7 A Glimpse at Veblen Hierarchies

What have we accomplished in section 6.5? If one examines carefully the proofs of propositions 6.23, 6.24, 6.27, 6.28, 6.31, 6.35, and definition 6.21, one discovers that the conditions that make everything go through are the fact that $\alpha \mapsto \omega^\alpha$ is a normal function φ such that $0 < \varphi(0)$. This suggests the following generalization.

Definition 7.1 Given *any* normal function φ such that $0 < \varphi(0)$, mimicking definition 6.21, we define the hierarchy $\{\varphi^0_\alpha\}_{\alpha\in\mathcal{O}}$ of functions such that,

- $\varphi^0_0 = \varphi$, and for every $\alpha > 0$,
- φ^0_α enumerates the set $\{\eta \mid \varphi^0_\beta(\eta) = \eta, \text{ for all } \beta < \alpha\}$ of common fixed points of the functions φ^0_β for all $\beta < \alpha$.

We have what is called a *Veblen hierarchy* (a concept due to Veblen [53]), and according to our previous remark, the following properties hold.

Theorem 7.2 (Veblen Hierarchy theorem) Denoting each function φ^0_α as $\varphi^0(\alpha, -)$, each $\varphi^0(\alpha, -)$ is a normal function, and the function $\varphi^0(-, 0) : \alpha \mapsto \varphi^0(\alpha, 0)$ is also a normal function such that $0 < \varphi^0(0, 0)$.

But since $\varphi^0(-, 0)$ satisfies the conditions for building a Veblen hierarchy, we can iterate the process just described in definition 7.1. For this, following Larry Miller [34], it is convenient to define an operator Δ_1 on normal functions, the *1-diagonalization operator*, defined as follows.

Given a normal function φ such that $0 < \varphi(0)$, $\Delta_1(\varphi)$ is the normal function enumerating the fixed points of $\varphi^0(-, 0)$.

Note that in a *single step*, Δ_1 performs the Ω iterations producing the Veblen hierarchy $\{\varphi^0_\alpha\}_{\alpha<\Omega}$! (where Ω denotes the first uncountable ordinal, i.e., the order type of $\mathcal{O}$). Using the operator Δ_1, we can define a sequence $\{\varphi^1_\beta\}_{\beta<\Omega}$ of normal functions, and so, a sequence of Veblen hierarchies – or a doubly indexed sequence of normal functions – $\{\varphi^1_\beta(\gamma, -)\}_{\beta,\gamma<\Omega}$ defined as follows:

- $\varphi^1_0 = \varphi$,
- $\varphi^1_{\beta'} = \Delta_1(\varphi^1_\beta)$, and
- φ^1_β is the normal function enumerating $\bigcap_{\gamma<\beta} range(\varphi^1_\gamma)$, for a limit ordinal β.

But $\beta \mapsto \varphi^1_\beta(0)$ (also denoted $\varphi^1(-, 0)$) is also a normal function such that $0 < \varphi^1_0(0)$. Hence, we can define an operator Δ_2 enumerating the fixed points of $\beta \mapsto \varphi^1_\beta(0)$, and build

a hierarchy. But we can iterate the operator Δ into the transfinite! This leads to the following definition.

Definition 7.3 Given a normal function φ such that $0 < \varphi(0)$, we define by simultaneous induction the Ω-indexed sequence $\{\Delta_\alpha\}_{\alpha<\Omega}$ of diagonalization operators and the doubly Ω-indexed sequence $\{\varphi^\alpha_\beta\}_{\alpha,\beta<\Omega}$ of normal functions as follows.

- $\Delta_0(\psi)$ enumerates the fixed points of the normal function ψ;
- $\Delta_{\alpha'}(\varphi) = \Delta_0(\varphi^\alpha(-,0))$ enumerates the fixed points of $\varphi^\alpha(-,0) : \beta \mapsto \varphi^\alpha_\beta(0)$;
- $\Delta_\alpha(\varphi)$ enumerates $\bigcap_{\gamma<\alpha} range(\Delta_\gamma(\varphi))$, for a limit ordinal α;
- $\varphi^\alpha_0 = \varphi$;
- $\varphi^\alpha_{\beta'} = \Delta_\alpha(\varphi^\alpha_\beta)$;
- φ^α_β enumerates $\bigcap_{\gamma<\beta} range(\varphi^\alpha_\gamma)$, for a limit ordinal β.

It is convenient to keep track of the diagonalization level (the index α) and the number of iterations of diagonalizations of level α (the index β) by using indices beyond Ω. Indeed, using the families $\{\varphi^\alpha_\beta\}_{\alpha,\beta<\Omega}$ and the representation of the ordinals in base Ω, it is possible to extend our original Ω-indexed hierarchy $\{\varphi(\beta,-)\}_{\beta<\Omega}$ (dropping the superscript 0 in φ^0) to an Ω^Ω-indexed hierarchy $\{\varphi(\delta,-)\}_{\delta<\Omega^\Omega}$. Let us first consider the simple case where $\alpha = 1$.

Using the fact that every ordinal $\delta < \Omega^2$ is uniquely expressed as $\delta = \Omega\beta_1 + \beta_2$ for some ordinals $\beta_1, \beta_2 < \Omega$, we can extend the Ω-indexed hierarchy $\{\varphi(\beta,-)\}_{\beta<\Omega}$ to an Ω^2-indexed hierarchy $\{\varphi(\delta,-)\}_{\delta<\Omega^2}$ as follows. For any $\beta_1, \beta_2 < \Omega$, we let

$$\varphi(\Omega\beta_1 + \beta_2, -) = (\varphi^1_{\beta_1})^0_{\beta_2}.$$

With this convention applied to the function $\omega(-) : \alpha \mapsto \omega^\alpha$ and the Ω^2-indexed sequence $\{\omega(\delta,-)\}_{\delta<\Omega^2}$, note that $\omega^1_1 = \Delta_1(\omega(-)) = \Delta_0(\omega^0(-,0))$ is denoted by $\omega(\Omega,-)$, and $\omega(\Omega,0) = \Gamma_0$ denotes the least fixed point of $\omega^0(-,0)$. Similarly, $\omega^2_1 = \Delta_2(\omega(-)) = \Delta_0(\omega^1(-,0))$ is denoted by $\omega(\Omega^2,-)$, and $\omega(\Omega^2,0)$ denotes the least fixed point of $\omega^1(-,0)$.

In general, since every ordinal $\delta < \Omega^\Omega$ is uniquely expressed as

$$\delta = \Omega^{\alpha_1}\beta_1 + \cdots + \Omega^{\alpha_n}\beta_n$$

for some ordinals $\alpha_n < \ldots < \alpha_1 < \Omega$ and $\beta_1, \ldots, \beta_n < \Omega$, we can regard the multiply Ω-indexed sequence

$$\{(\cdots(\varphi^{\alpha_1}_{\beta_1})\cdots)^{\alpha_n}_{\beta_n})\}_{\alpha_n<\ldots<\alpha_1<\Omega,\beta_1,\ldots,\beta_n<\Omega}$$

as an Ω^Ω-indexed sequence $\{\varphi(\delta, -)\}_{\delta<\Omega^\Omega}$, if we put

$$\varphi(\Omega^{\alpha_1}\beta_1 + \cdots + \Omega^{\alpha_n}\beta_n, -) = (\cdots(\varphi_{\beta_1}^{\alpha_1})\cdots)_{\beta_n}^{\alpha_n}).$$

Hence, a constructive ordinal notation system for the ordinals less than $\varphi(\Omega^\Omega, 0)$, the least fixed point of $\delta \mapsto \varphi(\delta, 0)$ $(\delta < \Omega^\Omega)$, can be obtained using the families

$$\{(\cdots(\varphi_{\beta_1}^{\alpha_1})\cdots)_{\beta_n}^{\alpha_n})\}_{\alpha_n<\ldots<\alpha_n<\Omega,\beta_1,\ldots,\beta_n<\Omega}.$$

It is possible to go farther using Bachmann-Isles hierarchies, but we are already quite dizzy, and refer the reader to Larry Miller's paper [34]. Readers interested in the topic of ordinal notations should consult the very nice expository articles by Crossley and Bridge Kister [5], Miller [34], and Pohlers [42], and for deeper results, Schütte [46] and Pohlers [41].

8 Normal Form For the Ordinals $< \Gamma_0$

One of the most remarkable properties of Γ_0 is that the ordinals less than Γ_0 can be represented in terms of the functions $+$ and φ. First, we need the following lemma.

Lemma 8.1 Given an additive principal ordinal γ, if $\gamma = \varphi(\alpha, \beta)$, with $\alpha \leq \gamma$ and $\beta < \gamma$, then $\alpha < \gamma$ iff γ is not strongly critical.

Proof. By proposition 6.31, we have $\gamma \leq \varphi(\gamma, 0)$. By proposition 6.28, since $\alpha \leq \gamma$ and $\beta < \gamma \leq \varphi(\gamma, 0)$, we have $\gamma = \varphi(\alpha, \beta) < \varphi(\gamma, 0)$ iff $\alpha < \gamma$. By proposition 6.34 and proposition 6.31, γ is not critical iff $\gamma < \varphi(\gamma, 0)$, iff $\alpha < \gamma$ from above. □

We can now prove the fundamental normal form representation theorem for the ordinals less than Γ_0.

Theorem 8.2 For every ordinal α such that $0 < \alpha < \Gamma_0$, there exist unique ordinals $\alpha_1, \ldots, \alpha_n, \beta_1, \ldots, \beta_n$, $n \geq 1$, with $\alpha_i, \beta_i < \varphi(\alpha_i, \beta_i) \leq \alpha$, $1 \leq i \leq n$, such that

(1) $\alpha = \varphi(\alpha_1, \beta_1) + \cdots + \varphi(\alpha_n, \beta_n)$, and

(2) $\varphi(\alpha_1, \beta_1) \geq \ldots \geq \varphi(\alpha_n, \beta_n)$.

Proof. Using the Cantor Normal Form for the (countable) ordinals (proposition 6.20), there are unique ordinals $\eta_1 \geq \ldots \geq \eta_n$, $n \geq 1$, such that

$$\alpha = \omega^{\eta_1} + \cdots + \omega^{\eta_n}.$$

Each ordinal ω^{η_i} is an additive principal ordinal, and let $\gamma_i = \omega^{\eta_i}$. By Proposition 6.32, for every additive principal ordinal γ_i, there exist unique $\alpha_i, \beta_i \in \mathcal{O}$ such that, $\alpha_i \leq \gamma_i$, $\beta_i < \gamma_i$,

and $\gamma_i = \varphi(\alpha_i, \beta_i)$. Since for each ordinal γ_i, we have $\gamma_i \leq \alpha < \Gamma_0$, and Γ_0 is the least strongly critical ordinal, by proposition 8.1, $\alpha_i < \gamma_i$. Since $\gamma_i \leq \alpha$, $\alpha_i < \gamma_i$, and $\beta_i < \gamma_i$, we have $\alpha_i < \alpha$ and $\beta_i < \alpha$. Property (2) follows from the fact that $\eta_1 \geq \ldots \geq \eta_n$ implies that $\gamma_1 \geq \ldots \geq \gamma_n$ (since $\gamma_i = \omega^{\eta_i}$). □

We need a few more properties of the ordinals less than Γ_0 before we establish the connection between Γ_0 and Kruskal's theorem.

Lemma 8.3 For all $\alpha, \beta < \Gamma_0$, if $\alpha \leq \beta$, then

$$\alpha \leq \beta \leq \beta + \alpha \leq \varphi(\beta, \alpha),$$

and if $\alpha \leq \beta$ and $\beta < \varphi(\alpha, \beta)$, then

$$\beta + \alpha \leq \varphi(\alpha, \beta) \leq \varphi(\beta, \alpha).$$

Proof. That $\alpha \leq \beta \leq \beta + \alpha$ is easy to show. If $\alpha = 0$, since by proposition 6.31, $\beta \leq \varphi(\beta, 0)$, we have $\beta + 0 = \beta \leq \varphi(\beta, 0)$. If $0 < \alpha = \beta$, we have shown earlier that $\alpha < \varphi(\alpha, \alpha)$ (in the proof of proposition 6.32), and since $\varphi(\alpha, \alpha)$ is an additive principal ordinal, we also have $\alpha + \alpha < \varphi(\alpha, \alpha)$. If $0 < \alpha < \beta$, by proposition 6.29, we have $\beta \leq \varphi(0, \beta)$, and by proposition 6.31, we have $\beta \leq \varphi(\beta, 0)$. By strict monotonicity of φ_β, since $\alpha > 0$, we have $\beta < \varphi(\beta, \alpha)$. Hence, $\alpha < \beta < \varphi(\beta, \alpha)$. By proposition 6.28, $\varphi(0, \beta) < \varphi(\beta, \alpha)$, since $\beta < \varphi(\beta, \alpha)$. Hence,

$$\beta + \alpha \leq \varphi(0, \beta) + \varphi(\beta, \alpha) = \varphi(\beta, \alpha),$$

since $\varphi(0, \beta) < \varphi(\beta, \alpha)$ and $\varphi(\beta, \alpha)$ is an additive principal ordinal.

Now assume $\alpha \leq \beta$ and $\beta < \varphi(\alpha, \beta)$. If $\alpha = 0$, since by proposition 6.29, $\beta \leq \varphi(0, \beta)$, we have $\beta + 0 = \beta \leq \varphi(0, \beta)$. If $0 < \alpha = \beta$, the proof is identical to the proof of the previous case. If $0 < \alpha < \beta$, then by proposition 6.28, $\varphi(0, \beta) < \varphi(\alpha, \beta)$, since $\beta < \varphi(\alpha, \beta)$. We can also show that $\alpha < \varphi(\alpha, \beta)$ as in the previous case (since $\beta > 0$), and we have

$$\beta + \alpha \leq \varphi(0, \beta) + \varphi(\alpha, \beta) = \varphi(\alpha, \beta),$$

since $\varphi(0, \beta) < \varphi(\alpha, \beta)$ and $\varphi(\alpha, \beta)$ is an additive principal ordinal. The fact that $\varphi(\alpha, \beta) \leq \varphi(\beta, \alpha)$ if $\alpha \leq \beta$ was shown in proposition 6.31. □

It should be noted that if $\alpha \leq \beta$, when $\beta = \varphi(\alpha, \beta)$ (which happens when $\beta \in Cr(\alpha')$), the inequality $\beta + \alpha \leq \varphi(\alpha, \beta)$ is *incorrect*. This minor point noted at the very end of Simpson's paper [47, page 117] is overlooked in one of Smoryński's papers [51, page 394]. In the next section, we will correct Smoryński's defective proof (Simpson's proof is also defective, but he gives a glimpse of a "repair" at the very end of his paper, page 117).

By theorem 8.2, the ordinals less than Γ_0 can be defined recursively as follows.

Lemma 8.4 For every ordinal $\gamma < \Gamma_0$, either

(1) $\gamma = 0$, or

(2) $\gamma = \beta + \alpha$, for some ordinals $\alpha, \beta < \gamma$ such that $\alpha \leq \beta$, or

(3) $\gamma = \varphi(\alpha, \beta)$, for some ordinals $\alpha, \beta < \gamma$.

Proof. The proof follows immediately from theorem 8.2 by induction on n in the decomposition $\gamma = \varphi(\alpha_1, \beta_1) + \cdots + \varphi(\alpha_n, \beta_n)$. □

In case (3), we cannot guarantee that $\alpha \leq \beta$, and we have to consider the three subcases $\alpha < \beta$, $\alpha = \beta$, and $\alpha > \beta$. Actually, we can reduce these three cases to two if we replace $<$ by $\leq$.

This recursive representation of the ordinals $< \Gamma_0$ is the essence of the connection between Γ_0 and Kruskal's theorem explored in section 9.

Lemma 8.4 shows that every ordinal $\alpha < \Gamma_0$ *can be* represented in terms of 0, +, and φ, but this representation has some undesirable properties, namely that different notations can represent the same ordinal. In particular, for some $\alpha \leq \beta < \Gamma_0$, we may have $\beta = \varphi(\alpha, \beta)$ (which happens when $\beta \in Cr(\alpha')$). For example, $\epsilon_0 = \varphi(0, \epsilon_0)$ (since $\epsilon_0 = \varphi(1, 0)$). It would be desirable to have a representation similar to that given by lemma 8.2, but for a function ψ such that $\alpha < \psi(\alpha, \beta)$ and $\beta < \psi(\alpha, \beta)$, for all $\alpha, \beta < \Gamma_0$. Such a representation is possible, as shown in Schütte [46, Section 13.7, page 84-92]. The key point is to consider ordinals γ that are *maximal α-critical*, that is, maximal with respect to the property of belonging to some $Cr(\alpha)$.

Definition 8.5 An ordinal $\gamma \in \mathcal{O}$ is *maximal α-critical* iff $\gamma \in Cr(\alpha)$ and $\gamma \notin Cr(\beta)$ for all $\beta > \alpha$.

By proposition 6.22 and proposition 6.23, $\gamma \in Cr(\alpha)$ iff $\varphi_\beta(\gamma) = \gamma$ for all $\beta < \alpha$. Thus, γ is maximal α-critical iff $\varphi_\alpha(\gamma) \neq \gamma$. However, because φ_α is the ordering function of $Cr(\alpha)$, we know from proposition 6.9 that $\delta \leq \varphi_\alpha(\delta)$ for all δ, and so, γ is maximal α-critical iff $\gamma = \varphi_\alpha(\beta)$ for some $\beta < \gamma$. It follows from proposition 6.32 that for every principal additive number γ, there is some $\alpha \leq \gamma$ such that γ is maximal α-critical.

Definition 8.6 The function ψ_α is defined as the ordering function of the maximal α-critical ordinals.

We also define $\psi(\alpha, \beta)$ by letting $\psi(\alpha, \beta) = \psi_\alpha(\beta)$. It is possible to give a definition of ψ in terms of φ, as shown in Schütte [46].

Lemma 8.7 The function ψ defined such that

$$\psi(\alpha,\beta) = \begin{cases} \varphi(\alpha,\beta+1), & \text{if } \beta = \beta_0 + n \text{ and } \varphi(\alpha,\beta_0) = \beta_0, \\ & \text{for some } \beta_0 \text{ and } n \in \mathbf{N}; \\ \varphi(\alpha,\beta), & \text{otherwise.} \end{cases}$$

is the ordering function of the maximal α-critical ordinals for every α.

We list the following properties of ψ without proof, referring the reader to Schütte [46] for details.

Lemma 8.8 For every additive principal number γ, there are unique $\alpha, \beta \leq \gamma$ such that $\gamma = \psi(\alpha,\beta)$.

Lemma 8.9 (1) If $\gamma = \psi(\alpha,\beta)$, then $\alpha < \gamma$ iff γ is not strongly critical.

(2) $\beta < \psi(\alpha,\beta)$ for all α, β.

Lemma 8.10 $\psi(\alpha_1,\beta_1) < \psi(\alpha_2,\beta_2)$ holds iff either

(1) $\alpha_1 < \alpha_2$ and $\beta_1 < \psi(\alpha_2,\beta_2)$, or

(2) $\alpha_1 = \alpha_2$ and $\beta_1 < \beta_2$, or

(3) $\alpha_2 < \alpha_1$ and $\psi(\alpha_1,\beta_1) \leq \beta_2$.

It should be noted that the set of maximal α-critical ordinals is unbounded, but it is not closed, because the function ψ_α is *not* continuous. However, this is not a problem for representing the ordinals less than Γ_0.

Since Γ_0 is the least strongly critical ordinal, by lemma 8.9, we have the following corollary.

Lemm 8.11 For all $\alpha, \beta < \Gamma_0$, we have

(1) $\alpha < \psi(\alpha,\beta)$, and

(2) $\beta < \psi(\alpha,\beta)$.

Using lemma 8.8, we can prove another version of the normal form theorem 8.2 for the ordinal less than Γ_0, using ψ instead of φ.

Theorem 8.12 For every ordinal α such that $0 < \alpha < \Gamma_0$, there exist unique ordinals $\alpha_1, \ldots, \alpha_n, \beta_1, \ldots, \beta_n$, $n \geq 1$, with $\alpha_i, \beta_i < \psi(\alpha_i,\beta_i) \leq \alpha$, $1 \leq i \leq n$, such that

(1) $\alpha = \psi(\alpha_1,\beta_1) + \cdots + \psi(\alpha_n,\beta_n)$, and

(2) $\psi(\alpha_1, \beta_1) \geq \ldots \geq \psi(\alpha_n, \beta_n)$.

The advantage of the reprentation given by theorem 8.12 is that it is now possible to design a system of notations where distinct notations represent distinct ordinals, and ψ satisfies the subterm property of lemma 8.11. Such a notation system will be given in section 11.

9 Kruskal's Theorem and Γ_0

The connection between Γ_0 and Kruskal's theorem lies in the fact that there is a close relationship between the embedding relation $\preceq$ on trees (definition 4.11) and the well-ordering $\leq$ on $\mathcal{O}(\Gamma_0)$ (recall that $\mathcal{O}(\Gamma_0)$ is the set of all ordinals $< \Gamma_0$).

We shall restrict our attention to tree domains, or equivalently assume that the set of labels contains a single symbol. Let T denote the set of all finite tree domains, which, for brevity are also called trees. In this case, by a previous remark, it is easy to show that $\preceq$ is a partial order. We shall exhibit a function $h : T \to \mathcal{O}(\Gamma_0)$ from the set of finite trees to the set of ordinals less that Γ_0, and show that h is (1). surjective, and (2). preserves order, that is, if $s \preceq t$, then $h(s) \leq h(t)$ (where $\preceq$ is the embedding relation defined in definition 4.11). It will follow that Kruskal's theorem (theorem 4.12) implies that $\mathcal{O}(\Gamma_0)$ is well-ordered by $\leq$, or put slightly differently, Kruskal's theorem implies the validity of transfinite induction on Γ_0. In turn, the provability of transfinite induction on large ordinals is known to be proof-theoretically significant. As first shown by Gentzen, one can prove the consistency of logical theories using transfinite induction on large ordinals. As a consequence, Kruskal's theorem in not provable in fairly strong logical theories, in particular some second-order theories for which transfinite induction up to Γ_0 is not provable.

We now give the definition of the function h mentioned above. In view of the recursive characterization of the ordinals $< \Gamma_0$, it is relatively simple to define a surjective function from T to $\mathcal{O}(\Gamma_0)$. However, making h order-preserving is more tricky. As a matter of fact, this is why lemma 8.3 is needed, but beware! Simpson defines a function h using five recursive cases, but points out at the end of his paper that there is a problem, due to the failure of the inequality $\beta + \alpha \leq \varphi(\alpha, \beta)$ [47, page 117]. Actually, a definition with fewer cases can be given, and Smoryński defines a function h using four recursive cases [51]. Unfortunately, Smoryński's definition also makes use of the erroneous inequality [51, page 394]. We give what we believe to be a repaired version of Smoryński's definition of h (using five recursive cases).

Remark. We do not know whether a definition using the function ψ of the previous section can be given. Certainly a surjective function can be defined using ψ, but the difficult

part is to insure monotonicity.

Definition 9.1 The function $h : T \to \mathcal{O}(\Gamma_0)$ from the set of finite trees to the set of ordinals less that Γ_0 is defined recursively as follows:

(0) $h(t) = 0$, when t is the one-node tree.

(1) $h(t) = h(t/1)$, if $rank(t) = 1$, i.e, the root of t has only one successor.

(2) $h(t) = \beta + \alpha$, if $rank(t) = 2$, where α is the least element of $\{h(t/1), h(t/2)\}$ and β is the largest.

(3) $h(t) = \varphi(\alpha, \beta)$, if $rank(t) = 3$, where $\alpha \leq \beta$ are the two largest elements of the set $\{h(t/1), h(t/2), h(t/3)\}$, and $\beta < \varphi(\alpha, \beta)$.

(4) $h(t) = \beta + \alpha$, if $rank(t) = 3$, where $\alpha \leq \beta$ are the two largest elements of the set $\{h(t/1), h(t/2), h(t/3)\}$, and $\beta = \varphi(\alpha, \beta)$.

(5) $h(t) = \varphi(\beta, \alpha)$, if $rank(t) \geq 4$, where $\alpha \leq \beta$ are the two largest elements of the set $\{h(t/1), h(t/2), \ldots, h(t/k)\}$, with $k = rank(t)$.

The following important theorem holds.

Theorem 9.2 The function $h : T \to \mathcal{O}(\Gamma_0)$ is surjective and monotonic, that is, for every two finite tree s, t, if $s \preceq t$, then $h(s) \leq h(t)$.

Proof (sketch). The fact that h is surjective follows directly from the recursive definition shown in lemma 8.4. Note that clause (1) and (4) are not needed for showing that h is a surjection, but they are needed to ensure that h is well defined and preserves order. By clause (0), $h(t) = 0$, for the one-node tree t. Clause (2) is used when $\gamma = \beta + \alpha$, with $\alpha, \beta < \gamma$ and $\alpha \leq \beta$. Clause (3) is used when $\gamma = \varphi(\alpha, \beta)$ with $\alpha, \beta < \gamma$ and $\alpha \leq \beta$, and clause (5) is used when $\gamma = \varphi(\beta, \alpha)$ with $\alpha, \beta < \gamma$ and $\alpha \leq \beta$.

The proof that if $s \preceq t$, then $h(s) \leq h(t)$ proceeds by cases, using induction on trees, corollary 6.30, and lemma 8.3. The only delicate case arises when $rank(s) = 2$, $rank(t) = 3$, and, assuming that $h(t/1) \geq h(t/2) \geq h(t/3)$ and $h(s/1) \geq h(s/2)$, we have $h(t/1) = \varphi(h(t/2), h(t/1))$, $s/1 \preceq t/1$ and $s/2 \preceq t/2$. By the induction hypothesis, $h(s/1) \leq h(t/1)$ and $h(s/2) \leq h(t/2)$, and since $h(s) = h(s/1) + h(s/2)$ and $h(t) = h(t/1) + h(t/2)$, we have $h(s) \leq h(t)$. If $h(t/1) < \varphi(h(t/2), h(t/1))$, then $h(t) = \varphi(h(t/2), h(t/1))$, and by proposition 8.3, $h(s) = h(s/1) + h(s/2) \leq h(t/1) + h(t/2) \leq \varphi(h(t/2), h(t/1)) = h(t)$. The other cases are left to the reader. □

Theorem 9.2 implies that there exist total orderings of order type Γ_0 extending the partial order $\preceq$ on (finite) trees. DeJongh and Parikh [6] proved that the maximum (sup)

of all the total extensions is attained, and they computed the maximum for certain of the (Higman) orderings. The ordinals associated with various orderings on trees arising in the theory of rewriting systems have been investigated by Dershowitz and Okada [9], Okada and Takeuti [38], and Okada [37, 39, 40].

Theorem 9.2 also has the following important corollary.

Lemma 9.3 Kruskal's theorem implies that $\mathcal{O}(\Gamma_0)$ is well-ordered by $\leq$.

Proof. Assume that there is some infinite sequence $(\alpha_i)_{i\geq 1}$ of ordinals in $\mathcal{O}(\Gamma_0)$ such that $\alpha_{i+1} < \alpha_i$ for all $i \geq 1$. By theorem 9.2, since h is surjective, there is an infinite sequence of trees $(t_i)_{i\geq 1}$ such that $h(t_i) = \alpha_i$ for all $i \geq 1$. By Kruskal's theorem (theorem 4.12), there exist $i, j > 0$ such that $i < j$ and $t_i \preceq t_j$. By theorem 9.2, we have $\alpha_i = h(t_i) \leq h(t_j) = \alpha_j$, contradicting the fact that $\alpha_j < \alpha_i$. Hence, $\mathcal{O}(\Gamma_0)$ is well-ordered by $\leq$. $\square$

Let us denote by $WO(\Gamma_0)$ the property that $\mathcal{O}(\Gamma_0)$ is well-ordered by $\leq$, and by $WQO(T)$ the property that the embedding relation $\preceq$ is a *wqo* on the set T of finite trees. $WQO(T)$ is a formal statement of Kruskal's theorem.

For every formal system $\mathcal{S}$, if the proof that $(WQO(T) \supset WO(\Gamma_0))$ (given in lemma 9.3) can be formalized in $\mathcal{S}$ and $WO(\Gamma_0)$ is not provable in $\mathcal{S}$, then $WQO(T)$ is not provable in $\mathcal{S}$. In the next section, we briefly describe some subsystems of 2^{nd}-order arithmetic in which Kruskal's theorem and its miniature versions are not provable.

10 The Subsystems ACA_0, ATR_0, Π^1_1-CA_0, of Second-Order Arithmetic

Harvey Friedman has shown that $WO(\Gamma_0)$ is not provable in some relatively strong subsystems of 2^{nd}-order arithmetic, and therefore, Kruskal's theorem is not provable in such systems. Friedman also proved similar results for some finite (first-order) miniaturizations of Kruskal's theorem. In particular, these first-order versions of Kruskal's theorem are not provable in Peano's arithmetic, since transfinite induction up to ϵ_0 is not provable in Peano's arithmetic, due to a result of Gentzen. We now provide some details on these subsystems of 2^{nd}-order arithmetic.

Second-order arithmetic can be formulated over a two-sorted language with *number variables* $(m, n, \ldots)$ and *set variables* $(X, Y, \ldots)$. We define *numerical terms* as terms built up from number variables, the constant symbols 0, 1, and the function symbols + (addition) and $\cdot$ (multiplication). An *atomic formula* is either of the form $t_1 \doteq t_2$, or $t_1 < t_2$, or $t_1 \in X$, where t_1 and t_2 are numerical terms. A *formula* is built up from atomic formulae using

$\wedge, \vee, \supset, \equiv, \neg$, number quantifiers $\forall n, \exists n$, and set quantifiers $\forall X, \exists X$. We say that a formula is *arithmetical* iff it does not contain set quantifiers.

All systems of second-order arithmetic under consideration include standard axioms stating that $\langle \mathbf{N}, 0, 1, +, \cdot, < \rangle$ is an ordered semi-ring. The real power of a system of second-order arithmetic is given by the form of its *induction axioms*, and the form of its *comprehension axioms*.

For the systems under consideration, the induction axiom is

$$[0 \in X \wedge \forall m(m \in X \supset m + 1 \in X)] \supset \forall n(n \in X),$$

where X is a set variable. This form of induction is often called *restricted induction*, in contrast with the principle of *full induction* stated as

$$[\varphi(0) \wedge \forall m(\varphi(m) \supset \varphi(m + 1))] \supset \forall n \varphi(n),$$

where φ is an arbitrary 2^{nd}-order formula. Apparently, Friedman initiated the study of subsystems of 2^{nd}-order arithmetic with restricted induction (this explains the subscript 0 after the name of the systems ACA, ATR, or Π^1_1-CA).

The system Π^1_∞-CA_0, also known as Z_2, or *second-order arithmetic*, has comprehension axioms of the form

$$\exists X \forall n(n \in X \equiv \varphi(n)),$$

where φ is any 2^{nd}-order formula φ in which X is not free. This is a very powerful form of comprehension axioms. Susbystems of Z_2 are obtained by restricting the class of formulae for which comprehension axioms hold.

The system ACA_0 is obtained by restricting the comprehension axioms to *arithmetical formulae* in which X is not free (ACA stands for Arithmetical Comprehension Axioms). It turns out that ACA_0 is a conservative extension of (first-order) Peano Arithmetic (PA). A weak form of König's lemma is provable in ACA_0, and a fairly smooth theory of continuous functions and of sequential convergence can be developed. For example, Friedman proved that the Bolzano/Weierstrass theorem (every bounded sequence of real numbers contains a convergent subsequence) is provable in ACA_0. In fact, Friedman proved the stronger result that no set existence axioms weaker than those of ACA_0 are sufficient to establish the Bolzano/Weierstrass theorem. For details, the reader is referred to Simpson [48].

The system ATR_0 contains axioms stating that arithmetical comprehension can be iterated along any countable well ordering (ATR stands for Arithmetical Transfinite Recursion). A precise formulation of the axiom ATR can be found in Friedman, McAloon, and

Simpson [16] (see also Feferman [14]), but it is not essential here. The system ATR_0 permits a convenient development of a large part of ordinary mathematics, including, the theory of continuous functions, the Riemann integral, the theory of countable fields, the topology of complete separable metric spaces, the structure theory of separable Banach spaces, a good theory of countable well orderings, Borel sets, analytic sets, and more.

The system Π^1_1-CA_0 is obtained by allowing comprehension axioms in which φ is any Π^1_1-formula in which X is not free. This is a system even stronger that ATR_0, whose axioms imply many mathematical results in the realm of algebra, analysis, classical descriptive set theory, and countable combinatorics.

The systems ACA, ATR and Π^1_1-CA allow full induction rather than restricted induction. It might be interesting to mention that the least ordinals for which transfinite induction cannot be proved in ACA_0 and ATR_0 are respectively ϵ_0 and Γ_0. Such an ordinal has also be determined for Π^1_1-CA_0, but the notation system required to describe it is beyond the scope of this paper. In contrast, the least ordinals for which transfinite induction cannot be proved in ACA and ATR are respectively ϵ_{ϵ_0} and Γ_{ϵ_0}.

We now return to the connections with Γ_0 and Kruskal's theorem. Friedman has shown that $WO(\Gamma_0)$ is not provable in ATR_0 (Friedman, McAloon, and Simpson [16]). He also showed that $(WQO(T) \supset WO(\Gamma_0))$ is provable in ACA_0. Since ACA_0 is a subsystem of ATR_0, we conclude that $WQO(T)$ is not provable in ATR_0. This is already quite remarkable, considering that a large part of ordinary mathematics can be done in ATR_0. But Friedman also proved that the miniature version $LWQO(T)$ of Kruskal theorem given in theorem 5.1 is not provable in ATR_0, an even more remarkable result. The proof of this last result is given in Simpson [47].

There is one more "tour de force" of Friedman that we have not mentioned! Harvey Friedman has formulated an extension of the miniature version of Kruskal's theorem (using a gap condition), and proved that this version of Kruskal's theorem is not provable in Π^1_1-CA_0. The proof can be found in Simpson [47]. There are also some connections bewteen this last version of Kruskal's theorem and certain ordinal notations due to Takeuti known as ordinals diagrams. These connections ae investigated in Okada and Takeuti [38], and Okada [39, 40].

11 A Brief Introduction to Term Orderings

This section is a brief introduction to term orderings. These orderings play an important role in computer science, because they are the main tool for showing that sets of rewrite rules are finite terminating (Noetherian). In turn, Noetherian sets of rewrite rules play a

fundamental role in automated deduction in equational logic. Indeed, one of the major techniques in equational logic is to complete a given set of equations E to produce an equivalent set R of rewrite rules which has some "good" properties, namely to be confluent and Noetherian. A number of procedures that attempt to produce such a set R of rewrite rules from a set E of equations have been designed. The first such procedure is due to Knuth and Bendix [27], but there are now many kinds of completion procedures. For more details on completion procedures, we refer the reader to Dershowitz [11] and Bachmair [2].

There are many classes of term orderings, but an important class relevant to our considerations is the class of simplification orderings, because Kruskal's theorem can be used to prove the well-foundedness of these orderings. For a comprehensive study of term orderings, the reader is referred Dershowitz's excellent survey [7] and to Dershowitz's fundamental paper [8].

Given a set of labels Σ, the notion of a tree was defined in definition 4.2. When considering rewrite rules, we usually assume that Σ is a ranked alphabet, that is, that there is a ranking function $r : \Sigma \to \mathbf{N}$ assigning a natural number $r(f)$, the *rank* (or *arity*) of f, to every $f \in \Sigma$. We also have a countably infinite set $\mathcal{X}$ of variables, with $r(x) = 0$ for every $x \in \mathcal{X}$, and we let $T_\Sigma(\mathcal{X})$ be the set of all trees (also called Σ-*terms*, or *terms*) $t \in T_{\Sigma \cup \mathcal{X}}$ such that, for every tree address $u \in dom(t)$, $r(t(u)) = rank(t/u)$. In other words, the rank of the label of u is equal to the rank of t/u (see definition 4.3), the number of immediate successors of u.

Given a tree t, we let $Var(t) = \{x \in \mathcal{X} \mid \exists u \in dom(t),\ t(u) = x\}$ denote the set of variables occurring in t. A *ground term* t is a term such that $Var(t) = \emptyset$.

Definition 11.1 A *set of rewrite rules* is a binary relation $R \subseteq T_\Sigma(\mathcal{X}) \times T_\Sigma(\mathcal{X})$ such that $Var(r) \subseteq Var(l)$ whenever $\langle l, r\rangle \in R$.

A rewrite rule $\langle l, r\rangle \in R$ is usually denoted as $l \to r$. The notions of tree replacement and substitution are needed for the definition of the rewrite relation induced by a set of rewrite rules.

Definition 11.2 Given two trees t_1 and t_2 and a tree address u in t_1, the *result of replacing* t_2 at u in t_1, denoted by $t_1[u \leftarrow t_2]$, is the function whose graph is the set of pairs

$$\{(v, t_1(v)) \mid v \in dom(t_1),\ u \text{ is not a prefix of } v\} \cup \{(uv, t_2(v)) \mid v \in dom(t_2)\}.$$

Definition 11.3 A *substitution* is a function $\sigma : \mathcal{X} \to T_\Sigma(\mathcal{X})$, such that, $\sigma(x) \neq x$ for only finitely many $x \in \mathcal{X}$. Since $T_\Sigma(\mathcal{X})$ is the free Σ-algebra generated by $\mathcal{X}$, every substitution $\sigma : \mathcal{X} \to T_\Sigma(\mathcal{X})$ has a unique homomorphic extension $\widehat{\sigma} : T_\Sigma(\mathcal{X}) \to T_\Sigma(\mathcal{X})$. In the sequel, we will identify σ and its homomorphic extension $\widehat{\sigma}$, and denote $\widehat{\sigma}(t)$ as $t[\sigma]$.

Definition 11.4 Given a substitution σ, the *domain* of σ is the set of variables $D(\sigma) = \{x \mid \sigma(x) \neq x\}$. Given a substitution σ, if its domain is the set $\{x_1, \ldots, x_n\}$, and if $t_i = \sigma(x_i)$, $1 \leq i \leq n$, then σ is also denoted by $[t_1/x_1, \ldots, t_n/x_n]$.

Definition 11.5 A substitution σ is a *renaming* iff $\sigma(x)$ is a variable for every $x \in D(\sigma)$, and σ is injective. Let $R \subseteq T_\Sigma(\mathcal{X}) \times T_\Sigma(\mathcal{X})$ be a set of rewrite rules. A rewrite rule $s \to t$ is a *variant* of a rewrite rule $u \to v \in R$ iff there is some renaming ρ with domain $Var(u) \cup Var(v)$ such that $s = u[\rho]$ and $t = v[\rho]$.

Definition 11.6 Let $\longrightarrow$ be a binary relation $\longrightarrow \subseteq T_\Sigma(\mathcal{X}) \times T_\Sigma(\mathcal{X})$. (i) The relation $\longrightarrow$ is *monotonic* (or *stable under the algebra structure*) iff for every two terms s, t and every function symbol $f \in \Sigma$, if $s \longrightarrow t$ then $f(\ldots, s, \ldots) \longrightarrow f(\ldots, t, \ldots)$.

(ii) The relation $\longrightarrow$ is *stable* (under substitution) if $s \longrightarrow t$ implies $s[\sigma] \longrightarrow t[\sigma]$ for every substitution σ.

Definition 11.7 Let $R \subseteq T_\Sigma(\mathcal{X}) \times T_\Sigma(\mathcal{X})$ be a set of rewrite rules. The relation $\longrightarrow_R$ over $T_\Sigma(\mathcal{X})$ is defined as the smallest stable and monotonic relation that contains R. This is the *rewrite relation* associated with R.

This relation is defined explicitly as follows: Given any two terms $t_1, t_2 \in T_\Sigma(\mathcal{X})$, then

$$t_1 \longrightarrow_R t_2$$

iff there is some variant $l \to r$ of some rule in R, some tree address α in t_1, and some substitution σ, such that

$$t_1/\alpha = l[\sigma], \quad \text{and} \quad t_2 = t_1[\alpha \leftarrow r[\sigma]].$$

We say that a rewrite system R is Noetherian iff the relation $\longrightarrow_R$ associated with R is Noetherian.

Now, our goal is to describe some orderings that will allow us to prove that sets of rewrite rules are Noetherian. First, it is convenient to introduce the concept of a strict ordering.

Definition 11.8 A *strict ordering* (or *strict order*) $\prec$ on a set A is a transitive and irreflexive relation (for all a, $a \not\prec a$.)

Given a preorder (or partial order) $\preceq$ on a set A, the strict ordering $\prec$ associated with $\preceq$ is defined such that $s \prec t$ iff $s \preceq t$ and $t \not\preceq s$. Conversely, given a strict ordering $\prec$,

the partial ordering $\preceq$ associated with $\prec$ is defined such that $s \preceq t$ iff $s \prec t$ or $s = t$. The converse of a strict ordering $\prec$ is denoted as $\succ$.

We now introduce the important concepts of simplification ordering, and reduction ordering. Let Σ be a set of labels (in most cases, a ranked alphabet).

Definition 11.9 A strict order $\prec$ on T_Σ satisfying conditions

(1) $s \prec f(\ldots, s, \ldots)$, and

(2) $f(\ldots) \prec f(\ldots, s, \ldots)$,

is said to have the *subterm property* and the *deletion property*.

A *simplification ordering* $\prec$ is a strict ordering that is monotonic and has the subterm and deletion property.[1]

A *reduction ordering* $\prec$ is a strict ordering that is monotonic, stable under substitution, and such that $\succ$ is well-founded.

With a slight abuse of language, we will also say that the converse $\succ$ of a strict ordering $\prec$ is a simplification ordering (or a reduction ordering). The importance of term orderings is shown by the next fundamental result.

Lemma 11.10 A set of rules R is Noetherian if and only if there exists a reduction ordering $\succ$ on $T_\Sigma(\mathcal{X})$ such that $l \succ r$ for every $l \to r \in R$.

Unfortunately, it is undecidable in general if an arbitrary system R is Noetherian since it is possible to encode Turing machines using a system of two rewrite rules, and this would imply the decidability of the halting problem (see Dershowitz [7]). The importance of simplification orderings is shown by the next theorem.

Theorem 11.11 (Dershowitz) If Σ is finite, then every simplification ordering on T_Σ is well-founded.

Proof. This is a consequence of proposition 4.8, which uses Kruskal's tree theorem. □

In practice, we want theorem 11.11 to apply to simplification orderings on $T_\Sigma(\mathcal{X})$, but since $\mathcal{X}$ is infinite, there is a problem. However, we are saved because we usually only care about terms arising in derivations.

---

[1] When Σ is a ranked alphabet, the deletion property is superfluous.

Definition 11.12 An ordering $\succ$ is *well-founded for derivations* iff $\succ \cap \xrightarrow{*}_R$ is well-founded for every finite rewrite system R.

Since $Var(r) \subseteq Var(l)$ for every $l \to r \in R$, every derivation of a finite rewrite system involves only finitely many symbols. Thus, as corollary of the above theorem we have:

Corollary 11.13 (Dershowitz) Every simplification ordering is well-founded for derivations.

Warning: There exists rewrite systems whose termination cannot be shown by any total simplification ordering as shown by the following example.

Example 11.14

$$f(a) \to f(b)$$
$$g(b) \to g(a)$$

Next, we are going to describe two important classes of simplification orderings, the recursive path ordering, and the lexicographic path ordering. But first, we need to review the definitions of the lexicographic ordering and the multiset ordering.

Definition 11.15 Given n partially ordered sets $(S_i, \prec_i)$ (where each $\prec_i$ is a strict order, $n > 1$), the *lexicographic order* $\prec_{lex}$ on the set $S_1 \times \cdots \times S_n$ is defined as follows. Let $\langle a_1, \ldots, a_n \rangle$ and $\langle b_1, \ldots, b_n \rangle$ be members of $S_1 \times \cdots \times S_n$. Then

$$\langle a_1, \ldots, a_n \rangle \prec_{lex} \langle b_1, \ldots, b_n \rangle$$

if and only if there exists some i, $1 \leq i \leq n$, such that $a_i \prec_i b_i$, and $a_j = b_j$ for all j, $1 \leq j < i$.

We now turn to multiset orderings. Multiset orderings have been investigated by Dershowitz and Manna [10], and Jouannaud and Lescanne [24].

Definition 11.16 Given a set A, a *multiset* over A is an unordered collection of elements of A which may have multiple occurrences of identical elements. More formally, a multiset over A is a function $M : A \to \mathbf{N}$ (where $\mathbf{N}$ is the set of natural numbers) such that an element $a \in A$ has exactly n occurrences in M iff $M(a) = n$. In particular, a does not belong to M when $M(a) = 0$, and we say that $a \in M$ iff $M(a) > 0$.

The *union* of two multisets M_1 and M_2, denoted by $M_1 \cup M_2$, is defined as the multiset M such that for all $a \in A$, $M(a) = M_1(a) + M_2(a)$.

Let $(S, \prec)$ be a partially ordered set (where $\prec$ is a strict order), let M be some finite multiset of objects from S, and finally let n, $n'_1, \ldots, n'_k \in S$. Define the relation $\Leftarrow_m$ on finite multisets as

$$M \cup \{n'_1, \ldots, n'_k\} \Leftarrow_m M \cup \{n\},$$

where $k \geq 0$ and $n'_i \prec n$ for all i, $1 \leq i \leq k$.

The multiset ordering $\prec_{M(S)}$ is simply the transitive closure $\overset{+}{\Leftarrow}_m$.

In other words, $N' \prec_{M(S)} N$ iff N' is produced from a finite multiset N by removing one or more elements and replacing them with any finite number of elements, each of which is strictly smaller than at least one element removed. For example, $\{4, 4, 3, 3, 1\} \prec \{5, 3, 1, 1\}$, where $\prec$ is the multiset ordering induced by the ordering $<$ of the natural numbers.

It is easy to show that for any partially ordered set $(S, \preceq)$, we have associated partially ordered sets $(M(S), \preceq_{M(S)})$ (where $M(S)$ is the set of all finite multisets of members of S), and $(S^n, \preceq_{lex})$ for $n > 0$. Furthermore $\preceq$ is total (respectively, well-founded) iff $\preceq_{lex}$ (for any n) is total (respectively, well-founded).

Using König's lemma, we can also show the following useful result.

Lemma 11.17 If $\preceq$ is well-founded (respectively, total) on S, then $\preceq_{M(S)}$ is well-founded (respectively, total) on $M(S)$.

There is an interesting connection between the multiset ordering and ordinal exponentiation. Given a well ordering $\preceq$ on a set S, it is well know that there is a unique ordinal α and a unique order-preserving bijection $\varphi : S \to \alpha$.

The connection is that $(M(S), \prec_{M(S)})$ is order-isomorphic to ω^α. Indeed, the function $\psi\colon M(S) \to \omega^\alpha$ defined such that $\psi(\emptyset) = 0$, and

$$\psi(\{m_1, \ldots, m_k\}) = \omega^{\varphi(m_1)} + \cdots + \omega^{\varphi(m_k)},$$

where $\varphi(m_1) \geq \ldots \geq \varphi(m_k)$ is the nonincreasing sequence enumerating $\varphi(\{m_1, \ldots, m_k\})$,[2] is easily shown to be an order-isomorphism.

The lexicographic ordering and the multiset ordering can also be defined for preorders. This generalization will be needed for defining *rpo* and *lpo* orderings based on preorders.

Definition 11.18 Given n preordered sets $(S_i, \preceq_i)$ $(n > 1)$, the *lexicographic preorder* $\preceq_{lex}$ on the set $S_1 \times \cdots \times S_n$ is defined as follows:

$$\langle a_1, \ldots, a_n \rangle \preceq_{lex} \langle b_1, \ldots, b_n \rangle$$

[2] In the theory of ordinals, the sum $\omega^{\varphi(m_1)} + \cdots + \omega^{\varphi(m_k)}$ is a *natural sum*.

if and only if there exists some i, $1 \leq i \leq n$, such that $a_i \preceq_i b_i$, and $a_j \approx_j b_j$ for all j, $1 \leq j < i$.[3]

Definition 11.19 Let $(S, \preceq)$ be a preordered set, let M be some finite multiset of objects from S, and finally let $n, n'_1, \ldots, n'_k \in S$. Define the relation $\Leftarrow_m$ on finite multisets as

$$M \cup \{n'_1, \ldots, n'_k\} \Leftarrow_m M \cup \{n\},$$

where either $k = 1$ and $n \approx n'_1$, or $k \geq 0$ and $n'_i \prec n$ for all i, $1 \leq i \leq k$.[4]

The multiset preorder $\preceq_{M(S)}$ is the transitive closure $\overset{+}{\Leftarrow}_m$.

Two finite multisets M_1 and M_2 are equivalent ($M_1 \approx_{M(S)} M_2$) iff they have the same number of elements, and every element of M_1 is equivalent to some element of M_2 and vice versa. It is easy to show that for any preordered set $(S, \preceq)$ we have associated preordered sets $(M(S), \preceq_{M(S)})$ (where $M(S)$ is the set of all finite multisets of members of S), and $(S^n, \preceq_{lex})$ for $n > 0$. Furthermore $\preceq$ is total (respectively, well-founded) iff $\preceq_{lex}$ (for any n) is total (respectively, well-founded).

Using König's lemma, we can also show that lemma 11.17 holds for preorders.

Lemma 11.20 If $\preceq$ is a well-founded preorder (respectively, total) on S, then $\preceq_{M(S)}$ is well-founded (respectively, total) on $M(S)$.

A naive ordering on terms based on the notion of lexicographic order is as follows.

For any given ordering $\succ$ on Σ we say that

$$s = f(s_1, \ldots, s_n) \succ^{tlex} g(t_1, \ldots, t_m) = t$$

iff either

(i) $f \succ g$; or

(ii) $f = g$ and $\langle s_1, \ldots, s_n \rangle \succ^{tlex}_{lex} \langle t_1, \ldots, t_n \rangle$,

where $\succ^{tlex}_{lex}$ is the lexicographic extension of $\succ^{tlex}$ to n-tuples of terms (the success of this recursive definition depends on the fact that we use the lexicographic extension over terms *smaller* than s and t).

It is easy to show by structural induction on terms that *tlex* is total on ground terms whenever the $\succ$ is total on Σ, but it has a severe defect: it is not well-founded. For example,

---

[3] As usual, the equivalence $\approx$ associated with a preorder $\preceq$ is defined such that $a \approx b$ iff $a \preceq b$ and $b \preceq a$.

[4] As usual, given a preorder $\preceq$, the strict order $\prec$ is defined such that $a \prec b$ iff $a \preceq b$ and $b \not\preceq a$.

if $a \succ f$ then we have $a \succ^{tlex} fa \succ^{tlex} f^2a \succ^{tlex} \ldots$. The problem arises since it is possible for a term to be strictly smaller than one of its subterms.

The most powerful forms of reduction orderings are based on the relative syntactic simplicity of two terms, i.e., on the notion of a simplification ordering. Although there are many types of simplification orderings, one of the most elegant and useful is the *recursive path ordering*, for short, *rpo*.

Definition 11.21 Let $\preceq$ be a preorder on Σ. The *recursive path ordering* $\preceq_{rpo}$ on $T_\Sigma(\mathcal{X})$, for short, *rpo*, is defined below. Actually, we give a simultaneous recursive definition of $\succeq_{rpo}$, $\succ_{rpo}$, and $\approx_{rpo}$, where $s \succ_{rpo} t$ iff $s \succeq_{rpo} t$ and $s \not\preceq_{rpo} t$, and $s \approx_{rpo} t$ iff $s \succeq_{rpo} t$ and $s \preceq_{rpo} t$.

Then, $f(s_1, \ldots, s_n) \succeq_{rpo} g(t_1, \ldots, t_m)$ holds iff one of the conditions below holds:

(i) $f \approx g$ and $\{s_1, \ldots, s_n\} \succeq_{rpo}^{mult} \{t_1, \ldots, t_m\}$; or
(ii) $f \succ g$ and $f(s_1, \ldots, s_n) \succ_{rpo} t_i$ for *all* i, $1 \leq i \leq m$; or
(iii) $s_i \succeq_{rpo} g(t_1, \ldots, t_m)$ for *some* i, $1 \leq i \leq n$,

where $\succeq_{rpo}^{mult}$ is the extension of $\succeq_{rpo}$ to multisets,[5]

Note that since the preorder $\preceq$ is only defined on Σ, variables are regarded as incomparable symbols. In (ii), the purpose of the condition $f(s_1, \ldots, s_n) \succ_{rpo} t_i$ for *all* i, is to insure that $f(s_1, \ldots, s_n) \succ_{rpo} g(t_1, \ldots, t_m)$.

Theorem 11.22 (Dershowitz, Lescanne) The relation $\succ_{rpo}$ is a simplification ordering stable under substitution. Furthermore, if the strict order $\succ$ is well-founded on Σ, then $\succ_{rpo}$ is well-founded, even when Σ is infinite.

Proof sketch. Proving that *rpo* is a simplification ordering is laborious, especially transitivity. The complete proof can be found in Dershowitz [8]. In order to prove that $\succ_{rpo}$ is well-founded when $\succ$ is well-founded on Σ, it is tempting to apply proposition 4.8 to the preorders $\ll$ and $\preceq_{rpo}$, where $\ll$ is defined such that $s \ll t$ iff $root(s) \preceq root(t)$, since the conditions of this lemma hold. Unfortunately, $\preceq$ is not a *wqo*. However, we can use the idea from theorem 4.10 to extend $\preceq$ to a total well-founded ordering $\leq$. Then, by theorem 4.7, the embedding preorder $\preceq_\leq$ induced by $\leq$ (see definition 4.6) is a *wqo*, and thus, it is well-founded. We can now apply proposition 4.8, which shows that $\leq_{rpo}$ (the *rpo* induced by $\leq$) is well-founded. Finally, we prove by induction on terms that $\leq_{rpo}$ contains $\preceq_{rpo}$, which proves that $\succ_{rpo}$ itself is well-founded. □

[5] Other authors define $\succ_{rpo}^{mult}$ as the multiset extension of the strict order $\succ_{rpo}$, and $s \succeq_{rpo}^{mult} t$ iff $s \succ_{rpo}^{mult} t$ or $s = t$. Our definition is more general.

A proof not involving Kruskal's theorem, but using Zorn's lemma, is given in Lescanne [29]. Of course, a strict order on a finite set is always a *wqo*, and the significance of the second part of the theorem is that it holds even when Σ is infinite.

Example 11.23 Consider the following set of rewrite rules to convert a proposition to disjunctive normal form:

$$\begin{aligned}
\neg(P \vee Q) &\longrightarrow \neg P \wedge \neg Q,\\
\neg(P \wedge Q) &\longrightarrow \neg P \vee \neg Q,\\
P \wedge (Q \vee R) &\longrightarrow (P \wedge Q) \vee (P \wedge R),\\
(P \vee Q) \wedge R &\longrightarrow (P \wedge R) \vee (Q \wedge R),\\
\neg\neg P &\longrightarrow P,\\
P \vee P &\longrightarrow P,\\
P \wedge P &\longrightarrow P.
\end{aligned}$$

This system can be easily shown to be Noetherian using the *rpo* induced by the following ordering on the set of operators: $\neg \succ \wedge \succ \vee$.

It is possible to show that $\succeq_{rpo}$ is total on ground terms whenever $\succ$ is total on Σ. It is also possible to define reduction orderings which are total on ground terms; the problem with $\succeq_{rpo}$ is that it is not a partial order in general, but only a preorder, i.e., the equivalence relation $\approx_{rpo}$ is not necessarily the identity. For example, for any $\succ$ we have $f(a,b) \approx_{rpo} f(b,a)$ but clearly $f(a,b) \neq f(b,a)$. It is easy to show by structural induction on terms, and using only clause (i) of the definition of *rpo* that for any two ground terms $s = f(s_1, \ldots, s_n)$ and $t = g(t_1, \ldots, t_m)$, we have $s \approx_{rpo} t$ iff $f \approx g$ and $s_i \approx_{rpo} t_{\pi(i)}$, for $1 \leq i \leq n$, where π is some permutation of the set $\{1, \ldots, n\}$. (In other words, $s \approx_{rpo} t$ iff s and t are equal up to equivalence of symbols, and up to the permutation of the order of the terms under each function symbol, where the permutation of subterms arises by the comparison of multisets of subterms in clause (i) of the definition.)

This motivates the following definition.

Definition 11.24 For any ordering $\succ$ on Σ, let the term ordering $\succ_{rpol}$ be defined such that $s \succ_{rpol} t$ iff either $s \succ_{rpo} t$ or s and t are ground, $s \approx_{rpo} t$, and $s \succ^{tlex} t$.

Clearly for any total $\succ$ on Σ this is a reduction ordering total on ground terms, since $\succeq_{rpo}$ is total on ground terms and if $s \succeq_{rpo} t$ and $s \preceq_{rpo} t$ then, since $\succ^{tlex}$ is total on ground terms, we must have either $s \succ^{tlex} t$ or $s \prec^{tlex} t$.

Thus, any time the underlying ordering on Σ is total we have a total ordering on T_Σ, even though the ordering may not be total on $T_\Sigma(\mathcal{X})$. This is a major problem with

term orderings: in order to preserve stability under substitution, they must treat variables as incomparable symbols. Thus equations such as commutative axioms (e.g. $f(x,y) \doteq f(y,x)$) can never be oriented.

Warning: It is possible that for R and S rewrite systems on disjoint sets of function (and constant) symbols, both R and S are Noetherian, but $R \cup S$ is not, as shown by the following example due to Toyama.

Example 11.25

$$\begin{aligned} R &= \{f(0,1,z) \to f(z,z,z)\} \\ S &= \{g(x,y) \to x \\ &\qquad g(x,y) \to y\} \end{aligned}$$

Observe that the term $f(g(0,1),g(0,1),g(0,1))$ rewrites to itself:

$$\begin{aligned} f(g(0,1),g(0,1),g(0,1)) &\longrightarrow f(0,g(0,1),g(0,1)) \\ &\longrightarrow f(0,1,g(0,1)) \\ &\longrightarrow f(g(0,1),g(0,1),g(0,1)). \end{aligned}$$

Another interesting kind of term ordering is the *lexicographic path ordering* due to Kamin and Levy.

Definition 11.26 Let $\preceq$ be a preorder on Σ. The *lexicographic path ordering* $\preceq_{lpo}$ on $T_\Sigma(\mathcal{X})$, for short, *lpo*, is defined below. Actually, we give a simultaneous recursive definition of $\succeq_{lpo}$, $\succ_{lpo}$, and $\approx_{lpo}$, where $s \succ_{lpo} t$ iff $s \succeq_{lpo} t$ and $s \not\preceq_{lpo} t$, and $s \approx_{lpo} t$ iff $s \succeq_{lpo} t$ and $s \preceq_{lpo} t$.

Then, $f(s_1, \ldots, s_n) \succeq_{lpo} g(t_1, \ldots, t_m)$ holds iff one of the conditions below holds:

(i) $f \approx g$, $s_1 \approx_{lpo} t_1, \ldots, s_{i-1} \approx_{lpo} t_{i-1}, s_i \succeq_{lpo} t_i$, and $s \succ_{lpo} t_{i+1}, \ldots, s \succ_{lpo} t_n$, for some i, $1 \leq i \leq n$, with $s = f(s_1, \ldots, s_n)$ and $m = n$; or

(ii) $f \succ g$ and $f(s_1, \ldots, s_n) \succ_{lpo} t_i$ for *all* i, $1 \leq i \leq m$; or

(iii) $s_i \succeq_{lpo} g(t_1, \ldots, t_m)$ for *some* i, $1 \leq i \leq n$.

Note that since the preorder $\preceq$ is only defined on Σ, variables are regarded as incomparable symbols. Also, condition (i) is sometimes stated as:

(i') $f \approx g$, $\langle s_1, \ldots, s_n \rangle \succeq^{lex}_{lpo} \langle t_1, \ldots, t_n \rangle$, $m = n$, and $f(s_1, \ldots, s_n) \succ_{lpo} t_i$ for all i, $1 \leq i \leq n$, where $\succeq^{lex}_{lpo}$ is the lexicographic extension of $\succeq_{lpo}$ on n-tuples.[6]

[6] Other authors define $\succ^{lex}_{lpo}$ as the lexicographic extension of the strict order $\succ_{lpo}$, and $s \succeq^{lex}_{lpo} t$ iff $s \succ^{lex}_{lpo} t$ or $s = t$. Our definition is more general.

It is easily seen that (i) and (i') are equivalent. In (i), the purpose of the conditions $s \succ_{lpo} t_{i+1}, \ldots, s \succ_{lpo} t_n$ is to insure that $f(s_1, \ldots, s_n) \succ_{lpo} g(t_1, \ldots, t_m)$ iff $s_i \succ_{lpo} t_i$. Similarly, in (ii), the purpose of the condition $f(s_1, \ldots, s_n) \succ_{lpo} t_i$ for *all* i, is to insure that $f(s_1, \ldots, s_n) \succ_{lpo} g(t_1, \ldots, t_m)$.

Theorem 11.27 (Kamin, Levy) The relation $\succ_{lpo}$ is a simplification ordering stable under substitution. Furthermore, if the strict order $\succ$ is well-founded on Σ, and equivalent symbols have the same rank, then $\succ_{lpo}$ is well-founded, even when Σ is infinite.

Proof. The proof uses the techniques used in theorem 11.22 (Kruskal's theorem). □

As in the previous theorem on *rpo*, the significance of the second part of the theorem is that it holds even when Σ is infinite.

Example 11.28 Consider the following set of rewrite rules for free groups (Knuth and Bendix [27]).

$$\begin{aligned}
(x * y) * z &\longrightarrow x * (y * z),\\
1 * x &\longrightarrow x,\\
I(x) * x &\longrightarrow 1,\\
I(x) * (x * y) &\longrightarrow y,\\
I(1) &\longrightarrow 1,\\
x * 1 &\longrightarrow x,\\
I(I(x)) &\longrightarrow x,\\
x * I(x) &\longrightarrow 1,\\
x * (I(x) * y) &\longrightarrow y,\\
I(x * y) &\longrightarrow I(y) * I(x).
\end{aligned}$$

This system can be easily shown to be Noetherian using the *lpo* induced by the following ordering on the set of operators: $I \succ * \succ 1$.

It is possible to combine *lpo* and *rpo* (Lescanne [32]). It is also possible to define *semantic path orderings* (Kamin, Levy), as opposed to the above *precedence orderings*. Semantic path orderings use orderings on T_Σ rather than orderings on Σ (see Dershowitz [7]).

The relative strength and the ordinals associated with these orderings have been studied by Okada and Dershowitz [37, 9]. For instance, given a strict ordering $\prec$ on a finite set Σ of n elements, then T_Σ under $\prec_{rpo}$ is order-isomorphic to $\varphi_n(0)$, the first n-critical

ordinal.[7] In particular, there is a very natural representation of the ordinals less than ϵ_0 in terms of nested multisets of natural numbers. It is even possible to define an *rpo* whose order-type is Γ_0 (see Dershowitz [7]), if we allow terms to serve as labels.[8]

Okada has showed that it is possible to combine the multiset and lexicographic ordering to obtain term orderings subsuming both the *rpo* and *lpo* ordering, and also obtain a system of notations for the ordinals less than Γ_0 (see Okada [37], and Dershowitz and Okada [9]). Such systems are inspired by some earlier work of Ackermann [1], and we briefly describe one of them.

Let C be a set of constants, and F a set of function symbols (we are not assuming that symbols in F have a fixed arity).

Definition 11.29 For any $n > 0$, the set $A_n(F, C)$ of *generalized Ackermann terms* is defined inductively as follows:

(1) $c \in A_n(F, C)$ whenever $c \in C$.

(2) $f(t_1, \ldots, t_n) \in A_n(F, C)$ whenever $f \in F$ and $t_1, \ldots, t_n \in A_n(F, C)$.

The terms defined by (1) and (2) are called *connected terms*.

(3) $t_1 \# \cdots \# t_m \in A_n(F, C)$, whenever $t_1, \ldots, t_m$ are connected terms in $A_n(F, C)$ ($m \geq 2$).[9]

Given a set $\Sigma = C \cup F$ of labels, note that the set of trees T_Σ can be viewed as a subset of $A_1(F, C)$, using the following representation function:

$rep(c) = c$, when $c \in C$, and

$rep(f(t_1, \ldots, t_m)) = f(rep(t_1) \# \cdots \# rep(t_n))$.

Given a preorder $\preceq$ on $C \cup F$, we define a preorder $\preceq_{ack}$ on $A_n(F, C)$ as follows.

Definition 11.30 The *Ackermann ordering* $\preceq_{ack}$ on $A_n(F, C)$ is defined below. Actually, we give a simultaneous recursively definition of $\succeq_{ack}$, $\succ_{ack}$, and $\approx_{ack}$, where $s \succ_{ack} t$ iff $s \succeq_{ack} t$ and $s \not\preceq_{ack} t$, and $s \approx_{ack} t$ iff $s \succeq_{ack} t$ and $s \preceq_{ack} t$.

(1) If $s, t \in C$, then $s \succeq_{ack} t$ iff $s \succeq t$. If $s \in C$ and $t \notin C$, then $t \succ_{ack} s$ (and $t \not\succeq_{ack} s$).

(2) Let $s = f(s_1, \ldots, s_n)$ and $t = g(t_1, \ldots, t_n)$. Then, $s \succeq_{ack} t$ iff one of the conditions below holds:

---

[7] In this case, Σ is not a ranked alphabet. We allow the symbols in Σ to have varying (finite) ranks.

[8] These terms are formed using a single symbol $\star$ that can assume any finite rank.

[9] Compared to the definition in Dershowitz and Okada [9], we require that $t_1, \ldots, t_m$ are connected terms. This seems cleaner and does not seem to cause any loss of generality.

(i) $f \approx g$, $s_1 \approx_{ack} t_1, \ldots, s_{i-1} \approx_{ack} t_{i-1}, s_i \succeq_{ack} t_i$, and $s \succ_{ack} t_{i+1}, \ldots, s \succ_{ack} t_n$, for some i, $1 \leq i \leq n$; or

(ii) $f \succ g$ and $f(s_1, \ldots, s_n) \succ_{ack} t_i$ for *all* i, $1 \leq i \leq n$; or

(iii) $s_i \succeq_{ack} g(t_1, \ldots, t_n)$ for *some* i, $1 \leq i \leq n$.

(3) Let $s = s_1 \# \cdots \# s_m$ (or $s = s_1$) and $t = t_1 \# \cdots \# t_p$ (or $t = t_1$). Then, $s \succeq_{ack} t$ iff

$$\{s_1, \ldots, s_m\} \succeq_{ack}^{mult} \{t_1, \ldots, t_p\},$$

where $\succeq_{ack}^{mult}$ is the multiset extension of $\succeq_{ack}$.

The following results are stated in Okada [37], and Dershowitz and Okada [9].

Theorem 11.31 (1) If the strict order $\succ$ is well-founded on $C \cup F$, then $\succ_{ack}$ is well-founded on $A_n(F, C)$.

(2) The multiset extension of *rpo* is identical to $\succ_{ack}$ on $A_1(F, C)$.

Proof. The proof of (1) uses the techniques used in theorem 11.22 (Kruskal's theorem). The proof of (2) is straightforward. □

Equivalently, part (2) of theorem 11.31 says that the restriction of $\succeq_{ack}$ to connected terms in $A_1(F, C)$ is identical to *rpo* (we use the representation of terms given by the function *rep* described earlier).

Finally, as noted by Okada, $\langle A_2(\{\psi\}, \{0\}), \preceq_{ack} \rangle$ provides a system of notations for the ordinals less than Γ_0. This is easily seen using theorem 8.12. To show that $\preceq_{ack}$ corresponds to the ordering on the ordinals less than Γ_0, we use lemma 8.11 and lemma 8.10. We can even define a bijection *ord* bewteen the equivalence classes of $A_2(\{\psi\}, \{0\})$ modulo $\approx_{ack}$ and the set of ordinals less than Γ_0 as follows:

$ord(\psi(s, t)) = \psi(ord(s), ord(t))$,

$ord(s_1 \# \cdots \# s_m) = \alpha_1 + \cdots + \alpha_m$,

where $\alpha_1 \geq \ldots \geq \alpha_m$ is the sequence obtained by ordering $\{ord(s_1), \ldots, ord(s_m)\}$ in nonincreasing order.

12 A Glimpse at Hierarchies of Fast and Slow Growing Functions

In this section, we discuss briefly some hierarchies of functions that play an important role in logic because they provide natural classifications of recursive functions according to their computational complexity. It is appropriate to discuss these classes of functions now,

because we have sufficient background about constructive ordinal notations at our disposal. When restricted to the ordinals less than ϵ_0, these hierarchies provide natural rate-of-growth and complexity classifications of the recursive functions which are *provably total* in Peano's arithmetic. In particular, for two of these hierarchies, F_{ϵ_0} and H_{ϵ_0} dominate every such function (for all but finitely many arguments). Thus, the statement "F_{ϵ_0} is total recursive" is true, but not provable in Peano's arithmetic. The relationship with Kruskal's theorem is that the function Fr mentioned in the discussion following theorem 5.2 dominates F_{ϵ_0} (for all but finitely many arguments). In fact, Fr has the rate of growth of a function F_α where α is considerably larger that Γ_0! The results of this section are presented in Cichon and Wainer [4], and Wainer [54], and the reader is referred to these papers for further details.

For ease of understanding, we begin by defining hierarchies indexed by the natural numbers. There are three classes of hierarchies.

1. *Outer iteration hierarchies.*

Let $g\colon \mathbf{N} \to \mathbf{N}$ be a given function. The hierarchy $(g_m)_{m\in\mathbf{N}}$ is defined as follows: For all $n \in \mathbf{N}$,

$$\begin{aligned} g_0(n) &= 0, \\ g_{m+1}(n) &= g(g_m(n)). \end{aligned}$$

The prime example of this kind of hierarchy is the *slow-growing hierarchy* $(G_m)_{m\in\mathbf{N}}$ based on the successor function $g(n) = n + 1$. This hierarchy is actually rather dull when the G_m are indexed by finite ordinals, since $G_m(n) = m$ for all $n \in \mathbf{N}$, but it is much more interesting when the index is an infinite ordinal.

2. *Inner iteration hierarchies.*

Again, let $g\colon \mathbf{N} \to \mathbf{N}$ be a given function. The hierarchy $(h_m)_{m\in\mathbf{N}}$ is defined as follows: For all $n \in \mathbf{N}$,

$$\begin{aligned} h_0(n) &= n, \\ h_{m+1}(n) &= h_m(g(n)). \end{aligned}$$

The prime example of this kind of hierarchy is the *Hardy hierarchy* $(H_m)_{m\in\mathbf{N}}$ based on the successor function $g(n) = n+1$. This hierarchy is also rather dull when the H_m are indexed by finite ordinals, since $H_m(n) = n+m$ for all $n \in \mathbf{N}$, but it is much more interesting when the index is an infinite ordinal.

3. *Fast iteration hierarchies.*

Let $g\colon \mathbf{N} \to \mathbf{N}$ be a given increasing function. The hierarchy $(f_m)_{m\in\mathbf{N}}$ is defined as follows: For all $n \in \mathbf{N}$,

$$\begin{aligned} f_0(n) &= g(n), \\ f_{m+1}(n) &= f_m^n(n), \end{aligned}$$

where $f_m^n(x) = f_m(f_m(\ldots(f_m(x))\ldots))$, the nth iterate of f_m applied to x. The prime example of this kind of hierarchy is the *Grzegorczyk hierarchy* $(F_m)_{m\in\mathbf{N}}$ based on the successor function $g(n) = n + 1$. This hierarchy is not dull even when the F_m are indexed by finite ordinals. Indeed, $F_1(n) = 2n$, $F_2(n) = 2^n n$, and

$$2^{2^{\cdot^{\cdot^{\cdot^{2^n}}}}}\Big\}^n < F_3(n).$$

In order to get functions growing even faster than those obtained so far, we extend these hierarchies to infinite ordinals. The trick is to diagonalize at limit ordinals. However, this presuposes that for each limit ordinal α under consideration, we already have a particular predefined increasing sequence $\alpha[0], \alpha[1], \ldots, \alpha[n], \ldots$, such that $\alpha = \bigsqcup_{n\in\mathbf{N}} \alpha[n]$, a so-called *fundamental sequence*. The point of ordinal notations is that they allow the definition of standard fundamental sequences. This is particularly simple for the ordinals less than ϵ_0, where we can use the Cantor normal form.

For every limit ordinal $\delta < \epsilon_0$, if $\delta = \alpha + \beta$, then $\delta[n] = \alpha + \beta[n]$, if $\delta = \omega^{\alpha+1}$, then $\delta[n] = \omega^\alpha n$ (i.e. $\omega^\alpha + \cdots + \omega^\alpha$ n times), and when $\delta = \omega^\alpha$ for a limit ordinal α, then $\delta[n] = \omega^{\alpha[n]}$. For ϵ_0 itself, we choose $\epsilon_0[0] = 0$, and $\epsilon_0[n+1] = \omega^{\epsilon_0[n]}$.

Fundamental sequences can also be assigned to certain classes of limit ordinals larger than ϵ_0, but this becomes much more complicated. In particular, this can be done for limit ordinals less than Γ_0, using the normal form representation given in theorem 8.2.

Assuming that fundamental sequences have been defined for all limit ordinals in a given subclass $\mathcal{I}$ of $\mathcal{O}$, we extend the definition of the hierarchies as follows.

Definition 12.1 *Outer iteration hierarchies.*

Let $g\colon \mathbf{N} \to \mathbf{N}$ be a given function. The hierarchy $(g_\alpha)_{\alpha\in\mathcal{I}}$ is defined as follows: For all $n \in \mathbf{N}$,

$$\begin{aligned} g_0(n) &= 0, \\ g_{\alpha+1}(n) &= g(g_\alpha(n)), \\ g_\alpha(n) &= g_{\alpha[n]}(n), \end{aligned}$$

where in the last case, α is a limit ordinal. The prime example of this kind of hierarchy is the *slow-growing hierarchy* $(G_\alpha)_{\alpha\in\mathcal{I}}$ based on the successor function $g(n) = n + 1$. This time, we can show that for any n, $g_\alpha(n) = g^{G_\alpha(n)}(0)$, and $G_{\alpha+\beta}(n) = G_\alpha(n) + G_\beta(n)$, from which it follows that $G_{\omega^\alpha}(n) = n^{G_\alpha(n)}$. This means that if α is represented in Cantor normal form, then $G_\alpha(n)$ is the result of replacing ω by n troughout the Cantor normal form! Thus, we have

$$G_{\epsilon_0[n]}(n) = n^{n^{\cdot^{\cdot^{\cdot^{n^n}}}}\Big\}n-1}.$$

Definition 12.2 *Inner iteration hierarchies.*

Again, let $g\colon \mathbf{N} \to \mathbf{N}$ be a given function. The hierarchy $(h_\alpha)_{\alpha\in\mathcal{I}}$ is defined as follows: For all $n \in \mathbf{N}$,

$$\begin{aligned} h_0(n) &= n, \\ h_{\alpha+1}(n) &= h_\alpha(g(n)), \\ h_\alpha(n) &= h_{\alpha[n]}(n), \end{aligned}$$

where in the last case, α is a limit ordinal. The prime example of this kind of hierarchy is the *Hardy hierarchy* $(H_\alpha)_{\alpha\in\mathcal{I}}$ based on the successor function $g(n) = n+1$ (Hardy [20]). It is easy to show that $h_{\alpha+\beta}(n) = h_\alpha(h_\beta(n))$, and so $h_{\omega^{\alpha+1}}(n) = h^n_{\omega^\alpha}(n)$.

Definition 12.3 *Fast iteration hierarchies.*

Let $g\colon \mathbf{N} \to \mathbf{N}$ be a given increasing function. The hierarchy $(f_\alpha)_{\alpha\in\mathcal{I}}$ is defined as follows: For all $n \in \mathbf{N}$,

$$\begin{aligned} f_0(n) &= g(n), \\ f_{\alpha+1}(n) &= f^n_\alpha(n), \\ f_\alpha(n) &= f_{\alpha[n]}(n), \end{aligned}$$

where $f^n_\alpha(x) = f_\alpha(f_\alpha(\ldots(f_\alpha(x))\ldots))$, the nth iterate of f_α applied to x, and in the last case, α is a limit ordinal.

The prime example of this kind of hierarchy is the extended *Grzegorczyk hierarchy* $(F_\alpha)_{\alpha\in\mathcal{I}}$ based on the successor function $g(n) = n + 1$. It is interesting to note that Ackermann's function has rate of growth roughly equivalent to that of F_ω.

It is not difficult to show that $f_\alpha(n) = h_{\omega^\alpha}(n)$. Thus, even though the fast-growing hierarchy seems to grow faster than the inner iteration hierarchy, the h-hierarchy actually "catches up" with the f-hierarchy at ϵ_0, in the sense that

$$f_{\epsilon_0}(n-1) \le h_{\epsilon_0}(n) \le f_{\epsilon_0}(n+1).$$

Given two functions $f, g: \mathbf{N} \to \mathbf{N}$, we say that *$g$ majorizes f (or that g dominates f)* iff there is some $k \in \mathbf{N}$ such that $g(n) > f(n)$ for all $n \geq k$. It is shown in Buchholz and Wainer [3] that F_β majorizes F_α and that H_β majorizes H_α if $\beta > \alpha$. This property can also be shown for the slow-growing hierarchy. Buchholz and Wainer [3] also show that every recursive function provably total in Peano's arithmetic is majorized by some $F_{\alpha+1}$ in the fast-growing hierarchy up to ϵ_0, and that every F_α for $\alpha < \epsilon_0$ is recursive and provably total in *PA*. It follows that F_{ϵ_0} is recursive, but *not* provably total in *PA*. Going back to the function Fr associated with Friedman miniature version of Kruskal's theorem (theorem 5.2), Friedman has shown that Fr majorizes F_{Γ_0}, and in fact, Fr has the rate of growth of a function F_α where α is considerably larger that Γ_0!

We noted that the h-hierarchy catches up with the f-hierarchy at ϵ_0. It is natural to ask whether the slow-growing hierarchy catches up with the fast-growing hierarchy. At first glance, one might be skeptical that this could happen. But large ordinals are tricky objects, and in fact there is an ordinal α such that the slow-growing hierarchy catches up with the fast-growing hierachy.

Theorem 12.4 (Girard) There is an ordinal α such that G_α and F_α have the same rate of growth, in the sense that

$$G_\alpha(n) < F_\alpha(n) < G_\alpha(an + b),$$

for some simple linear function $an + b$. □

This remarkable result was first proved by Girard [17]. The ordinal α for which G_α and F_α have the same rate of growth is nonother than *Howard's ordinal*, another important ordinal occurring in proof theory. Unfortunately, we are not equipped to describe it, even with the apparatus of the normal functions $\varphi(\alpha, \beta)$. Howard's ordinal is greater than Γ_0, and it is denoted by $\varphi_{\epsilon_{\Omega+1}+1}(0)$, where Ω is the least uncountable ordinal, and $\epsilon_{\Omega+1}$ is the least ϵ-number after Ω (so $\epsilon_{\Omega+1} = \Omega^{\Omega^{\Omega^{\cdot^{\cdot^{\cdot}}}}}$). Alternate proofs of this result are given in Cichon and Wainer [4], and Wainer [54] (among others). A fairly simple description of Howard's ordinal is given in Pohlers [41].

Before closing this section, we cannot resist mentioning Goodstein sequences [18], another nice illustration of the representation of ordinals less than ϵ_0 in Cantor normal form.

Let n be any fixed natural number, and consider any natural number a such that

$$a < (n+1)^{(n+1)^{\cdot^{\cdot^{\cdot^{(n+1)}}}}} \Big\}(n+1).$$

We express a in *complete base* $n+1$ by first writing $a = m_0 + m_1(n+1) + \ldots + m_k(n+1)^{a_k}$, where $m_i \leq n$, and $a_i < a_{i+1}$, and then recursively writing each a_i in complete base $n+1$, until all the exponents are $\leq n$. Given a, denote by $rep(a, n+1)$ its associated representation in complete base $n+1$. Given a number a and its representation $rep(a, n+1)$, we denote by $shiftrep(a, n+1)$ the result of replacing $n+1$ by $n+2$ throughout the representation $rep(a, n+1)$, and by $|shiftrep(a, n+1)|$ the numerical value of this new term.

Definition 12.5 The *Goodstein sequence starting with* $a \geq 0$ is defined as follows. Choose n as the least number such that

$$a < (n+1)^{(n+1)^{\cdot^{\cdot^{\cdot^{(n+1)}}}}} \Big\}(n+1).$$

Set $a_0 = a - 1$, and $a_{k+1} = |shiftrep(a_k, n+k+1)| - 1$.

In the above definition, $a - b$ is the usual difference between a and b when $a \geq b$, and it is equal to 0 otherwise. Thus, we obtain a_{k+1} from a_k by changing $n+k+1$ to $n+k+2$ in the representation $rep(a_k, n+k+1)$ of a_k and subtracting 1 from this new value.

Theorem 12.6 (Goodstein, Kirby and Paris) Every Goodstein sequence terminates, that is, there is some k such that $a_k = 0$. Furthermore, the function *Good* such that *Good*$(a) =$ the least k such that $a_k = 0$ is recursive, but it majorizes the function H_{ϵ_0} from the Hardy Hierarchy.

Proof. The proof that every Goodstein sequence terminates is not that difficult. The trick is to associate to each a_k an ordinal $\alpha_k < \epsilon_0$ obtained by replacing $n+k+1$ by ω throughout $rep(a_k, n+k+1)$. Then, it is easy to see that $\alpha_{k+1} < \alpha_k$, and thus, the sequence a_k reaches 0 for some k. The second part of the theorem is due to Kirby and Paris [26]. Another relatively simple proof appears in Buchholz and Wainer [3]. □

Since H_{ϵ_0} is not provably recursive in *PA*, Goodstein's theorem is a statement that is true but not provable in *PA*.

Readers interested in combinatorial independence results are advised to consult the beautiful book on Ramsey theory, by Graham, Rothschild, and Spencer [19].

13 Constructive Proofs of Higman's Lemma

If one looks closely at the proof of Higman's lemma (lemma 3.2), one notices that the proof is not constructive for two reasons:

(1) The proof proceeds by contradiction, and thus it is not intuitionistic.

(2) The definition of a minimal bad sequence is heavily impredicative, as it involves universal quantification over **all** bad sequences.

Thus, it is natural, and as it turns out, quite challenging, to ask whether it is possible to give a constructive (and predicative) proof of Higman's lemma.

In a remarkable (and short) paper, Friedman [15] introduces a new and simple technique, *the A-translation*, which enables him to give simple proofs of the fact that first-order classical Peano arithmetic and classical higher-order arithmetic are conservative over their respective intuitionistic version over Π^0_2-sentences. His technique also yields closure under Markov's rule for several intuitionistic versions of arithmetic (if $\neg\neg\exists x\varphi$ is provable, then $\exists x\varphi$ is also provable, where x is a numeric variable, and φ is a primitive recursive relation). Using Friedman's A-translation technique, it follows that there is an intuitionistic impredicative proof of Higman's lemma. However, it would still be interesting to see whether a constructive (predicative) proof can be extracted *directly* from the classical proof, and Gabriel Stolzenberg was among the first researchers to propose this challenge, and eventually solve it. It turns out that (at least) two constructive (predicative) proofs of a constructive version of Higman's lemma have been given independently by Richman and Stolzenberg [45], and Murthy and Russell [35]. Steve Simpson has proven a related result for the Hilbert's basis theorem [49], and his proof technique seems related to some of the techniques of Richman and Stolzenberg. The significance of having a constructive proof is that one gets an algorithm which, given a constructively (and finitely presented) infinite sequence, yields the lefmost pair of embedded strings. Murthy and Russell [35] do extract such an algorithm using the NuPRL proof development system. The next challenge is to find a constructive proof of Kruskal's theorem.

Acknowledgment: I wish to thank Robert Constable, Thierry Coquand, Nachum Dershowitz, Jean-Yves Girard, Pierre Lescanne, Anil Nerode, Mitsu Okada, Wayne Snyder, Rick Statman, and Gabriel Stolzenberg, for helpful comments and for pointing out related work.

14 References

[1] Ackermann, W. Konstruktiver Aubfau eines Abschnitts der zweiten Cantorschen Zalhenklasse. *Math. Zeit.* 53, 403-413 (1951).

[2] Bachmair, L. *Canonical Equational Proofs*. John Wiley and Sons, New York (1990).

[3] Buchholz, W., and Wainer, S.S. Provably Computable Functions and the Fast Growing Hierarchy. *Logic and Combinatorics*, edited by S. Simpson, Comtemporary Math-

ematics, Vol. 65, AMS (1987), 179-198.

[4] Cichon, E.A., and Wainer, S.S. The Slow-Growing and the Grzegorczyk Hierarchies. *J. of Symbolic Logic* 48(2) (1983), 399-408.

[5] Crossley, J.N., and Bridge Kister, J. Natural Well-Orderings. *Arch. math. Logik* 26 (1986/1987), 57-76.

[6] DeJongh, D.H.J., and Parikh, R. Well partial orderings and hierarchies. *Indagationes Mathematicae* 14 (1977), 195-207.

[7] Dershowitz, N. Termination of Rewriting. *J. Symbolic Computation* (3), 1-2 (1987), 69-116.

[8] Dershowitz, N. Orderings for Term-Rewriting Systems. *TCS* 17(3) (1982), 279-301.

[9] Dershowitz, N., and Okada, M. Proof-theoretic techniques for term rewriting theory. *3rd Annual Symposium on Logic In Computer Science*, IEEE Computer Society, Edinburgh, Scotland, July 1988, 104-111.

[10] Dershowitz, N., and Manna, Z. Proving termination with multiset orderings. *Communications of the ACM* 22, 465-476 (1979).

[11] Dershowitz, N. Completion and its Applications. In *Resolution of Equations in Algebraic Structures*, Vol. 2, Aït-Kaci and Nivat, editors, Academic Press, 31-85 (1989).

[12] Dickson, L.E. Finiteness of the odd perfect and primitive abundant numbers with n distinct prime factors. *Am. J. Math* 35, (1913), 413-426.

[13] Feferman, S. Systems of Predicative Analysis. *J. of Symbolic Logic* 29 (1964), 1-30.

[14] Feferman, S. Proof Theory: A Personnal Report. In [52], 447-485.

[15] Friedman, H. Classically and intuitionistically provably recursive functions. *Higher set theory* (G.H. Müller and Dana S. Scott, editors), Lecture Notes in Mathematics, Vol. 699, Springer-Verlag, Berlin (1978), 21-28.

[16] Friedman, H., McAloon, K., and Simpson, S. A finite combinatorial principle which is equivalent to the 1-consistency of predicative analysis. *Logic Symposion I (Patras, Greece, 1980)*, G. Metakides, editor, North-Holland, Amsterdam, (1982), 197-230.

[17] Girard, J.Y. Π^1_2-logic. *Annals of Mathematical Logic* 21 (1981), 75-219.

[18] Goodstein, R.L. On the restricted ordinal theorem. *J. of Symbolic Logic* 9 (1944), 33-41.

[19] Graham, R.L., Rothschild, B.L., and Spencer, J.H. *Ramsey Theory*, John Wiley & Sons, Inc., 2nd edition, pp. 196 (1990).

[20] Hardy, G.H. A theorem concerning the infinite cardinal numbers. *Quarterly J. Math.* 35 (1904), 87-94.

[21] Harrington, L.A. et al. *Harvey Friedman's Research on the Foundations of Mathematics*. Harrington, Morley, Ščedrov, and Simpson, Editors, North-Holland (1985).

[22] Higman, G. Ordering by divisibility in abstract algebras. *Proc. London Math. Soc.* (3), 2 (1952), 326-336.

[23] Janet, M. Sur les systèmes d'équations aux dérivées partielles. *J. de Mathématiques* III(8) (1920).

[24] Jouannaud, J.P., and Lescanne, P. On multiset orderings. *Information Processing Letters* 15(2), 57-63 (1982).

[25] Kaplanski, I. Ph.D. thesis, 1941.

[26] Kirby L., and Paris, J. Accessible independence results from Peano arithmetic. *Bull. London Math. Soc.* 14 (1982), 285-293.

[27] Knuth, D.E. and Bendix, P.B., "Simple Word Problems in Univeral Algebras," in *Computational Problems in Abstract Algebra*, Leech, J., ed., Pergamon Press (1970).

[28] Kruskal, J.B. Well-quasi-ordering, the tree theorem, and Vázsonyi's conjecture. *Trans. American Math. Soc.* 95 (1960), 210-225.

[29] Lescanne, P. Some properties of decomposition ordering. A simplification ordering to prove termination of rewriting systems. *RAIRO Informatique Théorique* 16(4), 331-347 (1982).

[30] Lescanne, P. Well rewrite orderings. Extended Abstract, Centre de Recherche en Informatique de Nancy, France (1989).

[31] Lescanne, P. Well quasi orderings in a paper by Maurice Janet. *Bulletin of the EATCS*, No. 39, (1989), 185-188.

[32] Lescanne, P. On the recursive decomposition ordering with lexicographical status and other related orderings. To appear in *Journal of Automated Reasoning* (1990).

[33] Levy, J.J. "Kruskaleries et Dershowitzeries", unpublished notes (1981).

[34] Miller, L.W. Normal functions and constructive ordinal notations. *J. of Symbolic Logic* 41(2) (1976), 439-459.

[35] Murthy, C.R., and Russell, J.R. A constructive proof of Higman's lemma. *5th Annual Symposium on Logic In Computer Science*, IEEE Computer Society, Philadelphia, PA, 257-267, June 4-7, 1990.

[36] Nash-Williams, C. St. J. A. On well-quasi-ordering finite trees. *Proc. Cambridge Phil. Soc.* 59 (1963), 833-835.

[37] Okada, M. Ackermann's ordering and its relationship with ordering systems of term rewriting theory. *Proceedings of the 24th Allerton Conference on Communication, Control, and Computing*, Monticello, ILL (1986).

[38] Okada, M., and Takeuti. G. On the theory of quasi ordinal diagrams. *Logic and Combinatorics*, edited by S. Simpson, Comtemporary Mathematics, Vol. 65, AMS (1987), 295-307.

[39] Okada, M. Kruskal-type theorems on labeled finite trees in term-rewriting theory, graph theory, and proof theory. Manuscript (1987).

[40] Okada, M. Quasi-ordinal diagrams and Kruskal-type theorems on labeled finite trees. Manuscript (1987).

[41] Pohlers, W. *Proof Theory, an Introduction.* Lecture Notes in Mathematics No. 1407, Springer Verlag (1989).

[42] Pohlers, W. Proof theory and ordinal analysis. Preprint, MSRI, Berkeley, California (1989)

[43] Puel, L. Bon préordres sur les arbres associés à des ensembles inévitables et preuves de terminaison de systèmes de réécriture. Thèse d'Etat, (1987), Université de Paris VII.

[44] Puel, L. Using unavoidable sets of trees to generalize Kruskal's theorem. Technical Report 86-4, Laboratoire d'Informatique de l'Ecole Normale Supérieure, Paris, France (1986).

[45] Richman, F., and Stolzenberg, G. Well quasi-ordered sets. Technical report submitted for publication, Northeastern University, Boston MA, and Harvard University, Cambridge, MA, April 1990.

[46] Schütte, K. *Proof Theory.* Springer-Verlag (1977).

[47] Simpson, S.G. Nonprovability of certain combinatorial properties of finite trees. In [21], 87-117.

[48] Simpson, S.G. Which set existence axioms are needed to prove the Cauchy/Peano theorem for ordinary differential equations? *J. of Symbolic Logic* 49(3) (1984), 783-802.

[49] Simpson, S.G. Ordinal numbers and the Hilbert basis theorem. *Journal of Symbolic Logic* 53 (1988), 961-964.

[50] Smoryński, C. "Big" News From Archimedes to Friedman. In [21], 353-366.

[51] Smoryński, C. The Varieties of Arboreal Experience. In [21], 381-397.

[52] Takeuti, G. *Proof Theory*. Studies in Logic, Vol. 81, North-Holland, Amsterdam, Second Edition (1987).

[53] Veblen, O. Continuous increasing functions of finite and transfinite ordinals. *Transactions of the American Mathematical Society*, Vol. 9 (1908), 280-292.

[54] Wainer, S.S. Slow Growing Versus Fast Growing. *J. of Symbolic Logic* 54(2) (1989), 608-614.