From Constructive Mathematics to Computable Analysis via the Realizability Interpretation

Vom Fachbereich Mathematik
der Technischen Universität Darmstadt
zur Erlangung des akademischen Grades eines
Doktors der Naturwissenschaften (Dr. rer. nat.)
genehmigte

Dissertation

von

Dipl.-Math. Peter Lietz
aus Mainz

Referent:	Prof. Dr. Thomas Streicher
Korreferent:	Dr. Alex Simpson
Tag der Einreichung:	22. Januar 2004
Tag der mündlichen Prüfung:	11. Februar 2004

Darmstadt 2004
D17

Hiermit versichere ich, dass ich diese Dissertation selbständig verfasst und nur die angegebenen Hilfsmittel verwendet habe.

Peter Lietz

Abstract

Constructive mathematics is mathematics without the use of the principle of the excluded middle. There exists a wide array of models of constructive logic. One particular interpretation of constructive mathematics is the realizability interpretation. It is utilized as a metamathematical tool in order to derive admissible rules of deduction for systems of constructive logic or to demonstrate the equiconsistency of extensions of constructive logic. In this thesis, we employ various realizability models in order to logically separate several statements about continuity in constructive mathematics.

A trademark of some constructive formalisms is predicativity. Predicative logic does not allow the definition of a set by quantifying over a collection of sets that the set to be defined is a member of. Starting from realizability models over a typed version of partial combinatory algebras we are able to show that the ensuing models provide the features necessary in order to interpret impredicative logics and type theories *if and only if* the underlying typed partial combinatory algebra is equivalent to an untyped pca.

It is an ongoing theme in this thesis to switch between the worlds of classical and constructive mathematics and to try and use constructive logic as a method in order to obtain results of interest also for the classically minded mathematician. A classical mathematician can see the value of a solution algorithm as opposed to an abstract proof of the existence of a solution, but he or she would not insist on a constructive correctness proof for that algorithm. We introduce a class of formulae which is supposed to capture this pragmatic point of view. The class is defined in such a way that existence statements have a strong status, yet the correctness of an operation need only be proved classically. Moreover, this theory contains only classically true formulae. We pose the axiomatization of this class of formulae as an open problem and provide partial results.

Like ordinary recursion theory, computable analysis is a branch of classical mathematics. It applies the concept of computability to entities of analysis by equipping them with a generalization of Gödelizations called representations. Representations can be organized into a realizability category with rich logical properties. In this way, natural representations of spaces can be found by categorically interpreting the description of the underlying set of a space. Computability and non-computability results can be and are shown on an abstract, logical level.

Finally, we turn to another application of realizability models, the field of strong normalization proofs for type theoretic frameworks. We will argue why we think that the modified realizability topos is not suited for this purpose and propose an alternative.

Zusammenfassung

Konstruktive Mathematik ist Mathematik ohne die Verwendung des Prinzips *tertium non datur*. Es gibt eine Vielzahl unterschiedlicher Modelle für die konstruktive Logik. Eine bestimmte Interpretation der konstruktiven Mathematik ist die Realisierbarkeits-Interpretation. Sie findet Anwendung als ein metamathematisches Werkzeug welches es gestattet, zulässige Regeln oder Äquikonsistenzaussagen für logische Kalküle nachzuweisen. In dieser Dissertation verwenden wir die Realisierbarkeits-Interpretation zum Zwecke der Separierung verschiedener Aussagen über Stetigkeit im Rahmen der konstruktiven Mathematik.

Eine Eigenart einiger konstruktiver Formalismen ist die Prädikativität. Prädikative Logik verbietet die Definition einer Menge durch Quantifikation über eine Familie von Mengen, welche die zu definierende Menge enthält. Ausgehend von Realisierbarkeits-Modellen über einer getypten Version partieller kombinatorischer Algebren zeigen wir, dass die zugehörigen Modelle die nötigen Eigenschaften zur Interpretation imprädikativer Logik und Typ Theorie genau dann besitzen, wenn die zugrundeliegende partielle kombinatorische Algebra äquivalent ist zu einer ungetypten pca.

Der Wechsel zwischen den Welten der konstruktiven und der klassischen Mathematik und der Versuch, die konstruktive Logik als eine Methode zu benutzen, um Resultate zu erzielen, welche von Interesse für klassische Mathematiker sind, ist ein wiederkehrendes Thema dieser Dissertation. Ein klassischer Mathematiker erkennt sehr Wohl den Wert eines Lösungsalgorithmus im Vergleich zu einem bloßen Beweis der Existenz einer Lösung, aber er wird gewöhnlich nicht darauf bestehen, dass der Korrektheitsbeweis für den Algorithmus konstruktiv ist. Wir führen eine Klasse von Formeln ein, welche diesen pragmatischen Blickwinkel erfassen soll. Die Klasse ist derart definiert, dass Existenzaussagen einen starken Status haben, Korrektheitsbeweise für Operationen jedoch nur klassisch geführt werden brauchen. Die besagte Klasse von Formeln enthält nur klassisch wahre Formeln. Wir stellen die Axiomatisierung dieser Klasse als ein Problem und bieten Teilergebnisse.

Wie die Rekursionstheorie, so ist auch die berechenbare Analysis ein Zweig der klassischen Mathematik. Die berechenbare Analysis wendet das Konzept der Berechenbarkeit an auf Größen der Analysis, indem sie sie mit einer Verallgemeinerung von Gödelisierungen, genannt Darstellungen, ausstattet. Darstellungen können in einer Realisierbarkeits-Kategorie mit reichhaltigen logischen Eigenschaften zusammengefasst werden. Auf diesem Wege können natürliche Darstellungen von Räumen gefunden werden, indem man die Beschreibung der unterliegenden Menge der Räume kategoriell interpretiert. Berechenbarkeits- und Nicht-Berechenbarkeitsresultate können so auf einer abstrakten, logischen Ebene hergeleitet werden.

Zuletzt wenden wir uns einer weiteren Anwedung der Realisierbarkeitsmodelle zu, nämlich dem Gebiet der starken Normalisierungbeweise für typtheoretische Kalküle. Wir legen dar warum wir denken, dass der modified realizability topos nicht das geeignete Modell für diesen Zweck ist und schlagen eine Alternative vor.

Acknowledgements

I would like to gratefully acknowledge the many fruitful discussions that have contributed to this thesis. I am greatly indebted to: Andrej Bauer, Lars Birkedal, Vasco Brattka, Martín Escardó, Helge Glöckner, Karl Heinrich Hofmann, Martin Hofmann, Shin-ya Katsumata, Klaus Keimel, John Longley, Matias Menni, Jaap van Oosten, Matthias Schröder, Klaus Weihrauch, Alex Simpson, Eike Ritter, Peter Schuster, Helmut Schwichtenberg, Dana Scott and Bas Spitters.

In particular, I would like to thank Andrej Bauer, whose two visits to Darmstadt have meant times of great motivation and good progress. The regular correspondence with Andrej Bauer, John Longley, Jaap van Oosten, Matthias Schröder, Alex Simpson and Bas Spitters has always been very stimulating and valuable to me and is much appreciated.

I would like to express my gratitude for invitations to present my work at the IT-University of København, the FernUniversität Hagen, and the Ludwig Maximilians Universität in München.

I have profited from having been allowed to participate in the "Postgraduate Course in the Theory of Computation" of the "Laboratory for Foundations of Computer Science" at the University of Edinburgh.

I would like to thank my officemates in Edinburgh and Darmstadt and my fellow members of the research unit "Logik und mathematische Grundlagen der Informatik" at the Technische Universität Darmstadt.

Finally and foremost, I would like to thank my supervisor Thomas Streicher for the great many things I have learnt from him and for his enduring support and patience.

I gratefully acknowledge receipt of the *Promotionsstipendium nach dem hessischen Gesetz zur Förderung von Nachwuchswissenschaftlern* (PhD–grant according to the Hessian Law for the furtherance of young scientists) during the time from July 1998 to September 2000.

This document was typeset with LaTeX 2$_\varepsilon$ using the KOMA–Script documentclass scrbook, Kristoffer Rose's XY–pic package, Paul Taylor's prooftree package and the AMS-packages.

Contents

Introduction

Constructive mathematics is mathematics without the use of the principle of the excluded middle or, equivalently, without the proof principle of *reductio ad absurdum*. While it is often harder and sometimes impossible to prove a classical theorem constructively, a constructive proof, once obtained, is much more informative. From a constructive existence proof, for instance, one can extract a witness for the existence statement, which is in general impossible for a classical proof. One might say that in constructive logic, the status of existence statements is much stronger than in classical logic. Moreover, there exists a wide array of models of constructive logic. One particular interpretation of constructive mathematics (actually, a theme with many variations) is the realizability interpretation. The realizability interpretation can be seen as a concrete incarnation of the Brouwer Heyting Kolmogorof interpretation of the logical connectives. It is utilized as a metamathematical tool in order to derive admissible rules of deduction for systems of constructive logic or to demonstrate the equiconsistency of extensions of constructive logic. It is an ongoing theme in this thesis to switch between the worlds of classical and constructive mathematics and to try and use constructive logic as a method in order to obtain results of interest also for the classically minded mathematician. A classical mathematician can see the value of a solution algorithm as opposed to an abstract proof of the existence of a solution, but he or she would not insist on a constructive correctness proof for that algorithm. In section 2.1.3 we introduce a class of formulae which is supposed to capture this pragmatic point of view. The class is defined in such a way that existence statements have a strong status, yet the correctness of an operation need only be proved classically. Moreover, this theory contains only classically true formulae. We pose the axiomatization of this class of formulae as an open problem and provide partial results.

A trademark of some constructive formalisms is predicativity. Predicativism means that, for any set, the collection of subsets of that set is not an entity in its own right but each subset has to be defined by refering only to previously defined subsets. This in particular prevents the definition of a set by quantifying over a collection of sets that the set to be defined is a member of. This is perceived by predicavists as a vicious circle. Starting from realizability models over a typed version of partial combinatory algebras we are able to show that the ensuing models provide the features necessary in order to interpret impredicative logics and type theories *if and only if* the underlying typed partial combinatory algebra is equivalent to an untyped pca.

One aspect of constructive mathematics as opposed to classical mathematics is that it can be extended in various ways incompatible with classical reasoning. A continuity principle is an axiom which states that every function defined on a metric space pertaining to some specified class of spaces to another metric space is continuous. This is generally incompatible with classical logic, but perfectly (equi-)consistent with constructive logic. As possible classes we can for example choose the class of complete separable metric spaces or the class of complete totally bounded spaces. We exhibit realizability models that satisfy the weaker and fail to satisfy the stronger principle, thereby demonstrating that the implication between these principles is strict. Moreover we give an easy model validating the statement “all functions from $\mathbb{R}$ to $\mathbb{R}$ are sequentially continuous” and falsifying the statement “all functions from $\mathbb{R}$ to $\mathbb{R}$ are continuous”.

Like ordinary recursion theory, computable analysis is a branch of classical mathematics. It applies the concept of computability to entities of analysis. In addition to whether or not there exists a solution to a mathematical problem, in computable mathematics one is interested in the question whether, given that the input data is computable, there is a computable solution. Furthermore one asks whether one can uniformly compute a solution from the input data. In order to give meaning to sentences such as “the real number x is computable” or “some function f is computable” one has to introduce some computability structure. Amongst the several, non-equivalent approaches, we shall concentrate on the approach taken by Weihrauch and Kreitz. It utilizes a generalization of the notion of Gödel numberings, called representations. Representations can be organized into a category with rich logical properties, which is essentially a realizability category. In this way, natural representations of spaces can be found by categorically interpreting the description of the underlying set of a space. Computability and non-computability results can be shown on an abstract, logical level.

Finally, we turn to another application of realizability models, the field of strong normalization proofs for type theoretic frameworks. We will argue why we think that the modified realizability topos is not suited for this purpose and propose an alternative.

1 Constructive Mathematics

This chapter attempts to give a brief overview over several schools of constructive mathematics. The most important branches of constructivism for the purposes of this thesis are E. Bishop's constructivism and L.E.J. Brouwer's intuitionism.

The introduction to the underlying philosophy of constructivist schools is followed by a section on formalized mathematics and an analysis of the principles employed by the various schools.

In the last section of this chapter we give a short introduction to categorical semantics of constructive logic.

1.1 General philosophy of constructivism

Constructive Mathematics is mathematics without the law of the excluded middle or, equivalently, without the rule of reductio ad absurdum. That is, in order to prove an existence statement $\exists x.A$ in constructive mathematics, it is not sufficient to demonstrate that the non-existence $\forall x.\neg A$ is absurd. Likewise, in order to prove a disjunction $A \vee B$ it is not sufficient to show that $\neg A \wedge \neg B$ is absurd. This self imposed restriction results in a finer distinction of concepts that would be equivalent under the reign of classical logic.

Another aspect of constructivism, often, but not necessarily, combined with constructive logic is predicativism. In predicative logic it is not admissible to, in order to construct one set, quantify over a collection of sets of which the set *in statu nascendi* is a member of. If one does not adopt the philosophy that sets are entities that have an ideal existence, then such a construction of a set would constitute a vicious circle.

Both restrictions of ordinary mathematics, while being not far from each other in spirit, can be and are applied independently and in each of the possible combinations. While it is more difficult (and sometimes impossible) to prove a theorem constructively, the existence of a constructive proof has stronger implications than that of a classical proof.

Consistency or enhanced safety from contradiction, by the way, is not amongst the reasons to drop the law of the excluded middle. Due to the negative translation of Gödel and Gentzen, it is just as hard to prove $0 = 1$ in Heyting arithmetic as it is in Peano arithmetic.

1.2 Branches of constructivism

We shall, in the following three sections, give a very brief account of the main branches of constructive mathematics. A more extensive comparative overview can be found in [BR87] or, amongst many other things, in [TvD88a, TvD88b] or [Bee85]. The following sections touch on each of the constructive schools from the point of view of this thesis, they do not serve as a real introduction. Also they lack all due historical information.

1.2.1 Bishop's constructive mathematics

The definitive introductory presentation of the ideas of Bishop's mathematics is the *constructivist manifesto*, the first chapter of the seminal book "Foundations of constructive analysis" [Bis67]. Although not formalized, it is rather clear in practice what are the valid constructions and derivations in Bishop's mathematics (BISH for short).

Bishop style mathematics is based on constructive logic. As Bishop puts it: "when a man proves a positive integer to exist, he should show how to find it". This rules out the unrestricted use of indirect proofs.

Bishop's mathematics is also predicative. As he writes: "A set is not an entity which has an ideal existence: a set exists only when it has been defined. To define a set we prescribe, at least implicitly, what we (the constructing intelligence) must do in order to construct an element of the set, and what we must do in order to show that two elements are equal." [Bis67, p.2]. This implies that, in defining a set, one must not quantify over some collection of *ideal* sets, but may only refer to previously defined sets. Bishop refers to sets and existence as used in classical mathematics as ideal sets and ideal existence as opposed to constructive sets and constructive existence. Sets may have a hypothetical status in order to allow a proof generic for all possible sets, but quantification over all sets or even only all subsets of a given set in order to define a new set is not allowed.

The set that Bishop's mathematics starts out with is the set of natural numbers. Given two sets, it is permissible to construct the set of ordered pairs of elements of these, or the set of functions between these. Furthermore, given a set, one can form the subset of all elements satisfying some given property. Quotient sets are not part of Bishop's mathematics, instead, every set is equipped with a defined notion of equality. Equality is in most cases derived from a notion of apartness. This is typical for constructive mathematics: in order to be able to prove some result, the hypotheses have to be expressed in a positive way. The property that two elements of, say, a metric space are apart is stronger constructively than the property of merely being not equal. Instead, equality can be defined as non-apartness. If equality on some set is defined as the negation of apartness, which is frequently the case, the apartness relation is called *tight*. Indeed, apartness is the key notion in the constructive redevelopment of topology (see [BSV02]). Classical topology is not applicable, here, as the dualism between open and closed sets is intrinsically based on non-constructive reasoning.

It is to be emphasized that constructive mathematics is not about imposing restrictions and hindering mathematical development but about maintaining a higher standard for valid proofs. If theorems fail to be provable constructively, they do so for a good reason.

Equally important, one should keep in mind that Bishop style mathematics is not a logical discipline, but a flavor of ordinary mathematics in its own right. Therefore, the main activity in the field consists in proving theorems rather then demonstrating the unprovability of theorems or making other metamathematical observations. Nevertheless, it is essential to have some clear indicators of when some theorem A is unprovable, constructively. In classical mathematics, in this situation one usually tries, and often succeeds, to prove $\neg A$. This method will, however, succeed less often in constructive mathematics. The reason is that in contrast to classical mathematics, in constructive mathematics a lot more statements, even mathematically natural ones, are undecided by the theory.

Omniscience principles

The principle of the excluded middle is undecided in Bishop style mathematics. It is neither adopted, as in classical mathematics, nor does Bishop's mathematics contain any principle that refutes instances of the principle of the excluded middle, as does, for instance, Bouwer's intuitionism. However, as intuitionism is an equiconsistent extension of Bishop style mathematics, instances of the principle of the excluded middle that are refuted by intuitionism cannot be proved in Bishop's mathematics.

One such instance has been taunted by Bishop the *principle of omniscience*:

PO For a set A either all elements of A have the property P or there is an element of A with the property not P.

A particular choice of a set A and a property P, for which this principle is not part of constructive mathematics, is the so called *limited principle of omniscience*:

LPO Let (a_n) be a binary sequence. Then either there is a k such that $a_k = 1$ or $a_k = 0$ for all k.

A number of theorems of classical mathematics are intrinsically non-constructive, which can be demonstrated by deriving LPO from them. Another principle, weaker than LPO but yet non-constructive and often used for the refutation of theorems, is the *lesser limited principle of omniscience*:

LLPO Let (a_n) be a binary sequence with at most one 1. Then either $a_{2k} = 0$ for all k or $a_{2k+1} = 0$ for all k.

Constructive versions of classical theorems

There is quite a number of classical theorems that are provable constructively. Amongst these are for instance the fundamental theorem of algebra or the Picard-Lindelof theorem. Other classical theorems are non-constructive, but are provable subject to modifications. The law of trichotomy for the reals

$$\forall r \in \mathbb{R}.\, r < 0 \vee r = 0 \vee r > 0$$

is equivalent to LPO relative to BISH and therefore non-constructive. Another example of a classical theorem that entails LPO is the bounded completeness of the reals, i.e. the statement that every inhabited subset of the reals that is bounded above has a supremum. In the other direction, LPO itself proves that every sequence of reals that is bounded above has a supremum.

As to the first example, the theorem

$$\forall x, y \in \mathbb{R}.\ x < y \to \forall r \in \mathbb{R}.\ r < y \vee x < r$$

serves as a constructive substitute for trichotomy. As to the second example, every inhabited subset $P \subseteq \mathbb{R}$ that is bounded above and has the additional property that for all $y, z \in \mathbb{R}$ such that $y < z$ either $P < z$ or there is an $x \in P$ with $y < x$ does have a supremum, constructively. Such a form of constructive theorem is called an *equal conclusion substitute*, as only the hypothesis of the theorem is altered. The additional assumption is void from a classical point of view but essential for a constructive proof. It can be shown that the image of a uniformly continuous function defined on a totally bounded metric space meets the additional requirement. This fact is important for the constructive definition of Banach spaces like $\mathrm{C}[0,1]$. On the other hand, the operator norm of the dual of a Banach space can not be defined, constructively, unless the original space is finite dimensional.

The failure of bounded completeness of the reals is also the culprit for the fact that the distance of some inhabited subset of a metric space to a point is not guaranteed to exist. Subsets, whose distances to any point exist are called *located*. Locatedness is quite a central notion in constructive mathmatics and often has to be required as an extra assumption for a theorem to work, see for instance [Spi02]. These kinds of subtleties are levelled off by the use of classical logic.

The principles

$$\forall r \in \mathbb{R}.\, r \leq 0 \vee r \geq 0 \quad \text{and} \quad \forall x, y \in \mathbb{R}.\ xy = 0 \to (x = 0 \vee y = 0)$$

are each equivalent to LLPO relative to BISH We shall see that the former principle is a consequence of the intermediate value theorem. Define

$$f_r : [0,1] \longrightarrow \mathbb{R} \qquad f_r(x) = \begin{cases} -1 + 3(1+r)x & \text{for } x < \frac{1}{3} \\ r & \text{for } \frac{1}{3} \leq x < \frac{2}{3} \\ -2 + 3r + 3(1-r)x & \text{for } \frac{2}{3} \leq x \end{cases}$$

In the light of the constructive failure of $x < a \vee x \geq a$ one would of course have to argue why the piecewise definition of f_r is admissible. A constructive argument is that given functions

$$f : \mathbb{R}_{\leq a} \longrightarrow \mathbb{R} \quad \text{and} \quad g : \mathbb{R}_{\geq a} \longrightarrow \mathbb{R} \quad \text{such that} \quad f(a) = g(a)$$

we can constructively define a pasted function h as

$$h : \mathbb{R} \longrightarrow \mathbb{R} \qquad x \mapsto f(\min(x, a)) + g(\max(x, a)) - f(a)$$

A more general justification for the pasting of continuous functions can also be given but does involve the constructive Tietze extension theorem (see [BB85, Theorem (6.6)]).

Now regarding the intermediate value theorem, a given zero z of the function f_r would allow us, using the constructive substitute for the law of trichotomy, to decide whether $z < \frac{2}{3}$ or $\frac{1}{3} < z$ and hence to decide whether $r \geq 0$ or $r \leq 0$. This demonstrates the non-constructivity of the intermediate value theorem.

There are two constructive substitutes for the intermediate value theorem, one with equal hypothesis and one with equal conclusion, The latter has the extra assumption that the function f is locally non-zero, i.e. between any two numbers in the domain of definition, there is an argument for which the function yields a result apart from zero. This hypothesis is met e.g. by non-constant polynomials. The equal hypothesis substitute has the weaker conclusion that, for every $\varepsilon > 0$, there is an $x \in [0, 1]$ such that $|f(x)| < \varepsilon$.

Given two real numbers a, b, their maximum and their minimum exist. However, it is not possible to pick a maximal number amongst a and b, as this again would allow us to decide whether $a - b \leq 0$ or $a - b \geq 0$. As a consequence, a uniformly continuous function on a closed interval, albeit having a supremum (i.e., a least upper bound), need not have a maximum (i.e., need not attain a maximal value).

A subset P of some set X is called *decidable* if $\forall x \in X.\, x \in P \vee \neg x \in P$. The only subsets of $\mathbb{R}$ that can be proven to be decidable are $\emptyset$ and $\mathbb{R}$. On the other hand, the equality relation and the relations $\leq$ and $<$ are decidable subsets of both $\mathbb{N} \times \mathbb{N}$ and $\mathbb{Q} \times \mathbb{Q}$. As the principle of the excluded middle is not at our disposition, this has to be proven by natural induction.

Continuity

Several classically equivalent notions of continuity fall apart in Bishop style mathematics. Let $f : \mathbb{R} \longrightarrow \mathbb{R}$ be a function. If f is uniformly continuous on every compact interval, then f is continuous (i.e. ε-δ-continuous) in every point. If f is continuous then it is sequentially continuous (i.e. the image of a convergent sequence is convergent). Unfortunately, neither of these implications can be reversed (see [Ish92] for a detailed analysis of continuity in BISH).

When classically equivalent notions diversify under a constructive examination, it is crucial to choose the most meaningful amongst them. It cannot be shown constructively that every continuous function defined on a compact interval is uniformly continuous. Also, it turns out that continuity alone is too weak a property in the purely constructive setting: One cannot show that a continuous function defined on a compact interval is Riemann-integrable or even just bounded. Therefore, the more meaningful notion of continuity for functions defined on subsets of Euclidean space is that of locally uniform continuity, and hence we define $\mathsf{C}(\mathbb{R}^n)$ to be the set of all real-valued functions on $\mathbb{R}^n$ that are uniformly continuous on every bounded subset.

Choice

The use of choice principles in Bishop style mathematics goes as far as including the axiom of dependent choice and the axiom of unique choice. Although Bishop writes that "a choice function exists because a choice function is *implied by the very meaning of existence*"[BB85, A Constructivist Manifesto], it still has to be verified that the choice function respects the defined equality relation. This reveals an intensional view that Bishop has on mathematics (see also Bridges comment at the end of *loc. cit.*).

1.2.2 Brouwer's intuitionism

Like Bishop's constructivism, Brouwer's intuitionism (INT) is not a formalized theory, in its original conception it is not even based on mathematical logic. This, however, did not prevent Brouwer's disciple Heyting from giving a formalized account of intuitionistic logic [Hey30]. A lively introduction to intuitionism is [Hey56], another extensive treatment is [Dum77].

Brouwer's intuitionism is an extension of Bishop's constructivism, that is, every argument valid in BISH is also valid in INT. Brouwer takes a more fundamental stance than Bishop in that he breaks with classical mathematics. Unlike Bishop style mathematics, Brouwer's intuitionism is not a subset of classical mathematics. It contains axioms that negate instances of the principle of the excluded middle and is hence incompatible with classical mathematics.

Infinitely prodeeding sequences

A fundamental notion in Intuitionism is the notion of an *infinitely prodeeding sequence* of natural numbers. Such a sequence may, but need not, be determined by a law. In order to talk about sequences, both finite and infinite ones, we have got to fix some notation. We choose to use the symbols introduced in [TvD88a, p.186]. Finite sequences are denoted as $\langle a_1, a_2, \ldots a_n \rangle$, in particular the empty sequence is denoted as $\langle\rangle$. The concatenation of the sequences a and x is denoted by $a * x$, where a

is a finite sequence and x is either finite or infinite. The initial segment of length n of the sequence α is denoted by $\overline{\alpha}n$. We write $a \prec a'$ if a is a prefix of a'. If $a \prec a'$, we say that a is an ascendent of a' and a' is a descendent of a. The prefix of length $\mathrm{lth}(a) - 1$ of a is called the immediate ascendent of a, a sequence that a is the immediate ascendent of is called an immediate descendent of a. If the finite sequence a is an element of some set S of finite sequences then any descendent of a that is a member of S is called an S-descendent.

Definition 1.2.1. A decidable set of finite sequences that contains the empty sequence and that with any sequence contains all its prefixes is called a *tree*. A *spread* is a tree such that every finite sequence in S has an immediate S-descendent. That is, a tree S is a spread if $\forall a \in S \exists n \in \mathbb{N}.\, a * \langle n \rangle \in S$. Finally, a *fan* is a *finitely branching spread*, i.e. a spread in which each node has a finite number of immediate descendents.

Every spread determines a set of infinite sequences. We say that the infinite sequence α belongs to the spread S if all initial segments of α are elements of S. By abuse of language we call the thus defined set of sequences S as well. The sense of the statement $x \in S$, depends on whether x is a finite or an infinite sequence. We usually denote finite sequences by small roman and infinite sequences by small greek letters.

Example 1.2.2. Let $(r_n)_{n\in\mathbb{N}}$ be some standard bijective enumeration of the set $\mathbb{Q}$. We define the spread S as the set of all finite sequences $\langle a_1, \dots a_n\rangle$ such that $|r_{\alpha(k)} - r_{\alpha(k+1)}| \leq 2^{-(n+1)}$ for all $1 \leq k < n$. Then for each $\alpha \in S$ the rational sequence $(q_{\alpha(n)})_{n\in\mathbb{N}}$ is a Cauchy sequence. The condition of $\alpha, \alpha' \in S$ representing the same real number does as well depend on the initial segments, alone: The sequences α and α' represent the same real number if and only if $|r_{\alpha(n)} - r_{\alpha'(n)}| \leq 2^{-n}$ for all $n \in \mathbb{N}$.

Example 1.2.3. Let F be the set of $\{-1, 0, 1\}$-valued finite sequences. Obviously, F is defined by a a fan. With each infinite sequence $\alpha \in F$ we associate the real number $\sum_{k=0}^{\infty} \alpha_k 2^{-\alpha_k - 1}$. The set of real numbers that is represented by a sequence in F is exactly the closed interval $[-1, 1]$. Two sequences α, α' represent the same number if and only if $|\sum_{k=0}^{n} \alpha_k 2^{-\alpha_k-1} - \sum_{k=0}^{n} \alpha'_k 2^{-\alpha_k-1}| \leq 2^{-n}$ for all $n \in \mathbb{N}$, that is, two sequences represent the same number if and only if all their pairs of initial segments satisfy some decidable property.

The continuity principle

Two principles govern the nature of infinitely proceeding sequences in intuitionism, the *continuity principle* and the principle of *bar induction*. The continuity principle as taken from [Hey56] (with a slight modification) reads as follows.

Principle 1.2.4 (Continuity). *Let S be a spread and let Φ be a mapping of the set of infinite sequences defined by S into the natural numbers. Then there is a function which computes $\Phi\alpha$ for any infinite sequence α in S from some sufficiently long finite initial segment of α. Moreover, it is decidable for a given finite sequence $a \in S$ whether it is sufficiently long in order to compute $\Phi\alpha$ for any α beginning with a.*

The set of all infinite sequences of natural numbers is known as the *Baire space*. The Baire space topology is defined by the following base. For every finite sequence a, the set of all inifinite sequences that have a as an initial segment is declared a basic open set. The continuity principle implies in particular that every function from a spread to the natural numbers is continuous with respect to the topology inherited from the Baire space topology. A function that, applied to some finite sequence, yields either the information that the sequence is too short for computing a result or else gives $\Phi\alpha$ for all sequences α that a is an initial segment of. Such a function is called a *neighborhood function*. The continuity principle reflects one part of Brouwer's intuition on infinitely proceeding sequences, namely that as the procession of the sequences cannot be predicted, any integer-valued operation acting on sequences can only use a finite amount of information in order to produce a result.

Bar induction

The other main principle of intuitionism is the principle of *bar induction*. Let S be a spread and $P \subseteq S$, then P is a *bar* for a if for every infinite sequence α in S that has a as a prefix, some prefix $a * b$ of α is in P. In that situation, we say a is *barred* with respect to P. The set P is called a bar if it is a bar for the empty sequence $\langle\rangle$.

Principle 1.2.5 (Bar Induction). *Let S be a spread and let P be a subset of S meeting the following properties.*

1. *For all finite sequences a, b, whenever $a \in P$ and $a * b \in S$ then $a * b \in P$.*
2. *For each $\alpha \in S$ there is a $k \in \mathbb{N}$ such that $\overline{\alpha}k \in P$.*
3. *For all $a \in S$, if all immediate S-descendents of a are in P, then so is a.*

Then $\langle\rangle \in P$ (and hence $a \in P$ for all $a \in S$).

The first property states that P is a monotone subset, the second property states that P is a bar and the third property states that P propagates towards the top in the described manner.

The most important direct application of the principle of bar induction is the *fan theorem*. The fan theorem is a positive formulation of König's Lemma. König's Lemma is a classical theorem that states that every infinite, finitely branching tree has an infinite path. While König's Lemma is intrinsically classical, its contrapositive variant, the fan theorem, makes more sense, constructively.

Theorem 1.2.6 (Fan). *Let S be a fan and let P be a subset of S. If P is a bar, then there is a natural number n such that for every infinite sequence in S, an initital segment with length less than n can be found that is an element of P.*

The continuity principle and the fan theorem can be used to show that all real-valued functions on the reals are continuous and all real-valued functions defined on a closed interval are uniformly continuous. Hence, in contrast to Bishop's mathematics, in intuitionism one does not have to distinguish between various forms of continuity but in fact *all* functions on the reals (or any complete separable metric space for that matter) are continuous in the strongest possible sense.

It is to be noted that the bar theorem is classically provable. The continuity principle, on the other hand, is expressedly inconsistent with classical logic in that it refutes instances of LPO and even LLPO.

Another consequence of the fan theorem is that every continuous, real-valued function defined on the unit interval that has only positive values, has a positive infimum. Bishop's mathematics alone is too weak to prove this, in fact, the existence of a uniformly continuous positive-valued function that has infimum zero is undecided in Bishop's mathematics.

This is an appetizer rather than a complete introduction to Intuitionism. Creative subject arguments for instance, will be completely neglected in this thesis. For our purposes we will unduly simplify matters and identify Intuitionism with Bishop's mathematics augmented with the continuity principle and the principle of bar induction.

1.2.3 Markov's constructive recursive mathematics

The third major school of constructivism is Markov's constructive recursive mathematics (CRM). We shall touch upon CRM only very briefly for the sake of completeness, as it is of less importance for the purposes of this thesis.

For $e, n \in \mathbb{N}$, we denote by $e.n$ the result of applying the e^{th} partial recursive function (with respect to some fixed admissible numbering) to the argument n. The formula $e.n\downarrow$ expresses that the e^{th} partial recursive function terminates when applied to n.

The two main principles governing CRM are *Church's Thesis* and *Markov's Principle.*

Principle 1.2.7 (Church's Thesis). *Assume $\forall x \in \mathbb{N} \exists y \in \mathbb{N}.\ A(x, y)$. Then there exists some $e \in \mathbb{N}$ such that $\forall x \in \mathbb{N}.\ e.x\downarrow \wedge A(x, e.x)$.*

Principle 1.2.8 (Markov's Principle). *Assume $\forall x \in \mathbb{N}.\ A(x) \vee \neg A(x)$ and $\neg\neg\exists x \in \mathbb{N}.\ A(x)$. Then $\exists x \in \mathbb{N}.\ A(x)$.*

Church's Thesis expresses that every operation on the natural numbers is effective. Markov's Principle has the consequence that an algorithm, whose non-termination is absurd, actually terminates.

Although Intuitionism contains theorems that are incompatible with classical logic, it is otherwise relatively close to classical mathematics in certain respects. This can hardly be said of CRM.

Not only does CRM refute LPO and LLPO, but whereas in INT it can only be shown that not every bounded, monotone real sequence converges, CRM actually allows to derive the *existence* of a so called *Specker sequence*, i.e. a bounded, increasing real sequence whose elements are eventually bounded away from every real (and which is hence divergent).

Another pathology that can be derived using Church's Thesis is the existence of a continuous, unbounded (and hence not uniformly continuous) real function defined on the unit interval.

On the positive side, in CRM it can be shown that any real-valued function defined on the reals (or any complete separable metric space) is continuous. A consequence of Markov's Principle is that apartness on the reals is just inequality (which cannot be proved in INT or BISH).

1.3 Formalized constructive logic

As we intend to describe models of constructive mathematics, we are forced to use formal systems of constructive mathematics. In order to facilitate the description of the models, we shall use explicit contexts for variable declarations and hypotheses in our formalization.

1.3.1 Constructive many-sorted predicate logic with equality

Signatures

A *signature* $\Sigma = (\mathcal{S}, \mathcal{F}, \mathcal{R}, \mathrm{ar})$ for many-sorted predicate logic consists of a set $\mathcal{S}$ of sorts, a set $\mathcal{F}$ of function symbols, a set $\mathcal{R}$ of relation symbols and arity functions $\mathrm{ar}_{\mathcal{F}} : \mathcal{F} \longrightarrow \mathcal{S}^* \times \mathcal{S}$ and $\mathrm{ar}_{\mathcal{R}} : \mathcal{R} \longrightarrow \mathcal{S}^*$. We will mostly refer to the sorts as *types*.

Terms

We assume that $\mathcal{V}$ is a countable set of variables. A *context* Γ is a finite list $\Gamma \equiv x_1 : s_1, \ldots, x_n : s_n$ where $x_i \in \mathcal{V}$ and $s_i \in \mathcal{S}$ and the x_i are pairwise distinct. The notation $\Gamma \vdash t : s$ means that t is a term of type s with respect to the variable declarations of the context Γ. This style of typing is called Church style.

The term formation rules are

$$\frac{}{\Gamma \vdash x : s}\,(x : s \in \Gamma) \qquad \frac{\Gamma \vdash t_1 : s_1 \quad \cdots \quad \Gamma \vdash t_n : s_n}{\Gamma \vdash f(t_1, \ldots, t_n) : s}\,(\mathrm{ar}_{\mathcal{F}}(f) = (s_1, \ldots, s_n, s))$$

Formulae

The set of atomic formulae is defined by the following rules.

$$\frac{\Gamma \vdash t : s \quad \Gamma \vdash t' : s}{\Gamma \vdash t =_s t' : \mathsf{Prop}} \qquad \frac{\Gamma \vdash t_1 : s_1 \quad \cdots \quad \Gamma \vdash t_n : s_n}{\Gamma \vdash R(t_1, \ldots, t_n) : \mathsf{Prop}}\,(\mathrm{ar}_{\mathcal{F}}(f) = (s_1, \ldots, s_n))$$

The set of all formulae is defined by the above and the following rules.

$$\frac{\Gamma \vdash \varphi : \mathsf{Prop} \quad \Gamma \vdash \psi : \mathsf{Prop}}{\Gamma \vdash \varphi \wedge \psi : \mathsf{Prop}} \qquad \frac{}{\Gamma \vdash \bot : \mathsf{Prop}}$$

$$\frac{\Gamma \vdash \varphi : \mathsf{Prop} \quad \Gamma \vdash \psi : \mathsf{Prop}}{\Gamma \vdash \varphi \to \psi : \mathsf{Prop}} \qquad \frac{\Gamma \vdash \varphi : \mathsf{Prop} \quad \Gamma \vdash \psi : \mathsf{Prop}}{\Gamma \vdash \varphi \vee \psi : \mathsf{Prop}}$$

$$\frac{\Gamma, x : s \vdash \varphi : \mathsf{Prop}}{\Gamma \vdash \forall x : s.\ \varphi : \mathsf{Prop}} \qquad \frac{\Gamma, x : s \vdash \varphi : \mathsf{Prop}}{\Gamma \vdash \exists x : s.\ \varphi : \mathsf{Prop}}$$

We use $\neg\varphi$ as a shorthand for $\varphi \to \bot$ and $\varphi \leftrightarrow \psi$ as a shorthand for $(\varphi \to \psi) \wedge (\psi \to \varphi)$.

Sequents

A sequent is an expression of the form $\Gamma \mid \Theta \vdash \varphi$ where Γ is a context, Θ is a list of propositions and φ is a proposition, both with respect to the context Γ. The sequent is to be interpreted as: under the variable declarations of Γ, the proposition φ follows from the propositions enlisted in Θ. If $\Gamma \vdash \varphi_1, \ldots, \Gamma \vdash \varphi_n, \Gamma \vdash \varphi$ are well formed propositions, then $\Gamma \mid \varphi_1 \cdots \varphi_n \vdash \varphi$ is a well formed sequent. Although one is primarily interested in single sentences, sequents provide a convenient notation for the fact that a proposition holds, subject to a set of variable declarations and hypotheses.

Proof rules

The set of provable sequents is defined inductively by the following rules.

The axiom rule:

$$\frac{}{\Gamma \mid \Theta \vdash \varphi}\,(ax) \quad (\text{if } \varphi \in \Theta)$$

Introduction and elimination rules for logical connectives:

$$\frac{\Gamma \mid \Theta \vdash \bot}{\Gamma \mid \Theta \vdash \varphi}\,(\bot E)$$

$$\frac{\Gamma \mid \Theta \vdash \varphi \quad \Gamma \mid \Theta \vdash \psi}{\Gamma \mid \Theta \vdash \varphi \wedge \psi}\,(\wedge I) \qquad \frac{\Gamma \mid \Theta \vdash \varphi \wedge \psi}{\Gamma \mid \Theta \vdash \varphi}\,(\wedge E_1) \qquad \frac{\Gamma \mid \Theta \vdash \varphi \wedge \psi}{\Gamma \mid \Theta \vdash \psi}\,(\wedge E_2)$$

$$\frac{\Gamma \mid \Theta \vdash \varphi \vee \psi \quad \Gamma \mid \Theta, \varphi \vdash \theta \quad \Gamma \mid \Theta, \psi \vdash \theta}{\Gamma \mid \Theta \vdash \theta}\,(\vee E) \qquad \frac{\Gamma \mid \Theta \vdash \varphi}{\Gamma \mid \Theta \vdash \varphi \vee \psi}\,(\vee I_1) \qquad \frac{\Gamma \mid \Theta \vdash \psi}{\Gamma \mid \Theta \vdash \varphi \vee \psi}\,(\vee I_2)$$

$$\frac{\Gamma \mid \Theta, \varphi \vdash \psi}{\Gamma \mid \Theta \vdash \varphi \to \psi}\,(\to I) \qquad \frac{\Gamma \mid \Theta \vdash \varphi \quad \Gamma \mid \Theta \vdash \varphi \to \psi}{\Gamma \mid \Theta \vdash \psi}\,(\to E)$$

Introduction and elimination rules for quantifiers:

$$\frac{\Gamma, x : s \mid \Theta \vdash \varphi}{\Gamma \mid \Theta \vdash \forall x : s.\, \varphi}\,(\forall I)\ \ (x \text{ not free in } \varphi) \qquad \frac{\Gamma \mid \Theta \vdash \forall x : s.\, \varphi \quad \Gamma \vdash t : s}{\Gamma \mid \Theta \vdash \varphi[t/x]}\,(\forall E)$$

$$\frac{\Gamma \mid \Theta \vdash \varphi[t/x] \quad \Gamma \vdash t : s}{\Gamma \mid \Theta \vdash \exists x : s.\, \varphi}\,(\exists I)$$

$$\frac{\Gamma \mid \Theta \vdash \exists x : s.\, \varphi \quad \Gamma, x : s \mid \Theta, \varphi \vdash \psi}{\Gamma \mid \Theta \vdash \psi}\,(\exists E)\ \ (x \text{ not free in } \Theta, \psi)$$

Rules for equality:

$$\frac{}{\Gamma \vdash x =_s x}\,(refl)\ \ (x : s \in \Gamma) \qquad \frac{\Gamma \mid \Theta \vdash t =_s t' \quad \Gamma \mid \Theta \vdash \varphi[t/x]}{\Gamma \mid \Theta \vdash \varphi[t'/x]}\,(repl)$$

Structural rules

The following structural rules are admissible

$$\frac{\Gamma \mid \Theta \vdash \psi}{\Gamma \mid \Theta, \varphi \vdash \psi}\,(w) \qquad \frac{\Gamma \mid \varphi_1, \ldots, \varphi_i, \varphi_{i+1}, \ldots, \varphi_n \vdash \psi}{\Gamma \mid \varphi_1, \ldots, \varphi_{i+1}, \varphi_i, \ldots, \varphi_n \vdash \psi}\,(ex) \qquad \frac{\Gamma \mid \Theta, \varphi, \varphi \vdash \psi}{\Gamma \mid \Theta, \varphi \vdash \psi}\,(c)$$

The above rules are called weakening, exchange and contraction. Analogues of the weakening and exchange rules with respect to the type theoretic context are admissible, too. In practice, we will often dispense with the explicit context in order to not distract too much attention from the actual contents of a formula or a sequent. Nevertheless, the chosen presentation of many-sorted predicate logic will motivate its categorical semantics.

1.3.2 Heyting arithmetic

A very basic system of arithmetic based on intuitionistic predicate logic with equality is *Heyting arithmetic* (HA). Its signature consists of a single type whose inhabitants are thought of as natural numbers. Furthermore, no relation symbols other than equality occur. Finally, the set of function symbols holds an element for each definition of a primitive recursive function. The set of axioms contains all definitional equations for these functions.[1] We shall denote by F_n the set of n-ary function symbols. The family $(F_n)_{n\in\mathbb{N}}$ is defined inductively by the following rules.

1. $0 \in \mathbf{F}_0$
2. $S \in \mathbf{F}_1$
3. $\mathrm{p}_k^{(m)} \in \mathbf{F}_m$ for $m, k \in \mathbb{N}, k < m$
4. If $t \in \mathbf{F}_n$ and $s_1, \ldots, s_n \in \mathbf{F}_m$ then $\mathrm{Comp}[t, s_1, \ldots, s_n] \in \mathbf{F}_m$ for $m, n \in \mathbb{N}$
5. If $t \in \mathbf{F}_n$ and $s \in \mathbf{F}_{n+2}$ then $\mathrm{R}[t, s] \in \mathrm{F}_{n+1}$

As for the axioms of Heyting arithmetic, the following equations describe the behaviour of the primitive recursive constructors.

$$\mathrm{p}_k^{(m)}(x_1, \ldots, x_m) = x_k$$
$$\mathrm{Comp}[t, s_1, \ldots, s_n](x_1, \ldots, x_m) = t(s_1(x_1, \ldots, x_m), \ldots, s_n(x_1, \ldots, x_m))$$
$$\mathrm{R}[t, s](x_1, \ldots, x_n, 0) = t(x_1, \ldots, x_n)$$
$$\mathrm{R}[t, s](x_1, \ldots, x_n, \mathrm{S}k) = s(x_1, \ldots, x_n, \mathrm{R}[t, s](x_1, \ldots, x_n, k), k)$$

The following axiom expresses that zero is not a successor.

$$0 \neq \mathrm{S}x$$

Finally, for every formula $A(x)$ there is an instance of the induction principle.

$$A(0) \wedge (\forall x.\ A(x) \to A(\mathrm{S}x)) \to \forall x.\ A(x)$$

The classical version of Heyting arithmetic, i.e. Heyting arithmetic augmented with *tertium non datur* is called *Peano arithmetic.*

[1] Heyting arithmetic can be alternatively presented using only 0, S, + and ×, and the definitional equations for + and ×. The presentation using symbols for *all* definitions of primitive recursive functions is a definitional extension of the former.

1.3.3 Elementary analysis

Heyting arithmetic provides no linguistic means for expressing properties of real numbers or infinite sequences of natural numbers. *Elementary analysis(EL)* is an extension of Heyting arithmetic that has two sorts: $\mathbb{N}$ and $\mathbb{N} \to \mathbb{N}$. In addition to the constants of Heyting arithmetic, elementary analysis has a constant r of arity $\mathbb{N} \times (\mathbb{N} \to \mathbb{N}) \times \mathbb{N} \longrightarrow \mathbb{N}$. In addition to the term formation rules of predicate logic, we have the rules of abstraction and application

$$\frac{\Gamma, x : \mathbb{N} \vdash t : \mathbb{N}}{\Gamma \vdash \lambda x.\, t : \mathbb{N} \to \mathbb{N}}\,(abs) \qquad \frac{\Gamma \vdash t : \mathbb{N} \to \mathbb{N} \quad \Gamma \vdash t' : \mathbb{N}}{\Gamma \vdash tt' : \mathbb{N}}\,(app)$$

As to equations, in addition to those of Heyting arithmetic, we have the type theoretic rules

$$\frac{\Gamma, x : \mathbb{N} \vdash t : \mathbb{N} \quad \Gamma \vdash t' : \mathbb{N}}{\Gamma \vdash (\lambda x.\, t)t' =_{\mathbb{N}} t[t'/x]}\,(\beta) \qquad \frac{\Gamma \vdash t : \mathbb{N} \to \mathbb{N}}{\Gamma \vdash \lambda x.\, tx =_{\mathbb{N}\to\mathbb{N}} t}\,(\eta) \quad (\text{for } x \notin \Gamma)$$

and the recursor equations

$$\frac{\Gamma \vdash t : \mathbb{N} \quad \Gamma \vdash u : \mathbb{N} \to \mathbb{N}}{\Gamma \vdash \mathsf{r}(t, u, 0) =_{\mathbb{N}} t} \qquad \frac{\Gamma \vdash t, t' : \mathbb{N} \quad \Gamma \vdash u : \mathbb{N} \to \mathbb{N}}{\Gamma \vdash \mathsf{r}(t, u, \mathrm{S}t') =_{\mathbb{N}} u(\mathsf{p}(\mathsf{r}(t, u, t'), t'))}$$

where $\mathsf{p} \in \mathbf{F}_2$ is a primitive recursive, surjective pairing function.

We extend the induction scheme to all formulae of EL and finally we assume a restricted axiom of choice

$$\forall x : \mathsf{N}\, \exists y : \mathsf{N}.\, A(x, y) \to \exists \alpha : \mathsf{N} \to \mathsf{N}\, \forall x.\, A(x, \alpha(x))$$

where $A(x, y)$ is a quantifier-free formula containing only equations of ground type.

We will denote the classical version of elementary analysis by $\mathrm{EL}^{(\mathrm{c})}$.

1.3.4 Heyting arithmetic with higher types

The finite type hierarchy over $\mathbb{N}$ is the set of types freely generated from $\mathbb{N}$ by means of $\times$ and $\to$. As a convention, $\times$ binds stronger than $\to$ and $\to$ associates to the right.

The set $\mathcal{S}$ of sorts of *extensional Heyting arithmetic with higher types* (E-HA$^\omega$) consists of the finite type hierarchy over $\mathbb{N}$. For $s, s' \in \mathcal{S}$, we have constants

$$\mathsf{p}_0^{(s,s')}, \qquad \mathsf{p}_1^{(s,s')}, \qquad \mathsf{p}^{(s,s')}$$

of type

$$s \times s' \to s, \qquad s \times s' \to s', \qquad s \to s' \to s \times s',$$

respectively. When the types are clear from the context or irrelevant, we feel free to omit the superscripts. Pairing and unpairing is governed by the equations

$$\mathsf{p}_0(\mathsf{p}(x,y)) = x, \qquad \mathsf{p}_1(\mathsf{p}(x,y)) = y, \qquad \mathsf{p}(\mathsf{p}_0(z),\mathsf{p}_1(z)) = z.$$

As in the case of elementary analysis, we extend the term formation rules of predicate logic with rules for abstraction and application,

$$\frac{\Gamma, x : s \vdash t : s'}{\Gamma \vdash \lambda x : s.\, t : s \to s'}\,(abs) \qquad\qquad \frac{\Gamma \vdash t : s \to s' \quad \Gamma \vdash t' : s}{\Gamma \vdash tt' : s'}\,(app)$$

accompanied by the respective equations.

$$\frac{\Gamma, x : s \vdash t : s' \quad \Gamma \vdash t' : s}{\Gamma \vdash (\lambda x : s.\, t)t' =_{s'} t[t'/x]}\,(\beta) \qquad\qquad \frac{\Gamma \vdash t : s \to s'}{\Gamma \vdash \lambda x.\, tx =_{s\to s'} t}\,(\eta) \quad (\text{for } x \notin \Gamma)$$

In Heyting arithmetic and elementary analysis, all definitions of primitive recursive functions were introduced at the level of function symbols. In contrast, in E-HA$^\omega$, we provide the constant $0 \in \mathbb{N}$, the the function symbol $S \in \mathbb{N} \to \mathbb{N}$ and for each $s \in \mathcal{S}$ the function symbol $\mathsf{r}_s : s \times (s \times \mathbb{N} \to s) \to (\mathbb{N} \to s)$. The constant r_s is a recursor over type s, governed by the equations

$$\frac{\Gamma \vdash t : s \quad \Gamma \vdash u : s \times \mathbb{N} \to s}{\Gamma \vdash \mathsf{r}(t,u)(0) =_s t} \qquad\qquad \frac{\Gamma \vdash t : s \quad \Gamma \vdash u : s \times \mathbb{N} \to s \quad \Gamma \vdash t' : \mathbb{N}}{\Gamma \vdash \mathsf{r}(t,u)(\mathrm{S}t') =_s u(\mathsf{r}(t,u)(t'), t')}$$

Finally we state the axiom that 0 is not a successor

$$0 \neq \mathrm{S}x$$

and the induction scheme

$$A(0) \wedge (\forall x.\ A(x) \to A(\mathrm{S}x)) \to \forall x.\ A(x)$$

This finishes the definition of E-HA$^\omega$. For the definition of HA$^\omega$, we simply drop the extensionality rule (η).

As previously mentioned, in our formulation of (E-)HA$^\omega$ (and EL) we have extended the term language of multi-sorted predicate logic in that we allow λ-abstraction. If we wanted to stay strictly within the realm of traditional predicate calculus, at least in the case of (E-)HA$^\omega$, we could have used combinators instead. We have decided not to do so, as λ-terms allow a more intuitive and natural notation. Note also that extensionality is commonly formulated as an axiom stating that equality on functions is equivalent to argumentwise equality. Our formulation featuring the η-rule of λ-calculus is easily seen to be equivalent.

The *level* of a type is inductively defined by

$$\text{level}(\mathbb{N}) = 0 \qquad \text{level}(s \times s') = \max(\text{level}(s), \text{level}(s'))$$

$$\text{level}(s \to s') = \max(\text{level}(s) + 1, \text{level}(s'))$$

Given some type s, we can introduce the type s^* of finite sequences of elements of s as a notational convenience. We set

$$\text{level}(s^*) = \text{level}(s).$$

Obviously, for each s, a bijection between s and s^* can be defined.

We could have defined elementary analysis to be the fragment of E-HA$^\omega$ that uses only types of level 0 and 1. This would allow for a slightly more natural notation and presentation. However, in order not to complicate the description of the various interpretations of EL to be defined in the next chapter, it is more convenient to keep the language as small as possible. As the extension of elementary analysis with all types of type level 0 and 1 is a *definitional* extension, we will freely use this extra notation.

The classical versions of E-HA$^\omega$ and HA$^\omega$ are called E-PA$^\omega$ and PA$^\omega$

1.3.5 Higher order arithmetic

All previously mentioned extensions of Heyting arithmetic are *conservative extensions*, i.e., they prove the same theorems of Heyting arithmetic as Heyting arithmetic itself. This is not the case for higher order arithmetic. Higher order arithmetic (HAH) results from E-HA$^\omega$ by adding the kind Ω of propositions as a type and extending the set of types to all types freely generated from $\mathbb{N}$ and Ω by means of $\times$ and $\to$. Functions and functional relations are linked by the *axiom of unique choice.*

$$(\forall x : s\ \exists! y : s'.\ A(x,y)) \;\longrightarrow\; (\exists f : s \to s'\ \forall x : s.\ A(x, f(x)))$$

The extensionality of entailment axiom, which equates logical equivalence and equality on the type of propositions, is not standardly assumed. Introducing propositions as a type allows for quantification over truth values and predicates. In particular, it allows one to define a predicate by quantifying over *all* predicates. Hence, in contrast to HA, EL and (E-)HA$^\omega$, the system HAH is impredicative.

1.4 Constructive Principles

In this section we are going to review some common priciples that are used in various systems of constructive mathematics. In section 1.4.4 we will examine the dependencies and inconsistencies between combinations of these principles.

1.4.1 Definitions

In the following, we will use the small greek letters to denote infinite sequences of numbers, small roman letters from the beginning of the alphabet to denote finite sequences of numbers and small roman letters from the end of the alphabet to denote natural numbers.

Notation

Let $\alpha, \beta : \mathbb{N} \to \mathbb{N}$, $a, b : \mathbb{N}^*$, $n : \mathbb{N}$. Then

$$\begin{array}{rcl}
\langle\rangle : \mathbb{N}^* & = & \text{the empty sequence.} \\
\langle n \rangle : \mathbb{N}^* & = & \text{the sequence of length 1 with entry } n. \\
a * \alpha : \mathbb{N} \to \mathbb{N} & = & \text{the concatenation of } a \text{ and } \alpha. \\
a * b : \mathbb{N}^* & = & \text{the concatenation of } a \text{ and } b. \\
a \frown n : \mathbb{N}^* & = & a * \langle n \rangle. \\
\mathrm{lth}(a) : \mathbb{N} & = & \text{the length of } a. \\
\overline{\alpha} n : \mathbb{N}^* & = & \text{the initial segment of length } n \text{ of } \alpha.
\end{array}$$

The statement $\alpha \in a$ means $\forall k < \mathrm{lth}(a).\ \alpha k = ak$. In other words, a finite sequence is identified with its set of infinite extensions.

For a set A of finite sequences we write $\alpha \in A$ to express that all initial segments of α are elements of A, in other words, $\alpha \in A$ if and only if $\forall k.\ \overline{\alpha} k \in A$.

A function $\gamma : \mathbb{N}^* \to \mathbb{N}$ describes a partial continuous operation from $\mathbb{N} \to \mathbb{N}$ to $\mathbb{N}$ in the following manner. We write

$$\gamma(\alpha) = n \quad \text{for} \quad \exists k.\ \gamma(\overline{\alpha} k) = n + 1 \wedge \forall i < k.\ \gamma(\overline{\alpha} i) = 0.$$

The meaning of $\gamma(\_)$ thus depends on whether the argument is a finite or infinite sequence. The value of $\gamma(\alpha)$, if it exists, depends only on some finite initial segment of α. If $\gamma(\overline{\alpha} k) = 0$, then $\overline{\alpha} k$ is not sufficiently long in order for γ to compute the result. The formula $\gamma(\alpha)\downarrow$ expresses that $\gamma(\_)$ terminates when applied to α, in other words, it expresses that there exists a $k \in \mathbb{N}$ such that $\gamma(\overline{\alpha} k) > 0$. The function γ is called a *neighborhood function* for the operation $\gamma(\_)$.

The definition of $\gamma(\_)$ can be used in order to let γ define a partial continuous operation from $\mathbb{N} \to \mathbb{N}$ to $\mathbb{N} \to \mathbb{N}$. We write

$$\gamma | \alpha = \beta \quad \text{for} \quad \forall n.\ \gamma(\langle n \rangle * \alpha) = \beta n.$$

For $e, n \in \mathbb{N}$, we denote by $e.n$ the result of applying the e^{th} partial recursive function (with respect to some fixed admissible numbering) to the argument n. The formula $e.n\downarrow$ expresses that the e^{th} partial recursive function terminates when applied to n. An alternative notation for $e.n$ is the Kleene-bracket notation $\{e\}(n)$.

Choice Principles

The axiom of choice

$\mathsf{AC}(X,Y) \qquad (\forall x : X\ \exists y : Y.\ A(x,y)) \longrightarrow \exists f : X \to Y\ \forall x : X.\ R(x, f(x))$

expresses that every total relation on $X \times Y$ admits a choice function. By $\mathsf{AC}(X)$ we denote that for all Y, $\mathsf{AC}(X,Y)$ holds. The principle $\mathsf{AC}(\mathbb{N})$ is called *countable choice* or *number choice*, the principle $\mathsf{AC}(\mathbb{N} \to \mathbb{N})$ is called *function choice*.

The axiom of unique choice

$\mathsf{AC!}(X,Y) \qquad (\forall x : X\ \exists! y : Y.\ A(x,y)) \longrightarrow \exists f : X \to Y\ \forall x : X.\ R(x, f(x))$

expresses that every functional relation on $X \times Y$ is the graph of a function.

The axiom of dependent choice

$$\begin{aligned}&\mathsf{DC}(X) \qquad (\forall x : X\ \exists y : X.\ A(x,y))\\ &\qquad \longrightarrow \forall x : X\ \exists f : \mathbb{N} \to X.\ f(0) = x \wedge \forall n : \mathbb{N}.\ R(f(n), f(n+1))\end{aligned}$$

allows to construct an infinite sequence of pairwise related elements of X starting with an arbitrary $x \in X$. By DC we denote that for all X, $\mathsf{DC}(X)$ holds. Dependent choice is more powerful than countable choice.

Continuity Principles

Continuity principles conflict with classical reasoning but are certainly compatible with constructive mathematics. Continuity principles state that every function or operation between some specified (classes of) spaces is continuous.

Let (X,d) and (X',d') be metric spaces. Then the continuity principles

$\mathsf{CP}(X,X')$ Every function from X to X' is ε-δ-continuous.

and

$\mathsf{CP}_{\mathsf{seq}}(X,X')$ Every function from X to X' is sequentially continuous.

state the continuity (resp. sequential continuity) of all functions from X to X'.

For the particular case that $X = \mathbb{N} \to \mathbb{N}$ and $X' = \mathbb{N}$, the following combinations of continuity with choice are widely used. The weak principle of continuity states that every operation from $\mathbb{N} \to \mathbb{N}$ to $\mathbb{N}$ is continuous.

$$\begin{aligned}&\mathsf{WC\text{-}N} \qquad (\forall \alpha : \mathbb{N} \to \mathbb{N}\ \exists n : \mathbb{N}.\ A(\alpha, n))\\ &\qquad \longrightarrow (\forall \alpha : \mathbb{N} \to \mathbb{N}\ \exists m, n : \mathbb{N}\ \forall \beta : \mathbb{N} \to \mathbb{N}.\ \beta \in \overline{\alpha}m \longrightarrow A(\beta, n))\end{aligned}$$

The strong principle of continuity states that every operation from $\mathbb{N} \to \mathbb{N}$ to $\mathbb{N}$ is given by a neighborhood function.

CONT $\quad (\forall \alpha : \mathbb{N} \to \mathbb{N}\ \exists n : \mathbb{N}.\ A(\alpha, n))$
$\quad \longrightarrow (\exists \gamma : \mathbb{N}^* \to \mathbb{N}\ \forall \alpha : \mathbb{N} \to \mathbb{N}.\ \gamma(\alpha) \downarrow \wedge A(\alpha, \gamma(\alpha)))$

By [TvD88a, 4.1.4], the continuity principle for α ranging over an arbitrary spread is derivable from the respective continuity principle stated for the full spread. Clearly, the strong principle of continuity entails the weak principle of continuity. The strong principle of continuity enables one to actually decide whether a given finite sequence contains enough information to compute the result for all its infinite extensions. We will meet a generalization of CONT in Theorem 2.1.18.

We shall now formulate two weakenings of WC-N. Let 2 be the two element set $\{0, 1\}$. The following principle of continuity states that every operation from the full binary fan $\mathbb{N} \to 2$ to $\mathbb{N}$ is continuous.

$\mathsf{WC_{cp}\text{-}N}$ $\quad (\forall \alpha : \mathbb{N} \to 2\ \exists n : \mathbb{N}.\ A(\alpha, n))$
$\quad \longrightarrow (\forall \alpha : \mathbb{N} \to 2\ \exists m, n : \mathbb{N}\ \forall \beta : \mathbb{N} \to 2.\ \beta \in \overline{\alpha} m \longrightarrow A(\beta, n))$

By [TvD88a, 4.7.5], the continuity principle for α ranging over an arbitrary fan is derivable from the respective continuity principle stated for the full binary fan.

Now, let $\mathbb{N}^+ = \{\alpha : \mathbb{N} \to 2 \mid \forall m, n.\, \alpha(m) = 1 \wedge m < n \to \alpha(n) = 0\}$, $\mathbb{N}^+$ is the one point compactification of $\mathbb{N}$ (in the sense of [BB85, (6.6)]). The following principle of continuity states that every operation from the fan $\mathbb{N}^+$ to $\mathbb{N}$ is continuous.[2]

$\mathsf{WC_{seq}\text{-}N}$ $\quad (\forall \alpha : \mathbb{N}^+\ \exists n : \mathbb{N}.\ A(\alpha, n))$
$\quad \longrightarrow (\forall \alpha : \mathbb{N}^+\ \exists m, n : \mathbb{N}\ \forall \beta : \mathbb{N}^+.\ \beta \in \overline{\alpha} m \longrightarrow A(\beta, n))$

It is actually sufficient to require continuity at the constant zero sequence.

$\mathsf{WC_{seq}\text{-}N'}$ $\quad (\forall \alpha : \mathbb{N}^+\ \exists n : \mathbb{N}.\ A(\alpha, n))$
$\quad \longrightarrow (\exists m, n : \mathbb{N}\ \forall \beta : \mathbb{N}^+.\ \beta \in 0^m \longrightarrow A(\beta, n))$

WC-N entails $\mathsf{WC_{cp}\text{-}N}$, as the full binary fan is a retract of the full spread. Likewise, $\mathsf{WC_{cp}\text{-}N}$ entails $\mathsf{WC_{seq}\text{-}N}$, as $\mathbb{N}^+$ is a retract of the full binary fan. The relationship between the variants of WC-N and CP will be further discussed in the sections 1.4.3 and 2.3.

Church's Thesis

In theoretical computer science, Church's thesis referes to the observation that a great many different models of computation give rise to the same class of computable functions. In constructive mathematics, however, Church's thesis is referred to as the principle that *all* operations on the natural numbers are computable. Although regarded a misnomer by many, we will adhere to this terminology.

CT $\quad (\forall n \exists m.\ A(n, m)) \to (\exists e \forall n.\ e.n \downarrow \wedge A(n, e.n))$

We will meet a generalization of CT in Theorem 2.1.6.

[2] To our knowledge, the first reference in literature, where a continuity principle w.r.t. $\mathbb{N}^+$ is used, is [BS03].

Bar Induction

The principle of bar induction equates two notions of well foundedness for trees. Amongst the various possible formulations, we choose to present the principle of *monotone bar induction*. For a complete discussion see e.g. [Dum77, chapter 3].

BI Let $P \subseteq \mathbb{N}^*$ such that
(1) $\forall a, b : \mathbb{N}^*.\ P(a) \to P(a * b)$
(2) $\forall a : \mathbb{N}^*.\ (\forall x : \mathbb{N}.\ P(a \frown x)) \to P(a)$
(3) $\forall \alpha : \mathbb{N} \to \mathbb{N}\ \exists n : \mathbb{N}.\ P(\overline{\alpha} n)$.
Then $P(\langle\rangle)$.

As is the case for continuity principles, the principle of bar induction relativized to a spread follows from the principle of bar induction for the full spread.

The Fan Theorem

The fan theorem is a consequence of the principle of bar induction. It expresses the Heine-Borel compactness of Cantor space.

FAN $$(\forall \alpha : \mathbb{N} \to 2\ \exists n : \mathbb{N}.\ A(\overline{\alpha} n)) \longrightarrow \exists m : \mathbb{N}\ \forall \alpha : \mathbb{N} \to 2\ \exists n : \mathbb{N}.\ n \leq m \wedge A(\overline{\alpha} n)$$

Again, the fan theorem for α ranging over an arbitrary fan follows from the above case.

Markov's Principle

One consequence of Markov's principle

M $$((\forall n : \mathbb{N}.\ A(n) \vee \neg A(n)) \wedge \neg\neg\exists n : \mathbb{N}.\ A(n)) \longrightarrow \exists n : \mathbb{N}.\ A(n)$$

is that an algorithm that cannot diverge, actually terminates.

Double Negation Shift

The double negation shift schema is

DNS $$(\forall x.\ \neg\neg A(x)) \longrightarrow \neg\neg\forall x.\ A(x),$$

where x is a variable of arbitrary type. Its main use is that in (E-)HA$^\omega$, one can derive the negative translation of some choice principle from that choice principle itself and DNS. This is due to the fact that the negative translation of a formula is equivalent to its double negation relative to DNS, see [Spe62].

Independence of Premise

The independence of premise schema is given by

IP $(\neg B \to \exists x.\, A(x)) \longrightarrow \exists x.\, \neg B \to A(x),$

where x is a variable of arbitrary type not free in B.

Boundedness Principle

A set $S \subseteq \mathbb{N}$ is called pseudobounded whenever for each sequence $\gamma : \mathbb{N} \to S$ there exists an $n \in \mathbb{N}$ such that for all $k \geq n$ the inequality $\gamma(k) < k$ holds. Hajime Ishihara introduced the principles

BD Every inhabited pseudobounded subset of $\mathbb{N}$ is bounded

and

BD-N Every countable pseudobounded subset of $\mathbb{N}$ is bounded,

where a set is called countable if it is the image of a sequence. The notion was introduced in [Ish92].

1.4.2 Status with respect to classcial mathematics

The principles introduced in the previous section fall into three categories. The first category is constituted by the continuity principles and Church's thesis. All these principles are incompatible with classical mathematics in that they refute instances of the principle of the excluded middle.

The second category is constituted by the choice principles, the principle of bar induction and the fan theorem. The choice principles are not derivable in classical mathematics, but compatible with classical mathematics. The principle of bar induction (and hence the fan theorem) can be proved in classical mathematics augmented with dependent choice (see [HK66]).

The last category consists of Markov's principle, double negation shift, independence of premise and the boundedness principle. These can be proven in (or rather, become trivial in the realm of) classical logic without choice.

1.4.3 Consequences of constructive principles in analysis

We shall now point out some of the consequences that the constructive principles introduced in section 1.4.1 have on constructive analysis.

Representations

In order to make these principles applicable to common mathematical objects, we have to define representations of these objects in terms of spreads and fans. A metric space (X, d) is a *complete separable metric space (csm)* if it is Cauchy-complete and there exists a dense sequence of elements of X.

Proposition 1.4.1 (Representation of csm-spaces). *Let (X, d) be a csm-space. Then there exist decidable predicates $T \subseteq \mathbb{N}^*$ and $R \subseteq \mathbb{N}^* \times \mathbb{N}^*$ and a map ρ from (the set of infinite sequences defined by) T to X such that*

1. *T is a spread*
2. *For $\alpha, \beta \in T$, $\rho(\alpha) = \rho(\beta)$ if and only if aRb for all finite initial segments a and b of α and β, respectively.*
3. *The map $\rho : T \to X$ is a quotient map with respect to the Baire-topology on T.*

Proof. [TvD88b, 7.2.3] and [TvD88b, 7.2.4]. □

See example 1.2.2 for a representation of the space of real numbers equipped with the euclidean topology meeting the properties stated in Proposition 1.4.1.

A metric space (X, d) is a *complete totally bounded metric space (ctb)* if it is Cauchy-complete and for each $\varepsilon > 0$ there exists a finitely indexed ε-net, i.e. a finitely indexed cover by ε-balls. Classically, a metric space is complete and totally bounded if and only if it is compact. Amongst all classically equivalent characterizations, completeness and total boundedness proves to be the most useful notion for constructive analysis. When (X, d) is a ctb-space, it can be represented by a fan.

Proposition 1.4.2 (Representation of ctb-spaces). *Let (X, d) be a ctb-space. Then there exist decidable predicates $T \subseteq \mathbb{N}^*$ and $R \subseteq \mathbb{N}^* \times \mathbb{N}^*$ and a map ρ from (the set of infinite sequences defined by) T to X such that*

1. *T is a fan*
2. *For $\alpha, \beta \in T$, $\rho(\alpha) = \rho(\beta)$ if and only if aRb for all finite initial segments a and b of α and β, respectively.*
3. *The map $\rho : T \to X$ is a quotient map with respect to the Baire-topology on T.*

Proof. [TvD88b, 7.4.2] and [TvD88b, 7.4.3]. □

See example 1.2.3 for a representation of the real unit interval equipped with the euklidean topology meeting the properties stated in Proposition 1.4.2.

Continuity principles

Theorem 1.4.3.

(i) $\mathsf{WC\text{-}N}$ *entails* $\mathsf{CP}(X, X')$ *for all complete separable metric spaces* X.

(ii) $\mathsf{WC_{cp}\text{-}N}$ *entails* $\mathsf{CP}(X, X')$ *for all complete totally bounded metric spaces* X

(iii) $\mathsf{WC_{seq}\text{-}N}$ *entails* $\mathsf{CP_{seq}}(X, X')$ *for all complete separable metric spaces* X

Proof. (i) See [TvD88b, 7.2.7]. Actually, in the reference given, the separability of X' is required. This requirement can be lifted by the following argument. As $\mathbb{R}$ is separable, the validity of the principle $\mathsf{CP}(X, \mathbb{R})$ holds. Now, $\mathsf{CP}(X, X')$ for arbitrary X' follows, as for each $f : X \to X'$ and $x_0 \in X$, the function $x \mapsto d'(f(x), f(x_0))$ is continuous by $\mathsf{CP}(X, \mathbb{R})$.). (ii) This follows from the proof of [TvD88b, 7.2.7] and the fact, that the representation can be defined on a fan, as shown in Proposition 1.4.2. (iii) One can use the proof for the implication $1 \Rightarrow 3$ in [BS03, Proposition 4.4]. The only application of $\mathsf{AC}(\mathbb{N} \to \mathbb{N}, \mathbb{N})$ in that proof can be replaced by $\mathsf{WC_{seq}}$. □

Conversely, $\mathsf{AC}(\mathbb{N} \to \mathbb{N}, \mathbb{N})$ and $\mathsf{CP}(\mathbb{N} \to \mathbb{N}, \mathbb{N})$ imply $\mathsf{WC\text{-}N}$, $\mathsf{AC}(\mathbb{N} \to 2, \mathbb{N})$ and $\mathsf{CP}(\mathbb{N} \to 2, \mathbb{N})$ imply $\mathsf{WC_{cp}\text{-}N}$, and $\mathsf{AC}(\mathbb{N}^+, \mathbb{N})$ and $\mathsf{CP}(\mathbb{N}^+, \mathbb{N})$ imply $\mathsf{WC_{seq}\text{-}N}$.

Church's thesis and Markov's principle

The famous Kreisel-Lacombe-Schoenfield-Tsejtin theorem states that

Theorem 1.4.4. CT *and* M *entail* $\mathsf{CP}(X, X')$ *for all complete separable metric spaces.*

Proof. See [TvD88b, 7.2.11]. □

Interestingly, the use of Markov's principle is essential, here, as CT alone is not sufficient in order to derive this result, see [Bee75].

On the other hand, CT has some consequences that one might find pathological or at least counterintuitive.

Theorem 1.4.5. *Under the assumption of* CT *it can be shown that there exist*

(i) a uniformly continuous function $f : [0, 1] \to \mathbb{R}$ *such that* $\forall x : \mathbb{R}.\ f(x) > 0$ *and* $\inf(f) = 0$.

(ii) a continuous function $f : [0, 1] \to \mathbb{R}$ *which is unbounded (and hence not uniformly continuous).*

Proof. See [TvD88a, 6.4.4] □

A consequence of Markov's principle is that in every metric space (X, d), if $x, y \in X$ are not equal (i.e. $\neg x = y$) then x and y are apart (i.e. $d(x, y) > 0$).

Bar induction and the fan theorem

The most notable consequences of the fan theorem (and hence of bar induction) in constructive analysis are

Theorem 1.4.6. *Under the assumption of the fan theorem*

(i) Every complete totally bounded space is Heine-Borel compact (i.e. for every open cover of X there exists a finitely indexed subcover).

(ii) Every continuous function defined on a ctb space is uniformly continuous.

(iii) For every continuous real valued function f defined on a ctb space X, if $\forall x : X.\ f(x) > 0$ then $\inf(f) > 0$.

Proof. (i) follows from the fact that by the fan theorem, the set of infinite sequences defined by any fan is Heine-Borel compact and from the fact that by Proposition 1.4.2, X is a continuous image of a fan. (ii) and (iii) follow directly from (i). □

The boundedness principle

The boundedness principle $\mathsf{BD\text{-}N}$ has a great number of interesting equivalents (see e.g. [Ish01]) in analysis such as

- Every sequentially continuous mapping of a separable metric space into a metric space is continuous.
- Every sequentially continuous mapping of a complete separable metric space into a metric space is continuous.
- Banach's inverse mapping theorem
- The open mapping theorem
- The closed graph theorem
- The Banach-Steinhaus theorem
- The Hellinger-Toeplitz theorem
- The sequential completeness of the space of test-functions

Remark 1.4.7. Note that in the presence of $\mathsf{BD\text{-}N}$, the principles $\mathsf{CP}(X, Y)$ and $\mathsf{CP}_{\mathsf{seq}}(X, Y)$ are equivalent for every separable metric space X.

1.4.4 Inconsistencies between constructive principles

Pure constructive mathematics can be extended with a variety of constructive principles. Some of these are, however, mutually incompatible. We will mention only those inconsistencies that are relevant for the rest of this thesis.

Theorem 1.4.8. *In E-HA$^\omega$, the following holds.*

(i) $\mathsf{CT} + \mathsf{AC}(\mathbb{N} \to \mathbb{N}, \mathbb{N}) \vdash \bot$

(ii) $\mathsf{CP}(\mathbb{N} \to \mathbb{N}, \mathbb{N}) + \mathsf{AC}((\mathbb{N} \to \mathbb{N}) \to \mathbb{N}, \mathbb{N}) \vdash \bot$

(iii) $\mathsf{FAN} + \mathsf{CT} \vdash \bot$

(iv) $\mathsf{IP} + \mathsf{CT} + \mathsf{M} \vdash \bot$

Proof. (i) see [TvD88b, 9.6.8], (ii) see [TvD88b, 9.6.11], (iii) see [TvD88a, 4.7.6] or 1.4.5 (ii) and 1.4.6 (ii), (iv) see [Tro98, 1.13(iii)] □

Remark 1.4.9. For (i) and (ii) in the previous theorem, the extensionality axiom is essential.

1.5 Categorical Logic

Although category theory [Mac71] has originated from algebra and topology, there is also a strong link between category theory and constructive logic. On the one hand, categories with suitable properties are models for various flavors of logic, on the other hand, syntactic entities can often be organized into a categorical structure.

Many-sorted predicate logic is interpreted by assigning to each sort s an object $[\![s]\!]$. A proposition

$$x_1 : s_1, \ldots, x_n : s_n \vdash \varphi$$

is interpreted as a subobject $[\![\varphi]\!]$ of $[\![s_1]\!] \times \cdots \times [\![s_n]\!]$. A sequence

$$x_1 : s_1, \ldots, x_n : s_n \mid \varphi_1, \ldots, \varphi_m \vdash \varphi$$

is validated if and only if the infimum $[\![\varphi_1]\!] \wedge \cdots \wedge [\![\varphi_m]\!]$ of the interpretations of subobjects $\varphi_1, \ldots, \varphi_m$ is contained in the subobject $[\![\varphi]\!]$ of $[\![s_1]\!] \times \cdots \times [\![s_n]\!]$. The interpretations of the logical connectives and quantifiers and type theoretic constructions can be conveniently and stringently described in categorical terms by means of limits, colimits and other adjoints. Therefore, if some category is suitable for the interpretation of some logic or type theory, the interpretation is determined up to isomorphism. This allows for a very concise description of categorical models. See [Pit00] for a gentle introduction to categorical logic and [Jac99] for an extensive source.

2 Realizability Interpretations

Realizability interpretations can be seen as concrete instantiation of the *Brouwer-Heyting-Kolmogorov* explanation of the intuitionistic meaning of the logical connectives. Although the BHK explanation leaves the basic notions of *proof* and *construction* open, one may instantiate these with concrete mathematical notions. See [Tro98] for an article on the state of the art in realizability and [Oos02] for a historical essay.

2.1 Formalized realizability interpretation

Concrete instantiations of the BHK interpretation can be formalized in various systems of constructive mathematics, thus giving rise to interesting metamathematical information like the admissibility of certain rules and equiconsistency results.

2.1.1 Numerical realizability

In this section we recall the definition of Kleene's numerical realizability as introduced in [Kle45] and some of its basic properties. In numerical realizability, the notions of proof object and construction are instantiated as numbers and partial recursive function application, respectively.

Convention 2.1.1. *In Heyting arithmetic, any disjunction $A \vee B$ can be equivalently expressed as $\exists x.\, (x = 0 \to A) \wedge (x > 0 \to B)$. As dropping $\vee$ as a primitive allows us to make some definitions more concise, we shall treat disjunction as a derived notion for the rest of this section.*

The underlying idea of the realizability interpretation is to declare a proof object for a disjunction to be a pair consisting of an indicator of which of the disjuncts is to be proved and a proof object for that disjunct. Likewise, the proof object for an existentially quantified statement is a pair consisting of an actual witness for the existence statement and a proof object for the statement, instantiated with this very witness.

Definition 2.1.2 (Numerical realizability).
The *numerical realizability interpretation* is an inductively defined translation of formulas A of Heyting arithmetic into formulas $e\ \underline{\mathtt{rn}}\ A$ of Heyting arithmetic, where e is a fresh variable not occuring in A,

$$
\begin{array}{lcl}
e \ \underline{\texttt{rn}} \ \bot & \equiv & \bot \\
e \ \underline{\texttt{rn}} \ t = s & \equiv & t = s \\
e \ \underline{\texttt{rn}} \ A \wedge B & \equiv & \mathsf{p}_0 \, e \ \underline{\texttt{rn}} \ A \ \wedge \ \mathsf{p}_1 \, e \ \underline{\texttt{rn}} \ B \\
e \ \underline{\texttt{rn}} \ A \to B & \equiv & \forall a. \ a \ \underline{\texttt{rn}} \ A \ \to \ e.a \downarrow \wedge \ e.a \ \underline{\texttt{rn}} \ B \\
e \ \underline{\texttt{rn}} \ \forall x. \ A & \equiv & \forall x. \ e.x \downarrow \wedge \ e.x \ \underline{\texttt{rn}} \ A \\
e \ \underline{\texttt{rn}} \ \exists x. \ A & \equiv & \mathsf{p}_1 \, e \ \underline{\texttt{rn}} \ A[\mathsf{p}_0 \, e/x]
\end{array}
$$

with $\mathsf{p}_0, \mathsf{p}_1$ being the primitive recursive projection functions w.r.t. some fixed primitive recursive paring function p.

Definition 2.1.3 (Numerical realizability combined with truth).
The *numerical realizability combined with truth* interpretation translates a formula A into a formula $e \ \underline{\texttt{rnt}} \ A$. Its definition differs from that of $\underline{\texttt{rn}}$ only in the clause for implication.

$$e \ \underline{\texttt{rnt}} \ A \to B \quad \equiv \quad (\forall a. \ a \ \underline{\texttt{rnt}} \ A \ \to \ e.a \downarrow \wedge \ e.a \ \underline{\texttt{rnt}} \ B) \wedge (A \to B)$$

We shall now single out some classes of formulas that have special properties with respect to the numerical realizability interpretation.

Definition 2.1.4 (Special classes of formulae of HA).

1. A formula is *negative* if it does not contain $\exists$.
2. A formula is *almost negative* if it contains $\exists$ only directly in front of prime formulae.
3. A formula A is *stable* if $\mathrm{HA} \vdash \neg\neg A \to A$.
4. The $\underline{\texttt{rn}}$-conservative class ($\mathrm{CC}(\underline{\texttt{rn}})$) consists of all formulae A such that whenever $B \to C$ is a subformula of A, then B is almost negative.

Negative formulas are equivalent to their negative translations. Therefore, Peano arithmetic is conservative over Heyting arithmetic with respect to negative formulas. Almost negative formulas are equivalent to their realizability interpretations, i.e. they are *self-realizing*. Another important aspect of almost negative formulas is the fact that the realizability interpretation translates formulas into the almost negative fragment of arithmetic. A formula is stable if and only if it is equivalent to a negated formula.

Proposition 2.1.5 (Self-realizing formulas).

(i) The formulae $e \ \underline{\texttt{rn}} \ A$ *and* $\exists e. \ e \ \underline{\texttt{rn}} \ A$ *are equivalent to almost negative formulae.*

(ii) If A is almost negative, then $\mathrm{HA} \vdash A \leftrightarrow (\exists e. \ e \ \underline{\texttt{rn}} \ A)$.

Proof. See [Tro73, 3.2.10, 3.2.11]. The realizer for (ii) is essentially an unbounded search algorithm. It is crucial, that every prime formula of Heyting arithmetic is decidable. □

The class of realizable formulas exceeds the class of theorems of HA. In particular, even formulas that are inconsistent with classical logic can be realized. The following axiomatization of realizable formulas was found by Troelstra.

Theorem 2.1.6 (Soundness and characterization of $\underline{\texttt{rn}}$).

(i) $\mathrm{HA} + \mathsf{ECT}_0 \vdash A \quad \Longleftrightarrow \quad \mathrm{HA} \vdash \exists e.\, e\ \underline{\texttt{rn}}\ A$

(ii) If A is closed, then

$\mathrm{HA} + \mathsf{ECT}_0 \vdash A \quad \Longleftrightarrow \quad$ *there exists* $n \in \mathbb{N}$ *such that* $\mathrm{HA} \vdash \bar{n}\ \underline{\texttt{rn}}\ A$*,*

where ECT_0 *is the extended Church's thesis*

$\mathsf{ECT}_0 \qquad (\forall x.\, D(x) \to \exists y.\, A(x,y)) \to (\exists e \forall x.\, D(x) \to e.x\downarrow \wedge\, A(x, e.x))$

for almost negative $D(x)$

and $\bar{n}$ *is the numeral associated with a natural number* n*. Moreover, the number* n *can be computed from a derivation of* A*.*

Proof. See [Tro98, 1.11] □

We will now take a look at the properties of the realizability combined with truth interpretation.

Proposition 2.1.7.

(i) $\mathrm{HA} \vdash (e\ \underline{\texttt{rnt}}\ A) \to A$

(ii) If $A \in \mathrm{CC}(\underline{\texttt{rn}})$ *then* $\mathrm{HA} \vdash e\ \underline{\texttt{rn}}\ A \leftrightarrow e\ \underline{\texttt{rnt}}\ A$*.*

(iii) If $A \in \mathrm{CC}(\underline{\texttt{rn}})$ *and* $\mathrm{HA} + \mathsf{ECT}_0 \vdash A$ *then* $\mathrm{HA} \vdash A$*.*

Proof. (i) is proved by induction on the structure of A. (ii) is proved by induction on the structure of A. The crucial case is implication. We assume that $\mathrm{HA} \vdash e\ \underline{\texttt{rn}}\ B \leftrightarrow e\ \underline{\texttt{rnt}}\ B$ and $\mathrm{HA} \vdash e\ \underline{\texttt{rn}}\ C \leftrightarrow e\ \underline{\texttt{rnt}}\ C$. We must show that $B \to C$ follows from $e\ \underline{\texttt{rn}}\ (B \to C)$. As A is in $\mathrm{CC}(\underline{\texttt{rn}})$, B is almost negative. From B we can conclude $\exists x.\, x\ \underline{\texttt{rn}}\ B$. If $x\ \underline{\texttt{rn}}\ B$ then $e.x\downarrow \wedge\, e.x\ \underline{\texttt{rn}}\ C$. By the induction hypothesis, $e.x\downarrow \wedge\, e.x\ \underline{\texttt{rnt}}\ C$ and by (i), C follows. (iii) By the axiomatization of realizability, $\mathrm{HA} \vdash \exists x.\, x\ \underline{\texttt{rn}}\ A$. By (ii), $\mathrm{HA} \vdash \exists x.\, x\ \underline{\texttt{rnt}}\ A$ and by (i), $\mathrm{HA} \vdash A$. □

In the light of the previous proposition, the characterization part of the following theorem is of course trivial.

Theorem 2.1.8 (Soundness and characterization of $\underline{\texttt{rnt}}$).

(i) $\mathrm{HA} \vdash A \quad\Longleftrightarrow\quad \mathrm{HA} \vdash \exists e.\ e\ \underline{\texttt{rnt}}\ A$

(ii) If A is closed, then

$\mathrm{HA} \vdash A \quad\Longleftrightarrow\quad$ *there exists* $n \in \mathbb{N}$ *such that* $\mathrm{HA} \vdash \bar{n}\ \underline{\texttt{rnt}}\ A$.

The number n can be computed from a derivation of A.

The soundness and characterization theorems hold also with respect to certain extensions of HA. In particular they hold for $\mathrm{HA} + \mathsf{M}$.

Proposition 2.1.9. *The soundness and characterization theorems hold mutatis mutandis for* $\mathrm{HA} + \mathsf{M}$ *instead of* HA.

Proof. See [Tro98, 1.12]. □

Corollary 2.1.10. $\mathrm{HA} + \mathsf{M} + \mathsf{ECT}_0$ *is conservative over* $\mathrm{HA} + \mathsf{M}$ *with respect to* $\mathrm{CC}(\underline{\texttt{rn}})$ *formulas.*

Proof. Assume $\mathrm{HA} + \mathsf{M} + \mathsf{ECT}_0 \vdash A$ for $A \in \mathrm{CC}(\underline{\texttt{rn}})$. It follows from the soundness and characterization theorem for $\mathrm{HA}+\mathsf{M}$ that $\mathrm{HA}+\mathsf{M} \vdash \exists e.\ e\ \underline{\texttt{rn}}\ A$. As $A \in \mathrm{CC}(\underline{\texttt{rn}})$, $\mathrm{HA} + \mathsf{M} \vdash \exists e.\ e\ \underline{\texttt{rnt}}\ A$ and thus $\mathrm{HA} + \mathsf{M} \vdash A$. □

Corollary 2.1.11. $\mathrm{HA} + \mathsf{M} + \mathsf{ECT}_0$ *is equiconsistent with* HA.

Proof. Assume $\mathrm{HA} + \mathsf{M} + \mathsf{ECT}_0 \vdash 0 = 1$. As $0 = 1$ is in $\mathrm{CC}(\underline{\texttt{rn}})$, by Corollary 2.1.10, $\mathrm{HA} + \mathsf{M} \vdash 0 = 1$. As $0 = 1$ is negative, $\mathrm{HA} \vdash 0 = 1$ follows from the negative translation (which validates M). □

As we have seen, the $\underline{\texttt{rn}}$-interpretation is a useful tool in order to obtain equiconsistency results. We shall now turn our attention to applications of the $\underline{\texttt{rnt}}$-interpretation. The $\underline{\texttt{rnt}}$-interpretation of a $\forall\exists$ quantifier combination reads as follows.

$$e\ \underline{\texttt{rnt}}\ \forall x \exists y.\ A(x,y) \quad\equiv\quad \forall x.\ e.x\downarrow \wedge\ \mathsf{p}_1(e.x)\ \underline{\texttt{rnt}}\ A(x, \mathsf{p}_0(e.x))$$

As a consequence of the soundness theorem, if HA proves a sentence $\forall x \exists y.\ A(x,y)$, then there is an $n \in \mathbb{N}$ such that $\mathrm{HA} \vdash \forall x.\ \bar{n}.x\downarrow \wedge\ \mathsf{p}_1(\bar{n}.x)\ \underline{\texttt{rnt}}\ A(x, \mathsf{p}_0(\bar{n}.x))$ and as $e\ \underline{\texttt{rnt}}\ A$ entails A, HA proves $\forall x.\ \bar{m}.x\downarrow \wedge\ A(x, \bar{m}.x)$ for $m = \lambda^* a.\ \mathsf{p}_0(n.a)$. One can view the formula $A(x,y)$ as specifying the set of acceptable outputs y for the input x. If one can prove constructively, that for each input x there exists an output y meeting the specification $A(x,y)$, then by the above reasoning there exists a program m together with a constructive correctness proof.

We summarize the direct consequences of the soundness of the $\underline{\texttt{rnt}}$-interpretation.

Corollary 2.1.12 (Admissible rules of HA**).**

(i) For A containing at most x, y free, if $\mathrm{HA} \vdash \forall x \exists y.\ A(x, y)$ *then there is some* $n \in \mathbb{N}$ *such that* $\mathrm{HA} \vdash \forall x.\ \bar{n}.x \downarrow \wedge A(x, \bar{n}.x)$.

(ii) For A containing at most x free, if $\mathrm{HA} \vdash \exists x.\ A(x)$ *then there is some* $n \in \mathbb{N}$ *such that* $\mathrm{HA} \vdash A(\bar{n})$.

(iii) For A, B closed, if $\mathrm{HA} \vdash A \vee B$ *then* $\mathrm{HA} \vdash A$ *or* $\mathrm{HA} \vdash B$.

Proof. Immediate consequences of Theorem 2.1.8. □

The properties (i), (ii) and (iii) of the above corollary are called *Church's rule*, the *explicit definability property (EDP)* and the *disjunction property (DP)*.[1] As all formalizations of arithmetic considered in this thesis possess Rosser sentences R, neither of these properties holds for their classical counterparts, as $R \vee \neg R$ is classically provable, but neither is R nor $\neg R$.

We could have as well used the $\underline{\texttt{rn}}$-interpretation in order to derive Church's rule for $\mathrm{HA} + \mathsf{ECT}_0$. The correctness of the algorithm n would in that case be provable only in $\mathrm{HA} + \mathsf{ECT}_0$ instead of HA. The provability of correctness in HA can be guaranteed, if we require $A \in \mathrm{CC}(\underline{\texttt{rn}})$. However, in that case ECT_0 would not contribute to the proof of $\forall x \exists y.\ A(x, y)$, as the latter formula is as well in $\mathrm{CC}(\underline{\texttt{rn}})$. We therefore prefer the $\underline{\texttt{rnt}}$-interpretation over the $\underline{\texttt{rn}}$-interpretation for the purpose of extracting algorithms from proofs, as it applies to general formulas A.

2.1.2 Function realizability

In this section we recall the definition of Kleene's function realizability as introduced in [Kle65, Kle69] and some of its basic properties. We will omit those proofs which parallel those of the previous section.

In function realizability, the notions of proof object and construction are instantiated as functions and partial continuous function application, respectively.

Convention 2.1.13. *As in Heyting arithmetic, in elementary analysis $\vee$ is definable from $\exists$ and will be treated as a derived notion in this section. Furthermore, we will regard equality on function type $\alpha = \beta$ as defined by $\forall x.\ \alpha x = \beta x$. This has the advantage that the prime formulae of elementary analysis with respect to this presentation are decidable. Finally, we will omit typing annotations and instead use roman letters for numbers and greek letters for infinite sequences.*

[1] As disjunction can be defined from existence, DP follows from EDP. The converse is also true for recursively enumerable extensions of HA, as shown in [Fri77].

Definition 2.1.14 (Function realizability).
The *function realizability interpretation* is an inductively defined translation of a formula A of elementary analysis into a formula $\alpha \ \underline{\mathtt{rf}}\ A$ of elementary analysis, where α is a fresh variable not occuring in A,

$$
\begin{array}{lcl}
\alpha \ \underline{\mathtt{rf}}\ \bot & \equiv & \bot \\
\alpha \ \underline{\mathtt{rf}}\ t = s & \equiv & t = s \\
\alpha \ \underline{\mathtt{rf}}\ A \wedge B & \equiv & \mathsf{j}_0\, \alpha \ \underline{\mathtt{rf}}\ A \ \wedge \ \mathsf{j}_1\, \alpha \ \underline{\mathtt{rf}}\ B \\
\alpha \ \underline{\mathtt{rf}}\ A \to B & \equiv & \forall \beta.\ \beta \ \underline{\mathtt{rf}}\ A \to \alpha|\beta \downarrow \wedge\ \alpha|\beta \ \underline{\mathtt{rf}}\ B \\
\alpha \ \underline{\mathtt{rf}}\ \forall x\ A & \equiv & \forall x.\ \alpha x \ \underline{\mathtt{rf}}\ A \\
\alpha \ \underline{\mathtt{rf}}\ \exists x\ A & \equiv & \alpha^+ \ \underline{\mathtt{rf}}\ A[\alpha 0/x] \\
\alpha \ \underline{\mathtt{rf}}\ \forall \beta\ A & \equiv & \forall \beta\ \alpha|\beta \downarrow \wedge\ \alpha|\beta \ \underline{\mathtt{rf}}\ A \\
\alpha \ \underline{\mathtt{rf}}\ \exists \beta\ A & \equiv & \mathsf{j}_1 \alpha \ \underline{\mathtt{rf}}\ A[\mathsf{j}_0 \alpha/\beta]
\end{array}
$$

with $\mathsf{j}_0, \mathsf{j}_1$ being the projection functions w.r.t. some fixed paring function j on $\mathbb{N} \to \mathbb{N}$ and α^+ being the tail of α (i.e. $\alpha = \langle \alpha 0 \rangle * \alpha^+$).

Definition 2.1.15 (Function realizability combined with truth).
The *function realizability combined with truth* interpretation translates a formula A into a formula $\alpha \ \underline{\mathtt{rft}}\ A$. Its definition differs from that of $\underline{\mathtt{rf}}$ only in the clause for implication.

$$\alpha \ \underline{\mathtt{rft}}\ A \to B \quad \equiv \quad (\forall \beta.\ \beta \ \underline{\mathtt{rft}}\ A \to \alpha|\beta \downarrow \wedge\ \alpha|\beta \ \underline{\mathtt{rft}}\ B) \wedge (A \to B)$$

Definition 2.1.16 (Special classes of formulae of EL).

1. A formula is *negative* if it does not contain $\exists$.
2. A formula is *almost negative* if it contains $\exists x$ and $\exists \alpha$ only directly in front of prime formulae.
3. A formula A is *stable* if $\mathrm{EL} \vdash \neg\neg A \to A$.
4. The $\underline{\mathtt{rf}}$-conservative class ($\mathrm{CC}(\underline{\mathtt{rf}})$) consists of all formulae A such that whenever $B \to C$ is a subformula of A, then B is almost negative.

Proposition 2.1.17 (Self realizing formulae).

(i) The formulae $\alpha \ \underline{\mathtt{rf}}\ A$ and $\exists \alpha.\ \alpha \ \underline{\mathtt{rf}}\ A$ are equivalent to almost negative formulae.

(ii) If A is almost negative, then $\mathrm{EL} \vdash A \leftrightarrow (\exists \alpha.\ \alpha \ \underline{\mathtt{rf}}\ A)$.

Proof. In order to see (ii), it is essential to notice that firstly, prime formulas are decidable and secondly, for every prime formula it is provable that if $A(\alpha)$ then there exists an $n \in \mathbb{N}$ such that for all $\beta \in \overline{\alpha}n$, $A(\beta)$ holds. If A is prime, then a realizer for $\exists \alpha . A$ can be defined essentially via an unbounded search through all infinite sequences of the form $a * \mathrm{const}_0$, where $a \in \mathbb{N}^*$ and const_0 is the infinite constant zero sequence. $\square$

Theorem 2.1.18 (Soundness and characterization of $\underline{\texttt{rf}}$).

(i) $\mathrm{EL} + \mathsf{GC} \vdash A \quad \Longleftrightarrow \quad \mathrm{EL} \vdash \exists\alpha.\ \alpha\ \underline{\texttt{rf}}\ A$

(ii) If A is closed, then
$\mathrm{EL}{+}\mathsf{GC} \vdash A \quad \Longleftrightarrow \quad$ *there exists* $n \in \mathbb{N}$ *such that* $\mathrm{EL} \vdash \{\bar{n}\}\ \textit{total} \wedge \{\bar{n}\}\ \underline{\texttt{rf}}\ A$,

where GC *is the generalized continuity principle*

$$\mathsf{GC} \qquad (\forall\alpha.\ D(\alpha) \to \exists\beta.\ A(\alpha,\beta)) \to (\exists\gamma\,\forall\alpha.\ D(\alpha) \to \gamma.\alpha\downarrow \wedge\, A(\alpha,\gamma.\alpha))$$

for almost negative $D(\alpha)$,

$\bar{n}$ *is the numeral associated with the natural number* n *and* $\{k\}$ *is the* k^{th} *partial recursive function. Moreover, the number* n *can be computed from a derivation of* A.

Proposition 2.1.19.

(i) $\mathrm{EL} \vdash (\alpha\ \underline{\texttt{rft}}\ A) \to A$

(ii) If $A \in \mathrm{CC}(\underline{\texttt{rf}})$ *then* $\mathrm{EL} \vdash \alpha\ \underline{\texttt{rf}}\ A\ \leftrightarrow\ \alpha\ \underline{\texttt{rft}}\ A$.

(iii) If $A \in \mathrm{CC}(\underline{\texttt{rf}})$ *and* $\mathrm{EL} + \mathsf{GC} \vdash A$ *then* $\mathrm{EL} \vdash A$.

Theorem 2.1.20 (Soundness and characterization of $\underline{\texttt{rft}}$).

(i) $\mathrm{EL} \vdash A \quad \Longleftrightarrow \quad \mathrm{EL} \vdash \exists\alpha.\ \alpha\ \underline{\texttt{rft}}\ A$

(ii) If A is closed, then
$\mathrm{EL} \vdash A \quad \Longleftrightarrow \quad$ *there exists* $n \in \mathbb{N}$ *such that* $\mathrm{EL} \vdash \{\bar{n}\}\ \mathrm{total}\ \wedge\ \{\bar{n}\}\ \underline{\texttt{rft}}\ A$.

The number n *can be computed from a derivation of* A.

The soundness and characterization theorems hold also with respect to certain extensions of EL, in particular for $\mathrm{EL} + \mathsf{BI}$.

Proposition 2.1.21. *The soundness and characterization theorems hold mutatis mutandis for* $\mathrm{EL} + \mathsf{BI}$ *instead of* EL.

Proof. It is stated in [Tro98, 2.9] that soundness and characterization hold in presence of BI_D (see *loc. cit.*). As on the one hand, BI implies BI_D, and on the other hand BI_D and GC imply BI (see [Dum77, Theorem 3.8]), the claim follows. □

We skip the equiconsistency results for intuitionism that can be obtained via the $\underline{\texttt{rf}}$-interpretation and summarize the direct consequences of the soundness of the $\underline{\texttt{rnt}}$-interpretation.

Corollary 2.1.22 (Admissible rules of EL).

(i) *For A containing at most α, β free, if* $\mathrm{EL} \vdash \forall\alpha\exists\beta.\ A(\alpha, \beta)$ *then there is some* $n \in \mathbb{N}$ *such that* $\mathrm{EL} \vdash \{\bar{n}\}\ \mathrm{total}\ \wedge\ \forall\alpha.\ \{\bar{n}\}|\alpha\downarrow \wedge A(\alpha, \{\bar{n}\}|\alpha)$.

(ii) *For A containing at most α free, if* $\mathrm{EL} \vdash \exists\alpha.\ A(\alpha)$ *then there is some* $n \in \mathbb{N}$ *such that* $\mathrm{EL} \vdash A(\{\bar{n}\})$.

(iii) *For A, B closed, if* $\mathrm{EL} \vdash A \vee B$ *then* $\mathrm{EL} \vdash A$ *or* $\mathrm{EL} \vdash B$.

Proof. Immediate consequences of Theorem 2.1.20. □

2.1.3 Classically provably realizable formulas

In this section we will examine classes of formulas, whose realizability can be proved classically.

Classically provably numerically realizable formulas of HA

The class of formulas whose $\underline{\mathtt{rn}}$-realizability can be proved constructively was axiomatized in Theorem 2.1.6. The following proposition gives an answer to the question of what more formulas can be proved realizable if we use classical arithmetic. As classical arithmetic does not have the explicit definability property, two separate questions arise.

Proposition 2.1.23.

(i) $\mathrm{PA} \vdash \exists n.\ n\ \underline{\mathtt{rn}}\ A$ *if and only if* $\mathrm{HA} + \mathsf{ECT}_0 + \mathsf{M} \vdash \neg\neg A$

(ii) *Let A be closed.*
Then $\mathrm{PA} \vdash \bar{n}\ \underline{\mathtt{rn}}\ A$ *for some* $n \in \mathbb{N}$ *if and only if* $\mathrm{HA} + \mathsf{ECT}_0 + \mathsf{M} \vdash A$

Proof. (i) see [Tro73, 3.2.25]. (ii) The formula $\bar{n}\ \underline{\mathtt{rn}}\ A$ is almost negative, hence there exists a negative formula B such that $\mathrm{HA} + \mathsf{M} \vdash (\bar{n}\ \underline{\mathtt{rn}}\ A) \leftrightarrow B$. As PA is conservative over HA with respect to negative formulas, we have $\mathrm{HA} \vdash B$ and thus $\mathrm{HA} + \mathsf{M} \vdash \bar{n}\ \underline{\mathtt{rn}}\ A$. By the axiomatization part of Proposition 2.1.9, this is equivalent to $\mathrm{HA} + \mathsf{ECT}_0 + \mathsf{M} \vdash A$. For the reverse direction, the existence of some $n \in \mathbb{N}$ such that $\mathrm{HA} + \mathsf{M} \vdash \bar{n}\ \underline{\mathtt{rn}}\ A$ follows from the soundness part of Proposition 2.1.9. Consequently, $\mathrm{PA} \vdash \bar{n}\ \underline{\mathtt{rn}}\ A$. □

The answer to (i) of the above question is not actually an axiomatization of the class of provably realizable formulas but of the class of their double negations. We are, however, more interested in the case described in (ii), namely, the class of formulas, whose realizability can be proven for a concretely given realizer, as opposed to the class of formulas for which the existence of a realizer is classically provable.

By definition, the $\underline{\texttt{rnt}}$-interpretations entails truth, i.e.

$$\mathrm{HA} \vdash (\exists e.\ e\ \underline{\texttt{rnt}}\ A) \to A$$

so asking for an axiomatization of $\underline{\texttt{rnt}}$-realizable formulas is of course pointless, as long as we interpret in HA. But this is no longer so if we interpret in PA. The class of those formulas, whose $\underline{\texttt{rnt}}$-realizability can be proven classically has the property that $\forall\exists$ quantifier combinations are required to be algorithmically tractable, yet the corresponding correctness proof is only required to be classical. Put differently, the class of formulas is strictly larger than the class of theorems of HA, yet enjoys Church's rule, EDP and DP and stays within the bounds of PA. This property makes this class of formulas interesting from a classical, recursion theoretic point of view. As far as we know, the following is an open problem.

Problem 2.1.24. *Axiomatize the class of closed formulae A such that*

$$\mathrm{PA} \vdash \bar{n}\ \underline{\texttt{rnt}}\ A \textit{ for some } n \in \mathbb{N}.$$

The axiomatization problem can be naturally extended to open formulas. We wish to give an axiomatization for the class of formulas such that

$$\mathrm{PA} \vdash \forall \vec{x}.\ \bar{n}.\vec{x} \downarrow \wedge\ \bar{n}.\vec{x}\ \underline{\texttt{rnt}}\ A \textit{ for some } n \in \mathbb{N},$$

where $\vec{x}$ is the list of free variables in A.

In the following we shall give examples of realizable and non-realizable formulas and give axiomatization with respect to special classes of formulas.

Lemma 2.1.25. $\mathrm{HA} \vdash (e\ \underline{\texttt{rnt}}\ \neg A) \leftrightarrow \neg A$

Proof. By definition, $e\ \underline{\texttt{rnt}}\ \neg A$ if and only if $(\forall a.\ (a\ \underline{\texttt{rnt}}\ A) \to (e.a \downarrow \wedge\ e.a\ \underline{\texttt{rnt}}\ \bot)) \wedge \neg A$, which is equivalent to $(\forall a.\ (a\ \underline{\texttt{rnt}}\ A) \to \bot) \wedge \neg A$. As $(a\ \underline{\texttt{rnt}}\ A) \to A$, this is equivalent to $\neg A$. □

Proposition 2.1.26.

(i) There exists an $n \in \mathbb{N}$ such that $\mathrm{PA} \vdash \bar{n}\ \underline{\texttt{rnt}}\ M$ *for each instance M of Markov's principle.*

(ii) There exists an $n \in \mathbb{N}$ such that $\mathrm{PA} \vdash \bar{n}\ \underline{\texttt{rnt}}\ D$ *for each instance D of the double negation shift principle.*[2]

[2] I would like to gratefully acknowledge helpful suggestions regarding the axiomatization problem made by Jaap van Oosten and Dana Scott. Dana Scott suggested to add the double negations of all theorems of Peano arithmetic as axioms, Jaap van Oosten suggested to add principles that guarantee that every stable formula is equivalent to a negative formula. The double negation shift principle accomplishes the latter. The former is a direct consequence as, relative to DNS, the double negation of a formula is equivalent to its negative translation.

(iii) There is an instance I of IP *such that for all $k \in \mathbb{N}$, PA $\nvdash \bar{k}$ $\underline{\texttt{rnt}}$ I.*

(iv) There is an instance L of LPO *such that for all $k \in \mathbb{N}$, PA $\nvdash \bar{k}$ $\underline{\texttt{rnt}}$ L.*

Proof. (i) Let $M \equiv ((\forall x.\ A(n) \vee \neg A(x)) \wedge \neg\neg\exists x.\ A(x)) \to \exists x.\ A(x)$ and assume e_1 $\underline{\texttt{rnt}}$ $\forall x.\ A(x) \vee \neg A(x)$ and e_2 $\underline{\texttt{rnt}}$ $\neg\neg\exists x.\ A(x)$. The former gives rise to some e such that $\forall x.\ (\mathsf{p}_0(e.x) = 0 \to \mathsf{p}_1(e.x)$ $\underline{\texttt{rnt}}$ $A) \wedge (\mathsf{p}_0(e.x) > 0 \to \mathsf{p}_1(e.x)$ $\underline{\texttt{rnt}}$ $\neg A)$ and the latter is equivalent to $\neg\neg\exists x.\ A(x)$ by Lemma 2.1.25. A realizer can be constructed by unbounded search using e as a halting condition. The termination of this algorithm is provable in PA. Finally, M itself is provable in PA. Moreover, n can be chosen independently of the particular instance of M.
(ii) Let $D \equiv (\forall x.\ \neg\neg A) \to (\neg\neg\forall x.\ A)$. As e $\underline{\texttt{rnt}}$ $\neg\neg\forall x.\ A$ is equivalent to $\neg\neg\forall x.\ A$ by Lemma 2.1.25, and as D is a theorem of PA, any code for a provably total recursive function will do as a realizer.
(iii) We can use the same instance of IP that is used in [Tro98, 1.13(iii)] in order to show the inconsistency of $\mathsf{CT} + \mathsf{M} + \mathsf{IP}$, namely $((\neg\neg\exists y.\ Txxy) \to \exists z.\ Txxz) \to (\exists z.\ (\neg\neg\exists y.\ Txxy) \to Txxz)$. As the hypothesis is an instance of M, the universal closure of the conclusion, $\forall x \exists z.\ (\neg\neg\exists y.\ Txxy) \to Txxz)$ would be realizable. A realizer would however give rise to a decision algorithm for the halting set, which is absurd.
(iv) The formula $(\exists y.\ Txxy) \vee (\neg\exists y.\ Txxy)$ is an instance of LPO. A realizer for $\forall x.\ (\exists y.\ Txxy) \vee (\neg\exists y.\ Txxy)$ would again give rise to a decision algorithm for the halting set. □

Markov's principle and the double negation shift principle are validated by the interpretation under discussion. In HA, every negative formula is almost negative and stable. Markov's principle can be thought of as the principle responsible for making every almost negative formula equivalent to a negative formula, while the double negation shift principle can be thought of as making every stable formula equivalent to a negative formula. It is not known, whether these constitute a complete axiomatization, although it seems very unlikely.

The axiomatization problem appears to be very hard. The main obstacle may be that, whereas in the case of $\underline{\texttt{rn}}$-realizability the formula x $\underline{\texttt{rn}}$ A is always almost negative, we are not able to abstractly describe the class of formulas that are equivalent to a formula of the form x $\underline{\texttt{rnt}}$ A.

The following proposition states that $\mathrm{HA} + \mathsf{M} + \mathsf{DNS}$ *is* a complete axiomatization for two (however) very restricted classes of formulae.

Proposition 2.1.27. *Let A be closed.*

(i) If A is stable, then there exists $n \in \mathbb{N}$ such that PA $\vdash \bar{n}$ $\underline{\texttt{rnt}}$ A if and only if PA $\vdash A$ if and only if HA + $\mathsf{DNS} \vdash A$.

(ii) If $A \in \mathrm{CC}(\underline{\texttt{rn}})$, then there exists $n \in \mathbb{N}$ such that PA $\vdash \bar{n}$ $\underline{\texttt{rnt}}$ A if and only if HA + $\mathsf{M} \vdash A$.

Proof. (i) The first equivalence follows immediately from Lemma 2.1.25, as A is equivalent to a negated formula. The second equivalence follows from the fact that $\mathrm{HA} + \mathsf{DNS} \vdash A^G \leftrightarrow \neg\neg A$, where A^G is the negative translation of A.
(ii) As $A \in \mathrm{CC}(\underline{\mathtt{rn}})$, $\mathrm{PA} \vdash \bar{n}\ \underline{\mathtt{rnt}}\ A$ if and only if $\mathrm{PA} \vdash \bar{n}\ \underline{\mathtt{rn}}\ A$. By Proposition 2.1.23, this is the case if and only if $\mathrm{HA} + \mathsf{M} + \mathsf{ECT}_0 \vdash A$. As $\mathrm{HA} + \mathsf{M} + \mathsf{ECT}_0$ is conservative over $\mathrm{HA} + \mathsf{M}$ with respect to $\mathrm{CC}(\underline{\mathtt{rn}})$ by Corollary 2.1.10, $\mathrm{HA} + \mathsf{M} \vdash A$. The reverse direction follows from the soundness of the $\underline{\mathtt{rnt}}$-interpretation and the fact that M is realized. □

We have thus far examined, in how far theorems of PA are realized. As $\underline{\mathtt{rnt}}$-realizability entails truth, principles that are not theorems of PA are not realizable. In fact, by (i) of the previous proposition, negations of instances of principles that are refuted by PA *are* realized. In particular, there are instances E of CT such that $\neg E$ is realized.

Classically provably function realizable formulas of EL

An analogous set of questions can be posed for the $\underline{\mathtt{rf}}$- and $\underline{\mathtt{rft}}$-interpretations. We will omit those proofs which closely resemble the respective proofs for numerical realizability, which can be found in the previous section.

Proposition 2.1.28.

(i) $\mathrm{EL}^{(c)} \vdash \exists\alpha.\ \alpha\ \underline{\mathtt{rf}}\ A$ *if and only if* $\mathrm{EL} + \mathsf{GC} + \mathsf{M} \vdash \neg\neg A$

(ii) Let A be closed.
Then $\mathrm{EL}^{(c)} \vdash \{\bar{n}\}$ total $\wedge$ $\{\bar{n}\}\ \underline{\mathtt{rf}}\ A$ *for some* $n \in \mathbb{N}$
if and only if $\mathrm{EL} + \mathsf{GC} + \mathsf{M} \vdash A$.

Proof. It is crucial to notice that also in EL, the principle M suffices to show that every almost negative formula is equivalent to a negative formula. The rest is analogous to the proof of Proposition 2.1.23. □

Problem 2.1.29. *Axiomatize the class of closed formulae A such that*

$$\mathrm{EL}^{(c)} \vdash \{\bar{n}\} \text{ total } \wedge\ \{\bar{n}\}\ \underline{\mathtt{rf}}\ A \text{ for some } n \in \mathbb{N}.$$

As in problem 2.1.24, the axiomatization problem can be naturally extended to open formulas.

Proposition 2.1.30.

(i) There exists an $n \in \mathbb{N}$ such that $\mathrm{EL}^{(c)} \vdash \{\bar{n}\}$ total $\wedge$ $\{\bar{n}\}\ \underline{\mathtt{rft}}\ M$ *for each instance M of Markov's principle.*

(ii) There exists an $n \in \mathbb{N}$ such that $\mathrm{EL}^{(c)} \vdash \{\bar{n}\}$ total $\wedge$ $\{\bar{n}\}\ \underline{\mathtt{rft}}\ D$ *for each instance D of the double negation shift principle.*

(iii) *There is an instance I of* IP *such that for all $k \in \mathbb{N}$,* $\mathrm{EL}^{(c)} \nvdash \{\bar{k}\}$ total $\wedge$ $\{\bar{k}\}\ \underline{\texttt{rft}}\ I$.

(iv) *There is an instance L of* LPO *such that for all $k \in \mathbb{N}$,* $\mathrm{EL}^{(c)} \nvdash \{\bar{k}\}$ total $\wedge$ $\{\bar{k}\}\ \underline{\texttt{rft}}\ L$.

Proposition 2.1.31. *Let A be closed.*

(i) *If A is stable, then* $\mathrm{EL}^{(c)} \vdash \{\bar{n}\}\ \underline{\texttt{rft}}\ A$ *for some $n \in \mathbb{N}$ if and only if* $\mathrm{EL}^{(c)} \vdash A$ *if and only if* $\mathrm{EL} + \mathsf{DNS} \vdash A$.

(ii) *If $A \in \mathrm{CC}(\underline{\texttt{rf}})$, then* $\mathrm{EL}^{(c)} \vdash \{\bar{n}\}\ \underline{\texttt{rft}}\ A$ *for some $n \in \mathbb{N}$ if and only if* $\mathrm{EL}+M \vdash A$.

Again, the class of formulae under discussion here is strictly larger than EL, satisfies Church's rule, the continuity rule, EDP and DP, but stays within the bounds of $\mathrm{EL}^{(c)}$. It is worth noticing that it also includes BD-N.

Proposition 2.1.32. *Every instance of the boundedness principles* BD-N *is classically* $\underline{\texttt{rft}}$*-realized.*

Proof. The boundedness principle can be formalized as

$$(\forall\alpha\exists n\forall k > n.\ \beta(\alpha(k)) < k) \rightarrow (\exists N\forall x.\ \beta(x) < N)$$

Let $H \equiv \forall\alpha\exists n\forall k > n.\ \beta(\alpha(k)) < k$ and $C \equiv \exists N\forall x.\ \beta(x) < N$. Note that the boundedness principle $H \rightarrow C$ is an implication between the two $\mathrm{CC}(\underline{\texttt{rf}})$ formulas H and C. By Proposition 2.1.19 (ii),

$$\begin{array}{ll} & \gamma\ \underline{\texttt{rft}}\ H \rightarrow C \\ \equiv & (\forall\delta.\ \delta\ \underline{\texttt{rft}}\ H \rightarrow \gamma|\delta\downarrow \wedge\ \gamma|\delta\ \underline{\texttt{rft}}\ C) \wedge (H \rightarrow C) \\ \Leftrightarrow & (\forall\delta.\ \delta\ \underline{\texttt{rf}}\ H \rightarrow \gamma|\delta\downarrow \wedge\ \gamma|\delta\ \underline{\texttt{rf}}\ C) \wedge (H \rightarrow C) \\ \Leftrightarrow & (\gamma\ \underline{\texttt{rf}}\ (H \rightarrow C)) \wedge (H \rightarrow C) \end{array}$$

By the above, there exists an $n \in \mathbb{N}$ such that $\mathrm{EL}^{(c)}$ proves that $\{\bar{n}\}$ $\underline{\texttt{rft}}$-realizes (the universal closure of) BD-N if and only if there exists an $n \in \mathbb{N}$ such that $\mathrm{EL}^{(c)}$ proves that $\{\bar{n}\}$ $\underline{\texttt{rf}}$-realizes BD-N and $\mathrm{EL}^{(c)}$ proves BD-N. By Proposition 2.1.28, the former is the case if and only if EL + GC + M proves BD-N. As shown in [Ish92, Proposition 3], EL + WC-N suffices to prove BD-N. The latter is also the case, as BD-N is a theorem of $\mathrm{EL}^{(c)}$. The classical proof of BD-N requires an instance of the axiom of choice for prime formulas, which is contained in the definition of EL. □

Corollary 2.1.33. *Let A be a formula of the form $B \rightarrow C$, where B and C are in $\mathrm{CC}(\underline{\texttt{rf}})$. Then there exists an $n \in \mathbb{N}$ such* $\mathrm{EL}^{(c)}$ *proves that $\{n\}$ $\underline{\texttt{rft}}$-realizes the universal closure of A if and only if A is a theorem of both* EL + GC + M *and* $\mathrm{EL}^{(c)}$.

Proof. The claim follows from the proof of Proposition 2.1.32. □

2.2 Categorical Realizability Semantics

One might give realizability interpretations for E-HA$^\omega$ and higher order arithmetic in the same style as for HA and EL, now. We shall, however, take a different route. We define categories of sets equipped with a realizability structure. These categories will, as we shall observe, have good properties that allow for the interpretation of constructive arithmetic. The advantage is that, once we have defined the category, the rest of the interpretation is fixed up to isomorphism. This allows for a more concise description of realizability models, compared to those given in 2.1.2 and 2.1.14, where we needed a clause for every logical connective and quantifier. If one wants to retain the metamathematical flavor of realizability interpretation one can express the validity in the categorical model in a sufficiently expressive logical system.

Mind that a complete change of paradigm happens, here. Whereas in the previous section we described the interpretation of every feature of some formal system of arithmetic, we will now define the categorical model first and then observe which features and hence which formal system of arithmetic finds an interpretation in this model. In particular, we will give a characterization of those categorical models that allow for an interpretation of higher order arithmetic in Theorem 2.2.20.

We will also make an abstraction as to the notion of realizers. Whereas we have used numbers and functions as models of computation in the previous section, we will use the more abstract notion of a *partial combinatory algebra* (pca), here. Both notions of realizers employed in the previous section are pca's. As an extension, we allow the realizers to have types, an idea that originates from Kreisel's modified realizability (see [Kre59]). While both typed combinatory algebra and partial combinatory algebra are known for some time, their combination has been introduced only recently by John Longley in [Lon99b].

2.2.1 Typed Partial Combinatory Algebras

In this section we review the notions of *partial combinatory algebra* and *typed partial combinatory algebra* and give some examples.

Definition 2.2.1. A *partial combinatory algebra (pca)* is a set A equipped with a partial binary application that has elements $\mathsf{s}, \mathsf{k} \in A$ such that for all $a, b, c \in A$

$$\mathsf{s}\,a\,b\downarrow, \qquad \mathsf{s}\,a\,b\,c \succeq a\,c\,(b\,c), \qquad \mathsf{k}\,a\,b = b$$

Notice that $e = e'$ means that both sides are defined and equal and $e \succeq e'$ (as in [FS91]) means that whenever e is defined then e' is defined and equal to e. As a convention, application associates to the left and is often denoted by mere juxtaposition, as above. In the following, for any expression e the statement $e \Vdash x$ is supposed to imply that e is defined.

Remark 2.2.2. Combinatory completeness holds for every pca A, i.e., for any polynomial e (a term built up from typed variables and constants) and any variable x there is a polynomial $\lambda^* x.\, e$ whose variables are those of e excluding x such that $\lambda^* x.\, e \downarrow$ and $(\lambda^* x.\, e)\, a \succeq e[a/x]$ for all $a \in A$ and all valuations of the free variables. Conversely, every partial applicative structure that is combinatory complete is a pca.

Example 2.2.3.

(i) The set $\mathbb{N}$, equipped with partial recursive application $(x, y) \mapsto \{x\}(y)$, where $\{\cdot\}$ is an admissible numbering of the set of partial recursive functions, is a pca. We denote this pca by K_1 (for first Kleene algebra).

(ii) The set $\mathbb{N} \to \mathbb{N}$ equipped with partial continuous function application $|$ as defined in section 1.4.1 is a pca. We denote this pca by K_2 (for second Kleene algebra).

Remark 2.2.4. One often encounters a stronger definition of pca's, with $\mathsf{s}\, a\, b\, c \succeq a\, c\, (b\, c)$ replaced by the requirement

$$\mathsf{s}\, a\, b\, c \simeq a\, c\, (b\, c),$$

where $e \simeq e'$ means that e is defined if and only if e' is defined, in which case both are equal. Both K_1 and K_2 are also pca's in this stronger sense, we will however not need this extra requirement for our results. Moreover, although an s combinator for K_2 satisfying the stronger requirement exists, it has to be forced to diverge on certain arguments, which unnecessarily complicates its definition.

Definition 2.2.5. A *typed partial combinatory algebra (tpca)* is a non-empty set $\mathcal{T}$ of types together with

1. binary operations $\times$ and $\to$ on $\mathcal{T}$,
2. a set $|T|$ of realizers of type T for every $T \in \mathcal{T}$
3. a partial application function $\cdot_{S,T} : |S \to T| \times |S| \longrightarrow |T|$ for all $S, T \in \mathcal{T}$

such that for all $S, T, U \in \mathcal{T}$ there are elements

$$\begin{array}{lll} \mathsf{k}_{S,T} \in |S \to T \to S|, & \mathsf{s}_{S,T,U} \in |(S \to T \to U) \to (S \to T) \to (S \to U)| & \\ \mathsf{pair}_{S,T} \in |S \to T \to S \times T|, & \mathsf{fst}_{S,T} \in |S \times T \to S|, & \mathsf{snd}_{S,T} \in |S \times T \to T| \end{array}$$

satisfying

$$\begin{array}{lll} \mathsf{k}\, a\, b = a, & \mathsf{s}\, a\, b \downarrow, & \mathsf{s}\, a\, b\, c \succeq a\, c\, (b\, c) \\ \mathsf{fst}\, (\mathsf{pair}\, a\, b) = a, & \mathsf{snd}\, (\mathsf{pair}\, a\, b) = b & \end{array}$$

for all appropriately typed a, b, c.

As usual, $\to$ associates to the right and $\times$ binds stronger than $\to$, type annotations are frequently omitted. In the following, for any expression e the statements $e \in |T|$ and $e \Vdash x$ are supposed to imply that e is defined.

Remark 2.2.6. The notion of typed pca was introduced by John Longley in [Lon99b] under the name *partial combinatory type structure*. Although typed combinators as well as (untyped) partial combinatory algebra have been studied for a long time, surprisingly enough a combination of both had not been considered before. Our definition differs from that given in [Lon99b] only in the respect that Longley requires that every type is inhabited, whereas we decided not to do so in view of some of the following examples. However, as soon as there is some type T there also is an inhabited type, as $T \to T$ is inhabited by $\mathsf{i} = \mathsf{skk}$.

The notion of tpca should be viewed as open, as it is possible to extend it with additional features such as sum types, arithmetic, recursion etc.

Remark 2.2.7. A typed version of combinatory completeness holds for every tpca, i.e., for any polynomial $e : T$ (a well-typed term built up from typed variables and constants) and any variable $x : S$ there is a polynomial $\lambda^* x : S.\, e : S \to T$ whose variables are those of e excluding x such that $\lambda^* x : S.\, e \downarrow$ and $(\lambda^* x : S.\, e)\, a \succeq e[a/x]$ for all $a \in |T|$ and all valuations of the free variables.

Examples of tpca's are abundant. First of all, any pca can be made into a tpca with one single type. Examples arising from typed functional programming languages are studied in [Lon99a].

Example 2.2.8. Let $\mathcal{A}$ be a pca.

(i) We define the tpca $\mathcal{P}(\mathcal{A})$ as follows. The underlying set of types is the power set of $\mathcal{A}$ and for any type S let $|S| = S$. For types S, T we define

$$\begin{aligned} |S \times T| &= \{a \in \mathcal{A} \mid \mathsf{p}_0\, a \in |S| \text{ and } \mathsf{p}_1\, a \in |T|\} \\ |S \to T| &= \{a \in \mathcal{A} \mid \forall b \in |S|.\, a\, b \in |T|\} \end{aligned}$$

where $\mathsf{p}, \mathsf{p}_0, \mathsf{p}_1 \in \mathcal{A}$ are appropriate combinators for pairing and projection. The typed application operations are defined as restrictions of the application operation in $\mathcal{A}$ to the respective subsets and, therefore, are total.

(ii) If we restrict the set of types to the nonempty subsets of $\mathcal{A}$ we obtain another tpca which we call $\mathcal{P}^+(\mathcal{A})$.

(iii) Let $a \in \mathcal{A}$ be such that $\mathsf{p}aa = a$ and $\forall x \in \mathcal{A}.\ ax = a$. Then we can further restrict the set of types to those subsets of $\mathcal{A}$ containing a and call the resulting tpca $\mathcal{P}^a(\mathcal{A})$.

(iv) Finally, we can restrict the set of types to the set of finite types over some chosen ground type. For appropriate choices of $\mathcal{A}$ this gives rise to HRO (hereditary recursive operations) and ICF (intensional continuous functionals) as discussed in [Tro73]).

Example 2.2.9. Interesting examples arise from cartesian closed categories. Let $\mathbb{C}$ be a CCC. We obtain a tpca as follows. Define the set of types to be the set of objects[3] of $\mathbb{C}$ and for any type T let $|T|$ be the set $\mathbb{C}(1, T)$ of global sections of T. For instance, the cartesian closed category of algebraic lattices gives rise to a notion of realizability investigated in [BBS98]. Other interesting examples arise from the cartesian closed category $\mathbf{PER}(\mathcal{A})$ of partial equivalence relations on a pca $\mathcal{A}$. In analogy to Example 2.2.8 (ii) and (iii) one can consider the subcategory $\mathbf{PER}^+(\mathcal{A})$ of nonempty per's and the subcategory $\mathbf{PER}^a(\mathcal{A})$ of those per's containing some suitable fixed element a in their carrier, as both are exponential ideals. Finally one can restrict to subcategories of finite types over some chosen ground type thus obtaining HEO and ECF (see [Tro73]).

Example 2.2.10.

(i) Fix a set $\mathcal{B}$ of base types and define the set of types as the set freely generated from $\mathcal{B}$ by $\times$ and $\to$. For any type T define $|T|$ to be the set of closed simply typed λ–terms of type T modulo β or $\beta\eta$ equality. Application is induced by the application of λ–calculus.

(ii) Let the set of types be those of Gödel's system T (with product types) and for each type T let $|T|$ be the set of closed terms of system T of type T modulo an appropriate conversion equality or observational equivalence.

(iii) Similarly one may take the closed types of Girard's system F (polymorphic λ–calculus) and for $|T|$ the closed terms of type T modulo β or $\beta\eta$ equality.

Remark 2.2.11. Notice that for example 2.2.8 (ii) and (iii) (but not (i)!) it suffices to require that $\mathcal{A}$ is a *conditional* pca (see [HO93]), in which the axiom $\mathsf{s}\,a\,b\downarrow$ need not hold. In a sense, the introduction of types makes up for the lack of definedness of $\mathsf{s}\,a\,b$ in $\mathcal{A}$. In Example 2.2.8(iii) instead of requiring that all types contain some fixed element a we can require that all types contain some fixed *right–absorptive set* Θ of $\mathcal{A}$ (see [HO93]).

Correspondingly, in Example 2.2.9 for the cases of $\mathbf{PER}^+(\mathcal{A})$ and $\mathbf{PER}^a(\mathcal{A})$ (but not $\mathbf{PER}(\mathcal{A})$!) it suffices that $\mathcal{A}$ is a conditional pca. Furthermore, instead of the category $\mathbf{PER}^a(\mathcal{A})$ we can consider the cartesian closed category $\mathbf{PER}^\Theta(\mathcal{A})$ of those per's R satisfying $\theta R \theta'$ for all $\theta, \theta' \in \Theta$ for some right–absorptive set Θ.

In Example 2.2.9 the category $\mathbb{C}$ need not be cartesian closed. Instead, one may take for $\mathbb{C}$ a weak version of a partial cartesian closed category as in [Bir99].

2.2.2 Universal Types

Definition 2.2.12. Let $\mathcal{T}$ be a tpca. A type $U \in \mathcal{T}$ is called *universal* if for any type $T \in \mathcal{T}$ there are realizers $\mathsf{e}_T \in |T \to U|$ and $\mathsf{r}_T \in |U \to T|$ such that

[3] ignoring size matters for the moment

$\forall x \in |T|.\, \mathsf{r}_T(\mathsf{e}_T\, x) = x$ i.e. every type is a partial[4] retract of the type U.

A universal type U in a tpca gives rise to a pca as in particular $U \to U$ is a retract of U. The underlying set of the pca is $|U|$ and for $x, y \in |U|$ application is defined as $\mathsf{r}_{U\to U} xy$. This structure is combinatory complete. For any untyped U-polynomial e and variables $x_1, \ldots, x_n$ define $\lambda^* x_1 \cdots x_n.\, e$ as $\mathsf{e}_{U\to U}\, \lambda^* x_1 : U. \cdots \mathsf{e}_{U\to U}\, \lambda^* x_n : U.\, e'$, where e' is e with applications MN replaced by $\mathsf{r}_{U\to U} MN$. In particular, combinators s and k can be defined for U in this way.

In [Lon99b] a 2-category of tpca's and applicative morphisms is introduced, containing the 2-category of pca's familiar from [Lon94] as a full subcategory. We will not repeat the definition of the 2-category here but rather state the induced notion of equivalence for typed pca's as a definition, although it is a bit lengthy.

Definition 2.2.13. Two tpca's $\mathcal{T}$ and $\mathcal{S}$ are *equivalent* whenever there are functions $f : \mathcal{T} \to \mathcal{S}$ and $g : \mathcal{S} \to \mathcal{T}$ together with a function $u_T : |T| \to |f(T)|$ for every $T \in \mathcal{T}$ and a function $v_S : |S| \to |g(S)|$ for every $S \in \mathcal{S}$ satisfying the following properties. For every pair $T, T' \in \mathcal{T}$ there is some $q_{T,T'} \in |f(T \to T') \to f(T) \to f(T')|$ such that

$$q_{T,T'}\, u_{T\to T'}(a)\, u_T(b) \succeq u_{T'}(ab) \quad \text{for all} \quad a \in |T \to T'|, b \in |T|.$$

Likewise, for every pair $S, S' \in \mathcal{S}$ there is some $r_{S,S'} \in |g(S \to S') \to g(S) \to g(S')|$ such that

$$r_{S,S'}\, v_{S\to S'}(a)\, v_S(b) \succeq v_{S'}(ab) \quad \text{for all} \quad a \in |S \to S'|, b \in |S|.$$

Furthermore, for every $T \in \mathcal{T}$ there are elements $i_T \in |T \to g(f(T))|$ and $j_T \in |g(f(T)) \to T|$ satisfying

$$i_T\, a = v_{f(T)}(u_T(a)) \quad \text{and} \quad a = j_T\, v_{f(T)}(u_T(a))$$

for all $a \in |T|$ and for every $S \in \mathcal{S}$ there are elements $k_S \in |S \to f(g(S))|$ and $l_S \in |f(g(S)) \to S|$ satisfying

$$k_S\, a = u_{g(S)}(v_S(a)) \quad \text{and} \quad a = l_S\, u_{g(S)}(v_S(a))$$

for all $a \in |S|$.

For instance, the tpca's $\mathcal{P}^0(K_1)$ and $\mathcal{P}^+(K_1)$ are equivalent. The following proposition says that a tpca is equivalent to a pca if and only if it has a universal type.

Proposition 2.2.14. *If a tpca $\mathcal{T}$ has a universal type U, then it is equivalent to the induced pca on U. Conversely, if a tpca is equivalent to a pca, then it has a universal type.*

[4] meaning that r_T may be partial although e_T is still total

Proof. straightforward exercise. □

One consequence of the existence of a universal type is the existence of fixpoint combinators for all types.

Proposition 2.2.15 (Shin-ya Katsumata). *Let $\mathcal{T}$ be a tpca with a universal type $U \in \mathcal{T}$. Then for every type $T \in \mathcal{T}$ there is an element $\mathsf{Y}_T \in |(T \to T) \to T|$ satisfying $\forall f \in |T \to T|.\, (\mathsf{Y}_T f) \succeq f(\mathsf{Y}_T f)$.*

Proof. Since U is universal, $U \to (T \to T) \to T$ is a retract of U with realizers $\mathsf{r} : U \to (U \to (T \to T) \to T)$ and $\mathsf{e} : ((U \to (T \to T) \to T) \to U$. Let

$$\mathsf{A} = \lambda^* x : U \lambda^* y : T \to T.\, y(rxxy) : U \to (T \to T) \to T)$$

and $\mathsf{Y}_T = \mathsf{A}(\mathsf{eA}) : (T \to T) \to T$. Then for any $f : T \to T$,

$$\begin{array}{rcll} \mathsf{Y}_T f & = & \mathsf{A}(\mathsf{eA}) f & \text{(by definition of } \mathsf{Y}_T) \\ & = & (\lambda^* x : U \lambda^* y : T \to T.\, y(\mathsf{r}xxy))(\mathsf{e}A) f & \text{(by definition of } \mathsf{A}) \\ & \succeq & f(\mathsf{r}(\mathsf{eA})(\mathsf{eA}) f) & \text{(by remark 2.2.7)} \\ & = & f(\mathsf{A}(\mathsf{eA}) f) & \text{(since } \mathsf{r}(\mathsf{e}(x)) = x) \\ & = & f(\mathsf{Y}_T f) & \text{(by definition of } \mathsf{Y}_T \text{ again)} \end{array}$$

□

Corollary 2.2.16. *Let $\mathcal{T}$ be a tpca with an s combinator satisfying $\mathsf{s}\, a\, b\, c \simeq a\, c\, (b\, c)$ with a universal type $U \in \mathcal{T}$. Then for every type $T \in \mathcal{T}$ there is an element $\mathsf{Y}_T \in |(T \to T) \to T|$ satisfying $\forall f \in |T \to T|.\, f(\mathsf{Y}_T f) \simeq (\mathsf{Y}_T f)$.*[5]

Proof. The claim follows from the proof of Proposition 2.2.15 and the fact that if $\mathsf{s}\, a\, b\, c \simeq a\, c\, (b\, c)$ then $(\lambda^* x : S.\, e)\, a \simeq t[a/x]$. □

The existence of a universal type is a rather strong requirement for a tpca. As shown in the following proposition, a lot of natural examples of tpca's do not have a universal type.

Proposition 2.2.17.

(i) The tpca $\mathcal{P}(\mathcal{A})$ has no universal type.

(ii) The tpca's $\mathcal{P}^+(\mathcal{A})$ and $\mathcal{P}^a(\mathcal{A})$ have no universal type unless $\mathcal{A}$ has only one element.

[5] At the time of writing the article [LS02], we were only able to prove this weaker statement about the existence of fixpoint combinators and posed the existence of a fixpoint combinator with the properties described in 2.2.15 as a question. Although it is well known that Turing's fixpoint combinator has the desired property in the untyped case, it is not obvious how to transfer this result into the typed case. Upon reading our article, Shin-ya Katsumata communicated his solution to us.

(iii) *The tpca* $\mathcal{P}^\Theta(\mathcal{A})$ *has no universal type unless* $\mathcal{A} = \Theta$.

(iv) *The tpca's HRO and ICF have no universal type.*

(v) *The tpca induced by a well-pointed CCC has a universal type if and only if there is an object* U *such that any object* A *in the category is a retract of* U.

(vi) *The tpca induced by the cartesian closed category of algebraic lattices has no universal type.*

(vii) *The tpca* $\mathbf{PER}(\mathcal{A})$ *has no universal type*

(viii) *The tpca's* $\mathbf{PER}^+(\mathcal{A})$ *and* $\mathbf{PER}^a(\mathcal{A})$ *have no universal type unless* $\mathcal{A}$ *has only one element.*

(ix) *The tpca's HEO and ECF have no universal type*

(x) *None of the tpca's in example 2.2.10 has a universal type.*

Proof. Statements (i), (ii) and (iii) hold for cardinality reasons. Let us suppose $\mathcal{T}$ is one of the tpca's $\mathcal{P}(\mathcal{A})$, $\mathcal{P}^+(\mathcal{A})$, $\mathcal{P}^a(\mathcal{A})$ or $\mathcal{P}^\Theta(\mathcal{A})$. Assume there exists a universal type $U \in \mathcal{T}$. Then for any $T \in \mathcal{T}$, $T = \{\mathsf{r}_T\, u \mid u \in U\}$. Thus, the number of types were bounded by the cardinality of $\mathcal{A}$ which is not the case.

Similarly for (vii) and (viii). Statements (iv), (ix) and (x) hold because of Proposition 2.2.15. All application functions are total but there are realizable endomaps without fixpoint. Statement (v) is only the expansion of the definition of universal type. Finally, (vi) follows from (v) as there clearly is no universal algebraic lattice for cardinality reasons. □

2.2.3 Realizability Models

Now we generalize various constructions of realizability models from pca's to typed pca's.

Definition 2.2.18. Let $\mathcal{T}$ be a tpca.

(i) The category $\mathbf{Asm}(\mathcal{T})$ of *assemblies over* $\mathcal{T}$ is defined as follows. An object X consists of an underlying set I together with an underlying type $T \in \mathcal{T}$ and a realizability relation $\Vdash_X \subseteq |T| \times I$ such that $\forall i \in I. \exists t \in |T|.\, t \Vdash_X i$. For any $i \in I$ we define the set of realizers of i, $\|i\|_X$ to be the set $\{t \in |T| \mid t \Vdash_X i\}$. A morphism from X to X' is a map $f : I \longrightarrow I'$ which is tracked by some $a \in |T \to T'|$, i.e. $\forall i \in I. \forall b \in |T|.\, b \Vdash_X i \to a\, b \Vdash_{X'} f(i)$.

(ii) The category $\mathbf{D}(\mathcal{T})$ of *discrete sets over* $\mathcal{T}$ is the full subcategory of $\mathbf{Asm}(\mathcal{T})$ on those objects $X = (I, T, \Vdash_X)$ which satisfy $\forall t \in |T|.\forall i, i' \in I.\, t \Vdash_X i \wedge t \Vdash_X i' \rightarrow i = i'$.[6]

(iii) The fibration $\mathbf{UFam}(\mathbf{D}(\mathcal{T}))$ over $\mathbf{Asm}(\mathcal{T})$ of *uniform families of discrete sets* induced by $\mathcal{T}$ is defined as follows. An object in the total category is an assembly $X = (I, T, \Vdash_X)$ together with an I-indexed familiy of discrete sets $(D_i)_{i\in I}$ over one common type S. A morphism from $(X, S, (D_i)_{i\in I})$ to $(X', S', (D'_i)_{i\in I'})$ consists of a morphism of assemblies $f : X \longrightarrow X'$ together with a family of morphisms of discrete sets $(f_i : D_i \longrightarrow D'_{f(i)})_{i\in I}$ which is uniformly tracked, i.e. there is a realizer $a \in |T \rightarrow S \rightarrow S'|$ such that $\forall i \in I.\forall b \in |T|.\, b \Vdash_X i \rightarrow a\, b \Vdash f_i$. The fibration itself is given by the projection functor sending $(X, S, (D_i)_{i\in I})$ to X and $(f, (f_i)_{i\in I})$ to f.

(iv) The hyperdoctrine $\mathcal{H}(\mathcal{T})$ over **Set** is defined as follows. An element in the fibre over a set I is a type $T \in \mathcal{T}$ together with a function $\varphi : I \longrightarrow \mathcal{P}(|T|)$. The entailment $\varphi \vdash_I \varphi'$ holds if there is some $a \in |T \rightarrow T'|$ such that $\forall i \in I.\forall b \in \varphi(i).\, a\, b \in \varphi'(i)$. Reindexing is given by precomposition.

(v) The *realizability category over* $\mathcal{T}$, $\mathbf{RC}(\mathcal{T})$, is defined as $\mathbf{Set}[\mathcal{H}(\mathcal{T})]$, i.e. the category whose objects are formal partial equivalence relations w.r.t. $\mathcal{H}(\mathcal{T})$ and whose morphisms are formal strict functional relations (see [Pit99] for the exact definition).

One can prove that all constructions of Definition 2.2.18 applied to equivalent tpca's yield equivalent (fibred) categories. For tpca's with just one type, the constructions above coincide with the well known constructions for pca's. Next we discuss which of the properties of realizability models lift to the typed case.

The categories $\mathbf{Asm}(\mathcal{T})$ and $\mathbf{D}(\mathcal{T})$ are regular and locally cartesian closed. The global sections functor $\Gamma : \mathbf{Asm}(\mathcal{T}) \longrightarrow \mathbf{Set}$ has a right adjoint $\nabla : \mathbf{Set} \longrightarrow \mathbf{Asm}(\mathcal{T})$ which can be described concretely as follows. Let $T \in \mathcal{T}$ be a type inhabited by $* \in |T|$. Then for any set I define ∇I as the assembly with underlying set I, underlying type T and $\|i\|_{\nabla_I} = \{*\}$ for all $i \in I$. An assembly $X = (I, T, \Vdash_X)$ is called *codiscrete* iff it is isomorphic to ∇I, i.e. iff there exists a $t \in |T|$ such that $t \Vdash_X i$ for all $i \in I$. Note that an object of $\mathbf{Asm}(\mathcal{T})$ is in $\mathbf{D}(\mathcal{T})$ if and only if it is internally orthogonal to $\nabla(2)$. Furthermore, the fibration $\mathbf{UFam}(\mathbf{D}(\mathcal{T}))$ over $\mathbf{Asm}(\mathcal{T})$ is a split fibration equivalent to the fibration of *discrete families* in $\mathbf{Asm}(\mathcal{T})$, i.e. those families that are internally orthogonal to $\nabla(2)$. Thus it follows immediately that the fibration is complete and a fibred CCC (see [HRR90]). However, there need not exist

[6] Note that this category is frequently called the category of modest sets in the untyped case. In calling it the category of discrete sets we follow the argument given in [OS00, Section 2.2].

a generic family.[7]

The hyperdoctrine $\mathcal{H}(\mathcal{T})$ is a model for the $(\top, \wedge, \rightarrow, \forall, \exists, =)$–fragment of first order predicate logic. However, in general $\mathcal{H}(\mathcal{T})$ need not be a tripos (see [HJP80, Pit81]) as there need not exist a generic predicate allowing for an interpretation of higher order logic. Moreover, the fibres need not have finite suprema allowing for an interpretation of $(\bot, \vee)$.

The category $\mathbf{RC}(\mathcal{T})$ is locally cartesian closed and exact. It contains $\mathbf{Asm}(\mathcal{T})$ as a full reflective subcategory closed under subobjects and the inclusion preserves the locally cartesian closed structure. The category $\mathbf{RC}(\mathcal{T})$ is the *exact completion of the regular category* $\mathbf{Asm}(\mathcal{T})$ and was in fact defined this way in [Lon99b]. Every object of $\mathbf{RC}(\mathcal{T})$ is covered via a regular epi by an object of $\mathbf{Asm}(\mathcal{T})$ (see [FS91, Car95, CV98]). In general $\mathbf{RC}(\mathcal{T})$ need not be a topos.

2.2.4 Impredicativity entails Untypedness

We will show now that a tpca $\mathcal{T}$ has a universal type, i.e. is essentially untyped, if and only if the ensuing realizability models of Definition 2.2.18 have the usual impredicative features.

For this purpose we need the following somewhat technical lemma.

Lemma 2.2.19. *Suppose there exists a morphism $m : A \longrightarrow X$ in $\mathbf{Asm}(\mathcal{T})$ such that for every subobject $n : B \longrightarrow Y$ of a codiscrete Y there is a pullback diagram*

$$\begin{array}{ccc} B & \longrightarrow & A \\ {\scriptstyle n}\downarrow & \lrcorner & \downarrow{\scriptstyle m} \\ Y & \longrightarrow & X \end{array}$$

Then $\mathcal{T}$ has a universal type.

Proof. Let U be the type of A. We claim that U is universal. For an arbitrary type $T \in \mathcal{T}$ let Δ_T be the assembly with underlying set $|T|$, underlying type T and $\|t\|_{\Delta_T} = \{t\}$ for all $t \in |T|$. Let $i : \Delta_T \longrightarrow \nabla|T|$ be the morphism whose underlying map is the identity on $|T|$. As i is monic and $\nabla|T|$ is codiscrete there is a pullback diagram

$$\begin{array}{ccc} \Delta_T & \xrightarrow{f} & A \\ {\scriptstyle i}\downarrow & \lrcorner & \downarrow{\scriptstyle m} \\ \nabla|T| & \xrightarrow[u]{} & X \end{array}$$

[7]Notice that we call an object X in the total category of a fibration *generic* whenever for each object Y in the total category there is a cartesian morphism from Y to X. We do not require uniqueness in any form.

due to our assumption on m. Let $\mathsf{e}_T \in |T \to U|$ be a realizer of f. Note that e_T is defined for all $t \in |T|$. Let D be the assembly with underlying set $|T|$, underlying type U and $\|t\|_D = \{\mathsf{e}_T t\}$ for all $t \in |T|$. Let $i' : D \longrightarrow \nabla|T|$ and $f' : D \longrightarrow A$ be the morphisms whose underlying maps are those of i and f, respectively. Therefore, $u \circ i' = m \circ f'$. Let $g : D \longrightarrow \Delta_T$ be the mediating morphism, i.e. $i \circ g = i'$ and $f \circ g = f'$. Notice that $g(t) = t$ for all $t \in |T|$, as $i \circ g = i'$. Let r_T be a realizer of g. Then for all $t \in |T|$ we have $\mathsf{r}_T(\mathsf{e}_T t) \Vdash_{\Delta_T} g(t) = t$ and, therefore, $\mathsf{r}_T(\mathsf{e}_T t) = t$, as desired. □

Theorem 2.2.20. *Let $\mathcal{T}$ be a tpca. The following are equivalent.*

(i) $\mathcal{T}$ *has a universal type*

(ii) $\mathbf{Asm}(\mathcal{T})$ *has a generic mono*

(iii) $\mathcal{H}(\mathcal{T})$ *is a tripos*

(iv) $\mathbf{RC}(\mathcal{T})$ *is a topos*

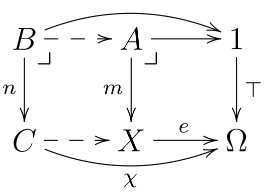
(v) $\mathbf{UFam}(\mathbf{D}(\mathcal{T}))$ *over* $\mathbf{Asm}(\mathcal{T})$ *has a generic family.*

Proof. If $\mathcal{T}$ has a universal type then it is equivalent to a pca. As all constructions considered in (ii) - (v) preserve equivalence, the induced structures are equivalent to ones induced by a pca. For these the properties (ii) - (v) are well known and, therefore, condition (i) implies conditions (ii) - (v).

Assume (ii) and let m be a generic mono. As any mono arises as a pullback of m, this holds in particular for subobjects of codiscrete objects. Therefore, condition (i) follows by Lemma 2.2.19.

Clearly (iii) implies (iv).

Assume (iv). Let $\top : 1 \longrightarrow \Omega$ be a subobject classifier in $\mathbf{RC}(\mathcal{T})$. Then Ω is covered by an assembly X via an epi $e : X \longrightarrow \Omega$. Let $m : A \longrightarrow X$ be the pullback of $\top$ along e. As $\mathbf{Asm}(\mathcal{T})$ is closed under subobjects A is an assembly, too. We show that all subobjects of codiscrete objects arise as pullbacks of m. Let $n : B \longrightarrow C$ be a subobject of a codiscrete assembly C. As C is projective in $\mathbf{RC}(\mathcal{T})$ the classifying morphism $\chi : C \longrightarrow \Omega$ for n factors through e and, therefore, in the diagram



the left square is a pullback, too. Thus, by Lemma 2.2.19 the tpca $\mathcal{T}$ has a universal type. Hence (iv) entails (i).

Finally, assume (v). Remember that the fibration $\mathbf{UFam}(\mathbf{D}(\mathcal{T}))$ is equivalent to the fibration of discrete families in $\mathbf{Asm}(\mathcal{T})$. As any mono is a discrete family this implies that every mono appears as a pullback of a fixed morphism in $\mathbf{Asm}(\mathcal{T})$. Thus by Lemma 2.2.19 condition (i) follows. □

In the proof of (v) ⇒ (i) we have actually shown more. Whenever there is a subfibration of the codomain fibration of $\mathbf{Asm}(\mathcal{T})$ that contains all monos and has a generic family, then $\mathcal{T}$ has a universal type.

Moreover, whenever there is an arbitrary fibration over $\mathbf{Asm}(\mathcal{T})$ that contains the subobject fibration of $\mathbf{Asm}(\mathcal{T})$ as a *definable* subfibration and has a generic family, then $\mathcal{T}$ has a universal type. This is because a definable subfibration as opposed to an arbitrary subfibration inherits the property of having a generic family. See [Bor94] for a treatment of definability in fibrations.

But, equally important, from Theorem 2.2.20 and Proposition 2.2.17 we get quite a few examples of realizability models which are not impredicative. We just discuss those examples which have already been considered in other contexts.

As the category **ALat** of algebraic lattices does not have a universal object the exact completion of the regular category $\mathbf{Asm}(\mathbf{ALat})$ is not a topos. Observe that as $\mathbf{Asm}(\mathbf{ALat})$ is the regular completion of the category **Sp** of topological spaces (see [Ros99]) it follows immediately that the exact completion of **Sp** as a lex category is not a topos. Moreover, the fibration $\mathbf{UFam}(\mathbf{D}(\mathbf{ALat}))$ has no generic family and, therefore, is not a model of polymorphic λ–calculus or even Calculus of Constructions. This confirms a conjecture raised in [BBS98].

Next notice that the category of $\neg\neg$–separated objects of the modified realizability topos **Mod** (see [Gra81, Oos97b]) is equivalent to the category $\mathbf{Asm}(\mathcal{P}^0(K_1))$ where K_1 is the 1st Kleene algebra and $\mathcal{P}^0(K_1)$ is the tpca defined in Example 2.2.8. In this case the underlying type of an assembly is called its *set of potential realizers*.

More generally, whenever $\mathcal{A}$ is a conditional pca and $\Theta \subseteq \mathcal{A}$ is a right–absorptive set then the category of $\neg\neg$–separated objects in the modified realizability topos over $(\mathcal{A}, \Theta)$ (see [HO93, OR94]) is equivalent to the category $\mathbf{Asm}(\mathcal{P}^\Theta(\mathcal{A}))$ (see Remark 2.2.11). As tpca's of the form $\mathcal{P}^\Theta(\mathcal{A})$ do not have a universal type it follows from Theorem 2.2.20 that there is no generic family for the fibration of discrete families in the category of $\neg\neg$–separated objects of a modified realizability topos. Another consequence of Theorem 2.2.20 is that a modified realizability topos does not arise as the exact completion of its regular subcategory of $\neg\neg$–separated objects as the exact completion is not a topos.

All these considerations apply as well to extensional versions of modified realizability as considered in [Gra81] replacing $\mathcal{P}^0(K_1)$ by $\mathbf{PER}^0(K_1)$.

We have seen that one can build realizability models over arbitrary typed pca's, i.e. typed models of computation, such as for example models of Gödel's system T. Traditional untyped pca's can be considered as particular instances of typed pca's, namely those with a universal type. It is well known that realizability over untyped pca's gives rise to models of impredicative type theory. We have shown that this implication can be reversed in the sense that a realizability model over a typed pca

$\mathcal{T}$ is impredicative if and only if $\mathcal{T}$ admits a universal type, i.e. is essentially untyped.

There are, however, examples of typed pca's without a universal type whose categories of assemblies are nevertheless equivalent to categories of $\neg\neg$–separated objects of appropriate toposes. For example, the category of assemblies over $\mathcal{P}^0(K_1)$) turns out to be equivalent to the category of $\neg\neg$–separated objects of the modified realizability topos introduced independently by Hyland and Grayson in the early eighties and studied in greater detail by van Oosten [Oos97b, Oos91]. Another example is the category of assemblies over $\mathbf{PER}^0(K_1)$, which is equivalent to the category of $\neg\neg$–separated objects of the extensional modified realizability topos of [Gra81]. But it is not clear in general when a category of assemblies over a tpca without a universal type can be reconstructed as the category of $\neg\neg$–separated objects for some topos. For instance, the categories of assemblies over the tpca's listed in Example 2.2.10 are not categories of $\neg\neg$–separated objects of *any* topos. The category of assemblies over the tpca described in Example 2.2.10 (i) has no NNO. The categories of assemblies over the tpca's described in Example 2.2.10 (ii) and (iii) fail to validate sentences of the form $\forall x : \mathbb{N} \exists y : \mathbb{N}.\, T(e, x, y)$ for appropriately chosen $e \in \mathbb{N}$, which hold in (the category of $\neg\neg$–separated objects of) any topos.

2.2.5 Discussion of related work

L. Birkedal [Bir99] has proved a result analogous to ours but for his *wcpc*'s instead of typed pca's. Birkedal's *wcpc*'s are a bit more general than tpca's as the former cover also the relative realizability toposes of [Bir99]. His proof relies on A. Pitts' [Pit99] characterization of those hyperdoctrines which give rise to a topos when applying the "tripos–to–topos construction". This use of Pitts' characterization theorem was inspired by discussions with the author of this thesis in the autumn of 1999 and an early version of [LS02], in which we proved our result using Pitts' result of [Pit99].

When reading a draft version of the paper "An abstract look at realizability" [dMRR99] and trying to understand some proof therein it appeared to us that we can reuse an older result of ours, namely that a tpca $\mathcal{T}$ has a universal type if and only if $\mathbf{Asm}(\mathcal{T})$ has a generic mono, with almost no modification. The point was that we saw that firstly, a subobject classifier in $\mathbf{RC}(\mathcal{T})$ gives rise to a mono m in $\mathbf{Asm}(\mathcal{T})$ from which one may obtain every subobject of a codiscrete object ∇I via pullback and secondly, that from such a mono m we can construct a universal type in $\mathcal{T}$ in literally the same way as from a generic mono in $\mathbf{Asm}(\mathcal{T})$. Using this method one can prove the respective results of [Bir99] and [dMRR99] in a simpler way even though these are formulated for more general settings than just tpca's. In particular, the intensional notion of computation employed in [dMRR99] is given by just a category $\mathbb{C}$ together with a faithful functor U from $\mathbb{C}$ to the category of sets and partial functions both satisfying some elementary conditions but without any assumptions about the existence of weak partial function spaces. Using a result from [CR98] it follows that $\mathbb{C}$ has weak dependent products whenever the ensuing realizability category is locally cartesian closed.

One also should mention the work of M. Menni. In [Men99] he gave a very nice characterization of those finite limit categories $\mathbb{C}$ whose exact completion is a topos. In the subsequent paper [Men00] under the assumption that the global sections functor $\Gamma : \mathbb{C} \to \mathbf{Set}$ has a full and faithful right adjoint ∇ (a so–called "chaotic situation") he could simplify the characterization of [Men99] as follows: the exact completion of $\mathbb{C}$ is a topos if and only if $\mathbb{C}$ has weak dependent products and a generic mono. As partitioned assemblies over a typed pca (or arising from the more general situations considered in [Bir99, dMRR99]) are instances of a "chaotic situation" the result of [Men00] provides a conceptual explanation for the central role that generic monos play—explicitly or implicitly—in our work and the above mentioned papers.

2.3 Separating models for continuity axioms

The dependencies between the continuity principles introduced in section 1.4.1 can be charted as follows.

$$\begin{array}{ccccc} \mathsf{WC\text{-}N} & \Longrightarrow & \mathsf{WC_{cp}\text{-}N} & \Longrightarrow & \mathsf{WC_{seq}\text{-}N} \\ \Downarrow & & \Downarrow & & \Downarrow \\ \mathsf{CP}(\mathbb{N} \to \mathbb{N}, \mathbb{N}) & \Longrightarrow & \mathsf{CP}(\mathbb{N} \to 2, \mathbb{N}) & \Longrightarrow & \mathsf{CP}(\mathbb{N}^+, \mathbb{N}) \end{array}$$

We will demonstrate that all of these implications are strict by exhibiting separating models. All of these models can be extend to toposes, which shows that even in higher order arithmetic, none of the above implications can be reversed.

We start by giving a model that refutes all principles in the top row and satisfies all principles in the bottom row.

Proposition 2.3.1.

(i) The category $\mathbf{Asm}(K_1)$ *validates Church's thesis and (when interpreted with a classical metalogic) Markov's principle.*

(ii) The category $\mathbf{Asm}(K_1)$ *validates* $\mathsf{CP}(\mathbb{N} \to \mathbb{N}, \mathbb{N})$.

(iii) The category $\mathbf{Asm}(K_1)$ *refutes* $\mathsf{WC_{seq}\text{-}N}$.

Proof. (i) follows from Proposition 2.1.23, as the $\underline{\mathtt{rn}}$-interpretation is the formalized version of the interpretation in $\mathbf{Asm}(K_1)$. (ii) The principle $\mathsf{CP}(\mathbb{N} \to \mathbb{N}, \mathbb{N})$ follows from CT and M by the KLST theorem (see Theorem 1.4.4). (iii) As a consequence of Church's thesis, $\forall \alpha : \mathbb{N}^+ \exists n : \mathbb{N}.\ \alpha = \{n\}$. By applying $\mathsf{WC_{seq}\text{-}N}$ we get $\exists m, n : \mathbb{N}\ \forall \alpha : \mathbb{N}^+.\ \alpha \in 0^m \to \alpha = \{n\}$. We conclude $\exists m : \mathbb{N}\ \forall \alpha : \mathbb{N}^+.\ \alpha \in 0^m \to \alpha = 0^\omega$, which is absurd. $\square$

As $\mathbf{Asm}(K_1)$ validates the strongest of the principles in the bottom row and refutes the weakest of the principles in the top row of the above diagram, the vertically depicted implications can be seen to be strict. Moreover, as $\mathbf{Asm}(K_1)$ is the subcategory of $\neg\neg$-separated objects of the effective topos $\mathcal{E}ff$ (see [Hyl82]), the implications cannot be reversed even in higher order arithmetic.

We will now give two models that both separate the principles in the leftmost column from those in the middle column.

Definition 2.3.2.

(i) Let $\omega\mathbf{ALat}$ be the category of ω-algebraic lattices and continuous (with respect to the Scott-topology) maps.

(ii) Let $\mathbf{Dom}$ be the category of coherently complete Scott-domains, i.e. the category of ω-algebraic, coherently complete dcpo's with a least element and continuous (with respect to the Scott-topology) maps.

Proposition 2.3.3.

(i) The category $\mathbf{Asm}(\mathbf{Dom})$ *validates* $\mathsf{AC}(\sigma, \tau)$ *for all* σ, τ *in the finite type hierarchy over* $\mathbb{N}$.

(ii) The category $\mathbf{Asm}(\mathbf{Dom})$ *validates* $\mathsf{WC}_{\mathsf{cp}}\text{-}\mathsf{N}$.

(iii) The category $\mathbf{Asm}(\mathbf{Dom})$ *refutes* $\mathsf{CP}(\mathbb{N} \to \mathbb{N}, \mathbb{N})$.

Proof. (i) As the interpretations of the finite types over $\mathbb{N}$ in $\mathbf{Asm}(\mathbf{Dom})$ are discrete objects, these can be identified with the induced partial equivalence relations on the respective domains of realizers. The domain D_ι of realizers for the type $\mathbb{N}$ is $\mathbb{N}_\perp$. For finite types σ, τ the domain $D_{\sigma\to\tau}$ of realizers for $\sigma \to \tau$ is the domain $D_\tau^{D_\sigma}$. The induced partial equivalence relation on D_σ coincides with the binary totality relation $\sim_\sigma$ on D_σ. Let $z \in D_{\sigma\to\tau}$ such that for all $x \in D_\sigma$, if $x \sim_\sigma x$ then $zx \sim_\tau zx$. By a result of Y.L. Ershov (see [Ers75] and also [LM84]), it is the case that for all $x, y \in D_\sigma$, if $x \sim_\sigma y$ then $zx \sim_\tau zy$. In other words, every *a priori* intensional operation from σ to τ is extensional and thus $\mathsf{AC}(\sigma, \tau)$ holds in for all finite types σ, τ over $\mathbb{N}$ in $\mathbf{Asm}(\mathbf{Dom})$.
(ii) As $\mathsf{AC}(\mathbb{N} \to 2, \mathbb{N})$ holds in the model under discussion, it suffices to show $\mathsf{CP}(\mathbb{N} \to 2, \mathbb{N})$. The interpretation of the finite type hierarchy in $\mathbf{Asm}(\mathbf{Dom})$ yields the Kleene-Kreisel hierarchy of countable functionals. It is well known that in the hierarchy of countable functionals there exists a functional $\Phi : ((\mathbb{N} \to 2) \to \mathbb{N}) \to \mathbb{N})$ that has the property that for all $F : (\mathbb{N} \to 2) \to \mathbb{N}$ and all $\alpha, \beta : \mathbb{N} \to 2$,

$$\text{if} \quad \overline{\alpha}(\Phi(F)) = \overline{\beta}(\Phi(F)) \quad \text{then} \quad F(\alpha) = F(\beta),$$

in other words, $\Phi(F)$ is a uniform modulus of continuity for F. Therefore, *a fortiori* $\mathsf{CP}(\mathbb{N} \to 2, \mathbb{N})$ holds.

(iii) The principles $\mathsf{CP}(\mathbb{N} \to \mathbb{N}, \mathbb{N})$ and $\mathsf{AC}((\mathbb{N} \to \mathbb{N}) \to \mathbb{N}, \mathbb{N})$ entail the existence of a "modulus of continuity at const_0" functional $\Psi : ((\mathbb{N} \to \mathbb{N}) \to \mathbb{N}) \to \mathbb{N})$, which is absurd by [TvD88b, Corollary 6.11]. □

The category **Dom** is well-pointed and has $\mathbb{T}^\omega$ (and $\mathbb{N} \to \mathbb{N}_\perp$) as universal objects as shown in [Plo78]. Therefore, the category $\mathbf{Asm}(\mathbf{Dom})$ is the subcategory of $\neg\neg$-separated objects of the topos $\mathbf{RC}(\mathbf{Dom}) = \mathbf{RC}(\mathbb{T}^\omega)$, which demonstrates that the implications cannot be reversed in higher order arithmetic, either. Note also that the use of typed realizability allowed for an easier and more transparent proof, here, compared to working with the (albeit equivalent) category $\mathbf{Asm}(\mathbb{T}^\omega)$.

Proposition 2.3.4.

(i) The category $\mathbf{Asm}(\omega\mathbf{ALat})$ *validates* $\mathsf{AC}(\sigma, \tau)$ *for all* σ, τ *in the finite type hierarchy over* $\mathbb{N}$.

(ii) The category $\mathbf{Asm}(\omega\mathbf{ALat})$ *validates* $\mathsf{WC}_{\mathsf{cp}}\text{-}\mathsf{N}$.

(iii) The category $\mathbf{Asm}(\omega\mathbf{ALat})$ *refutes* $\mathsf{CP}(\mathbb{N} \to \mathbb{N}, \mathbb{N})$.

Proof. (i) See [TvD88b, 9.4.10–9.4.13].
(ii) The interpretation of the finite type hierarchy in $\mathbf{Asm}(\omega\mathbf{ALat})$ yields the Kleene-Kreisel hierarchy of countable functionals. The rest of the argument follows the proof of Proposition 2.3.3 (ii).
(iii) follows by the same argument as the proof of 2.3.3 (iii). □

The category $\omega\mathbf{ALat}$ is well-pointed and has $\mathcal{P}(\omega)$ as a universal object. Therefore, the category $\mathbf{Asm}(\omega\mathbf{ALat})$ is the subcategory of $\neg\neg$-separated objects of the topos $\mathbf{RC}(\omega\mathbf{ALat}) = \mathbf{RC}(\mathcal{P}(\omega))$, which demonstrates that the implications cannot be reversed in higher order arithmetic, either.

Remark 2.3.5. The subcategory $\mathbf{D}(\omega\mathbf{ALat})$ of discrete objects of the category $\mathbf{Asm}(\omega\mathbf{ALat})$ is equivalent to Scott's category **Equ** of countably based *equilogical spaces*. As the interpretation of HA^ω takes place in this subcategory, Proposition 2.3.4 holds for the category of equilogical spaces, as well.

We will now give two models that both separate the principles in the middle column from those in the rightmost column. For this purpose, we will make use of a form of extensional realizability.

Proposition 2.3.6.

(i) The category $\mathbf{Asm}(\mathbf{PER}^0(K_1))$ *validates* $\mathsf{AC}(\sigma, \tau)$ *for all* σ, τ *in the finite type hierarchy over* $\mathbb{N}$.

(ii) The category $\mathbf{Asm}(\mathbf{PER}^0(K_1))$ *validates* $\mathsf{WC}_{\mathsf{seq}}\text{-}\mathsf{N}$.

(iii) The category $\mathbf{Asm}(\mathbf{PER}^0(K_1))$ *refutes* $\mathsf{CP}(\mathbb{N} \to 2, \mathbb{N})$.

Proof. (i) By the very definition of extensional realizability, the intensional operations from σ to τ coincide with the extensional operations and thus $\mathsf{AC}(\sigma, \tau)$ holds for all finite types σ, τ over $\mathbb{N}$ in $\mathbf{Asm}(\mathbf{PER}^0(K_1))$.
(ii) As $\mathsf{AC}(\mathbb{N}^+, \mathbb{N})$ holds in the model under discussion, it suffices to show $\mathsf{CP}(\mathbb{N}^+, \mathbb{N})$. As CT is not valid in extensional realizability, we have to argue "externally". By the KLS theorem of recursion theory (see [Rog67]), there is an algorithm computing from every realizer of an effective operation F from $\mathbb{N}^+$ to $\mathbb{N}$ a modulus of continuity n at const_0, i.e. an $n \in \mathbb{N}$ such that for all $\alpha \in 0^n$, $F\alpha = F\mathrm{const}_0$. The set const_0 is complemented in $\mathbb{N}^+$ by the finite set $C = \{0^k 10^\omega \mid k < n\}$. Therefore, the *smallest* modulus of continuity of F at const_0 can be computed by evaluating F on all elements of C. Now, the smallest modulus of continuity functional *is* extensional and therefore a realizer for $\mathsf{CP}(\mathbb{N}^+, \mathbb{N})$ in extensional realizability.
(iii) By a theorem of Kleene, there exists a primitive recursive tree $T_K \subseteq 2^*$ (the *Kleene tree*) which is well-founded with respect to recursive branches (in the sense that every recursive 01-valued function eventually leaves the tree), but contains arbitrarily long finite sequences (see e.g. [TvD88b, 4.7.6]). This result can be used to establish the fact that $\mathbb{N} \to 2$ and $\mathbb{N} \to \mathbb{N}$ are homeomorphic (w.r.t. the product topology) in extensional realizability. Therefore, $\mathsf{CP}(\mathbb{N} \to 2, \mathbb{N})$ entails $\mathsf{CP}(\mathbb{N} \to \mathbb{N}, \mathbb{N})$. But $\mathsf{CP}(\mathbb{N} \to \mathbb{N}, \mathbb{N})$ is refuted by $\mathsf{AC}((\mathbb{N} \to \mathbb{N}) \to \mathbb{N}, \mathbb{N})$, as shown in the proof of Proposition 2.3.3 (iii). $\square$

Although $\mathbf{RC}(\mathbf{PER}^0(K_1))$ is not a topos by Theorem 2.2.20, the category $\mathbf{Asm}(\mathbf{PER}^0(K_1))$ *is* the subcategory of $\neg\neg$-separated objects of the *extensional realizability topos* introduced independently in [Gra81] and [Pit81], see also [Oos97a]. Therefore, again, the implications can not be reversed in higher order arithmetic, either.

The idea of searching for a model with the properties described in Proposition 2.3.6 was stimulated by reading an early version of [BS03] and discussions with the authors. Conversely, our Proposition 2.3.6 found a first application in [BS03, Example 4.12], which states the following.

Proposition 2.3.7. *In the extensional realizability topos, the statement*

all functions from $\mathbb{R}$ *to* $\mathbb{R}$ *are sequentially continuous*

holds, whereas the statement

all functions from $\mathbb{R}$ *to* $\mathbb{R}$ *are* ε-δ-*continuous*

does not hold.

As all models mentioned in this section except for $\mathbf{Asm}(K_1)$ and $\mathcal{E}\mathit{ff}$ separate continuity principles which are equivalent relative to Ishihara's boundedness principle, all of these models falsify the latter.

Corollary 2.3.8. *The boundedness principle* $\mathsf{BD\text{-}N}$ *does not hold in the categories* $\mathbf{Asm}(\mathbf{Dom})$*,* $\mathbf{Asm}(\omega\mathbf{ALat})$*,* $\mathbf{Asm}(\mathbf{PER}^0(\mathcal{P}(\omega)))$ *and* $\mathbf{Equ}$*. Moreover,* $\mathsf{BD\text{-}N}$ *does not hold in the toposes* $\mathbf{RC}(\mathbf{Dom})$*,* $\mathbf{RC}(\omega\mathbf{ALat})$ *and the extensional realizability topos.*

Proof. $\mathsf{BD\text{-}N}$ and $\mathsf{CP}(\mathbb{N}^+, \mathbb{N})$ entail $\mathsf{CP}(\mathbb{N} \to \mathbb{N}, \mathbb{N})$. □

The only previously known countermodel of $\mathsf{BD\text{-}N}$ in the literature was Beeson's *fp-realizability*, a notion of realizability combined with formalized provability introduced in [Bee75]. Is is our hope that the further and above all simpler countermodels exhibited here can be put to good use as a test bed for constructive mathematics in the following sense: Ishihara has proved that a number of mathematical statements are equivalent to $\mathsf{BD\text{-}N}$, and thus not provable in pure constructive mathematics. If the exact status of some new statement A with respect to $\mathsf{BD\text{-}N}$ is not yet settled, that is, neither has a purely constructive proof of A, nor a proof of $\mathsf{BD\text{-}N}$ from A been found, then the interpretation in one of the models mentioned here would constitute an alternative handle for the purpose of refuting the provability of A in pure constructive mathematics.

3 Computable Analysis

Computable Analysis is a branch of classical mathematics in which one studies the computability and complexity aspects of analytic theorems. One therefore has to extend classical recursion theory to structures with continuum cardinality. The beginnings of computable analysis can be traced back as far as [Tur36]. Although there exist several non-equivalent approaches to computable analysis, in this thesis we will concentrate on the *Type–2 theory of effectivity (TTE)* approach initiated by Kreitz and Weihrauch (see [Wei00] for a complete presentation).

3.1 Introduction

We shall briefly recall some basic notions and facts of computable analysis.

3.1.1 Numberings

Following Church's thesis[1], we call a function $f : \mathbb{N}^n \rightharpoonup \mathbb{N}$ computable if f is a partial recursive function. In order to extend the notion of computability to other countable sets, we introduce *numberings* of these sets. A numbering of a set X is a partial surjection $\nu : \mathbb{N} \rightharpoonup\!\!\!\rightharpoonup X$. Given numbered sets $(X_1, \nu_1), \ldots, (X_n, \nu_n)$ and (Y, μ), a partial function $f : X_1 \times \cdots \times X_n \rightharpoonup Y$ is computable if and only if there exists a partial recursive function $r : \mathbb{N}^n \rightharpoonup \mathbb{N}$ such that for all $\vec{x} = (x_1, \ldots, x_n) \in \mathbb{N}^n$, $f(\nu_1(x_1), \ldots, \nu_n(x_n)) \preceq \mu(r(\vec{x}))$. This situation is depicted in the following diagram.

$$\begin{array}{ccc} \mathbb{N}^n & \xrightarrow{\quad r \quad} & \mathbb{N} \\ {\scriptstyle \nu_1 \times \cdots \times \nu_n} \downarrow & \preceq & \downarrow {\scriptstyle \mu} \\ X_1 \times \cdots \times X_n & \xrightarrow[\quad f \quad]{} & Y \end{array}$$

For instance, let $\{\cdot\} : \mathbb{N} \twoheadrightarrow \mathcal{PR}$ be some standard numbering of the set of partial recursive functions and let $\mathbb{N}$ be numbered by the identical function. Then the functions

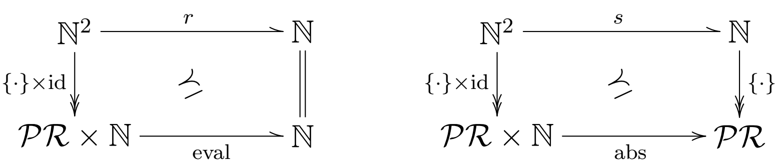
$$\mathrm{eval} : \mathcal{PR} \times \mathbb{N} \rightharpoonup \mathbb{N} \quad (\varphi, x) \mapsto \varphi(x)$$

and

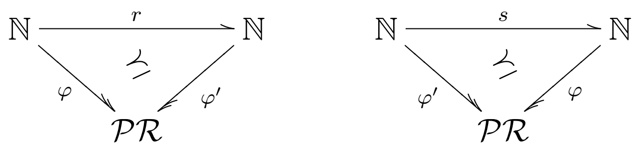
$$\mathrm{abs} : \mathcal{PR} \times \mathbb{N} \longrightarrow \mathcal{PR} \quad (\varphi, x) \mapsto \lambda y.\, \varphi(\mathsf{p}(x, y))$$

[1] not to be confused with the principle CT, introduced in Subsection 1.4.1

(where p is some primitive recursive pairing function) are computable, that is, there exist partial recursive functions $r, s : \mathbb{N}^2 \longrightarrow \mathbb{N}$ such that



Moreover, it is an elementary recursion theoretic fact that, if φ, φ' are numberings of $\mathcal{PR}$ such that the functions eval and abs are computable with respect to both φ and φ', then φ and φ' are equivalent in the sense that there exist partial recursive functions $r, s : \mathbb{N} \longrightarrow \mathbb{N}$ such that $\varphi(x) \preceq \varphi'(r(x))$ and $\varphi'(x) \preceq \varphi(s(x))$ for all $x \in \mathbb{N}$.



The same holds *mutatis mutandis* for numberings of the set of total recursive functions. A numbering meeting these requirement is called an *admissible Gdel numbering.*

3.1.2 Representations

In order to study computability aspects in analysis, numberings can not be employed, as most interesting structures in analysis are of continuum cardinality. We therefore use the elements of *Baire space* $\mathbb{B} = \mathbb{N} \to \mathbb{N}$ in order to represent elements of spaces with continuum cardinality. As is to be expected, topology will play a greater role in the theory of representations than it does in the theory of numberings. We standardly equip Baire space with the product topology, also referred to as initial segment topology or simply Baire topology.

Definition 3.1.1 (Baire space). Baire space ($\mathbb{B}$) is the set of all functions from $\mathbb{N}$ to $\mathbb{N}$, equipped with the initial segment topology, i.e. the subsets of the form $\{a\alpha \mid \alpha \in \mathbb{B}\}$ for $a \in \mathbb{N}^*$ form a basis for the topology.

Even though Baire space is uncountable, there is a self-evident notion of computability for partial functions $r : \mathbb{B} \longrightarrow \mathbb{B}$. We call a partial function $r : \mathbb{B} \longrightarrow \mathbb{B}$ *(continuously) realizable* if it is realized by some $\gamma \in \mathbb{B}$, i.e. $f(\alpha) \simeq \gamma \,|\, \alpha$ for all $\alpha \in \mathbb{B}$. The realizable functions are exactly the continuous functions with a G_δ domain of definition. Every partial continuous function $r : \mathbb{B} \longrightarrow \mathbb{B}$ has a realizable extension. We call a function $r : \mathbb{B} \longrightarrow \mathbb{B}$ *computably realizable* if it is realized by some computable function $\alpha \in \mathbb{B}$.

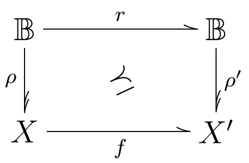
Now that the notions are settled for Baire space, we can extend them to represented spaces.

Definition 3.1.2.

(i) A representation of a set X is a partial surjection $\rho : \mathbb{B} \rightharpoonup X$.

(ii) An element $x \in X$ is called computable with respect to the representation ρ if there is a computable function $\alpha \in \mathbb{B}$ such that $\rho(\alpha) = x$.

(iii) Let (X, ρ) and (X', ρ') represented sets. Then a partial function $f : X \rightharpoonup X'$ is called *(continuously) realizable* if there is a a realizable function $r : \mathbb{B} \rightharpoonup \mathbb{B}$ such that



(iv) In the above situation, the function is called *computably realizable* if there is a computable funtion $r : \mathbb{B} \rightharpoonup \mathbb{B}$ such that the above diagram commutes.

(v) Let $R \subseteq X \times X'$. Then the relation R is called realizable / computably realizable if there is a realizable / computably realizable function $r : \mathbb{B} \rightharpoonup \mathbb{B}$ such that $R(\rho\alpha, \rho'(r\alpha))$ for all $\alpha \in \mathrm{dom}(\rho)$.

(vi) Two representations ρ, ρ' of the same set X are called *topologically equivalent / computably equivalent* if and only if $\mathrm{id}_X : (X, \rho) \longrightarrow (X, \rho')$ and $\mathrm{id}_X : (X, \rho') \longrightarrow (X, \rho)$ are both realizable / computably realizable.

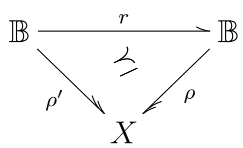
3.1.3 Admissibility

The notion of representation is of course rather general. A goodness to fit criterium that in fact determines the representation of a topological space up to topological equivalence is *admissibility.* The notion of admissibility was introduced by Weihrauch for countably based T_0-spaces and extended to arbitrary spaces by Schrder in [Sch02].

Definition 3.1.3. Let X be a topological space. A representation $\rho : \mathbb{B} \rightharpoonup X$ is called *admissible* if

1. The map $\rho : \mathbb{B} \rightharpoonup X$ is partial continuous.

2. For every partial continuous $\rho' : \mathbb{B} \rightharpoonup X$, there is a partial continuous map $r : \mathbb{B} \rightharpoonup \mathbb{B}$ such that $\rho'(\alpha) \preceq \rho(r(\alpha))$ for all $\alpha \in \mathbb{B}$.



A signed digit representation of the real numbers is an example for an admissible representation. It is well known that with respect to the unsigned decimal digit representation, multiplication by three is not realizable. By the following theorem, it follows that the unsigned digit representation is an example of a non-admissible representation.

Theorem 3.1.4. *Let (X, ρ_X) and (Y, ρ_Y) be admissibly represented topological spaces. Then a function $f : X \rightharpoonup Y$ is continuously realizable if and only if it is sequentially continuous.*

Proof. See [Sch02]. □

A subset O of a topological space X is sequentially open if, whenever a sequence (x_n) of elements of X converges to some $x \in O$, then all but finitely many elements of the sequence (x_n) are in O. The sequentially open sets of a topological space form a topology, which is (possibly) finer than the original topology. By $\mathrm{seq}(X)$ we denote the *sequentialization of* X, that is the set X, equipped with the topology of sequentially open sets. A topological space is sequential if it coincides with its sequentialization, i.e., if every sequentially open set is open.

Theorem 3.1.5. *Given a partial surjection $\rho : \mathbb{B} \rightharpoonup X$ then*

(i) The map $\rho : \mathbb{B} \rightharpoonup \mathrm{seq}(X)$ is continuous if and only if $\rho : \mathbb{B} \rightharpoonup X$ is.

(ii) The map $\rho : \mathbb{B} \rightharpoonup \mathrm{seq}(X)$ is an admissible representation if and only if $\rho : \mathbb{B} \rightharpoonup X$ is.

(iii) If ρ is admissible, then the topology $\mathrm{seq}(X)$ coincides with the final topology induced by ρ.

Proof. See [Sch02]. □

Remark 3.1.6. By the above theorem, the property of ρ being an admissible representation is independent of the topology on X in the following sense: if $\rho : \mathbb{B} \rightharpoonup X$ is admissible, then ρ is admissible with respect to *all* topologies on X such that ρ is continuous, that is, with respect to all topologies that are coarser than the final topology induced by ρ. Thus, the notion of admissibility makes sense even for represented sets as opposed to represented spaces. Note also that, up to sequentialization, the original topology on X can be recovered from the final topology induced by ρ.

The class of admissible representations is closed under various constructions. We cite the two closure properties needed in this thesis.

Theorem 3.1.7.

(i) Let (X, ρ) be an admissible representation. Then for every $S \subseteq X$ the restriction $(S, \rho|^S)$ of ρ to S is an admissible representation.

(ii) *Let* (X, ρ_X), (Y, ρ_Y) *be a admissible representations Then the representation* $[\rho_X, \rho_Y]$ *of the cartesian product of* X *and* Y *(defined in [Wei00, Definition 3.3.3]) is admissible.*

(iii) *Let* (X, ρ_X), (Y, ρ_Y) *be representations and assume that* (Y, ρ_Y) *is admissible. Then the representation* $[\rho_X \to \rho_Y]_X$ *of the space of continuously realizable maps (defined in [Wei00, Definition 3.3.13]) is admissible. The final topology induced by* $[\rho_X \to \rho_Y]_X$ *is the sequential topology whose convergence relation is defined as*

$$(f_n) \longrightarrow f \quad :\Longleftrightarrow \quad \textit{for all } (x_n) \subseteq X \textit{ if } (x_n) \longrightarrow_X x \textit{ then } (f_n(x_n)) \longrightarrow_Y f(x)$$

Proof. (i) and (ii) are trivial. (iii) see [Sch03, Proposition 4.2.5]. □

3.2 Category theoretic approach to representations

In this section we will make use of the fact that the category of representations and continuously realizable maps has a rich logical structure. The standard constructions known from computable analysis can be rediscovered via their universal properties. An abstract proof of Schrder's theorem that every T_0–quotient of a subset of Baire space has an admissible representation can be conducted. Finally, a logical characterization of admissible representations will be given.

3.2.1 The category of representations

Let us denote by **Rep** the category of represented sets and continuously realizable maps. By Definition 3.1.2, a representation of a set X is defined to be a partial surjection $\rho_X : \mathbb{B} \rightharpoonup X$. Let us define the relation $\Vdash_X \subseteq \mathbb{B} \times X$ by $\alpha \Vdash_X x :\Longleftrightarrow \rho_X(\alpha) = x$. Then $(X, \Vdash_X)$ is a discrete set over the pca K_2. Conversely, if $(X, \Vdash_X)$ is a discrete set, then the inverse relation ρ_X of $\Vdash_X$ is a partial surjection $\rho_X : \mathbb{B} \rightharpoonup X$.

Now, let (X, ρ_X) and (Y, ρ_Y) be represented spaces, and let $(X, \Vdash_X)$ and $(Y, \Vdash_Y)$ be the corresponding discrete sets. Then a function $f : X \longrightarrow Y$ is continuously realizable with respect to ρ_X and ρ_Y if and only if it is a morphism from $(X, \Vdash_X)$ to $(Y, \Vdash_Y)$ in $\mathbf{D}(K_2)$. That is, the categories **Rep** and $\mathbf{D}(K_2)$ are equivalent. Categories of discrete sets are well known to have very pleasent logical properties. See [Bau00] for an extensive treatment. As we try to derive results for computable analysis, which is a classical discipline, we use classical set theory with choice as our metalogic. The interpretation in $\mathbf{Rep} \simeq \mathbf{D}(K_2)$ validates the generalized continuity principle and Markov's principle by Proposition 2.1.28 and it validates bar induction by Proposition 2.1.21. That is, **Rep** is in particular a model of Intuitionism as described in Section 1.2.2.

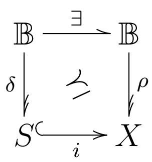
The standard representations of $\mathbb{N}$ and $\mathbb{R}$ are the category theoretic natural number object and real number object, respectively.[2] Standard representations of the product of two represented spaces and the space of continuously realizable maps between two represented spaces have been defined in [Wei00, Definition 3.3.3] and [Wei00, Definition 3.3.13], respectively. These particular choices of representations are well justified from a category theoretic point of view.

Proposition 3.2.1. *Let (X, ρ_X) and (Y, ρ_Y) be represented sets. Then the standard representation $[\rho_X, \rho_Y]$ of the set-theoretic product $X \times Y$ is the category theoretic product $(Y, \rho_Y) \times (X, \rho_X)$ and the standard representation $[\rho_X \to \rho_Y]_X$ of the space of continuously realizable maps from X to Y is the category theoretic exponent $(Y, \rho_Y)^{(X, \rho_X)}$.*

Proof. The universal property of the representation $[\rho_X, \rho_Y]$ follows from [Wei00, Lemma 3.3.4]. The universal property of the representation $[\rho_X \to \rho_Y]_X$ follows from [Wei00, Lemma 3.3.14 and Theorem 3.3.15]. □

Remark 3.2.2. In the light of the preceeding proposition, Theorem 3.1.7 (iii) states that the admissible representations form an exponential ideal in the category of representations.

In order to familiarize ourselves with the internal logic of the category of representations, we shall have a look at what subobjects look like concretely. A subobject of a represented space (X, ρ) is a represented space (S, δ) and a continuously realizable injection $i : S \rightarrowtail X$. Without loss of generality, we can (and will henceforth) assume that S is a subset of X and i is the subset inclusion of S into X.



A subobject S of an object X is called *stable* if the sentence

$$\forall x \in X.\ \neg\neg(x \in S) \to (x \in S)$$

holds in the internal logic and *dense* if the sentence

$$\forall x \in X.\ \neg\neg(x \in S)$$

[2]By real number object, we refer to the Cauchy construction of the real numbers as opposed to the Dedekind construction, as the category $\mathbf{D}(K_2)$ provides no means to interpret higher order arithmetic. Nevertheless, the Dedekind construction of the real numbers, interpreted in the topos $\mathbf{RC}(K_2)$, which contains the category $\mathbf{D}(K_2)$ as the subcategory of discrete objects, will yield an object isomorphic to the Cauchy construction, as the axiom of countable choice holds in $\mathbf{RC}(K_2)$.

holds in the internal logic of the category. In the case of the category of represented sets, the stable subobjects of (X, ρ) are up to isomorphism exactly those subobjects (S, δ), $S \subseteq X$, where δ is the restriction $\rho|^S$ of ρ to S. The dense subobjects of (X, ρ) are exactly those subobjects (S, δ) where $S = X$.

For instance, the interpretation of

$$\{x \in \mathbb{R} \mid x < 0 \vee x \geq 0\} \subseteq \mathbb{R}$$

is a dense subset of $\mathbb{R}$, as $\forall x \in \mathbb{R}. \neg\neg(x < 0 \vee x \geq 0)$ holds in the category of representations (as it is provable constructively). On the other hand, the subobject is not stable, as $\forall x \in \mathbb{R}. x < 0 \vee x \geq 0$ does not hold. On the other hand, $\mathbb{R}_{>0}$ and $\mathbb{R}_{\geq 0}$ are stable subobjects of $\mathbb{R}$ (the proof of the stability of $\mathbb{R}_{>0}$ requires classical metalogic), but obviously not dense subobjects.

Remark 3.2.3. In the light of the preceeding comments, Theorem 3.1.7 (i) states that the class of admissible representations is closed under stable subobjects in the category of representations.

We will now define an admissible representation of a non T_1-space that plays an important role in the categorical approach to the theory of representations.

Definition 3.2.4. The Sierpinski space Σ is the set $\{\bot, \top\}$, equipped with the topology $\{\emptyset, \{\top\}, \{\bot, \top\}\}$.

The Sierpinski space has the important property that for every topological space X, the constinuous functions $f : X \longrightarrow \Sigma$ correspond exactly to the open subsets of X. An admissible representation of Σ can be defined by

$$\rho(\alpha) = \begin{cases} \top & \text{if } \exists x : \mathbb{N}. \alpha(x) > 0 \\ \bot & \text{if } \forall x : \mathbb{N}. \alpha(x) = 0 \end{cases}$$

The standard representation of Σ is definable in the internal logic of the category of representations as the quotient set

$$\Sigma = 2^{\mathbb{N}} / \sim$$

where $2 = \{0, 1\}$ and $\alpha \sim \beta$ is defined as $(\forall x : \mathbb{N}.\ \alpha x = 0) \leftrightarrow (\forall x : \mathbb{N}.\ \beta x = 0)$. The quotient exists, as the relation $\sim$ is $\neg\neg$-stable. The object Σ is a dominance in the sense of Rosolini, see [Bau00, Section 5.3].

3.2.2 Logical aspects of admissibility

The following theorem by Schröder characterizes those topological spaces that have an admissible representation.

Theorem 3.2.5 (Schröder). *A sequential space has an admissible representation if and only if it is T_0 and is the topological quotient of a subset of Baire space.*

Proof. See [Sch03, Theorem 3.2.4]. Actually, the cited theorem states that a sequential space has an admissible representation if and only if it is T_0 and is the topological quotient of a countably based T_0–space. As every countably based T_0–space is a quotient of a subspace of Baire space, both statements are easily seen to be equivalent. □

By Remark 3.1.6, the admissibility of a representation does not depend on the topology on X, but is a property of the represented set (X, ρ_X), alone. We will reprove the harder direction of Schröder's theorem, using the internal logic of the category of represented sets and concepts of synthetic domain theory (cf. [Hyl91]). As a corollary, we obtain a logical characterization of the admissibile representations.

Theorem 3.2.6. *Every T_0–quotient of a subspace of Baire space has an admissible representation.*

Proof. A quotient of a subset of Baire space is nothing but a represented set, equipped with its final topology. For a represented set X, regardless of whether the representation is admissible, by the universal property of the final topology, the set of morphisms from X to $\mathbf{\Sigma}$ corresponds to the set of open sets of X with respect to the final topology induced by ρ_X. Thus, the object $\mathbf{\Sigma}^X$ is a representation of the set $\mathcal{O}(X)$ of open subsets of X which is admissible by Remark 3.2.2to be defined . The convergence relation on $\mathcal{O}(X)$ can be described as follows: a sequence (O_n) of open subsets of X converges to $O \in \mathcal{O}(X)$ if for every sequence (x_n) in X, whenever (x_n) converges to x (with respect to the final topology on X induced by ρ_X) and $x \in O$, then there exists $n_0 \in \mathbb{N}$ such that $x_n \in O_n$ for all $n \geq n_0$.

Let $\eta_X : X \longrightarrow \mathbf{\Sigma}^{\mathbf{\Sigma}^X}$ be the transpose of the evaluation function $\varepsilon_X : \mathbf{\Sigma}^X \times X \longrightarrow \mathbf{\Sigma}$. For every $x \in X$, $\eta_X(x)(O) = \top$ if and only if $x \in O$. Therefore, the map η_X is a mono if and only if the final topology on X is T_0.

Given a represented set X, we can factor the map η_X as follows,

$$\begin{array}{ccc} X & \xrightarrow{\eta_X} & \mathbf{\Sigma}^{\mathbf{\Sigma}^X} \\ & {\scriptstyle e}\searrow \quad \nearrow{\scriptstyle \iota} & \\ & \overline{X} & \end{array}$$

where $\overline{X} = \{\varphi \in \Sigma^{\Sigma^X} \mid \neg\neg\exists x : X.\, \eta_X x = \varphi\}$. As $\mathbf{\Sigma}$ is an admissible representations and as admissible representations form an eponential ideal by Remark 3.2.2, $\mathbf{\Sigma}^{\mathbf{\Sigma}^X}$ is admissible. Being a stable subobject of an admissible representation, $\overline{X}$ is an admissible representation, too, by Remark 3.2.3. The morphism $e : X \longrightarrow \overline{X}$ is epic, as its image is dense. Moreover, it is monic if η_X is. Therefore, if X is T_0, then the underlying map of the morphism e is a bijection.

It remains to be shown that e is also a homeomorphism from X to $\overline{X}$ with respect to the final topologies induced by ρ_X and $\rho_{\overline{X}}$, respectively. This is tantamount to the underlying map of $\Sigma^e : \Sigma^{\overline{X}} \longrightarrow \Sigma^X$ being a bijection. In fact, we can show that Σ^e is always an isomorphism. A morphism $f : A \longrightarrow B$ such that Σ^f is an isomorphism is called Σ-*equable*. We will demonstrate the Σ–equability of Σ^e by showing separately that Σ^e is monic and split epic.

As to the fact that Σ^e is monic, we prove in the internal logic of **Rep** that Σ^e is one–one. Let $f, f' \in \Sigma^{\overline{X}}$ such that $f \circ e = f' \circ e$. We have to show that $f = f'$. Let $z \in \overline{X}$. Then

$$(\exists x : X.e(x) = z) \rightarrow f(z) = f'(z)$$

As equality on Σ is stable, it is also true that

$$(\neg\neg\exists x : X.e(x) = z) \rightarrow f(z) = f'(z)$$

But the above hypothesis is true by definition for all elements of $\overline{X}$ and so in particular for z.

The fact that Σ^e is split epic follows by pure category theory. The contravariant functor $\Sigma^{(\_)}$ is self–adjoint on the right. The triangular equation applied to this adjunction yields

$$\Sigma^{\eta_X} \circ \eta_{\Sigma^X} = \mathrm{id}_{\Sigma^X}$$

As $\eta_X = \iota \circ e$, by functoriality of $\Sigma^{(\_)}$ it is the case that $\Sigma^e \circ \Sigma^\iota = \Sigma^{\eta_X}$ and therefore By the triangular equation of this adjunction

$$\Sigma^e \circ \Sigma^\iota \circ \eta_{\Sigma^X} = \mathrm{id}_{\Sigma^X}$$

which finishes the proof. □

The real advantage of the purely logical description of the admissible representation corresponding to some arbitrary representation is that we can now express the admissibility of a representation in terms of the internal language of **Rep**.

Theorem 3.2.7. *An object X in* **Rep** *is an admissible representation if and only if*

$$\forall \varphi : \Sigma^{\Sigma^X}. (\neg\neg\exists x : X.\, \eta_X x = \varphi) \rightarrow (\exists! x : X.\, \eta_X x = \varphi)$$

is satisfied in the internal logic of **Rep**.

Proof. The formula expresses that the morphism $e : X \longrightarrow \overline{X}$ is an iso. This is both sufficient and (because of the uniqueness of admissible representations of a topological space) necessary. □

We will now describe the object Σ^X more concretely for the case that X is a complete separable metric space. Mind that in the presence of the continuity principle WC-N, any two metrics rendering X a complete separable metric space can be shown to be topologically equivalent. This means that with respect to the internal logic of **Rep**, being a csm is a property of the set X alone, and not a property of some additional structure on X. We already know that the morphisms from X to Σ correspond to the open subsets of X with respect to the final topology induced by ρ_X. In the internal logic of the category of representations, we can prove that if X is a csm, then the elements of Σ^X can be identified with those subsets of X that are countable unions of open balls.

Proposition 3.2.8. *Let (X, d) be a csm with a dense sequence (q_i) and let $S \subseteq X$ be a subset of X. Consider the following two statements about the subset S.*

(1) There exists $\alpha : \mathbb{N} \longrightarrow \mathbb{N}$ with

$$\forall x : X.\, x \in S \longleftrightarrow \exists n : \mathbb{N}.\, \alpha n > 0 \wedge x \in B(q_{\mathsf{p}_0(\alpha n - 1)}, 2^{-\mathsf{p}_1(\alpha n - 1)}).$$

(2) There exists $\chi : X \longrightarrow \Sigma$ such that

$$\forall x : X.\, x \in S \longleftrightarrow \chi(x) = \top.$$

The implication (1) $\Rightarrow$ (2) *holds constructively. Moreover, χ is uniquely determined by S. As to the reverse direction, under the additional assumption of* CONT*, the implication* (2) $\Rightarrow$ (1) *holds constructively.*

Proof. The membership of some $x \in X$ in the union of the balls $B(q_{\mathsf{p}_0(\alpha n)}, 2^{-\mathsf{p}_1(\alpha n)})$ can be expressed as a Σ^0_1–statement, i.e., given x and α, one can construct a sequence β auch that $x \in S$ if and only if there is some $n \in \mathbb{N}$ such that $\beta(n) > 0$. This is tedious, but straightforward. As to the converse direction, for the space (X, d) there exists a representation ρ with the properties described in Proposition 1.4.1. By CONT, we can find a neighborhood function for $\chi \circ \rho$. From such a neighborhood function, one can construct an appropriate enumeration of open balls with the help of [TvD88b, 7.2.4 (v)]. □

As a first application of the logical characterization of admissible representations we show constructively that whenever X is a csm, then it satisfies the characterization formula for admissible representations.

Theorem 3.2.9. *The statement*

$$(X, d) \textit{ is a csm} \quad \rightarrow \quad \forall \varphi : \Sigma^{\Sigma^X}.\, (\neg\neg\exists x : X.\, \eta_X x = \varphi) \rightarrow (\exists! x : X.\, \eta_X x = \varphi)$$

is constructively provable under the assumption of Markov's principle.

Proof. For the sake of brevity, we will write η for η_X. Let $\varphi \in \mathbf{\Sigma}^{\mathbf{\Sigma}^X}$ and assume $\neg\neg\exists x : X.\, \eta x = \varphi$. Furthermore, let $(q_n)_{n\in\mathbb{N}}$ be a dense sequence in X. By Proposition 3.2.8, every open ball $B(c, r)$ in X induces an element of $\mathbf{\Sigma}^X$, to which we can apply $\varphi \in \mathbf{\Sigma}^{\mathbf{\Sigma}^X}$. We will construct a sequence $(x_k)_{k\in\mathbb{N}}$ such that $\varphi(B(x_k, 2^{-k})) = \top$ and $d(x_k, x_{k+1}) < 2^{-k}$ for all $k \in \mathbb{N}$.

We define the sequence $(q_n^{(0)})_{n\in\mathbb{N}}$ by setting $q_n^{(0)} = q_n$ for all $n \in \mathbb{N}$. The family $(B(q_n^{(0)}, 2^0))_{n\in\mathbb{N}}$ covers X. Hypothetically, from the assumption $\exists x : X.\, \eta x = \varphi$ we can conclude

$$\exists n : \mathbb{N}.\, \varphi(B(q_n^{(0)}, 2^0)) = \top,$$

as $(\eta x)(B(c, r)) = \top \leftrightarrow x \in B(c, r)$. Therefore, our assumption $\neg\neg\exists x : X.\, \eta x = \varphi$ entails

$$\neg\neg\exists n : \mathbb{N}.\, \varphi(B(q_n^{(0)}, 2^0)) = \top,$$

which, however, by Markov's principle yields

$$\exists n : \mathbb{N}.\, \varphi(B(q_n^{(0)}, 2^0)) = \top.$$

Set $x_0 = q_n^{(0)}$ for the n found above.

Assume now that $x_k \in X$ has been found such that $\varphi(B(x_k, 2^{-k})) = \top$. In order to find x_{k+1}, first assume hypothetically that $\exists x : X.\, \eta x = \varphi$. For such x one can show $(\eta x)(B(x_k, 2^{-k})) = \varphi(B(x_k, 2^{-k})) = \top$, which is equivalent to $x \in B(x_k, 2^{-k})$. In particular, we can conclude $\exists x : X.\, x \in B(x_0, 2^{-k}) \wedge \eta x = \varphi$ from the temporary hypothesis. Thus, from the main assumption of this proof we can conclude

$$\neg\neg\exists x : X.\, x \in B(x_k, 2^{-k}) \wedge \eta x = \varphi$$

Let $(q_n^{(k+1)})_{n\in\mathbb{N}}$ be some enumeration of all elements of the sequence $(q_n)_{n\in\mathbb{N}}$ that are in $B(x_k, 2^{-k})$. The family $(B(q_n^{(k+1)}, 2^{-(k+1)}))_{n\in\mathbb{N}}$ covers $B(x_k, 2^{-k})$. By the same argument as before, we can conclude

$$\exists n : \mathbb{N}.\, \varphi(B(q_n^{(k+1)}, 2^{-(k+1)})) = \top.$$

Set $x_{k+1} = q_n^{(k)}$ for the n found above.

By dependent choice, we obtain a sequence (x_k) such that $\varphi(B(x_k, 2^{-k})) = \top$ and $d(x_k, x_{k+1}) < 2^{-k}$ for all $k \in \mathbb{N}$. Set $\hat{x} = \lim_{k\to\infty} x_k$. We claim that $\eta\hat{x} = \varphi$. Assume $\exists x : X.\, \eta x = \varphi$. For such x, as $(\eta x)(B(x_k, 2^{-k})) = \varphi(B(x_k, 2^{-k})) = \top$, we conclude $x \in B(x_k, 2^{-k})$, i.e. $d(x, x_k) < 2^{-k}$ for all $k \in \mathbb{N}$. Therefore $x = \hat{x}$ and thus $\eta\hat{x} = \eta x = \varphi$. As equality on $\mathbf{\Sigma}^{\mathbf{\Sigma}^X}$ is stable, $\eta\hat{x} = \varphi$ and the unicity of $\hat{x}$ follow simultaneously from the assumption $\neg\neg\exists x : X.\, \eta x = \varphi$. □

Remark 3.2.10. As a consequence of Theorem 3.2.7 and 3.2.9, whenever (the syntactic description of) a set can be shown to be a csm with respect to the internal logic of the category **Rep**, then its interpretation in **Rep** is an admissible representation. In particular, this is true for every set such that it is provable in Intuitionism that the set can be equipped with a metric rendering it a csm.

Many important spaces used in functional analysis are not metrizable but fall into the following, more general class of spaces.

Definition 3.2.11 (Inductive Limit of CSMs). Let (X, d) be a metric space and let $(X_n)_{n\in\mathbb{N}}$ be a sequence of subsets of X such that

1. $X_n \subseteq X_{n+1}$
2. (X_n, d) is a complete separable metric space

for all $n \in \mathbb{N}$. The inductive limit topology on X is defined as follows. A set $O \subseteq X$ is declared open if and only if $O \cap X_n$ is open in (X_n, d) for all $n \in \mathbb{N}$.

Note that in the above definition we assume that all subsets X_n are csm's with respect to the restriction of one and the same metric d on X. We will meet examples of inductive limits of csm's in Section 3.3.3. In the following proposition, we describe the object $\mathbf{\Sigma}^X$ if X is an inductive limit of csm's.

Proposition 3.2.12. *Let (X, d) be a metric space and let $(X_n)_{n\in\mathbb{N}}$ be an increasing sequence of subsets of X such that $(q_i^{(n)})_{i\in\mathbb{N}}$ is a dense sequence in (X_n, d) and (X_n, d) is complete for all $n \in \mathbb{N}$. Let $S \subseteq X$ be a subset of X. Consider the following two statements about the subset S.*

(1) There exists an $\alpha : \mathbb{N} \times \mathbb{N} \longrightarrow \mathbb{N}$ with

$$\forall n : \mathbb{N}\, \forall x : X_n.\, x \in S \cap X_n$$
$$\longleftrightarrow \exists k : \mathbb{N}.\, \alpha(n, k) > 0 \wedge x \in B(q^{(n)}_{\mathsf{p}_0(\alpha(n,k)-1)}, 2^{-\mathsf{p}_1(\alpha(n,k)-1)}).$$

(2) There exists a $\chi : X \longrightarrow \mathbf{\Sigma}$ such that

$$\forall x : X.\, x \in S \longleftrightarrow \chi(x) = \top.$$

The implication (1) $\Rightarrow$ (2) *holds constructively. Moreover, χ is uniquely determined by S. As to the reverse direction, under the additional assumption of* CONT, *the implication* (2) $\Rightarrow$ (1) *holds constructively.*

Proof. From (1) we can conclude by Proposition 3.2.8 that there is a family of maps $\chi_n : X_n \longrightarrow \mathbf{\Sigma}$ characterizing the sets $S \cap X_n$. This family of maps gives rise to a unique $\chi : X \longrightarrow \mathbf{\Sigma}$ such that each χ_n is the restrictions of χ to X_n. For the converse direction, simply apply Proposition 3.2.8 to the family of restrictions of χ to all X_n. □

It would be desirable to prove an extension of Theorem 3.2.9 to inductive limits of csm's. The proof is invariably harder in this case, as it would have to necessarily involve to show the convergence of a net or a filter (as opposed to the convergence of a sequence). For the time being, we therefore have to content ourselves with the following theorem.

Theorem 3.2.13. *Let X be an object of* **Rep**. *If (X, d) is an inductive limit of csm's (X_n, d) with respect to the internal logic of* **Rep**, *then X is admissible.*

Proof. The formula $\forall x : X\, \exists n : \mathbb{N}.\, x \in X_n$ holds with respect to the internal logic of **Rep**. The interpretation in **Rep** of the existential quantifier corresponds exactly to the definition of the representation δ in line 3 of [Sch02, Theorem 19] if we let δ_n be the interpretation of X_n. □

3.3 Categorical approach to computable analysis

3.3.1 Motivation

In the previous two sections we have talked mainly about *continuous* realizability of functions. In computable analysis, one examines which functions and operations in classical analysis are *computable* with respect to the standard representations on the spaces involved. The standard representations are defined in such a way that the canonical operations on these spaces are computable.

The standard representation $\rho_\mathbb{R}$ on the space of real numbers equipped with the euclidean topology is defined as in example 1.2.2. Its domain of definition is (defined by) a spread. The standard representation $\rho_\mathbb{C}$ on the complex numbers is induced by the product representation $[\rho_\mathbb{R}, \rho_\mathbb{R}]$ on $\mathbb{R}^2$.

With respect to these representations, the computable elements of $\mathbb{R}$ are exactly the computable real numbers as defined by Turing. These form a real closed field, containing amongst others all algebraic numbers, π and e. As to computable functions from $\mathbb{R}$ to $\mathbb{R}$ (or $\mathbb{C}$ to $\mathbb{C}$) with respect to the standard representations, polynomials with computable coefficients are computable functions, so is the exponential function and hence so are the trigonometric functions. For $n \in \mathbb{N}$, $n > 0$, define $R \subseteq \mathbb{C}^n \times \mathbb{C}$ by

$$R(a_0, \ldots, a_{n-1}, a) \quad \rightleftharpoons \quad a_0 a^0 + a_1 a^1 + \cdots + a_{n-1} a^{n-1} + a^n = 0$$

The relation R is computably realizable, that is, there is a computable realizer which computes a realizer for a zero of the polynomial from realizers of its coefficients. Different realizers for the same coefficients will possibly result in a realizer for a different zero of the polynomial.

We would like to achieve this type of result in an abstract way, by interpreting constructive mathematics in categorical realizability models. Moreover, we would like to systematically obtain representations of spaces with desirable properties by interpreting an appropriate constructive description of the underlying set of the space in a suitable realizability category. The connection between constructive analysis and computable analysis has already been pointed out in [Tro92].

3.3.2 Models

The categories we will use are variants of the category of assemblies over the second Kleene algebra K_2.

Definition 3.3.1.

(i) An object of the category $\mathbf{Asm}(K_2)$ is a set X together with a relation $\Vdash_X \subseteq \mathbb{B} \times X$ such that for all $x \in X$ there is an $\alpha \in \mathbb{B}$ such that $\alpha \Vdash_X x$. A morphism from $(X, \Vdash_X)$ to $(Y, \Vdash_Y)$ is a function $f : X \longrightarrow Y$ such that there exists $\gamma \in \mathbb{B}$ such that if $\alpha \Vdash_X x$ then $\gamma|\alpha \Vdash_X f(x)$.

(ii) The objects of the category $\mathbf{Asm}(K_2, K_2^{\text{eff}})$ are those of $\mathbf{Asm}(K_2)$. A morphism from $(X, \Vdash_X)$ to $(Y, \Vdash_Y)$ is a function $f : X \longrightarrow Y$ such that there exists a *computable* $\gamma \in \mathbb{B}$ such that if $\alpha \Vdash_X x$ then $\gamma|\alpha \Vdash_X f(x)$.

(iii) An object of the category $\mathbf{Asm}_t(K_2)$ is a set X together with a relation $\Vdash_X \subseteq \mathbb{B} \times X$. A morphism from $(X, \Vdash_X)$ to $(Y, \Vdash_Y)$ is a function $f : X \longrightarrow Y$ such that there exists $\gamma \in \mathbb{B}$ such that if $\alpha \Vdash_X x$ then $\gamma|\alpha \Vdash_X f(x)$.

(iv) The objects of the category $\mathbf{Asm}_t(K_2, K_2^{\text{eff}})$ are those of $\mathbf{Asm}_t(K_2)$. A morphism from $(X, \Vdash_X)$ to $(Y, \Vdash_Y)$ is a function $f : X \longrightarrow Y$ such that there exists a *computable* $\gamma \in \mathbb{B}$ such that if $\alpha \Vdash_X x$ then $\gamma|\alpha \Vdash_X f(x)$.

The category $\mathbf{Asm}(K_2)$ is the semantic version of the <u>`rf`</u>–interpretation. It is equivalent to the category of multirepresentations (cf. [Sch03]) and continuously realizable maps. Its full subcategory $\mathbf{D}(K_2)$ is equivalent to the category of representations. The category $\mathbf{Asm}(K_2, K_2^{\text{eff}})$ is called the *relative realizability category over* K_2 *and* K_2^{eff} (cf. [Bir99]). This category is equivalent to the category of multirepresentations and computably realizable maps. The inclusion functor of $\mathbf{Asm}(K_2, K_2^{\text{eff}})$ into $\mathbf{Asm}(K_2)$ is logical. The category $\mathbf{Asm}_t(K_2)$ is the semantic version of the <u>`rft`</u>–interpretation. Its subcategory $\mathbf{Asm}_t(K_2, K_2^{\text{eff}})$ is a combination of relative and truth–realizability. Again, the inclusion functor of $\mathbf{Asm}_t(K_2, K_2^{\text{eff}})$ into $\mathbf{Asm}_t(K_2)$ is logical.

As computable analysis is a branch of classical mathematics, we must make sure that the sets and functions we reason about correspond to actual sets and functions in the world of classical mathematics. While the internal logics of $\mathbf{Asm}(K_2)$ and $\mathbf{Asm}(K_2, K_2^{\text{eff}})$ validate principles like the continuity principle, which are incompatible with classical mathematics, this is not the case for their truth–variants $\mathbf{Asm}_t(K_2)$ and $\mathbf{Asm}_t(K_2, K_2^{\text{eff}})$. As the forgetful functor $U : \mathbf{Asm}_t(K_2) \longrightarrow \mathbf{Set}$ is logical, the interpretation of the description of a set will result in just that set, equipped with some realizability relation. On the other hand, the category $\mathbf{Asm}(K_2)$ is well suited for reasoning about the intrinsic topology and the admissibility of some object. Fortunately, for many spaces of interest in analysis the interpretations in both categories will coincide.

Proposition 3.3.2. *For a predicate $R \subseteq \mathbb{B} \times \mathbb{B}$ definable in elementary analysis (EL), assume that*

1. $R \in \mathrm{CC}(\underline{\mathtt{rf}})$
2. *EL proves that R is a partial equivalence relation.*
3. *EL proves $\neg\neg R(x, y) \rightarrow R(x, y)$.*

Then the interpretations of the quotient set $\mathbb{B}/R$ in all categories of Definition 3.3.1 yield computably equivalent realizability structures.

Proof. The proposition follows from the fact that the toposes in question are the semantic variants of the $\underline{\mathtt{rf}}$ and $\underline{\mathtt{rft}}$ interpretations and that these coincide for $\mathrm{CC}(\underline{\mathtt{rf}})$–formulas. Condition 3 guarantees that the subquotient can be interpreted in these categories. □

Proposition 3.3.3. *Assume that EL proves that (X, d) is a csm. Then the interpretations of (X, d) in all categories of Definition 3.3.1 yield computably equivalent realizability structures.*

Proof. As shown implicitly in [TvD88b, Proposition 7.2.4], there is a partial equivalence relation R on $\mathbb{B}$ which is negative such that X can be identified with $\mathbb{B}/R$. Every negative formula in particular belongs to $\mathrm{CC}(\underline{\mathtt{rf}})$. Moreover, every negative formula is stable. □

The previous proposition can be generalized to inductive limits of csm's.

Proposition 3.3.4. *Assume that EL proves that (X, d) is an inductive limit of csm's with respect to the sequence of subsets $(X_n)_{n\in\mathbb{N}}$. Then the interpretations of (X, d) in all categories of Definition 3.3.1 yield computably equivalent realizability structures.*

Proof. For all $n \in \mathbb{N}$, let $(q_i^{(n)})_{i\in\mathbb{N}}$ be a dense sequence in X_n. Similar to Section 7.2.3 of [TvD88b], there is a function $\alpha : \mathbb{N}^5 \longrightarrow \mathbb{Q}$ such that

$$\left| d(q_i^{(n)}, q_{i'}^{(n')}) - \alpha(n, i, n', i', k) \right| < 2^{-k}$$

A sequence β is in the domain $|R|$ of R, the partial equivalence relation to be defined, if and only if

$$\forall k > 0.\, \alpha(\beta 0, \beta k, \beta 0, \beta(k+1), k+1) < 2^{-k+1}$$

Two sequences β, β' of $|R|$ are equivalent with respect to R if

$$\forall k > 0.\, \alpha(\beta 0, \beta k, \beta' 0, \beta' k, k) < 2^{-k+3}$$

The partial equivalence relation R is hence negative. Moreover, $\mathbb{B}/R$ can be identified with X by associating with a sequence $\beta \in |R|$ the limit of the sequence $(q^{(\beta 0)}_{\beta(i+1)})_{i \in \mathbb{N}}$ in $X_{\beta 0}$. □

We will summarize our results in the following theorem.

Theorem 3.3.5. *Assume that EL proves that (X, d) is an inductive limit of csm's with respect to the sequence of subsets $(X_n)_{n \in \mathbb{N}}$. Then*

(i) The interpretations of X in all categories of Definition 3.3.1 yield computably equivalent realizability structures. In particular, the underlying set of the interpretation of X in $\mathbf{Asm}(K_2)$ *and* $\mathbf{Asm}(K_2, K_2^{\text{eff}})$ *coincides with the classical interpretation.*

(ii) The transpose of the realizability relation on the interpretation of X is an admissible representation.

(iii) The final topology induced by the representation is the inductive limit topology.

Proof. (i) is the contents of Proposition 3.3.4.
(ii) As the interpretation of EL in $\mathbf{Asm}(K_2)$ takes place in the subcategory of discrete objects, the transpose of the realizability relation is a partial surjective function. By Theorem 3.2.13, this function is an admissible representation.
(iii) follows from Proposition 3.2.12. □

The above theorem states that in many interesting cases, in particular for all separable Banach spaces,[3] we do not have to define a representation *ad hoc*, but the realizability interpretation systematically provides us with a good representation. As all separable Banach spaces are csm's, in the next section we shall have a closer look at two examples of inductive limits of csm's, which are not metrizable.

3.3.3 Examples of inductive limits of csm's

Dual spaces

By the dual of a normed space we mean the set of all continuous linear functionals. In constructive analysis it cannot be proved that every such functional has a norm, as this would require bounded completeness of the real numbers, which is not constructively true. We can, however, prove that the dual of a separable normed space is an inductive limit of csm's. We take the following definition from [BB85, Section 7.6].

[3] at least those where this property can be demonstrated constructively

Definition 3.3.6. Let X be a separable normed space. For a dense sequence $(x_n)_{n\in\mathbb{N}}$, we define the corresponding *double norm* on X^* by

$$|||u||| = \sum_{n=1}^{\infty} \frac{|u(x_n)|}{1+||x_n||}$$

While we cannot define a norm on the dual space X^* of a normed space X, we can constructively define the unit ball of X^*, i.e. the set of functionals that map the unit ball of X into the unit ball of $\mathbb{C}$. Now we make use of the fact that a constructive version of the Banach-Alaoglu theorem holds in Bishop's mathematics.

Theorem 3.3.7. *The unit ball B^* of the dual X^* of a separable normed space X is complete and totally bounded with respect to the double norm.*

Proof. See [BB85, Chapter 7, Theorem (6.7)] for a constructive proof. □

This means in particular that the dual of a separable normed space is the inductive limit of a sequence of csm's, as, if B^* is the unit ball of X^*, then $X^* = \bigcup_{n\in\mathbb{N}} nB^*$.

The question arises whether the final topology induced by the admissible representation (which coincides with the inductive limit topology induced by the sequence of csm's) is equal to the usual weak* topology on X. For infinite dimensional separable Banach spaces this is *never* the case, as shown in [HS96]. A consequence of Theorem 2.5 of *loc. cit.* is that the weak* topology on X^* is *not* sequential if X has infinite dimension. However, as the application function $X \times X^* \longrightarrow \mathbb{C}$ is sequentially continuous with respect to the weak* topology on X^* by the Banach–Steinhaus theorem (see [Rud73, Theorem 2.17]), the notions of *sequential* convergence of the weak*-topology and the inductive limit topology coincide, and hence the latter is the sequentialization of the former.

Test functions

A similar situation arises in the case of the space of Schwartz test functions $\mathcal{D}$, that is the set of infinitely differentiable functions with compact support. It is easy to see that this set is an inductive limit of csm's, constructively. To this end, for every $k \in \mathbb{N}$ one defines $C_k^\infty \subseteq \mathcal{D}$ to be the set of those test functions, whose support is contained in the interval $[-k, k]$. Each C_k^∞ is a Frchet space and hence a csm. Furthermore, $\mathcal{D}$ is the union of all spaces C_k^∞. Therefore, the function realizability interpretation yields an admissible representation of $\mathcal{D}$ and the final topology induced by this representation is the inductive limit topology on $\mathcal{D}$. This topology is not identical with the usual vector space topology on $\mathcal{D}$, as the latter has been shown to be non-sequential by T. Shirai. Nevertheless, the notions of sequential convergence of the usual topology on $\mathcal{D}$ and the inductive limit topology coincide, and hence the latter is the sequentialization of the former.

The space ℓ^p

As we have seen, if one can constructively prove that X is a Banach space or an inductive limit of csm's, then the realizability interpretation of the description of the *underlying set alone* will be interpreted as an admissible representation of that space with respect to the expected topology. How can this be, if one and the same set can carry several meaningful vector space topologies? For instance, we can meaningfully equipp the space ℓ^p with its norm, or with the weak topology. Can we define admissible representations for both spaces logically? The answer is yes. The crucial point is that in constructive mathematics, classically indistinguishable variations in the description of a set can have a big impact. We will find an admissible representation of the Banach space ℓ^p as the result of the interpretation of the set

$$\{(x_n)_{n\in\mathbb{N}} \mid \sum_{n=0}^{\infty} |x_n|^p \text{ is convergent}\}$$

whereas we find an admissible representation of the space ℓ^p equipped with the weak topology as the result of the interpretation of the set

$$\{(x_n)_{n\in\mathbb{N}} \mid \sum_{n=0}^{\infty} |x_n|^p < \infty\}.$$

By the constructive failure of the bounded completeness of the real numbers, a limit of the series $\sum_{n=0}^{\infty} |x_n|^p$ need not exist, even if an upper bound for the sequence of partial sums is known.

3.3.4 From constructive mathematics to computable analysis

We have already seen how one can obtain useful representations for a wide class of spaces by applying realizability semantics. In this section we will obtain computability results from constructive existence proofs. Let (X, ρ_X) and (Y, ρ_Y) be represented sets. The representations give rise to objects $(X, \Vdash_X)$ and $(Y, \Vdash_Y)$ of $\mathbf{Asm}_t(K_2, K_2^{\mathrm{eff}})$. Furthermore, every relation $R \subseteq X \times Y$ corresponds uniquely to a stable subobject of $(X, \Vdash_X) \times (Y, \Vdash_Y)$ which we will denote by R as well. In the above situation, the relation R is computably realizable with respect to ρ_X and ρ_Y in the sense of definition 3.1.2 if and only if the formula $\forall x : X \exists y : Y. R(x, y)$ holds with respect to the internal logic of $\mathbf{Asm}_t(K_2, K_2^{\mathrm{eff}})$.

This fact allows us to relate constructive mathematics and computable analysis. We can perform the proof of $\forall x : X\, \exists y : Y. R(x, y)$ on an abstract level and do not have to manipulate realizers concretely. For instance, recall the example of finding a complex zero of a non-constant polynomial mentioned in section 3.3.1. The computable realizability of the relation follows immediately from the existence of constructive proofs of the Fundamental Theorem of Algebra.

On the other hand, it can be shown that a relation R is *not* computably realizable by deriving a statement from the assumption $\forall x : X \exists y : Y.\, R(x, y)$ which is known to be false in $\mathbf{Asm}_t(K_2, K_2^{\mathrm{eff}})$. Useful such statements are omniscience principles like LPO or LLPO.

We will now look at some mini-examples from basic numerical mathematics. We use the following notations.

$$\begin{aligned} \mathsf{M}(n) &: \text{ real } n \times n \text{ matrices} \\ \mathsf{O}(n) &: \text{ orthogonal } n \times n \text{ matrices} \\ \Delta(n) &: \text{ upper triangular } n \times n \text{ matrices} \end{aligned}$$

One can prove in classical mathematics that every matrix A can be written as a product $A = QR$ of an orthogonal matrix Q and an upper triangular matrix R. Is it the case that we can *compute* such matrices Q and R from A?

The set $\mathsf{M}(n) = \mathbb{R}^{n \cdot n}$ is interpreted as the standard representation of $\mathbb{R}^{n \cdot n}$. Its subsets $\mathsf{O}(n)$ and $\Delta(n)$ are defined by equations and thus stable. Therefore, their interpretations are the corestrictions of the standard representation of $\mathbb{R}^{n \cdot n}$ to the respective subsets, as expected. Finally, the relation $R(A, Q, R) \leftrightarrow A = QR$ is stable.

From the assumption

$$\forall A \in \mathsf{M}(n) \exists Q \in \mathsf{O}(n) \exists R \in \Delta(n).\, A = QR$$

we can infer

$$\forall x \in \mathbb{R}^n \exists H \in \mathsf{O}(n).\, Hx \in \mathbb{R}\mathrm{e}_1$$

by looking only at the first column of each matrix. The latter statement entails the principle LLPO for the following reason. Given a sequence α such that $\alpha(n) = 1$ for at most one argument n, one can define the vector $\sum_{n=0}^{\infty} a_n$ where $a_n = 2^{-n} e_2$ if n is even and $a_n = 2^{-n} e_1$ if n is odd. Now $H^{-1} e_1$ is either off the line $\mathbb{R}e_1$ or off the line $\mathbb{R}e_2$, which gives us the information whether α is zero for all odd or all even arguments.

On the other hand

$$\forall A \in \mathsf{Gl}(n) \exists Q \in \mathsf{O}(n) \exists R \in \Delta(n).\, A = QR$$

holds because

$$\forall x \in \mathbb{R}^n \setminus \{0\} \exists H \in \mathsf{O}(n).\, Hx \in \mathbb{R}\mathrm{e}_1$$

holds and the whole decomposition can be assembled from this transformation. Furthermore, a version of the original theorem with the weaker conclusion that R is upper triangular only up to an arbitrary $\varepsilon > 0$ can be proved and hence computed by an algorithm.

A similar situation occurs for the Gauss factorization. It is not possible to compute from an arbitrary matrix A a permutation matrix P, a lower diagonal matrix L and an upper diagonal matrix U such that $PA = LU$. However, if we either require A to be invertible or else allow the matrices L and U to be lower and upper diagonal up to ε, a decomposition can be computed.

Our last example will be concerned with the computability of eigenvectors. We can prove classically that every real symmetric matrix has an eigenvector. Nevertheless, the assumption

$$\forall M \in \mathsf{Sym}(2) \exists v \in \mathbb{R}^2.\ v \in \mathrm{EV}(M)$$

entails LLPO. This can be seen as follows. Let $a, b \in \mathbb{R}$ such that $ab = 0$. From

$$\exists \lambda \in \mathbb{R} \exists v \in \mathbb{R}^2.\ ||v|| = 1, \begin{pmatrix} 1+a & b \\ b & 1 \end{pmatrix} v = \lambda v$$

it is not too hard to infer $a = 0$ or $b = 0$. The implication $ab = 0 \rightarrow (a = 0 \vee b = 0)$ is an equivalent of LLPO. Accordingly, there exists no algorithm computing an eigenvalue for an arbitrary symmetric matrix, and we were able to show this on a purely logical level.

4 (Appendix) Semantic Strong Normalization

In order to get a semantic strong normalization proof for a type theory to work as proposed in [HO93, OR94], it is necessary to build a model of the respective type theory based on a right–absorptive conditionally partial combinatory algebra (SN_*, Θ) consisting of (equivalence classes of) strongly normalizing λ–terms with a distinguished constant $*$.

We recall the definition of a conditional pca.

Definition 4.0.8. A *conditionally partial combinatory algebra (c-pca)* is a set A equipped with a partial binary application that has elements $\mathsf{s}, \mathsf{k} \in A$ such that for all $a, b, c \in \mathcal{A}$

$$\mathsf{s}\, a\, b\, c \simeq a\, c\, (b\, c), \qquad \mathsf{k}\, a\, b = b$$

Whereas we believe that one could relax the condition on s without any harm by requiring merely $\mathsf{s}\, a\, b\, c \succeq a\, c\, (b\, c)$, we will refrain from doing so for now.

Remark 4.0.9. Conditional combinatory completeness holds for every c-pca A, i.e., for any polynomial e (a term built up from typed variables and constants) and any variable x there is a polynomial $\lambda^* x.\, e$ whose variables are those of e excluding x such that $(\lambda^* x.\, e)\, a \simeq e[a/x]$ for all $a \in A$ and all valuations of the free variables. That means, $\lambda^* x.\, e$ is only guaranteed to exist if $e[a/x]$ exists for some a.

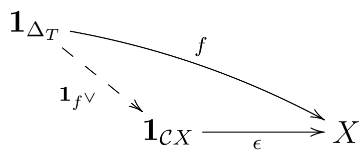
Definition 4.0.10. A subset $\Theta \subseteq \mathcal{A}$ of a c-pca is called an *ideal* if for all $a \in \mathcal{A}$ and all $\theta \in \Theta$, the result θa exists and is an element of Θ. A c-pca equipped with a distinguished ideal is called *right–absorptive*.

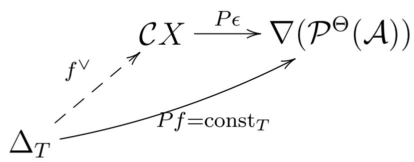
As we have seen in section 2.2, there is no generic family for the fibration of discrete families in $\mathbf{Asm}(\mathcal{P}^{\Theta}(\mathcal{A}))$ and, therefore, this fibration cannot be used for a semantic strong normalization proof for an impredicative type theory. There are, however, fibrations built upon the notion of right–absorptive c-pca that do have a generic family. In [HO93] the fibration of *PER–extension pairs* over **Set** is introduced. This fibration is a complete fibred CCC with a generic family and gives rise to a semantic strong normalization proof for system F.

In [OR94] this fibration is extended to the fibration of PER–extension pairs over $\mathbf{Asm}(\mathcal{P}^{\Theta}(\mathcal{A}))$. Roughly, the difference between the fibration of PER–extension pairs over $\mathbf{Asm}(\mathcal{P}^{\Theta}(\mathcal{A}))$ on the one hand and the fibration of uniform families of discrete sets over $\mathbf{Asm}(\mathcal{P}^\Theta(\mathcal{A}))$ on the other hand is that in the former there need not be one single set of potential realizers common to all elements of a family but the set of potential realizers may depend on the index. Therefore, the fibration of PER–extension pairs over $\mathbf{Asm}(\mathcal{P}^\Theta(\mathcal{A}))$ does have a generic family. However, as we shall see this fibration does not have *Lawvere comprehension* any more (i.e. the terminal object functor of the fibration has no right–adjoint). Accordingly, it cannot arise as the fibration of propositions induced by a model of the Calculus of Constructions in $\mathbf{Asm}(\mathcal{P}^\Theta(\mathcal{A}))$ in contradiction to what is claimed in [OR94].

Theorem 4.0.11. *The fibration of PER–extension pairs over* $\mathbf{Asm}(\mathcal{P}^\Theta(\mathcal{A}))$ *does not have Lawvere comprehension unless* $\mathcal{A} = \Theta$*.*

Proof. Assume that $\mathcal{C}$ is a right–adjoint of the terminal object functor. Let X be the family over $\nabla(\mathcal{P}^\Theta(\mathcal{A}))$ that associates with any $T \in \mathcal{P}^\Theta(\mathcal{A})$ the PER–extension pair $\langle =_T, T\rangle$. Let U be the set of potential realizers of $\mathcal{C}X$. Now, for any $T \in \mathcal{P}^\Theta(\mathcal{A})$ let $f = (f_t)_{t\in T} : \mathbf{1}_{\mathcal{C}X} \to X$ be the morphism over $\mathrm{const}_T : \Delta_T \to \nabla(\mathcal{P}^\Theta(\mathcal{A}))$ defined by $f_t(*) = \{t\}$. As $\mathcal{C}$ is right-adjoint to $\mathbf{1}$ the diagram can be completed as indicated. From the realizers of $f^\vee$ and ϵ we obtain elements $\mathsf{e}_T \in |T \to U|$ and $\mathsf{r}_T \in |U \to T|$ satisfying $\mathsf{r}_T(\mathsf{e}_T t) = t$ for all $t \in |T|$. By Proposition 2.2.17 this is a contradiction.





□

Instead of the modified realizability topos over a right–absorptive c-pca $(\mathcal{A}, \Theta)$ we propose to use a different topos. It can be defined by a tripos that has $\Sigma = \{(A, P) \mid A \subseteq P \subseteq \mathcal{A},\ A = \emptyset \Rightarrow P = \emptyset,\ A \neq \emptyset \Rightarrow \Theta \subseteq P\}$ as truth values. Entailment is defined as for the tripos representing the topos $\mathbf{Eff}^\to$. The category of $\neg\neg$–separated objects of this topos does indeed provide a model of the Calculus of Constructions. Moreover, it avoids the major complication caused by the fact that in the modified realizability topos, the global sections functor is *not* left-adjoint to the inclusion of $\neg\neg$-sheaves.

Bibliography

[Bau00] A. Bauer. *The Realizability Approach to Computable Analysis and Topology.* PhD thesis, School of Computer Science, Carnegie Mellon University, Pittsburgh, 2000.

[BB85] E. Bishop and D.S. Bridges. *Constructive Analysis.* Springer-Verlag, 1985.

[BBS98] A. Bauer, L. Birkedal, and D.S. Scott. Equilogical spaces. Revised February 2001. To appear in Theoretical Computer Science, September 1998.

[Bee75] M.J. Beeson. The nonderivability in intuitionistic formal systems of theorems on the continuity of effective operations. *Journal of Symbolic Logic*, 40(3), 1975.

[Bee85] M.J. Beeson. *Foundations of constructive mathematics.* Springer Verlag, Berlin, 1985.

[Bir99] L. Birkedal. *Developing Theories of Types and Computability.* PhD thesis, School of Computer Science, Carnegie Mellon University, 1999.

[Bis67] E. Bishop. *Foundations of Constructive Analysis.* McGraw-Hill, New York, 1967.

[Bor94] F. Borceux. *Handbook of Categorical Algebra 2, Categories and Structures*, volume 51 of *Encyclopedia of Mathematics and its Applications.* Cambridge University Press, 1994.

[Bou66] N. Bourbaki. *Espaces vectorielles topologiques*, volume XV. Hermann, 1966.

[BR87] D.S. Bridges and F. Richman. *Varieties of Constructive Mathematics*, volume 97 of *London Mathematical Society Lecture Note Series.* Cambridge University Press, Cambridge, 1987.

[BS03] A. Bauer and A. Simpson. Two constructive embedding-extension theorems with applications to continuity principles and to banach-mazur computability. submitted manuscript, 2003.

[BSV02] D.S. Bridges, P. Schuster, and L. Vîţă. Apartness, topology, and uniformity: a constructive view. *Mathematical Logic Quarterly*, 48(Suppl. 1):16–28, 2002.

[Car95] A. Carboni. Some free constructions in realizability and proof theory. *Journal of Pure and Applied Algebra*, 103:117–148, 1995.

[CR98] A. Carboni and G. Rosolini. Locally cartesian closed exact completions. *Journal of Pure and Applied Algebra*, 1998. to appear.

[CV98] A. Carboni and E.M. Vitale. Regular and exact completions. *Journal of Pure and Applied Algebra*, 125:79–116, 1998.

[dMRR99] F. de Marchi, E.P. Robinson, and G. Rosolini. An abstract look at realizability. Submitted for publication, November 1999.

[Dum77] M. Dummett. *Elements of Intuitionism*. Oxford University Press, Oxford, 1977.

[Ers75] Y.L. Ershov. Theorie der numerierungen ii. *Zeitschrift für Mathematische Logik und Grundlagen der Mathematik*, 21:473–584, 1975.

[Fri77] H. Friedman. On the derivability of instantiation properties. *Journal of Symbolic Logic*, 42:506–514, 1977.

[FS91] P.J. Freyd and A. Scedrov. *Categories, Allegories*. North-Holland, 1991.

[Gra81] R. J. Grayson. Modified realizability toposes. Handwritten notes from Münster University, July 1981.

[Hey30] A. Heyting. Die formalen regeln der intuitionistischen logik. In *Sitzungsberichte der Preußischen Akademie der Wissenschaften. Physikalisch-mathematische Klasse*, pages 42–56, 1930.

[Hey56] A. Heyting. *Intuitionism, An Introduction*. North-Holland, 1956.

[HJP80] J.M.E. Hyland, P.T. Johnstone, and A.M. Pitts. Tripos theory. *Math. Proc. Camb. Phil. Soc.*, 88:205–232, 1980.

[HK66] W.A. Howard and G. Kreisel. Transfinite induction and bar induction of types zero and one, and the role of continuity in intuitionistic analysis. *Journal of Symbolic Logic*, 31(3), 1966.

[HO93] J.M.E. Hyland and C.-H. L. Ong. Modified realizability toposes and strong normalization proofs. In J.F. Groote and M. Bezem, editors, *Typed Lambda Calculi and Applications*, volume 664 of *Lecture Notes in Computer Science*, pages 179–194. Springer-Verlag, 1993.

[HRR90] J.M.E. Hyland, E.P. Robinson, and G. Rosolini. The discrete objects in the effective topos. *Proceedings of the London Mathematical Society*, 60, 1990.

[HS96] A.J. Humphreys and S.G. Simpson. Separable banach space theory needs strong set existence axioms. *Transactions of the American Mathematical Society*, 348:4231–4255, 1996.

[Hyl82] J.M.E. Hyland. The effective topos. In A.S. Troelstra and D. van Dalen, editors, *The L.E.J. Brouwer Centenary Symposium*, pages 165–216. North-Holland, 1982.

[Hyl91] J.M.E. Hyland. First steps in synthetic domain theory. In A. Carboni, C. Pedicchio, and G. Rosolini, editors, *Conference on Category Theory 1990*, volume 1488 of *Lecture Notes in Mathematics*, pages 131–156. Springer Verlag, 1991.

[Ish92] H. Ishihara. Continuity properties in constructive mathematics. *Journal of Symbolic Logic*, 57(2):557–565, 1992.

[Ish01] H. Ishihara. Sequentially continuity in constructive mathematics. In C.S. Calude, M.J. Dinneen, and S. Sburlan, editors, *Combinatorics, Computability and Logic*, Discrete Mathematics and Theoretical Computer Science, pages 5–12, London, 2001. Springer. Proceedings of DMTCS 2001.

[Jac99] B. Jacobs. *Categorical logic and type theory.* Studies in logic and foundations of mathematics. North-Holland, Amsterdam, 1999.

[Kle45] S.C. Kleene. On the interpretation of intuitionistic number theory. *Journal of Symbolic Logic*, 10:109–124, 1945.

[Kle65] S.C. Kleene. Logical calculus and realizability. *Acta Philosophica Fennica*, 18:71–80, 1965.

[Kle69] S.C. Kleene. *Formalized Recursive Functionals and Formalized Relizability*, volume 89 of *Memoirs of the American Mathematical Society.* American Mathematical Society, 1969.

[Kre59] G. Kreisel. Interpretation of analysis by means of functionals of finite type. In *Constructivity in Mathematics*. North Holland, Amsterdam, 1959.

[LM84] G. Longo and E. Moggi. The hereditary partial effective functionals and recursion theory in higher types. *Journal of Symbolic Logic*, 49:1319–1332, 1984.

[Lon94] J. Longley. *Realizability Toposes and Language Semantics*. PhD thesis, University of Edinburgh, 1994.

[Lon99a] J. Longley. Matching typed and untyped realizability. In L. Birkedal, J. van Oosten, G. Rosolini, and D.S. Scott, editors, *Workshop on Realizability Semantics and Applications*, volume 23 of *Electronic Notes in Computer Science*. Elsevier, 1999. Available from http://www.elsevier.nl/gej-ng/31/29/23/92/27/show/Products/notes/index.htt.

[Lon99b] J. Longley. Unifying typed and untyped realizability. Manuscript (available from `http://www.dcs.ed.ac.uk/home/jrl/unifying.txt`), 1999.

[LS02] P. Lietz and T. Streicher. Impredicativity entails untypedness. *Mathematical Structures in Computer Science*, 12(3):335–347, 2002.

[Mac71] S. Mac Lane. *Categories for the Working Mathematician*. Springer Verlag, Berlin, 1971.

[Men99] M. Menni. A characterization of the left exact categories whose exact completions are toposes. Submitted for publication, 1999.

[Men00] M. Menni. Generic proofs and chaotic situations (extended abstract). Submitted for publication, 2000.

[Oos91] J. van Oosten. *Exercises in realizability*. PhD thesis, University of Amsterdam, 1991.

[Oos97a] J. van Oosten. Extensional realizability. *Annals of Pure and Applied Logic*, 84:317–349, 1997.

[Oos97b] J. van Oosten. The modified realizability topos. *Journal of Pure and Applied Algebra*, 116:273–289, 1997.

[Oos02] J. van Oosten. Realizability: A historical essay. *Mathematical Structures in Computer Science*, 12(3):239–263, 2002.

[OR94] C.-H. L. Ong and E. Ritter. A generic strong normalization proof: application to the calculus of constructions (extended abstract). In *Computer Science Logic: 7th Workshop, CSL'93, Swansea, UK, Sep. 1993, Selected Papers*, volume 832 of *Lecture Notes in Computer Science*, pages 261–279. Springer Verlag, 1994.

[OS00] J. van Oosten and A. Simpson. Axioms and (counter)examples in synthetic domain theory. *Annals of Pure and Applied Logic*, 104(1–3):233–278, 2000.

[Pit81] A.M. Pitts. *The Theory of Triposes*. PhD thesis, Cambridge University, 1981.

[Pit99] A.M. Pitts. Tripos theory in retrospect. In L. Birkedal, J. van Oosten, G. Rosolini, and D.S. Scott, editors, *Workshop on Realizability Semantics and Applications*, volume 23 of *Electronic Notes in Computer Science*. Elsevier, 1999.

[Pit00] A.M. Pitts. Categorical logic. In S. Abramsky, D. M. Gabbay, and T. S. E. Maibaum, editors, *Handbook of Logic in Computer Science, Volume 5. Algebraic and Logical Structures*, chapter 2. Oxford University Press, 2000.

[Plo78] G.D. Plotkin. $\mathbb{T}^\omega$ as a universal domain. *J. Computer and System Sciences*, 17:209–236, 1978.

[Rog67] H. Rogers, Jr. *Theory of Recursive Functions and Effective Computability*. McGraw-Hill Book Company, 1967.

[Ros99] G. Rosolini. Equilogical spaces and filter spaces, 1999. to appear.

[Rud73] W. Rudin. *Functional Analysis*. McGraw Hill Series in Higher Mathematics. McGraw Hill, New York, 1973.

[Sch02] M. Schröder. Extended admissibility. *Theoretical Computer Science*, 284(2):519–538, 2002.

[Sch03] M. Schröder. *Admissible Representations for Continuous Computations*. PhD thesis, FernUniversität – Gesamthochschule Hagen, 2003.

[Spe62] C. Spector. Provably recursive functionals of analysis: a consistency proof of analysis by an extension of principles formulated in current intuitionistic mathematics. In J.C.E. Dekker, editor, *Proceedings of Symposia in Pure Mathematics*, volume 5, pages 1–27. American Mathematical Society, 1962.

[Spi02] B. Spitters. Located operators. *Mathematical Logic Quarterly*, 48(Suppl. 1):107–122, 2002.

[Tro73] A.S. Troelstra, editor. *Metamathematical Investigation of Intuitionistic Arithmetic and Analysis*. Springer Verlag, 1973. With contributions by A.S. Troelstra, C.A. Smoryński, J.I. Zucker and W.A. Howard.

[Tro92] A.S. Troelstra. Comparing the theory of representations and constructive mathematics. In E. Börger, G. Jäger, H. Kleine Büning, and M.M. Richter, editors, *Computer Science Logic. 5th workshop, CSL '91*, pages 382–395. Springer Verlag, 1992.

[Tro98] A.S. Troelstra. Realizability. In S.R. Buss, editor, *Handbook of Proof Theory*, pages 407–473. North-Holland, 1998.

[Tur36] A. Turing. On computable numbers with an application to the Entscheidungsproblem. *Proc. London Math. Soc.*, 42:230–265, 1936.

[TvD88a] A.S. Troelstra and D. van Dalen. *Constructivism in mathematics. An introduction. Volume I.* Number 121 in Studies in Logic and the Foundations of Mathematics. North-Holland, Amsterdam, 1988.

[TvD88b] A.S. Troelstra and D. van Dalen. *Constructivism in mathematics. An introduction. Volume II.* Number 123 in Studies in Logic and the Foundations of Mathematics. North-Holland, Amsterdam, 1988.

[Wei00] K. Weihrauch. *Computable Analysis*. Springer, Berlin, 2000.

Curriculum Vitae

Peter Lietz

	geboren am 1. November 1970 in Mainz
1990	Allgemeine Hochschulreife, Staatliches Gymnasium am Kurfürstlichen Schloß, Mainz
1991 – 1998	Studium der Mathematik mit Wahlpflichtfach Informatik, Technische Universität Darmstadt
1994 – 1995	Erasmus-Aufenthalt am Instituto Superior Técnico der Universidade Técnica de Lisboa, Portugal
1998	Hochschulabschluss als Diplom-Mathematiker, Diplomarbeit "A Fibrational Theory of Geometric Morphisms"
1998 – 2004	Doktorand am Fachbereich Mathematik der TU Darmstadt, Arbeitsgruppe Logik und mathematische Grundlagen der Informatik
1998 – 1999	Teilnahme am Postgraduate Course in the Theory of Computation des Laboratory for Foundations of Computer Science an der University of Edinburgh
	zeitweise gefördert durch ein Stipendium des Landes Hessen zeitweise wissenschaftlicher Mitarbeiter