{LK} A$ by axiom with the term $A^x, \bullet A^{\alpha} \vdash [\alpha]x$ empty.

3.3 Reduction and the inversion principle

The litmus test of the calculus lies in our ability to provide an account of equality on derivations that is conservative over the account of section 1.3, (ie. for s, t terms

of the lambda calculus, $\Gamma \vdash_\lambda s = t : A$ iff $\Gamma \vdash_{\lambda\mu} s = t : A$) and which supports a demonstration of Belnapian duality.

We shall begin by defining a new form of substitution, and use it to support the inversion principle. We shall the show that the simple Belnapian duality can be shown by a simple extension of Prawitz's proof. We shall then provide an account of reduction that is justified by this extended inversion principle, and show that normal forms of the lambda calculus are also normal forms of this extended form of reduction. A proof of Church–Rosser and strong normalisation will then vindicate the above claims.

The first step then, is to state our new inversion principle. The need for such a step is necessitated by the observation that the contraversion rule may introduce a formula governed by any logical connective. Thus, whilst it is not a true introduction rule, it is necessary to show that elimination rules can 'recover' content from a derivation whose principle subderivation arises by a contraversion.

As it happens, this is quite easy to achieve: we add a new 'contravertive part' to the inversion principle, which is defined as follows:

DEFINITION 112 (INVERSION PRINCIPLE)
Contravertive decomposition Let d be a derivation whose last rule is the contraversion of a denial governed by $\otimes$. Then it is possible to assemble from the parts of the immediate subderivation a derivation whose last rule is the introduction rule governing $\otimes$, and whose only other new rules are instances of the matching elimination rule.

To prove this, we need a new form of substitution:

DEFINITION 113 The *mixed substitution* $s[\alpha := f(-)/\beta]$ is defined for assertion identifiers α, β and single-ended contexts $f(-)$ recursively. In the following scheme is not capture-free: we depend upon the bound identifier convention to keep the calculus out of mischief:

$$\begin{aligned}
x[\alpha := f(-)/\beta] &\hat{=} x \\
(\lambda x : A.s)[\alpha := f(-)/\beta] &\hat{=} \lambda x : A.(s[\alpha := f(-)/\beta]) \\
\mathsf{ev}(s,t)[\alpha := f(-)/\beta] &\hat{=} \mathsf{ev}(s[\alpha := f(-)/\beta], t[\alpha := f(-)/\beta]) \\
\langle s,t\rangle[\alpha := f(-)/\beta] &\hat{=} \langle s[\alpha := f(-)/\beta], t[\alpha := f(-)/\beta]\rangle \\
\mathsf{outl}(s)[\alpha := f(-)/\beta] &\hat{=} \mathsf{outl}(s[\alpha := f(-)/\beta]) \\
\mathsf{outr}(s)[\alpha := f(-)/\beta] &\hat{=} \mathsf{outr}(s[\alpha := f(-)/\beta]) \\
(\mu\gamma^A.e)[\alpha := f(-)/\beta] &\hat{=} \mu\gamma^A.(e[\alpha := f(-)/\beta]) \\
([\alpha]s)[\alpha := f(-)/\beta] &\hat{=} [\alpha]f(s[\alpha := f(-)/\beta]) \\
([r]s)[\alpha := f(-)/\beta] &\hat{=} [r][s[\alpha := f(-)/\beta]], \text{ where } \alpha \not\equiv r
\end{aligned}$$

PROPOSITION 114 (MIXED SUBSTITUTION LEMMA)
Let $\Gamma, \alpha : \bullet A \vdash_{\lambda\mu} s : C$ and let $\Gamma \vdash f(x : A) : B$ be a context. Then $\Gamma, \beta : \bullet B \vdash_{\lambda\mu} s[\alpha := f(-)/\beta]$.

We are now in a position to show that the rules satisfy the inversion principle for $\wedge, \supset$:

1. (Contraversion for $\wedge$) Let $\Gamma \vdash_{\lambda\mu} \mu\alpha^{A\wedge B}.e : A \wedge B$. Then
$$\Gamma \vdash_{\lambda\mu} \langle \mu\beta_0^A.e[\alpha := \mathsf{outl}(-)/\beta_0], \mu\beta_1^B.e[\alpha := \mathsf{outr}(-)/\beta_1]\rangle : A \wedge B;$$

2. (Contraversion for $\supset$) Let $\Gamma \vdash_{\lambda\mu} \mu\alpha^{A\supset B}.e : A \supset B$. Then
$$\Gamma \vdash_{\lambda\mu} \lambda x^A.\mu\beta^B.e[\alpha := \mathsf{ev}(-, x)/\beta] : A \supset B.$$

There are two possible ways that we may define reduction rules which introduce the above contravertive decompositions as term equalities. The first is to allow the conclusion of a contraversion to be a minima, and permit the general scheme:
$$\frac{\Gamma \vdash f(x : A) : B \quad \Gamma, \alpha : \bullet A \vdash e \ \mathsf{empty}}{\Gamma \vdash f(\mu\alpha^A.[\alpha]e) \to^c \mu\beta^B.e[\alpha := f(-)/\beta] : B}$$
where $f(-)$ arises by the application of the appropriate elimination rule, and whose principal premiss is a variable. The contravertive part of the inversion lemma for each of $\wedge$, $\supset$ then follows from an eta reduction followed by this reduction which we call a *beta-form zeta conversion*.

Alternatively, we may define the reduction to be precisely that of the contravertive part of the inversion lemma, yielding the *eta-form zeta conversion* rules:

1. ($\zeta\wedge$) If $\Gamma, \alpha : \bullet A \wedge B \vdash_{\lambda\mu} e$ empty, then
$$\mu\alpha^{A\wedge B}.e \to^c_\zeta \langle \mu\beta_0^A.e[\alpha := \mathsf{outl}(-)/\beta_0], \mu\beta_1^B.e[\alpha := \mathsf{outr}(-)/\beta_1]\rangle;$$

2. ($\zeta\supset$) If $\Gamma, \alpha : \bullet A \supset B \vdash_{\lambda\mu} e$ empty, then
$$\mu\alpha^{A\supset B}.e \to^c_\zeta \lambda x^A.\mu\beta^B.e[\alpha := \mathsf{ev}(-, x)/\beta].$$

Whilst the beta-form rules are more elegant, and cast more light upon the extended duality in the calculus, they substantially complicate the proof of strong normalisation, and so we shall instead adopt the eta-form of zeta conversion.

The definition of maxima elimination is exactly the same as that in section 1.2, as the beta rule is defined by compatible closure. The definition of minima decomposition depends upon the definition of path, and in the presence of the new rules, this assumes a quite different form.

The essential innovation is that we consider paths not to be sequences of formula occurrences, but instead sequences of segments, where a segment is a sequence of formula occurrences bearing the same formula. The formula occurrences of a segment are related by contraversions and antitheses in the derivation. As we shall see in the conclusion, the notion of segment is a more complex form of that introduced by Prawitz[7] to show his analogous property for disjunctions.

[7]Chapter three of*Natural Deduction* [PrawitzD:natd].

DEFINITION 115

1. The *indirect premisses* of a contraversion in a derivation are the set of assertive premisses to antitheses whose reject premiss arises by a denial in the assumption packet bound by the contraversion. A *vacuous contraversion* is a contraversion with no indirect premisses;

2. The *indirect subderivations* of a contraversion are the subderivations above the indirect premisses of that contraversion. In other words, d is an indirect subderivation of $\mu\alpha^A.e$ when $[\alpha]d \in \mathsf{st}_\epsilon(e)$;

3. A *segment* of d is a sequence of formulae occurrences $\langle A_1, \ldots, A_n \rangle$ in d where the following properties are satisfied:

   (a) A_1 is not the conclusion of a contraversion;
   (b) A_n is not the indirect premiss of a contraversion;
   (c) For each $i \in \{1, \ldots, n-1\}$ A_i is the indirect premiss to a contraversion of which A_{i+1} is the conclusion.

   A formula occurrence is simple if it occurs only in segments of length 1.

4. A *path* in d is a sequence of segments $\langle \sigma_i \rangle$ in d such that for $i \leq i < n$, the last formula of σ_i is a premiss to a logical rule of which the first formula of σ_{i+1} is the conclusion. A path is *maximal* if it is not a strict subsequence of another path.

EXAMPLE 116 We give a proof of Peirce's law in which each formula is numbered:

$$\dfrac{{}^6[\bullet A]^\alpha \qquad \dfrac{{}^4(A \supset B) \supset A^z \qquad \dfrac{\dfrac{{}^1[\bullet A]^\alpha \quad {}^2[A]^x}{{}^3B}\,\mu\beta}{{}^5A \supset B}\,\lambda x}{{}^7A}}{{}^8A}\,\mu\alpha$$

The formulae 2 and 8 form a segment, as formula 2 is the indirect premiss of the contraversion to which formula 8 is the conclusion, which is to say that formula 2 appears as the affirmative premiss of the antithesis whose rejective premiss arises by the denial which is bound by the contraversion to which formula 8 is the conclusion. Similarly the formulae 7 and 8 form a segment. The maximal paths of the above derivation are $\langle\langle 2, 8\rangle\rangle$, $\langle\langle 3\rangle, \langle 5\rangle\rangle$ and $\langle 4, \langle 7, 8\rangle\rangle$.

PROPOSITION 117

1. There is a unique segment associated with any affirmative formula occurrence that is not the conclusion of a contraversion;

2. Every affirmative formula occurrence occurs in some principal path.

Contravertive decompositions are defined as follows: if $d \equiv \mu\alpha^A.e$ where A is a formula governed by $\otimes$, then $d \to^c_\zeta d'$ where d' is the alternative proof of $\Gamma \vdash_{\lambda\mu} A$. The relation $\to^1_\zeta$ is the compatible closure of $\to^c_\zeta$. A formula occurrence is a *zeta minima* if the subderivation above it is a contravertive decomposition.

Minima decompositions are defined so as to avoid the situation that a formula is more than one of a maxima, eta minima, or zeta minima. This is guaranteed by insisting that an eta minima must be a simple formula occurrence.

DEFINITION 118

1. A formula occurrence is *minimal* if
   (a) It belongs to the only segment on a principal path;
   (b) It belongs to the first segment on a principal path and it is the principal premiss of an introduction;
   (c) It belongs to the last segment on a principal path and it is the conclusion of an elimination;
   (d) It is the conclusion of an elimination and the principal premiss of an introduction.
2. A formula occurrence is an *eta minima* if it is a minima, it is governed by a connective, and it is a simple formula occurrence.
3. A *reducible formula occurrence*, or simply a *reducible* is one that is one of a maxima, an eta minima or a zeta minima.

Note that, confusingly, a zeta minima need not be a minima.

It is possible to describe two additional conversions, which involve only the new structural conversions, which we call the *mu conversions*. These are not required for the lowering of maxima or minima, but instead they shorten the length of segments in derivations. There are two mu conversions, the *mu-beta* and *mu-eta* conversions, so called due to an underlying similarity between these conversions and the $\Rightarrow\beta$ and $\Rightarrow\eta$ conversions

1. We allow the shorthand $[\alpha := r]$ to be shorthand for the mixed substitution operator $[\alpha := -/r]$. The mu-beta redex is given
$$[r]\mu\alpha.e \to^c_\mu e[\alpha := -/r]$$
2. Let s be a term such that $\alpha \notin \mathsf{FV}(s)$. The mu-eta redex is given
$$\mu\alpha[\alpha]s \to^c_\mu s$$

In each of the cases mu-beta and mu-eta, we shall say the associated reducible formula occurrence is the conclusion of the eliminated contraversion.

We may summarise the conversions of the lambda-mu calculus, eight in all, and the restrictions on eta conversion as follows:

1. The beta redex–contractum pairs of the lambda calculus are rule of the lambda-mu calculus, and the one-step beta reductions are defined to be the compatible closure in the same way;

2. The eta redex–contractum pairs of lambda-mu are also the same as those of the lambda-calculus. One-step eta reductions are subject to an additional restriction, that mu-abstractions are never eta redexes;

3. The zeta redex–contractum pairs are given

$$\mu\alpha^{A\wedge B}.e \to^c_\zeta \langle \mu\beta_0^A.e[\alpha := \mathsf{outl}(-)/\beta_0], \mu\beta_1^B.e[\alpha := \mathsf{outr}(-)/\beta_1]\rangle$$
$$\mu\alpha^{A\supset B}.e \to^c_\zeta \lambda x^A.\mu\beta^B.e[\alpha := \mathsf{ev}(-, x)/\beta]$$

   As with beta reduction, one-step zeta reduction is the compatible closure of '$\to^c_\zeta$';

4. The mu redex–contractum pairs are given

$$[r]\mu\alpha^A.e \to^c_\mu e[\alpha := -/r]$$
$$\mu\alpha^A.[\alpha]s \to^c_\mu s, \text{ if } \alpha \notin \mathsf{FV}(s)$$

   As with the beta conversions, the one-step reductions for the mu conversions are given by compatible closure.

PROPOSITION 119 (DIAMOND PROPERTY)
One-step reduction satisfies the diamond property.

PROOF Under the notion of overlapping redex that we described for the lambda calculus, we can find four new critical pairs in the lambda-mu calculus, with the general forms

1. $\mu\alpha.[\alpha]\mu\beta.e$, where $\alpha \notin \mathsf{FV}(e)$;
2. $[r]\mu\alpha.[\alpha]s$, where $\alpha \notin \mathsf{FV}(s)$;
3. $\mu\alpha^{A\wedge B}.[\alpha]s$, where $\alpha \notin \mathsf{FV}(s)$, and
4. $\mu\alpha^{A\supset B}.[\alpha]s$, where $\alpha \notin \mathsf{FV}(s)$.

In the first two parts, contracting either redex leads to the same term up to alpha equivalence, whilst in the last two parts there are two cases to consider. If applying the mu-eta reduction yields a term that admits an eta expansion, then it is possible to obtain this expanded term by applying the zeta redex followed by a mu-eta reduction, whilst if the result of applying the mu-eta reduction is not eta expandable, then we obtain this term by a zeta rewrite, a mu-eta rewrite and then a beta rewrite.

However, in the presence of zeta rewrites, with the attendant operation of mixed substitution, it is possible for reductions to destroy redexes associated with residual formula occurrences that are not immediately adjacent to the reduced formula. Zeta reductions can affect other reducible formulae in only one way, by placing

an elimination rule on the path between the affirmative premiss of an antithesis and the application of the antithesis itself. Consequently we extend the notion of overlapping redex to cover these redexes. We also note that a single zeta reduction may conflict with many other redexes, since for a given zeta redex $\mu\alpha^{A\otimes B}.e$, each subterm of $e\,[\alpha]s_i$ may overlap with a redex.

There are no critical pairs involving zeta and beta reducible formulae, and there are none involving two zeta reducibles, but there are conflicts between both zeta reducibles and eta reducibles for each of the type formers, and between zeta reducibles governed by each type and a mu-beta reducible. The critical pairs are

1. $\mu\alpha^{A\wedge B}.e$ where there is a subterm $[\alpha]s$ in e where s does not arise by a $\wedge\mathcal{I}$ rule;
2. $\mu\alpha^{A\supset B}.e$ where there is a subterm $[\alpha]s$ in e where s does not arise by a $\supset\mathcal{I}$ rule;
3. $\mu\alpha^{A\wedge B}.e$ where there is a subterm of the form $[\alpha]\mu\beta^{A\wedge B}.e'$ in e;
4. $\mu\alpha^{A\supset B}.e$ where there is a subterm of the form $[\alpha]\mu\beta^{A\supset B}.e'$ in e;

These conflicts are easily resolved. □

We shall now move onto proving the normal form theorem. As with the proof of section 1.2, we start by developing an account of the depth of formula occurrences, culminating in a depth coherence lemma, before proving the theorem proper.

The definitions of junction, branch and order are unaffected by the extensions to the calculus.[8] To define the depth measure, we will need to alter the definition of assumption binding to include contraversions of implicative type, as these may be reduced to $\supset\mathcal{I}$ assumption bindings. We shall also allow for the fact that formulae occurrences of the form $A\supset B$ may be eta minima, and that a contraversion of an assumption of a formula $A_1\supset(\ldots(A_n\supset B))$ may be reduced to a derivation containing assumption bindings bearing each formula A_i.

DEFINITION 120 An *assumption binding* is a subderivation of a given derivation which may bind an affirmative assumption packet. In addition to $\supset\mathcal{I}$, contraversions of an implicative formula are also assumption bindings, as are occurrences of implicative formula in minima. A formula A is bound by an assumption binding when it satisfies one of the following conditions:

1. For an $\supset\mathcal{E}$ assumption binding with conclusion $A\supset B$;
2. For a contraversion of an implicative assumption B, where there are formula $\{A_1,\ldots,A_n,C\}$ with $B\equiv A_1\supset(\ldots(A_n\supset C))$ and $A\equiv A_i$ for some i;
3. If the conclusion is a minima B in the ambient derivation and there are formula $\{A_1,\ldots,A_n,C\}$ with $B\equiv A_1\supset(\ldots(A_n\supset C))$ and $A\equiv A_i$ for some i.

[8]This has the effect that we have two parallel analyses of structure of each CND derivation, the branch structure and the path structure, which coincide for derivations of NJ.

DEFINITION 121

1. The depth of a formula occurrence is the sum of the depths of the junctions where a minor premiss appears on the terminal thread beneath the formula occurrence, and the sum of the depths of the $\wedge\eta$ and $\wedge\zeta$ reducibles occurring on this terminal thread;

2. The depth of an assumption binding is calculated for the assumption binding independently (see 1.2). For a $\supset\mathcal{E}$ assumption binding it is the maximum of all the depths of the assumptions bound by the concluding rule, whilst for a contraversion it is one plus this value;

3. The depth of a junction is one plus the depth of all the assumption bindings which bind the same formula as the minor premiss (from $\supset\mathcal{E}$ junctions) and whose conclusion does not appear on the terminal thread beneath the conclusion of the $\supset\mathcal{E}$ rule;

LEMMA 122

1. There is a unique depth associated with each formula occurrence;
2. Every formula occurrence in a given branch shares the same depth, and that depth is at least as large as its order.

The proof of this lemma is almost exactly the same as that in section 1.2.

DEFINITION 123

1. The one-step reduction relation on the $\lambda\mu$ calculus is defined

$$\rightarrow^1 \hat{=} \rightarrow^1_\beta \cup \rightarrow^1_\eta \cup \rightarrow^1_\zeta;$$

2. A *reduction sequence* is a sequence of terms $\langle s_i \rangle$ which may be either finite or infinite, with the property that if s_{i+1} is in the sequence then $s_i \rightarrow^1 s_{i+1}$;

3. A *full reduction sequence* is a finite reduction sequence whose last term is a normal form;

4. The *measure* of a formula occurrence is an ordinal $u{\cdot}\omega{+}v$ where u is the degree of its formula and v is the depth of the formula occurrence. The measure of a derivation is defined to be the natural sum of all the measures of its reducible formula occurrences.

LEMMA 124 Suppose $d \rightarrow^1 d'$. Then the measure of d' is strictly less than that of d.

PROOF In the lemma matching the simple normal form there were two reductions to consider. Here there are eight:

1. Maxima eliminations (or the $\wedge\beta$ and $\supset\beta$ rewrites);
2. Minima decompositions for $\wedge$ (ie. $\wedge\eta$ rewrites);
3. Minima decompositions for $\supset$ (ie. $\supset\eta$ rewrites);

4. Contravertive decompositions for $\wedge$ (ie. $\wedge\zeta$ rewrites);
5. Contravertive decompositions for $\supset$ (ie. $\supset\zeta$ rewrites);
6. Structural reductions (or the $\mu\beta$ and $\mu\eta$ rewrites).

In each case we must name the direct and indirect residuals, and show that

1. Each direct residual is the unique residual of its antecedent, and has measure no more than that of the antecedent;
2. Each indirect residual has measure strictly less than that of its antecedent;
3. No formula occurrence is reducible in the residual derivation that is not either the residual of a reducible in the antecedent, or has measure strictly less than that of the reduced reducible.

As before, these are adequate to establish the lemma in each of the eight cases.

1. By analogy with the previous lemma. The change in the definition of depth does not affect the argument, and it is clear that there are no new eta or zeta reducibles. Both of these reductions can create new $\mu\beta$ reducibles and the $\supset\beta$ reduction can create a new $\mu\eta$ reducible, but they are of measure less than the eliminated reducible;

2. ($\wedge\eta$ rewrite) The rewrite is of the form:

$$\begin{array}{c} \vdots\, d_0 \\ A\wedge B \\ \vdots\, d_1 \\ C \end{array} \quad \to^1_\eta \quad \begin{array}{c} \dfrac{\begin{array}{c}\vdots\, d_0\\ A\wedge B\end{array}}{A} \quad \dfrac{\begin{array}{c}\vdots\, d_0\\ A\wedge B\end{array}}{B} \\ \hline A\wedge B \\ \vdots\, d_1 \\ C \end{array}$$

   The direct residuals are those of d_1, and the indirect residuals are the formula occurrences of d_0. The $\wedge$-depth of the indirect residuals clearly decreases, and so we have the first clause, and the direct residual clause is straightforward (given that we note that the depth of assumption packets in d_1 may reduce). We may have new $\wedge\eta$ reducibles in the formula occurrences A, B, but they are of strictly lower measure than the decomposed $A\wedge B$ in the antecedent.

3. ($\supset\eta$ rewrite) The rewrite is of the form:

$$\begin{array}{c} \vdots\, d_0 \\ A\supset B \\ \vdots\, d_1 \\ C \end{array} \quad \to^1_\eta \quad \begin{array}{c} \dfrac{\dfrac{\begin{array}{c}\vdots\, d_0\\ A\supset B\end{array} \quad [A]^x}{B}}{A\supset B}\,(x) \\ \vdots\, d_1 \\ C \end{array}$$

All formula occurrences of d_0 and d_1 are direct residuals. To see that their depth does not increase we must check that

(a) The new $\supset\mathcal{E}$ junction does not increase the depth of any residual;

(b) The new $\supset\mathcal{I}$ assumption packet does not increase the depth of any junction from antecedent to residual.

The first part is immediate since there are no residuals on the new branch. The second part follows from the observation that the depth of the assumption packet is 1, and no junction has depth less than 2.

We observe that the only possible new reducibles correspond to the formula occurrences A, B which have strictly lower measure than the reduced antecedent.

4. ($\wedge\zeta$ rewrite) The rewrite is of the form:

$$\begin{array}{c}
\dfrac{[\bullet A \wedge B]^{\beta} \quad \begin{array}{c}\vdots\, d_0\\ A \wedge B\end{array}}{\begin{array}{c}\vdots\, e\\ \dfrac{}{A \wedge B}\,\mu\beta\\ \vdots\, d_1\\ C\end{array}}\,/
\end{array}
\quad \rightarrow^{1}_{\zeta} \quad
\begin{array}{c}
\dfrac{\dfrac{[\bullet A]^{\alpha} \quad \begin{array}{c}\vdots\, d_0\\ \dfrac{A \wedge B}{A}\end{array}}{\begin{array}{c}\vdots\, e\\ \dfrac{}{A}\,\mu\alpha\end{array}}\,/ \quad \dfrac{[\bullet B]^{\alpha'} \quad \begin{array}{c}\vdots\, d_0\\ \dfrac{A \wedge B}{B}\end{array}}{\begin{array}{c}\vdots\, e\\ \dfrac{}{B}\,\mu\alpha'\end{array}}\,/}{\begin{array}{c}A \wedge B\\ \vdots\, d_1\\ C\end{array}}
\end{array}$$

The indirect residuals are the formula occurrences of d_0 and e, and the direct residuals are the formula occurrences of d_1. An analogous argument to that of $\wedge\eta$ shows that the required bound on the residuals holds. However we must also consider the possible new reducibles: there are three of $A \wedge B$, and two each of A and B. Reflection on the definition of depth shows that the sum of their depths is less than before.

5. ($\supset\zeta$ rewrite) The rewrite is of the form:

$$\begin{array}{c}\dfrac{[\bullet A\supset B]^{\alpha}\quad \begin{array}{c}\vdots\, d_0\\ A\supset B\end{array}}{}\,/\\ \vdots\, e\\ \dfrac{}{A\supset B}\,\mu\alpha\\ \vdots\, d_1\\ C\end{array}\quad \to^1_{\zeta}\quad \begin{array}{c}\dfrac{[\bullet B]^{\beta}\quad \dfrac{\begin{array}{c}\vdots\, d_0\\ A\supset B\end{array}\quad [A]^x}{B}\,/}{}\\ \vdots\, e\\ \dfrac{}{B}\,\mu\beta\\ \dfrac{}{A\supset B}\,(x)\\ \vdots\, d_1\\ C\end{array}$$

All the formula occurrences of d_0, e and d_1 are direct residuals, and there are no indirect residuals. Thus the same arguments as in $\supset\eta$ deals with these clauses. The new junctions we have created are also trivial for the same reasons as in the $\supset\eta$ case. It remains to check firstly that the new lambda assumption packet and mu assumption packet have depth less than the old mu assumption packet, and that the potential new beta, eta and zeta reducibles have measures less than the decomposed reducible. These are established by inspection.

6. ($\mu\beta$ and $\mu\eta$ rewrites) The rewrites are of the form:

$$\begin{array}{c}\dfrac{[\bullet]^{\beta}\quad \begin{array}{c}\vdots\, e\\ \dfrac{}{A}\,\mu\alpha\end{array}}{}\,/\\ \vdots\, d\\ B\end{array}\quad \to^1_{\mu}\quad \begin{array}{c}\vdots\, e\\ \cdot\\ \vdots\, d\\ B\end{array}$$

and

$$\begin{array}{c}\dfrac{[\bullet A]^{\alpha}\quad \begin{array}{c}\vdots\, d_0\\ A\end{array}}{A}\,\mu\alpha/\\ \vdots\, d_1\\ B\end{array}\quad \to^1_{\mu}\quad \begin{array}{c}\vdots\, d_0\\ A\\ \vdots\, d_1\\ B\end{array}$$

In both of these cases it is clear that there are no indirect residuals, and that we eliminate a contraversion without introducing anything that might increase the measure of any direct residual. In the first case it is possible that the rule separating d from e in the residual might be the site of a $\mu\beta$ or $\mu\eta$ reducible, in which case we notice that it existed in the antecedent.

In the second case we note that the formula occurrence A in the residual derivation might be a beta, eta or zeta reducible. If it is a zeta reducible, then it

existed in the antecedent. If it is a beta or eta reducible, then we note that its depth is less than the rewritten reducible.

□

THEOREM 125 (NORMAL FORM)
Let d be a derivation of $\Gamma \vdash A$ in $\lambda\mu$. Then every reduction sequence whose first term is d is finite and is an initial subsequence of a full reduction sequence.

As with the simple normal form theorem of section 1.2, this theorem follows easily from the previous lemma, and the well-foundedness of the ordinal ω^ω. In view of the diamond property, we have an immediate corollary.

COROLLARY 126 (CHURCH–ROSSER)
One-step reduction is Church–Rosser in lambda-mu.

PROPOSITION 127 (PRINCIPAL PATH LEMMA)
If d is a normal affirmative derivation of classical natural deduction, then any principal path in d consists of the conjunctive catenation of an analytic and a synthetic path.

Furthermore the analytic and synthetic paths consist entirely of singleton segments except in their last and first parts respectively.

PROOF The analogous proof of section 1.2 can be generalised from formula occurrences to segments. It remains to show that, apart from the exceptional segment, there are no contraversions on a principal path.

This is shown easily enough, since all of the other segments are not atomic, and so would admit a contravertive decomposition. □

The property relating to segments can be rephrased as saying that in a normal derivation we can only have contraversions whose conclusion is atomic. The same applies to antitheses where there are no open denials, but in the case where there are, an antithesis may appear at the last segment of an analytic thread, as in the following example:

$$\dfrac{\bullet A \wedge B^{\alpha} \quad \dfrac{A^x \quad B^y}{A \wedge B}}{C}\,\mu\beta/$$

PROPOSITION 128 (THE SUBFORMULA PROPERTY)
If d is a normal derivation then the bare formula in all formula occurrences of d are subformulae of the bare formula in an open assumption or of the conclusion.

PROOF The proposition is proven by induction over branches of d, but the proof is somewhat complicated by the fact that branches need not be principal threads in classical natural deduction.

Before we describe the inductive proof, we must first establish the following two propositions:

1. An assertion of order i is either discharged by an $\supset\mathcal{E}$ rule whose conclusion is a formula occurrence of order no more than i, or it contains a formula in Γ;
2. The last formula of every synthetic thread is either the last formula of the branch, or it is the affirmative premiss of an antithesis whose rejective premiss is an open denial in Γ. (*)

The first proposition is established just as in section 1.2. The second proposition is established by noting that the formula occurrence can't be atomic as it is the conclusion of some introduction rule. But if it is bound, then there must be a contravertive reducible formula occurrence in the derivation, contradicting normality.

We now endeavour to establish that all formula occurrences of order i are subformulae of $\{\phi^* \mid \phi \in \Gamma \cup \{A\}\}$ (where ϕ^* is the bare formula in ϕ) by induction. We observe that it suffices to show the result for:

1. The first formula of a non-trivial analytic thread or order i;
2. The last formula of a synthetic thread of order i,

since all other formula occurrences of a derivation are subformula of one or other of these (for some i).

The last formula of a synthetic thread is either the last formula of the branch, or it is a subformula of some open assumption in Γ by (*) above. To complete this case we must show the last formula of the branch to be a subformula of $\Gamma \cup \{A\}$.

If the branch is of order 0, then this formula is the conclusion. Otherwise it is the minor premiss which is on a lesser branch, and so by the induction hypothesis we are done.

The first formula of a non-trivial analytic thread must arise by an assertion (since if it arose by a contraversion it would be reducible), and so either it is an open assumption or it is bound. If it is bound by an assumption packet whose conclusion has order strictly less than it, then we may apply the induction hypothesis. If it is bound by an assumption packet whose conclusion is of the same order, then by the argument for synthetic threads above, we are done. □

COROLLARY 129 (CONSISTENCY)
There is no derivation d which justifies $\vdash X$ where X is a schematic letter.

PROOF Following the logic of the consistency proof of section 1.2, d cannot be an assertion or the conclusion of an introduction or elimination rule. For a contradiction, assume the last rule of d is a contraversion. Then there is a normal derivation e justifying $\alpha : \bullet X \vdash e$ empty.

Such an e must be an antithesis of the form $[\alpha]d'$ (because a different assumption packet would be open, and $\perp \notin \mathsf{sf}(\{X\})$). But the same argument that we used first of all shows that the last rule of d' cannot be an assertion, introduction or elimination, nor can it be a contraversion, since this would create a $\mu\beta$ redex in d. □

Finally, we shall note that the terms of the $\lambda\mu$ calculus admit constructor extractor decompositions like those of the lambda calculus.

PROPOSITION 130 Every affirmative term has a unique constructor–extractor decomposition.

We note that, since there are no contraversions or antitheses in constructors, introducing the new classical strength term formers only affects the grounds upon which we may derive an atomic proposition from more logically complex propositions, at least when we consider cut-free proofs using only the $\supset$ and $\wedge$ connectives. We shall see the significance of this point in the conclusion.

3.4 Recursion

Recall that we said in section 1.5 that for the lambda calculus there was no clearly decisive grounds for choosing between the left and right recursive forms. Now we shall introduce the term formers of the lambda-mu calculus, the admission of zeta reduction changes this, because in a certain sense we may say that the left-recursive form is well-founded.

This gives two advantages to the left-recursive form: firstly, it is easier to characterise the left-recursive form in a way that justifies inductive reasoning, and secondly the left-recursive form admits a desirable zeta reduction that cannot be applied to the right-recursive form without losing confluence.

Both flavours of PRA^{ω}_{μ} admit the same type inference rules for the term formers as we used for the extensions to the lambda calculus. Our conversion rules for the two type formers are given as follows[9], where the type T is the inferred type of the cut-formula:

1. Left recursion

$$\begin{aligned}
\mathsf{Rec}(\underline{0}, a, f) &\rightarrow^{c}_{\beta} a \\
\mathsf{Rec}(\mathsf{succ}(s), a, f) &\rightarrow^{c}_{\beta} \mathsf{Rec}(s, \mathsf{ev}(f, \underline{0}, a), \lambda n^{\mathbb{N}}.\mathsf{ev}(f, \mathsf{succ}(n))) \\
\mathsf{Rec}(\mu\alpha^{\mathbb{N}}.e, a, f) &\rightarrow^{c}_{\zeta} \mu\beta^{T}.e[\alpha := \mathsf{Rec}(-, a, f)/\beta] \\
\mathsf{succ}(\mu\alpha^{\mathbb{N}}.e) &\rightarrow^{c}_{\zeta} \mu\beta^{\mathbb{N}}.e[\alpha := \mathsf{succ}(-)/\beta]
\end{aligned}$$

2. Right recursion

$$\begin{aligned}
\mathsf{Rec}(\underline{0}, a, f) &\rightarrow^{c}_{\beta} a \\
\mathsf{Rec}(\mathsf{succ}(s), a, f) &\rightarrow^{c}_{\beta} \mathsf{ev}(f, s, \mathsf{Rec}(s, a, f) \\
\mathsf{Rec}(\mu\alpha^{\mathbb{N}}.e, a, f) &\rightarrow^{c}_{\zeta} \mu\beta^{T}.e[\alpha := \mathsf{Rec}(-, a, f)/\beta]
\end{aligned}$$

[9] We omit a treatment of the strict-recursive flavour due to the reasons we cited in chapter 2.

The right recursive form cannot admit the zeta reduction on '$\mathsf{succ}(\mu\alpha^{\mathbb{N}}.e)$' as with this reduction the term '$\mathsf{Rec}(\mathsf{succ}(\mu\alpha^{\mathbb{N}}.[\beta]\underline{1}), \underline{1}, \lambda n^{\mathbb{N}}.\lambda z^{T}.\underline{0})$' is a bad critical pair.

PROPOSITION 131 Both of the theories of $\mathrm{PRA}^{\omega}_{\mu}$ satisfy the diamond property.

PROOF We have to consider the critical pairs that arise by adding the new recursor. For the right-recursive form there are none, whilst for the left-recursive form we have the critical pair:

$$\mathsf{Rec}(\mathsf{succ}(\mu\alpha^{\mathbb{N}}.e), a, f)$$

To show that the two contractions have the same normal form, it is necessary to show that the two substitutions

$$e[\alpha := \mathsf{Rec}(-, \mathsf{ev}(f, \underline{0}, a), \lambda n^{\mathbb{N}}.\mathsf{ev}(f, \mathsf{succ}(n)))/\beta]$$
$$(e[\alpha := \mathsf{succ}(-)/\gamma])[\gamma := \mathsf{Rec}(-, a, f)]$$

have identical results, which requires a structural induction to establish. □

The other weakness of the right-recursive form causes no difficulty in the formulation of $\mathrm{PRA}^{\omega}_{\mu}$, but it does cause difficulties when we introduce the right recursive form into type theory. We shall examine this matter further in section 4.1, but let us briefly examine the problem here. Consider the family of terms V_k

$$V_k \mathrel{\hat{=}} \mathsf{Rec}(\underline{k}, a, f)$$

Then we have

$$\begin{aligned}
&V_0 \to^* a\\
&V_1 \to^* \mathsf{ev}(f, \underline{0}, a)\\
&V_2 \to^* \mathsf{ev}(f, \underline{1}, \mathsf{ev}(f, \underline{0}, a))\\
&V_3 \to^* \mathsf{ev}(f, \underline{2}, \mathsf{ev}(f, \underline{1}, \mathsf{ev}(f, \underline{0}, a)))\\
&\quad \cdots
\end{aligned}$$

It is due to the unfolding nature of recursion that we obtain the possibility of reasoning about such families of terms inductively, and both the left- and right-recursive schemes are seen to be essentially value-equivalent since they agree on the normal form of each V_k.

However, with the move to $\mathrm{PRA}^{\omega}_{\mu}$ a difficulty emerges with the right-recursive form; we have closed normal forms of $\mathbb{N}$ which are not canonical forms, such as the term:

$$d \mathrel{\hat{=}} \mu\alpha^{\mathbb{N}}.[\alpha]\mathsf{succ}(\mu\beta^{\mathbb{N}}.[\alpha]\underline{0})$$

If we evaluate the term $\mathsf{Rec}(\mathsf{succ}(d), a, f)$ in one step we obtain

$$\mathsf{ev}(f, d, \mathsf{Rec}(d, a, f))$$

and so the outermost function f receives as its first argument a term which cannot be considered a natural number, and the term $V_{\mathsf{succ}^k(d)}$ will evaluate to term with k such troublesome applications. With the left-recursive form, this does not happen, as we build up 'from below' so to speak, and the term f is only ever applied to a canonical inhabitant of $\mathbb{N}$. This justifies describing the left-recursive form as being a *well-founded recursion*, since the evaluation is 'built-up' in the progression $V_0, V_1, \ldots$. Of course, in the presence of the zeta rule on'$\mathsf{succ}(\mu\alpha)$' we do not have to contend with deviant natural numbers at all.

It now only remains to show that the two flavours of $\mathrm{PRA}^{\omega}_{\mu}$ are strongly normalising. Unfortunately, due to a disanalogy between the zeta rule for $\Rightarrow$ and the zeta rule $= -\mathrm{z}$ that we shall introduce in the next chapter, we cannot define the equality type by encoding into a Π-type as we did for the intuitionistic type theory, and so have to introduce a 'dummy' type former Eq whose only use is to prove strong normalisation for the theories of the next chapter. We briefly describe the type former Eq as follows:

$$\frac{\Gamma \vdash s : A}{\Gamma \vdash \mathsf{refl}(s) : \mathsf{Eq}\, A}\ \mathsf{Eq}\,\mathcal{I} \qquad \frac{\Gamma \vdash s : \mathsf{Eq}\, A \quad \Gamma \vdash t : A \supset B}{\Gamma \vdash \mathsf{eqev}(s,t) : B}\ \mathsf{Eq}\,\mathcal{E}$$

There is a beta and a zeta reduction defined for this type, and no eta reduction. They are defined as follows

$$\begin{aligned}
&\mathsf{eqev}(\mathsf{refl}(s), t) \to^{c}_{\beta} \mathsf{ev}(t, s)\\
&\mathsf{eqev}(\mu\alpha^{A}.e, t) \to^{c}_{\zeta} \mu\beta^{B}.e[\alpha := \mathsf{eqev}(-, t)/\beta]\\
&\qquad\qquad\qquad \text{where } t \text{ admits the type } A \supset B
\end{aligned}$$

We note that they satisfy the diamond property.

As for the theory PRA^{ω}, we shall prove strong normalisation by defining an appropriate reducibility predicate and then showing inductively that all terms satisfy it. We employ three new techniques in this proof:

1. *Strengthening of reducibility predicates* There are two parts to the new predicate, part reducibility, one part which is defined inductively on the type as before, and a new part which appeals to a notion of reducing substitution, which we can reason about by induction. The extension is needed to show that reducibility is closed under the contraversions at higher types. We also appeal to two still stronger predicates, ramified reducibility, and full reducibility which are needed to show that terms formed using the so-called wild term formers succ, Rec and eqev are part reducible;

2. *Shoehorn lemma* We show that the fully reducible terms at any context are the image of the fully reducible terms of another context under a reducible substitution. This is needed to show that part reducibility is closed under lambda abstraction;

3. *Wild substitutions and full reducibility* This is required to show that the succ, Rec and eqev term formers are closed. We cannot simply strengthen our reducibility predicate further, as we cannot directly show that the full reducibility predicate satisfies CR 3.

We shall appeal to a distinction between two kinds of substitution that was not introduced in the earlier version of this thesis; between *disjoint* and *sequential* substitutions. In sequential substitutions, the kind we have usually appealed to, we apply the first substitution, then apply the second, and so on, whilst with disjoint substitutions we apply the substitutions in parallel and do not iterate them.

We shall also appeal to relative strong normalisation arguments to avoid considering eta and mu-beta reduction in proving CR 1 to CR 3. Although these are not new techniques, they are indispensable in the current proof: mu-beta reductions are incompatible with our proof of CR 3, whilst eta reductions are incompatible with our proof of the shoehorn lemma.

We also introduce a symbol $\rightarrow^+$ which expresses the relation $\rightarrow^* \cap \not\equiv$. We will restrict our attention to the left-recursive form in this section, we note that the right-recursive form can be shown analogously.

DEFINITION 132 We say that $s \rightarrow_l^1 t$ if $s \rightarrow^1 t$ but not $s \rightarrow_\mu^1 t$. $\rightarrow_l^*$ is defined to be the reflexive transitive closure of $\rightarrow_l^1$.

PROPOSITION 133

1. $\rightarrow_\mu^1$ is strongly normalising;
2. If $s \rightarrow_\mu^1 t$ and $t \rightarrow_l^1 u$ then there is a term t' such that $s \rightarrow_l^+ t'$ and $t' \rightarrow_\mu^* u$.

PROOF Part 1 follows immediately from the observation that the number of contraversions in a term is strictly reduced by any mu-eta or mu-beta reduction.

In the case of part 2, note that if the second reduction is a residual of the first, then the two redexes are not overlapping, and so the reductions may be performed in any order (though of course substitutions may change the number of such reductions, possibly to zero). Since no substitutions of terms arise from mu reductions, the number of logical reductions that arise when we reverse the order is exactly one.

We show the proposition in the case of that the logical reduction is not a residual of the mu reduction by case analysis. To each pair of reductions, we describe the sequences of reductions that 'solve' the reduction diamond:

1. (a) $\mu\eta$ then n. r. (that is, non residual) $\times\eta$: solved by $\times\zeta$ then $\mu\eta$;
   (b) $\mu\eta$ then n. r. $\Rightarrow\eta$: solved by $\Rightarrow\zeta$ then $\mu\eta$;
   (c) $\mu\beta$ then n. r. eta: no possibilities;
2. (a) $\mu\eta$ then n. r. $\times\beta$: solved by $\times\zeta$, two $\times\beta$ then $\mu\eta$;
   (b) $\mu\eta$ then n. r. $\Rightarrow\beta$: solved by $\Rightarrow\zeta$, two $\Rightarrow\beta$ then $\mu\eta$;
   (c) $\mu\eta$ then n. r. $\mathbb{N}\beta$: solved by $\mathbb{N}\zeta$, $\mathbb{N}\beta$ then $\mu\eta$;

(d) $\mu\eta$ then n. r. $\mathsf{Eq}\ \beta$: solved by $\mathsf{Eq}\ \zeta$, $\mathsf{Eq}\ \beta$ then $\mu\eta$;
(e) $\mu\beta$ then n. r. beta: no possibilities;

3. (a) $\mu\eta$ then n. r. zeta: no possibilities;
   (b) $\mu\eta$ then n. r. $\mathbb{N}\zeta$: solved by $\mathbb{N}\zeta$ twice then $\mu\eta$;
   (c) $\mu\eta$ then n. r. $\mathsf{Eq}\ \zeta$: solved by $\mathsf{Eq}\ \zeta$ twice then $\mu\eta$;
   (d) $\mu\beta$ then n. r. $\times\zeta$ or $\Rightarrow\zeta$: no possibilities;

□

COROLLARY 134 (RELATIVE NORMALISATION I)
If $\rightarrow_l^1$ is strongly normalising, then so is $\rightarrow^1$.

PROOF The chain bound lemma (from section 1.5) allows us to associate with any finite sequence of reductions of PRA_μ^ω, a sequence with as many logical reductions, and with the same start and end point in which all logical reductions occur beneath any mu-eta or mu-beta reductions. □

DEFINITION 135 We say that $s \rightarrow_L^1 t$ if $s \rightarrow_\beta^1 t$ or $s \rightarrow_\zeta^1 t$. $\rightarrow_L^*$ is defined to be the reflexive transitive closure of $\rightarrow_L^1$.

PROPOSITION 136
1. $\rightarrow_\eta^1$ is strongly normalising.
2. If $s \rightarrow_L^1 t$ and $t \rightarrow_\eta^1 u$, then there are terms t', u' such that either $s \equiv t'$ or $s \rightarrow_\eta^1 t'$ and $t' \rightarrow_L^+ u'$, where $u \rightarrow_\eta^* u'$;
3. If $s \rightarrow_L^1 t$ and s possesses no eta minima, then neither does t;
4. If $s \rightarrow^1 t \rightarrow_\eta^* u$ and u possesses no eta minima, then there is a term t' such that $s \rightarrow_\eta^* t' \rightarrow_L^+ u$.

PROOF The proof of part one is just as before, by noting that a measure on the minima of each derivation is reducing. Part two is similar to that in section 1.5, though there is a tricky case: a $\times$ or $\Rightarrow$ zeta reduction may create a non-residual eta redex. In each case this may be rearranged as two eta reductions followed by a zeta and a beta reduction. Part three is by inspection, and part four is a corollary of parts one to three. □

COROLLARY 137 (RELATIVE NORMALISATION II)
If $\rightarrow_L^1$ is strongly normalising, then so is $\rightarrow_l^1$.

PROOF This part follows by the same reasoning on reduction sequences of $\rightarrow_l^1$ as we carried out in section 1.5. □

PROPOSITION 138

1. $\rightarrow^1_\mu$ is strongly normalising;
2. If $s \rightarrow^1_\mu t$ and $t \rightarrow^1_l u$, then there is a term t' such that either $s \equiv t'$ or $s \rightarrow^1_l t'$ and $t' \rightarrow^*_\mu u$.

PROOF The first part is an immediate consequence that the number of contraversions decreases by one with any mu abstraction. The second part is shown by a case analysis that is rather easier than that for the first relative normalisation argument. □

It remains to show that the system with only beta and zeta reductions is strongly normalising. In what follows we shall assume that there is a type associated with each open variable, which allows us the convenience of applying reducibility predicates to terms and not judgements.

DEFINITION 139

1. We say that a substitution $\sigma :: \Gamma \rightarrow \Gamma'$ is a *reducing substitution* if
   (a) $\sigma \equiv \alpha := \mathsf{outl}(-)/\beta$, and there is a telescope Δ such that $\Gamma \equiv \Delta \cup \{\alpha : \bullet A \times B\}$, and $\Gamma' \equiv \Delta \cup \{\beta : \bullet A\}$;
   (b) $\sigma \equiv \alpha := \mathsf{outr}(-)/\beta$, and there is a telescope Δ such that $\Gamma \equiv \Delta \cup \{\alpha : \bullet A \times B\}$, and $\Gamma' \equiv \Delta \cup \{\beta : \bullet B\}$;
   (c) $\sigma \equiv \alpha := \mathsf{ev}(-, x)/\beta$, and there is a telescope Δ such that $\Gamma \equiv \Delta \cup \{\alpha : \bullet A \Rightarrow B\}$, and $\Gamma' \equiv \Delta \cup \{\beta : \bullet B, x : A\}$;
   (d) $\sigma \equiv \alpha := \mathsf{eqev}(-, \lambda x^A.x)/\beta$, and there is a telescope Δ such that $\Gamma \equiv \Delta \cup \{\alpha : \bullet \mathsf{Eq}\ A\}$, and $\Gamma' \equiv \Delta \cup \{\beta : \bullet A\}$
2. The *open subterms* of s are the terms t such that either $[\mathcal{D}_\perp]t \in \mathsf{st}(s)$, or $[\alpha]t \in \mathsf{st}(s)$ and $\alpha \in \mathsf{FV}(s)$;
3. A term is *canonical* if it has one of the following forms: $\langle s, s' \rangle$, $\lambda x^A.s$, $*$, $\mathsf{refl}(s)$ or $\underline{k}$. It is *non-canonical* if it is not canonical. Note that contraversions are non-canonical[10].

DEFINITION 140

1. Suppose $\Gamma \vdash s : A$. We say that s is *partly reducible*, or *PR* if we have the two conditions
   (a) Firstly we have a condition that depends upon the open rejective variables: for every reducing substitution $\sigma :: \alpha \rightarrow \Gamma'$ we have that $s[\sigma]$ is partly reducible;
   (b) Secondly, we have a condition on the type A:
      i. If A is atomic, s is strongly normalising;

[10] If we had beta-form zeta conversions, then the neutral terms could not encompass contraversions, which would complicate the proof somewhat. Fortunately we do not need the tiresome condition that open subterms are non-canonical in the specification of CR 3.

ii. If $A \equiv B \times B'$ then $\mathsf{outl}(s)$ and $\mathsf{outr}(s)$ are partly reducible;
iii. If $A \equiv B \Rightarrow B'$ then s is partly reducible if for all ramified reducible t (described below) we have that $\mathsf{ev}(s,t)$ is partly reducible;
iv. If $A \equiv \mathsf{Eq}\ B$, $\mathsf{eqev}(s, \lambda x^A.x)$ is partly reducible.

2. A term is *ramified reducible*, or *RR* if all of its subterms are partly reducible.

3. The *grounding substitutions* are a finite set of sequential substitutions defined for the typed rejective variables of any term:

$$\begin{aligned}
\mathsf{gss}(\alpha : \bullet A) &\hat{=} \{\cdot\},\ \text{if } A \equiv \mathbb{N} \text{ or } A \equiv \mathsf{Eq}\ B \\
\mathsf{gss}(\alpha : \bullet A \times B) &\hat{=} \{(\alpha := \mathsf{outl}(-)/\beta, \sigma) \mid \sigma \in \mathsf{gss}(\beta : \bullet A)\} \\
&\quad \cup \{(\alpha := \mathsf{outr}(-)/\gamma, \sigma) \mid \sigma \in \mathsf{gss}(\gamma : \bullet B)\} \\
\mathsf{gss}(\alpha : \bullet A \Rightarrow B) &\hat{=} \{(\alpha := \mathsf{ev}(-, s)/\beta, \sigma) \mid \sigma \in \mathsf{gss}(\beta : \bullet A) \wedge s \text{ RR}\}
\end{aligned}$$

The *simple grounding substitutions* for the typed rejective variables of a term are defined in the same way, except that in the last clause we permit only the substitutions $\alpha := \mathsf{ev}(-, x)/\beta, \sigma)$;

4. The *A-class substitutions* of a term are defined to be the set of substitutions of ramified reducible terms for the open affirmative variables of that term;

5. The *B-class substitutions* of a term are defined to be the product of the A-class substitutions of that set with the grounding substitutions of its open rejective variables.

REMARK 141
The two conditions required to show part reducibility are defined in such a way that they appeal to the part reducibility of other terms. We might be concerned that this recursive appeal is well-founded, since the part of the second condition a function space type can introduce new rejective variables. However if we formulate the *reducibility measure* to be $a\omega + b$ where a is the degree of the type of the conclusion, and b is the sum of the degrees of the open rejective open variables, then we can see that the terms we appeal to have strictly lower measure.

PROPOSITION 142 Let $\Gamma \vdash s : A$.

(CR 1) If s is partly reducible then it is strongly normalising;

(CR 2) If s is partly reducible and $s \to_L^* s'$, then s' is partly reducible;

(CR 3) If s is non-canonical, and if for every term s' with $s \to_L^1 s'$ we have that s' is partly reducible, then s is partly reducible.

PROOF CR 1 is established by the same arguments used to establish CR 1 in section 1.5. For the other two properties we shall require a lemma

LEMMA 143 Let σ be a reducing substitution, and let $s \to_L^1 s'$. Then $s[\sigma] \to_L^1 s'[\sigma]$.

To see this, simply note that no maxima or zeta minima overlap with reducing substitutions[11].

CR 2 is established by induction on the reducibility measure over terms. We establish the two properties in turn.

1. Consider any reducing substitution $\sigma :: \Gamma \to \Gamma'$. We need to show that for any redex $s \to_L^1 s'$ that $\Gamma' \vdash s'[\sigma]$ is partly reducible. Since s is, then by the definition of full reducibility $s[\sigma]$ is, and so by the induction hypothesis we have our conclusion.

2. Consider the type A:

   (a) A is atomic. Then every term obtained by reducing s is strongly normalising, since s is, and so the second condition is satisfied. Thus we have that $\Gamma \vdash s : A$ is partly reducible;

   (b) $A \equiv B \times B'$. Our argument follows the same form as that for section 1.5;

   (c) $A \equiv B \Rightarrow B'$. Again, our argument follows the same form as that for section 1.5;

   (d) $A \equiv \mathsf{Eq}\, B$. If $s \to_L^* s'$ then $\mathsf{eqev}(s, \lambda x^B.x) \to^* \mathsf{eqev}(s', \lambda x^B.x)$. Since $\Gamma \vdash s : \mathsf{Eq}\, B$ is partly reducible, we have that $\Gamma \vdash \mathsf{eqev}(s, \lambda x^B.x) : B$ is and so $\Gamma \vdash \mathsf{eqev}(s, \lambda x^B.x) : B$ satisfies the second condition by the induction hypothesis. Since we have already shown that it satisfies the first condition, we have that $\Gamma \vdash s' : A$ is partly reducible.

CR 3 is established by a similar induction, establishing the two conditions are satisfied in turn.

1. Let $\sigma : \Gamma \to \Gamma'$ be a reducing substitution. Consider any one-step reduction $s \to_L^1 s'$. By our lemma above we have that $s[\sigma] \to_L^1 s'[\sigma]$: since each $s'[\sigma]$ is partly reducible we may appeal to the induction hypothesis to establish that $s[\sigma]$ is partly reducible by CR 3;

2. The second condition is established in much the same way as CR 3 is established in section 1.5. As for CR 2, the case of $A \equiv \mathsf{Eq}\, B$ causes no particular difficulties.

□

The following lemma is also very useful for proving that terms are fully reducible.

LEMMA 144 Let $\alpha : \bullet A$ be an open rejective variable of some term s. If for all ρ in the simple grounding substitution set for the α we have that $s[\rho]$ is partly reducible, then so is s.

[11] Both eta reductions and the mu-beta reductions are affected, however.

PROOF The key to this proof lies in the correspondence between the redexes of s and $s[\sigma]$, where σ is a reducing substitution of s, which may be summarised:

- There is an natural injection from the redexes of s to those of $s[\sigma]$ (the reader may wish to work out how it can be defined recursively);
- If $s \to_L^1 s'$ then $s[\sigma] \to^1 s'[\sigma]$ where the redex contracted in the latter part is given by the above injection.

We can show that this correspondence may be extended to simple grounding substitutions by induction, and then the result follows by showing that to each of the terms t to whose FR-ness the definition of FR depends upon, there is a similar correspondence from redexes of t to redexes of $t[\rho]$.

We need to show each condition in turn. For the first condition, we have that for reducing substitutions $\sigma :: \alpha : \bullet A \to \beta : \bullet B$ we can immediately appeal to our induction hypothesis, since the composition of the grounding substitutions in $s[\sigma]$ of β with σ is a grounding substitution in s of α.

For other reducing substitutions, say σ, we note that $s[\sigma, \rho] \equiv s[\rho, \sigma]$, and so we can note the injection of redexes from s to $s[\rho]$ implies an injection of redexes from $s[\sigma]$ to $s[\sigma, \rho]$, and so the former is true if the latter is.

The second condition follows from applying a similar manipulation to $\mathsf{outl}(s)$, $\mathsf{outr}(s)$ and $\mathsf{ev}(s, t)$. □

We need the following definition and proposition to prove the shoehorn lemma.

DEFINITION 145 Let $\alpha : \bullet A$ be an open rejective variable of some term s. If s' is obtained by replacing some (zero, one or more) α-open subterms of s by the variable $x : A$ then s' is an *α/x-erasing* of s.

LEMMA 146 The α/x-erasings of any partly reducible term s are partly reducible.

PROOF We note that a similar correspondence exists between the redexes of a given term and the terms obtained by applying substitutions in the simple grounding substitution set for the α to the α/x-erasings. An easy induction on the reducibility measure of terms obtains our result. □

PROPOSITION 147 (SHOEHORN LEMMA)
Let s be partly reducible, and let $\sigma :: \Gamma' \to \Gamma$ be some reducing substitution whose range is a rejective variable in s. Then there is a partly reducible term t such that $t[\sigma] \to^* s$.

PROOF We devise an inverse substitution τ to σ as follows:

1. If $\sigma \equiv \alpha := \mathsf{outl}(-)/\beta$ then $\tau \mathrel{\hat{=}} \beta := \langle -, x\rangle/\alpha$;
2. If $\sigma \equiv \alpha := \mathsf{outr}(-)/\beta$ then $\tau \mathrel{\hat{=}} \beta := \langle x, -\rangle/\alpha$;
3. If $\sigma \equiv \alpha := \mathsf{ev}(-, x)/\beta$ then $\tau \mathrel{\hat{=}} \beta := \lambda y. - /\alpha$;
4. If $\sigma \equiv \alpha := \mathsf{eqev}(-, \lambda x.x)/\beta$ then $\tau \mathrel{\hat{=}} \beta := \mathsf{refl}(-)/\alpha$.

t is then defined to be $s[\tau]$. We have immediately that $t[\sigma] \to_L^* s$. We also have a one-to-one correspondence between redexes of s and those of t, and if $s \to_L^1 s'$ then for some t' we have that $t' \to_L^1 s'[\tau]$.

Consider the simple grounding substitutions ρ on α in t. Simple formula manipulation yields that for each such ρ, $t[\rho]$ reduces either to the result of applying a simple grounding substitution to s or to the result of applying such to an α/x-erasing of s. □

PROPOSITION 148 (REDUCIBILITY I)

1. If $\Gamma \vdash s : A$ and $\Gamma \vdash s' : A'$ are partly reducible, then so is $\Gamma \vdash \langle s, s' \rangle : A \times A'$;
2. If $\Gamma, x : A \vdash s : B$ and for all judgements $\Gamma \vdash t : A$ which are ramified reducible we have that $\Gamma \vdash s[x := t] : A$ is partly reducible, then $\Gamma \vdash \lambda x^A.s : A \Rightarrow B$ is partly reducible;
3. If $\Gamma \vdash s : A$ is partly reducible, then $\Gamma \vdash \mathsf{refl}(s) : \mathsf{Eq}\ A$ is also.

PROOF Each part is shown by induction on $d(\Gamma)$ and by induction on the reduction chain bound of the subterms.

Part one. There are two conditions the judgement must satisfy to by partly reducible. The second condition is shown just as in the proof of Girard–Lafont–Taylor (and requiring the inner induction hypothesis). For the first condition, let $\sigma :: \Gamma \to \Gamma'$ be any reducing substitution. By assumption $s[\sigma]$ and $s'[\sigma]$ must by partly reducible, and so their pairing must be too by the induction hypothesis.

Part two. The second condition follows that of Girard–Lafont–Taylor as before. To show the first condition we need to show that for every ramified reducible $\Gamma' \vdash t : A$ that $\Gamma' \vdash s[\sigma][x := t] : A$ is partly reducible. By the shoehorn lemma we may associate to each such t a partly reducible judgement $\Gamma \vdash u : A$ such that $u[\sigma] \to_L^* t$, and so $s[\sigma][x := u[\sigma]] \to^* s[\sigma][x := t]$. We may apply the induction hypothesis to obtain that $\lambda x.s[\sigma]$ is partly reducible.

Part three. Left as an exercise to the reader (it is along the same lines as part one). □

PROPOSITION 149 (REDUCIBILITY II)
If $\Gamma, \alpha : \bullet A \vdash s : B$, $\Gamma, \alpha : \bullet A \vdash r : \bullet B$ and $s[\sigma]$ is partly reducible for each $\sigma \in \mathsf{gss}(\alpha : \bullet A)$, then $\mu\alpha^A.[r]s$ is PR.

PROOF Suppose $\Gamma \vdash r : \bullet C$. We then prove the proposition by establishing it by a triple induction. The outermost variant is the sum of the height of all the types of open subterms of s, the second variant is the height of the type A, and the third variant is the number $\mathsf{rcb}(s)$. We shall proceed by case analysis on the type A; the base cases for the first and third variants are simply that certain kinds of redex cannot occur in the following analysis:

1. A is atomic: All of the relevant redexes of $\mu\alpha^A.[r]s$ are redexes of s, and so by an appeal to CR 3 (the term is non-canonical) and the inner induction hypothesis we have that $\mu\alpha^A.[r]s$ is fully reducible;

2. $A \equiv B \times B'$: we have that s is fully reducible if $\mathsf{outl}(s)$ and $\mathsf{outr}(s)$ are (since there is an obvious surjection from the redexes that obtain from applying simple grounding substitutions to s and to $\mathsf{outl}(s)$ and $\mathsf{outr}(s)$). The argument that these two terms are fully reducible is similar in each case, so without loss of generality I shall consider only $\mathsf{outl}(s)$. We aim to show that the result of contracting any possible redex obtains a fully reducible term:

   - Base $\times\beta$ redex: not possible as $\mu\alpha^A.[r]s$ is non-canonical;
   - We contract a redex of s to obtain $\mu\alpha[r]s'$: by appeal to CR 3, in a similar manner to atomic case;
   - Base $\times\zeta$ redex: it follows from the previous proposition and the first induction hypothesis that if $\Gamma \vdash s[\alpha := \mathsf{outl}(-)/\beta]$ and $s[\alpha := \mathsf{outr}(-)/\gamma]$ are the object of fully reducible judgements, then

     $$\left\langle \mu\beta^B.[r]s[\alpha := \mathsf{outl}(-)/\beta], \mu\gamma^{B'}.[r]s[\alpha := \mathsf{outr}(-)/\gamma] \right\rangle$$

     is. These cases follow from the first condition on fully reducible judgements.

3. $A \equiv B \Rightarrow B'$: because the open subterms are the same between the two terms, we have that s is fully reducible if $\mathsf{outl}(s)$ and $\mathsf{outr}(s)$ are. The argument that these two terms are fully reducible is similar in each case, so without loss of generality I shall consider only $\mathsf{outl}(s)$. We aim to show that the result of contracting any possible redex obtains a fully reducible term:

   - Base $\Rightarrow\beta$ redex: not possible as $\mu\alpha^A.[r]s$ is non-canonical;
   - We contract a redex of s to obtain $\mu\alpha[r]s'$: by appeal to CR 3, in a similar manner to atomic case;
   - Base $\Rightarrow\zeta$ redex: we can show that $\mu\beta^{B'}.([r]s)[\alpha := \mathsf{ev}(-, x)]$ is fully reducible by reasoning similar to the $\times\zeta$ case. Additionally we require that $\lambda x^B.\mu\beta^{B'}.([r]s)[\alpha := \mathsf{ev}(-, x)]$ if partly reducible, for which we require that the the result of applying all substitutions $x := u$, where u is RR, is fully reducible. But by assumption we have that s is closed under all grounding substitutions, and so this case holds as well.

□

The results we have so far derived are adequate to show that terms that do not involve the succ, Rec and eqev term formers, what we call the *wild term formers*, are strongly normalising, by showing by structural induction on terms that the result of applying any B-class substitution to the given term is part reducible. To extend this result to these term formers requires a stronger form of reducibility predicate, which we call full reducibility. To define it, we also introduce a notion of reducibility related to ramified reducibility, called co-RR. It is our aim to show that all ramified reducible terms are fully reducible, and that the class of fully reducible

terms is closed under $\to^1_L$, from which we can extend the above induction to show that all terms are ramified reducible.

DEFINITION 150

1. A term s is *co-RR* if one of the following conditions hold:

   (a) s is RR;
   (b) For some co-RR term t, $t \to^1_\beta s$;
   (c) For some co-RR term t, $s \in \mathsf{st}(t)$.

   A term is RR* if it is obtained from just the first two parts of the above scheme. The *C-class substitutions* on a term are the disjoint substitutions of co-RR terms for open affirmative variables;

2. The *unary wild substitutions* are as follows

   (a) $\alpha := \mathsf{succ}(-)/\beta$;
   (b) $\alpha := \mathsf{Rec}(-, a, f)/\beta$ where a and f are RR;
   (c) $\alpha := \mathsf{eqev}(-, t)/\beta$ where t is RR;

3. The *wild substitutions* are obtained from a number of disjoint unary wild substitutions (possibly zero), that is we do not substitute into the terms obtained from the previous substitutions.

4. A term s is *fully reducible*, or *FR*, if

   (a) For all wild substitutions σ, $s[\sigma]$ is PR;
   (b) For every $\mu\alpha^A.[r]t \in \mathsf{st}(s)$, and for every wild substitution σ and C-class substitution ρ on t, $t[\sigma, \rho]$ is FR.

   The *D-class substitutions* are the same as the B-class substitutions except that we restrict ourselves to terms that are both FR and RR.

PROPOSITION 151 (REDUCIBILITY III)

1. If s is FR and co-RR, and $s \to^* t$, then t is FR.
2. (a) If s is FR, and σ is a wild substitution, then $\mathsf{succ}(s[\sigma])$ is PR;
   (b) If $\mathsf{succ}(s)$ is FR then s is FR;
3. If s is FR, a and f are RR* and FR, and σ is a wild substitution, then $\mathsf{Rec}(s[\sigma], a, f)$ is PR;
4. If s is FR, f is RR* and FR, and σ is a wild substitution, then $\mathsf{eqev}(s[\sigma], f)$ is PR.

PROOF By a simultaneous induction on the sum of the chain bounds of terms appearing in the condition, in each case.

- Part one. We only need to show this for $\to_L^1$, obtaining the full strength of the proposition by an appeal to the induction hypothesis. We note that there is a surjection from the redexes of $s[\sigma]$ for wild substitution σ to those of s, and so by CR 2, the first condition holds for the residual.

  Condition two holds immediately for beta reductions, since co-RR is closed under beta reduction (this is the only reason we need the restriction to RR terms), and all beta substitutions occurring in a co-RR term will be of co-RR terms. Thus it remains only to show that the last FR-ness condition is conserved under the three wild zeta reductions. The second condition is obtained from the other three parts by noting that the result of a wild substitution is co-RR since all of its parts are. The last condition depends upon showing that the result of applying a wild substitution instance then a wild substitution, then a C-class substitution is FR: by rearrangement of substitutions and our induction applied to the subterm (which must have strictly smaller rcb) we obtain this.

- Parts two, three and four are shown by similar means. We shall, for the sake of brevity, consider only part three (concerning Rec) in detail.

  Subpart (a) is shown by an induction on $\mathsf{snf}(s) \cdot \omega + \mathsf{rcb}(s) + \mathsf{rcb}(a) + \mathsf{rcb}(f)$, (where $\mathsf{snf}(\cdot)$ is the size of the normal form of the argument) and an appeal to CR 3. A reduction $\mathsf{Rec}(s[\sigma], a, f) \to_L^1 t$ matches one of the following cases:

  1. A redex of $s[\sigma]$, a or f is converted. Then we appeal to the induction hypothesis, where in the case of $s[\sigma]$ we must show that the two reductions can be commuted. This follows by a similar argument to that in part 1.
  2. $s \equiv \underline{0}$ and $t \equiv a$. a is PR by assumption;
  3. $s \equiv \mathsf{succ}(n)$ and $t \equiv \mathsf{Rec}(n, \mathsf{ev}(f, \underline{0}, a), \lambda x^{\mathbb{N}}.\mathsf{ev}(f, \mathsf{succ}(x)))$. Each of the last two arguments are easily shown to be RR (obviously the numerals $\underline{k}$ are RR). It is easy to show that n is FR; from this we may by appeal to the induction hypothesis (part 2(b)) to show that the term is PR (since $\mathsf{snf}(n) < \mathsf{snf}(\mathsf{succ}(n))$);
  4. $s \equiv \mu\alpha^{\mathbb{N}}.e$ and $t \equiv \mu\beta^{A}.e[\alpha := \mathsf{Rec}(-, a, f)/\beta, \sigma]$. There are two subcases to consider
     (a) $e \equiv [\alpha]s'$. By the third condition on FR of s, s' is FR, and so by the outer induction, 2(b) applies, and so $\mathsf{Rec}(s'[\alpha := \mathsf{Rec}(-, a, f)], a, f)$ is PR, and so the whole term is by Reducibility II;
     (b) Otherwise we may appeal directly to Reducibility II.

  In subpart (b), there are two conditions needed to show full reducibility is satisfied. The first is exactly subpart (a) with t, whilst the second is vacuous.

□

PROPOSITION 152 (REDUCIBILITY IV)

1. Let s be a RR term, and let ρ be a D-class substitution. Then $s[\rho]$ is both RR and FR;
2. Let s be some term and let ρ be a B-class substitution. Then $s[\rho]$ is RR.

PROOF Before we begin, we need to show that the FR terms are closed under the following substitutions: $\alpha := \mathsf{succ}(-)/\beta$, $\alpha := \mathsf{Rec}(-, x, y)/\beta$ and $\alpha := \mathsf{eqev}(-, x)/\beta$, which may be established by an easy induction on the number of contraversions in terms, appealing to Reducibility III.

Each part is shown by structural induction over all terms. In each part it suffices to show that s is PR, since the premisses to the induction the yield that s is RR.

1. (a) $s \equiv x$: By assumption;
   (b) $s \equiv \mathsf{outl}(t)$ or $s \equiv \mathsf{outr}(t)$: both properties follow immediately from the respective definitions;
   (c) $s \equiv \langle t_0, t_1 \rangle$: PR follows from Reducibility I, and taming follows from applying the above lemma to $\mathsf{outl}\langle t_0, t_1 \rangle$ and $\mathsf{outr}\langle t_0, t_1 \rangle$;
   (d) $s \equiv \mathsf{ev}(t_0, t_1)$: The second property on FR terms is immediate. To show the first property, we need that $t_1[\rho, \sigma]$ is RR for all wild substitutions σ, a property that follows easily by the property we showed at the beginning, and rearrangement of substitutions;
   (e) $s \equiv \lambda x.s$: PR follows by a similar argument to that of Reducibility I. To show taming we need that for any wild σ and RR u that $t[\rho, \sigma, x := u]$ is PR. But since $x \notin \mathsf{ran}(\sigma)$, $t[\rho, \sigma, x := u] \equiv t[\rho, x := u, \sigma]$, and so this follows by the induction hypothesis;
   (f) $s \equiv \mu\alpha.[r]s'$: PR is an immediate consequence of Reducibility II. For the term to be FR our second condition follows by the induction hypothesis, whilst to show the first condition we need a case analysis on possible substitutions and an appeal to Reducibility III (in the case that $\alpha \not\equiv r$). Details are left as an exercise;
   (g) $s \equiv \underline{0}$: immediate;
   (h) $s \equiv \mathsf{succ}(t)$, $s \equiv \mathsf{Rec}(t, a, f)$ or $s \equiv \mathsf{eqev}(t, f)$: Both parts follow easily from Reducibility III;
2. In the light of part one, B-class and D-class substitutions coincide (in the case where ρ is the identity substitution). Thus we can establish this proposition by showing by structural induction that each case is PR, which is quite elementary.

□

THEOREM 153 The calculus $\mathrm{PRA}^{\omega}_{\mu}$ is strongly normalising and Church–Rosser.

PROOF Given that the above argument can be shown in a similar way for $\mathrm{PRA}^{\omega}_{\mu}$ (only the proof of part 2 of Reducibility III differs), and given the two relative normalisation results, we obtain the theorem for $\mathrm{PRA}^{\omega}_{\mu}$ in a similar manner to that for $\mathrm{PRA}^{\omega}_{\mu}$.

Since the identity substitution is a B-class lambda substitution, we have that all terms are RR, and so are PR. By CR 1 we have that all terms are strongly normalising, and so by the diamond property it is also Church–Rosser. □

REMARK 154 The move from the proof for PRA^{ω} to that for $\mathrm{PRA}^{\omega}_{\mu}$ suggests a general way of formulating generalisations of logical relations for the lambda calculus to those for lambda-mu, by considering not just the type of the term, but the finite set consisting of the type of the term and the types of all its open subterms.

3.5 An alternative formulation

Recall from section 1.4 the definition of the semantic content of proofs in terms of types:

$$\mathsf{Prf}(A \wedge B) \cong \mathsf{Prf}(A) \times \mathsf{Prf}(B)$$
$$\mathsf{Prf}(A \supset B) \cong \mathsf{Prf}(A) \Rightarrow \mathsf{Prf}(B)$$

It is possible to formulate a dual definition which captures the semantic content of refutations:

$$\mathsf{Ref}(A \wedge B) \cong \mathsf{Ref}(A) + \mathsf{Ref}(B)$$
$$\mathsf{Ref}(A \supset B) \cong \mathsf{Prf}(A) \times \mathsf{Ref}(B)$$

where the type '$A + B$' contains the disjoint union of the types of A and B.

This allows us to formulate 'introduction' rules for $\bullet A \wedge B$ and $\bullet A \supset B$:

$$\frac{\bullet A}{\bullet A \wedge B}\bullet\wedge 1 \qquad \frac{\bullet B}{\bullet A \wedge B}\bullet\wedge 2$$

$$\frac{A \quad \bullet B}{\bullet A \supset B}\bullet\supset$$

In the presence of the structural rules of lambda-mu, these rules have the same logical strength as the elimination rules. Furthermore these rules admit an account of reduction that is closely related to that of lambda-mu.

We shall devote the remainder of this chapter to a brief outline of this system, which is intended only to look at our treatment of logical duality for lambda-mu from a slightly different perspective. We shall call this calculus $\lambda\mu^*$, and we shall not have occasion to use this calculus again in this work.

$\lambda\mu^*$ has exactly the same structural rules as $\lambda\mu$ and its logical rules are divided into assertive and rejective rules. These are given:

$$\frac{\Gamma \vdash s_0 : A \quad \Gamma \vdash s_1 : b}{\Gamma \vdash \langle s_0, s_1 \rangle : A \times B}$$

$$\frac{\Gamma \vdash r : \bullet A}{\Gamma \vdash \mathsf{inl}(r) : \bullet A \times B} \qquad \frac{\Gamma \vdash r : \bullet A}{\Gamma \vdash \mathsf{inl}(r) : \bullet A \times B}$$

$$\frac{\Gamma x : A \vdash s : B}{\Gamma \vdash \lambda x^A.s : A \Rightarrow B} \qquad \frac{\Gamma \vdash s : A \quad \Gamma \vdash r : \bullet B}{\langle s, r\rangle : \bullet A \Rightarrow B}$$

The reduction rules associated with the structural rules are given:

$$\begin{aligned} [r]\mu\alpha.e &\to^c_\mu e[\alpha := r] \\ \mu\alpha.[\alpha]s &\to^c_\mu s, \text{ if } \alpha \notin \mathsf{FV}(s) \end{aligned}$$

where the substitution on the variables α is defined just as substitution on assertive variables.

$$\begin{aligned} [\mathsf{inl}(r)]\langle s_0, s_1\rangle &\to^c_\beta [r]s_0 \\ [\mathsf{inr}(r)]\langle s_0, s_1\rangle &\to^c_\beta [r]s_1 \\ [\langle s, r\rangle]\lambda x^A.t &\to^c_\beta [r]t[x := s] \\ s &\to^c_\eta \langle \mu\alpha^A.[\mathsf{inl}(\alpha)]s, \mu\beta^B.[\mathsf{inr}(\beta)]s\rangle \\ s &\to^c_\eta \lambda y^A.\mu\beta^B.[\langle y, \beta\rangle]s \end{aligned}$$

where the first eta rule applies if s admits a type governed by $\times$, and the second rule applies if s admits a type governed by $\Rightarrow$. There are no zeta rules.

The one-step reduction for $\lambda\mu^*$ is obtained by compatible closure for the mu and beta rules, and obtained for the eta rule by permitting the redex to occur anywhere in the term except as the affirmative premiss to an antithesis, or where the redex arises by an introduction rule.

It is possible to provide a translation from $\lambda\mu$ to $\lambda\mu^*$ (if we are a little glib about the types attached to bound identifiers):

$$\begin{aligned} [\![x]\!] &\hat{=} x \\ [\![\alpha]\!] &\hat{=} \alpha \\ [\![[r]s]\!] &\hat{=} [[\![r]\!]][\![s]\!] \\ [\![\mu\alpha^A.e]\!] &\hat{=} \mu\alpha^A.[\![e]\!] \\ [\![\langle s, t\rangle]\!] &\hat{=} \langle [\![s]\!], [\![t]\!]\rangle \\ [\![\mathsf{outl}(s)]\!] &\hat{=} \mu\alpha.[\mathsf{inl}(\alpha)]s \\ [\![\mathsf{outr}(s)]\!] &\hat{=} \mu\alpha.[\mathsf{inr}(\alpha)]s \\ [\![\lambda x^A.s]\!] &\hat{=} \lambda x^A.[\![s]\!] \\ [\![\mathsf{ev}(s, t)]\!] &\hat{=} \mu\alpha.[\langle t, \alpha\rangle]s \end{aligned}$$

The reader will easily see that the translation redex-contractum pairs for mu and eta redexes in lambda-mu are redex–contractum pairs of $\lambda\mu^*$. For beta redexes the translation of the contractum is obtained by applying the respective beta contraction followed by a mu contraction to the translation of the redex. Beta form zeta

rules translate to mu redexes, whilst eta form zeta rules translate to the composition of an eta contraction and a mu contraction.

It is not the case that translating a normal form of lambda-mu obtains a normal form in $\lambda\mu^*$, since long synthetic threads of the former map to a number of mu-beta redexes. However it is fair to say that the logical content of the two calculi is the same, up to a few mu rules, and this is vindicated by the existence of a reverse translation whose composition with the above is an identity up to mu-beta and mu-eta conversion.

What is gained from the translation from lambda-mu to $\lambda\mu^*$? The first advantage is that synthetic threads of $\lambda\mu^*$ increase in logical complexity as we move down them, so following our brief mention of this in section 3.1 we see that they share with the sequent calculus a more evident logical duality, and a more easily formulated measure of logical complexity.

The second advantage is that this formulation has a suggestive similarity with an illuminating continuation semantics of the lambda calculus presented in the paper 'Continuation Semantics: Abstract machines and control operators' by Streicher and Reus [StreicherT:consam][12]. Readers familiar with this work should see the relationship between the rejective rules and the representation of continuations.

Nonetheless the calculus lambda-mu was chosen over that of $\lambda\mu^*$ to introduce the classical propositional calculus, principally because $\lambda\mu^*$ does not appear to extend well to inductive types, and also because the theory of $\lambda\mu^*$ has an unattractive asymmetry, which can be seen in the absence of any derivation justifying '$\bullet A \wedge B, A \vdash \bullet B$'.

It is possible to formulate an extension to the calculus to avoid this asymmetry, obtaining the calculus $\lambda\mu^{**}$. We allow a new, dual form of contraversion, which we call *negative contraversion*:

$$\frac{\Gamma, x : A \vdash e \ \mathsf{empty}}{\overline{\mu}x : \bullet A.e : \bullet A}$$

Negative contraversion admits dual rules to the mu rules, which we call the *mu-bar rules*:

$$[\overline{\mu}x : \bullet A.e]s \rightarrow^c_{\overline{\mu}} e[x := s]$$
$$\overline{\mu}x : \bullet A.[r]x \rightarrow^c_{\overline{\mu}} r, \text{ if } x \notin \mathsf{FV}(r)$$

Unfortunately this calculus does not satisfy the Church–Rosser property, due to the existence of a bad critical pair:

$$[\overline{\mu}z : \bullet C.[\alpha]x]\mu\gamma : C.[\beta]y$$

which admits '$[\alpha]x$' and '$[\beta]y$' as its normal forms. Note that the critical pair requires no term formers governed by logical connectives.

[12]Readers not familiar with this work are strongly encouraged to seek it out

Chapter 4

Classical proofs II: arithmetic

A natural reaction to the picture of excellent health developed for the proof theory of the classical propositional calculus is that we will be able to extend it directly to the predicate calculus. Unfortunately, things are not quite so straightforward, but it will be instructive to first proceed as if they were. The calculus we develop this way shall be called the *naïve account* of *classical type theory* or NTT.

4.1 The naïve theory

Formulae, terms and telescopes

We extend the grammar of term candidates to distinguish between affirmative, rejective and empty term candidates, just as we did in chapter 3. We do not introduce any new type formers, but we introduce a new kind of degree 2 assumption:

$$\mathcal{A} ::= X \text{ type} \mid x : T \mid \alpha : \bullet T$$

and we introduce the type constants $\bot$ and $\top$. The rejective variables α also occur in the domain of a telescope. Telescopes are lists of assumptions as before.

The affirmative term candidates are obtained by extending the existing notion of term, and the rejective and empty term candidates are introduced as follows:

$$\begin{aligned} s &::= \ldots \mid * \mid \mu\alpha : T.e \\ r &::= \alpha \mid \mathcal{D}_\bot \\ e &::= [r]s \end{aligned}$$

Judgements

Where we had one kind of core judgement on terms, now we have three, respectively on affirmative, rejective and empty terms. Our judgements are:

$$\Gamma \vdash_{\text{NTT}} s : A$$
$$\Gamma \vdash_{\text{NTT}} r : \bullet A$$
$$\Gamma \vdash_{\text{NTT}} e \ \mathsf{empty}$$

where s, r and e range over affirmative, rejective and empty term candidates respectively. We also have three kinds of convertive equality on terms:

$$\Gamma \vdash_{\text{NTT}} s = s' : A$$
$$\Gamma \vdash_{\text{NTT}} r = r' : \bullet A$$
$$\Gamma \vdash_{\text{NTT}} e = e' \ \mathsf{empty}$$

Thus where there were six judgements of ITT, there are ten for NTT.

The valid judgements are those that are conclusions of a valid derivation with the rules of NTT.

Substitutions and definitions

Our new form of substitution, mixed substitution, from chapter 3 can be applied to our type theory to obtain the substitution operator $\mathsf{subst}(-, \sigma)$ where σ ranges over mixed substitutions. We obtain the class of general substitutions by extending the core substitutions to include a third judgement:

$$\frac{\Gamma, x : B \vdash f(x) : A \quad \Gamma \vdash r : \bullet B}{\alpha := f(-^B)/r :: \Gamma, \alpha : \bullet A \to \Gamma}$$

where $f(-)$ is a unary parameterised term.

We shall not have cause to use rejective assumptions in parameterised definitions, but there is nothing in principle to stop us from using them; they are defined exactly as they were in ITT.

In presenting the rules of NTT and then of CTT we shall make use of implicit parameterised definitions as we did for ITT, but we shall introduce a new twist. A premiss of the presentation of a rule may take the form '$\Gamma \vdash f(\Gamma') \equiv g : A$' which is introducing a parameterised term f, insisting that it takes the form g, in which variables of both Γ and Γ' may occur.

EXAMPLE 155
The $\mathbb{N} - \mathrm{z}$ rule is defined as follows:

$$\frac{\Gamma, \alpha : \bullet\mathbb{N} \vdash e \ \mathsf{empty} \quad \Gamma \vdash C(n^{\mathbb{N}}) \ \mathsf{type} \quad \Gamma \to r(n : \mathbb{N}) \mathrel{\hat{=}} \mathrm{R}^{\mathbb{N}}(n, a, (i, z)f) : C(n)}{\Gamma \vdash r(\mu\alpha^{\mathbb{N}}.e) = \mu\beta^{C(\mu\alpha:\mathbb{N}.e)}.e[\alpha := r(-)/\beta] : C(\mu\alpha^{\mathbb{N}}.e)} \ \mathbb{N} - \mathrm{z}$$

The third premiss may be equivalently rendered '$\Gamma, n : \mathbb{N} \vdash \mathrm{R}^{\mathbb{N}}(n, a, (i, z)f) : C(n)$'; it also indicates that we must recast two expressions occurring in the conclusion as follows: firstly the expression '$r(\mu\alpha^{\mathbb{N}}.e)$' is to be understood as '$\mathrm{R}^{\mathbb{N}}(\mu\alpha : \mathbb{N}.e, a, (i, z)f)$' and secondly that the incomplete substitution expression '$[\alpha := f(-)/\beta]$' should be understood by the concrete substitution '$[\alpha := \mathrm{R}^{\mathbb{N}}(-, a, (i, z)f)/\beta]$'.

Structural rules

We add a new class of structural rule, making a total of six classes:

1. Formation of telescopes;
2. Introduction of assumptions;
3. Convertive equality is an equivalence;
4. Convertive compatibility;
5. Alpha conversion and type congruence;
6. Classical structural rules

There are two new rules in class 1 and one each in classes 2 and 5,

$$\frac{\Gamma \vdash A \ \mathsf{type} \quad \alpha \notin \mathsf{dom}(\Gamma)}{\Gamma, \alpha : \bullet A \ \mathsf{tel}} \ \mathrm{tel-den}$$

$$\frac{\Gamma = \Gamma' \ \mathsf{tel} \quad \Gamma \vdash A = A' \ \mathsf{type} \quad \alpha \notin \mathsf{dom}(\Gamma)}{\Gamma, \alpha : \bullet A = \Gamma, \alpha : \bullet A \ \mathsf{tel}} \ \mathrm{tel-den-eq}$$

$$\frac{\Gamma, \alpha : \bullet A, \Gamma' \ \mathsf{tel}}{\Gamma, \alpha : \bullet A, \Gamma' \vdash \alpha : \bullet A} \ \mathrm{hyp-den}$$

$$\frac{\Gamma \vdash A = A' \ \mathsf{type} \quad \Gamma, \alpha : \bullet A \vdash e \ \mathsf{empty} \quad \beta \notin \mathsf{dom}(\Gamma)}{\Gamma \vdash \mu\alpha : A.e = \mu\beta : A'.e[\alpha := -^{A}/\beta]} \ \mathrm{alpha-mu}$$

two new rules in class 3,

$$\frac{\Gamma \vdash r : \bullet A}{\Gamma \vdash r = r : \bullet A} \ \mathrm{eq-rej-r}$$

$$\frac{\Gamma \vdash r = r' : \bullet A \quad \Gamma \vdash s = s' : A}{\Gamma \vdash [r]s = [r']s' \ \mathsf{empty}} \ \mathrm{eq-aff-rej-emp}$$

and the four rules that constitute class 6:

$$\frac{\Gamma \vdash r : \bullet A \quad \Gamma \vdash s : A}{\Gamma \vdash [r]s \ \mathsf{empty}} \ \mathrm{mu-anti} \qquad \frac{\Gamma, \alpha : \bullet A \vdash e \ \mathsf{empty}}{\Gamma \vdash \mu\alpha : A.e : A} \ \mathrm{mu-abs}$$

$$\frac{\Gamma \vdash r : \bullet A \quad \Gamma, \alpha : \bullet A \vdash e \ \mathsf{empty}}{\Gamma \vdash [r]\mu\alpha : A.e = e[\alpha := -^{A}/r] \ \mathsf{empty}} \ \mathrm{mu-beta} \qquad \frac{\Gamma, \alpha : \bullet A \vdash s : A \quad \alpha \notin \mathsf{FV}(s)}{\Gamma \vdash \mu\alpha : A.[\alpha]s = s : A} \ \mathrm{mu-beta}$$

which correspond to antithesis, contraversion, mu-beta reduction and mu-alpha reduction respectively. Denial corresponds to the new class 2 rule, $\mathrm{hyp-den}$.

Zeta rules

Finally, we require a new class of rules, corresponding to the zeta elimination rules. For Π and Σ, these are familiar. For equality types and $\mathbb{N}$, we cannot provide a justification of the decomposition part of the inversion principle as we would require for an eta form zeta rule. Therefore we provide beta form analogues:

$$\frac{\Gamma \vdash \mu\alpha^{\Pi x:A.B}.e : \Pi x^A.B}{\Gamma \vdash \mu\alpha^{\Pi x:A.B}.e = \lambda x^A.\mu\beta^B.e[\alpha := \mathsf{ev}(-,x),\beta] : \Pi x^A.B}\ \Pi - \mathrm{z}$$

$$\frac{\Gamma \vdash \mu\alpha^{\Sigma x:A.B}.e : \Sigma x^A.B}{\Gamma \vdash \mu\alpha^{\Sigma x:A.B}.e = \left\langle \mu\beta^A.e[\sigma_1], \mu\beta^{B[\sigma_2]}.e[\sigma_3] \right\rangle : \Sigma x^A.B}\ \Sigma - \mathrm{z}$$

$$\begin{aligned} \text{where } \sigma_1 &\mathrel{\hat{=}} \alpha := \mathsf{outl}(-)/\beta \\ \sigma_2 &\mathrel{\hat{=}} x := \mathsf{outl}(\mu\alpha : (\Sigma x : A.B).e) \\ \sigma_3 &\mathrel{\hat{=}} \beta := \mathsf{outr}(-)/\beta \end{aligned}$$

$$\frac{\Gamma, \alpha^{s=_A t} \vdash e\ \mathsf{empty} \quad \Gamma \vdash C(x^A, y^A, p^{s=_A t})\ \mathsf{type} \quad \Gamma \to f(p^{s=_A t}) \mathrel{\hat{=}} \mathrm{R}^{=}(p,(u)f) : C(s,t,p)}{\Gamma \vdash f(\mu\alpha^{s=_A t}.e) = \mu\beta^{C(s,t,\mu\alpha.e)}.e[\alpha := f(-)/\beta] : C(s,t,\mu\alpha.e)}\ = -\mathrm{z}$$

$$\frac{\Gamma, \alpha : \bullet\mathbb{N} \vdash e\ \mathsf{empty} \quad \Gamma \vdash C(n^{\mathbb{N}})\ \mathsf{type} \quad \Gamma \to r(n:\mathbb{N}) \mathrel{\hat{=}} \mathrm{R}^{\mathbb{N}}(n,a,(i,z)f) : C(n)}{\Gamma \vdash r(\mu\alpha^{\mathbb{N}}.e) = \mu\beta^{C(\mu\alpha:\mathbb{N}.e)}.e[\alpha := r(-)/\beta] : C(\mu\alpha^{\mathbb{N}}.e)}\ \mathbb{N} - \mathrm{z}$$

Propositional constants

Finally we allow a new class of rule to deal with the propositional constants we considered in chapter 3. These rules only have type former and introduction rules attached to them.

$\top$ **type constant**

$$\frac{\Gamma\ \mathsf{tel}}{\Gamma \vdash \top\ \mathsf{type}}\ \mathrm{const} - \top - \mathrm{f} \qquad \frac{\Gamma\ \mathsf{tel}}{\Gamma \vdash * : \top\ \mathsf{type}}\ \mathrm{const} - \top - \mathrm{i}$$

$\bot$ **type constant**

$$\frac{\Gamma\ \mathsf{tel}}{\Gamma \vdash \bot\ \mathsf{type}}\ \mathrm{const} - \bot - \mathrm{f} \qquad \frac{\Gamma\ \mathsf{tel}}{\Gamma \vdash \mathcal{A}_\bot : \bullet\bot\ \mathsf{type}}\ \mathrm{const} - \bot - \mathrm{i}$$

Failure of the theory NTT

We can define the notion of reduction for NTT in a manner analogous to that for ITT. A point worthy of note is that the conversions associated with the class 6 structural

rules (ie. those responsible for classical strength provability), are considered alongside the logical conversions, and not with the structural conversions, as they are not computationally trivial.

The following can be proved by a method used later in this chapter.

PROPOSITION 156 The conversion theory of NTT is strongly normalising and Church–Rosser.

Unfortunately, the theory is not sound.

PROPOSITION 157

1. $\vdash_{\mathrm{NTT}} 0 =_{\mathbb{N}} 1$ true;
2. The conversion theory of NTT violates subject reduction.

PROOF Define terms $F(-)$, $G(-)$ and d as follows:

$$F(n:\mathbb{N}) \mathrel{\hat{=}} \mathrm{R}^{\mathbb{N}}(n, \underline{0}, (-,-)\underline{1}) : \mathbb{N}$$
$$G(n:\mathbb{N}) \mathrel{\hat{=}} \mathrm{R}^{\mathbb{N}}(n, \underline{0}, (i^{\mathbb{N}}, -)\,\mathrm{R}^{\mathbb{N}}(i, \underline{1}, (-,-)\underline{1}))$$
$$d \mathrel{\hat{=}} \mu\alpha^{\mathbb{N}}.[\alpha]\mathsf{succ}(\mu\beta^{\mathbb{N}}.[\alpha]\underline{0})$$

We can show that

1. $\vdash F(d) = \underline{1} : \mathbb{N}$;
2. $\vdash G(d) = \underline{0} : \mathbb{N}$;
3. $n : \mathbb{N} \vdash F(n) =_{\mathbb{N}} G(n)$ true.

the first two following from elementary formula manipulation, and the third being proven inductively as follows.

Let $C(n:\mathbb{N}) \mathrel{\hat{=}} F(n) =_{\mathbb{N}} G(n)$ type. Then $\vdash \mathsf{refl}(\underline{0}) : C(\underline{0})$ follows by formula manipulation, and similarly we have that $i : \mathbb{N} \vdash \mathsf{refl}(\underline{1}) : C(\mathsf{succ}(i))$. These are the auxiliary premisses to the rule $\mathbb{N}-\mathrm{e}$; for our principal premiss use the variable $n : \mathbb{N}$.

Part one is an elementary corollary of the three cited propositions: substitute d into the proof of $n : \mathbb{N} \vdash F(n) =_{\mathbb{N}} G(n)$ true we have obtained. Reduction of the zeta redex at the root obtains the term candidate

$$\mu\alpha^{C(d)}.[\alpha]H(\mu\beta^{\mathbb{N}}, [\alpha]H(\underline{0}))$$
$$\text{where } H(n:\mathbb{N}) \mathrel{\hat{=}} \mathrm{R}^{\mathbb{N}}(n, \mathsf{refl}(\underline{0}), (i,-)\mathsf{refl}(\mathsf{succ}(i))) : C(n)$$

which cannot be typed, obtaining failure of subject reduction. □

Reflections upon the unsound theory

One might place the blame for the unsoundness of NTT squarely upon those conversion rules which violates the subject reduction, namely the rules $\Sigma-\mathrm{z}$, $\mathbb{N}-\mathrm{z}-\mathrm{e}$ and possibly $=-\mathrm{z}$[1]. However, in the absence of these rules, we obtain additional

[1] I have not been able either to find a $=-\mathrm{z}$ redex which violates subject reduction, or to show that it is sound.

unwanted normal forms for our theory, to supplement such inhabitants as the closed normal form

$$\vdash_{\mathrm{NTT}} \mu\alpha^{\mathbb{N}}.[\alpha]\mathsf{succ}(\mu\beta^{\mathbb{N}}.[\alpha]\underline{0}) : \mathbb{N}$$

inhabiting the natural numbers. These closed normal forms which are not canonical forms we call *deviant terms*: in the presence of such terms we do not know how to claim that logical harmony prevails. Indeed the existence of the deviant natural number above suggests that we need the rule:

$$\frac{\Gamma, \alpha : \bullet\mathbb{N} \vdash e \ \mathsf{empty}}{\Gamma \vdash \mathsf{succ}(\mu\alpha^{\mathbb{N}}.e) = \mu\beta^{\mathbb{N}}.e[\alpha := \mathsf{succ}(-)/\beta] : \mathbb{N}} \ \mathbb{N} - \mathrm{z} - \mathsf{succ}$$

PROPOSITION 158 We can find deviant inhabitants of type $\mathbb{N}$ in the absence of any one of the following rules:

1. $\mathbb{N} - \mathrm{z} - \mathsf{succ}$;
2. $\mathbb{N} - \mathrm{z} - \mathrm{e}$;
3. $= -\mathrm{z}$;
4. $\Sigma - \mathrm{z}$.

PROOF All of the following are normal forms in the absence of the indicated conversion.

1. $\vdash \mu\alpha^{\mathbb{N}}.[\alpha]\mathsf{succ}(\mu\beta^{\mathbb{N}}.[\alpha]\underline{0}) : \mathbb{N}$;
2. $\vdash \mu\alpha^{\mathbb{N}}.[\alpha]\,\mathrm{R}^{\mathbb{N}}(\mu\beta^{\mathbb{N}}.[\alpha]\underline{0}, \underline{0}, (i, z)\underline{1}) : \mathbb{N}$;
3. $\vdash \mu\alpha^{\mathbb{N}}.[\alpha]\,\mathrm{R}^{=}(\mu\beta^{0=_{\mathbb{N}}1}.[\alpha]\underline{0}, (x^{\mathbb{N}})x) : \mathbb{N}$;
4. $\vdash \mu\alpha^{\mathbb{N}}.[\alpha]\mathsf{outr}(\mu\beta^{\Sigma x \in \mathbb{N}.x=x}.[\alpha]\underline{0})$.

□

With the exception of the rule $\mathbb{N} - \mathrm{z} - \mathsf{succ}$, and possibly $= -\mathrm{z}$, all of the above three conversions are responsible for the failure of subject reduction; the rule $\mathbb{N} - \mathrm{z} - \mathsf{succ}$ together with $\mathbb{N} - \mathrm{z} - \mathrm{e}$ allows us to show $\vdash \underline{0} = \underline{1} : \mathbb{N}$, since the underlying rewrite mechanism is not Church–Rosser as we showed in section 3.4.

Consequently we do not seem to be able to accept or reject any of the above rewrites: all of the sixteen possibilities suffer from at least one of the following afflictions:

1. Failure of subject reduction;
2. Failure of Church–Rosser;
3. $\vdash 0 =_{\mathbb{N}} 1 \ \mathsf{true}$;
4. Existence of deviant inhabitants of $\mathbb{N}$.

Indeed the theory with all of the zeta rules except $= -z$ suffers from all of the above!

Furthermore it is not possible to give a conversion rules for the strong sum type former in the presence of the classical structural rules without making our type theory unsound in some way. This can be shown in quite general way following an argument of Thierry Coquand[2]. To give this general characterisation, let us say that an *arithmetic type theory* is a theory possessing the structural rules of ITT and the type former, introduction and elimination rules for Π, $\mathbb{N}$ and $=_A$, but leave the conversion rules unspecified. We use this characterisation to talk about the possible limits of expressivity of type theory.

REMARK 159
If an arithmetic type theory satisfies subject reduction, the Church–Rosser property, strong normalisation, and the canonical form property, then the argument used for ITT ensures consistency; the reader should satisfy himself that no three alone suffice. We shall therefore call arithmtic type theories which satisfy all four of these properties *well-behaved*.

PROPOSITION 160 If an arithmetic type theory contains the strong sum type former and validates the principle of the excluded middle, then it is not sound.

PROOF We show that a canonical form theorem would give a decision procedure for any proposition, which contradicts the existence of such undecidable statements as the Gödelian consistency sentence for the theory. Let A be any closed proposition whose truth-value we wish to determine. The predicate $\chi_A(n^{\mathbb{N}})$:

$$n : \mathbb{N} \vdash \chi_A(n) \equiv n =_{\mathbb{N}} \underline{0} \Leftrightarrow A \text{ prop}$$

codes up A possessing a truth-value as a predicate. It is left as an easy exercise for the reader to verify that there is a term d, making use of strong existence and classical logic, which satisfies

$$\vdash d : \exists n \in \mathbb{N}.\chi_A(n)$$

Were the canonical form theorem valid for Λ, then the normal form of d would match either $\langle \underline{0}, d_1 \rangle$ where $\vdash d_1 : \chi_A(\underline{0})$ or $\langle \mathsf{succ}(\underline{k}), d_2 \rangle$ where $\vdash d_1 : \chi_A(\mathsf{succ}(\underline{k}))$. Since $\chi_A(\underline{0})$ is provably equivalent to A and $\chi_A(\mathsf{succ}(\underline{k}))$ is provably equivalent to $A \supset 0 =_{\mathbb{N}} 1$, the normal form determines the truth-value of the formula A. But since the theory satisfies strong normalisation we have an effective method for calculating the canonical form of d, contradicting the undecidability of A. □

[2]'Computational content of classical logic' [CoquandT:comccl]; his argument here is derived from his joint work with Franco Barbanera and Stefano Berardi in 'Computational content of the axiom of choice'[CoquandT:comcac].

Our calculus of classical type theory, which we shall introduce in the next section, is founded upon three measures to eliminate these pathological behaviours. Firstly, we replace the right recursive form of $\mathrm{R}^{\mathbb{N}}$ by the left recursive form so that we are able to introduce the rule '$\mathsf{succ}(\mu\alpha^{\mathbb{N}}.e) \to^* \mu\beta^{\mathbb{N}}.e[\alpha := \mathsf{succ}(-)/\beta]$' and thus eliminate the deviant points of $\mathbb{N}$, without permitting the derivability of arithmetic absurdity. These properties follow from the satisfaction of the Church–Rosser property for the underlying left-recursive form of $\mathrm{PRA}^{\omega}_{\mu}$.

Secondly, we restrict the zeta rules to operate only upon vacuous contraversions, we regain the subject reduction property, since it is possible to give a vacuous contraversion any type we please. It is not obvious that this measure is sufficient to eliminate deviant terms from the calculus, and the proof of this result is the cornerstone of our adequacy result for the calculus we are about to introduce.

Thirdly, we must abandon the strong sum type former. This is unfortunate: it means that we cannot show such desirable results as the axiom of choice, and also the strong sum operator is convenient for the coding of relations into types. However we can still handle existentials synthetically, by means of the $\exists x \in A.\phi(x) \mathrel{\hat{=}} \neg\forall x \in A.\neg\phi(x)$ encoding where the synthetic type former $\neg$ is defined $\neg A \mathrel{\hat{=}} A \supset \bot$. This formulation is sufficient to define such relations as $<$ by analogy with the treatment we gave before.

4.2 The calculus CTT

Our type theory, *classical type theory*, or CTT, has the same structural rules as the theory NTT, but lacks the rules associated with the type former Σ, and our conversion rules differ at the $\mathbb{N}$ type and $=_A$ type, as we discussed in the last section.

$\mathbb{N}$ type former:

$$\frac{\Gamma, n : \mathbb{N} \vdash r(n) \equiv \mathrm{R}^{\mathbb{N}}(s, a, (i, z)t) : C(n)}{\Gamma \vdash r(\underline{0}) = a : C(\underline{0})}\ \mathbb{N}-\mathrm{c}-\underline{0}$$

$$\frac{\Gamma \vdash s : \mathbb{N} \quad \Gamma, n : \mathbb{N} \vdash \mathrm{R}^{\mathbb{N}}(s, a, (i^{\mathbb{N}}, z^{C(i)})t) : C(n)}{\begin{array}{c}\mathrm{R}^{\mathbb{N}}(\mathsf{succ}(s), a, (i, z)t) = \mathrm{R}^{\mathbb{N}}(s, a', (j, z^{C(j)})t') : C(\mathsf{succ}(s)) \\ \text{where } a' \mathrel{\hat{=}} t[i := \underline{0}, z := a] \\ t' \mathrel{\hat{=}} t[i := \mathsf{succ}(j)]\end{array}}\ \mathbb{N}-\mathrm{c}-\mathsf{succ}$$

$$\frac{\Gamma, \alpha : \bullet\mathbb{N} \vdash e\ \mathsf{empty} \quad \Gamma, n : \mathbb{N} \vdash r(n) \equiv \mathrm{R}^{\mathbb{N}}(s, a, (i, z)t) : C(n) \quad \beta \notin \mathsf{dom}(\Gamma)}{\Gamma \vdash r(\mu\alpha^{\mathbb{N}}.e) = \mu\beta^{C(\mu\alpha.e)}.e[\alpha := r(-)/\beta] : C(\mu\alpha.e)}\ \mathbb{N}-\mathrm{z}-\mathrm{e}$$

$$\frac{\Gamma, \alpha : \bullet\mathbb{N} \vdash e\ \mathsf{empty}}{\Gamma \vdash \mathsf{succ}(\mu\alpha^{\mathbb{N}}.e) = \mu\beta^{\mathbb{N}}.e[\alpha := \mathsf{succ}(-)/\beta] : \mathbb{N}}\ \mathbb{N}-\mathrm{z}-\mathsf{succ}$$

$=_A$ type former:

$$\frac{\Gamma \vdash s =_A t \ \mathsf{type} \quad \Gamma, x : A \vdash r(x) \equiv \mathrm{R}^{=}(x,(a)d) : C(x,x,\mathsf{refl}(x))}{r(\mathsf{refl}(s)) = d[a := s] : C(s,t,\mathsf{refl}(s))} = -\mathrm{c}$$

$$\frac{\Gamma, \alpha : \bullet s =_A t \vdash e \ \mathsf{empty} \quad \Gamma, x : A \vdash r(x) \equiv \mathrm{R}^{=}(x,(a)d) : C(x,x,\mathsf{refl}(x))}{r(\mu\alpha.e) = \mu\beta.e[\alpha := r(-)/\beta] : C(s,t,\mathsf{refl}(s))} = -\mathrm{z}$$

4.3 Type soundness

With the restriction on substitutions and zeta reductions, this section presents no new difficulties over the theory of ITT; all of the parts are established with only inessential variations of the proofs of section 2.3.

PROPOSITION 161

1. If Γ, Γ' tel then Γ tel;
2. If $\Gamma, \mathcal{A}, \Gamma'$ then $\mathsf{dom}(\mathcal{A}) \notin \mathsf{dom}(\Gamma')$ and $\mathsf{dom}(\mathcal{A}) \notin \mathsf{ran}(\Gamma)$;
3. (Weakening) If $\Gamma, \Gamma'' \vdash \mathcal{J}$, $\mathsf{dom}(\Gamma) \cap \mathsf{dom}(\Gamma') = \emptyset$, and Γ, Γ' tel then $\Gamma, \Gamma', \Gamma'' \vdash \mathcal{J}$;
4. (Exchange) If $\Gamma, \mathcal{A}, \mathcal{A}', \Gamma' \vdash \mathcal{J}$ and $\Gamma, \mathcal{A}'$ tel then $\Gamma, \mathcal{A}', \mathcal{A}, \Gamma' \vdash \mathcal{J}$;
5. If $\Gamma \vdash s : A$ or $\Gamma \vdash s = s' : A$ then s satisfies the bound identifier convention.

LEMMA 162 (SUBTERM LEMMA)
Let d justify $\Gamma \vdash s : A$, and let $s' \in \mathsf{st}(s)$. Then there is a telescope candidate Γ' and a type candidate A such that $\Gamma, \Gamma' \vdash s' : A'$ is the conclusion of some sub-inference of d.

LEMMA 163 (SUBSTITUTION LEMMA)
Let $\Gamma \vdash s : A$ and let $\sigma :: \Gamma \to \Gamma'$ be a substitution in which the only substitution operations occurring on rejective variables are on variables of function space type. Then $\Gamma' \vdash \mathcal{J}[\sigma]$.

PROPOSITION 164 (SUBJECT REDUCTION)

1. If $\Gamma \vdash A = A'$ type then $\Gamma \vdash A$ type and $\Gamma \vdash A'$ type;
2. (a) If $\Gamma \vdash s = s' : A$ then $\Gamma \vdash s : A$ and $\Gamma \vdash s' : A$;
   (b) If $\Gamma \vdash r = r' : \bullet A$ then $\Gamma \vdash r : \bullet A$ and $\Gamma \vdash r' : \bullet A$;
   (c) If $\Gamma \vdash e = e'$ empty then $\Gamma \vdash e$ empty and $\Gamma \vdash e'$ empty;
3. If $\Gamma \vdash s : A$ then $\Gamma \vdash A$ type.

PROOF We note that there are two new class 3 rules which may be used to infer a term equality, but they are quite elementary cases. So too with the new alpha conversions and the $\mathrm{mu} - \mathrm{alpha}$ and $\mathrm{mu} - \mathrm{beta}$ rules. Type judgements are the same, and the proofs go through by analogy with proposition **??** □

PROPOSITION 165

1. (a) If $\Gamma = \Gamma'$ tel then Γ tel;
   (b) If $\Gamma \vdash \mathcal{J}$ and $\Gamma = \Gamma'$ tel then $\Gamma' \vdash \mathcal{J}$;
2. If $\Gamma \vdash \mathcal{J}$ then Γ tel;
3. If $\Gamma \vdash s : A$ or $\Gamma \vdash r : \bullet A$ then $\Gamma \vdash A$ type.

4.4 Conversion theory

We follow quite a similar treatment compared to section 2.6, in showing that CTT is strongly normalising and Church–Rosser.

DEFINITION 166

1. We write $\Gamma \vdash s \to^c s' : A$, or that s converts to s', if there is an inference of the judgement $\Gamma \vdash s = s' : A$ whose last rule is a conversion rule governed by a type former, or occurs in the class 6 structural rules;

2. We write $\Gamma \vdash s =_\alpha s' : A$, for *alpha equivalence*, if there is an inference of the judgement $\Gamma \vdash s = s' : A$ in which none of the conversion rules governed by type formers, and none of the conversion rules of class 6 appear (ie. an inference which only contains the structural rules of classes 1 to 5);

3. We write $\Gamma \vdash s \to^1 s' : A$ if there is a context-redex decomposition $s \equiv f(r)$ where $\Gamma \vdash f((\Delta)x : B) : A, \Gamma, \Delta \vdash r \to^c c : B$ and $s' \equiv f(c)$, which satisfies an additional restriction in the case of the $\Pi - \mathrm{c} - \mathrm{eta}$ conversion. This condition is that r is not arise by the $\Pi -$ i rule or by a contraversion, and that the variable x in $f(x)$ is not the principal premiss if a $\Pi - e$ rule;

4. We write $\Gamma \vdash s \to^* s' : A$ if either $s \equiv s'$ or there is a sequence of term candidates $s_1, \ldots, s_n$ such that $s_1 \equiv s$, $s_n \equiv s'$ and for $1 \leq i < n$, $\Gamma \vdash s_i \to^1 s_{i+1} : A$.

PROPOSITION 167

1. If $\Gamma \vdash s \to^* s' : A$ then $\Gamma \vdash s = s' : A$;
2. $\Gamma \vdash s \to^c s' : A$ iff there is an inference of the judgement $\Gamma \vdash s = s' : A$ in which none of the class 3, 4 or 5 structural rules appear.

Our proof of strong normalisation proceeds by translation.

DEFINITION 168 We define the *dependency forgetting translation* of judgements of CTT into judgements of $\lambda\mu\mathbb{N}_l$:

1. Firstly there is the translation on types:

$$\begin{aligned}
[\![X]\!] &\hat{=} X, \quad [\![\mathbb{N}]\!] \hat{=} \mathbb{N} \\
[\![\bot]\!] &\hat{=} \bot, \quad [\![\top]\!] \hat{=} \top \\
[\![\Pi x : A.B]\!] &\hat{=} [\![A]\!] \Rightarrow [\![B]\!] \\
[\![s =_A t]\!] &\hat{=} [\![A]\!]
\end{aligned}$$

2. Then the translation on hypotheses:

$$\begin{aligned}
[\![X \ \mathsf{type}]\!] &\hat{=} \emptyset \\
[\![x : A]\!] &\hat{=} \{x : [\![A]\!]\} \\
[\![\alpha : \bullet A]\!] &\hat{=} \{\alpha : \bullet[\![A]\!]\}
\end{aligned}$$

3. and then the translation on term candidates:

$$\begin{aligned}
[\![x]\!] &\hat{=} x \qquad [\![\alpha]\!] \hat{=} \alpha \\
[\![[r]s]\!] &\hat{=} [[\![r]\!]][\![s]\!] \\
[\![\mu\alpha^A.e]\!] &\hat{=} \mu\alpha^{[\![A]\!]}.[\![e]\!] \\
[\![\lambda x^A.s]\!] &\hat{=} \lambda x^{[\![A]\!]}.[\![s]\!] \\
[\![\mathsf{ev}(s,t)]\!] &\hat{=} \mathsf{ev}([\![s]\!],[\![t]\!]) \\
[\![\mathsf{refl}(s)]\!] &\hat{=} \mathsf{refl}([\![s]\!]) \\
[\![\mathrm{R}^{=}(s,(x^A)t)]\!] &\hat{=} \mathsf{eqev}(\lambda x^{[\![A]\!]}.[\![t]\!],[\![s]\!]) \\
[\![\mathsf{zero}]\!] &\hat{=} \mathsf{zero} \qquad [\![\mathsf{succ}(n)]\!] \hat{=} \mathsf{succ}([\![n]\!]) \\
\left[\!\!\left[\mathrm{R}^{\mathbb{N}}(s,a,(i^{\mathbb{N}},z^{C(i)})f)\right]\!\!\right] &\hat{=} \mathsf{Rec}([\![s]\!],[\![a]\!],\lambda i^{\mathbb{N}}.\lambda z^{[\![C(i)]\!]}.[\![f]\!]) \\
[\![*]\!] &\hat{=} * \qquad [\![\mathcal{D}_\bot]\!] \hat{=} \mathcal{A}_\bot
\end{aligned}$$

PROPOSITION 169 (THE DIAMOND PROPERTY)
If $\Gamma \vdash s \to^1 t_1 : A$ and $\Gamma \vdash s \to^1 t_2 : A$ then there is a term candidate u such that $\Gamma \vdash t_1 \to^* u : A$ and $\Gamma \vdash t_2 \to^* u : A$.

PROOF By analogy with $\lambda\mu\mathbb{N}_l$, just as we did for ITT. □

PROPOSITION 170
1. If $\Gamma \vdash_{\mathrm{CTT}} s : A$ then $\bigcup\{[\![\mathcal{A}]\!] \mid \mathcal{A} \in \Gamma\} \vdash_{\lambda\mu\mathbb{N}_l} [\![s]\!] : [\![A]\!]$;
2. If $\Gamma \vdash_{\mathrm{ITT}} s \to^1 s' : A$ then $[\![s]\!] \to^* [\![s']\!]$ in $\lambda\mu\mathbb{N}_l$ and $[\![s]\!] \not\equiv [\![s']\!]$.

COROLLARY 171 The relation $\to^1$ is strongly normalising.

COROLLARY 172 The relation $\to^1$ is Church–Rosser.

PROPOSITION 173

1. If $\Gamma \vdash s =_\alpha s' : A$ and $\Gamma \vdash s \to^1 t : A$ then there is a term candidate t' such that $\Gamma \vdash s' \to^1 t' : A$ and $\Gamma \vdash t =_\alpha t' : A$;
2. Alpha equivalence is decidable.

PROOF Our proof of part one follows the same lines as the analogous proposition in section 2.6. Our algorithm for part two is extended to include relabelling of rejective identifier bindings, and the proof of correctness follows the same form as that of ITT. □

THEOREM 174 The equality relation on term candidates is decidable.

PROOF The proof follows the same form as that of ITT. □

COROLLARY 175 Let s and t be alpha-distinct normal forms. Then for no telescope Γ and type A do we have $\Gamma \vdash s = t : A$.

Canonical form theorem

Recall that a term is a canonical form if it matches one of $\lambda x^A.s$, $\underline{k}$, $\mathsf{refl}(s)$ or $*$.

DEFINITION 176

1. The *head subterms* of a term s is a subset of $\mathsf{st}(s)$ consisting of:

   (a) Just the term itself if it is a canonical form or a variable;
   (b) The term itself together with the head subterms of s' if s matches one of $\mathsf{ev}(s', t)$, $\mathrm{R}^{\mathbb{N}}(s', t, (i, z)u)$ or $\mathrm{R}^{=}(s', (x)t)$;
   (c) The term itself together with the head subterms of s' in case $s \equiv \mathsf{succ}(s')$ and s' is not a canonical form;
   (d) The term itself together with the head subterms of s' if $s \equiv \mu\alpha^A.[r]s'$.

   Note that the head subterms are all distinct, and as occurrences of s are arranged from left to right in the term, rather like Russian dolls;

2. We call a term *head reducible* if one of its head subterms is a redex. We say that a term is in *head normal form*, or *HNF*, if it is not head reducible. Finally a term is in *quasi HNF* if it is in head normal form and has the additional property that all of its subterms of the form $\mathsf{refl}(t)$ are normal forms.

3. If there is a variable amongst the head subterms of a term, we call this its *head variable*. Note that it is necessarily unique, and that unlike in ITT it is possible for a term to possess a head variable and yet be head reducible.

4. If there are any contraversions amongst the head subterms, then these are called *head contraversions*.

The following lemma is quite elementary to show:

LEMMA 177

1. If a term has a head variable, then it is a free variable;
2. If a head contraversion is vacuous, then it is either vacuous, or it is the principal subterm of a head redex.

THEOREM 178 (QUASI HNF THEOREM)
If a term is in quasi HNF then one of the following applies:

1. It is a canonical form;
2. It has a head variable;
3. It possesses unbound rejective variables.

PROOF Assume that none of these apply, and that all subterms of the form $\mathsf{refl}(t)$ are normal forms. We shall show that the term possesses a head redex.

Since the term does not have a head variable, by inspection of the definition of head subterms we see that the rightmost subterm must be a canonical form. The next rightmost subterm (ie. the head subterm immediately to its left) must either be a redex, or it must match one of the cases:

1. $\mu\alpha^A.[r]\lambda x^B.s$: r must be a rejective variable, and since it isn't free it must be bound in a head contraversion, which must be a head zeta redex;
2. $\mu\alpha^A.[r]\underline{k}$: $\alpha \notin \mathsf{FV}(\underline{k})$, and so either $r \equiv \alpha$ in which case this is a $\mu\eta$ redex, or it is a vacuous contraversion, which by the lemma above must form part of a head redex;
3. $\mu\alpha^A.[r]\mathsf{refl}(s)$: We shall show that $\alpha \notin \mathsf{FV}(s)$, which by the same argument as in the above case shows that there must be a head redex.

   Since, by the condition above, s must be a normal form and so the type of $\mathsf{refl}(s)$ will be $t_0 =_B t_1$ where the normal forms of t_0 and t_1 is s. Thus if $\alpha \in \mathsf{FV}(s)$ then α occurs in the type of $\mathsf{refl}(s)$, and so it occurs in the type of r. But this is impossible, as α is discharged, so whether $r \equiv \alpha$ must occur either in the type A or in the type of β.

□

REMARK 179 We observe that the condition on normalcy of subterms of the form $\mathsf{refl}(s)$ is necessary, since otherwise the theorem would admit the following counterexample:

$$\mu\alpha^{\underline{0}=_{\mathbb{N}}\underline{0}}.[\alpha]\mathsf{refl}(\mathsf{ev}(\lambda x^{\mathbb{N}}.\underline{0}, \mu\beta^{\mathbb{N}}.[\alpha]\mathsf{refl}(\underline{0})))$$

which is a closed head normal form but not a canonical form.

COROLLARY 180 (CANONICAL FORM THEOREM)
All closed normal forms are canonical forms.

COROLLARY 181 $\not\vdash 0 =_{\mathbb{N}} 1\ \mathsf{true}$

4.5 Representability

In section 2.4 we showed how we could encode the theories of intuitionistic equational logic, PRA$^\omega$ and Heyting Arithmetic into ITT. Now we shall show how it is possible to obtain classical analogues for these theorems. First let us prove that double-negation holds.

PROPOSITION 182 Let $\Gamma \vdash A\,\mathsf{type}$. Then $\Gamma, x : \neg\neg A \vdash \mu\alpha^A.[\mathcal{D}_\perp]\mathsf{ev}(x, \lambda y^A.\mu\epsilon^\perp.[\alpha]y) : A$.

Equational logic The main technical results allowing us to carry out the coding in ITT were:

1. To show that the equality type is a type former;
2. To show Leibniz's rule holds;
3. To show that functions conserve propositional equality.

All of these properties are equivalent to the inhabitation of certain types in CTT: these are in each case presentable by the same term as for the analogous proposition of ITT.

Classical equational logic has the same formulae as intuitionistic equational logic, and has all of the axioms plus the axiom scheme $\neg\neg\phi \supset \phi$. Since we can interpret all of the axioms and axiom schemas, we obtain the following proposition direct analogue of corollary 68.

PROPOSITION 183 If ϕ is a theorem of classical equational logic, then $\Gamma \vdash_{\text{CTT}} [\![\phi]\!]\,\mathsf{true}$, where Γ consists only of those assumptions needed to declare the variables and function letters of ϕ. Furthermore it is a theorem of the subsystem of CTT with only the Π and $=$ type formers, and the rule $\mathbb{N} - \mathrm{f}$.

We mentioned that it was possible to handle existential quantification by means of the following axiom schemas:

1. $\phi[x := s] \supset \exists x.\phi$;
2. $\exists x.\phi \supset ((\forall x.\phi) \supset \psi) \supset \psi$.

With the coding $[\![\exists x.\phi]\!] \triangleq \neg\Pi x^{\mathbb{N}}.\neg[\![\phi]\!]$ it is possible to find witnesses to the instances of the above scheme.

PRA and related theories It is perfectly easy to translate the left-recursive form of each of the theories PRA, PRA$^\omega$ and PRA$^\omega_\mu$ into 'logic-free' CTT.

Peano Arithmetic Recall that the crucial lemmas necessary for our interpretation of Heyting Arithmetic in Peano Arithmetic were as follows.

1. Definability of pred, add and mul, together with the equations specifying their evaluative behaviour;
2. Demonstration that $\underline{0} =_{\mathbb{N}} \underline{1} \supset \phi$ obtains for each instance ϕ that is used in the encoding of the formulae of Heyting Arithmetic;
3. Proof of Peano's axioms with the weak fourth axiom.

Recall that we had some trouble with the defining equations in the left-recursive form of PRA. However these difficulties evaporate in the presence of the $= -\mathrm{e}$ recursor, and all of the usual right-recursive specifying equations can be proven, which the reader is invited to confirm.

EXERCISE 184

1. (a) $\vdash \forall m^{\mathbb{N}}.\mathsf{pred}(\mathsf{succ}(m)) =_{\mathbb{N}} m$ true;
   (b) $\vdash \mathsf{pred}(\underline{0}) =_{\mathbb{N}} \underline{0}$ true;
2. (a) $\vdash \forall m^{\mathbb{N}}.\mathsf{add}(\underline{0}, m) =_{\mathbb{N}} m$ true;
   (b) $\vdash \forall m^{\mathbb{N}}.\forall n^{\mathbb{N}}.\mathsf{add}(\mathsf{succ}(m), n) \mathrel{\hat{=}} \mathsf{succ}(\mathsf{add}(m, n))$ true;
   (c) $\vdash \forall m^{\mathbb{N}}.\forall n^{\mathbb{N}}.\mathsf{add}(m, n) \mathrel{\hat{=}} (\mathsf{add}(n, m))$ true;
3. (a) $\vdash \forall m^{\mathbb{N}}.\mathsf{mul}(\underline{0}, m) =_{\mathbb{N}} \underline{0}$ true;
   (b) $\vdash \forall m^{\mathbb{N}}.\forall n^{\mathbb{N}}.\mathsf{mul}(\mathsf{succ}(m), n) \mathrel{\hat{=}} \mathsf{add}((\mathsf{mul}(m, n), n)$ true;
   (c) $\vdash \forall m^{\mathbb{N}}.\forall n^{\mathbb{N}}.\mathsf{mul}(m, n) \mathrel{\hat{=}} (\mathsf{mul}(n, m))$ true;

Similarly we have no difficulty with proving that Peano's axioms, with the weak formulation of the fourth axiom, hold in CTT. Just as with classical equational logic, Peano Arithmetic is obtained from Heyting Arithmetic by adding the axiom scheme $\neg\neg\phi \supset \phi$. There is a little twist with this axiom: we must show not $\neg\neg[\![\phi]\!] \supset [\![\phi]\!]$, but $\neg_a \neg_a [\![\phi]\!] \supset [\![\phi]\!]$, where $\neg_a(A) \mathrel{\hat{=}} A \supset \underline{0} =_{\mathbb{N}} \underline{1}$.

PROPOSITION 185 If $A \mathrel{\hat{=}} [\![\phi]\!]$ for some proposition ϕ of Peano Arithmetic, then $\neg_a \neg_a A \supset A$.

PROOF The following proof tree can be translated into a term of the required type:

$$\dfrac{\dfrac{\dfrac{[\bullet\underline{0} =_{\mathbb{N}} \underline{1}]^{\beta} \qquad \dfrac{[\neg_a \neg_a [\![\phi]\!]]^{h} \qquad \dfrac{\dfrac{[\bullet A]^{\alpha}[A]^{x}}{\underline{0} =_{\mathbb{N}} \underline{1}}\,\mu\gamma}{\neg_a A}\,\lambda x}{\underline{0} =_{\mathbb{N}} \underline{1}}}{\underline{0} =_{\mathbb{N}} \underline{1}}\,\mu\beta}{A}\,(*)}{\neg_a \neg_a A \supset A}\,\lambda h$$

where (*) denotes the application of the derivation $\underline{0} =_{\mathbb{N}} \underline{1} \supset [\![\phi]\!]$. □

We obtain the following analogue of corollary 77.

PROPOSITION 186 If ϕ is a theorem of PA, then $\Gamma \vdash_{\text{CTT}} [\![\phi]\!]\ \mathsf{true}$, where Γ consists only of enough set-valued assumptions to bind the variables and function letters of ϕ.

Strong existential quantifier Recall that we described three uses of the Σ type in ITT:

1. To represent existential quantification;
2. To encode relations as type;
3. To encode signatures, subsets and quotient types.

We have seen that we can represent existential quantification in CTT, using the representation '$\exists x^A.\phi(x) \mathrel{\hat{=}} \neg(\forall x^A.\neg\phi(x))$'. However it is not possible to show the axiom of choice for this type. This is not altogether surprising: the rules of Σ allow us to show the existence property, and we could hardly demand this of a classical proof theory. We shall revisit this matter in the conclusion.

Another consequence is that CTT does not allow us to define existential properties in the same way, such as, for example our definition of the less-than relation $\mathsf{LT}(a^{\mathbb{N}}, b^{\mathbb{N}}) \mathrel{\hat{=}} \Sigma i^{\mathbb{N}}.a =_{\mathbb{N}} \mathsf{succ}(\mathsf{plus}(i, b))$. We could code these up be means of test functions, as are used in recursive function theory, but it would be more desirable to code them as types.

Recall the specification of the type LQ representing 'less than or equal to':

1. $\vdash \forall m^{\mathbb{N}}.\mathsf{LQ}(\underline{0}, m)\ \mathsf{true}$;
2. $\vdash \forall m^{\mathbb{N}}, n^{\mathbb{N}}.\mathsf{LQ}(m, n) \supset \mathsf{LQ}(\mathsf{succ}(m), \mathsf{succ}(n))\ \mathsf{true}$;
3. $\vdash \forall l^{\mathbb{N}}, m^{\mathbb{N}}, n^{\mathbb{N}}.\mathsf{LQ}(l, m) \supset \mathsf{LQ}(m, n) \supset \mathsf{LQ}(l, n)\ \mathsf{true}$;
4. $\vdash \forall m^{\mathbb{N}}, n^{\mathbb{N}}.\mathsf{LQ}(\mathsf{succ}(m), \mathsf{succ}(n)) \supset \mathsf{LQ}(m, n)\ \mathsf{true}$;
5. $\vdash \forall m^{\mathbb{N}}, n^{\mathbb{N}}.\mathsf{LQ}(m, n) \supset \mathsf{LQ}(n, m) \supset m =_{\mathbb{N}} n\ \mathsf{true}$.

Can we find a candidate for our type in CTT? Two candidates suggest themselves. The simpler type is based an inversion of 'strictly more than':

$$\mathsf{LQ}(a : \mathbb{N}, b : \mathbb{N}) \mathrel{\hat{=}} \forall i^{\mathbb{N}}.(a =_{\mathbb{N}} \mathsf{add}(\mathsf{succ}(b), i) \supset 0 =_{\mathbb{N}} 1)$$

However it does not appear possible to give a witness to the third, transitivity equation. Also, we do seem to need to code up negation in the arithmetic way (ie. by means of $\neg_a A \mathrel{\hat{=}} A \supset 0 =_{\mathbb{N}} 1$), in order to obtain the last property. The more complex type is based upon an arithmetic rereading of De Morgan's law for the existential:

$$\mathsf{LQ}(a : \mathbb{N}, b : \mathbb{N}) \mathrel{\hat{=}} (\forall i^{\mathbb{N}}.\mathsf{add}(a, i) =_{\mathbb{N}} b \supset 0 =_{\mathbb{N}} 1) \supset 0 =_{\mathbb{N}} 1$$

The trickiest part of this is the proof of transitivity, though it can be proven with only intuitionistic arithmetic. It can be obtained by currying from the following term witness:

$$a : \mathbb{N}, b : \mathbb{N}, c : \mathbb{N}, u : \mathsf{LQ}(a,b), v : \mathsf{LQ}(b,c) f : \forall i^{\mathbb{N}}.\mathsf{add}(a,i) =_{\mathbb{N}} c \vdash \\ \mathsf{ev}(u, \lambda j^{\mathbb{N}}.\lambda p^{a+j=b}.\mathsf{ev}(v, \lambda k^{\mathbb{N}}.\lambda q^{b+k=c}.\mathsf{ev}(f, \mathsf{add}(j,k), *) : 0 =_{\mathbb{N}} 1$$

The first part can be constructed by induction using this result, and the second and fourth parts are almost immediate.

In general we would expect to be able to code up the same kinds of relations in CTT as in ITT, but we should not be surprised to find that there may be certain properties of relations in ITT that are not conserved by their best analogue in CTT.

Finally, a brief mention of the coding of signatures, subsets and quotients in CTT. The use of the strong sum operator is essential to the behaviour of these features; however the strong sum is used only at the top level in these features, and there is nothing to stop us defining a strong sum operator in our constructive meta-theory. Consequently there is no reason to expect the use of these features to be affected by the absence of the strong connective in CTT.

4.6 Consistency and harmony

Recall that we distinguished between four strengths of consistency claim, namely:

1. There are uninhabited types;
2. Arithmetic is consistent (ie. $\nvdash 0 =_{\mathbb{N}} 1$ true);
3. We can provide a static semantics for the theory;
4. The conversion theory of the calculus is sound.

It is easy to show that there are uninhabited types: Smith's 'proof-irrelevance' semantics carries over perfectly well[3] to CTT, showing that negations of equality type are unprovable.

The canonical form theorem also allows us to show the unprovability of the judgement $\vdash 0 =_{\mathbb{N}} 1$ true just as it did for ITT. Furthermore, as we have shown, there is an associated notion of reduction for CTT which is strongly normalising and Church–Rosser.

However there are problems facing the presentation of a direct semantics, and so, for the time being at least, we must depend upon the reduction semantics for the soundness and consistency of our calculus.

[3]We simply interpret assumptions of type $\alpha : \bullet A$ by $x :_{\alpha}: \neg_{\mathbb{B}}\phi_T(A)$, and we must show that if $\Gamma \vdash s$ empty that the conjunctions of the truth-values associated with Γ is ff.

Failure of extensional static semantics The difficulty with CTT is that functions (ie. terms inhabiting Π types) cannot be given an extensional interpretation. To see this consider the following functions, definable in ITT: $f \triangleq \lambda x^{\mathbb{N}}.\underline{0}$ and $g \triangleq \lambda x^{\mathbb{N}}.\,\mathrm{R}^{\mathbb{N}}(x, \underline{0}, (-,-)\underline{0})$. We have for these functions

$$\vdash_{\Lambda} \forall i : \mathbb{N}.\mathsf{ev}(f, i) =_{\mathbb{N}} \mathsf{ev}(g, i)\ \mathsf{true}$$

where Λ is either ITT or CTT, and so f and g have the same denotation in the direct semantics we gave for ITT. However the following term of CTT can distinguish them:

$$F \triangleq \lambda h^{\mathbb{N} \Rightarrow \mathbb{N}}.\mu\alpha^{\mathbb{N}}.[\alpha]\mathsf{ev}(h, \mu\beta^{\mathbb{N}}.[\alpha]\underline{1})$$

since $\vdash_{\mathrm{CTT}} \mathsf{ev}(F, f) = \underline{0} : \mathbb{N}$ and $\vdash_{\mathrm{CTT}} \mathsf{ev}(F, g) = \underline{1} : \mathbb{N}$. Indeed it is possible to construct a witness to $0 =_{\mathbb{N}} 1\ \mathsf{true}$ from the following extensionality axiom:

$$X\ \mathsf{type}, Y\ \mathsf{type} \vdash \mathsf{ext}_{X,Y} : \forall f : \Pi(X, Y).\forall g : \Pi(X, Y).$$
$$(\forall a : X.\mathsf{ev}(f, a) =_{Y} \mathsf{ev}(g, a)) \supset f =_{\Pi} g$$

The phenomenon responsible for the failure of extensionality in the above example is that it is possible to distinguish between strict and non-strict functions, ie. between functions that examine their arguments and functions that do not. Consequently this distinction is forced upon any static semantics for CTT.

Harmony Justification of harmony for CTT builds upon our justification of harmony for ITT and differs from it, excepting the absence of the strong sum type former, in just one respect: to justify analytic harmony for CTT we need to show that the contravertive decomposition clause for the inversion principle is satisfied for the three connectives.

For the Π type former this justification is straightforward and is a direct analogue of that given in chapter three; we note that the rule $\Pi - \mathrm{z}$ suffices to show that logical harmony extends unproblematically to constructive harmony.

For the $=_A$ and $\mathbb{N}$ type formers we face a problem: we restricted the zeta rules to cover only vacuous contraversions, the step required to maintain subject reduction. We must find an alternative means of satisfying contravertive decomposition in the non-vacuous cases.

Our justification in this case is somewhat unsatisfactory: we appeal to the quasi head-normal form theorem in these cases, yielding that each left-hand side is equivalent to either a canonical form or a term with unbound rejective assumptions. In the former case, analytic harmony is justified as usual by means of beta conversion. In the latter case, we simply do not maintain harmony prevails: that is harmony only prevails in the portions of the term calculus which do not have open rejective assumptions.

Ugly though this caveat is, the portions of the calculus which satisfy harmony are quite sufficient to represent all of the formal systems discussed in our section on representability. Thus the parts of the calculus that do not satisfy harmony can

be viewed in a Hilbertian light: they are not completely meaningful according to the standards of semantic clarity that we have set ourselves, but their presence is needed in our type theory according to the way we have constructed it. Given this restriction, the form of the semantics for the type formers is the same as in chapter two, though of course the classical structural rules change the relationship of hypothetical contribution to assertoric content.

Conclusions

We introduced the picture of the semantic justification of logic and arithmetic that lies behind the work of this thesis as a three-tiered hierarchy, consisting in the first place of the general theses about the nature of meaning and truth associated with an internalist approach to semantics, in the second place of Gentzen's idea that the meaning of a logical connective resides in the rules associated with it and Belnap's crucial insight about the need for harmony between the grounds for asserting a proposition and the consequences that flow from it, which together we have taken to constitute logical formalism. Lastly, we have the technical ideas originating in the work of Prawitz which show us how we can show that Belnap's criteria are satisfied for a proof theory in a natural deduction style.

My aim in these conclusions is twofold: firstly to cast critical light on the central claims of this thesis, namely to have successfully applied this model to justify both the fragment of classical logic based on the connectives of implication, conjunction (with or without the logical constants) and then the extension of this justification to a type-theoretic model of Peano Arithmetic, and secondly to make some comments about this model as an alternative to the Tarskian model theoretic semantics as a conceptual foundation for work in mathematical logic.

Let us begin by reviewing the main features of the three-tiered account. We note that, whilst the later levels depend upon the earlier levels (at least in so far as we have developed them here: alternative grounds for the applications of these ideas may of course be found), the earlier levels do not fix the later levels. So we might accept the need for what I have described as an internal approach to semantics, but seek another aspect of use to inferential role in providing an account of meaning. Similarly it would be possible to accept the account of logical formalism here, but seek to provide the justification of Belnap's criteria in terms of the sequent calculus and not natural deduction, as indeed Belnap set out to do.

We shall begin by considering the relationship between Belnap's requirement of logical harmony and its justification by Prawitz's inversion principle. Table 4.1 summarises the formal relationships.

The first two rows table 4.1 correspond to the clauses needed to justify intuitionistic propositional logic. The third row corresponds to the additional requirement needed to justify logical harmony in the case of classical propositional logic. We argued in chapter three that this new clause was needed since the distinct form of detour introduced by the classical structural rules was a contraversion appearing

Belnap's harmony requirement	Clause of Prawitz's inversion principle	Induced conversion rule
Analytic	Eliminative	β
Synthetic	Decompositional	η
Analytic	Contravertive	ζ
—	—	μ

Table 4.1: Relating harmony, the inversion principle and conversion relations

as the principal premiss of an elimination rule; consequently the new rule should be regarded as of the same kind as the introduction rules from the point of view of our division of harmony into analytic and synthetic parts.

The last row describes conversions that are not required for connectives that satisfy the strong decompositional clause (due to the consequent interderivability of the eta and beta form of that clause), but which are required to justify analytic harmony if only the weaker form is satisfied. They are not directly associated with any particular connective, and so they are explicitly a matter of justification by analytic or synthetic harmony.

What does harmony give us? We identified two motivations; firstly, and most importantly, harmony is required for us to claim coherence for our deductive practice, and in particular recognise this as an assertoric practice. Another motivation becomes important if we wish to claim that the meaning of the logical constants resides in either the introduction or elimination rules: harmony yields the criteria for the matching non–meaning–constituting rule. If we wish to derive the elimination rule from the introduction rule, then the elimination rule is the strongest rule consistent with analytic harmony. Vica versa, the criteria is the strongest rule consistent with synthetic harmony.

The threat to deductive practice posed by the failure of logical harmony lies in the possibility of non-conservativity. The division of the requirement of harmony into an analytic part (showing how a local peak can be lowered) and a synthetic part (showing how assumptions of higher type may be considered to be in the natural form of that type, and the transparency of the type to detour elimination in other types), leads to a similar division of non-conservation results into downwards and upwards non-conservativity.

Failure of analytic harmony risks downwards non-conservation, such as Prior's example in section 1.1, and in turn may directly threaten the consistency of the theory. Furthermore it seems unlikely that there could be any sense in which the semantics of such a system could be regarded as compositional, since substitution in such a system cannot be regarded as elementary.

Failure of synthetic harmony risks the more subtle threat to coherence of upwards non-conservation. We gave an example in section 2.7 inspired by quantum logic, and showed how the failure of synthetic harmony interferes with the nor-

mal elimination of detours. The satisfiability of a particular judgement involving a connective failing synthetic harmony may therefore depend upon the presence of detours not formally connected with the judgement being proved.

This threatens a kind of holism, which is usually held to be incompatible with a compositional semantics. The holism concerns satisfiability, however, but the sense in which we claim compositionality for our system is in terms of our two-factors. The holism remarked upon would be a threat to a claim of compositionality based upon truth-conditions: the move from truth-conditions to inference-conditions is therefore associated with an increase in the ease in which we can claim compositionality to hold, an important point which I have not seen discussed in the literature.

We admit a type former, $\mathbb{N}$, into our account of ITT and CTT which does not satisfy synthetic harmony. I believe that the best way to regard the underlying disorder in the semantics is to say that, while our compositional semantics *specifies* the meaning of $\mathbb{N}$, it does not *fully determine* that meaning. We say that the meaning is specified since for any purported derivation involving $\mathbb{N}$ we can say whether or not the use of the rules associated with $\mathbb{N}$ is correctly made. We say that the meaning is not fully determined since we cannot in advance circumscribe the set of possible techniques that may be necessary to prove or disprove a formula involving quantification over $\mathbb{N}$ (say): the set of detours available in a given proof theory may always be extended by adding reflection principles.

An important issue in the overall picture is that we were forced to go beyond Prawitz's formulation of the inversion principle to justify harmony. We immediately needed to add a clause dual to Prawitz's eliminative clause in order to justify synthetic harmony. In chapter two we need to extend the principle first with an equality requirement in order to claim a form of constructive harmony, and then with commuting conversion to justify synthetic harmony for indirect elimination rules. Finally we needed to appeal to the canonical form theorem to justify analytic harmony for propositional harmony in the case of propositional equality and the natural numbers. As a consequence of all of these amendments, the relationship of the inversion principle to logical and constructive harmony is not particularly tidy.

Now let us turn to an assessment of the state of this account as a possible foundational approach. What interest has been shown in this account? In the philosophical community the writings of Michael Dummett and Dag Prawitz on the semantics of language, logic and mathematics have had an impact both wide and deep. However the more formal aspects of the logical formalist approach have received, by comparison, very little attention since the mid seventies. Certainly Dummett's *Logical Basis of Metaphysics* [DummettMAE:logbm] has discussed at a deep level the significance of the proof-theoretic justification of the logical laws, but the technical aspects are not a great advance upon the early writings of Dag Prawitz. The work of Martin-Löf and his followers constitute almost all of the important new ideas in the field: however, as I argued in the introduction, the rejection of compositionality and the lack of investigation of harmony are significant flaws in this approach.

The divide between the approach I have taken in this work and the approach of

Martin-Löf should not be over-emphasised, however: as we have seen in chapter two, the key to our demonstration of global harmony depends upon the head-normal form theorem, the result that essentially constitutes the soundness result for Martin-Löf, and our further discussion of semantics is more or less built on top of this theorem.

Furthermore, there is something of a programme in the type-theoretic community to clarify and extend the expressivity of intuitionistic type theory. Important results so far of this programme are:

1. Peter Aczel's representability result, translating theorems of CZF into ITT and vica-versa. CZF, or constructive Zermelo–Frankel set theory, is an axiomatisation of a constructive cumulative hierarchy in the same language as ZF set theory, but using intuitionistic logic. Harvey Friedman has shown that in the presence of the principle of the excluded middle, the axioms of CZF are equivalent to the axioms of ZFC.

2. Anton Setzer has provided an extension of Martin-Löf's type theory with Mahlo universes [SetzerA:extmlt], a theory strong enough to give a well-foundedness proof of countable ordinals less than Γ_0.

3. Jeremy Avigad has shown how it is possible to use type theory with Mahlo Universes to provide a functional interpretation of the subsystem of second-order classical arithmetic, ATR_0 [AvigadJ:relbia, AvigadJ:reaifc].

These kinds of results are of great relevance to the thesis that the internalist semantics generally, and the three-tiered view particularly, provides a practicable alternative to the Tarskian view as a basis for mathematics. However this thesis requires a further kind of support: the theses we have wished to maintain require a more demanding demonstration of soundness of the internal semantics. Also it is important to ask how much of the above kind of result really contribute to giving a genuine rival to the cumulative hierarchy as a foundation to the Tarskian view of mathematics. The following matters are of importance, and the above thesis depends upon their resolution.

1. Peano Arithmetic is a beginning in providing an internalist semantics of mathematics, but is it enough? There seem to be three possible classes of answer to the question of 'how much is enough?'

   Constructivism We should aim only in providing an account of the mathematics traditionally regarded as constructive. This seems to be a common view amongst people interested in internalist semantics, but, pace Dummett, in my opinion it is wrong-headed: it argues that classical mathematics should be abandoned for reasons of a lack of success in providing a sound philosophical basis for it: but this rather shows that our best philosophical justifications fall short, not the incoherence of classical mathematics. Paul Bernays has made this point well:

> "In fact the current discussion about the foundations of mathematics does not have its origins in a predicament of mathematics itself. Mathematics is in a completely satisfactory state of methodological certainty. In particular, the concern caused by the paradoxes of set theory has long been overcome, ever since it was discovered that for the avoidance of the contradictions encountered, one only needs restrictions that do not encroach in the least on the claims of mathematical theories on set theory.
>
> "The problematic, the difficulties, and the differences of opinion begin rather at the point where one inquires not simply about the mathematical facts, but rather about the grounds of knowledge and the delimitation of mathematics. These questions of a philosophical nature have received a certain urgency since the transformation the methodological approach to mathematics experienced at the end of the nineteenth century." *On Hilbert's Thoughts Concerning the Grounding of Arithmetic* [MancosuP:frobh]

Justifying the cumulative hierarchy At the opposite extreme, we can regard the formal systems we need to justify as encompassing the whole of set theory. The ease with which the set-theoretic cumulative hierarchy allows mathematics to be interpreted along the lines described by Tarski reamins something of an embarrassment to the internalist programme. Naturally a satisfactory account of the cumulative hierarchy according to the three-tiered account would provide a compelling defence of internal semantics against the accusation that it artificially constrains mathematics.

Intermediate positions It is quite possible to expect that there will be irresolvable difficulties facing an attempted justification of set theory within the kind of internalist semantic framework described here, especially interpreting impredicativist mathematics within classical logic. If such a justification proves untenable , and if we reject the restriction to constructive mathematics, what options remain open to us?

Since little of mathematics makes full use of set theory, it may be acceptable to target a weaker theory, perhaps predicative mathematics.

Regardless of what position we choose to adopt, the tasks facing us remain the same: to justify the principles we do provide, and to provide adequate grounds for rejecting the mathematics we leave unjustified. What kind of results might justify the latter? This is a rather difficult issue: if any grounds at all exist for this, surely they must lie in a demonstration of incoherence based upon some kind of intolerable non-conservation results.

2. We have claimed that the price of violating harmony is a kind of formal incoherence in our deductive practice characterised by some form of non-

conservativity. For analytical harmony this implication is sufficient (ie. analytic harmony entails the subformula property, so ensuring conservativity), but we have not shown the necessity of these results: indeed the strong existence type former in the presence of the principle of the excluded middle appears to be a counter-example. For synthetic harmony we have a counter-example to sufficiency (our connective $\wedge^*$ of section 2.6), and necessity has not been shown.

Consequently the relationship between harmony and coherence is not quite as we introduced it, and some clarification of the precise relationship is in order.

3. We diagnosed that the $\mathbb{N}$ type former violates synthetic harmony, and drew a parallel between the resulting non-conservativity and the well-known extensibility of mathematical formal systems.

   We also note that only certain type formers are responsible for this extensibility: for example, the Σ type former extends the theory conservatively, in the case of ITT, but the Π type former does not.

   To better understand the issues attached to the violation of synthetic harmony, it is desirable to provide a thorough investigation of this phenomenon, perhaps by providing a number of such extensible concepts. We are particularly interested in providing precise grounds for distinguishing between violations of synthetic harmony that can be attributed to the extensibility of mathematical concepts, and ones that are symptoms of pathological behaviour (eg. such as I take quantum disjunction to be).

4. What is the right notion of extensional equality? We have presented arguments in chapter two arguing for Martin-Löf's account of intensional equality: that it provides the right basis for the characterisation of equational logic on the basis of introduction and elimination rules, and that it satisfies cut-elimination. We have also argued that it fulfils an objective of Gödel's: it provides a plausible theory of intensional equality at higher types[4] In fact this is not quite right: there is a serious problem with intensional equality that prejudices my claim for it to be the 'right' notion of equality on the internalist account: it fails to reflect the term equality rules under propositional equality in the same way that convertive equality does.

   The general desiderata may be characterised so: let us consider $\Pi - e$ for example. If we have that $\Gamma \vdash s =_{\Pi x \in A.B(x)} t$ true and $\Gamma \vdash s' =_A t'$ true, then the term reflection property is satisfied for $\Pi - e$ is $\Gamma \vdash \mathsf{ev}(s, s') =_{B(s)} \mathsf{ev}(t, t')$ true, as indeed it is. This rule can be seen to be satisfied for the Π and $=_A$ type formers. It fails, however, for the $\Sigma - \mathrm{i}$ rule, and I have not succeeded in justifying the principle for the $\mathbb{N} - \mathrm{e}$ rule.

[4]See Troelstra's introduction.

> So it is important to ask: can we provide a strengthened account of propositional equality for ITT satisfying rule-reflection and compatible with the given reductive semantics?

Difficulties at the limits of a theory need not be seen as signs to its unworkability: on the contrary they may be the spur that drives its growth and stimulates fruitful relationships with other theories. The issues touched upon by this work are many: philosophy of language, foundations of mathematics, semantics of logic and theoretical computer science. I hope this will encourage readers, regardless of their degree of sympathy with my central thesis, to see these subjects as complementary and deserving of novel approaches.

Appendix A

A summary of $\mathbf{PRA}^{\omega}_{\mu}$ and CTT

A.1 The calculus $\mathbf{PRA}^{\omega}_{\mu}$

Syntax of types terms and assumptions

The types of PRA^{ω}_{μ} are given by the following grammar.

$$
\begin{aligned}
T ::= & X \mid \mathbb{N} \mid \top \mid \bot \mid \mathsf{Eq} \\
& \mid T \Rightarrow T \mid T \times T
\end{aligned}
$$

where X is a schematic letter.

The terms are given:

$$
\begin{aligned}
s ::= & x \mid \mu\alpha^{T}.e \\
& \mid \lambda x^{T}.s \mid \mathsf{ev}(s,s) \\
& \mid \langle s,s\rangle \mid \mathsf{outl}(s) \mid \mathsf{outr}(s) \\
& \mid \underline{0} \mid \mathsf{succ}(s) \mid \mathsf{Rec}(s,s,s) \\
& \mid \mathsf{refl}(s) \mid \mathsf{eqev}(s,s) \\
e ::= & [r]s \\
r ::= & \alpha \mid \mathcal{D}_{\bot}
\end{aligned}
$$

where x, α are variable letters, and T is a type.

Variable contexts are finite sets of assertive and rejective assumptions of the form $x : A$ or $\alpha : \bullet A$ in which no variable is associated with more than one type.

Inference rules

Structural rules

$$\frac{x : A \in \Gamma}{\Gamma \vdash x : A} \qquad \frac{\alpha : \bullet A \in \Gamma}{\Gamma \vdash \alpha : \bullet A}$$

$$\frac{\Gamma \vdash r : \bullet A \quad \Gamma \vdash s : A}{\Gamma \vdash [r]s \ \mathsf{empty}} \qquad \frac{\Gamma, \alpha : \bullet A \vdash e \ \mathsf{empty}}{\Gamma \vdash \mu\alpha^A.e : A}$$

Logical rules

$$\frac{\Gamma \vdash s : A \quad \Gamma \vdash t : B}{\Gamma \vdash \langle s, t\rangle : A \times B}$$

$$\frac{\Gamma \vdash s : A \times B}{\Gamma \vdash \mathsf{outl}(s) : A} \qquad \frac{\Gamma \vdash s : A \times B}{\Gamma \vdash \mathsf{outr}(s) : B}$$

$$\frac{\Gamma, x : A \vdash s : B}{\Gamma \vdash \lambda x^A.s : A \Rightarrow B} \qquad \frac{\Gamma \vdash s : A \Rightarrow B \quad \Gamma \vdash t : A}{\Gamma \vdash \mathsf{ev}(s, t) : B}$$

Then the rules for the two propositional constants.

$$\frac{}{\Gamma \vdash * : \top} \qquad \frac{}{\Gamma \vdash \mathcal{D}_{\perp} : \bullet\perp}$$

Arithmetic rules

$$\frac{}{\Gamma \vdash \mathsf{zero} : \mathbb{N}} \qquad \frac{\Gamma \vdash n : \mathbb{N}}{\Gamma \vdash \mathsf{succ}(n) : \mathbb{N}}$$

$$\frac{\Gamma \vdash s : \mathbb{N} \quad \Gamma \vdash a : A \quad \Gamma \vdash f : \mathbb{N} \Rightarrow A \Rightarrow A}{\mathsf{Rec}(s, a, f) : A}$$

$$\frac{\Gamma \vdash s : A}{\Gamma \vdash \mathsf{refl}(s) : \mathsf{Eq}\ A} \qquad \frac{\Gamma \vdash s : \mathsf{Eq}\ A \quad \Gamma \vdash t : A \supset B}{\Gamma \vdash \mathsf{eqev}(s, t) : B}$$

Conversion rules

We give the elementary conversion steps by themselves first, and describe the derived reduction relations giveing the structural, logical and arithmetic conversions.

Structural conversions There are two structural conversions, called mu-type conversions in $\lambda\mu$, a category of conversion that does not exist in λ.

$$[r]\mu\alpha^A.e \to^c_{\mu} e[\alpha := r]$$
$$\mu\alpha^A.[\alpha]s \to^c_{\mu} \quad \text{if } x \notin \mathsf{FV}(s)$$

Logical conversions There are reductions of beta, eta- and zeta- type governed by $\times$ and $\Rightarrow$. The notation $\mathsf{ev}(f, a, b)$ is short for $\mathsf{ev}(\mathsf{ev}(f, a), b)$.

1. Beta conversions.

$$\begin{aligned}
\mathsf{outl}(\langle s, t\rangle) &\to^c_\beta s \\
\mathsf{outr}(\langle s, t\rangle) &\to^c_\beta t \\
\mathsf{ev}(\lambda x^A.s, t) &\to^c_\beta s[x := t]
\end{aligned}$$

2. Eta conversions. These conversions are only permitted if they satisfy subject reduction.

$$\begin{aligned}
s &\to^c_\eta \langle \mathsf{outl}(s), \mathsf{outr}(s)\rangle \\
s &\to^c_\eta \lambda x^A.\mathsf{ev}(s, x)
\end{aligned}$$

3. Zeta conversions.

$$\begin{aligned}
\mu\alpha^{A\wedge B}.e &\to^c_\zeta \langle \mu\beta_0^A.e[\alpha := \mathsf{outl}(-)/\beta_0], \mu\beta_1^B.e[\alpha := \mathsf{outr}(-)/\beta_1]\rangle \\
\mu\alpha^{A\supset B}.e &\to^c_\zeta \lambda x^A.\mu\beta^B.e[\alpha := \mathsf{ev}(-, x)/\beta]
\end{aligned}$$

Arithmetic conversions There are reductions of beta- and zeta- type governed by the types $\mathbb{N}$ and Eq in PRA^{ω}_{μ}. The reduction rules come in two flavours, reflecting the left- and right- recursive forms of the $\mathbb{N}$ beta conversions:

1. Left-recursive form.

   (a) Beta conversions.

$$\begin{aligned}
\mathsf{Rec}(\underline{0}, a, f) &\to^c_\beta a \\
\mathsf{eqev}(\mathsf{refl}(s), t) &\to^c_\beta \mathsf{ev}(t, s) \\
\mathsf{Rec}(\mathsf{succ}(s), a, f) &\to^c_\beta \mathsf{Rec}(s, \mathsf{ev}(f, \underline{0}, a), \lambda n^{\mathbb{N}}.\mathsf{ev}(f, \mathsf{succ}(n)))
\end{aligned}$$

   (b) Zeta conversions.

$$\begin{aligned}
\mathsf{Rec}(\mu\alpha^{\mathbb{N}}.e, a, f) &\to^c_\zeta \mu\beta^T.e[\alpha := \mathsf{Rec}(-, a, f)/\beta] \\
\mathsf{succ}(\mu\alpha^{\mathbb{N}}.e) &\to^c_\zeta \mu\beta^{\mathbb{N}}.e[\alpha := \mathsf{succ}(-)/\beta] \\
\mathsf{eqev}(\mu\alpha^A.e, t) &\to^c_\zeta \mu\beta^B.e[\alpha := \mathsf{eqev}(-, t)/\beta]
\end{aligned}$$

2. Right-recursive form.

(a) Beta conversions.

$$\begin{aligned}
\mathsf{Rec}(\underline{0}, a, f) &\to^{c}_{\beta} a \\
\mathsf{eqev}(\mathsf{refl}(s), t) &\to^{c}_{\beta} \mathsf{ev}(t, s) \\
\mathsf{Rec}(\mathsf{succ}(s), a, f) &\to^{c}_{\beta} \mathsf{ev}(f, s, \mathsf{Rec}(s, a, f))
\end{aligned}$$

(b) Zeta conversions.

$$\begin{aligned}
\mathsf{Rec}(\mu\alpha^{\mathbb{N}}.e, a, f) &\to^{c}_{\zeta} \mu\beta^{T}.e[\alpha := \mathsf{Rec}(-, a, f)/\beta] \\
\mathsf{eqev}(\mu\alpha^{A}.e, t) &\to^{c}_{\zeta} \mu\beta^{B}.e[\alpha := \mathsf{eqev}(-, t)/\beta]
\end{aligned}$$

Other reduction relations We obtain the one-step reduction relation for beta-, zeta- and mu- forms as the compatible closure on the redex rewrite relation. For eta-form reduction, it is obtained only for certain contexts corresponding to *eta minimal formula occurrences*; see the definition in section 3.3.

We also define appeal to the transitive closure of these rewrites and the symmetric transitive closure is regarded as a form of denotational equality on terms.

A.2 The calculus CTT

Structural rules

The structural rules of CTT are organised into six classes.

1. Formation of telescopes;
2. Introduction of assumptions;
3. Convertive equality is an equivalence;
4. Convertive compatibility;
5. Alpha conversion and type congruence;
6. Classical structural rules.

Class 1. Formation of telescopes

$$\frac{}{\cdot\ \mathsf{tel}}\,\mathrm{tel-emp} \qquad \frac{}{\cdot = \cdot\ \mathsf{tel}}\,\mathrm{tel-emp-eq}$$

$$\frac{\Gamma\ \mathsf{tel} \quad X \notin \mathsf{dom}(\Gamma)}{\Gamma, X\ \mathsf{tel}}\,\mathrm{tel-ty} \qquad \frac{\Gamma = \Gamma'\ \mathsf{tel} \quad X \notin \mathsf{dom}(\Gamma)}{\Gamma, X\ \mathsf{type\ tel} = \Gamma', X\ \mathsf{type\ tel}}\,\mathrm{tel-ty-eq}$$

$$\frac{\Gamma \vdash A\ \mathsf{type} \quad x \notin \mathsf{dom}(\Gamma)}{\Gamma, x : A\ \mathsf{tel}}\,\mathrm{tel-ass}$$

$$\frac{\Gamma = \Gamma' \ \mathsf{tel} \quad \Gamma \vdash A = A' \ \mathsf{type} \quad x \notin \mathsf{dom}(\Gamma)}{\Gamma, x : A = \Gamma', x : A' \ \mathsf{tel}} \ \mathrm{tel-ass-eq}$$

$$\frac{\Gamma \vdash A \ \mathsf{type} \quad \alpha \notin \mathsf{dom}(\Gamma)}{\Gamma, \alpha : \bullet A \ \mathsf{tel}} \ \mathrm{tel-den}$$

$$\frac{\Gamma = \Gamma' \ \mathsf{tel} \quad \Gamma \vdash A = A' \ \mathsf{type} \quad \alpha \notin \mathsf{dom}(\Gamma)}{\Gamma, \alpha : \bullet A = \Gamma', \alpha : \bullet A' \ \mathsf{tel}} \ \mathrm{tel-den-eq}$$

Class 2. Introduction of assumptions

$$\frac{\Gamma, X \ \mathsf{type}, \Gamma' \ \mathsf{tel}}{\Gamma, X \ \mathsf{type}, \Gamma' \vdash X \ \mathsf{type}} \ \mathrm{hyp-ty}$$

$$\frac{\Gamma, x : A, \Gamma' \ \mathsf{tel}}{\Gamma, x : A, \Gamma' \vdash x : A} \ \mathrm{hyp-ass} \qquad \frac{\Gamma, \alpha : \bullet A, \Gamma' \ \mathsf{tel}}{\Gamma, \alpha : \bullet A, \Gamma' \vdash \alpha : \bullet A} \ \mathrm{hyp-den}$$

Class 3. Convertive equality is an equivalence relation

$$\frac{\Gamma \vdash A \ \mathsf{type}}{\Gamma \vdash A = A \ \mathsf{type}} \ \mathrm{eq-ty-r} \qquad \frac{\Gamma \vdash A = B \ \mathsf{type} \quad \Gamma \vdash A = C \ \mathsf{type}}{\Gamma \vdash B = C \ \mathsf{type}} \ \mathrm{eq-ty-st}$$

$$\frac{\Gamma \vdash s : A}{\Gamma \vdash s = s : A} \ \mathrm{eq-tm-r} \qquad \frac{\Gamma \vdash s = t : A \quad \Gamma \vdash s = u : A}{\Gamma \vdash t = u : A} \ \mathrm{eq-tm-st}$$

$$\frac{\Gamma \vdash r : \bullet A}{\Gamma \vdash r = r : \bullet A} \ \mathrm{eq-rej-r} \qquad \frac{\Gamma \vdash r = r' : \bullet A \quad \Gamma \vdash s = s' : A}{\Gamma \vdash [r]s = [r']s' \ \mathsf{empty}} \ \mathrm{eq-emp}$$

Class 4. Convertive compatibility

$$\frac{\Gamma \vdash s : A \quad \Gamma \vdash A = A' \ \mathsf{type}}{\Gamma \vdash s : A'} \ \mathrm{ty-conv}$$

$$\frac{\Gamma \vdash s = s' : A \quad \Gamma, x : A \vdash B \ \mathsf{type} \quad (*)}{\Gamma \vdash B[x := s] = B[x := s'] \ \mathsf{type}} \ \mathrm{tm-ty-conv}$$

$$\frac{\Gamma \vdash s = s' : A \quad \Gamma, x : A \vdash t : B \quad (*)}{\Gamma \vdash t[x := s] = t[x := s'] : B[x := s]} \ \mathrm{tm-tm-conv}$$

The premiss $(*)$ displayed for the $\mathrm{tm-ty-conv}$ and $\mathrm{tm-tm-conv}$ rules indicates that there is a positivity requirement for these rules. For $\mathrm{tm-ty-conv}$ we insist that the substituted terms are positive, and for $\mathrm{tm-tm-conv}$ we insist that *either* the substituted terms are positive, *or* that they are not substituted into any type, including types attached to bound variables.

Class 5. Alpha conversion and type congruence

$$\frac{\Gamma \vdash A = A' \ \mathsf{type} \quad \Gamma, x : A \vdash s : B = B' \quad y \notin \mathsf{dom}(\Gamma)}{\Gamma \vdash \Pi x^A.B = \Pi y^{A'}.B'[x := y] \ \mathsf{type}} \ \mathrm{alpha-ty-\Pi}$$

$$\frac{\Gamma \vdash A = A' \ \mathsf{type} \quad \Gamma, x : A \vdash B \ \mathsf{type} \quad y \notin \mathsf{dom}(\Gamma)}{\Gamma \vdash \lambda x^A.s = \lambda y^{A'}.s[x := y] : \Pi x^A.B} \ \mathrm{alpha-\Pi}$$

$$\frac{\Gamma, i : \mathbb{N} \vdash C(i) = C'(i) \ \mathsf{type} \quad \Gamma \vdash \mathrm{R}^{\mathbb{N}}(s, a, (i^{\mathbb{N}}, z^{C(i)})f) : C(s) \quad j, w \notin \mathsf{dom}(\Gamma)}{\Gamma \vdash \mathrm{R}^{\mathbb{N}}(s, a, (i^{\mathbb{N}}, z^{C}(i))f) = \mathrm{R}^{\mathbb{N}}(s, a, (j^{\mathbb{N}}, w^{C'(j)})f[i := j, z := w]) : C(s)} \ \mathrm{alpha-\mathbb{N}}$$

$$\frac{\Gamma \vdash A = A' \ \mathsf{type} \quad \Gamma \vdash \mathrm{R}^{=}(p, (a^A)f) : C \quad x \notin \mathsf{dom}(\Gamma)}{\Gamma \vdash \mathrm{R}^{=}(p, (a^A)f) = \mathrm{R}^{=}(p, (x^{A'})f) : C} \ \mathrm{alpha-=}$$

$$\frac{\Gamma \vdash A = A' \ \mathsf{type} \quad \Gamma, \alpha : \bullet A \vdash e \ \mathsf{empty} \quad \beta \notin \mathsf{dom}(\Gamma)}{\Gamma \vdash \mu\alpha^A.e = \mu\beta^{A'}.e[\alpha := -/\beta] : A} \ \mathrm{alpha-mu}$$

Class 6. Classical structural rules

$$\frac{\Gamma \vdash r : \bullet A \quad \Gamma \vdash s : A}{\Gamma \vdash [r]s \ \mathsf{empty}} \ \mathrm{mu-anti} \qquad \frac{\Gamma, \alpha : \bullet A \vdash e \ \mathsf{empty}}{\Gamma \vdash \mu\alpha^A.e : A} \ \mathrm{mu-abs}$$

$$\frac{\Gamma \vdash r : \bullet A \quad \Gamma, \alpha : \bullet A \vdash e = e' \ \mathsf{empty}}{\Gamma \vdash [r]\mu\alpha^A.e = e'[\alpha := -^A/r] \ \mathsf{empty}} \ \mathrm{mu-beta} \qquad \frac{\Gamma, \alpha : \bullet A \vdash s : A \quad \alpha \notin \mathsf{FV}(s)}{\Gamma \vdash \mu\alpha^A.[\alpha]s = s : A} \ \mathrm{mu-eta}$$

Type formers

The calculus CTT lacks a type former possessed by ITT, namely the Σ type former.

Π type

$$\frac{\Gamma \vdash A \ \mathsf{type} \quad \Gamma, x : A \vdash B \ \mathsf{type}}{\Gamma \vdash \Pi x : A.B \ \mathsf{type}} \ \Pi-\mathrm{f}$$

$$\frac{\Gamma, x : A \vdash s : B}{\Gamma \vdash \lambda x : A.s : \Pi x : A.B} \ \Pi-\mathrm{i}$$

$$\frac{\Gamma \vdash f : \Pi x : A.B \quad \Gamma \vdash t : A}{\Gamma \vdash \mathsf{ev}(f, t) : B[x := t]} \ \Pi-\mathrm{e}$$

$$\frac{\Gamma, x : A \vdash s : B \quad \Gamma \vdash t : A}{\Gamma \vdash \mathsf{ev}(\lambda x : A.s, t) = s[x := t] : B[x := t]} \ \Pi-\mathrm{c}$$

$$\frac{\Gamma \vdash s : \Pi x : A.B \quad y \notin \mathsf{dom}(\Gamma)}{\Gamma \vdash s = \lambda y : A.\mathsf{ev}(s, y) : \Pi x : A.B} \ \Pi-\mathrm{c-eta}$$

$$\frac{\Gamma, \alpha : \bullet\Pi x : A.B \vdash e \ \mathsf{empty} \quad y \notin \mathsf{dom}(\Gamma)}{\Gamma \vdash \mu\alpha^A.e = \lambda y : A.e[\alpha := \mathsf{ev}(-, y)/\beta] : \Pi x : A.B}\ \Pi - \mathrm{z}$$

$\mathbb{N}$ **type**

$$\frac{\Gamma \ \mathsf{tel}}{\Gamma \vdash \mathbb{N}\ \mathsf{type}}\ \mathbb{N} - \mathrm{f}$$

$$\frac{\Gamma \ \mathsf{tel}}{\Gamma \vdash \underline{0} : \mathbb{N}}\ \mathbb{N} - \mathrm{i} - \underline{0} \qquad \frac{\Gamma \vdash n : \mathbb{N}}{\Gamma \vdash \mathsf{succ}(n) : \mathbb{N}}\ \mathbb{N} - \mathrm{i} - \mathsf{succ}$$

$$\frac{\Gamma \vdash s : \mathbb{N} \quad \Gamma \vdash C(n : \mathbb{N})\ \mathsf{type} \quad \Gamma \vdash a : C(\mathsf{zero}) \quad \Gamma, i : \mathbb{N}, z : C(i) \vdash t : C(\mathsf{succ}(i))}{\mathrm{R}^{\mathbb{N}}(s, a, (i^{\mathbb{N}}, z^{C(i)})t) : C(s)}\ \mathbb{N} - \mathrm{e}$$

$$\frac{\Gamma, n : \mathbb{N} \vdash r(n) \equiv \mathrm{R}^{\mathbb{N}}(s, a, (i, z)t) : C(n)}{\Gamma \vdash r(\underline{0}) = a : C(\underline{0})}\ \mathbb{N} - \mathrm{c} - \underline{0}$$

$$\frac{\Gamma \vdash s : \mathbb{N} \quad \Gamma, n : \mathbb{N} \vdash \mathrm{R}^{\mathbb{N}}(s, a, (i^{\mathbb{N}}, z^{C(i)})t) : C(n)}{\mathrm{R}^{\mathbb{N}}(\mathsf{succ}(s), a, (i, z)t) = \mathrm{R}^{\mathbb{N}}(s, a', (j^{\mathbb{N}}, z^{C(j)})t') : C(\mathsf{succ}(s))}\ \mathbb{N} - \mathrm{c} - \mathsf{succ}$$

$$\text{where } a' \triangleq t[i := \underline{0}, z := a]$$
$$t' \triangleq t[i := \mathsf{succ}(j)]$$

$$\frac{\Gamma, \alpha : \bullet\mathbb{N} \vdash e \ \mathsf{empty} \quad \Gamma, n : \mathbb{N} \vdash r(n) \equiv \mathrm{R}^{\mathbb{N}}(s, a, (i, z)t) : C(n) \quad \beta \notin \mathsf{dom}(\Gamma)}{\Gamma \vdash r(\mu\alpha^{\mathbb{N}}.e) = \mu\beta^{C(\mu\alpha.e)}.e[\alpha := r(-)/\beta] : C(\mu\alpha.e)}\ \mathbb{N} - \mathrm{z} - \mathrm{e}$$

$$\frac{\Gamma, \alpha : \bullet\mathbb{N} \vdash e \ \mathsf{empty}}{\Gamma \vdash \mathsf{succ}(\mu\alpha^{\mathbb{N}}.e) = \mu\beta^{\mathbb{N}}.e[\alpha := \mathsf{succ}(-)/\beta] : \mathbb{N}}\ \mathbb{N} - \mathrm{z} - \mathsf{succ}$$

$=_A$ **types**

$$\frac{\Gamma \vdash s : A \quad \Gamma \vdash t : A \quad s, t\ \mathsf{positive}}{\Gamma \vdash s =_A t\ \mathsf{type}} = -\mathrm{f}$$

$$\frac{\Gamma \vdash s = t : A \quad s, t\ \mathsf{positive}}{\Gamma \vdash \mathsf{refl}(s) : s =_A t} = -\mathrm{i}$$

$$\frac{\Gamma \vdash p : s =_A t \quad \Gamma \vdash C(x : A, y : A, z : x =_A y)\ \mathsf{type} \quad \Gamma, x : A \vdash d : C(x, x, \mathsf{refl}(x))}{\mathrm{R}^{=}(p, (a : A)d) : C(s, t, p)} = -\mathrm{e}$$

$$\frac{\Gamma \vdash s =_A t\ \mathsf{type} \quad \Gamma, x : A \vdash \mathrm{R}^{=}(x, (a)d) : C(x, x, \mathsf{refl}(x))}{\Gamma \vdash r(\mathsf{refl}(s)) = d[a := s] : C(s, t, \mathsf{refl}(s))} = -\mathrm{c}$$

$$\frac{\Gamma, \alpha : \bullet s =_A t \vdash e \ \mathsf{empty} \quad \Gamma, x : A \vdash r(x) \equiv \mathrm{R}^{=}(x, (a)d) : C(x, x, \mathsf{refl}(x))}{\Gamma \vdash r(\mu\alpha.e) = \mu\beta.e[\alpha := r(-)/\beta] : C(s, t, \mathsf{refl}(s))} = -\mathrm{z}$$

Propositional constants

Finally there are two propositional constants $\bot$, $\top$. These rules only have type former and introduction rules attached to them.

$\top$ **type constant**

$$\frac{\Gamma\ \mathsf{tel}}{\Gamma \vdash \top\ \mathsf{type}}\ \mathrm{const} - \top - \mathrm{f} \qquad \frac{\Gamma\ \mathsf{tel}}{\Gamma \vdash * : \top}\ \mathrm{const} - \top - \mathrm{i}$$

$\bot$ **type constant**

$$\frac{\Gamma\ \mathsf{tel}}{\Gamma \vdash \bot\ \mathsf{type}}\ \mathrm{const} - \bot - \mathrm{f} \qquad \frac{\Gamma\ \mathsf{tel}}{\Gamma \vdash \mathcal{A}_{\bot} : \bullet\bot}\ \mathrm{const} - \bot - \mathrm{i}$$

Bibliography

LNCS and LNM refer to the series *Lecture Notes in Computer Science* and *Lecture Notes in Mathematics* respectively. Both are published by Springer-Verlag.

[Acz80] P. Aczel. Frege structures and the notions of proposition, truth and set. In J. Barwise, editor, *The Kleene symposium*, pages 31–59, North-Holland, 1980.

[Acz99] P. Aczel. The Russell–Prawitz modality. Unpublished.

[Afr92] H. Africk. Classical logic, Intuitionistic logic, and the Peirce rule. *Notre Dame Journal of Formal Logic*, 33(2):191–201, 1992.

[AO93] S. Abramsky and C.-H. L. Ong. Full abstraction in the lazy lambda calculus. *Information and Computation*, 105:159–269, 1993.

[Ari84] Aristotle. *Complete Works: Revised Oxford Translation*. Princeton University Press, 1984. Translated by Jonathan Barnes, 2 volumes.

[Ari89] Aristotle. *Prior Analytics*. Hackett, Indiana, 1989. Translation and commentary by Robin Smith.

[Aug] Saint Augustine. *Confessions*. Translation and commentary by E. B. Pusey. New York, Dutton, 1950.

[Aus62] J. L. Austin. *How to do Things with Words*. Oxford University Press, 1962.

[Avi96] J. Avigad. On the relationship between $\mathbf{ID}_{<\omega}$ and $\mathbf{ATR}_0$. *Journal of Symbolic Logic*, 61:768–779, 1996.

[Avi98] J. Avigad. A realizability interpretation for classical arithmetic. Submitted for publication, 1998.

[Bar84] H. Barendregt. *The Lambda Calculus: its syntax and semantics*. Studies in Logic and the foundations of mathematics, North-Holland, 1984. Second edition.

[Bar92] H. Barendregt. Lambda Calculi with Types. In S. Abramsky, D. M. Gabbay and T. S. E. Maibaum, editors, *Handbook of Logic in Computer Science*, volume II, pages 117–309. Oxford University Press, 1992.

[Bar77] J. Barwise, editor. *The Handbook of Mathematical Logic*. North-Holland, 1977.

[BB94] F. Barbanera and S. Berardi. A symmetric lambda calculus for 'classical' program extraction. *Theoretical Aspects of Computer Software*, LNCS 789:494–515, 1994.

[BBC95] S. Berardi, M. Bezem, and T. Coquand. On the computational content of the axiom of choice. *Typed Lambda Calculus and Applications*, LNCS 902, 1995.

[BE87] J. Barwise and J. Etchemendy. *The Liar*. Oxford University Press, 1987.

[BHS97] G. Barthe, J. Hatcliff, and M. H. Sorensen. A Notion of Classical Pure Type Systems. In S. Brookes, M. Main, A. Melton, and M. Mislove, editors, *Mathematical foundations of programming semantics* Volume 6 of Electronic Notes in Computer Science. Elsevier, 1997.

[Bee85] Michael Beeson. *Foundations of Constructive Mathematics*. Springer-Verlag, 1985.

[Bel62] Nuel D. Belnap. Tonk, plonk and plink. *Analysis*, 22:130–134, 1962. Reprinted in *Philosophical logic* [StrawsonPF:phil].

[BP83] P. Benacerraf and H. Putnam. *Philosophy of Mathematics*. Cambridge University Press, 1983.

[Bis67] E. Bishop. *Foundations of Constructive Analysis* McGraw-Hill, New York, 1967. A substantially revised second edition has been published under the title *Constructive Analysis*, Springer-Verlag, 1985.

[Bus95] S. R. Buss. On Herbrand's theorem. *Logic and Computational Complexity*, LNCS 960:195–209, 1995.

[Bus98] S. R. Buss, editor. *The Handbook of Proof Theory*. North-Holland, 1998.

[BW88] R. Bird and P. Wadler. *Introduction to Functional Programming*. Prentice Hall, 1988.

[CF58] H. B. Curry and R. Feys. *Combinatory Logic*. North Holland, Amsterdam, 1958. 2 volumes.

[Con86] R. L. Constable (chief author). *Implementing mathematics with the Nuprl proof development system*. Prentice-Hall, New Jersey, 1986.

[Coq90] T. Coquand Metamathematical investigations of a theory of constructions. In P. Oddifreddi, editor, *Logic in Computer Science*. Academic Press, New York, 1990.

[Coq93] T. Coquand. A semantics of evidence for classical arithmetic. *Journal of Symbolic Logic*, 60:325–337, 1993.

[Coq95] T. Coquand. Computational content of classical logic. Lecture notes from the 1995 CLiCS-II Summer School.

[CSH80] H. B. Curry, J. P. Seldin, and J. R. Hindley. *To H.B. Curry: essays on combinatory logic, lambda calculus and formalism*. Academic Press, London, 1980.

[Cur77] H. B. Curry. *Foundations of mathematical logic*. Constable, London, 1977. Originally published by Dover Publications in New York in 1977 and by McGraw-Hill in London.

[Cur91] P.-L. Curien. An abstract framework for environment machines. *Theoretical Computer Science*, 82:389–402, 1991.

[dB68] N. G. de Bruijn. The mathematical language AUTOMATH, its usage, and some of its extensions. *Symposium on Automatic Deduction*, LNM 125:29–61, 1968.

[dB95] N. G. de Bruijn. On the roles of types in mathematics. In *The Curry–Howard isomorphism* [deGrooteP:curhi], pages 27–54.

[dG92] P. de Groote. Denotations for classical proofs. *Logical Foundations of Computer Science*, LNCS 620:105–116, 1992.

[dG94a] P. de Groote. A CPS-translation of the $\lambda\mu$-calculus. *Colloquium on Trees in Algebra (CAAP '94)*, LNCS 787:85–99, 1994.

[dG94b] P. de Groote. On the relation between the $\lambda\mu$-calculus and the syntactic theory of sequential control. *Logic Programming and Automated Reasoning (LPAR '94)*, LNCS 822:31–43, 1994.

[dG95] P. de Groote. *The Curry–Howard isomorphism*, volume 8 of *Cahiers du Centre de Logique*. Academia, Louvaine-la-Neuve, 1995.

[DJ90] N. Dershowitz and J.-P. Jouannaud. Rewrite systems. In J. van Leeuwen, editor, *Handbook of Theoretical Computer Science*, volume B, pages 243–309, North-Holland, 1990.

[dQ88] R. J. G. B. de Queiroz. A proof-theoretic account of programming and the role of reduction rules. *Dialectica*, 42(4):265–282, 1988.

[dQM88] R. J. G. B. de Queiroz and T. S. E. Maibaum. Proof theory and computer programming. *Zeitschrift für Mathematische Logik und Grundlagen der Mathematik*, 36(5):389–414, 1990.

[Dra80] A. G. Dragalin, New forms of realizability and Markov's rule. First published in Russian in *Dokl. Akad. Nauk SSSR*, 251:534–537, 1980. Translation appears in *Sov. Math. Dokl.*, 21:461–464.

[Dum58] M. A. E. Dummett. Truth. *Proc. Aristotelian Society*, 59:141–162, 1958. Reprinted in *Philosophical logic* [StrawsonPF:phil], and in *Truth and other enigmas* [DummettMAE:truoe].

[Dum73] M. A. E. Dummett. *Frege: Philosophy of Language*. Duckworth, London, 1973.

[Dum75] M. A. E. Dummett. The philosophical basis of intuitionistic logic. In H. E. Rose and J. Sheperdson, editors, *Logic Colloquium '73*. North-Holland, 1975. Reprinted in *Truth and other enigmas* [DummettMAE:truoe].

[Dum76] M. A. E. Dummett. What is a theory of meaning II. In Gareth Evans and John McDowell, editors, *Truth and Meaning*. Oxford University Press, 1976. Reprinted in 'The Seas of Language' [DummettMAE:seal].

[Dum77] M. A. E. Dummett. *Elements of Intuitionism*. Oxford University Press, 1977.

[Dum78] M. A. E. Dummett. *Truth and other enigmas*. Duckworth, London, 1978.

[Dum90] M. A. E. Dummett. The source of the concept of truth. In G. Boolos, editor, *Meaning and Method: Essays in Honor of Hilary Putnam*. Cambridge, 1990. Reprinted in *The Seas of Language* [DummettMAE:seal].

[Dum91] M. A. E. Dummett. *The Logical Basis of Metaphysics*. Duckworth, London, 1991.

[Dum93] M. A. E. Dummett. *The Seas of Language*. Oxford University Press, 1993.

[Etc88] J. Etchemendy. Tarski on truth and logical consequence. *Journal of Symbolic Logic*, 53:51–79, 1988.

[Euc] Euclid. *The thirteen books of Euclid's Elements* Translation and commentary by Sir T. L. Heath. Dover, New York, 1956.

[FA98] S. Feferman and J. Avigad. Gödel's functional ("Dialectica") interpretation. In *The Handbook of Proof Theory* [BussSR:hanpt], pages 337–405, 1998.

[Fef77] S. Feferman. Theories of finite type related to mathematical practice. In *The Handbook of Mathematical Logic* [BarwiseJ:hanml], pages 913–971, 1977.

[Fef81] S. Feferman. A theory of variable types. In *Proc. fifth Latin American symposium on mathematical logic*, Revista Colombiana de Matematicas, 19:95–105, 1985.

[Fef88] S. Feferman. Hilbert's program relativized: Proof-theoretical and foundational reductions. In 'A symposium on Hilbert's program', *Journal of Symbolic Logic*, 53:364–383, 1988.

[Fef93] S. Feferman. Gödel's Dialectica interpretation and its two-way stretch. *Computational logic and proof theory*, LNCS 713:23–40, 1993.

[Fel88] M. Felleisen. The theory and practice of first-class prompts. In *Proc. 15th* ACM *Symposium on Principles of Programming Languages*, pages 180–190, ACM Press, 1988.

[Fen71] J. E. Fenstad, editor. *Proc. Second Scandinavian Logic Symposium*, volume 63 of *Studies in Logic and the Foundations of Mathematics*. North-Holland, 1971.

[FFKD86] M. Felleisen, D. P. Friedman, E. Kohlbecker, and Bruce Duba. Reasoning with continuations. In *Proc. Symposium on Logic in Computer Science*, pages 131–141. IEEE Computer Society Press, 1986.

[FH89] M. Felleisen and R. Hieb, *The revised report on the syntactic theories of sequential control and state*, Rice University technical report, Rice COMP TR 89–100, 1989.

[Fra87] Torkél Franzen. *Provability and Truth*, volume 9 of *Stockholm Studies in Philosophy*. Almqvist and Wiksell, 1987.

[Fre79] G. Frege. *Begriffschrift, eine der Arithmetischen Nachgebildete Formalsprache des Reinen Denkens*. In van Heijenoort [vanHeijenoortJ:frofg], 1879. Translation and commentary in *From Frege to Gödel* [vanHeijenoortJ:frofg], pages 1–82.

[Fre84] G. Frege. *The Foundations of Arithmetic*. Originally published as *Die Grundlagen der Arithmetik, eine logisch mathemtische Untersuchung über den Begriff der Zahl* in 1884. Translated by J. L. Austin. Blackwell, Oxford, 1950.

[Fre92] G. Frege. On *Sinn* and *Bedutung*. Originally published in 1892. Translation appeared in *Translations from the Philosophical Writings of Gottlob Frege*, edited P. Geach and M. Black. Blackwell, Oxford, 1952.

[Fre03] G. Frege. *The Basic Laws of Arithmetic*. Originally *Grundgesetze der Arithmetik, begriffsschriftlich abgeleitet* in two volumes, published 1893 and 1903. Translated by M. Furth, California, 1964.

[Fre23] G. Frege. Gedankengefüge. Unpublished manuscript. Appeared in his collected works.

[Fri78] Harvey Friedman. Classically and intuitionistically provable functions. *Higher Set Theory*, LNM 669:21–28, 1978.

[Gan80] R. O. Gandy. Proof of strong normalisation. In J. R. Hindley and J. P. Seldin, editors, *To H. B. Curry: Essays in Combinatory Logic, Lambda Calculus and Formalism*. Academic Press, 1980.

[Gan93] R. O. Gandy. Dialogues, Blass games and sequentiality for objects of finite type. Unpublished manuscript, 1993.

[Gen35] G. Gentzen. Investigations into logical deduction. First published as 'Untersuchungen über das logische schliessen' in *Mathematische Zeitschrift*, 39:176–210, 1935. Translation published in *Collected Papers* [GentzenG:colpgg], 1969.

[Gen38] G. Gentzen. New version of the consistency proof for elementary number theory. First published as 'Neue Fassung des Widerspruchsfreiheit für die reine Zahlentheorie' in *Forschungen zur Logik und zur Grundlegung der exakten Wissenschaften*, 4:19–44, 1938. Translation published in *Collected Papers* [GentzenG:colpgg], 1969.

[Gha95] Neil Ghani. Eta-equality for coproducts. *Typed Lambda Calculus and Applications (TLCA 95)*, LNCS 902:171–185, 1995.

[Gha97] Neil Ghani. Eta-expansions in dependent type theory — the calculus of constructions. *Typed Lambda Calculus and Applications (TLCA 97)*, LNCS 1210:164–180, 1997.

[Gir86] J.-Y. Girard. The system F of variable types, fifteen years later. *Theoretical Computer Science*, 45:159–192, 1986.

[Gir87a] J.-Y. Girard. Linear Logic. *Theoretical Computer Science*, 50(1):1–102, 1987.

[Gir87b] J.-Y. Girard. *Proof Theory and Logical Complexity*, volume 1. Bibliopolis, 1987.

[Gir91] J.-Y. Girard. A new constructive logic: Classical logic. *Mathematical Structures in Computer Science*, 1(3):255–296, 1991.

[Gir92] J.-Y. Girard. Light linear logic. unpublished, 1992.

[GL87] J.-Y. Girard and Yves Lafont. Linear Logic and lazy computation. *Conference on Functional and Logic Programming (CFLP 87)*, LNCS 250, 1987.

[GLT89] J.-Y. Girard, Yves Lafont, and Paul Taylor. *Proofs and Types*, volume 7 of *Cambridge Tracts in Theoretical Computer Science*. Cambridge University Press, 1989.

[Goe58] K. Goedel On a hitherto unutilised extension of the finitary standpoint. First published as 'Über eine bischer noch nicht benütze Erweiterung des finiten Standpunktes' in *Dialectica*, 12:280–287, 1958. Translation and commentary by A. S. Troelstra, published in *Collected Works* [GoedelK:colw], 1990. An earlier translation of W. Hodges and B. Watson is published together with a bibliography of subsequent work compiled by J. R. Hindley in *Journal of Philosophical Logic*, 9(2):133–142, 1980.

[Goe86] Kurt Gödel. *Collected Works*. Oxford University Press, New York, 1986. Edited by S. Feferman, J.W. Dawson, Jr., S. C. Kleene, G. H. Moore, R. M. Solovay, and J. van Heijenoort. Volumes I and II (1986 and 1990) cover the work published in his lifetime, and volume III (1995) is the first volume of his Nachlass.

[Gog94] H. Goguen. *A Type Operational Semantics for Type Theory*. PhD thesis, LFCS, University of Edinburgh. Published as technical report ECS-LFCS-94-304.

[Goo68] N. Goodman, *Intuitionistic arithmetic as a theory of constructions*. PhD thesis, Stanford University, 1968.

[Gor87] L. Gordeev. On cut elimination in the presence of Peirce rule. *Archiv fur Mathematische Logik und Grundlagenforsch*, 26(3-4):147–164, 1987.

[Gri90] T. Griffin. A formulae-as-types notion of control. In *Proc. 17th* ACM *Symposium on Principles of Programming Languages*. ACM press, 1990.

[GRR95] C. A. Gunter, D. Rémy, and J. G. Riecke. A generalisation of exceptions and control in ML-like languages. In *Proc.* ACM *Conference Functional Programming and Computer Architecture*, pages 12–23. ACM Press, 1995.

[Gun92] C. A. Gunter. *Semantics of Programming Languages: Structures and Techniques*. Foundations of Computing. MIT Press, 1992.

[Har85] L. A. Harrington, editor. *Harvey Friedman's research on the foundations of mathematics*, volume 117 of *Studies in logic and the foundations of mathematics*. North-Holland, 1985.

[Hil28] D. Hilbert. The Foundations of Mathematics. Originally 'Die Grundlagen der Mathematik', published Hamburg University, 1928. Translation and commentary in *From Frege to Gödel* [vanHeijenoortJ:frofg], pages 464–479, 1969.

[Hin69] J. R. Hindley. The principal type-scheme of an object in combinatory logic. *Transactions of the American Mathematical Society*, 146:29–60, 1969.

[Hof95a] M. Hofmann. Extensional concepts in intensional type theory. PhD thesis, Edinburgh University, 1995.

[Hof95b] M. Hofmann. Syntax and semantics of dependent types. Lecture notes from the 1995 CLiCS-II Summer School.

[How80] William A. Howard. The formulae–as–types notion of construction. In *To H. B. Curry: Essays on Combinatory Logic, Lambda Calculus and Formalism* [HindleyJR:curecl], pages 479–490, 1990.

[HS97] M. Hofmann and T. Streicher. Continuation models are universal for $\lambda\mu$-calculus. Accepted for publication, LICS'97.

[HS86] J. R. Hindley and J. P. Seldin. *Introduction to combinators and λ-calculus*. Cambridge University Press, 1986.

[HS80] J. R. Hindley and J. P. Seldin, editors. *To H. B. Curry: Essays on Combinatory Logic, Lambda Calculus and Formalism*. Academic Press, 1980.

[JG95] C. B. Jay and N. Ghani. The virtues of eta-expansion. *Journal of Functional Programming*, 5(2):135–154, 1995.

[Joh79] P. T. Johnstone. Conditions related to De Morgan's law. *Applications of Sheaves*, LNM 753:479–491, 1979.

[Kle52a] S. C. Kleene. *Introduction to Metamathematics*. D. van Nostrand, Princeton, New Jersey, 1952.

[Kle52b] S. C. Kleene. Permutability of inferences in Gentzen's calculi LK and LJ. *Memoirs of the American Mathematical Society*, 10:1–26, 1952.

[KMO92] S. A. Kurtz, J. C. Mitchell, and M. J. O'Donnell. Connecting Formal Semantics to Constructive Intuitions. *Constructivity in Computer Science*, LNCS 613:1–21, 1992.

[Koh98a] T.-W. Koh. Multiplicative linear type theories. DPhil thesis, Oxford University, 1998.

[Koh98b] T.-W. Koh and C.-H. L. Ong. Type theories for $\star$-autonomous categories. Submitted for publication, 1998.

[Kol25] A. N. Kolmogorov. On the principle of excluded middle. First published as 'O principe tertium non datur' in *Matematiceskij Sbornik*, 32:646–667, 1925. Translation and commentary by H. Wang published in *From Frege to Gödel* [vanHeijenoortJ:frofg], pages 414–437, 1967.

[Kol32] A. N. Kolmogorov. On the interpretation of intuitionistic logic. First published as 'Zur Deutung der intuitionistischen Logik' in *Mathematischen Zeitschrift*, 35:58–65, 1932. Translation published in *From Brouwer to Hilbert* [MancosuP:frobh], 1998.

[Kre58] G. Kreisel. Mathematical significance of consistency proofs. *Journal of Symbolic Logic*, 23:155–182, 1958.

[Kre62] G. Kreisel. Foundations of intuitionistic logic. In *Logic, Methodology and the Philosophy of science*, pages 198 – 210. North-Holland, 1962.

[Kre68] G. Kreisel. A survey of proof theory. *Journal of Symbolic Logic*, 33:321–388, 1968.

[Kre71] G. Kreisel. A survey of proof theory II. In *Proc. Second Scandinavian Logic Symposium* [FenstadJE:prossl], 1971.

[Kri90] J.-L. Krivine. *Lambda-Calculus, types and models*. Masson and Ellis Horwood, 1993. First published as 'Lambda-calcul, types et modèles', Masson, 1990.

[Kri94a] J.-L. Krivine. Classical logic, storage operators and second-order lambda-calculus. *Annals of Pure and Applied Logic*, 68:53–78, 1994.

[Kri94b] J.-L. Krivine. A general storage theorem for integers in call-by-name λ-calculus. *Theoretical Computer Science*, 129:79–94, 1994.

[Lai97] J. Laird. Full abstraction for functional languages with control. In *Proc.* IEEE *Symposium on Logic in Computer Science (LICS '97)*, IEEE Computer Society Press, 1997.

[Lam80] J. Lambek. From lambda calculus to cartesian closed categories. In *To H. B. Curry: Essays on Combinatory Logic, Lambda Calculus and Formalism* [HindleyJR:curecl], pages 363–402. Academic Press, London, 1980.

[Lau70] H. Lauchli. An abstract notion of realizability for which intuitionistic predicate calculus is complete. In A. Kino, J. Myhill, and R. E. Vesley, editors, *Intuitionism and Proof theory*. North-Holland, London, 1970.

[Law64] F. W. Lawvere. An elementary theory of the category of sets. *Proc. National Academy of Sciences of the United States of America*, 52:1506–1511, 1964.

[Law67] F. W. Lawvere. Category-valued higher-order logic. Presented at UCLA 1967 set theory symposium.

[Law69] F. W. Lawvere. Adjointness in foundations. *Dialectica*, 23:281–296, 1969.

[Lei82] D. Leivant. Reasoning about functional programs and complexity classes associated with type disciplines. In *Proc. Symp. Foundations of Computer Science*, pages 460–469, 1983.

[Lew72] D. Lewis. General semantics. In D. Davidson and G. Harman, editors, *emantics of Natural Language*, pages 169–218. Dordrecht, 1972.

[Luo94] Z. Luo. *Computation and reasoning*. Clarendon Press, Oxford, 1994.

[LO96] James Lipton and Michael J. O'Donnell. Some intutions behind realizability semantics for constructive logic: Tableaux and Laüchli countermodels. *Annals of Pure and Applied Logic*, 81:187–239, 1996.

[LRS93] Y. Lafont, B. Reus, and T. Streicher. Continuation semantics or expressing implication by negation. Technical Report TR 93-21, University of Munich, 1993.

[LS86] J. Lambek and P. J. Scott. *Introduction to Higher Order Categorical Logic*. Cambridge Studies in Advanced Mathematics Vol. 7. Cambridge University Press, 1986.

[Man98] P. Mancosu, editor. *From Brouwer to Hilbert: the debate on the foundations of mathematics in the 1920s*. Oxford University Press, New York, 1998.

[McB94] Muhammad Ali McBeth. *Combinatorial Number Theory*. Edwin Mellen, Lampeter, Dyfed, Wales, 1994.

[ML70] Per Martin-Löf. *Notes on Constructive Mathematics*. Almqvist and Wiksell, Stockholm, 1970.

[ML71] P. Martin-Löf, Hauptsatz for intuitionistic simple type theory. In *Logic, Methodology and philosophy of science* IV, pages 279–290. Studies in Logic and Foundations of Mathematics, vol. 74, North-Holland, 1973.

[ML73] P. Martin-Löf, An intuitionistic theory of types: predicative part. In *Logic Colloquium '73*, pages 73–118, North-Holland, 1975.

[ML73] P. Martin-Löf, About models for intuitionistic type theories and the notion of definitional equality. In *Proc. Third Scandinavian Logic Symposium*, pages 81–109. Studies in Logic and Foundations of Mathematics, vol. 82, North-Holland, 1975.

[ML84a] Per Martin-Löf. Constructive mathematics and computer programming. In L. J. Cohen, editor, *Logic, Methodology and Philosophy of Science VI*, pages 153–175. North-Holland, Amsterdam, 1984.

[ML84b] Per Martin-Löf. *Intuitionistic Type Theory*. Bibliopolis, Naples, 1984.

[ML96] Per Martin-Löf. On the meanings of the logical constants and the justifications of the logical laws. *Nordic Journal of Philosophical Logic*, 1:11–60, 1996.

[Mur91a] Chetan Murthy. An evaluation semantics for classical proofs. In *Proc. 5th* IEEE *Annual Symposium on Logic in Computer Science*. IEEE Computer Society Press, 1991.

[Mur91b] Chetan Murthy. *Extracting constructive content from classical proofs*. PhD thesis, Cornell University, 1991.

[NPS90] Bengt Nordstrom, Kent Petersson, and Jan M Smith. *Programming in Martin-Löf's type theory*. Oxford University Press, New York, 1990.

[Ong96] C.-H. L. Ong. A semantic view of classical proofs: type-theoretic, categorical, denotational characterizations. In *Proc. Eleventh Annual* IEEE *Symposium on Logic in Computer Science*, pages 230–241. IEEE Computer Society Press, 1996.

[OS97] C.-H. L. Ong and C. A. Stewart. A Curry–Howard foundation for functional computation with control. In *Proc. 24th* ACM *Symposium on Principles of Programming Languages*, pages 215 – 217. ACM Press, 1997.

[Par91] M. Parigot. Free deduction: an analysis of 'computations' in classical logic. *Proc. Russian Conference on Logic Programming*, LNCS 592 :361–380, 1991.

[Par92a] M. Parigot. Classical proofs as programs. Preprint of paper presented at the TYPES workshop (Båstad June 1992).

[Par92b] M. Parigot. $\lambda\mu$-calculus: an algorithmic interpretation of classical natural deduction. *Logic Programming and Automated Reasoning*, LNCS 624:190–201, 1992.

[Par93] M. Parigot. Strong normalization for second order classical natural deduction. In *Proc. Eighth Annual* IEEE *Symposium on Logic in Computer Science*, pages 39–46, IEEE Computer Society Press, 1993.

[Par97] M. Parigot. Proofs of strong normalisation for second order classical natural deduction. *Journal of Symbolic Logic*, 62(4):1461–1479, 1997.

[Pei85] C. S. Peirce. On the Algebra of Logic: a contribution to the philosophy of notation. *American journal of mathematics*, 7:180–202, 1885.

[Pit97] A. M. Pitts. Operationally-Based Theories of Program Equivalence. In P. Dybjer and A. M. Pitts, editors, *Semantics and Logics of Computation*, pages 241–298, Publications of the Newton Institute, Cambridge University Press, 1997.

[Plo75] Gordon Plotkin. Call-by-name, call-by-value and the lambda calculus. *Theoretical Computer Science*, 1:125–159, 1975.

[Plo77] Gordon Plotkin. LCF considered as a programming language. *Theoretical Computer Science*, 5:223–255, 1977.

[Pra65] Dag Prawitz. *Natural Deduction: a Proof-Theoretical Study*. Almquist and Wiskell, 1965.

[Pra71] Dag Prawitz. Ideas and results in proof theory. In *Proc. Second Scandinavian Logic Symposium* [FenstadJE:prossl], pages 237–309, 1971.

[Pra75] Dag Prawitz. Comments on Gentzen-type procedures and the classical notion of truth. *Proof theory symposium*, LNM 500:290–319, 1975.

[Pra77] Dag Prawitz. Meaning and proof: on the conflict between classical and intuitionistic logic. *Theoria*, 43:2–40, 1977.

[Pra78] Dag Prawitz. Proofs and the meaning and completeness of the logical constants. In J. Hintikka, editor, *Essays on mathematical and philosophical logic*, pages 25–40, D. Reidel, Dordrecht, 1978.

[Pri60] A. Prior. The runabout inference ticket. *Analysis*, 21:38–39, 1960. Reprinted in 'Philosophical logic' [StrawsonPF:phil].

[RS94] J. Rehof and M. H. Sorensen. The $\lambda_\triangle$ calculus. *Theoretical Aspects of Computer Software*, LNCS 789:516–542, 1994.

[Rey84] J. C. Reynolds. Polymorphism is not Set-Theoretic. *Semantics of Data Types*, LNCS 173:145–156, 1984.

[RP98] Eike Ritter and David Pym. On the semantics of classical disjunction. Submitted for publication, 1998.

[RW10] B. Russell and A. N. Whitehead. *Principia Mathematica*. Three volumes. Cambridge University Press, 1910–1913.

[Sch90] Ernst Schröder. *Vorlesungen ber die Algebra der Logik (exakte Logik)*. Originally published in German, three volumes, between 1890 and 1905. New edition published Chelsea, New York, 1966.

[Sch76] H. Schwichtenberg. Definierbare Funktionen im Lambda-Kalkul mit Typen. *Archiv Logik Grundlagenforsch*, 17:113–114, 1976. No english translation.

[Sch77] H. Schwichtenberg. Proof theory: Some applications of cut-elimination. In *The Handbook of Mathematical Logic* [BarwiseJ:hanml], pages 867–895, 1977.

[Sch95] H. Schwichtenberg. Proofs, lambda terms and control operators. In H. Schwichtenberg, editor, *Logic of Computation*. Lecture notes from the 1995 Marktoberdorf Summer School.

[Sco70] D. S. Scott. Constructive validity. *Symposium on Automated Deduction*, LNM 125:237–275, 1970.

[Sco77] D. S. Scott. Identity and existence in intuitionistic logic. *Applications of Sheaves*, LNM 753:660–696, 1977.

[Sco80] D. S. Scott. Lambda calculus: Some models, some philosophy. In J. Barwise, H. J. Keisler, and K. Kunen, editors, *The Kleene Symposium*, pages 223–265. North-Holland Publishing Company, 1980.

[Sea69] J. Searle. *Speech Acts: An Essay in the Philosophy of Language*. Cambridge University Press, New York, 1969.

[Set98] A. Setzer. Extending Martin-Lf Type Theory by one Mahlo-Universe. Accepted for publication, *Archiv für Mathematische Logik*. Available as preprint since 1998.

[Sie88] W. Sieg. Hilbert's program sixty years later. In 'A symposium on Hilbert's program', *Journal of Symbolic logic*, 53:338–348, 1988.

[Sim85] S. G. Simpson. Friedman's research on subsystems of second order arithmetic. In *Harvey Friedman's research on the foundations of mathematics* [HarringtonLA:harfrf], pages 137–159, 1985.

[Sim88] S. G. Simpson. Partial realizations of Hilbert's program. In 'A symposium on Hilbert's program', *Journal of Symbolic logic*, 53:349–363, 1988.

[Sim98] S. G. Simpson. *Subsystems of Second Order Arithmetic*. Perspectives in Mathematical Logic, Springer-Verlag, 1998.

[Smi88] J. M. Smith. On the independence of Peano's fourth axiom from Martin-Löf's type theory without universes. *Journal of Symbolic Logic*, 53:840–845, 1988.

[SU99] Z. Splawski and P. Urzyczyn. Type Fixpoints: Iteration vs. Recursion. To appear in *Proc. 4th ICFP*, Paris, France, 1999. Available as preprint.

[SR96] T. Streicher and B. Reus. Continuation semantics: abstract machines and control operators. To appear *Journal of Functional Programming*, 1998.

[SS78] D. J. Shoesmith and T. J. Smiley. *Multiple-Conclusion Logic*. Cambridge, 1978.

[Sta94] I. Stark. *Names and Higher-Order Functions*. PhD thesis, University of Cambridge, 1994. Published as Technical Report 363, University of Cambridge Computer Laboratory.

[Str67] P. F. Strawson. *Philosophical logic*. Oxford University Press, 1967.

[Sun86] Göran Sundholm. Proof theory and meaning. In D. Gabbay and F. Guenther, editors, *Handbook of Philosophical Logic*, volume III, pages 471–506, 1986.

[Sza69] M. E. Szabo, editor. *The Collected Papers of Gerhard Gentzen*. North-Holland, 1969.

[Tai67] W. W. Tait. Intesional interpretation of functionals of finite type I. *Journal of Symbolic Logic*, 32:198–212, 1967.

[Tai68] W. W. Tait. Normal derivability in classical logic. *The Syntax and Semantics of Infinitary Logic*, LNM 72:204–236, 1968.

[Tai94] W. W. Tait. The law of the excluded middle and the axiom of choice. In *Mathematics and mind*, pages 45–70. Oxford University Press, New York, 1994.

[Tar44] A. Tarksi. The semantic conception of truth and the foundations of semantics. *Philosophy and Phenomenological Research*, 4:341–375, 1944. Reprinted *Collected Works* [TarskiA:colw], 1986.

[Tar86] A. Tarksi. *Collected Works*. Birkhäuser, 1986. 4 volumes, edited by S. R. Givant and R. McKenzie.

[Ten96] N. Tennant. Negation, absurdity and contraraiety. In D. M. Gabbay and H. Wansing, editors, *Negation*. Kluwer Academic Press, 1996.

[Ten91] R. D. Tennent. *Semantics of Programming Languages*. International Series in Computer Science. Prentice Hall, 1991.

[Tro87] A. S. Troelstra. On the syntax of Martin-Löf's type theories. *Theoretical Computer Science*, 51:1–26, 1987.

[Tro98] A. S. Troelstra. Realizability. In *The Handbook of Proof Theory* [BussSR:hanpt], 1998.

[TvD88] A. S. Troelstra and D. van Dalen. *Constructivism in Mathematics: An Introduction*. North-Holland, 1988. 2 volumes.

[Ung92] A. M. Ungar. *Normalization, cut-elimination and the theory of proofs*. Number 28 in CSLI Lecture notes. Stanford, 1992.

[Upe92] Vladimir Upensky. Kolmogorov and mathematical logic. *Journal of Symbolic Logic*, 57:385–412, 1992.

[vH67] J. van Heijenoort, editor. *From Frege to Gödel: a Source Book in Mathematical Logic, 1879–1931*. Harvard University Press, Cambridge, MA, 1967.

[Wai95] S. S. Wainer. Basic proof theory with applications to computation. In Helmut Schwichtenberg, editor, *Logic of Computation*. Lecture notes from the 1995 Marktoberdorf Summer School.

[Wit53] L. Wittgenstein. *Philosophical Investigations*. Translated from the German by G. E. M. Anscombe. Blackwell, Oxford, 1953.

[Wri81] C. Wright. Dummett and revisionism. *Philosophical Quarterly* 31:47–67. Reprinted as 'Anti-realism and Revisionism' in *Realism, Meaning and Truth* [WrightC:reamt].

[Wri97] C. Wright. *Realism, Meaning and Truth*. Second edition. Blackwell, Oxford, 1997.

Index

affirmative judgement, 122
affirmative subterms, 123
affirmative terms, 122
alpha compatible, 103
alpha equivalence, 101, 166
analytic harmony, 33
analytic thread, 50
antecedent derivation, 41
antithesis, 121
arithmetic type theory, 163
assertion conditions, 14
assertional content, 111
assumption, 77
    term, 77
    type, 77
assumption binding, 44, 132
assumption packet, 36
atomic formulae, 34
axiomatic system, 118

bare formula, 120
beta redex, 57
branch, 44

calculus
    CTT, 188
    CND, 120
    LK, 115
    NJ, 38
    λ, 54
    $\lambda\mu$, 120
    $\lambda\mu^*$, 153
    $\lambda\mu^{**}$, 155
    $\mathrm{PRA}^{\omega}_{\mu}$, 139, 185
    PRA^{ω}, 66
    CTT, 164
    ITT, 75
    NTT, 157
canonical, 144
canonical form, 104
Church–Rosser, 63
classical type theory, 164
commuting conversions, 108
compatible closure, 58
connective
    arity, 34
conservativity, 33
constructive content, 111
constructors, 59
contentual contribution, 111
context, 58
contraversion, 121
contributed assumptions, 36
conversion
    beta, 57
    beta-form zeta, 128
    eta, 64
    eta-form zeta, 128
    mu-beta, 130
    mu-eta, 130
critical pairs, 63
cut formula, 37, 67

denial, 121
dependency forgetting translation, 102, 166
dependent Cartesian product, 74
dependent function space, 74
derivable, 38
derivation, 36
    above A, 40
    admissible, 36

antecedent, 41
below A, 40
conclusion, 36
contradiction, 121
deduction, 122
height, 37
last rule, 37
leaves, 37
natural, 24
open assumptions, 36
refutation, 122
residual, 41
structural, 24
descending chain, 47
detours, 42
ineliminable, 69
deviant terms, 162
diamond property, 63
direct elimination rule, 36
direct subderivations, 36
discharge, 37
discharged formula, 35
downwards conservativity, 108

eigenformulae, 35
eigenvariable conditions, 119
elimination chain, 47
full, 47
empty judgement, 122
empty subterms, 123
empty terms, 122
extended polynomials, 46
external semantics, 7
extractors, 59

formula, 34
attitude, 120
formula occurrence
eta minimal, 130
formula candidates, 76
formula occurrence
antecedent, 41
maximal, 41
minimal, 52, 130
reducible, 130
stands immediately above, 37
formula occurrences, 37
formulae–as–types, 5
free variables, 55

general substitutions, 79
governed by R, 37
governed by $\otimes$, 34

harmony, 14
head contraversions, 168
head normal form, 104, 168
head normal forms, 104
head reducible, 104, 168
head subterms, 104, 168
head variable, 104, 168
HNF, *see* head normal form
hypothetical judgements, 78
hypothetical premiss, 35

incomplete expressions, 30
indirect elimination rule, 36
indirect premisses, 129
indirect residuals, 49
indirect subderivations, 129
inference rule, 35
$\mathcal{R}$ relates A to B, 37
arity, 35
direct elimination, 36
eigenformulae, 35
indirect elimination, 36
instances, 35
introduction, 36
rules governing $\otimes$, 35
structural, 121
initial thread, 37
instance of A, 35
instantiation, 118
inversion principle
permutative decomposition, 109

junction, 44
justifies, 38

lambda-mu calculus, 122
left recursion, 67
logic
    classical minimal, 25, 118
    intuitionistic, 118
logical formalism, 15, 32
logical harmony, 33

main branch, 44
major premisses, 44
maxima, 41
maxima elimination, 41
measure, 47, 133
    degree, 34
    depth, 45
minima
    eta, 130
    zeta, 130
minima decomposition, 52
minimal, 52, 130
minor branch, 44
minor premisses, 44
mu conversions, 130
mu-bar rules, 156
muliple generality, 30

naïve type theory, 157
negative contraversion, 156
non-canonical, 144
normal form, 52

open subterms, 144
order, 44
overlapping redexes, 41

parameterised definition, 57
partly reducible, 144
path, 129
    maximal, 129
    principal, 50
primitive letters, 34
principal premiss, 36
principal thread, 50
proper names, 30
propositional equality, 85

quasi HNF, 168

ramified reducible, 145
range of a telescope, 77
recursion
    left, 67
    right, 67
    strict, 67
recursor term, 65
redex–contractum pair, 57
reducibility measure, 145
reducibility predicate
    co-RR, 150
    FR, 150
    fully reducible, 150
    PR, 144
    RR, 145
reduction sequence, 133
    full, 133
rejective covering, 125
rejective judgement, 122
rejective subterms, 123
rejective terms, 122
relative normalisation, 18
required formula, 35
residual derivation, 41
rewrite chain, 53
    full, 53
rewrite chain bound, 69
right recursion, 67
rule, *see* inference rule

schematic letter, 34
segment, 129
semantics
    two factor, 111
    verificationist, 111
sequent, 37
simple definition, 57
simple measure, 47
simple normal form, 42
simple premiss, 35
spanning thread, 37
strong normalisation, 42

subderivation, 37
subformulae, 34
subjects, 76
substitution candidates, 79
substitution operation, 56
substitutions
    A-class, 145
    B-class, 145
    C-class, 150
    D-class, 150
    grounding, 145
    mixed, 127
    reducing, 144
    simple grounding, 145
    unary wild, 150
    wild, 150
subterms, 55, 123
synthetic harmony, 33
synthetic thread, 50

telescope, 54
    domain, 54, 77
    range, 54
telescope candidates, 77
term
    α/x-erasing, 147
    term constant, 117
term assumption, 77
term candidates, 54, 76
term decomposition
    n-ary, 58
    constructor–recursor, 65
term formers
    wild, 149
term property
    canonical, 72
    neutral, 72
    non-canonical, 72
terminal thread, 37
theory
    intuitionistic type, 75
theory of an axiomatic system, 118
thread, 37
tonk, 32
type assumption, 77
type theory
    classical, 26, 157
    Martin-Löf's, 18
types, 54

upwards conservativity, 108

vacuous contraversion, 129
valuation of Γ, 100
verification transcendence, 13

weak normalisation, 42
well-founded recursion, 141