Chapter 3
Intuitionistic and classical arithmetic in all finite types

3.1 Intuitionistic and classical predicate logic

In the following we formulate an axiomatic system for intuitionistic first order predicate logic IL. The particular Hilbert-type axiomatization we choose is due to [133] and specially suited to carry out proof interpretations inductively over the proof tree. Of course, like any Hilbert-type system also this axiomatization is not very convenient for actually carrying out proofs for which a natural deduction style calculus is to be recommended. For a proof of the equivalence of these calculi see [366] (1.1.5–1.1.11).

Intuitionistic first order predicate logic without equality $\mathbf{IL}_{-=}$

I. **The language $\mathscr{L}(\mathbf{IL}_{-=})$ of $\mathbf{IL}_{-=}$:**
As logical constants we use $\land, \lor, \to, \bot$ (absurdity or 'falsum'), $\exists, \forall$. $\mathscr{L}(\mathrm{IL}_{-=})$ contains variables $x, y, z, \ldots$ (which can be free or bound). Furthermore, for any arity $n \geq 0$ we have (possibly empty) denumerable sets of function symbols $f_1, f_2, f_3, \ldots$ and (for $n \geq 1$) predicate symbols $P_1, P_2, P_3, \ldots$.
0-place function symbols are called constants and usually denoted by $c_1, c_2, c_3, \ldots$.

Terms:

(i) Variables and constants are terms.
(ii) If $t_1, \ldots, t_n$ are terms and f is an n-ary function symbol, then $f(t_1, \ldots, t_n)$ is a term.

Terms that do not contain any variables are called closed.

Formulas:

(i) If $t_1, \ldots, t_n$ are terms and P an n-ary predicate symbol, then $P(t_1, \ldots, t_n)$ is a (prime) formula. Moreover, $\bot$ is a (prime) formula.
(ii) If A, B are formulas, then $(A \land B)$, $(A \lor B)$ and $(A \to B)$ are formulas.

(iii) If A is a formula and x a variable, then $(\forall x A)$ and $(\exists x A)$ are formulas.

As usual, formulas which do not contain free variables (i.e. variables occurring not bound by any quantifier) are called closed or sentences.

Abbreviations:
$\neg A :\equiv A \to \bot, A \leftrightarrow B :\equiv (A \to B) \wedge (B \to A)$.

Conventions on parentheses: Negation and quantifiers bind stronger than $\vee, \wedge$ which bind stronger than $\to, \leftrightarrow$. Using this convention we can safely drop many parentheses around formulas, e.g. we simply write (dropping also outermost parentheses) $A \wedge B \to \neg C \vee D$ instead of $((A \wedge B) \to ((\neg C) \vee D))$.

II. **Axioms of $\mathbf{IL}_{-=}$:**

(i) $A \vee A \to A, A \to A \wedge A$ (axioms of contraction);
(ii) $A \to A \vee B, A \wedge B \to A$ (axioms of weakening);
(iii) $A \vee B \to B \vee A, A \wedge B \to B \wedge A$ (axioms of permutation);
(iv) $\bot \to A$ (ex falso quodlibet);
(v) $\forall x A \to A[t/x], A[t/x] \to \exists x A$, where t is free for x in A and $A[t/x]$ is the result of replacing every free occurrence of x in A by t (quantifier axioms).

III. **Rules of $\mathbf{IL}_{-=}$:**

(i)
$$\frac{A, A \to B}{B}, \quad \frac{A \to B, B \to C}{A \to C}$$
(modus ponens and syllogism);

(ii)
$$\frac{A \wedge B \to C}{A \to (B \to C)}, \quad \frac{A \to (B \to C)}{A \wedge B \to C}$$
(exportation and importation);

(iii)
$$\frac{A \to B}{C \vee A \to C \vee B} \text{ (expansion);}$$

(iv)
$$\frac{B \to A}{B \to \forall x A}, \quad \frac{A \to B}{\exists x A \to B}, \text{ where } x \text{ is not free in } B$$
(quantifier rules).

Remark 3.1. Most of the time we will use for notational simplicity the slightly imprecise notation '$A(t)$' instead of '$A[t/x]$'.

Classical first order predicate logic without equality $\mathbf{PL}_{-=}$ results from $\mathrm{IL}_{-=}$ by adding the law-of-excluded-middle (LEM) schema

$$A \vee \neg A.$$

The Brouwer-Heyting-Kolmogorov '(BHK)' proof interpretation of the intuitionistic logical constants (our exposition makes use of [318]).

This interpretation is an informal attempt to explain the meaning of the logical constants of $\mathrm{IL}_{-=}$ in terms of proof constructions. Here 'proof' is understood as 'verification by a construction' and not as a formal proof in some fixed deductive framework like HA below.

(i) There is no proof for $\perp$.
(ii) A proof of $A \wedge B$ is a pair (q,r) of proofs, where q is a proof of A and r is a proof of B.
(iii) A proof of $A \vee B$ is a pair (n,q) consisting of an integer n and a proof q which proves A if $n = 0$ and resp. B if $n \neq 0$.
(iv) A proof p of $A \to B$ is a construction which transforms any hypothetical proof q of A into a proof $p(q)$ of B.
(v) A proof p of $\forall x A(x)$ is a construction which produces for every construction c_d of an element d of the domain a proof $p(c_d)$ of $A(d)$.
(vi) A proof p of $\exists x A(x)$ is a pair (c_d, q), where c_d is the construction of an element d of the domain and q is a proof of $A(d)$.

Discussion: There is one problem with the BHK-interpretation: from a strictly constructive point of view one would like to have a constructive verification of 'p is a proof of A' in case this is true, i.e. one would like to recognize a proof if one sees it. For $(i),(ii),(iii),(vi)$ there is no problem with this requirement. But for the universal statements in $(iv),(v)$ one would need an additional clause as suggested by Kreisel in [247]:

(iv)' A proof p of $A \to B$ is a pair (r,q), where q is a construction which transforms any hypothetical proof s of A into a proof $q(s)$ of B and r is a proof which verifies that q is such a construction.
(v)' A proof p of $\forall x A(x)$ is a pair (r,q) where q is a construction which produces for every construction c_d of an element d of the domain a proof $q(c_d)$ of $A(d)$ and r is a proof of the fact that q is such a construction.

Remark 3.2. There are various ways to formalize the idea behind the BHK-interpretation which give rise to various forms of so-called realizability interpretations. The first version of realizability, the so-called recursive realizability, was introduced by Kleene in [193]. In this book we will focus on a typed variant of Kleene's type-free interpretation which is called 'modified realizability' and is due to Kreisel [244, 246].
Recently S. Artemov has developed a so-called 'logic of proofs' where 'proof' in the BHK-clauses is interpreted as 't is a proof (polynomial) for A' referring to some standard proof (not: provability) predicate e.g. for PA. Using this interpretation he proves a completeness result for intuitionistic propositional logic (see [4]).

Intuitionistic first order predicate logic with equality IL

IL results from $\mathrm{IL}_{-=}$ by adding a special binary predicate symbol $=$ to the language together with the

Equality axioms:

(i) $x = x$, $x = y \to y = x$, $x = y \wedge y = z \to x = z$.
(ii) $x_1 = y_1 \wedge \ldots \wedge x_n = y_n \to f(x_1, \ldots, x_n) = f(y_1, \ldots, y_n)$ for any n-ary function symbol f.
(iii) $x_1 = y_1 \wedge \ldots \wedge x_n = y_n \to \big(P(x_1, \ldots, x_n) \to P(y_1, \ldots, y_n)\big)$ for any n-ary predicate symbol P.

Classical first order predicate logic with equality PL results from IL by adding the law-of-excluded-middle schema

$$\mathrm{LEM}: \ A \vee \neg A.$$

3.2 Intuitionistic ('Heyting') arithmetic HA and Peano arithmetic PA

$\mathscr{L}$(HA) contains the logical constants of $\mathscr{L}$(IL), number variables $x, y, z, \ldots$, a constant 0 (zero), a unary function symbol S (successor), function symbols for all primitive recursive functions (more precisely for all derivations of primitive recursive functions).

Axioms and rules of HA:

(i) axioms and rules of IL (based on $\mathscr{L}$(HA)),
(ii) successor axioms:

$$\begin{cases} S(x) \neq 0, \\ S(x) = S(y) \to x = y, \end{cases}$$

(iii) defining equations for the primitive recursive functions,
(iv) axiom schema of complete induction

$$\mathrm{IA}: \ A(0) \wedge \forall x \big(A(x) \to A(S(x))\big) \to \forall x A(x)$$

for every formula $A \in \mathscr{L}$(HA).

Convention: We often write x' or $x+1$ for $S(x)$.

Remark 3.3. 1) In HA one can prove that $\bot \leftrightarrow 0 = 1$ and so we may identify $\bot$ with $0 = 1$, where $1 := S(0)$. Then the axioms involving falsum, namely $\bot \to A$ and $\neg S(x) = 0$ even become redundant: for $0 = 1 \to A$ it suffices to establish this for all prime formulas $s = t$ which follows by primitive recursion. $S(x) = 0 \to 0 = 1$ is proved similarly (exercise).

2) Instead of the axiom schema IA we could have formulated HA equivalently using the rule of induction

$$\text{IR:} \ \frac{A(0)\ ,\ A(x) \to A(S(x))}{A(x)}.$$

Exercise!

Lemma 3.4.

$$\text{HA} \vdash \forall x(x = 0 \vee x \neq 0).$$

Proof: Induction on x : If $x = 0$, then $x = 0 \vee x \neq 0$ by weakening.
For $S(x)$ we have $S(x) \neq 0$ and hence $S(x) = 0 \vee S(x) \neq 0$ again by weakening. □

Proposition 3.5. *The following rule of double induction is derivable in* HA*:*

$$\frac{A(x,0)\ ,\ A(0,y)\ ,\ A(x,y) \to A(S(x),S(y))}{A(x,y)}.$$

Proof: We leave the tedious proof as exercise resp. refer to the literature: [371]. □

In the following, let $+, \cdot, \overline{sg}, pd, \dot{-}, |\cdot - \cdot|$ be defined primitive recursively as follows (to bring these informal primitive recursions into the official format of primitive recursion one has to make use of projections to introduce dummy arguments, exercise):

$$x + 0 = x, x + S(y) = S(x + y);$$

$$x \cdot 0 = 0, x \cdot S(y) = x \cdot y + x;$$

$$\overline{sg}(0) = 1, \overline{sg}(S(x)) = 0;$$

$$pd(0) = 0, pd(S(x)) = x;$$

$$x \dot{-} 0 = x, x \dot{-} (S(y)) = pd(x \dot{-} y);$$

$$|x - y| = (x \dot{-} y) + (y \dot{-} x).$$

Remark 3.6. In the presence of the defining axioms for the primitive recursive functions including the predecessor function pd the successor axiom $S(x) = S(y) \to x = y$ actually becomes redundant since

$$S(x) = S(y) \to x = pd(S(x)) = pd(S(y)) = y.$$

Lemma 3.7. HA *proves the following basic facts:*

1) $x + y = 0 \leftrightarrow x = 0 \wedge y = 0.$
2) $x \cdot y = 0 \leftrightarrow x = 0 \vee y = 0.$
3) $\overline{sg}(x) = 0 \leftrightarrow x \neq 0.$
4) $\overline{sg}(x) \cdot y = 0 \leftrightarrow (x = 0 \to y = 0).$
5) $|x - y| = 0 \leftrightarrow x = y.$

Proof: Exercise (use lemma 3.4 and double induction, i.e. proposition 3.5). □

Proposition 3.8. *Let $A_0(\underline{x})$ be a quantifier-free formula of $\mathscr{L}$(HA) whose free variables are among $\underline{x} = x_1, \ldots, x_n$. Then there is an n-ary primitive recursive function symbol f of HA such that*

$$\mathrm{HA} \vdash \forall \underline{x}(f(\underline{x}) = 0 \leftrightarrow A_0(\underline{x})).$$

Proof: Immediate from lemma 3.7 since all prime formulas of $\mathscr{L}$(HA) are of the form $t = s$. □

Corollary 3.9. *Let A_0 be a quantifier-free formula of $\mathscr{L}$(HA). Then*

$$\mathrm{HA} \vdash A_0 \vee \neg A_0.$$

In particular, quantifier-free formulas are provably stable, i.e.

$$\mathrm{HA} \vdash \neg\neg A_0 \to A_0.$$

Proof: Lemma 3.4 and proposition 3.8. □

Classical ('Peano') arithmetic PA results from HA by adding the law-of-excluded-middle schema

$$\mathrm{LEM}: \ A \vee \neg A.$$

3.3 Extensional intuitionistic ('Heyting') and classical ('Peano') arithmetic in all finite types

In chapter 5 we will show, in particular, that a certain subclass C of the class of all total computable functions suffices to provide witnesses for all HA-provable sentences of the form $\forall x \exists y A(x,y)$. However, to describe this class we have to go beyond the primitive recursively defined functions contained in HA as HA proves $\forall x \exists y A(x,y)$-sentences such that for no primitive recursive function f, $\forall x A(x, f(x))$ is true over $\mathbb{N}$. E.g. HA can prove the totality of the so-called Ackermann function (defined in the exercises to this chapter) which is not primitive recursive (for the latter see e.g. [341]). In order to describe algorithmically a sufficiently rich (and actually optimal, though we will not prove this) such class of functions we need to consider so-called functionals of higher type defined by a generalized form of primitive recursion. Even in fragments we consider where the recursion to define functionals of higher types is restricted to yield only ordinarily primitive recursive functions $f : \mathbb{N} \to \mathbb{N}$ (or even functions of much lower complexity) to enrich the language with variables and quantifiers for functionals of all finite types is important to carry out the proof interpretations we are mainly interested in. To some extent higher types (though usually quite low) are also needed to formalize proofs in analysis.

The set **T** of all finite types (over $\mathbb{N}$) is generated inductively by the clauses

$$(i)\ 0 \in \mathbf{T},\ (ii)\ \rho, \tau \in \mathbf{T} \Rightarrow \tau(\rho) \in \mathbf{T}.$$

The type 0 is the type natural numbers. Objects of type $\tau(\rho)$ are functions which map objects of type ρ to objects of type τ.

Remark 3.10. Some authors write $(\rho)\tau$, $(\tau\rho)$ or $(\rho \to \tau)$ instead of $\tau(\rho)$. The notation $(\rho \to \tau)$ has the benefit of indicating directly the formation of a function space. Because of this we will also use it occasionally. Moreover, it visualizes the so-called Curry-Howard correspondence (or isomorphism, see [349] for a comprehensive treatment) between formulas (of the implicative fragment of intuitionistic propositional logic) and types as well as between proofs and terms (see below). The drawback is that complicated types get much longer to write than in our notation.

We often omit brackets which are uniquely determined and write e.g. $0(00)$ instead of $0(0(0))$.
One easily notices that every type $\rho \neq 0$ can uniquely be written as $\rho = 0(\rho_k)\ldots(\rho_1)$ for suitable k and types $\rho_1,\ldots,\rho_k$. We usually use $\rho = 0\rho_k\ldots\rho_1$ as shorthand for this if it is clear to which types $\rho_1,\ldots,\rho_k$ we refer so that there is no danger of confusion.

The set $\mathbf{P} \subset \mathbf{T}$ of pure types is defined by

$$(i)\ 0 \in \mathbf{P},\ (ii)\ \rho \in \mathbf{P} \Rightarrow 0(\rho) \in \mathbf{P}.$$

Pure types are often denoted by natural numbers:

$$0(n) := n+1 \text{ (e.g. } 00 = 1, 0(00) = 2).$$

The type level or degree $deg(\rho)$ of a type ρ is defined as

$$deg(0) := 0,\ deg(\tau(\rho)) := \max(deg(\tau), deg(\rho)+1)$$

(note that for pure types ρ, $deg(\rho)$ is just the number which denotes ρ).

Objects of type ρ with $deg(\rho) > 1$ are usually called functionals.
We sometimes write ‘$\tau \leq n$’ instead of ‘$deg(\tau) \leq n$’.

The language $\mathscr{L}$(E-HA$^\omega$) of E-HA$^\omega$ is based on a many-sorted version $\mathrm{IL}^{\omega}_{=}$ of $\mathrm{IL}_{=}$ which contains variables $x^\rho, y^\rho, z^\rho, \ldots$ and quantifiers $\forall x^\rho$, $\exists y^\rho$ for every type ρ. As constants E-HA$^\omega$ contains 0^0 (zero), S^{00} (successor), $\Pi^{\rho\tau\rho}_{\rho,\tau}$ (projector), $\Sigma_{\delta,\rho,\tau}$ (combinator of type $\tau\delta(\rho\delta)(\tau\rho\delta)$) and (simultaneous) recursor constants $\underline{R}_{\underline{\rho}} = (R_1)_{\underline{\rho}}, \ldots, (R_k)_{\underline{\rho}}$, where R_i has type $\rho_i(\rho_k 0\underline{\rho}^t)\ldots(\rho_1 0\underline{\rho}^t)\underline{\rho}^t 0$ for all $\delta, \rho, \tau, \underline{\rho}$ $(= (\rho_1)\ldots(\rho_k))$ in $\mathbf{T}$. Here we use the notation $\underline{\rho}^t := (\rho_k)\ldots(\rho_1)$.
Furthermore $\mathscr{L}$(E-HA$^\omega$) contains a binary predicate symbol $=_0$ for equality between objects of type 0.
Sometimes we write $t \in \rho$ to express that t is of type ρ.

Terms of E-HA$^\omega$ are built up by

(i) constants c^ρ and variables x^ρ of type ρ are terms of type ρ
(ii) if $t^{\tau\rho}$ is a term of type $\tau\rho$ and s^ρ is a term of type ρ, then $t(s)$ is a term of type τ.

Again, we often simply write $ts_1 \ldots s_k$ – or $t(s_1, \ldots, s_k)$ – instead of $t(s_1)\ldots(s_k)$. Some authors write (ts) instead of $t(s)$.

Formulas of E-HA$^\omega$ are built up by

(i) prime formulas (also called 'atomic formulas') $s =_0 t$ are formulas (where s^0, t^0 are terms of type 0);
(ii) if A, B are formulas, then also $(A \wedge B)$, $(A \vee B)$ and $(A \to B)$ are formulas;
(iii) if A is a formula and x^ρ a variable of type ρ, then also $(\forall x^\rho A)$ and $(\exists x^\rho A)$ are formulas.

We adopt the same conventions on parentheses as before.

Abbreviations:

1) Higher type equations $s =_\rho t$ between terms s,t of type $\rho = 0\rho_k \ldots \rho_1$ (where $k \geq 1$) are abbreviations for

$$\forall y_1^{\rho_1}, \ldots, y_k^{\rho_k} \big(sy_1 \ldots y_k =_0 ty_1 \ldots y_k\big),$$

where $y_1, \ldots, y_k$ are variables which don't occur in s,t.
2) As before: $\neg A :\equiv A \to \bot$, where $\bot :\equiv (0 =_0 1)$; $A \leftrightarrow B :\equiv (A \to B) \wedge (B \to A)$.

Axioms and rules of E-HA$^\omega$

(i) all axioms and rules of $\mathrm{IL}^{\omega}_{=}$;
(ii) equality axioms for $=_0$:
$x =_0 x$, $x =_0 y \to y =_0 x$, $x =_0 y \wedge y =_0 z \to x =_0 z$;
(iii) higher type extensionality:

$$E_\rho : \forall z^\rho, x_1^{\rho_1}, y_1^{\rho_1}, \ldots, x_k^{\rho_k}, y_k^{\rho_k} \big(\bigwedge_{i=1}^{k} (x_i =_{\rho_i} y_i) \to z\underline{x} =_0 z\underline{y}\big),$$

where $\rho = 0\rho_k \ldots \rho_1$;
(iv) successor axioms;
(v) induction schema

$$\text{IA:}\ A(0) \wedge \forall x^0 \big(A(x) \to A(Sx)\big) \to \forall x^0 A(x),$$

where $A(x^0)$ is an arbitrary formula of E-HA$^\omega$;
(vi) axioms for $\Pi_{\rho,\tau}, \Sigma_{\delta,\rho,\tau}$ and $\underline{R}_{\underline{\rho}}$:

$$(\Pi) : \ \Pi_{\rho,\tau} x^\rho y^\tau =_\rho x^\rho,$$

$$(\Sigma) : \ \Sigma_{\delta,\rho,\tau} xyz =_\tau xz(yz) \quad (x^{\tau\rho\delta}, y^{\rho\delta}, z^\delta),$$

$$(\underline{R}) : \begin{cases} (R_i)_{\underline{\rho}} 0 \underline{y}\,\underline{z} =_{\rho_i} y_i \\ (R_i)_{\underline{\rho}} (Sx^0) \underline{y}\,\underline{z} =_{\rho_i} z_i(\underline{R}_{\underline{\rho}} x \underline{y}\,\underline{z})x \ \text{ for } i = 1, \ldots, k, \end{cases}$$

where $\underline{\rho} = \rho_1, \ldots, \rho_k$, $\underline{y} = y_1, \ldots, y_k$, $\underline{z} = z_1, \ldots, z_k$ with y_i of type ρ_i and z_i of type $\rho_i 0 \underline{\rho}^t$.

Remark 3.11. 1) As a many-sorted system, the fact that the sorts $\tau\rho$ and ρ, τ are connected via the term formation rule stating that with $t^{\tau\rho}, s^{\rho}$ also $t(s)$ is a term (of type τ) strictly speaking needs to be expressed via application symbols $Ap_{\rho,\tau}$, where then $Ap_{\rho,\tau}(t^{\tau\rho}, s^{\rho})$ stands for $t(s)$. We suppress this cumbersome notation and simply write $t(s)$. However, we have to keep in mind that in order to specify a model for E-HA$^{\omega}$ (see section 3.6 below) we also have to give interpretations to $Ap_{\rho,\tau}$ (usually this will be the obvious set-theoretic application).

2) The reflexivity, symmetry and transitivity of the defined higher type equalities $=_{\rho}$ are derivable from the corresponding axioms for $=_0$. Using the extensionality axioms one can prove

$$x =_{\rho} y \wedge A(x) \to A(y)$$

by induction on the complexity of A (for the case of prime formulas one first proves by induction on the terms that $x =_{\rho} y \to r[x/z^{\rho}] =_{\tau} r[y/z^{\rho}]$) (exercise).

3) Instead of the axioms E_{ρ} for all types ρ we could have used equivalently $E_{\rho,\tau}$ for all $\rho, \tau \in \mathbf{T}$, where

$$E_{\rho,\tau}: \ \forall z^{\tau\rho}, x^{\rho}, y^{\rho} (x =_{\rho} y \to zx =_{\tau} zy)$$

(exercise).

Definition 3.12. Later on we will need also a 'weakly extensional' variant WE-HA$^{\omega}$ of E-HA$^{\omega}$, where the extensionality axioms E_{ρ} are weakened to a quantifier-free rule of extensionality

$$\text{QF-ER:} \ \frac{A_0 \to s =_{\rho} t}{A_0 \to r[s/x^{\rho}] =_{\tau} r[t/x^{\rho}]},$$

where A_0 is quantifier-free and $s^{\rho}, t^{\rho}, r^{\tau}$ are terms of WE-HA$^{\omega}$ ($\rho, \tau \in \mathbf{T}$ arbitrary).

Remark 3.13. 1) The special case of QF-ER with $\tau = 0$ already implies the general case.

2) Note that QF-ER allows one (by taking $A_0 :\equiv x =_0 y$) to derive full extensionality for equality of type 0, i.e.

$$\forall x^0, y^0 \big(x =_0 y \to r[x/z^0] =_{\tau} r[y/z^0]\big).$$

Also, QF-ER suffices to prove the following rule

$$\frac{A_0 \to s =_{\rho} t}{A_0 \to (B[s/x^{\rho}] \to B[t/x^{\rho}])},$$

where B is an arbitrary formula and s, t are free for x^{ρ} in B.

Warning: We will prove later (chapter 9) that WE-HA$^{\omega}$ does not satisfy the deduction theorem.

Remark 3.14. 1) One can show in WE-HA$^\omega$ that the simultaneous primitive recursors $\underline{R}_{\underline{\rho}}$ can in fact be reduced to the single recursors R_ρ (i.e. the case $k=1$) using appropriate embeddings of tuples of types in a suitable common higher type and tuple codings of functionals. For details see [366](1.6.17). Nevertheless, we prefer to include simultaneous primitive recursion as a primitive concept since it is used in this form in the soundness proofs of our proof interpretations in chapters 5 and 8. Moreover, we will later (chapter 17) extend our framework to new types where such a reduction does not seem to be possible anymore (unless one introduces product types explicitly which would be another alternative).
2) Occasionally, we will denote the set of all closed terms of WE-HA$^\omega$ by T and use T_n to denote the subset of closed terms involving only recursors R_ρ with $deg(\rho) \le n$.

In the following FV(t) (FV(A)) denotes the set of all free variables of t (A).

WE-HA$^\omega$ allows the definition of λ-abstraction in the following sense:

Lemma 3.15. *For every term $t[x^\rho]^\tau$ (here x refers to all occurrences of x in t) one can construct in* WE-HA$^\omega$ *a term $(\lambda x^\rho.t[x])$ of type $\tau\rho$ (with $FV(\lambda x^\rho.t[x]) = FV(t[x]) \setminus \{x\}$) such that*

$$\text{WE-HA}^\omega \vdash (\lambda x^\rho.t[x])(s^\rho) =_\tau t[s/x].$$

In contexts where is no danger of ambiguity, we omit the outer parentheses around $(\lambda x.t[x])$.

Proof: Define

$$\lambda x.x := \Sigma\Pi\Pi,$$
$$\lambda x.t := \Pi t, \text{ if } x \notin \text{FV}(t),$$
$$\lambda x.(ts) := \Sigma(\lambda x.t)(\lambda x.s), \text{ if } x \in \text{FV}(ts)$$

(here Π, Σ of suitable types). □

Notation: Instead of $\lambda x_1 \ldots \lambda x_k.t$ we often write $\lambda x_1, \ldots, x_k.t$.

Remark 3.16. It is easy to see that using R_0 and lemma 3.15 one can define all primitive recursive functions so that HA can be viewed as a subsystem of WE-HA$^\omega$. For details see [366](1.6.9). Note that the use of R_ρ in [366](1.6.9) can be replaced by R_0 if T_{ψ_1} and T_{ψ_2} are replaced by $T_{\Psi_1}(x_1,\ldots,x_n)$ and $\lambda u^0, z^0.T_{\Psi_2}(u,z,x_1,\ldots,x_n)$ respectively.

Proposition 3.17. *Let $A_0(\underline{x})$ be a quantifier-free formula of $\mathcal{L}$(WE-HA$^\omega$) whose free variables are contained among $\underline{x}$. Then one can construct a closed term t such that*

$$\text{WE-HA}^\omega \vdash \forall \underline{x}\big(t\underline{x} =_0 0 \leftrightarrow A_0(\underline{x})\big).$$

Proof: Analogously to the proof of proposition 3.8 using lemma 3.15 and the previous remark. □

Corollary 3.18. *Let* A_0 *be a quantifier-free formula of* $\mathcal{L}(\text{WE-HA}^\omega)$*. Then*

$$\text{WE-HA}^\omega \vdash A_0 \vee \neg A_0.$$

In particular, quantifier-free formulas are provably stable, i.e.

$$\text{WE-HA}^\omega \vdash \neg\neg A_0 \to A_0.$$

Proposition 3.19. *For each type* $\rho \in \mathbf{T}$ *there exists a closed term t (using only the recursor* R_0 *of type* 0*) such that*

$$\text{WE-HA}^\omega \vdash \forall x^0, y_1^\rho, y_2^\rho ([x =_0 0 \to txy_1y_2 =_\rho y_1] \wedge [x \neq_0 0 \to txy_1y_2 =_\rho y_2]).$$

Proof: Define

$$\chi(x^0, y_1^0, y_2^0) := R_0(x, y_1, \lambda n^0, m^0.y_2).$$

Then

$$\chi(0, y_1, y_2) =_0 y_1 \text{ and } \chi(S(x), y_1, y_2) =_0 y_2.$$

Hence, since $x \neq_0 0 \to x =_0 S(pd(x))$,

$$x \neq_0 0 \to \chi(x, y_1, y_2) =_0 y_2.$$

Let $\rho = 0\rho_k \dots \rho_1$ and $\underline{v}^{\underline{\rho}} := v_1^{\rho_1}, \dots, v_k^{\rho_k}$ be distinct variables. Now define

$$t := \lambda x^0, y_1^\rho, y_2^\rho, \underline{v}^{\underline{\rho}}.\chi(x, y_1\underline{v}, y_2\underline{v}).$$

□

Definition 3.20. IL^ω is $\text{IL}^\omega_{-=}$ together with $=_0$ and the equality axioms for $=_0$ and QF-ER. $\text{PL}^\omega_{(-=)}$ (WE-PA^ω, E-PA^ω) is the extension of $\text{IL}^\omega_{(-=)}$ (WE-HA^ω, E-HA^ω) obtained by adding the law-of-excluded-middle schema LEM (i.e. $A \vee \neg A$ for **arbitrary** formulas A).

The set-theoretic functionals denoted by the closed terms of E-HA^ω (see the model $\mathcal{S}^\omega$ defined in section 3.6 below) are called the ‘Gödel primitive recursive functionals of finite type’. They were introduced first in [133] (but see also [161]). In the exercises to this chapter we show that the Gödel primitive recursive functionals of type degree 1 form a larger class than the ordinary primitive recursive functions. In fact, the former class coincides with the provably recursive functions of Peano arithmetic PA: that this class contains the provably recursive functions of PA follows from Gödel’s ([133]) functional interpretation (see chapter 9). The other inclusion follows from work of Parsons ([299]) and others.
The combinators Π and Σ are due already to [323]. The correspondence between these combinators and the usual Hilbert-style axiomatization of the implicative fragment of intuitionistic propositional logic given by (the modus ponens rule and) the schemata

$$A \to (B \to A)$$
$$(A \to (B \to C)) \to ((A \to B) \to (A \to C))$$

as well as the correspondence between typed λ-terms and the natural deduction style formalization of that fragment are known as the 'Curry-Howard'-isomorphism (see [166, 349]).

We now consider various fragments of full arithmetic in all types which will be used later. Most of the techniques used in this book apply to all of these systems down to $\mathrm{G}_2\mathrm{A}^\omega$ (to be defined below) whose provably recursive functions are bounded by polynomials. This shows that the proof interpretations on which the applications to proof mining in analysis discussed in chapters further below are based do not produce any non-polynomial growth of numerical bounds by themselves. So if a given proof implicitly contains a bound say of polynomial growth, then the unwinding process will produce such a bound. However, we will usually only sketch how the results proved in this book for (W)E-HA$^\omega$, (W)E-PA$^\omega$ can be adapted to these fragments and refer to the literature for more details. This is in order to avoid to have to deal with too many formal systems (a general malaise in the area of proof theory) but also for the following important reason: weak formal systems are only needed to state **a-priori** that a certain proof allows one to extract data of certain low complexity because it can be formalized in such a weak system. For actual proof mining where such data are explicitly extracted one will instead work in a stronger framework to get an easier formalization of the proof. If the proof indeed contains numerical data of low complexity the proof mining procedure will produce such if the procedure is faithful. It is only to show the latter point why it is of relevance to verify that in principle all techniques used in the procedure applied can be adapted to work for such weak systems as well.

3.4 Fragments of (W)E-HA$^\omega$ and (W)E-PA$^\omega$

Let $\widehat{\text{(W)E-PA}}^{\omega}\!\restriction$ ($\widehat{\text{(W)E-HA}}^{\omega}\!\restriction$) be the fragment of (W)E-PA$^\omega$ ((W)E-HA$^\omega$) where we only have the recursor R_0 for type-0-recursion and the induction schema is restricted to the schema of quantifier-free induction

$$\text{QF-IA} : A_0(0) \land \forall x^0 \big(A_0(x) \to A_0(S(x))\big) \to \forall x^0 A_0(x),$$

where A_0 is quantifier-free and may contain parameters of arbitrary types.

The set-theoretic functionals denoted by the closed terms of $\widehat{\text{E-HA}}^{\omega}\!\restriction$ are called the 'Kleene primitive recursive functionals of finite type'. They were (for pure types) first introduced in [195] under the name of S1-S8 computable functionals. In contrast to the Gödel primitive recursive functionals, the Kleene primitive recursive functionals of type degree 1 are just the ordinary primitive recursive functions (see

e.g. [7] for a proof of this).
The systems $\widehat{\text{(W)E-HA}}^{\omega}{\upharpoonright}$ were introduced in [98] (see also [299]). Proposition 3.17, corollary 3.18 and proposition 3.19 hold analogously for $\widehat{\text{WE-HA}}^{\omega}{\upharpoonright}$ instead of WE-HA$^\omega$.

In the presence of the quantifier-free axiom of choice schema for numbers

$$\text{QF-AC}^{0,0}: \ \forall x^0 \exists y^0 A_0(x,y) \to \exists f^1 \forall x^0 A_0(x, f(x)) \ (A_0 \text{ quantifier-free}),$$

the schema of induction for Σ^0_1-formulas

$$\Sigma^0_1\text{-IA}: \ \exists y^0 A_0(0,y) \wedge \forall x^0 \big(\exists y A_0(x,y) \to \exists y A_0(S(x),y)\big) \to \forall x \exists y A_0(x,y),$$

where A_0 is quantifier-free (with parameters of arbitrary types), becomes derivable from QF-IA:

Proposition 3.21. $\widehat{\text{WE-PA}}^{\omega}{\upharpoonright} +$ QF-AC$^{0,0} \vdash \Sigma^0_1$ -IA.

Proof: Assume $\exists y_0 A_0(0,y_0)$ and $\forall x, y_1 \exists y_2 \big(A_0(x,y_1) \to A_0(S(x),y_2)\big)$. By QF-AC^{0,0} we get (using that x, y_1 can be coded into a single variable, see lemma 3.30 below)

$$\exists f \forall x, y_1 \big(A_0(x,y_1) \to A_0(S(x), f(x,y_1))\big).$$

Define

$$\begin{cases} \Phi(0,y,f) :=_0 y \\ \Phi(S(x),y,f) :=_0 f(x, \Phi(x,y,f)) \end{cases}$$

(note that this can be done by R_0). Then by QF-IA one easily shows that

$$\forall x A_0(x, \Phi(x, y_0, f))$$

for y_0 such that $A_0(0,y_0)$ and, therefore, $\forall x \exists y A_0(x,y)$. □

Whereas inspection of the proof above shows that the result also holds for the intuitionistic system $\widehat{\text{WE-HA}}^{\omega}{\upharpoonright} +$ QF-AC^{0,0}, the next result requires classical logic: QF-AC^{0,0} also allows one to prove the schema of Δ^0_1-comprehension

$$\Delta^0_1\text{-CA}: \ \forall x^0 \big(\exists y^0 A_0(x,y) \leftrightarrow \forall y^0 B_0(x,y)\big) \to \exists f^1 \forall x^0 \big(f(x) = 0 \leftrightarrow \exists y^0 A_0(x,y)\big),$$

where again parameters in all types are allowed (A_0, B_0 quantifier-free).

Proposition 3.22. $\widehat{\text{WE-PA}}^{\omega}{\upharpoonright} +$ QF-AC$^{0,0} \vdash \Delta^0_1$ -CA.

Proof: Exercise! □

As the proof of lemma 3.15 shows we still can define λ-abstraction in $\widehat{\text{(W)E-PA}}^{\omega}{\upharpoonright}$. Using λ-abstraction and R_0 one can define recursors $\widehat{R}_\rho$ such that

$$\begin{cases} \widehat{R}_\rho 0yz\underline{w} =_0 y\underline{w} \\ \widehat{R}_\rho (Sx)yz\underline{w} =_0 z(\widehat{R}_\rho xyz\underline{w})x\underline{w}, \end{cases}$$

where y is of type $\rho = 0\rho_k \dots \rho_1$, $\underline{w} = w_1^{\rho_1} \dots w_k^{\rho_k}$ and the type of z is $\rho 00$ (exercise).

The crucial difference between $\widehat{R}_\rho$ and the much stronger R_ρ is that in the case of the former $\widehat{R}_\rho xyz\underline{w}$ may only be used with the fixed set of parameters $\underline{w}$ in the recursion step while in the case of R_ρ we can use the whole functional $R_\rho xyz$.

3.5 Fragments corresponding to the Grzegorczyk hierarchy

We now define a hierarchy of systems $\mathrm{G}_n\mathrm{A}^\omega$ corresponding (w.r.t. the definable and provably total functions) to n-th level of the so-called Grzegorczyk hierarchy [147]. Following Ritchie [314] we base our definition on the n-th branch of the Ackermann function [1]:

Definition 3.23. Let $n \in \mathbb{N}$. We define (by recursion on n from the outside) the n-th branch of the Ackermann function $A_n : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$ by

$$A_0(x,y) := y' \quad \text{(Here and in the following } x' \text{ denotes the successor } Sx \text{ of } x\text{)},$$

$$A_{n+1}(x,0) := \begin{cases} x, \text{ if } n = 0 \\ 0, \text{ if } n = 1 \\ 1, \text{ if } n \geq 2, \end{cases} \qquad .$$

$$A_{n+1}(x,y') := A_n(x, A_{n+1}(x,y))$$

Remark 3.24. 1) $A_1(x,y) = x + y$, $A_2(x,y) = x \cdot y$, $A_3(x,y) = x^y$,

$A_4(x,y) = x^{x^{\cdot^{\cdot^{\cdot x}}}}$ (y times).

2) For each fixed $n \in \mathbb{N}$ the function A_n is primitive recursive. However, as first shown by Ackermann [1], the diagonal function $A(x) := A_x(x,x)$ is no longer primitive recursive (see e.g. [341]).

The **intuitionistic Grzegorczyk arithmetic $\mathrm{G}_n\mathrm{A}_i^\omega$ of level n in all finite types** and its **classical variant $_n\mathrm{A}^\omega$** :

The language $\mathscr{L}(\mathrm{G}_n\mathrm{A}_i^\omega)$ of $\mathrm{G}_n\mathrm{A}_i^\omega$ is the extension of $\mathscr{L}(\mathrm{IL}^\omega_{=})$ resulting from the addition of the constant 0^0, the projectors $\Pi_{\rho,\tau}$, the combinators $\Sigma_{\delta,\rho,\tau}$, function constants S^{00} (successor), $\max_0^{000}, \min_0^{000}, A_0^{000}, \dots, A_n^{000}$ and functional constants $\Phi_1^{001}, \dots, \Phi_n^{001}$, $\mu_b^{00(000)}$ (bounded μ-operator), $\tilde{R}_0$ (bounded recursion of type 0) of type $01(000)00$. Furthermore we have a predicate symbol $=_0$ for equality at type 0 and a predicate symbol $<_0$ for objects of type 0.

In addition to the axioms and rules of $\mathrm{IL}^{\omega}_{-=}$ the theory $\mathrm{G}_n\mathrm{A}^{\omega}_i$ contains the following (with $x \le_0 y$ being defined as $x < y \vee x = y$):

1) the Π, Σ-axioms (as in the case of E-HA$^{\omega}$).
2) the equality axioms for $=_0$:
   $x =_0 x$, $x =_0 y \to y =_0 x$, $x =_0 y \wedge y =_0 z \to x =_0 z$.
   $x_1 =_0 x_2 \wedge y_1 =_0 y_2 \wedge x_1 < y_1 \to x_2 < y_2$.
3) $<_0$-axioms: $\neg x <_0 0$, $x <_0 Sy \leftrightarrow x <_0 y \vee x =_0 y$, $x <_0 y \vee x =_0 y \vee y <_0 x$.
4) S-axioms: $Sx =_0 Sy \to x =_0 y$, $\neg 0 =_0 Sx$.
5) (max) : $\max_0(x,y) \ge_0 x$, $\max_0(x,y) \ge_0 y$, $\max_0(x,y) =_0 x \vee \max_0(x,y) =_0 y$.
6) (min) : $\min_0(x,y) \le_0 x$, $\min_0(x,y) \le_0 y$, $\min_0(x,y) =_0 x \vee \min_0(x,y) =_0 y$.
7) The defining recursion equations for $A_0, \ldots, A_n$ from the definition 3.23 above.
8) Defining recursion equations for $\Phi_1, \ldots, \Phi_n$:

$$\begin{cases} \Phi_1 f0 =_0 f0 \\ \Phi_1 fx' =_0 \max_0(fx', \Phi_1 fx) \end{cases}$$

   and

$$\begin{cases} \Phi_i f0 =_0 f0 \\ \Phi_i fx' =_0 A_{i-1}(fx', \Phi_i fx), \quad \text{for } i \ge 2. \end{cases}$$

9)

$$(\mu_b): \begin{cases} y \le_0 x \wedge f^{000}xy =_0 0 \to fx(\mu_b fx) =_0 0, \\ y <_0 \mu_b fx \to fxy \ne_0 0, \\ \mu_b fx =_0 0 \vee (fx(\mu_b fx) =_0 0 \wedge \mu_b fx \le_0 x). \end{cases}$$

10) Defining recursion equations for $\tilde{R}_0$ (bounded recursion of type 0):

$$\begin{cases} \tilde{R}_0 0yzv =_0 y \\ \tilde{R}_0 x'yzv =_0 \min_0(z(\tilde{R}_0 xyzv)x, vx). \end{cases}$$

11) All $\mathbb{N}, \mathbb{N}^{\mathbb{N}}, \mathbb{N}^{(\mathbb{N}^{\mathbb{N}})}$-true purely universal sentences $\forall \underline{x} A_0(\underline{x})$, where $\underline{x}$ is a tuple of variables whose types have a degree ≤ 2, where B^A denotes the set of all set-theoretic functions $f : A \to B$.
12) The quantifier-free extensionality rule QF-ER.

$\mathrm{G}_n\mathrm{A}^{\omega}$ is the variant of $\mathrm{G}_n\mathrm{A}^{\omega}_i$ with the law-of-excluded-middle schema $A \vee \neg A$ added. Analogously, we define $\mathrm{G}_\infty\mathrm{A}^{\omega}$
If we add $(E) = \bigcup_\rho \{(E_\rho)\}$ to $\mathrm{G}_n\mathrm{A}^{\omega}$,$\mathrm{G}_n\mathrm{A}^{\omega}_i$ we obtain theories which are denoted by E-$\mathrm{G}_n\mathrm{A}^{\omega}$, E-$\mathrm{G}_n\mathrm{A}^{\omega}_i$. $\mathrm{G}_n\mathrm{R}^{\omega}$ denotes the set of all closed terms of $\mathrm{G}_n\mathrm{A}^{\omega}$.
(E)-$\mathrm{G}_\infty\mathrm{A}^{\omega}_{(i)} := \bigcup\limits_{n\ge 1} \{$ (E)-$\mathrm{G}_n\mathrm{A}^{\omega}_{(i)}\}$ and $\mathrm{G}_\infty\mathrm{R}^{\omega} := \bigcup\limits_{n\ge 1} \{\mathrm{G}_n\mathrm{R}^{\omega}\}$.

Remark 3.25. 1) Our axioms contain w.r.t. $0,S,+,\cdot,<$ what is called Robinson's system Q which specifies (numeralwise) the meaning of $0,S,+,\cdot,<$ when interpreted in the standard model $\mathbb{N}$ and hence – when augmented by the schema of full induction – results in a system containing a version of Peano arithmetic (see e.g. [363] (I.9) for more information on this). Also the meaning of the other constants is – when interpreted in the full type-structure over $\mathbb{N}$ (to be defined below) – uniquely determined by the axioms.

2) The functionals Φ_1,Φ_2 and Φ_3 have the following meaning: $\Phi_1 fx = \max(f0,f1,\ldots,fx)$, $\Phi_2 fx = \sum_{y=0}^{x} fy$, $\Phi_3 fx = \prod_{y=0}^{x} fy$. In general, for $i \geq 2$, $\Phi_i fx$ is the iteration of the (i-1)-th branch A_{i-1} of the Ackermann function on the f-values $f0,\ldots,fx$.

3) The axioms on μ_b formalize that

$$\mu_b fx := \begin{cases} \min y \leq_0 x(fxy =_0 0), \text{ if such an } y \leq x \text{ exists,} \\ 0, \text{ otherwise.} \end{cases}$$

4) As in the case of $\widehat{\text{WE-HA}}^{\omega}\!\upharpoonright$ one can define λ-abstraction in $\mathrm{G}_n\mathrm{A}_i^{\omega}$ as well as bounded recursors $\tilde{R}_\rho$ that satisfy

$$\begin{cases} \tilde{R}_\rho 0yzv\underline{w} =_0 y\underline{w} \\ \tilde{R}_\rho x'yzv\underline{w} =_0 \min{}_0(z(\tilde{R}_\rho xyzv\underline{w})x\underline{w}, vx\underline{w}), \end{cases}$$

where $\rho = 0\rho_k\ldots\rho_1$ and $\underline{w} = w_1^{\rho_1}\ldots w_k^{\rho_k}$.

5) Our definition of $\mathrm{G}_n\mathrm{A}^{\omega}$ contains some redundancies (which however we want to keep for greater flexibility of our language): E.g. Φ_i $(i > 1)$ can be defined from $A_i,\tilde{R}_0,\min_0$ and Φ_1: With $f^M := \lambda x.\Phi_1 fx$ one can show that $\Phi_i fx \leq \Phi_i f^M x \leq A_i(f^M(x)+1,x+1)$. Hence Φ_i can be defined by $\tilde{R}_0$ using $A_i(f^M(x)+1,x+1)$ as boundary function v.

6) The axiom of quantifier-free induction

$$(*)\ \forall f^1,x^0\big(f0 =_0 0 \wedge \forall y < x(fy =_0 0 \to fy' =_0 0) \to fx =_0 0\big)$$

can be expressed as a universal sentence $\forall f^1,x^0 A_0$ by prop. 3.28 below and thus is an axiom of $\mathrm{G}_n\mathrm{A}_i^{\omega}$. $(*)$ implies every instance (with parameters of arbitrary type) of the schema of quantifier-free induction in the form (equivalent to QF-IA)

$$\forall x^0\big(A_0(0) \wedge \forall y < x(A_0(y) \to A_0(y')) \to A_0(x)\big)$$

since again by prop. 3.28 there exists a term t such that $tx =_0 0 \leftrightarrow A_0(x)$ for quantifier-free A_0. Now apply $(*)$ to $f := t$.

7) Our reason for including all true universal sentences as axioms in 11) is as follows: it is an old observation made by G. Kreisel that proofs of universal lemmas have no impact on the extraction of programs or bounds from proofs. In fact, all the extraction techniques developed in this book allow one to treat universal

lemmas (and most of the time even more general classes of lemmas) as axioms. For the systems (W)E-PA$^\omega$ etc. we make this explicit in our main theorems on program and bound extractions rather than adding them beforehand as axioms already to the definition of (W)E-PA$^\omega$ since, occasionally, we state 'foundational' corollaries on the proof-theoretic strength of systems, consistency proofs, conservation theorems over PA or PRA etc. that would be spoiled by this. However, for systems as weak as say $G_2A^\omega|\backslash$ such foundational issues are less relevant and to verify some basic properties of functions such as $\max(x,y), |x-y|$ and codings etc. would be very tedious to carry out using just QF-IA. From $n \geq 3$ on, usually the standard proofs of these facts known from primitive recursive arithmetic PRA go through even in the variant of G_nA^ω with the universal axioms 11) replaced by the schema QF-IA and we make free use of this in chapter 13. The reason for our restriction of the types in the universal axioms we add is that in some places in this book we deal with principles which are valid only in the type structure $\mathscr{M}^\omega$ of the so-called majorizable functionals due to [27] (which will be discussed further below) but not in the full type structure $\mathscr{S}^\omega$ of all set-theoretic functionals. Since both type structures coincide up to type 1 and for the type 2 the inclusion $\mathscr{M}_2^\omega \subset \mathscr{S}_2^\omega$ holds, the implication $\mathscr{S}^\omega \models \forall x^\rho A_0 \Rightarrow \mathscr{M}^\omega \models \forall x^\rho A_0$ is obvious **if** $deg(\rho) \leq 2$.

In the following we make free use of the fact that universal lemmas are included as axioms if true.

Already in $G_1A_i^\omega$ we can, using $\tilde{R}_0$ and trivial bounding functions, define the following functions:

Definition 3.26. 1) $prd(0) =_0 0,\ prd(x') =_0 x$ (predecessor),

2) $\begin{cases} sg(0) =_0 0, \\ sg(x') =_0 1, \end{cases}$ and $\begin{cases} \overline{sg}(0) =_0 1, \\ \overline{sg}(x') =_0 0. \end{cases}$

3) $\begin{cases} x \dot{-} 0 =_0 x \\ x \dot{-} y' =_0 prd(x \dot{-} y), \end{cases}$

4) $|x-y| =_0 \max(x \dot{-} y, y \dot{-} x)$ (symmetrical difference).

For the rest of this section we usually omit the type in $=_0$ and simply write $=$.

Remark 3.27. The following basic properties of the functions defined above are all purely universal and so (because of the universal axioms 11) trivially provable already in $G_1A_i^\omega$:
$sg(x) = 0 \leftrightarrow x = 0,\ \overline{sg}(x) = 0 \leftrightarrow x \neq 0,\ sg(x) \leq 1,\ \overline{sg}(x) \leq 1,\ prd(x) \leq x,$
$|x-y| = 0 \leftrightarrow x = y,\ x = 0 \vee x = S(prd(x)),\ \max(x,y) = 0 \leftrightarrow x = 0 \wedge y = 0,$
$\min(x,y) = 0 \leftrightarrow x = 0 \vee y = 0,\ \max(x,y) = y \leftrightarrow x \leq y,\ x < y \leftrightarrow Sx \dot{-} y = 0.$

Proposition 3.28. *Let $n \geq 1$. For each formula $A \in \mathscr{L}(G_nA^\omega)$ which contains no quantifiers except for bounded quantifiers of type 0 one can construct a closed term t_A in G_nA^ω such that*

$$\mathrm{G}_n\mathrm{A}_i^\omega \vdash \forall x_1^{\rho_1},\ldots,x_k^{\rho_k}\big(t_A x_1 \ldots x_k =_0 0 \leftrightarrow A(x_1,\ldots,x_k)\big),$$

where $x_1,\ldots,x_k$ contain all the free variables of A.

Proof: Induction on the logical structure of A using the remark above. Bounded quantifiers are captured by μ_b:

$$\mathrm{G}_n\mathrm{A}_i^\omega \vdash \exists y \leq_0 x A_0(x,y,\underline{a}) \overset{(\mu_b)}{\leftrightarrow} A_0(x,\mu_b(\lambda x,y.t_{A_0}xy\underline{a},x),\underline{a}).$$

Similarly for the bounded universal quantifier. □

Proposition 3.29. *Let $n \geq 1$ and $A_0(\underline{x}) \in \mathscr{L}(\mathrm{G}_n\mathrm{A}^\omega)$ be a quantifier-free formula, where $\underline{x} = x_1^{\rho_1}\ldots x_k^{\rho_k}$ contain all the free variables of A_0, and $t_1^{0\rho_k\ldots\rho_1}, t_2^{0\rho_k\ldots\rho_1}$ are closed terms of $\mathrm{G}_n\mathrm{A}^\omega$. Then there exists a closed term $\Phi^{0\rho_k\ldots\rho_1}$ in $\mathrm{G}_n\mathrm{A}^\omega$ such that*

$$\mathrm{G}_n\mathrm{A}_i^\omega \vdash \forall \underline{x}\left(\Phi\underline{x} =_0 \begin{cases} t_1\underline{x}, \text{ if } A_0(\underline{x}) \\ t_2\underline{x}, \text{ if } \neg A_0(\underline{x}). \end{cases}\right)$$

Proof: Define

$$\Phi := \lambda \underline{x}.\tilde{R}_0(t_{A_0}\underline{x})(t_1\underline{x})(\lambda y^0,z^0.t_2\underline{x})(\lambda y^0.t_2\underline{x}),$$

where t_{A_0} is as in the previous proposition. One easily verifies that Φ does the job. □

We now show how to encode pairs, tuples and finite sequences of numbers:

Definition 3.30 (and lemma). For $\mathbf{n \geq 2}$ we can define the well-known surjective Cantor pairing function j with its projections in $\mathrm{G}_n\mathrm{R}^\omega$:

$$j(x^0,y^0) := \begin{cases} \min u \leq_0 (x+y)^2+3x+y[2u =_0 (x+y)^2+3x+y] \text{ if existent} \\ 0^0, \text{ otherwise,} \end{cases}$$
$$j_1 z := \min x \leq_0 z[\exists y \leq z(j(x,y)=z)],$$
$$j_2 z := \min y \leq_0 z[\exists x \leq z(j(x,y)=z)].$$

Using j, j_1, j_2 we can define a coding of k-tuples for every **fixed** number k by

$$\nu^1(x_0) := x_0,\ \nu^2(x_0,x_1) := j(x_0,x_1),\ \nu^{k+1}(x_0,\ldots,x_k) := j(x_0,\nu^k(x_1,\ldots,x_k)),$$
$$\nu_1^1(x) := x \text{ and (for } k>1)\ \ \nu_i^k(x) := \begin{cases} j_1 \circ (j_2)^{i-1}(x), \text{ if } 1 \leq i < k \\ (j_2)^{k-1}(x), \text{ if } 1 < i = k. \end{cases}$$

Indeed, these functions satisfy the following properties of a surjective tuple coding:

$$\nu_i^k(\nu^k(x_1,\ldots,x_k)) = x_i \quad (1 \leq i \leq k)$$

and

$$\nu^k(\nu_1^k(x), \ldots, \nu_k^k(x)) = x.$$

Following essentially [366] we now extend this tuple coding to a coding of finite sequences:

$$\langle\rangle := 0,\ \langle x_0, \ldots, x_k\rangle := S(\nu^2(k, \nu^{k+1}(x_0, \ldots, x_k))).$$

Using $\tilde{R}_0$ we define functions $lth, \Pi(x,y) \in \mathrm{G}_n\mathrm{R}^\omega$ such that for every fixed k

$$lth(\langle\rangle) = 0, \quad lth(\langle x_0, \ldots, x_k\rangle) = k+1$$

and (for $x = \langle x_0, \ldots, x_m\rangle$)

$$\Pi(x,y) = \begin{cases} x_y, \text{ if } y \le m \\ 0^0, \text{ otherwise.} \end{cases}$$

Define

$$lth(x) := \begin{cases} 0^0, \text{ if } x =_0 0 \\ j_1(x \dot{-} 1) + 1, \text{ otherwise,} \end{cases}$$

$$\Pi(x,y) := \begin{cases} 0^0, \text{ if } lth(x) \le y \\ j_1 \circ (j_2)^{y+1}(x \dot{-} 1), \text{ if } 0 \le y < lth(x) \dot{-} 1 \\ (j_2)^{lth(x)}(x \dot{-} 1), \text{ if } lth(x) > 0 \wedge y = lth(x) \dot{-} 1 \end{cases}$$

To improve the readability we normally write $(x)_y$ instead of $\Pi(x,y)$.
That $\Pi(x,y)$ is definable even in $\mathrm{G}_2\mathrm{R}^\omega$, follows from the fact that the iteration $\varphi xy = (j_2)^y(x)$ of j_2 is definable in $\mathrm{G}_2\mathrm{R}^\omega$ since $\varphi xy \le x$ for all x, y so that we can use $\lambda y.x$ as bounding function.

We need $\mathrm{G}_3\mathrm{R}^\omega$ to define a coding of initial segments of **variable** length of a function f. Indeed, there is a functional $\Phi_{\langle\rangle} \in \mathrm{G}_3\mathrm{R}^\omega$ such that $\Phi_{\langle\rangle} fx = \langle f0, \ldots, f(x \dot{-} 1)\rangle$. Of course we cannot write $\langle f0, \ldots, f(x \dot{-} 1)\rangle$ for variable x. However the meaning of $\Phi_{\langle\rangle} fx$ can be expressed via $(\Phi_{\langle\rangle} fx)_y = fy$ for all $y < x$ (and $= 0$ for $y \ge x$).
To achieve this, we first define

$$\begin{cases} \tilde{f}0 = f0 \\ \tilde{f}x' = j(fx', \tilde{f}x) \end{cases}$$

in $\mathrm{G}_3\mathrm{R}^\omega$: One easily verifies (using $j(x,x) \le 4x^2$) that $\tilde{f}x \le 4^{3^x}\left(f^M x\right)^{2^x}$ for all x. Hence the definition of $\tilde{f}$ can be carried out by $\tilde{R}_0$ using as our bounding function

$$\lambda x. j\left(fx', 4^{3^x}\left(f^M x\right)^{2^x}\right) \in \mathrm{G}_3\mathrm{R}^\omega.$$

With $\tilde{j}(x,y) := j(y,x)$ we see that $\tilde{f}x$ means $\tilde{j}(\dots\tilde{j}(\tilde{j}(f0,f1),f2)\dots fx)$. Hence $\widehat{fx} := (\lambda y.f(x \dot{-} y))x$ has the meaning $j(f0,\dots j(f(x-2),j(f(x-1),fx))\dots)$.
Finally, we are now in the position to define $\Phi_{\langle\rangle} \in \mathrm{G_3R}^\omega$:

$$\Phi_{\langle\rangle} fx := \begin{cases} 0^0, \text{ if } x = 0 \\ \widehat{(f_x)}x + 1, \text{ otherwise,} \end{cases}$$

where

$$f_x y := \begin{cases} x, \text{ if } y = 0 \\ f(y \dot{-} 1), \text{ otherwise.} \end{cases}$$

Again for better readability, we usually write $\overline{f}x$ instead of $\Phi_{\langle\rangle} fx$.
Next we define a function $*$ in $\mathrm{G_3R}^\omega$ by

$$n * m := \Phi_{\langle\rangle}(fnm)(lth(n) + lth(m)),$$

where

$$(fnm)(k) := \begin{cases} (n)_k, \text{ if } k < lth(n) \\ (m)_{k \dot{-} lthn}, \text{otherwise.} \end{cases}$$

Then

$$\langle x_0,\dots,x_k\rangle * \langle y_0,\dots,y_m\rangle = \langle x_0,\dots,x_k,y_0,\dots,y_m\rangle.$$

Note that $\Phi_{\langle\rangle}$ and $*$ are not definable in $\mathrm{G_2R}^\omega$ since their definitions involve an iteration of the polynomial j.

Remark 3.31. 1) For detailed information on this as well as various other codings see [341] and also [101] (where j is called 'Cauchy's pairing function').
2) One easily shows that $(x+y)^2 + 3x + y$ is always even so that the case 'otherwise' in the definition of j never occurs and $2j(x,y) = (x+y)^2 + 3x + y$ for all x,y.

Definition 3.32. For arbitrary $\rho \in \mathbf{T}$ we define the relation $x_1 \geq_\rho x_2$ between functionals x_1,x_2 of type ρ by induction on ρ:

$$\begin{cases} x_1 \geq_0 x_2 \text{ is defined already,} \\ x_1 \geq_{\tau\rho} x_2 :\equiv \forall y^\rho (x_1 y \geq_\tau x_2 y). \end{cases}$$

$x_1 \leq_\rho x_2 :\equiv x_2 \geq_\rho x_1$.

Lemma 3.33. *Let* $\rho = \tau\rho_k \dots \rho_1$. *Then*

$$\mathrm{G_1A}_i^\omega \vdash x_1 \geq_\rho x_2 \leftrightarrow \forall y_1^{\rho_1},\dots,y_k^{\rho_k}(x_1\underline{y} \geq_\tau x_2\underline{y}).$$

We now for the first time make use of an important structural property of the closed terms of all our systems. This property of 'majorizability' which is due to W.A.

Howard will play key roles (in different ways) in the proofs of numerous results in this book. In the present chapter it is used to prove results on the growth of the definable functionals of $\mathrm{G}_n\mathrm{A}^\omega$. We will focus on the cases $n = 1,2,3$ as only those (actually only $n = 2,3$) are of practical interest (the whole hierarchy is treated in [207]).

Definition 3.34 (W.A. Howard [163]). We define the relation $x^*\ maj_\rho\ x$ (x^* majorizes x) between functionals of type ρ by induction on ρ:

$$\begin{cases} x^*\ maj_0\ x :\equiv x^* \geq_0 x, \\ x^*\ maj_{\tau\rho}\ x :\equiv \forall y^*, y(y^*\ maj_\rho\ y \to x^*y^*\ maj_\tau\ xy). \end{cases}$$

Lemma 3.35. $\mathrm{G}_1\mathrm{A}^\omega_i$ *proves:*

(i) $\tilde{x}^* =_\rho x^* \wedge \tilde{x} =_\rho x \wedge x^*\ maj_\rho\ x \to \tilde{x}^*\ maj_\rho\ \tilde{x}$.
(ii) $x^*\ maj_\rho\ x \wedge x \geq_\rho y \to x^*\ maj_\rho\ y$.
(iii) *For* $\rho = \tau\rho_k \ldots \rho_1$:

$$x^*\ maj_\rho\ x \leftrightarrow \forall y_1^*, y_1, \ldots, y_k^*, y_k \Big(\bigwedge_{i=1}^{k} (y_i^*\ maj_{\rho_i}\ y_i) \to x^*\underline{y}^*\ maj_\tau\ x\underline{y} \Big).$$

Proof: Induction on the type respectively on k. □

Remark 3.36. 1) The relation maj_ρ is a kind of hereditary form of $\geq_\rho$ (combined with monotonicity) and is (in contrast to $\geq_\rho$) a so-called logical relation in the sense of G. Plotkin which implies a nice behavior w.r.t. substitution (see lemma 3.35 (iii)). Because of this, results on the majorization of complex terms can be established directly by induction on the term structure without any use of normalization.
2) The previous lemma can be proved also in $\widehat{\text{WE-HA}}^\omega{\restriction}$ since only the transitivity of $\leq_0$ is used (which can be proved by QF-IA) but no general universal axioms 11).

Next we need some basic properties of the constants in $\mathrm{G}_n\mathrm{A}^\omega_i$:

Lemma 3.37. *Provably in* $\mathrm{G}_n\mathrm{A}^\omega_i$ *(if applicable) the following holds:*

1) $\Pi\ maj\ \Pi$ *and* $\Sigma\ maj\ \Sigma$.
2) $0\ maj\ 0$, $S\ maj\ S$, $A_i\ maj\ A_i$ $(i \leq 2)$, $A_j^* := \lambda x^0, y^0.\max(A_j(x,y),1)\ maj\ A_j$ $(j \geq 3)$.
3) $\min\ maj\ \min$, $\max\ maj\ \max$.
4) $\Phi_1^* := \lambda f^1, x^0.f(x)\ maj\ \Phi_1$, $\Phi_2^* := \lambda f^1, x^0.f(x)\cdot(x+1)\ maj\ \Phi_2$, $\Phi_j^* := \lambda f^1, x^0.A_j(f(x)+1, x+1)\ maj\ \Phi_j$ $(j \geq 3)$.
5) $\mu_b^* := \lambda f, x.x\ maj\ \mu_b$.
6) $\tilde{R}_0^* := \lambda x, y, z, v.\max_0(y, v(prd(x)))\ maj\ \tilde{R}_0$.

Proof: Exercise! □

Definition 3.38. 1) $G_nR^{\omega}_{-} \subset G_nR^{\omega}$ denotes the subset of all closed terms in G_nR^{ω} that are built up from $\Pi_{\rho,\tau}, \Sigma_{\delta,\rho,\tau}, A_0, \ldots, A_n, 0^0, S, prd, \min_0$ and $\max_0$ only (i.e. which do not contain occurrences of $\Phi_1, \ldots, \Phi_n, \tilde{R}_0$ or μ_b).
2) $G_nR^{\omega}_{-}[\Phi_1]$ is the set of all closed terms built up from $G_nR^{\omega}_{-}$ plus Φ_1.

Proposition 3.39. *Let $n \geq 1$. To each term $t^{\rho} \in G_nR^{\omega}$ one can construct by direct induction on the structure of t a term $t^{*\rho} \in G_nR^{\omega}_{-}$ such that*

$$G_nA^{\omega}_i \vdash t^* \; maj_{\rho} \; t.$$

Proof: 1. Replace every occurrence of $\tilde{R}_0$ in t by $\tilde{R}_0^*$, which is built up from Π, Σ (which are used for defining the λ–operator) and the monotone functions $\max_0$ and prd.
2. Replace all occurrences of $A_3, \ldots, A_n$ in t by $A_3^*, \ldots, A_n^*$.
3. Replace all occurrences of $\Phi_1, \ldots, \Phi_n, \mu_b$ in t by $\Phi_1^*, \ldots, \Phi_n^*, \mu_b^*$.
Let t^* be the term which results from t after having carried out 1.–3. By construction, $t^* \in G_nR^{\omega}_{-}$. Moreover, t^* is constructed by replacing every constant c in t by a closed term s_c^* such that $s_c^* \; maj \; c$ (lemma 3.37). Since t is built up from constants only this implies (using lemma 3.35) $t^* \; maj \; t$. □

Corollary to the proof:
One can even achieve that the majorizing term t^* does not contain $S, prd, \max_0$ or $\min_0$ (though this in general will give a less good bound if we use t^* as a bound for t^0): this follows using $\lambda x^0.x \; maj_1 \; prd$ and $A_1 \; maj \; \max_0, \min_0$ and – as majorants for $\tilde{R}_0$ and A_j $(j \geq 3)$ – $\tilde{R}^* := \lambda x, y, z, v.(y + v(x))$ and $A_j^*(x,y) := A_j(x+1,y)+1$, where '$x+1$' can be replaced by $A_0 0x$.

The majorizing term t^* constructed in proposition 3.39 has a much simpler form than t as it does not contain any functional of type degree > 1 except for the projectors and combinators Π and Σ. We know show that if t^* has type $(\leq)2$, then we can rewrite the term t^*x^1 even in the form $\widehat{t}[x]$, where $\widehat{t}[x]$ no longer contains any projector or combinator (but does contain x^1) :

Proposition 3.40. *Let $n \geq 1$ and $\rho = 0\rho_k \ldots \rho_1$ with $deg(\rho_i) \leq 1$ for $i = 1, \ldots, k$ (i.e. $deg(\rho) \leq 2$). Moreover, let $t^{\rho} \in G_nR^{\omega}_{-}$. Then one can construct (by 'logical' normalization, i.e. by carrying out all possible Π, Σ-reductions) a term $\widehat{t}[x_1^{\rho_1}, \ldots, x_k^{\rho_k}]$ such that*

1) *$\widehat{t}[x_1, \ldots, x_k]$ contains at most $x_1 \ldots, x_k$ as free variables,*
2) *$\widehat{t}[x_1, \ldots, x_k]$ is built up only from $x_1, \ldots, x_k, A_0, \ldots, A_n, S^1, 0^0, prd, \min_0, \max_0$,*
3) *$G_nA^{\omega}_i \vdash \forall x_1^{\rho_1}, \ldots, x_k^{\rho_k} (\widehat{t}[x_1, \ldots, x_k] =_0 tx_1 \ldots x_k)$.*

Proof: We perform reductions $\Pi st \rightsquigarrow s$ and $\Sigma str \rightsquigarrow sr(tr)$ inside of $tx_1 \ldots x_k$ as long as no further such reduction is possible and denote the resulting term by $\widehat{t}[x_1, \ldots, x_k]$. The well-known strong normalization theorem for typed combinatory logic (see e.g. [370] (theorem 1.2.18) or [349] (theorem 5.3.6)) ensures that this situation will always occur after a finite number of reduction steps. Since $\Pi xy = x$ and $\Sigma xyz =$

$xz(yz)$ are axioms of $\mathrm{G}_n\mathrm{A}_i^\omega$ the quantifier-free rule of extensionality yields (even without appeal to the universal axioms 11)

$$\mathrm{G}_n\mathrm{A}_i^\omega \vdash \forall x_1^{\rho_1},\ldots,x_k^{\rho_k}(\widehat{t}[x_1,\ldots,x_k] =_0 tx_1\ldots x_k).$$

We now prove that $\widehat{t}[x_1,\ldots,x_k]$ does not contain any longer any combinator Π, Σ : Suppose that on the contrary $\widehat{t}[x_1,\ldots,x_k]$ contains an occurrence of Σ (resp. Π). Then Σ (Π) must occur in the form $\Sigma,\Sigma t_1$ or $\Sigma t_1 t_2$ ($\Pi,\Pi t_1$) but not in the form $\Sigma t_1 t_2 t_3$ (resp. $\Pi t_1 t_2$) since in the latter case we could have carried out the reduction $\Sigma t_1 t_2 t_3 \rightsquigarrow t_1 t_3(t_2 t_3)$ (resp. $\Pi t_1 t_2 \rightsquigarrow t_1$) contradicting the construction of $\widehat{t}$. All the terms $s = \Sigma,\Sigma t_1,\Sigma t_1 t_2,\Pi,\Pi t_1$ have a type whose degree is ≥ 1. Hence s can occur in $\widehat{t}$ only in the form $r(s)$, where $r = \Sigma,\Sigma t_4,\Sigma t_4 t_5,\Pi$ or Πt_4 since these terms are the only reduced ones requiring an argument of type ≥ 1 that can be built up from $x_1^{\rho_1},\ldots,x_k^{\rho_k},\Sigma,\Pi,A_i,S^1,0^0$ and $\max_0$ (here we use that $deg(\rho_i) \leq 1$). We notice that the cases $r = \Sigma t_4 t_5$ and $r = \Pi t_4$ cannot occur since otherwise r(s) would allow a reduction of Σ resp. Π. Hence $r(s)$ is again a Π,Σ-term having a type of degree ≥ 1 and, therefore, has to occur within a term r' for which the same reasoning as for r applies and so on. Since $\widehat{t}$ is finite this process has to stop which gives a contradiction. □

Remark 3.41. Proposition 3.40 gets false if $\deg(\rho) = 3$: Define $\rho := 0(0(000))$ and $t^\rho := \lambda x^{0(000)}.x(\Pi_{0,0})$. Then $tx =_0 x(\Pi_{0,0})$ contains Π but no Π-reduction applies.

Corollary 3.42. *Let $n \geq 1$, ρ be as in proposition 3.40 and $t^\rho \in \mathrm{G}_n\mathrm{R}^\omega$. Then one can construct (by majorization and subsequent 'logical' normalization) a term $t^*[x_1^{\rho_1},\ldots,x_k^{\rho_k}]$ such that*

1) *$t^*[x_1,\ldots,x_k]$ contains at most $x_1\ldots,x_k$ as free variables,*
2) *$t^*[x_1,\ldots,x_k]$ is built up only from $x_1,\ldots,x_k,0^0,A_0,\ldots,A_n$.*
3) *$\mathrm{G}_n\mathrm{A}_i^\omega \vdash \lambda x_1,\ldots,x_k.t^*[x_1,\ldots,x_k]$ maj t, i.e.*

$$\forall x_1^*,x_1,\ldots,x_k^*,x_k\Big(\bigwedge_{i=1}^{k}(x_i^* \; maj_{\rho_i}\, x_i) \to t^*[x_1^*,\ldots,x_k^*] \geq_0 tx_1\ldots x_k.$$

Proof: The corollary follows from propositions 3.39 (and the corollary to its proof) and 3.40 together with lemma 3.35. □

We are now in the position to estimate the growth of the functions definable by terms in $\mathrm{G}_1\mathrm{R}^\omega$, $\mathrm{G}_2\mathrm{R}^\omega$ and $\mathrm{G}_3\mathrm{R}^\omega$. Note that we only need majorization and logical normalization for this.

Proposition 3.43. *The growth of the functions defined by closed terms t^1 of $\mathrm{G}_n\mathrm{R}^\omega$ $(n = 1,2,3)$ can be calibrated as follows:*

$$\begin{cases} t^1 \in \mathrm{G_1R}^\omega \Rightarrow \exists c_1, c_2 \in \mathbb{N}: \ \mathrm{G_1A}_i^\omega \vdash \forall x^0 (tx \leq_0 c_1 x + c_2) \ \textit{(linear growth)}, \\ t^1 \in \mathrm{G_2R}^\omega \Rightarrow \exists k, c_1, c_2 \in \mathbb{N}: \ \mathrm{G_2A}_i^\omega \vdash \forall x^0 (tx \leq_0 c_1 x^k + c_2) \\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \textit{(polynomial growth)}, \\ t^1 \in \mathrm{G_3R}^\omega \Rightarrow \exists k, c \in \mathbb{N}: \ \mathrm{G_3A}_i^\omega \vdash \forall x^0 (tx \leq_0 2_k^{cx}), \ \textit{where } 2_0^a = a, 2_{k'}^a = 2^{2_k^a} \\ \qquad\qquad\qquad\qquad\qquad\qquad \textit{(finitely iterated exponential growth)}. \end{cases}$$

The result can also be extended to tuples of number variables: for t^ρ with $\rho = 0\underbrace{(0)\ldots(0)}_{m-times}$ we have

$$\begin{cases} t^\rho \in \mathrm{G_1R}^\omega \Rightarrow \exists c_1, \ldots, c_{m+1} \in \mathbb{N}: \ \mathrm{G_1A}_i^\omega \vdash \forall x_1^0, \ldots, x_m^0 (t\underline{x} \leq_0 \sum_{i=1}^{m} c_i x_i + c_{m+1}), \\ t^\rho \in \mathrm{G_2R}^\omega \Rightarrow \exists p \in \mathbb{N}[x_1, \ldots, x_m]: \ \mathrm{G_2A}_i^\omega \vdash \forall \underline{x} (t\underline{x} \leq_0 p\underline{x}), \\ t^\rho \in \mathrm{G_3R}^\omega \Rightarrow \exists k, c_1, \ldots, c_m \in \mathbb{N}: \ \mathrm{G_3A}_i^\omega \vdash \forall \underline{x} (t\underline{x} \leq_0 2_k^{c_1 x_1 + \ldots + c_m x_m}). \end{cases}$$

The constants $c_i, k \in \mathbb{N}$ as well as the coefficients of p can be effectively computed from a given closed term t by majorization and normalization.

Proof: By corollary 3.42 one can in all three cases construct a term $\widehat{t}[\underline{x}]$ built up from $\underline{x}^0, 0^0, A_0, \ldots, A_n$ $(n = 1,2,3)$ such that $\widehat{t}[\underline{x}] \geq_0 t\underline{x}$ for all $\underline{x}$. For the particular cases this yields the following:
$n = 1$: Consider a term $t^\rho \in \mathrm{G_1R}^\omega$, where $\rho = 0\underbrace{(0)\ldots(0)}_{m}$. $\widehat{t}[x_1^0, \ldots, x_m^0]$ is built up from $x_1^0, \ldots, x_m^0, 0^0, A_0$ and A_1 only. Both $A_0(x_1, x_2) = 0 \cdot x_1 + 1 \cdot x_2 + 1$ and $A_1(x_1, x_2) = 1 \cdot x_1 + 1 \cdot x_2 + 0$ are functions having the form $c_1 x_1 + c_2 x_2 + c_3$ or – more generally – $c_1 x_1 + \ldots + c_k x_k + c_{k+1}$. Since substitution of such functions again yields a function which can be written in this form it follows that $\widehat{t}[x_1, \ldots, x_m] = c_1 x_1 + \ldots + c_m x_m + c_{m+1}$ for suitable constants $c_1, \ldots, c_{m+1}$.
$n = 2$: For $t^\rho \in \mathrm{G_2R}^\omega$, $\widehat{t}[x_1, \ldots, x_m]$ is built up from $x_1^0, \ldots, x_m^0, 0^0, A_0, A_1, A_2$. Since A_0, A_1 and A_2 are polynomials (in two variables) and substitution of polynomials in several variables yields a function which can be written again as a polynomial, it is clear that $\widehat{t}[x_1, \ldots, x_m] = p(x_1, \ldots, x_m)$ for a suitable polynomial in $\mathbb{N}[x_1, \ldots, x_m]$. In the case $m = 1$, $p(x)$ can be bounded by $c_1 x^k + c_2$ for suitable numbers c_1, c_2.
$n = 3$: Let $t^\rho \in \mathrm{G_3R}^\omega$. For $\tilde{A}_3(x,y) := A_3(\max_0(x,2), \max_0(y,2))$ the following holds

$$\tilde{A}_3 \ maj \ A_0, A_1, A_2, A_3.$$

Replace in $\widehat{t}[x_1, \ldots, x_m]$ all occurrences of A_i (with $i \leq 3$) by $\tilde{A}_3$ and denote the resulting term by $\tilde{t}[x_1, \ldots, x_m]$. Then $\tilde{t}[x_1, \ldots, x_m]$ can be bounded by y_k, where $y_0 := y$, $y_{k'} := y^{y_k}$ and $y := \max(x_1, \ldots, x_m, 2)$. and hence $\forall \underline{x} (2_{\tilde{k}}^{\underline{x}} \geq t\underline{x})$ for a suitable $\tilde{k} \geq k$, where $2_0^{\underline{x}} := x_1 + \ldots + x_m$ and $2_{k'}^{\underline{x}} := 2^{2_k^{\underline{x}}}$. □

In concrete extractions of bounds from given proofs one will use, of course, all kinds of auxiliary functions to get a sharper estimate than the above perspicuous, but crude, calibration in terms of $0, S, +, \cdot, \exp$ only.

We now show how that calibration extends to functionals of type ρ with $deg(\rho) \leq 2$. To keep the notational complexity low we only formulate things for $\rho = 1(1)$. Before we can state the result we need the following

Definition 3.44. A functional $\Phi^{1(1)}$ is called linear (polynomial, elementary recursive resp.) if it can be written as a term $\tilde{t}[f,x]$ which is built up only out of $x, f, 0^0, S^1, +$ $(x, f, 0^0, S^1, +, \cdot$ resp. $x, f, 0^0, S^1, +, \cdot, (\cdot)^{(\cdot)})$.

Remark 3.45. If $\Phi^{1(1)}$ is linear (polynomial, elementary recursive) and f is a linear (polynomial, elementary recursive) function, then $\lambda x.\Phi f x$ again is a linear (polynomial, elementary recursive) function.

Proposition 3.46. *Let $t^{1(1)}$ be a closed term of $\mathrm{G_1R^\omega}$. Then there exists a linear functional Φ given by some term $\tilde{t}[f,x]$ as above such that*

$$\mathrm{G_1A}^\omega_i \vdash \forall x^0, f^1 \left(\tilde{t}[f^M, x] \geq t f x\right),$$

where $f^M := \Phi_1 f$, i.e. $f^M(x) := \max\{f(0), \ldots, f(x)\}$.
Moreover, f^M can be replaced by h^M for any h satisfying $h \geq_1 f$.
Analogously for $t^{1(1)} \in \mathrm{G_2R^\omega}$ (resp. $t^{1(1)} \in \mathrm{G_3R^\omega}$), where then Φ is a polynomial (resp. elementary recursive) functional and $\mathrm{G_1A}^\omega_i$ is replaced by $\mathrm{G_2A}^\omega_i$ (resp. $\mathrm{G_3A}^\omega_i$).

Proof: The proposition follows from corollary 3.42 and the fact that

$$\mathrm{G_1A}^\omega_i \vdash h \geq_1 f \to h^M \, maj_1 \, f.$$

□

So, in particular, if $t^{1(1)} \in \mathrm{G_2R^\omega}$ and f is bounded by a polynomial $p \in \mathbb{N}[x]$, then tf again is bounded by some polynomial $r \in \mathbb{N}[x]$ (note that p always is monotone so that $p^M = p$). In fact, this holds even in a certain uniform sense:

Proposition 3.47. *Let $t^{1(1)} \in \mathrm{G_2R^\omega}$. Then one can construct a polynomial $q \in \mathbb{N}[x]$ such that*

$$\begin{cases} \textit{For every polynomial } p \in \mathbb{N}[x] \\ \textit{one can construct a polynomial } r \in \mathbb{N}[x] \textit{ such that} \\ \forall f^1 \left(f \leq_1 p \to \forall x^0 (t f x \leq_0 r(x))\right) \textit{ and } deg(r) \leq q(deg(p)) \end{cases}$$

The result also holds in the case where t has tuples $f_1^1, \ldots, f_k^1, x_1^0, \ldots, x_l^0$ of arguments with $f_1, \ldots, f_k \leq_1 p$ and $r \in \mathbb{N}[x_1, \ldots, x_l]$.

Proof: Let $p \in \mathbb{N}[x]$ and $\tilde{t}[f,x]$ be constructed to tf according to proposition 3.46. Then $\tilde{t}[p,x] \geq_0 t f x$ for all $f \leq_1 p$ and $\tilde{t}[p,x]$ is built up from $x, 0^0, A_0, A_1$ and p only.

As in the proof of proposition 3.43 one concludes that $\tilde{t}[p,x]$ can be written as a polynomial r in x. The existence of the polynomial q bounding the degree of r in the degree of p follows from the fact that the degree of a polynomial $p_1 \in \mathbb{N}[x_1,\ldots,x_m]$ obtained by substitution of a polynomial p_2 for one variable in a polynomial p_3 is $\leq \deg(p_2)\cdot\deg(p_3)$ and that $\deg(p_2+p_3)$, $\deg(p_2\cdot p_3) \leq \deg(p_2)+\deg(p_3)$. □

The analysis of the growth of the functionals definable in $\mathrm{G}_n\mathrm{A}^\omega$ made use for the first time of the important notion of majorizability from [163]. In connection with the definition of a model for E-PA$^\omega$ (and its extension by bar recursion, see chapter 11) a variant – called 'strong majorization' and denoted by 's-maj' – was introduced semantically in [27] (see the next section). This notion has the following syntactic counterpart:

Definition 3.48 (Bezem [27]). We define the relation x^* *s-maj*$_\rho$ x (x^* strongly majorizes x) between functionals of type ρ by induction on ρ:

$$\begin{cases} x^* \ s\text{-}maj_0\ x :\equiv x^* \geq_0 x, \\ x^* \ s\text{-}maj_{\tau\rho}\ x :\equiv \forall y^*, y(y^* \ s\text{-}maj_\rho\ y \to x^*y^* \ s\text{-}maj_\tau\ x^*y, xy). \end{cases}$$

The properties basic properties of *s-maj* are

Lemma 3.49. $\mathrm{G}_1\mathrm{A}^\omega_i$ *(as well as* $\widehat{\text{WE-HA}}^\omega{\restriction}$*) proves:*

(i) $\tilde{x}^* =_\rho x^* \wedge \tilde{x} =_\rho x \wedge x^*\ s\text{-}maj_\rho\ x \to \tilde{x}^*\ s\text{-}maj_\rho\ \tilde{x}$.
(ii) $x^*\ s\text{-}maj_\rho\ x \to x^*\ s\text{-}maj_\rho\ x^*$.
(iii) $x_1\ s\text{-}maj_\rho\ x_2 \wedge x_2\ s\text{-}maj_\rho\ x_3 \to x_1\ s\text{-}maj_\rho\ x_3$.
(iv) $x^*\ s\text{-}maj_\rho\ x \wedge x \geq_\rho y \to x^*\ s\text{-}maj_\rho\ y$.
(v) *For* $\rho = \tau\rho_k\ldots\rho_1$:

$$x^*\ s\text{-}maj_\rho\ x \leftrightarrow \forall y_1^*, y_1, \ldots, y_k^*, y_k \Big(\bigwedge_{i=1}^{k} (y_i^*\ s\text{-}maj_{\rho_i}\ y_i) \to x^*\underline{y}^*\ s\text{-}maj_\tau\ x^*\underline{y}, x\underline{y} \Big).$$

Proof: (i)–(iv) follow by induction on ρ (where we use (ii) in the proof of (iii)). (v) follows by induction on k using again (ii). □

Analogously to proposition 3.39 one proves:

Proposition 3.50. *For all* $n \geq 1$ *the following holds: To each term* $t^\rho \in \mathrm{G}_n\mathrm{R}^\omega$ *one can construct by induction on the structure of t (without normalization) a term* $t^{*\rho} \in \mathrm{G}_n\mathrm{R}^\omega_-$ *such that*

$$\mathrm{G}_n\mathrm{A}^\omega_i \vdash t^*\ s\text{-}maj_\rho\ t.$$

Proof: The same term as constructed in the proof of proposition 3.39 also works for *s-maj* (exercise). □

3.6 Models of E-PA$^\omega$

The full set-theoretic model:

We define the type-structure $\mathscr{S}^\omega$ of **all** set-theoretic functionals as follows:

$$\begin{cases} S_0 := \mathbb{N} \\ S_{\tau\rho} := \{ \text{all set-theoretic functionals } \varphi : S_\rho \to S_\tau \} \\ \mathscr{S}^\omega := \langle S_\rho \rangle_{\rho \in \mathbf{T}}. \end{cases}$$

The following proposition is obvious:

Proposition 3.51. *$\mathscr{S}^\omega$ is a model of* E-PA$^\omega$.

The model of all sequentially continuous functionals: The type-structure $\mathscr{C}^\omega$ of all (sequentially) continuous functionals was introduced in [321]. We first need some preparatory definitions:

Definition 3.52 (Kuratowski [258]). Let X be a set together with a relation of convergence '$\to$' between sequences (p_n) of X and elements $p \in X$. As usual we write '$p_n \to p$' instead of '$\to ((p_n), p)$'.
$(X, \to)$ is called a 'limit space' (short: 'L-space') if the following axioms are satisfied:

1) $p_n \to p$ implies that for every subsequence $(p_{k_n}) \quad (k_1 < k_2 < \ldots)$ of (p_n) also $p_{k_n} \to p$;
2) if $p_n = p$ for almost all $n \in \mathbb{N}$, then $p_n \to p$;
3) if not $p_n \to p$, then there exists a sequence $k_1 < k_2 < \ldots$ such that no subsequence of (p_{k_n}) converges to p;
4) if $p_n \to p$ and $p_n \to q$, then $p = q$.

Definition 3.53. Let $(X, \to_X)$ and $(Y, \to_Y)$ be two L-spaces. A function $f : X \to Y$ is called continuous if $f(p_n) \to_Y f(p)$ whenever $p_n \to_X p$. The set of all continuous functions from X to Y is denoted by $\mathscr{C}(X,Y)$.

On $\mathscr{C}(X,Y)$ one can define the following relation of convergence

$$f_n \to f :\equiv \forall (p_n), p \big(p_n \to_X p \Rightarrow f_n(p_n) \to_Y f(p) \big).$$

Lemma 3.54. $(\mathscr{C}(X,Y), \to)$ *is again an* L-*space.*

Proof: Exercise! □

Definition 3.55 (Scarpellini [321]). The type-structure of sequentially continuous functionals is defined as follows:

$$\begin{cases} C_0 := \mathbb{N}, \ \ p_n \to_0 p :\equiv \exists k \forall m > k (p_m = p); \\ C_{\tau\rho} := \mathscr{C}(C_\rho, C_\tau), \\ f_n \to_{\tau\rho} f :\equiv \forall (p_n) \in C_\rho^{\mathbb{N}}, p \in C_\rho \big(p_n \to_\rho p \Rightarrow f_n(p_n) \to_\tau f(p) \big). \\ \mathscr{C}^\omega := \langle C_\rho \rangle_{\rho \in \mathbf{T}}. \end{cases}$$

Remark 3.56. Since $(C_0, \to_0)$ clearly is an L-space, lemma 3.54 implies that $(C_\rho, \to_\rho)$ is an L-space for any $\rho \in \mathbf{T}$.

Proposition 3.57 (Scarpellini [321]). *$\mathscr{C}^\omega$ is a model of* E-PA$^\omega$.

Proof: For $0^0, \Pi, \Sigma, S^1$ this is rather straightforward. For the recursors one shows by induction on n that the functionals $\underline{R}n$ belong to $\mathscr{C}$ for all n. It then follows immediately that also the recursors $\underline{R}$ themselves belong to $\mathscr{C}$. For details see [321]. □

The model of all extensional hereditarily continuous functionals: A different type structure of continuous functionals, the so-called extensional hereditarily continuous functionals ECF$^\omega$, is based on a notion of ‘continuous functional’ due to [196] and [244]. As proved in [169], ECF$^\omega$ is in fact isomorphic to $\mathscr{C}^\omega$. The main proof-theoretic use of ECF$^\omega$ is due to the fact that the functionals in ECF$^\omega$ are represented by number theoretic functions $\alpha \in \mathbb{N}^\mathbb{N}$, their so-called associates, for which an equivalence relation $=_\rho$ is defined for each type so that the equivalence classes correspond to the functionals being represented. This makes it possible for formalize certain semantic arguments in systems with number and function quantifiers such as the so-called elementary intuitionistic analysis EL of Kreisel and Troelstra (see [366], which in turn is based on [197], for a comprehensive treatment of all this). We sketch here only the definition of ECF$^\omega$: The definition is based on two versions of so-called partial continuous function application (in the following $\alpha, \beta, \gamma, \ldots$ range over unary number theoretic functions and $x, y, z, \ldots$ over natural numbers):

Definition 3.58. We define $\alpha|\beta$ and $\alpha(\beta)$ by

$$(\alpha|\beta)(x) \simeq y :\equiv \alpha(\langle x\rangle * \overline{\beta}(\mu z[\alpha(\langle x\rangle * \overline{\beta}z) \neq 0])) \dot{-} 1 \simeq y,$$
$$\alpha(\beta) \simeq y :\equiv \alpha(\overline{\beta}(\mu z[\alpha(\overline{\beta}z) \neq 0])) \dot{-} 1 \simeq y.$$

$(\alpha|\beta)(x)$ and $\alpha(\beta)$ are partial recursive functionals in x, α, β resp. α, β. Thus, using some basic notation from ordinary recursion theory, there are codes $n_0, n_1 \in \mathbb{N}$ for corresponding oracle Turing machines such that

$$\{n_0\}(x, \alpha, \beta) \simeq y \leftrightarrow (\alpha|\beta)(x) \simeq y,$$
$$\{n_1\}(\alpha, \beta) \simeq y \leftrightarrow \alpha(\beta) \simeq y.$$

$\alpha|\beta \simeq \gamma$ is defined as $\forall x \in \mathbb{N}((\alpha|\beta)(x) \simeq \gamma(x))$.

Definition 3.59 (Kleene [195], Kreisel [244], Troelstra [366]). The type structure ECF$^\omega$ of all extensional hereditarily continuous functionals of finite type is defined by simultaneously declaring ECF_ρ and an extensional equality $=_\rho$:

$$\mathrm{ECF}_0 := \mathbb{N},\ \ x =_0 y :\equiv x,y \in \mathbb{N} \wedge x = y;$$

$$\mathrm{ECF}_{0(0)} := \mathbb{N}^{\mathbb{N}},\ \ \alpha =_1 \beta :\equiv \alpha,\beta \in \mathbb{N}^{\mathbb{N}} \wedge \forall x \in \mathbb{N}(\alpha(x) = \beta(x));$$

$$\mathrm{ECF}_{0(\rho)} :=$$

$$\{\alpha \in \mathbb{N}^{\mathbb{N}} : \forall \beta \in \mathrm{ECF}_\rho \exists x \in \mathbb{N}(\alpha(\beta) \simeq x) \wedge \forall \beta_1,\beta_2(\beta_1 =_\rho \beta_2 \to \alpha(\beta_1) \simeq \alpha(\beta_2)\},$$

$$\alpha_1 =_{0(\rho)} \alpha_2 :\equiv \alpha_1,\alpha_2 \in \mathrm{ECF}_{0(\rho)} \wedge \forall \beta \in \mathrm{ECF}_\rho(\alpha_1(\beta) \simeq \alpha_2(\beta)) \text{ for } \rho \neq 0;$$

$$\mathrm{ECF}_{\tau(0)} := \{\alpha \in \mathbb{N}^{\mathbb{N}} : \forall x \in \mathbb{N} \exists \gamma \in \mathrm{ECF}_\tau(\alpha|\lambda y.x \simeq \gamma)\},$$

$$\alpha_1 =_{\tau(0)} \alpha_2 :\equiv \alpha_1,\alpha_2 \in \mathrm{ECF}_{\tau(0)} \wedge \forall x \in \mathbb{N}(\alpha_1|\lambda y.x =_\tau \alpha_2|\lambda y.x) \text{ for } \tau \neq 0;$$

$$\mathrm{ECF}_{\tau(\rho)} :=$$

$$\{\alpha \in \mathbb{N}^{\mathbb{N}}{:}\forall \beta \in \mathrm{ECF}_\rho \exists \gamma \in \mathrm{ECF}_\tau(\alpha|\beta \simeq \gamma) \wedge \forall \beta_1,\beta_2(\beta_1 =_\rho \beta_2 \to \alpha|\beta_1 =_\tau \alpha|\beta_2)\},$$

$$\alpha_1 =_{\tau(\rho)} \alpha_2 :\equiv \alpha_1,\alpha_2 \in \mathrm{ECF}_{\tau(\rho)} \wedge \forall \beta \in \mathrm{ECF}_\rho(\alpha_1|\beta =_\tau \alpha_2|\beta) \text{ for } \rho,\tau \neq 0.$$

$\mathrm{ECF}^\omega = \langle \mathrm{ECF}_\rho \rangle_{\rho \in \mathbf{T}}$.

So in ECF^ω, the application operation $App_{\tau,\rho}$ between (representatives of) functionals in $\mathrm{ECF}_{\tau(\rho)}$ and ECF_ρ is not the set-theoretic application but defined depending on ρ,τ as follows

$$App_{0,0}(\alpha,x) := \alpha(x),$$

$$App_{0,\rho}(\alpha,\beta) := \alpha(\beta) \text{ for } \rho \neq 0,$$

$$App_{\tau,0}(\alpha,x) := \alpha|\lambda y.x \text{ for } \tau \neq 0,$$

$$App_{\tau,\rho}(\alpha,\beta) := \alpha|\beta \text{ for } \rho,\tau \neq 0.$$

The following proposition is shown in [366]:

Proposition 3.60. ECF^ω *is a model of* E-PA$^\omega$. *Moreover, as interpretations* $[t]_{\mathrm{ECF}^\omega}$ *for the closed terms* t^ρ $(\rho \neq 0)$ *of* E-PA$^\omega$ *one can take computable functions* $\alpha \in \mathrm{ECF}_\rho \cap REC$.

The model of all strongly majorizable functionals: The following type structure $\mathscr{M}^\omega$ of all strongly majorizable functionals was constructed in [27] making use of the variant *s-maj* of Howard's majorization relation *maj*.

Definition 3.61 (Bezem [27]). The type structure $\mathscr{M}^\omega$ of all hereditarily strongly majorizable set-theoretic functionals of finite type is defined as

$$\begin{cases} M_0 := \mathbb{N},\ \ n\ s\text{-}maj_0\ m :\equiv n \geq m \wedge n,m \in \mathbb{N}, \\ x^*\ s\text{-}maj_{\tau(\rho)}\ x :\equiv x^*,x \in M_\tau^{M_\rho} \wedge \forall y^*,y \in M_\rho(y^*\ s\text{-}maj_\rho\ y \to x^*y^*\ s\text{-}maj_\tau\ x^*y,xy), \\ M_{\tau(\rho)} := \left\{x \in M_\tau^{M_\rho}\ :\ \exists x^* \in M_\tau^{M_\rho}(x^*\ s\text{-}maj_{\tau(\rho)}\ x)\right\} \quad (\rho,\tau \in \mathbf{T}) \end{cases}$$

(Here $M_\tau^{M_\rho}$ denotes the set of all total set-theoretic mappings from M_ρ into M_τ).
$\mathscr{M}^\omega := \langle M_\rho \rangle_{\rho \in \mathbf{T}}$.

Remark 3.62. An easy induction on ρ shows that for $x^*, x \in M_\rho$ the relation *s-maj*$_\rho$ as defined in definition 3.61 coincides with the interpretation of the corresponding syntactic relation from definition 3.48 in the model $\mathscr{M}^\omega$, i.e.

$$\forall x^*, x \in M_\rho \, (x^* \, s\text{-}maj_\rho \, x \leftrightarrow x^* [s\text{-}maj_\rho]_{\mathscr{M}^\omega} x).$$

The relation defined in definition 3.61, however, also applies to functionals which prima facie are only in say $M_\tau^{M_\rho}$ and is used to determine whether such a functional actual is in $M_{\tau(\rho)}$. In the following lemmas we always refer to the relation from definition 3.61.

Lemma 3.63. 1) $x^* \, s\text{-}maj_\rho \, x \to \; x \in M_\rho \wedge x^* \, s\text{-}maj_\rho \, x^* \to x^*, x \in M_\rho$.
2) *For all* $\rho = \tau\rho_k \ldots \rho_1$ $(k \geq 1)$ *and all* $x^*, x : M_{\rho_1} \to (M_{\rho_2} \to \ldots \to M_\tau) \ldots)$ *the following holds*

$$x^* \, s\text{-}maj_\rho \, x \leftrightarrow$$
$$\forall y_1^*, y_1, \ldots, y_k^*, y_k \Big(\bigwedge_{i=1}^{k} y_i^* \, s\text{-}maj_{\rho_i} \, y_i \to x^* y_1^* \ldots y_k^* \, s\text{-}maj_\tau \, x^* y_1 \ldots y_k, x y_1 \ldots y_k \Big).$$

Proof: 1) is proved by induction on ρ. 2) follows by induction on k using 1). □

Remark 3.64. The implication '←' in the second claim of this lemma is used often to establish that a functional $x : M_{\rho_1} \to (M_{\rho_2} \to \ldots \to M_\tau) \ldots)$ actually is in $M_{\tau\rho_k \ldots \rho_1}$.

Definition 3.65. Let $x \in M_\rho^{M_0}$, where $\rho = 0\rho_k \ldots \rho_1$. Then we define

$$x^M(n) := \lambda \underline{v}. \max\{xi\underline{v} \; : \; i \leq n\},$$

where $\underline{v} = v_1^{\rho_1}, \ldots, v_k^{\rho_k}$.

Lemma 3.66. *Let* $x, \widehat{x} \in M_\rho^{M_0}$ *be such that*

$$\forall n \in \mathbb{N} \, (\widehat{x} n \; s\text{-}maj_\rho \; xn).$$

Then

$$\widehat{x}^M \; s\text{-}maj_{\rho 0} \; x^M, x$$

and hence $\widehat{x}^M, x^M, x \in M_{\rho 0}$.
In particular, this implies that $(\cdot)^M \; s\text{-}maj_{\rho 0(\rho 0)} \; (\cdot)^M \in M_{\rho 0(\rho 0)}$.
As a special case it follows that

$$x_1^* \; s\text{-}maj_\rho \; x_1 \wedge x_2^* \; s\text{-}maj_\rho \; x_2 \to \max\nolimits_\rho(x_1^*, x_2^*) \; s\text{-}maj_\rho \; \max\nolimits_\rho(x_1, x_2), x_1, x_2,$$

where $\max_\rho(x_1, x_2) := \lambda \underline{v}. \max_0(x_1\underline{v}, x_2\underline{v})$.

Proof: Let $\rho = 0\rho_k \dots \rho_1$ and $y_1^*, y_1, \dots, y_k^*, y_k$ be such that $\bigwedge_{i=1}^{k} (y_i^* \text{ s-maj}_{\rho_i}\ y_i)$. One easily shows by induction on n (using lemma 3.63.1) that

$$\forall n \forall m \leq n (\widehat{x}^M n \underline{y}^* \geq_0 \widehat{x}^M m \underline{y}, x^M m \underline{y}, x m \underline{y})$$

which by lemma 3.63.2) yields the claim. □

Remark 3.67. Note that $(\cdot)^M$ is definable as a closed term of type $\rho 0(\rho 0)$ already in $\mathrm{G_1A}^\omega$.

Corollary 3.68. $M_\rho^{M_0} = M_{\rho 0}$ *for each* $\rho \in \mathbf{T}$.

Proof: $x \in M_\rho^{M_0}$ implies that for each $n \in \mathbb{N}$ there exists an $x_n^* \in M_\rho$ with

$$x_n^* \ s\text{-}maj_\rho \ x(n).$$

Using $\mathrm{AC}^{0,\rho}$ on the meta-level we obtain a sequence $x^* \in M_\rho^{M_0}$ with

$$\forall n \in \mathbb{N} \ (x^* n \ s\text{-}maj_\rho \ xn).$$

Lemma 3.66 now yields that $(x^*)^M \ s\text{-}maj_{\rho 0} \ x \in M_{\rho 0}$. □

Proposition 3.69 (Bezem [27]). $\mathscr{M}^\omega$ *is a model of* E-PA$^\omega$.

Proof: Using lemma 3.63.2) it is immediate that $0, S, \Pi, \Sigma$ all majorize themselves and hence belong to $\mathscr{M}^\omega$. For the recursors $\underline{R}_{\underline{\rho}}$ one shows by induction on n that for all $\underline{y}^*, \underline{y} \in M_{\underline{\rho}}$ with $\underline{y}^* \ s\text{-}maj_{\underline{\rho}} \ \underline{y}$, i.e. $\bigwedge_{i=1}^{k} (y_i^* \ s\text{-}maj_{\rho_i} \ y_i)$ and $\underline{z}^*, \underline{z} \in M_{\underline{\rho} 0 \underline{\rho}}$ with $\underline{z}^* \ s\text{-}maj_{\underline{\rho} 0 \underline{\rho}} \ \underline{z}$ we have

$$(*) \ \forall n \in \mathbb{N} \ (\underline{R}_{\underline{\rho}} n \underline{y}^* \underline{z}^* \ s\text{-}maj_{\underline{\rho}} \ \underline{R}_{\underline{\rho}} n \underline{y} \underline{z}).$$

Let $n = 0$: Then $\underline{R}_{\underline{\rho}} 0 \underline{y}^* \underline{z}^* = \underline{y}^* \ s\text{-}maj_{\underline{\rho}} \ \underline{y} = \underline{R}_{\underline{\rho}} 0 \underline{y} \underline{z}$.

$n \mapsto n+1$: $\underline{R}_{\underline{\rho}}(n+1)\underline{y}^* \underline{z}^* = \underline{z}^*(\underline{R}_{\underline{\rho}} n \underline{y}^* \underline{z}^*) n \overset{3.63, I.H.}{s\text{-}maj_{\underline{\rho}}} \underline{z}(\underline{R}_{\underline{\rho}} n \underline{y} \underline{z}) n = \underline{R}_{\underline{\rho}}(n+1)\underline{y}\underline{z}$. This finishes the proof of $(*)$. Lemma 3.63.2 yields that

$$\forall n \in \mathbb{N} (\underline{R}_{\underline{\rho}} n \ s\text{-}maj \ \underline{R}_{\underline{\rho}} n).$$

By lemma 3.66 it now follows that

$$(R_i)^*_{\underline{\rho}} := (R_i)^M_{\underline{\rho}} \ s\text{-}maj \ (R_i)_{\underline{\rho}} \in M_{\rho_i(\underline{\rho} 0 \underline{\rho}) \underline{\rho} 0}.$$

□.

The next proposition shows that the three models $\mathscr{S}^\omega, \mathscr{C}^\omega$ and $\mathscr{M}^\omega$ start to differ from type 2 on:

Proposition 3.70. 1) $S_1 = C_1 = \mathrm{ECF}_1 = M_1 = \mathbb{N}^{\mathbb{N}}$.
2) *$C_2 \subset M_2 \subset S_2$, where both '$\subset$' are strict.*

Proof: 1) By definition we have that $S_0 = C_0 = \mathrm{ECF}_0 = M_0 = \mathbb{N}$ and $\mathrm{ECF}_1 = \mathbb{N}^{\mathbb{N}}$. It is trivial to observe that every $f \in \mathbb{N}^{\mathbb{N}}$ is continuous in the sense of $\mathscr{C}^\omega$. Since $\forall n \in \mathbb{N}\ (f(n) \geq f(n))$ we get by lemma 3.66 that f^M *s-maj*$_1$ $f \in M_1$.

2) We first show the inclusions. Since $M_2 \subseteq M_0^{M_1} \overset{1)}{=} S_0^{S_1} = S_2$ the second inclusion is trivial. Now let $x \in C_2$. Then (by 1) $x \in M_0^{M_1}$. We have to construct a majorant $x^* \in M_0^{M_1}$: $x \in C_2$ implies that x is a continuous function $\mathbb{N}^{\mathbb{N}} \to \mathbb{N}$ between the Baire space $\mathbb{N}^{\mathbb{N}}$ (with its usual metric) and $\mathbb{N}$ with the discrete metric. Now let $f \in \mathbb{N}^{\mathbb{N}}$. Then $K_f := \{g \in \mathbb{N}^{\mathbb{N}} : \forall n \in \mathbb{N}\ (g(n) \leq f(n))\}$ is a compact subspace of the Baire space. Hence x is uniformly continuous on K_f and so the following definition is well-defined: $x^*(f) := \max\{x(g) : g \in K_f\}$. We claim that x^* *s-maj*$_2$ x. Let f *s-maj*$_1$ g. Then $f(n) \geq g(n)$ for all $n \in \mathbb{N}$. Hence $g \in K_f$ and so $x^*(f) \geq x(g)$. Moreover, $x^*(f) \geq x^*(g)$ since $K_g \subseteq K_f$.
To show that $S_2 \not\subseteq M_2$ consider $x \in S_2$ defined as follows

$$x(f) := \begin{cases} n, & \text{for the least } n \text{ such that } f(n) = 0 \text{ if existent} \\ 0, & \text{otherwise.} \end{cases}$$

Suppose x^* *s-maj*$_2$ x for some $x^* \in S_2$. Consider the function $1_n(k) := 1$ for $k < n$ and $1_n(k) := 0$ for $k \geq n$ and let 1^1 be the constant-1 function. Then 1 *s-maj*$_1$ 1_n and hence $x^*(1^1) \geq x(1_n)$ for all $n \in \mathbb{N}$, but $x(1_n) = n$ which is a contradiction.
To show that $M_2 \not\subseteq C_2$ consider $x \in M_2$ defined as follows

$$x(f) := \begin{cases} 1, & \text{if } \exists n \in \mathbb{N}(f(n) = 0) \\ 0, & \text{otherwise.} \end{cases}$$

Clearly, $x \in M_2$ since x is majorized by the constant-1 functional. But also clearly $x \notin C_2$. □

Let $\mathscr{T}^\omega$ be a model of E-PA$^\omega$ and t a closed term of E-PA$^\omega$. Then we use $[t]_{\mathscr{T}^\omega}$ to denote the interpretation of t in $\mathscr{T}^\omega$. If t has a type of degree ≤ 2, then – intuitively – it is clear (using proposition 3.70.1) that $[t]_{\mathscr{S}^\omega} = [t]_{\mathscr{C}^\omega} = [t]_{\mathscr{M}^\omega}$. The following proposition confirms this:

Proposition 3.71. *Let t^2 be a closed term of* E-PA$^\omega$ *of type* 2
Then $[t]_{\mathscr{S}^\omega}(f) = [t]_{\mathrm{ECF}^\omega}(f) = [t]_{\mathscr{C}^\omega}(f) = [t]_{\mathscr{M}^\omega}(f)$ for all $f \in \mathbb{N}^{\mathbb{N}}$, where in the case of ECF$^\omega$ *'$\cdot(\cdot)$' refers to the partial continuous function application.*
This result also holds for closed terms t^ρ whose type ρ is of degree ≤ 2.

Proof: We give the details for $[t]_{\mathscr{S}^\omega}(f) = [t]_{\mathscr{C}^\omega}(f)$ (the claims $[t]_{\mathscr{S}^\omega}(f) = [t]_{\mathscr{M}^\omega}(f)$ and $[t]_{\mathscr{S}^\omega}(f) = [t]_{\mathrm{ECF}^\omega}(f)$ are proved analogously). We define a so-called logical relation (first studied systematically by G.Plotkin) $\approx_\rho$ by induction on ρ as follows:

$$x \approx_0 y :\equiv x, y \in S_0 = C_0 \wedge x = y,$$
$$x \approx_{\tau\rho} y :\equiv x \in S_{\tau\rho} \wedge y \in C_{\tau\rho} \wedge \forall u \in S_\rho, v \in C_\rho (u \approx_\rho v \to xu \approx_\tau yv).$$

One easily shows for all constants c^ρ of E-PA$^\omega$ that

$$[c]_{\mathscr{S}^\omega} \approx_\rho [c]_{\mathscr{C}^\omega}.$$

This immediately implies that for all closed terms t

$$[t]_{\mathscr{S}^\omega} \approx_\rho [t]_{\mathscr{C}^\omega}.$$

Now let $deg(\rho) \le 2$. It is clear that it suffices to consider the case $\rho = 2$. By proposition 3.70.1) we have that

$$x \approx_1 y \leftrightarrow x, y \in S_1 = M_1 = \mathbb{N}^\mathbb{N} \wedge \forall n \in \mathbb{N} (x(n) = y(n)).$$

In particular, $x \approx_1 x$ for all $x \in \mathbb{N}^\mathbb{N}$. Hence $[t]_{\mathscr{S}^\omega} \approx_2 [t]_{\mathscr{C}^\omega}$ implies that

$$\forall x \in \mathbb{N}^\mathbb{N} ([t]_{\mathscr{S}^\omega}(x) = [t]_{\mathscr{C}^\omega}(x)).$$

□

Remark 3.72. For $\mathscr{C}^\omega, \mathscr{M}^\omega, \mathscr{S}^\omega$, where we still have at least inclusions for the type 2, the proof above can even be used to show that for closed terms t^3 of E-PA$^\omega$ of type 3 we have

$$\forall x \in C_2 \subset M_2 \subset S_2 \ ([t]_{\mathscr{S}^\omega}(x) = [t]_{\mathscr{C}^\omega}(x)) = [t]_{\mathscr{M}^\omega}(x))$$

and

$$\forall x \in M_2 \subset S_2 \ ([t]_{\mathscr{S}^\omega}(x) = [t]_{\mathscr{M}^\omega}(x)).$$

Proof: Exercise!

3.7 Exercises, historical comments and suggested further reading

Exercises:

1) Convince yourself that the axioms and rules of $\mathrm{IL}_{-=}$ are sound under the BHK interpretation.
2) Convince yourself that $\neg\neg A \to A$ in general is not valid under the BHK interpretation.
3) Show that IA is derivable from IR. Compare the complexity of the induction formula $\tilde{A}(x)$ of the IR-instance needed to prove an IA-instance with induction formula $A(x)$ with that of $A(x)$.
4) ([161]) It is known that the function $\alpha : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$ defined by the equations

$$(*)\begin{cases}\alpha(0,y)=y'\\ \alpha(x',0)=\alpha(x,1)\\ \alpha(x',y')=\alpha(x,\alpha(x',y))\end{cases}$$

is not primitive recursive (in the sense of Kleene). In fact, α is a variant due to R. Peter of the well-known Ackermann function.

Show that α is definable in WE-HA$^\omega$ by a closed term $t^{0(0)(0)}$ using R_1 (i.e. WE-HA$^\omega$ proves the equations $(*)$ for t).

5) Prove the statements in remark 3.11.2) and 3.11.3).
6) Prove the statement in remark 3.14.
7) Prove proposition 3.22.
8) The so-called bounded collection principle CP is given by

$$\text{CP: } \forall a^0(\forall x <_0 a\exists y^0 A(x,y) \to \exists z^0\forall x <_0 a\exists y <_0 zA(x,y)),$$

where A is an arbitrary formula with arbitrary parameters allowed (including a) but with z not free in A.

a. Prove WE-HA$^\omega \vdash$ CP.
b. Proved $\widehat{\text{WE-HA}}^\omega\!\upharpoonright$+QF-AC$^{0,0} \vdash \Sigma^0_1$-CP, where Σ^0_1-CP is the restriction of CP to Σ^0_1-formulas (with arbitrary parameters of arbitrary types).
Show the same result for $G_nA^\omega_i$ instead of $\widehat{\text{WE-HA}}^\omega\!\upharpoonright$.

9) Prove the claim at the end of section 3.4.
10) Show that WE-PA$^\omega$+QF-AC^{0,1} proves the following: if Φ^2 restricted to $\{f : f \le_1 h\}$ is uniformly continuous (for any h), then Φ is pointwise continuous everywhere, i.e.

$$\forall\Phi^2\Big(\forall h^1\exists n^0\forall f,g \le_1 h(\overline{f}n=\overline{g}n \to \Phi f=\Phi g) \to \\ \forall f^1\exists n^0\forall g^1(\overline{f}n=\overline{g}n \to \Phi f=\Phi g)\Big).$$

Analogously for $\Phi^{1(1)}$.

11) Show that G_3A^ω+(IPP) proves the following finite form of Σ^0_1-comprehension

$$\forall n\exists j\forall i<n\big((j)_i=0 \leftrightarrow \exists yA_0(i,y)\big),$$

where A_0 is quantifier-free and may contain arbitrary parameters. Hint: Let $\nu(x)$ be the code of the binary sequence of length n satisfying

$$(\nu(x))_i=0 \leftrightarrow \exists y \le xA_0(i,y)$$

for all $i<n$. Now define the coloring $f(x):=\nu(x)$ of $\mathbb{N}$ by $\le \overline{1}n$-colors.

12) Use the previous exercise to show that

$$G_3A^\omega\text{+(IPP)} \vdash \Sigma^0_1\text{-IA}.$$

13) Prove lemma 3.37 (Hint: see [207] where this lemma is proved for s-maj instead of maj. With minor modifications the proof applies to maj as well).
14) Prove proposition 3.50.
15) Prove lemma 3.54.
16) Fill in the details of the proof of proposition 3.57.
17) Define the functional

$$\Phi(z^2,x^1,y^1) := \min n^0[x(n) =_0 y(n) \to zx =_0 zy].$$

Show that $\Phi \in \mathscr{C}^\omega$ and $\Phi \in \mathscr{S}^\omega$ but $\Phi \notin \mathscr{M}^\omega$ ([163]).
18) Show that in ECF^ω the so-called fan functional

$$\Phi_{\mathrm{FAN}}(x^2) := \min n\ [\forall y_1,y_2 \in B(\forall i < n(y_1 i = y_2 i) \to xy_1 = xy_2)],\ \text{where}$$

$B := \{x \in \mathbb{N}^{\mathbb{N}} : \forall n \in \mathbb{N}(xn \leq 1)\}$, exists and even has a recursive associate α (such functionals are called recursively countable; as shown by Tait (unpublished) and Gandy/Hyland [111] Φ_{FAN} is not Kleene computable, in the sense of his schemata S1-S9 [195], over ECF^ω).
19) For x^1 define

$$(\overline{x,n^0})(k^0) := \begin{cases} xk, & \text{if } k < n \\ 0^0, & \text{otherwise} \end{cases}$$

and

$$\mu_1(x^2,y^1) := \begin{cases} \min n^0[x(\overline{y,n}) =_0 xy], & \text{if } \exists n(x(\overline{y,n}) =_0 xy) \\ \text{undefined}, & \text{otherwise.} \end{cases}$$

Show that $\mu_1 \in \mathscr{C}^\omega$, but $\mu_1 \notin \mathscr{M}^\omega$ and $\mu_1 \notin \mathscr{S}^\omega$.
20) Define

$$\mu_1^t(x^2,y^1) := \begin{cases} \min n^0[x(\overline{y,n}) =_0 xy], & \text{if } \exists n(x(\overline{y,n}) =_0 xy) \\ 0^0 \text{ otherwise.} \end{cases}$$

Show that $\mu_1 \in \mathscr{S}^\omega$ and $\mu_1 \in \mathscr{C}^\omega$, but $\mu_1 \notin \mathscr{M}^\omega$.
21) Define

$$\mu_2(x^2,y^1) := \begin{cases} \min n^0[x(\overline{y,n}) < n], & \text{if } \exists n(x(\overline{y,n}) < n) \\ \text{undefined}, & \text{otherwise.} \end{cases}$$

Show that $\mu_2 \in \mathscr{M}^\omega$ and $\mu_2 \in \mathscr{C}^\omega$, but $\mu_2 \notin \mathscr{S}^\omega$
22) Show that $\mathscr{C}^\omega \models \text{QF-AC}^{1,0}$, where

$$\text{QF-AC}^{1,0} : \forall f^1 \exists x^0 A_0(f,x) \to \exists F^2 \forall f^1 A_0(f,F(f))$$

with A_0 quantifier-free.
23) Show that $\mathscr{M}^\omega \not\models \text{QF-AC}^{1,0}$ ([201]).

24) Show the following 'weak continuity' property of $\mathscr{M}^\omega$ ([22]):

$$\mathscr{M}^\omega \models \forall y^{0(\rho 0)}, x^{\rho 0} \exists n^0 \forall \tilde{x}^{\rho 0} \big(\overline{x,n} =_{\rho 0} \overline{\tilde{x},n} \to y\tilde{x} \le n\big).$$

Use this to give an alternative proof of the previous exercise.

25) Prove the claim in remark 3.72.

Historical comments and suggested further reading:
For an introduction to intuitionistic logic and Heyting arithmetic as well as constructivism in general we refer to van Dalen [372] (chapter 5) and the first volume of Troelstra-van Dalen [371]. Most of the material on HA and its finite type extensions is taken from Troelstra [366]. Primitive recursion in higher types (in the sense of the recursors R_ρ was first considered in Hilbert [161] and Gödel [133]. The fragments $\widehat{\text{(W)E-PA}}^\omega{\restriction}$ are studied in Feferman [98]. For pure types the functionals definable in these fragments correspond to the primitive recursive functionals defined by the schemata S1-S8 of Kleene [195]. The fragments $\text{G}_n\text{A}^\omega$ were introduced and studied first in Kohlenbach [207] from which most material of the corresponding section is taken. See Troelstra [365] for further information on the BHK-interpretation. A thorough treatment of the Curry-Howard isomorphism can be found in Sørensen-Urzyczyn [349]. For more information on (W)E-HA$^\omega$ and its numerous variants see Troelstra [366] and Troelstra-van Dalen [371].
The model $\mathscr{C}^\omega$ of sequentially continuous functionals was defined first in Scarpellini [321] (with further refinements in Scarpellini [322]) as a model of bar recursion (see chapter 11). The model ECF$^\omega$ of all extensional hereditarily continuous functionals is due (independently) to Kleene [196] and Kreisel [244]. For uses of ECF$^\omega$ in the proof theory of systems of intuitionistic analysis see again Troelstra [366]. The precise relationship between ECF$^\omega$ and $\mathscr{C}^\omega$ was clarified in Hyland [169]. For a comprehensive treatment of the recursion theory for ECF$^\omega$ see Normann [287]. The model $\mathscr{M}^\omega$ of all hereditarily strongly majorizable functionals is due to Bezem [27] (see also chapter 11).