A model theoretic analysis of the Church-Turing Thesis

by

Csaba Henk

Submitted to
Central European University
Department of Mathematics and its Applications

In partial fulfillment of the requirements for the degree of
Doctor of Philosophy in Mathematics and its Applications

Supervisors: Hajnal Andréka and István Németi

Budapest, Hungary

2009

Abstract

We present recursion theory in terms of hereditarily finite sets. We use this as a basis for a purely model theoretic definition of decidability, using the notion of an end extension. Finally we show an argument for why it is not necessary to postulate the Church-Turing Thesis as a standalone hypothesis, by outlining a much more fundamental and paradigmatic hyptothesis of modern mathematics and tracing back the Church-Turing Thesis to this latter hypothesis.

Acknowledgments

The listing below was written with the intent of being both sound and complete. There is really no point in losing credibility by leaving ends loose along the lines "and let me say thanks to the numerous unnamed one who helped me on the way..." – there weren't numerous one. Still, needless to say, mind is fallible, and at the threshold of Christ's age, I cannot say anything encouraging in this respect – sorrow omissions are not impossible.

I would like to express my gratitude to (in random order):

- Ali Enayat, the only one whom I can call a co-thinker as far as this topic is concerned;
- Gábor Sági, who never hesitated to spare his time and wit with me;
- Elvira Kadvány, without her sparkling energies and agility and kind support I would have had not the faintest chance to get stuff done;
- Gheorghe Morosanu, who has not stopped believing in me even when I had;
- László Csirmaz, who forced me years ago to send him a sketch of the core ideas of this work, without which I would have had a hard time to recall some of the most essential bits of my original ideas, and also for his excellent TEX preprocessor;
- Konrad Zdanowski, who kindly gave me a hand literature-wise;
- my former and current employers, Watt22 Kft. and Zresearch, Inc., who were really supportive and went the furthest in helping to organize my life in a way which made this work possible;
- the Frugalware team for providing me with an OS which Just Works™ (in 96% percent of times);
- Gábor László and Ibolya Horváth who have taken the burden of housekeeping matters off of my shoulder during the course of working on this material;
- Andréka Hajnal and István Németi, my spiritual stepmother and stepfather.

And last but not least, there is a person towards whom I don't think "expressing my gratitude" would be my only – and not even primary – intention; I'd rather say he is the one to whom I owe by completing the Thesis (apart from myself). He is Ervin Győri, who applied a truckload of his personal charisma and severe effort and trust in me in order to get me to Calgary as an exchange student so that I can start working on the Thesis. (And it did help on the long run but that's another story...)

And, oh, yet another person to whom I'd like to express something other than gratitude – my utter eternal love, despite all calamity. She is Anna Szemerey – sure thing, I am to express gratitude to her in several aspects, like teaching me hard lessons on how to face the mirror... but then again, that's another story and has no direct link with this material.

Contents

To my parents, with whom it is so hard to speak about Things That Matter.

Chapter 1

Introduction

I know I'm artificial
But don't put the blame on me
...
My existence is elusive
The kind that is supported
By mechanical resources

X-Ray Spex: Art-I-Ficial

1.1 The elusiveness of computation

Mathematics begat computation. Therefore it was an immediate and natural demand to do mathematics *about* computation. For this purpose, mankind needed a precise mathematical notion of computation. However, the notion of computation – or similar notions, like that of an algorithm or effective decidability – has a peculiar elusive nature: it is easy to grasp intuitively, but there is no obvious way to translate this intuition to mathematical terms. This led to several independent attempts to define formal frameworks of computation, like Turing machines, recursive functions, the Lambda Calculus, register machines, Markov's normal algorithms, just to mention some. As the mathematics of computing has matured, the following two facts have been observed:

- any two of the general purpose formal frameworks for computation has been proven to be equivalent;
- any particular algorithm or effective procedure invented by people was possible to formalize within these frameworks.

Finally a consensus arose that these formal frameworks serve their purpose well, and they are adequate formal counterparts of the intuitive notion of computation. This metamathematical statement has become known as the *Church-Turing Thesis.*[1]

At this point we can observe a methodological handicap. The above summarized way to the establishment of Church-Turing Thesis can be described in methodological terms as a case of *partial induction.* Partial induction is the basic tool for natural sciences – while there is no a priori evidence for the result of letting a piece of stone fall free, we are pretty sure that if we perform this action once again, the stone will hit the ground once again. We *believe* that the stone will hit the ground once again. This belief of the deterministic nature of free fall is strong enough that it seems to be worth to weave a scientific framework around it, which can be used for analyzing such events and predict their outcome. This is how partial induction works, and the case of gravity is quite analogous to the case of computing, as far as the applied methodology is concerned.

On the other hand, there is a great difference between the epistemological status of natural sciences and metamathematics (or philosophy in general). Nature is external to us. Our knowledge of nature is artificial: what we have at hand is a *perception* of nature, still we want to discuss nature as such and not its perception. What is then nature "as such"? Any answer to this is a *construction* of the one who answers. We have to construct notions which can be the subject matter of our thoughts when we intend to think about nature. To be able to qualify things in nature, we perform a somewhat arbitrary projection from the realm of perceptions to the realm of notions. This gives a degree of indirection when

[1]Sometimes also referred to as *Church Thesis*.

speaking of truth with regards to nature. Philosophy, metamathematics and mathematics operate directly with notions – truth with regards to them is much more direct. (By this we do not imply it would be easy or clear or unambiguous or knowable or free of paradoxical traps.)

This indirection, manifested in partial induction seems to be an inappropriate burden on the Church-Turing Thesis. There is a desire to find an adequately direct verification for it. However, the ways one can provide arguments for the Church-Turing Thesis is deeply influenced by the actual approach taken when defining the formal counterpart of computability.

The aim of this Thesis[2] is to put down a formalization of computability which is much better suited for such metamathematical investigations than traditional formalizations. We will precisely verify our formalization on the mathematical side – that is, we will prove its equivalence with the existing formal definitions of computability. We will also make an attempt to articulate why our definition is better suited for metamathematical purposes; we will even expose our own take on this problem and provide arguments for the Church-Turing Thesis which go beyond partial induction. (These latter arguments are of a metamathematical nature, therefore – unlike strictly mathematical proofs – are appropriate subjects of "fear, uncertainty and doubt".)

Below we will give a sketch of our approach to the problem, but let's see first how it was done so far.

1.2 Coping with elusiveness: the traditional approach

Traditionally computability is formalized by defining an ad hoc pseudo-syntactic generative scheme (or something quite similar to that). What do we mean by this?

Syntactic generative schemes are the usual tool for defining formal languages. A simple

[2]Thesis, as "Ph.D. Thesis", and not a "metamathematical postulate", as in Church-Turing Thesis.

example of it is the language of propositional calculus: we fix a set of propositional variables $X, Y, Z, \ldots$, and we declare:

- propositional variables and the symbols **T** and **F** are propositional formulas;
- if Φ, Ψ are propositional formulas, then $\neg(\Phi)$, $(\Phi) \wedge (\Psi)$, $(\Phi) \vee (\Psi)$ are also propositional formulas;
- the set of propositional formulas is the smallest set which has the above two properties.

Formal definitions of computation sound similar. Let's see, e.g., the definition of recursive functions. Recursive functions are functions of $\mathbb{N}^n \to \mathbb{N}$ for some $n \in \mathbb{N}$ (we call such functions *arithmetic functions* here), so that:

- projections, the characteristic function of the "less than" relation, addition and multiplication are recursive functions;
- multi-variable composition of recursive functions is a recursive function;
- the so-called μ operation performed on a recursive function yields a recursive function;
- the set of recursive functions is the smallest set which has the above properties.

Despite all resemblance, there is an essential difference between the above two definitions. Proper syntactic generative schemes, like the one of propositional formulas, consist of generative steps based on purely syntactic criteria (and therefore it's not hard to find out an effective decision procedure which tells us if a given word belongs to the set of words defined by the scheme). Now if we look at recursive functions... there must be a catch somewhere, because it is a well known result of recursion theory that the set of recursive functions is undecidable. The catch is the μ operation – it is a partial operation on arithmetic functions, and the undecidable problem here is whether the μ operation is applicable to a given arithmetic function. This is why we have the "pseudo" attribute over there – these pseudo-syntactic generative schemes specify some syntactic generative rules, but there is always a restriction on them via which a non-trivial model theoretic condition is hardcoded into the definition (and due to the above mentioned undecidability result, such a model theoretic condition *must* be involved).

These schemes are also ad hoc – none of them can be positioned as natural or canonical when compared to the others. (There might be aesthetic preferences, historical conventions or practical arguments which make us use some of them more frequently than the others, but from a metamathematical point of view, none is better than the others.) (Finally we added, "or something quite similar to these". We had to add this – for example, Turing machines aren't strictly defined by a generative scheme, but the basic traits are the same: there is an easy-to-grasp collection of finitary objects, of which some are pointed out as the ones which actually describe a computation with a sensible result, and the criterion for this selection wires model theory into the definition (in case of Turing machines, the so-called *halting problem* serves as such a criterion). And Turing machines are no less an ad hoc choice than the others.[3])

This ad hocness – in conjunction with the above discussed partial inductive method-

[3]Let's also pin it down quickly, some sources make mention of machines that are of a completely different kind. They use a terminology which broadly overlaps with ours, but they are disjoint from us content-wise. We – i.e., them and us – are speaking about different things, but using similar words. This can lead to some confusion if the reader is not careful enough. Just like friends of a girl, wanting to help her, after she has made a bet, somewhat tipsy, that she shall get married in a day – the friends have a common case but they might have different routes. One might want to try to denounce the bet, other might want to mobilize his very special connections, the third one might apply himself for the role of the groom.

Somewhat simplifying the situation: while we we want to argue for the Church-Turing Thesis, these people aim to refute it. Are we then engaged in intense fighting? Not at all. The catch is that we rely on different models of computation. My computer is... I don't have a computer at all, I speak about the purely mental concept of an algorithm. Their computer is a spaceship. They want to use it to get beyond the Church-Turing barrier, cf. [Syr08]. I hope they can get far.

ological "backend" – hurts people. There are attempts to remedy this. For example, it is enough just to read the title of [DG08] – *A natural axiomatization of computability and proof of Church's Thesis* – to see such an intent. But let's see how the authors tackle this problem. They collect some properties of a generic computational device, which are natural and intuitive by their opinion, they define yet another kind of formal machine on this basis, called the "abstract state machine", and they prove that it's equivalent with the well established models of computation:

> "*The Abstract State Machine Theorem states that every classical algorithm is behaviorally equivalent to an abstract state machine. This theorem presupposes three natural postulates about algorithmic computation. Here, we show that augmenting those postulates with an additional requirement regarding basic operations gives a natural axiomatization of computability and a proof of Church's Thesis* [...]"

A similar approach is taken by [Sie02] (although by means of a different computational model):

> "*These investigations aim to provide a characterization of computations by machines that is as general and convincing as that of computations by human computors given by Turing.*"

Not in lack of some malice, we can say that these two papers defeat each others' purpose by the mere fact that they coexist. If their proposed computational models really matched as tightly to the human intuition of computation as the respective authors claim, how could they end up with different results?

1.3 Coping with elusiveness: our take on it

As seen above, we can't get rid of model theory when we try to formalize computation. Applying the old wisdom – if you can't kill the dragon, ride the dragon – this gives a hint for a new approach: let's detach ourselves from generative schemes and abstract machine models, and rely only on what needs to be relied on anyway – model theory. Let's provide a formal notion of computability purely in terms of model theory. If we are doing it well, we succeed to get rid of something else: ad hocness.

In fact, it is easy to test how successful we are in this respect: if our formalization is not ad hoc, then it should correspond clearly to some intuition we have about computing. So, to prove that we are not selling a pig in a poke, let's start at the heart of the matter and unfold what the intuition is we refer to. For the sake of simplicity we will deal with computations which yield a truth value, that is, instead of speaking about effective procedures, which can have any kind of (finitary) object both as input and output, we will discuss *decision procedures* and *decidable properties*.[4]

Better said, instead of computations or procedures, we will use another naive notion. This notion will correspond to decision procedures by-and-large, but it will sport a non-procedural character. We can say that it is a "de-proceduralization" of the idea of computing.

We will speak of *finitary witnesses*. If we ask, "is formula ψ provable?", then a proof of ψ is a finitary witness which verifies the answer "yes". If we ask, "can the plane be tiled with a given tile kit?", then an enumeration of all $2n \times 2n$ tiling attempts (i.e., arbitrary mappings of the unit squares within $[-n, n] \times [-n, n]$ to tile patterns of the kit), equipped with a badly connected tile pair is a finitary witness which verifies the answer "no". If we ask, "is the graph G planar?", then an injective mapping of vertices of G to nodes of a finite grid such that the line segments corresponding to edges of G don't intersect, is a finitary witness which verifies that the answer is "yes"; and, referring to Kuratowski's Theorem, a topological subgraph of G which is isomorphic with the complete graph on five vertices or the complete bipartite graph on six vertices, is a finitary witness which verifies that

[4] This restriction doesn't hurt generality – an effective procedure P can be traced back to the following decision problem: "Is the object in question a pair $\langle i, o \rangle$ such that P produces o if it's given i as its input?"

the answer is "no". In general, by *an instance of property P* we mean P, together with a object x for which it makes sense to ask "does x have P?"; and by a finitary witness (for an instance of some property of finitary objects) we mean a finitary entity which lets us see whether the object in question has the property, "just by looking at their shape as drawn on paper by pencil".

Instead speaking of a "decision procedure for property P", we can speak of "having a finitary witness for each instance of P". If we have a decision procedure for P, then surely we also have finitary witnesses for each instance of P: the procedure, together its run on x, is a finitary witness for P and x. If we happen to know that there is a finitary witness for each instance of P, then by enumerating finitary witness candidates we get a decision procedure for P. (Can such an enumeration be done? It's an interesting question to which one feels like answering "yes", but this naive context is too swampy to commit ourselves to saying yes. This question can be settled by referring to 4.1.2.)

Returning to the realm of precise mathematics, the hand-wavy notion of a finitary witness can give us a hint for how to capture formal decidability in terms of model theory.

First of all, we decide that we work in V_ω, the structure of hereditarily finite sets, which is considered an universal universe of discourse for finitary objects. Properties become first-order formulas of the language $\langle \in \rangle$ (we will use the $\in$ symbol for the membership relation of the object language, as in "$V_\omega \models \neg(x \in x)$", and $\in$ for membership of the metalanguage, as in "$\{1, 2\} \in V_\omega$"). So far so good.

Now think of the above example problems and finitary witnesses. These can be formalized in V_ω – including objects (formulas, tile patterns and partial tilings, graphs and grids), the properties (provability, tiling problem, planarity) and witnesses (proofs, listings of tiling attempts, graph embeddings to grids and topological subgraphs). Let P be one of these problems (or any similar problem the reader can think of). Say by formalizing P in V_ω we get the formula $\varphi(x)$. Pick a finitary object of which we want to ask P, and which happens to have a finitary witness. Say these get represented as $h, w \in V_\omega$, respectively. The question is: how could we express "witness-ness" in model theoretic terms? If w is a self-contained, full-fledged witness, then it is expected to verify/refute φ for h as is, no matter if the rest of the world goes mad. But then, what is w, "as is"? Identity of sets is given by all the things "below" them – their elements, elements of their elements, and so on, until we get to the empty set at the bottom. So we need at least all the contents of the transitive closure of w, and of course, of h. Let $W \subseteq V_\omega$ be a transitive set containing h and w. As w tells everything about h with respect to φ, we expect $W \models \varphi[x/h]$ iff $V_\omega \models \varphi[x/h]$. We expect even more: we expect to get the same truth value even if the rest of the world goes mad. That is, it should be the same if we extend W with new elements, as far as the new elements don't interfere with the identity of h and w, i.e., don't inject elements under them. More formally, we consider those extensions $E \geq W$, which have the following property: if $g \in W$, $e \in E$ and we have $e \mathbin{\boldsymbol{\in}} g$, then $e \in W$ also holds. Such extensions are commonly called *end extensions*. So what we really want from W is the following:

$$\text{for any end extension } E \geq W,\ W \models \varphi[x/h] \iff E \models \varphi[x/h].$$

Along these considerations, we can introduce our candidate for the formal counterpart of a decision procedure:

Definition We say that $\varphi(x)$ is *finitely determined* if for all $h \in V_\omega$ there is a finite transitive subset W of V_ω such that $h \in W$, and for any end extension E of W, we have $W \models \varphi[x/h]$ iff $E \models \varphi[x/h]$. □

Just to make ourselves unambiguous, we recall the following definition:

Definition If $X \subseteq V_\omega$, and $\varphi(x)$ is a formula of one free variable, we say that φ *defines* X if for any $h \in V_\omega$, we have $h \in X$ iff $V \models \varphi[x/h]$. □

Finally we claim that our notion is suitable for giving a model theoretic definition of formal decidability:

Theorem *A set $X \subseteq V_\omega$ is decidable iff it can be defined by a finitely determined formula.*

(See 3.11 for the formal exposition of this theorem.)

1.4 Closing the gap

Let's recap: we have so far introduced the naive notion of a finitary witness, which sort of corresponds to the (naive) notion of a decision procedure, and a model theoretic definition of (formal) decidability. But what about the Church-Turing Thesis?

Our take on it is the following: we simply don't need the Church-Turing Thesis. At least, we don't need to postulate it as a standalone metamathematical principle. It just follows from something else – a much stronger and much more crucial metamathematical principle. A principle on which modern mathematics is based. It's a somewhat hidden paradigm – by an ironic accident of history, while the Church-Turing Thesis got a name after two of the "founding fathers" of modern mathematics, logics and computer science, and enjoys a big publicity, this other principle, despite being much more fundamental, is not only lacking a name, there wasn't even an attempt to find a profound and sustainable wording for it.

This principle is the following: *set theory captures mathematics.* Any naive mathematical construction can be interpreted within the frame of set theory. Set theory is the universal universe of mathematical objects. It is what lets contemporary mathematicians not be entangled with issues about the ontology of functions like our notable predecessors did:

> "*Throughout its early history, the study of trigonometric functions was linked with the very question: what is a function? In the 18th-century, d'Alembert suggested that a curve can only be called a function of a variable when it is governed by a single analytical expression throughout. Euler replied that one should accept more general functions, which may be represented by different laws in different intervals or even*

drawn freely by hand. 'Continuous' meant that a function obeyed a single analytical law, but everything points to the conclusion that both Euler and his interlocutors presupposed continuity, in the sense of Cauchy, for all of the functions they had in mind. The real issue was, then, whether one should admit arbitrary *(continuous) functions. Fourier, with his more sophisticated series in hand, was unequivocally for arbitrary functions, i.e., for admitting the idea that a function is any correspondence by which ordinates are assigned to abscissas; there was no need to assume that the correspondence ought to follow a common law. But he, again, seems to have assumed that one is talking about functions that are, in general, continuous in the modern sense.*

Dirichlet too was radically in favor of the conception of functions as arbitrary laws [...] *here, Dirichlet is defining the notion of a continuous function, and it has been discussed whether he ever seriously entertained the concept of a completely arbitrary function. One may safely assume that he did not see the need to develop a research program on discontinuous functions; this step would only be done after the publication of Riemann's work.* " [Fer07]

Set theory made possible to emerge those disciplines of modern mathematics which deal with abstract structures, like topology or group theory, where the basic objects are structures over arbitrary collections, by letting us not be distracted by questions like "but what exactly are the elements of a group?". This is what lets us not get lost when we want to prove such basic theorems of analysis like the equivalence of sequential continuity and continuity (in terms of epsilons and deltas). This is what lets us not be confused upon meeting weird phenomena like the Banach-Tarski paradox.

However, probably we were a bit cursory when we blamed it on an "ironic accident of history" that the reliance on the possibility of a set theoretic translation of mathematical ideas has remained apocryphal. While one can see informal references to this principle as common sense knowledge at several places – e.g. [Fit07] mentions:

"[. . .] *it is well-known that all mathematics can be developed within the framework of set theory. In particular, finitary mathematical objects fit within the hereditarily finite portion of set theory.*"

– if indeed we want to make this principle explicit in way which we are willing to defend and take responsibility for, we shall see ourselves facing major difficulties. If we say it as loosely as above – "set theory captures mathematics" –, or as [Fit07] said, we cannot avoid the classical paradoxes of naive set theory, like Russel's paradox. At least we have to distinguish between different kinds of mathematical entities: real numbers are not represented in set theory in the same way as set theory itself. We don't get further even if we prohibit explicit self-reference: the previous sentence remains true also if we replace set theory with group theory. And if we let ourselves pulled into this game, and make careful distinctions to not to fall into paradoxical traps, we lose the generality and the power of our statement: we are likely to get stuck with a special case of the original idea, and an elaborate construction of conditions and cases will not be as self-evident as a paradigm is expected to be. It will be ad-hoc (which adjective already enjoys a status which is not far from that of a swear word, as far as the current text is concerned).

A further problem is that there are mathematical entities for which simply there is no way to fit into set theory, like the category of categories, or even more massively layered constructions of category theory. Well, they don't fit, if by set theory we mean ZFC. We can detach ourselves from ZFC – but which set theory to refer to then? NBG? Or one of the extensions of set-theory introduced by categorists themselves to find a "home", like in [ML69]? And how could we ensure that mathematics won't grow over the chosen flavor of set theory? And how could we argue that our choice is not ad-hoc?

At least, there is an item of the list of possible difficulties which we can cross out: the fact that we don't know if ZFC (or other set theories) is consistent (and we will never know it, either because it is not, or because of Gödel's second incompleteness theorem) is definitely not a problem. It's not a problem, because we don't intend to set a new paradigm, what we want is to explore the current paradigm. If ZFC turned out to be inconsistent, we could either get on with a minor adjustment or we would face a paradigm shift, but that is a completely different story. According to the current paradigm, we rely on ZFC, and we don't intend to say it is an inappropriate base.

While we do have an idea about how to attack these difficulties, for now it will suffice that we sketched them. Let's just assume that we succeeded to state the idea of mathematics being captured by set theory in a suitable way; moreover, our statement is "graded", i.e. we are not only saying that mathematics in general can be represented in set theory, our statement also embraces the other part of [Fit07]'s thought: the mathematics of finitary entites can be represented in V_ω. This is enough to establish an argument for the Church-Turing Thesis, as follows.

Say we have a decision procedure for some problem (in the naive sense). It can also be presented as having a finitary witness for each instance of the problem. Due to (graded) set-theoretic capture, the way the finite witnesses verify/refute instances of the problem gets represented in V_ω as a first-order formula; because the witnesses are finitary, this representation shall be finitely determined.[5] And we already know that a definition by a finitely determined formula means formal decidability.

[5] We don't mean that a reference to the loose sketch by which we arrived to finitely determined formlulas from finitary witnessess suffices for accepting this. We will need the exposition of the set theoretic capture, too.

Chapter 2

Elements of recursion theory among hereditarily finite sets

We will expose elements of recursion theory in context of hereditarily finite sets. Classically this is done using arithmetic. The choice of hereditarily finite sets is motivated by the following considerations:

- The expressiveness of set theory. It's a natural process to represent mathematical objects in set theory – this was the primary reason set theory was created for, after all. (This is not a novel idea – [Fit07], [Świ03] has also chosen to use hereditarily finite sets for discussing recursion theory and limitation theorems, and this one was their main motif for doing so.)

- We will rely on finite subuniverses of our universe, so using a relational structure (where every subset is a substructure) is more convenient than using an algebraic structure.

- To be able to discuss the topics mentioned in 1.4, we rely on the very natural way the universe of hereditarily finite sets (the domain of finitary objects) extends to general set theory (a domain of mathematical objects in general).

We will use the first-order language $\langle \in \rangle$, where $\in$ is a binary relation symbol. We will pronounce $\in$ as "membership", and when we discuss set theory, we will assume the $\in$ denotes the set-theoretic membership relation.

Definition 2.1 V_ω is the first-order structure of language $\langle \in \rangle$ the universe of which is the set of all hereditarily finite sets. □

Definition 2.2 If $G_1 = \langle V_1, E_1 \rangle$, $G_2 = \langle V_2, E_2 \rangle$ are digraphs, we say G_2 *is an end extension of* G_1, or G_2 *end extends* G_1, or G_1 *is an initial segment of* G_2, if G_1 is an induced subgraph of G_2 (ie. $V_1 \subseteq V_2$, $E_1 = E_2 \cap (V_1 \times V_1)$), moreover if $v_1 \in V_1$, $v_2 \in V_2$, $\langle v_2, v_1 \rangle \in E_2$, then $v_2 \in V_1$. A graph embedding is an *initial embedding*, if its image is an initial segment of the target graph.

A digraph $G = \langle V, E \rangle$ is said to be *extensional* if each vertex of it is determined by the set of its ingoing neighbours, that is, if $u, v \in G$, and

$$\forall w \in G(\langle w, u \rangle \in E \iff \langle w, v \rangle \in E),$$

then $u = v$. □

Theorem 2.3 *In the category of digraphs, take the diagram which consists of finite extensional DAGs*[1] *and initial embeddings between them.* V_ω *is a colimit of this diagram.*

2.3 is a constructive definition for V_ω in disguise (that is, there are well-known practices for turning such categorical (co)limits to constructions), and we stated it in order to demonstrate that the existence and identity of V_ω doesn't depend on general set theory. So we can put it like this: we have given two definitions for V_ω, one in top-down and one in bottom-up manner.

There is a well-known correspondence between V_ω and $\mathbb{N}$, due to [Ack37]. Let us recall the *cumulative hierarchy*: $V_0 = \varnothing$, and for each $n \in \mathbb{N}$, $V_{n+1} = \mathcal{P}V_n$. These are transitive

[1] DAG stands for "directed acyclic graph".

sets (their elements are also their subsets) and $V_0 \subset V_1 \subset V_2 \ldots$, while $V_\omega = \bigcup_{n<\omega} V_n$. A set is said to be of *rank* n if it is in $V_{n+1} \setminus V_n$. This gives rise to the following ordering:

Definition 2.4 For each n, we define an ordering $<_n$ on V_n inductively: let $<_{n+1}$ be the lexicographic ordering of V_{n+1} with respect to $<_n$. (In other words, if $h \neq g \in V_{n+1}$, then we set $h <_{n+1} g$ if the maximum of their symmetric difference (according to $<_n$) is in g.) It's easy to see that $<_n$ and $<_{n+1}$ agree on V_n, therefore we can unify the $<_n$ relations, thus getting an ordering $<$ of V_ω. We will refer to this ordering as *Ackermann's ordering.*$\square$

Both the natural ordering on $\mathbb{N}$ and Ackermann's ordering on V_ω are discrete total orderings with a smallest element on a countably inifite set, and thus there is a unique isomorphism between them. This is *Ackermann's correspondence* $\mathsf{ack} : \mathbb{N} \to V_\omega$. Say $h \in V_\omega$; consider the characteristic function $\chi_h : \mathbb{N} \to \{0, 1\}$, for which $\chi_h(i) = 1$ if $\mathsf{ack}(i) \in h$ and 0 otherwise. χ_h can be thought of as the binary representation of some number n, that is, $n = \sum\{2^i : \mathsf{ack}(i) \in h\}$. Set $b(h) = n$. This way we defined $b : V_\omega \to \mathbb{N}$. We claim b is the inverse of ack. To see this, first of all, let's investigate the successor function (as of Ackermann's ordering) on V_ω. The successor of $h \in V_\omega$ can be obtained as follows: $h \cup \{m\} \setminus \{x : x < m\}$, where m is the minimal set in V_ω which is not in h. Therefore, if g is the successor of h, the sequence χ_g can be obtained from χ_h by the following rule: find the lowest index i for which $\chi_h(i) = 0$, change the value at i to 1, and for $j < i$ change the value to 0. This is exactly the same rule we increase a number by one in binary representation. So $b(\mathsf{succ}(h)) = \mathsf{succ}(b(h))$ (where succ refers to the respective successor operations), moreover $b(\varnothing) = 0$, which means b is the unique order-isomorphism from V_ω to $\mathbb{N}$. This also gives a characterization of ack: $\mathsf{ack}(i) \in \mathsf{ack}(n)$ iff the i-th bit in the binary representation of n is 1.

Ackermann's correspondence can be used for giving simple definitions for the usual notions of recursion theory in V_ω, most importantly for Σ_1 a.k.a. recursively enumerable sets and for Δ_1 a.k.a. recursive sets. We could just say that $X \subseteq V_\omega$ is recursively enumerable /

recursive, if its inverse image by Ackermann's correspondence is a recursively enumerable / recursive subset of $\mathbb{N}$. However, we are more ambitious than doing this: we intend to define these notions in the context of V_ω, in a self-contained manner, making no reference to their definition in $\mathbb{N}$.

First of all, let's see the definition of Σ_1 formulas.

Definition 2.5 Let L be a first-order language which contains the binary relation symbols $\lhd_1, \dots, \lhd_n$.

If φ is a formula of L, we can use the following abbreviations:

$$(\forall x \lhd_i y)\varphi \text{ for } \forall x\,(x \lhd_i y \to \varphi)\ (1 \le i \le n)$$

and

$$(\exists x \lhd_i y)\varphi \text{ for } \exists x\,(x \lhd_i y \wedge \varphi)\ (1 \le i \le n).$$

We will call these constructs *bounded universal quantification* and *bounded existential quantification*, respectively (we can speak of *bounded quantification* if we don't want to specify which quantifier is involved). A formula which is built up using the Boolean connectives $\wedge, \vee, \neg, \to$ and bounded quantification is called a *bounded formula*.

- A formula φ is a Σ_1 formula if it is of the form $\exists x_1 \exists x_2 \dots \exists x_k \psi$, where ψ is a bounded formula.

- A formula φ is a Σ formula if it can be obtained from quantifier free formulas by means of $\wedge, \vee$, existential quantification and bounded universal quantification.

If the relations $\lhd_1, \dots, \lhd_n$ cannot be derived unambiguously from context, we might use these notions with an extra annotation like "... wrt. $\lhd_1, \dots, \lhd_n$". When we speak of Σ (Σ_1) formulas of a language consisting of binary relations, we will mean it as Σ (Σ_1) wrt. all the relations of the language (unless stated otherwise). $\square$

Remark 2.6 Literally the set of Σ_1 formulas and the set of Σ formulas don't include each other. For example, if ψ is quantifier free, then $\neg(\forall x \lhd_1 y)\psi$ is a Σ_1 formula but not a Σ formula. However, it is easy to see that Σ_1 formulas are also Σ formulas up to logical equivalence (e.g., the above example is equivalent with $(\exists x \lhd_1 y)\neg\psi$, which is a Σ formula). The reverse is not true in general, as we will show later on (see 3.5). $\square$

The problem is that it is not obvious if the expressive power of Σ_1 formulas in V_ω (wrt. $\in$) and in $\mathbb{N}$ (wrt. $<$) is the same. To see that these are actually the same, we should express $<$ in V_ω by means of a Σ_1 formula (although simply Σ_1 is not enough, we will give the exact condition in 2.11), so that we can replace $<$-bounded quantifiers by some complex formula, where bounds are given in terms of $\in$. Also there are several very natural constucts in set theory to express ideas, which are bounded (at least concerning universal quantifiers), so they appear as something like a Σ or Σ_1 formula, but the bounds are not expressed in terms of $\in$, but, say, $\subset$. The theorem below is to get over these concerns.

Given theories Γ, Θ_1, Θ_2, we say that Θ_1 is *equivalent with* Θ_2 *wrt.* Γ, if for each $\vartheta_1 \in \Theta_1$ there is $\vartheta_2 \in \Theta_2$ such that $\Gamma \models \vartheta_1 \leftrightarrow \vartheta_2$ and for each $\vartheta_2 \in \Theta_2$ there is $\vartheta_1 \in \Theta_1$ such that $\Gamma \models \vartheta_1 \leftrightarrow \vartheta_2$. If $\mathfrak{A}$ is a first order structure, then we say that Θ_1 *is equivalent with* Θ_2 *in* $\mathfrak{A}$ if they are equivalent wrt. $\mathsf{Thm}\,\mathfrak{A}$. If $\mathfrak{A}$ is a first-order structure, and R is a relation symbol, then $\langle \mathfrak{A}, R\rangle$ will stand for the first-order structure we get from $\mathfrak{A}$ by extending its language with R (the interpretation of R will follow from the context in our use cases). (The theory of) $\langle \mathfrak{A}, R\rangle$ is a *conservative extension* of (the theory of) $\mathfrak{A}$, i.e., the extension doesn't interfere with the validity of original formulas.

Furthermore, let Tr be the following relation on V_ω:

$$\{\langle x, y\rangle : x \text{ is in the transitive closure of } y\}.$$

That is, $\mathsf{Tr}(a, b)$ holds iff there are sets $a_1, \ldots, a_n \in V_\omega$ such that $a = a_1 \in a_2 \in \ldots \in a_n = b$.

Theorem 2.7 *Take the following set of binary relations on* V_ω*:* $B = \{\in, <, \subset, \mathsf{Tr}\}$.

The following classes of formulas are equivalent in $\langle V_\omega, <, \subset, \mathsf{Tr}\rangle$*:*

- Σ_1 *formulas of the language* $\langle \boldsymbol{\in} \rangle$;

- Σ *formulas of the language* S, *where* $S \subseteq B$ *and* $\boldsymbol{\in} \in S$.

This theorem can be decomposed to the following lemmas:

Lemma 2.8 Σ_1 *and* Σ *formulas are equivalent in* V_ω.

Lemma 2.9 *Let* $\lhd_1, \ldots, \lhd_n$ *be binary relations on* V_ω *such that* Σ *formulas of the language* $\langle \boldsymbol{\in}, \lhd_1, \ldots, \lhd_n \rangle$ *are equivalent with the* Σ *formulas of the language* $\langle \boldsymbol{\in} \rangle$ *in* V_ω; *and let* $\lhd$ *be one of* $<, \subset, \mathsf{Tr}$.

Then Σ *formulas of* $\langle \boldsymbol{\in}, \lhd_1, \ldots, \lhd_n, \lhd \rangle$ *are also equivalent with* Σ *formulas of* $\langle \boldsymbol{\in} \rangle$ *in* V_ω.

Practically we will consider only the $\langle \boldsymbol{\in} \rangle$ vs. $\langle \boldsymbol{\in}, \lhd \rangle$ cases – making mention of $\lhd_1, \ldots, \lhd_n$ all the time would seriously degrade readability.

Proof (2.8). We have to show that if φ is Σ, then it can be transformed to a Σ_1 formula φ' such that $V_\omega \models \varphi \leftrightarrow \varphi'$ (as we noted, this is always possible in the other direction). The transformation can be done inductively, according to the structure of formulas. The only non-trivial case is when φ is of the form $(\forall x \boldsymbol{\in} y)\psi$. Due to the induction, we can assume ψ is some Σ_1 formula. If ψ is bounded, then φ is bounded, too, therefore it is Σ_1 as is. Otherwise ψ is of the form $\exists z\, \vartheta$, i.e. φ is $(\forall x \boldsymbol{\in} y)\exists z\, \vartheta$. Let φ' be $\exists u\, (\forall x \boldsymbol{\in} y)(\exists z \boldsymbol{\in} u)\vartheta$. We claim this φ' is suitable for us. φ' is surely Σ_1, and $\varphi' \to \varphi$ is a tautology, so what we actually have to show is $V_\omega \models \varphi \to \varphi'$.

Consider an evaluation of φ in V_ω which evaluates to "true". There are finitely many members $x_1, \ldots, x_k$ of the value of y. For each x_i there must be a z_i with which ϑ holds. Set $u = \{z_1, \ldots, z_k\}$. This u verifies φ' for this evaluation. Because this can be done for any positive evaluation of φ, we have $V_\omega \models \varphi \to \varphi'$. ■

Definition 2.10 We say $\varphi(x_1,\ldots,x_n)$ *defines* $X \subseteq \mathfrak{A}^n$ (in $\mathfrak{A}$) if for any $\vec{u} \in \mathfrak{A}$: $\vec{u} \in X$ iff $\mathfrak{A} \models \varphi[\vec{x}/\vec{u}]$.

Given a class of formulas Θ (of the language $\langle \in \rangle$ or some extension of it), and $X \subseteq V_\omega^n$, we can simply say "X *is* Θ" instead of "there is a $\vartheta \in \Theta$ such that ϑ defines X in V_ω".

We define $\Pi_1 = \{\neg\varphi : \varphi \in \Sigma_1\}$, and $\Pi = \{\neg\varphi : \varphi \in \Sigma\}$. (They are equivalent in V_ω, because Σ_1 and Σ are equivalent; hence when only equivalence is concerned, Π_1 and Π can be used interchangeably.) We say $X \subseteq V_\omega^n$ is Δ_1, if it is both Σ_1 and Π_1 (in other words, both of X and $V_\omega^n \setminus X$ are Σ_1). □

Lemma 2.11 *Let $\lhd \subseteq V_\omega \times V_\omega$. If $\lhd$ is Δ_1 and $\{\langle y, \{x : x \lhd y\}\rangle : y \in V_\omega\}$ is Σ wrt. $\in$, then Σ formulas of the language $\langle \in, \lhd \rangle$ are equivalent with Σ formulas of the language $\langle \in \rangle$ in V_ω.*

Proof. Let φ be a Σ formula of $\langle \in, \lhd \rangle$. We will transform φ to a Σ formula of $\langle \in \rangle$ inductively according to the structure of formulas.

If φ is quantifier free, we may assume negation is applied only to atomic formulas in it (that is, if $\neg\xi$ is a subformula of φ, then ξ is atomic). Let $\lhd$ be defined by σ and $V_\omega^2 \setminus \lhd$ be defined by τ, where σ, τ are Σ_1 formulas of $\langle \in \rangle$. Replace all subformulas of φ of the form $\neg x \lhd y$ with $\tau(x,y)$, and then replace all remaining occurrences of $x \lhd y$ with $\sigma(x,y)$. This way we got a formula φ' which is equivalent with φ in V_ω, and is a Σ formula of $\langle \in \rangle$.

If φ is of the form $(\forall x \lhd y)\psi$, where ψ is Σ wrt. $\in$, then let $\{\langle y, \{x : x \lhd y\}\rangle : y \in V_\omega\}$ be defined by ϑ, where ϑ is Σ wrt. $\in$. Then $\exists u\,(\vartheta(y,u) \wedge (\forall x \in u)\psi)$ is Σ wrt. $\in$, and is equivalent with φ in V_ω.

If φ is of the form $\exists x\,\psi$, $\psi_1 \vee \psi_2$ or $\psi_1 \wedge \psi_2$, where ψ, ψ_1 and ψ_2 are Σ wrt. $\in$, then φ is also Σ wrt. $\in$, as is. ■

So to see the rest of 2.9, we have to show that the relations $\subset$, Tr, $<$ fulfill the condition of 2.11.

Remark 2.12 Here we digress on the legitimacy of some practical shorthands we would like to use when writing formulas. On the one hand, sometimes we include defined relations in formulas, like "x is transitive", "$x \subseteq y$". These are simple to resolve: they have a straightforward formal meaning (e.g., in the above cases $(\forall y \in x)(\forall z \in y) z \in x$ and $(\forall z \in x) z \in y$, respectively), so one just have to substitute these formal definitions to the formula shorthand in question, with taking care to choose the auxiliary variables so that they don't interfere with other variables in the formula (e.g., by choosing variables which doesn't yet occur in the formula). When we use such shorthands, we will take care to use them in a way that the formal resolution of the relations doesn't interfere with the "status" of the formula (if it looks / is claimed to be Σ, it will be really Σ after resolving the shorthands). Usually this fact will be straightforward to see and we don't discuss it. (Yet in some cases we might see it necessary to discuss.)

Another abbreviation is function application: if we define a function F by a formula, we would like to use the shorthand $F(x)$. This needs a bit more discussion. Let us assume that the function $F(\vec{x})$ is defined by the Σ formula $\psi(\vec{x}, y)$. Let $R(u, v)$ be an atomic formula or a negated atomic formula (in our case we happen to have only binary atomic formulas, so the notation $R(u, v)$ is fine for us), and take the shorthand $R(F(\vec{x}), v)$. This can be written formally as $\exists y\, (\psi(\vec{x}, y) \wedge R(y, v))$, which is also a Σ formula (do this twice for $R(F(\vec{x}), F(\vec{x}')))$. So if we have a Σ formula of the language $\langle \in, F \rangle$ wrt. $\in$ (where we allow only variables to be the bounds in bounded quantifiers, i.e. we do not allow bounded quantifiers of the form $\exists y \in F(\vec{x})$ [2]), then it can be transformed to a Σ formula of the language $\langle \in \rangle$ like in the proof of 2.11: first transform it so that negation is applied only to atomic subformulas, and then use the above formalizations to get rid of F.

There is also a variant when some variable f of the language is guaranteed to be a

[2] In fact, we can as well allow such function bounded quantifiers, but it is not worth to get into this here, as in practice we won't use function shorthands as quantifier bounders. See 2.21 of a more formal take on this kind of formula transformation (including function bounded quantifiers as well).

function (in the set theoretic sense, i.e., a set of ordered pairs with unique second entries for each first entry), and we use shorthand $f(x)$ (where x is guaranteed to be in the domain of f). This can be traced back to the above case by replacing $f(x)$ with $\mathsf{Apply}(f, x)$, where Apply is the function which gives $f(x)$ if f is a function and x is in its domain, and $\varnothing$ otherwise. Moreover, $f(x)$ can be resolved by a bounded formula: Apply itself can be defined with a bounded formula, and the outer existential quantifier (as in $\exists y\ (\psi(\vec{x}, y) \wedge R(y, v)))$ can be bounded by (an ordered pair in) f.

Therefore we can use function shorthands in Σ formulas if the function in question is Σ itself, and we can use function shorthands of the form $f(x)$ in bounded formulas.[3] □

In particular, we can state the following Lemma:

Lemma 2.13 *If $F \subseteq V_\omega^n$ is a function and is Σ, then it is also Δ_1.*

Proof. Let F be defined by the Σ formula $\psi(x_1, \ldots, x_n)$. Then, as F is a function, $V_\omega \models \forall x_1 \ldots \forall x_{n-1}\ \exists!\ x_n \psi(x_1, \ldots, x_n)$, and therefore $\exists x\ \psi(x_1, \ldots, x) \wedge x \neq x_n$ is a Σ formula which defines $V_\omega^n \setminus F$. ■

We will utilize *ordinals*. Recall ordinals are transitive sets within which $\in$ is a total ordering. In V_ω (or in general, in founded set theories), the "ordering" part follows from comparability, so we can define ordinals by the bounded formula $\mathsf{Ord}(n)$

$$(\forall x \in n)(\forall y \in x) y \in n \wedge (\forall x \in n)(\forall y \in n)(x \in y \vee y \in x \vee y = x).$$

[3] There are two usual approaches to such "shorthands": either we take them really as shorthands, and after some discussion we leave it to the reader to convince herself that they can be appropriately resolved without interfering the mathematical context they occur in; or taking a more formal approach, we say that we always write proper formulas, just in some extended language, where switching to the extended language can be done without affecting the mathematical context due to the conservativeness of the extension. Although one might think that the latter approach is cleaner and more precise, we stick with the former one, because in our case it is not true that the latter approach doesn't affect the mathematical context: introducing function symbols changes substructures, and we rely on them.

A nice feature of ordinals is that they can be used for specifying minimums in a bounded manner. Take, e.g., transitive closure. The transitive closure t of x is the minimal transitive superset of x, i.e.

$$x \subseteq t \wedge t \text{ is transitive } \wedge \forall t' \left((x \subseteq t' \wedge t' \text{ is transitive}) \rightarrow t \subseteq t' \right).$$

Here we have a boundless universal quantification on t' and there seems no way to put a limit on it. However, if we have some property P for ordinals, we can specify the minimal ordinal which has P simply as $\mathsf{Ord}(n) \wedge P(n) \wedge (\forall y \in n) \neg P(n)$. So if P can be formalized as a Σ or bounded formula, this one will be a Σ or bounded formula, too. Using Ackermann's ordering, this construction can be "pulled up" to sets other than ordinals (as we will see it in 2.17).

$\mathsf{Ord}_\omega = \{n \in V_\omega : \mathsf{Ord}(n)\}$, ordered by $\in$, forms a discrete total ordering with a smallest element on a countably infinite set, and thus there is a unique order isomorphism $\mathsf{enum} : \langle \mathsf{Ord}_\omega, \in \rangle \rightarrow \langle V_\omega, < \rangle$. This is an ordinal indexed enumeration of V_ω. How could we describe the n-th set? Take the relation

$$En_3 = \{\langle n, e, x \rangle : \mathsf{Ord}(n),\ e = \mathsf{enum} \restriction n,\ e(n-1) = x\}.$$

Utilizing the rule how the Ackermann successor operation works, En_3 can be defined by a bounded formula:

Lemma 2.14 *If n is an ordinal, let n^- denote its predecessor ordinal, i.e.*

$$m = n^- \iff m \in n \wedge (\forall y \in n) \neg m \in y$$

. *Take the following properties of e, n and x:*

- $\mathsf{Ord}(n)$, *e is a function, the domain of e is n, $e(\varnothing) = \varnothing$;*
- *for each $v, w \in n$, $v \neq \varnothing$, where w is $\in$-minimal such that $\neg e(w) \in e(v^-)$, we have:* $(\forall q \in n)(e(q) \in e(v) \leftrightarrow (w = q \vee (w \in q \wedge e(q) \in e(v^-))))$;

- *$e(n^-) = x$;*

Using the standard set-theoretic definitions of a function, its domain, etc., it is straightforward to formalize the above properties in the language $\langle \in \rangle$ by means of bounded formulas. Let their conjunction be $\mathsf{Enum}(n, e, x)$. $\mathsf{Enum}(n, e, x)$ is a bounded definition of En_3.

Now we can prove a practical variant of 2.11.

Lemma 2.15 *Let $\lhd \subseteq <$. If $\lhd$ is Δ_1, then Σ formulas of the language $\langle \in, \lhd \rangle$ are equivalent with the Σ formulas of the language $\langle \in \rangle$ in V_ω.*

Proof. By 2.11, it is enough to show that $S_\lhd = \{\langle y, \{x : x \lhd y\}\rangle : y \in V_\omega\}$ is Σ wrt. $\in$. Since $\lhd$ is Σ,

$$\exists n\, \exists e\, (\exists u \in n)(\mathsf{Enum}(n, e, y) \wedge x = e(u) \wedge x \lhd y)$$

is a Σ formula as well. Since $\lhd \subseteq <$, it defines $S_\lhd$. ■

Proof (2.9 for $\subset$). As $x \subset y$ has a bounded definition, and $V_\omega \models x \subset y \to x < y$, this is an immediate consequence of 2.15. ■

Lemma 2.16 *Ackermann's ordering is Δ_1.*

Proof. It suffices to show that Ackermann's ordering is Σ – as it is a total ordering, we have $V_\omega \models \neg x < y \leftrightarrow (x = y \vee y < x)$, so a definition of $<$ by a Σ formula also provides a Σ definition for its complement.

And then $\exists n\, \exists e\, (\mathsf{Enum}(n, e, y) \wedge (\exists u \in n) e(u) = x)$ is a Σ_1 definition of $<$. ■

Proof (2.9 for $<$). The equivalence immediately follows from 2.16 and 2.15. ■

Now let's see how ordinals can be used for giving Σ formalizations for "minimal set such that…" style constructions.

Lemma 2.17 *Let $\psi(x, y)$ be a Σ formula, and let φ be the formula*

$$\psi(x, y) \wedge \forall y' \, (\psi(x, y') \rightarrow y \subset y').$$

Then φ is equivalent with a Σ formula in V_ω.

Proof. Given that $h \subset g$ implies $h < g$, the Σ formula $\varphi'(x, y)$:

$$\exists n \, \exists e \, (\mathsf{Enum}(n, e, y) \wedge \psi(x, y) \wedge (\forall q \in n)(\psi(x, e(q)) \rightarrow q \text{ is the predecessor ordinal of } n))$$

is equivalent with φ in V_ω. ■

Nb. the above construction resembles the μ operation of arithmetic recursion theory.

Proof (2.9 for Tr***).*** Take the bounded formula $\psi(x, t)$ which says "t is transitive and $x \subseteq t$". For any x, the minimal t such that $\psi(x, t)$ is the transitive closure of x. Hence by applying 2.17 with ψ, we get that the relation $\{\langle x, t \rangle : t \text{ is the transitive closure of } x\}$ is Σ. We can get $\mathsf{Tr}(x, y)$ by one existential quantification on transitive closure, so it is Σ. $\neg \, \mathsf{Tr}(x, y)$ can be directly expressed by the Σ formula

$$\exists t \, (t \text{ is transitive } \wedge y \subseteq t \wedge \neg x \in t).$$

So $\{\langle x, y \rangle : \mathsf{Tr}(x, y)\}$ is Δ_1. In turn, Tr fulfils the conditions of 2.11, which then provides the equivalence we seek. ■

Proof (2.9). We kindly ask the reader to look back and check that we have already shown the equivalence for each Σ variant in question. ■

Proof (2.7). This follows from 2.8 and the iterated application of 2.9. ■

This, in fact, means that Σ_1 formulas in V_ω are as expressive as in $\mathbb{N}$.

Lemma 2.18 *The arithmetic operations $+, \cdot$ (pulled from $\mathbb{N}$ via* ack*) are Δ_1 in V_ω.*

Proof. By 2.13, it is enough to show that they are Σ. First take arithmetic within ordinals: there addition and multiplication have natural definitions in terms of cardinal arithmetic, i.e., addition yields the cardinality of the disjoint union, and multiplication yields the cardinality of Cartesian product. The straightforward formalization of cardinal arithmetic produces Σ formulas. Mapping this to V_ω with enum and turn it to Σ formulas using Enum is left to the reader. ■

Theorem 2.19 Σ_1, Π_1, Δ_1 *in* V_ω *wrt.* $\in$ *correspond to* Σ_1, Π_1, Δ_1 *in* $\mathbb{N}$ *wrt.* $<$ *by* ack.

Proof. By 2.7 we can use "arithmetics style" bounds in V_ω (i.e., where the bound is wrt. Ackermann's ordering), and by 2.12, 2.18 we can also use the arithmetic operations in Σ formulas.

For the other direction we should show that pulling $\in$ over to $\mathbb{N}$ via ack yields a relation which is Δ_1 in $\mathbb{N}$ wrt. $<$. As we noted, a number i will be an "element" of some other number n iff the i-th bit of n in its binary representation is 1. So basically the problem boils down to showing that 2 based exponentiation is Δ_1. This is a very classical fact, it is discussed in details in any textbook on the topic.

For a compact self-contained account on these cross-definability questions, see sections 4.2, 4.3, 4.4 in [Fit07].[4] ■

Now we will give a direct, pseudo-syntactic definition for Δ_1 sets, analogously to the definion of recursive functions in $\mathbb{N}$.

Definition 2.20 If $\psi(\vec{x},\vec{y})$ is a formula, by *ψ-bounded quantification* we mean the constructs

$$\exists\vec{y}\,(\psi(\vec{x},\vec{y})\wedge\varphi)$$

[4] In the setup [Fit07] has chosen, the operator of element injection is part of the formal language of set theory, so his results are not a drop-in replacement for our ones, however, those facts we were now just referring to are covered in details by him.

and

$$\forall\vec{y}\,(\psi(\vec{x},\vec{y})\to\varphi);$$

we will abbreviate these as

$$(\exists\vec{y}\ :\psi(\vec{x},\vec{y}))\varphi$$

and

$$(\forall\vec{y}\ :\psi(\vec{x},\vec{y}))\varphi,$$

respectively.

We will say that ψ *is function-like (in* $\vec{y}$*)* if $V_\omega \models \forall\vec{x}\,\exists!\,\vec{y}\psi(\vec{x},\vec{y})$.

We define the following set RecHF of formulas of the language $\langle\in\rangle$:

- bounded formulas are in RecHF;
- if $\varphi,\psi\in$ RecHF, then $\neg\varphi$, $\varphi\wedge\psi$, $\varphi\vee\psi$, are also in RecHF;
- if $\psi(\vec{x},\vec{y})$ and φ are in RecHF, and ψ is function-like in $\vec{y}$, then $(\exists\vec{y}\ :\psi(\vec{x},\vec{y}))\varphi$ and $(\forall\vec{y}\ :\psi(\vec{x},\vec{y}))\varphi$ are also in RecHF;
- RecHF is the smallest set of formulas which has the above two property. □

Note the following interesting property of function-like bounded quantifiers: if $\psi(\vec{x},\vec{y})$ is function-like in $\vec{y}$, then $V_\omega\models(\exists\vec{y}\ :\psi(\vec{x},\vec{y}))\varphi\leftrightarrow(\forall\vec{y}\ :\psi(\vec{x},\vec{y}))\varphi$. That is, we can freely switch between $\exists$ and $\forall$, if they occur under a function-like bound.

Theorem 2.21 *The sets definable by* RecHF *formulas are exactly the* Δ_1 *sets.*

Proof. To see RecHF $\to\Delta_1$, it suffices to show that RecHF formulas can be transformed to Σ formulas (as RecHF is closed to negation, this also implies that they can be transformed to Π formulas).

Perform the transformation of $\varphi\in$ RecHF by the following procedure *sig*:

- if φ is bounded, then return φ;

- if φ is of the form $\psi * \vartheta$, where $*$ is one of $\wedge$, $\vee$, then return $sig(\psi) * sig(\vartheta)$;
- if φ is of the form $(Q\vec{y} : \psi)\vartheta$, where Q is one of $\exists$, $\forall$, then return $(\exists\vec{y} : sig(\psi))sig(\vartheta)$;
- if φ is of the form $\neg\neg\psi$, then return $sig(\psi)$;
- if φ is of the form $\neg(\psi \wedge \vartheta)$, then return $sig(\neg\psi) \vee sig(\neg\vartheta)$;
- if φ is of the form $\neg(\psi \vee \vartheta)$, then return $sig(\neg\psi) \wedge sig(\neg\vartheta)$;
- if φ is of the form $\neg(Q\vec{y} : \psi)\vartheta$, where Q is one of $\exists$, $\forall$, then return $(\exists\vec{y} : sig(\psi))sig(\neg\vartheta)$.

sig finally ends, as the number of connectives strictly decreases upon its recursive calls. A simple formula induction shows that it returns a Σ formula, and that it performs equivalent transformations (wrt. V_ω).

Now let X be a Δ_1 set, and we want to define it with a RecHF formula. (We will use the same idea as is used in $\mathbb{N}$ for showing that Δ_1 sets are recursive.) Let X be defined by the formula $\exists\vec{x}\,\sigma$, and $V_\omega \setminus X$ be defined by $\exists\vec{x}\,\tau$, where σ and τ are bounded (we might assume that they are prepended by the same set of existential quantifiers).

If we have two variable vectors $\vec{x}$, $\vec{y}$, then by $\vec{x}\,\vec{<}\,\vec{y}$ we mean that $\vec{x}$ is lesser than $\vec{y}$ according to the lexicographic order derived from $<$ (i.e., either $x_1 < y_1$, or $x_1 = y_1$ but $x_2 < y_2$, or $\ldots$). We can use the notation $(\forall\vec{x}\ \vec{<}\,\vec{y})\varphi$ as a shorthand for $(\forall\vec{x} : \vec{x}\,\vec{<}\,\vec{y})\varphi$; it is left to the reader to see that this can also be expressed in terms of proper single-variable bounded quantifiers.

So take the following formula ϑ: $\forall\vec{x}\,(((\sigma \vee \tau) \wedge (\forall\vec{x}\,'\,\vec{<}\,\vec{x})\neg(\sigma \vee \tau)) \rightarrow \sigma)$. This says "the smallest $\vec{x}$ with which either of σ or τ is fulfilled happens to fulfil σ". As σ and τ are exclusive, this is equivalent with $\exists\vec{x}\,\sigma$, i.e. defines X. Let us use ψ to refer to the premise of the inner part of ϑ, i.e. ϑ is $\forall\vec{x}\,(\psi \rightarrow \sigma)$. ϑ can be written as $(\forall\vec{x} : \psi)\sigma$. As the minimum condition in ψ uniquely determines the whole $\vec{x}$ vector, this is in fact a multi-variable quantification with a function-like bound, so ϑ is a RecHF definition of X. ■

Corollary 2.22 *A Δ_1 set can be defined with a formula of the form $(\exists \vec{x} : \psi)\sigma$, where ψ, σ are bounded, and ψ is function-like in $\vec{x}$.*

Proof. The RecHF definition of a Δ_1 set we constructed in the proof of 2.21 is of this particular form (except that we had there $\forall$ instead of $\exists$, but that makes no difference due to function-likeness). ∎

Note that RecHF formulas can also be directly transformed to the above form, by using logically equivalent transformations and the property of function-likeness, without making a reference to Ackermann's ordering.

Chapter 3

Recursion theory and end extensions

Definition 3.1 Let L be a language with a binary relation symbol $\lhd$, and $\mathfrak{A}$, $\mathfrak{B}$ be L-structures. We say $\mathfrak{B}$ *is an end extension of* $\mathfrak{A}$, or $\mathfrak{B}$ *end extends* $\mathfrak{A}$, or $\mathfrak{A}$ *is an initial segment of* $\mathfrak{B}$, if $\mathfrak{A} < \mathfrak{B}$ and for any $a \in \mathfrak{A}$, $b \in \mathfrak{B}$, if $\mathfrak{B} \models b \lhd a$, then $b \in \mathfrak{A}$. We will denote it with $\mathfrak{A} <_e \mathfrak{B}$. If $X \subseteq \mathfrak{A}$, we may speak of the *initial segment generated by* X, i.e. $\bigcap\{U \subseteq \mathfrak{A} : X \subseteq U,\ \mathfrak{A} \restriction U <_e \mathfrak{A}\}$.

If $\varphi(\vec{x})$ is an L-formula, $v : \vec{x} \to \mathfrak{A}$ is an evaluation, we say that $\mathfrak{A}$ *fixes* φ *in* v, if for any $\mathfrak{B}$, $\mathfrak{A} <_e \mathfrak{B}$ we have $\mathfrak{A} \models \varphi[v] \iff \mathfrak{B} \models \varphi[v]$. We say that v *is a positive evaluation for* φ *(in* $\mathfrak{A}$*)*, if $\mathfrak{A} \models \varphi[v]$ (in the other case, of course, we can say that v *is a negative evaluation for* φ).

If $\varphi(x_1, \ldots, x_n)$ is an L-formula, then we use the notation $[\varphi]_{\mathfrak{A}}$ for the n-ary relation defined by φ in $\mathfrak{A}$, that is, $\{\vec{a} \in \mathfrak{A}^n : A \models \varphi[\vec{x}/\vec{a}]\}$. We read out $[\varphi]_{\mathfrak{A}}$ as *the extension of* φ *in* $\mathfrak{A}$. □

End extensions can be used to give a model theoretic description of Σ definability. It is straightforward to see that a Σ formula is fixed in any structure for its positive evaluations, that is, if $\varphi(\vec{x})$ is a Σ formula (wrt. $\lhd$), and $v : \vec{x} \to \mathfrak{A}$ is a positive evaluation for φ, then $\mathfrak{A}$ fixes φ in v. In other words, $\mathfrak{A} <_e \mathfrak{B} \Rightarrow [\varphi]_{\mathfrak{A}} \subseteq [\varphi]_{\mathfrak{B}}$. To be even more succinct, we can refer to this property as "φ is positively preserved by end extensions". The interesting fact is that the converse holds, too:

Theorem 3.2 (Feferman-Marker) *Take the following two formula properties with regard to some formula $\varphi(\vec{x})$:*

1. *φ is Σ;*

2. *for any $v : \vec{x} \to \mathfrak{A}$ positive, $\mathfrak{A}$ fixes φ in v.*

Then 1 $\Rightarrow$ 2, and if the language L is countable, then also 2 $\Rightarrow$ 1.

This theorem is usually referred to as "Feferman's Theorem", because it was him who has shown it in [Fef68], but I also attributed it to Marker, because he was who has given a succinct, elegant model theoretic proof to it in [Mar84].

To getting warmed up, first we show characterizations of this fashion for bounded and Σ_1 formulas, then we will use these notions to redeem a debt of us: to show that Σ_1 formulas and Σ formulas are not equivalent in general.

The following theorems have no central role in the current discourse, so occasionally we allow ourselves to be sketchy.

Theorem 3.3 *Take the following three formula properties with regard to some formula $\varphi(\vec{x})$:*

1. *φ is bounded;*

2. *for any $v : \vec{x} \to \mathfrak{A}$, $\mathfrak{A}$ fixes φ in v.*

3. *for any $v : \vec{x} \to \mathfrak{A}$, the initial segment generated by the range of v fixes φ for v.*

Then 1 $\Rightarrow$ 2 $\iff$ 3, and if the language L is relational, then also {2,3} $\Rightarrow$ 1.

Proof.

$2 \to 3$: Apply 2 with the generated initial segment instead of the original $\mathfrak{A}$.

3 → 2: Fixation is preserved upwards.

1 → 2: Let $\mathfrak{B}$ an arbitrary end extension of $\mathfrak{A}$. By looking at the definition of "the truth value of φ in $\mathfrak{B}$", it is straightforward to see that upon determining this truth value, we make references only to (relations between) elements of $\mathfrak{A}$, as we have only bounded quantifiers. Hence $\mathfrak{A}$ fixes φ for v.

3 → 1: The conditions imply that relativizing φ to its free variables – that is, replace all occurrences of Qx in it with $(Qx : x < x_1 \vee \ldots \vee x < x_n)$ for $Q = \exists, \forall$ – doesn't affect its truth value in v. It it left to the reader to decompose such quantifiers to plain bounded ones. ■

Theorem 3.4 *Take the following two formula properties with regard to some formula $\varphi(\vec{x})$:*

1. *φ is Σ_1;*
2. *for any $v : \vec{x} \to \mathfrak{A}$ positive, there is a finitely generated initial segment of $\mathfrak{A}$ containing v which fixes φ for v.*

Then 1 ⇒ 2, and if the language L is relational, then also 2 ⇒ 1.

Proof.

1 → 2: Take φ; say, it is of the form $\exists \vec{y}\, \vartheta(\vec{y}, \vec{x})$, with ϑ bounded. By the definition of the truth value of a formula by an evaluation, $\mathfrak{A} \models \exists \vec{y}\, \vartheta(\vec{y}, \vec{x})[v]$ implies we will find an evaluation $v' : \vec{x} \cup \vec{y} \to \mathfrak{A}$ which extends v and $\mathfrak{A} \models \vartheta(\vec{y}, \vec{x})[v']$. Let I' be the initial segment generated by the range of v'. By 3.3, I' fixes ϑ for v', so in any $\mathfrak{B}$, $I' <_e \mathfrak{B}$ we will have $\mathfrak{B} \models \vartheta(\vec{y}, \vec{x})[v']$, and in turn $\mathfrak{B} \models \exists \vec{y}\, \vartheta(\vec{y}, \vec{x})[v]$. That is, I' is a suitable finitely generated initial segment.

2 → 1: There will be a bound N, that for any v positive evaluation there will be an initial segment generated by N elements which fixes φ in v. (We mean it as an *absolute* bound, not just with respect to a particular $\mathfrak{A}$!): were there no such a bound, we could

take a sequence of $\langle \mathfrak{A}_i, v_i \rangle$ pairs which are counterexamples for i as a bound; and then an by taking an ultraproduct of these, we could pick elements from there so that no finitely generated initial segment fixes φ in them, which is a contradiction.

Given N, introduce N new variables, relativize φ to its free variables plus the new variables, like we did it in the proof of 3.3, and prefix this formula with existential quantifiers for the new variables. We got this way a Σ_1 formula, and it's easy to see that it's equivalent with φ. ■

The differences in the characterizations of Σ and Σ_1 sort of predict the following statement. Nevertheless, we show it explicitly:

Theorem 3.5 *Σ_1 formulas and Σ formulas are not equivalent in general.*

Proof. Let L be $\langle <, Add \rangle$, where Add is a ternary relation symbol. We will consider Σ formulas wrt. $<$. Ordinals can be thought of as an L structure with $<$ interpreted as ordinal ordering and Add interpreted as $\{\langle x, y, z \rangle : x + y = z\}$ (by "+" we mean ordinal addition; note it is a partial operation if we restrict it to certain ordinals). Let φ be the formula $(\forall x < t) \exists y \, Add(t, x, y)$; it is a Σ formula. We claim it is not equivalent with any Σ_1 formula within ordinals.

$\omega + \omega \models \varphi[t/\omega]$. However, for all $1 \leq n < \omega$, $\omega + n \not\models \varphi[t/\omega]$. On the other hand, take a Σ_1 formula ψ for which the following also holds: $\omega + \omega \models \psi[t/\omega]$. By 3.4, then it gets fixed for $t \mapsto \omega$ in some finitely generated initial segment containing ω, ie. in $\omega + n$ for some $1 \leq n < \omega$. For such an n, $\omega + n \models \psi[t/\omega]$, therefore it is an ordinal wrt. φ and ψ show a different behaviour. ■

Definition 3.6 Let $\mathfrak{A}$ be an L-structure, φ an L-formula, and κ a cardinal.

We say that $\varphi(\vec{x})$ *is positively κ-semidetermined in* $\mathfrak{A}$, if for any positive evaluation $v : \vec{x} \mapsto \mathfrak{A}$ there is $W <_e \mathfrak{A}$ such that it includes the range of v, it fixes φ for v, and $|W| < \kappa$ (we might call such an initial segment a *κ-witness*). Similarly, $\varphi(\vec{x})$ *is negatively*

κ*-semidetermined in* $\mathfrak{A}$, if for any negative evaluation there is a κ-witness. φ is κ*-determined in* $\mathfrak{A}$ if for any evaluation of its free variables there is a κ-witness.

In case of $\kappa = \omega$, we can also use the expressions *positively / negatively finitely semidetermined*, *finite witness for an evaluation*, *finitely determined.*

These notions can also be applied to subsets of $\mathfrak{A}$, relations on $\mathfrak{A}$: *a subset is positively* κ*-semidetermined* if it can be defined by a positively κ-semidetermined formula, etc. $\square$

Remark 3.7 While it is clear that a formula is κ-determined iff it is both positively and negatively κ-semidetermined, this coincidence vanishes if we move over to subsets: in general, the fact that a subset of $\mathfrak{A}$ can be defined both by a positively and a negatively κ-semidetermined formula doesn't imply that it could also be defined by a κ-determined formula. $\square$

Let's turn back now to V_ω.

There is an interesting trade-off between doing recursion theory in $\mathbb{N}$ and in V_ω: as we have seen in the previous chapter, in V_ω we had to work hard to establish a basic toolkit, to introduce an order on it, and enable ourselves to safely define functions by means of taking the minimum of some property. These are for free in $\mathbb{N}$, the total ordering is part of the language. However, when it comes to encoding (finitary) mathematical objects, one has to rely on contrived number theoretical constructs in $\mathbb{N}$, around which it is hard to wrap one's mind. On the other hand, having paid our entrance fee, we can represent things very comfortably and conveniently in V_ω. We don't have to get into details when we seek a set-theoretical representation of sequences, relations, graphs and alike. We can be both hand-wavy and happy. We continue in this spirit.

Take an axiomatic approximation of V_ω, ie. a theory S which implies for any model of it that the model has an initial segment isomorphic to V_ω, and also $V_\omega \models S$. Recall that we have given two definitions to V_ω – one time as a bottom of the set-theoretic universe, and second time as the top of finite extensional DAGs. A suitable S can be derived from both approaches: it can be either produced by trimming down axioms of general set theory, or by devising axioms which ensure that building finite sets on atop of each other never stops, and doesn't go astray (extensionality, acyclicity, etc. is maintained on the run). Either way, the details are left to the reader; for us it suffices to know that S enforces V_ω as an initial segment, and that axioms of S are effectively known (there is an algorithm to decide if a formula belongs to S or not), to which the least problematic way is making S finite.

Then the machinery of predicate calculus, starting from the definition of the syntactic elements and and ending at defining provability from S, can be described in the usual way: appoint a set to be the set of symbols, find out a way for them to be indexed (we need this in order to be able to use an infinite collection of variables), specify syntax rules for sequences of symbols so that we can have formulas, specify rules for sequences of formulas by which they form a proof.

Having done with all this straightforward but tedious work, we can set up a Gödel style encoding of sets and syntactic entities. For this, first choose a definition scheme for hereditarily finite sets, for example:

- let $\gamma_\varnothing(x)$ be the formula $(\forall y_0 \in x) x \neq x$;
- if $u \in V_\omega$ is of rank $n > 0$, let $\gamma_u(x)$ be the formula

$$\bigwedge_{v \in u} (\exists y_n \in x) \gamma_v(y_n) \wedge (\forall y_n \in x) \bigvee_{v \in u} \gamma_v(y_n).$$

Then for each $u \in V_\omega$, u is the only set which satisfies $\gamma_u(x)$ in V_ω. This definition can be extended to syntactic entities: we have defined a way to represent variables, logical constants, formulas, proofs in V_ω, so for a formula φ let γ_φ be γ_f, where f is the hereditarily finite set which represents φ, and so on. These γ formulas let us speak about metatheoretical entities within the theory. However, as the γ formulas are in fact nullary functions, we rather use our shorthands according to 2.12, that is, we will use constant symbols like $\ulcorner u \urcorner$, $\ulcorner \varphi \urcorner$ for the set u and the formula φ.

So then, finally we have a bounded formula $\mathsf{Proof}_S(n,p,f)$ which intuitively means "p is a proof from S of length n and the last formula of it is f". This will be a proper formalization of a proof, ie. $S \vdash \varphi \iff V_\omega \models \exists n\, \exists p\ \mathsf{Proof}_S(n,p,\ulcorner\varphi\urcorner)$. Further constructs of interest:

- $\mathsf{G}(x)$ is the function which "names of the name of a set", i.e. $u \mapsto \ulcorner\gamma_u\urcorner$.
- If q is a variable, $\mathsf{Subs}_q(f,c,g)$ is the following proprerty: "f is $\ulcorner\varphi\urcorner$ for some formula φ, c is $\mathsf{G}(u)$ for some set u, and g is $\ulcorner\varphi(q/\ulcorner u\urcorner)\urcorner$". We will rather use it also in a functional style like $\mathsf{Subs}_q(f,c)$ (we may assign a dummy value to f, c if they don't happen to be the encodings of a formula and a set definition).

Formalizing the construction of γ_u and the syntactic operations on formulas, we can define these functions with Σ formulas – which then implies that they are also Δ_1, by 2.13, and that we can use them in Σ formulas, cf. 2.12.

For the interested reader, [Świ03] gives a nice account of all the hairy details, while succeeding to remain compact.[1]

Theorem 3.8 *The following families of formulas are equivalent in V_ω:*

1. *positively finitely semidetermined formulas;*
2. *formulas which are fixed by V_ω for all their positive evaluations;*
3. *Σ_1 formulas.*

Proof.

$3 \to 1$: This is immediate from 3.4, given that finitely generated initial segments of V_ω are finite.

[1]Just like [Fit07], [Świ03] also includes the element injection operator in the language, but that's just a convenience and does not make a great difference.

$1 \to 2$: If a formula gets fixed on a finite initial segment in some evaluation, then a fortiori gets fixed in V_ω.

$2 \to 3$: The idea is as follows: for a given $\varphi(x)$, take the formula which says "putting x into φ, we get a provable formula"; this is Σ, because provability has been phrased in such a manner; and the conditions will imply the equivalence with φ.

Consider the following situation (at this point we don't yet need to assume that φ is fixed for all its positive evaluations): $u \in V_\omega$, and V_ω fixes $\varphi(x)$ in $x \mapsto u$ where $V_\omega \models \varphi[x/u]$. Take the closed formula $\varphi(\ulcorner u \urcorner)$. Then $V_\omega \models \varphi(\ulcorner u \urcorner)$, and due to the fixation, this will be true in all end extensions of V_ω, including models of S. Therefore $S \models \varphi(\ulcorner u \urcorner)$, so by the completeness theorem, $S \vdash \varphi(\ulcorner u \urcorner)$, then in turn, $V_\omega \models \psi_u$, with ψ_u being $\exists n\, \exists p\ \mathsf{Proof}_S(n, p, \ulcorner \varphi(\ulcorner u \urcorner) \urcorner)$.

This can be done in the other direction, too, and this works regardless of fixation: assume that for some $u \in V_\omega$ $\ V_\omega \models \psi_u$. This means then that $S \vdash \varphi(\ulcorner u \urcorner)$, then $S \models \varphi(\ulcorner u \urcorner)$, and because $V_\omega \models S$, we get that $V_\omega \models \varphi(\ulcorner u \urcorner)$, i.e., $V_\omega \models \varphi[x/u]$.

Now assume that, as in 2, V_ω fixes $\varphi(x)$ in all positive u; then we get that for all $u \in V_\omega$, $V_\omega \models \varphi[x/u] \iff V_\omega \models \psi_u$.

We can take the following tweaked form of the above construction as follows:

$$\text{let } \psi(x) \text{ be:} \quad \exists n \exists p\, \mathsf{Proof}_S(n, p, \mathsf{Subs}_x(\ulcorner \varphi \urcorner, \mathsf{G}(x))).$$

For any particular u, applying $\mathsf{Subs}_x(\ulcorner \varphi \urcorner, \mathsf{G}(x))$ to u will give us back $\ulcorner \varphi(\ulcorner u \urcorner) \urcorner$, so

$$V_\omega \models \psi[x/u] \iff V_\omega \models \psi_u.$$

Putting these together: for $u \in V_\omega$,

$$V_\omega \models \varphi[x/u] \iff V_\omega \models \psi_u \iff V_\omega \models \psi[x/u],$$

ie. $V_\omega \models \varphi \leftrightarrow \psi$. And ψ is Σ. ■

Remark 3.9 My original proof included "brute force" direct construction to show the 1 $\rightarrow$ 3 implication (somewhat resembling the 2 $\rightarrow$ 3 proof given here, however, it was more complex as it involved the encodings of finite witnesses). It was Ali Enayat in [Ena04a] who pointed out that positive fixation by V_ω suffices for being Σ_1. $\square$

Now, in fact, we have already a purely model theoretic characterization of decidability: Σ_1 sets and positively finitely semidetermined sets are the same, and then, of course, Π_1 sets and negatively finitely semidetermined sets are also the same; decidable sets (Δ_1 sets) are the ones which are Σ_1 and Π_1 at the same time, so a set is decidable iff it is both positively finitely semidetermined and negatively finitely semidetermined.

However, this can be strengthened by showing that a set is decidable iff it is finitely determined.

Remark 3.10 (The Enayat correspondence) This result was discussed in the Foundations Mathematics mailing list in June-July 2004. It was Ali Enayat who has took the effort to evaluate this result. Finally his conclusion was:

> [...] Henk asked whether his characterization is new or not. In my judgment the characterization is new, in the sense that it has not appeared in print before, but I also believe that it follows from standard arguments and should therefore be considered folklore. [Ena04b]

He claimed that the above statement can be seen by standard arguments of recursivity theory / theoretical computer science, and he has also given a sketch (also in the mail [Ena04b]) how he thinks it can be shown easily. I did not accept his judgement and in subsequent postings I tried to show that his argumentation is not sufficient to show that decidable sets are finitely determined. However, probably I did not present my case clearly enough, as he did not made further comments on the issue.

Here I try to reconstruct Enayat's argument and make another attempt to point out how his simple approach is leaky. Apart from trying to settle that old debate, I would also like to save the reader from making the same mistake: his reasoning is quite plausible, and the reader should see why we have to choose a more complex route.

Enayat says the following: let D be a decidable set. As it is decidable, there is a Turing machine which outputs 1 on elements of D, and outputs 0 elsewhere. Operation of a Turing machine can be formalized by a bounded formula, i.e., for any Turing machine T there is a bounded formula $\tau_T(n, c, x, o)$, which states the following: "c is a computation of T (i.e., a sequence of states of T where subsequent states adhere to the transition rules of Turing machines), at the 0-th state T has (a representation of) x written on its tape, the n-th state of T is the halting state and at that point o is written on T's tape". Now let T be the machine which decides D. Then D can be defined by the formula $\delta(x)$: $\exists n\, \exists c\, \tau_T(n, c, x, 1)$. Enayat claims that this formula is finitely determined (which then would provide the implication we seek): for a given u, take all the objects (numbers or sets, depending on whether we work in $\mathbb{N}$ or in V_ω) which are referred to during the run of T on u until the halting state, and the initial segment I generated by them will fix δ in u.

This is half-right. It might be the case that I is suitable, I is a finitary witness in some informal sense – just not for δ. That is, I will probably contain everything which is needed to fix an *appropriately chosen* definition of D, but δ is not an appropriate one. This claim sounds a bit vague, as we haven't put down explicitly how exactly δ looks like. However, we had an intent to get to the finitely determined definition of D using "standard arguments". What I am saying is that using the "standardish" way(s) to describe a Turing machine are far from yielding a finitely determined δ.

I cannot present a sharp counterexample to show why δ is not finitely determined, for the same reason as above: we have not given an exact definition of τ_T, and in turn, of δ. I can just point out some fallacies involved in the "standard way" of the construction of τ_T. The problematic case is, of course, the one when u is not in D (given that δ is Σ_1, it is positively finitely semidetermined, and therefore will be fixed by I in u if u is in D). So assume $u \notin D$. Then, of course, $\exists n\, \exists c\, \tau_T(n, c, u, 0)$ holds, and such a c is a kind of finitary witness for u not being in D, but it has little help for us – we have picked δ to define D, and we should see how the falsity of $\delta(u)$ is preserved in end extensions of I. Well, let's try to deconstruct τ a little bit – we haven't defined it in details, but how can τ look like? (At this point we choose to operate in an arithmetical context, where $\mathbb{N}$ is regarded as a relational structure, i.e. addition and multiplication are thought of as ternary relations.) We can take the formula $Tr_T(s, s')$ which defines the state transition of T, the formula $Tp_T(s, q)$, which says that in state s q is written on T's tape, and the formula $H_T(s)$ which says s is a halting state for T. Then $\tau_T(n, c, x, o)$ can be written along the following lines:

$$Tp_T(c(0), x) \wedge Tp_T(c(n), o) \wedge H_T(c(n)) \wedge (\forall m\ : 0 < m \leq n) Tr_T(c(m-1), c(m)).$$

We can forge an end extension of I to subvert this definition. For example, we can add a pseudo-number n to I which we declare to be not comparable to any other element of the extension, then take machine states s, s' such that $Tp_T(s, u)$, and $H_T(s')$, $Tp_T(s', 1)$, and let c be the pseudo-computation $0 \mapsto s$, $n \mapsto s'$. With this c the initial and terminal conditions are fulfilled, because c is chosen so, and the transition conditions trivially hold, as there are no transitions at all: there is no m such that $m < n$. So there is a "computation" of T starting from u where "at the end" we get 1. The falsity of δ is not preserved in this end extension.

One can argue that a more careful definition of τ_T can prevent this trick. (Although the usual careful definition: "if we take the shortest computation of T from x, which ends in a halting state, we find it has 1 written on the tape at the end" has no much help here – it protects only from that type of "forgery" which is performed via reasonably arithmetics-like end extensions, that is, non-standard Peano models, where one can "count up to" numbers (even to non-standard ones) in some sense.[2]) Well, yes, maybe this particular trick. But end extensions of I can be *very* pathological. For an other choice of τ_T we can find out

[2]By the way, this particular definition is resistant to non-standard Peano end extensions as-is, without having to strengthen it with the minimization technique, because T halts on each input after finitely many steps, and the fact that it is stuck in a halting state with no change in the content of the tape inherits

other tricks. Not to mention the fact that the basic constructs of arithmetics and recursion theory will become senseless: in general, induction principle is lost, $<$ will not be an order anymore, addition and multiplication will not be functions, constructs which we use to uniquely define some object will not be unique anymore. So a "standard" definition of a computation of T will go wild, and we just cannot guarantee anything about it in an end extension, in particular, we cannot guarantee that in a pathological end extension the "result" (*some* result – uniqueness is lost!) of the computation on u is not 1.

What we can do is a systematic hardening of our toolset, of which we show that it can resist any pathology. This *can* be done, and this is the actual construction which leads to the desired finitely determined definition of D. But this technique goes beyond the point what can be called "standard recursion theoretic argumentation". □

Theorem 3.11 *Δ_1 sets are the same as finitely determined sets.*

Proof. That finitely determined sets are Δ_1 is an immediate consequence of 3.8: if $X \subset V_\omega$ is finitely determined, then it is both positively and negatively finitely semidetermined, ie. by 3.8, both Σ_1 and Π_1, ie., Δ_1.

For the other direction we have to show how to define a decidable set with a finitely determined formula. Let's start with introducing $\mathsf{HEnum}(n, e, x)$, the "hardened enum formula". It will be a variant of Enum (of 2.14), it will even be equivalent with Enum on V_ω.

The difference is that in HEnum we don't rely on any fact which was guaranteed for us by a set theoretic context – for example, in Enum we did not include constraints like $\neg y \in y$, in HEnum we will. Concretely, $\mathsf{HEnum}(n, e, x)$ states:

upwards. What I'm trying to point out is that we cannot expect in general that minimizing a Σ_1 definition of a Δ_1 set yields a finitely determined formula. Yet it is true that the actual construction in the proof of 3.11 can be regarded as a variant of this minimization technique – but not the one you can find in books, I think.

- n is transitive, and $\in$ on n is a discrete ordering with both a minimal and a maximal element;
- e is a function with domain n, and here we give a full characterization of what being a function means in terms of digraph properties;
- $e(\varnothing) = \varnothing$, e on the maximal element of n gives x, and if $k, m \in n$, with m being the successor of k, then $e(m)$ can be obtained from $e(k)$ using the Ackermann successor rule (which we pin down in details, as we speak of digraphs in general, not sets).

In any $\in$-structure $\mathfrak{A}$, if we have $\mathsf{HEnum}(n, e, x)$ for some elements of $\mathfrak{A}$, then either n is infinite (i.e., its incoming degree is infinite – we cannot speak of cardinality in general), and in this case $\{m : m \in n\}$ has an initial segment isomorphic to ω, the e-image of which is an initial segment of the range of e, isomorphic to V_ω; or there is $u \in V_\omega$ such that the structure $\langle \{v : v \leq u\}, \in, u \rangle$ is isomorphic with $\langle \mathrm{Range}(e), \in, x \rangle$. As we said, we also have $V_\omega \models \mathsf{Enum}(n, e, x) \leftrightarrow \mathsf{HEnum}(n, e, x)$.

Using 2.22, take a definition φ of the given Δ_1 set of the form $(\exists y_0 \dots y_m : \psi(x, \vec{y}))\sigma(x, \vec{y})$, where ψ, σ are bounded, and ψ is function-like in $\vec{y}$. We construct a "hardened" version χ of it as follows. Let $Seq_m(s)$ denote the property that s is a sequence of length $m+1$ (a function with domain $0, \dots, m$). (Again, in this context we choose a formalization of it which fully describes this property as a digraph property; and also for function application). So let χ be

$$\Big(\exists n, e, N, E, Y : \tag{3.1}$$

$$Seq_m(Y) \wedge \tag{3.2}$$

$$\mathsf{HEnum}(n, e, x) \wedge \mathsf{HEnum}(N, E, Y) \wedge \tag{3.3}$$

$$(e \subset E \vee E \subset e \vee e = E) \wedge \tag{3.4}$$

$$\psi(x, Y(0), \ldots, Y(m)) \wedge \tag{3.5}$$

$$\exists! k \in n \; e(k) = x \wedge \tag{3.6}$$

$$\exists! K \in N \big(Seq_m(E(K)) \wedge \psi(x, E(K)(0), \ldots, E(K)(m))\big)\Big) \tag{3.7}$$

$$\sigma(x, Y(0), \ldots, Y(m)). \tag{3.8}$$

In χ we introduce new variables (3.1), one of them is Y which we require to be a sequence of length $m+1$ (3.2), and the others provide a hardened enumeration for x and Y (3.3), these enumerations are comparable (i.e., one extends the other, (3.4)), and we not only require that members of Y satisfy ψ with x (3.5), but we also require Y be minimal in this respect (3.7). Let's refer to this hardened variant of ψ as ψ', and let $\sigma'(x, Y)$ be $\sigma(x, Y(0), \ldots, Y(m))$, i.e. χ can be written as $(\exists n,\ e, N, E, Y : \psi')\sigma'$.

It is clear that φ and χ are equivalent in V_ω: the (hardened) enumeration uniquely exists for any hereditarily finite set, and have the desired properties. Furthermore ψ' is function-like in n, e, N, E, Y, again due to the fact that the enumerations are well-defined. We claim that χ is finitely determined.

χ is positively finitely semidetermined, being Σ_1. What we should see that it's also negatively finitely semidetermined, that is, for an $u \in V_\omega$ such that $V_\omega \not\models \chi[x/u]$, there is an initial segment which fixes χ in $x \mapsto u$. Take the uniquely determined sets n_0, e_0, N_0, E_0, Y_0 with which $V_\omega \models \psi'[u, n_0, e_0, N_0, E_0, Y_0]$. We claim that the initial segment I generated by e and E will fix χ in u (it is clear that $u \in I$, so saying this makes sense; and of course, n, N and Y are also in I).

Let $\mathfrak{E}$ be an end extension of I. We would like to show that if for some n_1, e_1, N_1, E_1, Y_1 $\mathfrak{E} \models \psi'[u, n_1, e_1, N_1, E_1, Y_1]$ holds, then we have $\mathfrak{E} \not\models \sigma'[u, Y_1]$. This would immediately follow if Y_1 were uniquely determined, i.e., $Y_1 = Y_0$, because we know that $I \not\models \sigma'[u, Y_0]$, and due to the boundedness of σ, the rest of $\mathfrak{E}$ would not interfere with the evaluation. Alas, this uniqueness will not be true in general. Rather we intend to prove an isomorphism property as follows: with I_0 being the initial segment of $\mathfrak{E}$ generated by u, Y_0, and I_1 being the initial segment of $\mathfrak{E}$ generated by u, Y_1, the structures $\langle I_0, \in, u, Y_0\rangle$ and $\langle I_1, \in, u, Y_1\rangle$ are isomorphic. This is also sufficient for seeing that $\mathfrak{E} \not\models \sigma'[u, Y_1]$, as the result of an evaluation is isomorphism invariant.

So we will show this isomorphism property. First of all, n_1 and N_1 are (of) finite (incoming degree): if n_1 were infinite, then it would have a complete ω type initial segment, in which the "real index" n' of u would occur (i.e., the one for which

$$V'_\omega \models \mathsf{HEnum}[n', e_1 \restriction n', u]$$

holds, where V'_ω is the e_1-image of the ω-type initial segment of n_1), and then the maximal elements of n' and n_1 would be two different elements of n_1 mapped to u by e_1, which violates (3.6). Similarly, if N_1 were infinite, then it would have a complete ω type initial segment, which would have an N_0-th element N'_0, and $Seq_m(E(K)) \wedge \psi(x, E(K)(0), \ldots, E(K)(m))$ would hold with the maximal element of both of N'_0 and N_1 as K, violating (3.7).

So n_1 and N_1 are finite, and then the ranges of e_1 and E_1 are isomorphic to some initial segment of V_ω. Their union includes I_1, and by (3.4), their union is one of them – that is, we got that an initial segment of V_ω (up to isomorphism) includes I_1, which then implies I_1 itself is isomorphic to an initial segment of V_ω. See the following properties of some I', Y':

- I' end extends the initial segment of V_ω generated by u;
- I' is isomorphic to an initial segment of V_ω;
- $Y' \in I'$ and $\langle I', \in, u, Y'\rangle \models Seq_m(Y') \wedge \psi(u, Y'(0), \ldots, Y'(m))$;

- I' is generated (as an initial segment) by u, Y'.

Both of the pairs I_0, Y_0 and I_1, Y_1 fulfil these properties, and, because of the function-likeness of ψ in $\vec{y}$, there is exactly one initial segment of V_ω with these properties (namely I_0). So the initial segment of V_ω to which I_1 is isomorphic turned out to be I_0. We have proved the desired isomorphism property. ■

Chapter 4

The set theoretic paradigm of mathematics

4.1 Manifestos[1]

This section will feature a sort of "cultural anthropology" of modern mathematics – we will not argue, just declare and illustrate. We just try to portray mathematics and metamathematics as it is perceived through the spectacles of set theory. Can we argue about the validity of such a portrayal on a common ground – or validity it a moot issue, as a portrayal is valid on its own right, as an external projection of the subject who has created the it? (In which case it still can be the submitted to aestethical investigations.) Or somewhere in between, our portrayal is expected to be basically correct but aestethics matters, too? We sidestep these questions.

[1]We suggest the reader to revisit Section 1.4, as this section relies on that material

4.1.1 The Short Manifesto: Take it, your badge

Each mathematical concept can be formalized in some kind of set theory.

Fine, kind reader. You have read it. It is a sentence. It makes sense. It is not saying anything unfamiliar. We have even quoted Fitting in the Introduction saying something like this (and surely we could quote others after doing some research). What is the point then? What is the novelty here? What made it worth to be typed out all caps? Well... the point in typing out it all caps is nothing else but to have it typed out all caps. When African American civil right movements have appeared on the scene, the novel thing was not the fact that they acknowledged that they are black, but that they said it loud and were proud of it.

Recall the issues related to such claims we have mentioned in the Introduction: the worry of being paradoxical and the worry of getting stuck with ad hoc formalizations which fail to be general. Worries which can make even great mathematicians to mumble words in the hope of remaining unheard. When we see Fitting saying "*It is well-known that all mathematics can be developed within the framework of set theory*", we have an impression that feels like adding "*... but I rather leave it to those guys who are cleverer than me, or better educated than me, or has specialized to such things, because it is a touchy topic and it should be given a proper treatment by knowledgeable people*".

We took care to word the Manifesto in a safe way. We don't have to worry about paradoxes and losing generality. See the slight but crucial difference between Fitting's wording and ours: he speaks of *the* framework of set theory, we speak of *some kind* of set theory. We can allow here being vague. Two centuries of Foundations behind us have left a mark on our mathematical consciousness. We do know what a set theory is. No, we are not going back to the era of naive set theory when people were thinking that there is a generally understood (but not formally defined) concept of sets – what we say that there is a generally understood (but not formally defined) concept of formal set theories.

Here is a simple test by which the reader can test herself if she knows what a formal set theory is in general: does the term “post-ZFC set theory” makes sense to you? Surely, there is no such thing, but can you imagine what it could be? Imagine that someone succeeds to find some inconsistency in ZFC – not a bad one, just a tiny one. Imagine that there is a “bug” in ZFC, which renders it inconsistent, but doesn’t force us to throw it away and never look back; it just has to be fixed. Can you imagine how that amended ZFC could look like? If yes, then you know what a set theory is.

By speaking in informal terms about formal theories we escape self-reference, the genesis of all paradoxes. And by not sticking to a particular formal theory we don’t lose generality.

So, this Manifesto is like those rectangular web badges by which people promote the blog engine of their choice on their blogs: “Powered by WordPress”, or “I power Blogger”. Now we can embrace “Powered by Set Theory” – quote it, spread it, blog it, shout it from the rooftops: yes, we can.

4.1.2 The Long Manifesto: The Thing you need to hit the nail on the head

Like the guy who gets his first can of Axe spray when he is just done with the shootings of the hottest newest Axe ad, for which he was recruited to be the poster boy – let’s become familiar with that what we are selling with our shiny new badges.

1. Mathematical language states properties about objects.[2]

2. Formal set theories are theories of some variant of first-order logic which aspire to describe an universe of mathematical objects.[3]

[2]The distinction between objects and properties lies at the linguistic level, it is not an ontological one. What is an object and what is a property: depends on context. Consider the role of “$<$” in “$3 < 5$” and “$<$ is a discrete ordering with a minimal element but without a maximal element”.

[3]We do not require (an intent of) universality from set theories. Partial universes are fine, too. In par-

3. Mathematical objects can be represented in an appropriately chosen formal set theory, as sets. Definitions of objects are represented as bounded formulas.

4. Properties of objects can be represented in an appropriately chosen formal set theory, as first-order formulas. This representation captures the intension of properties, not just their extension.[4]

5. Objects are constructed as a bottom-up hierarchy of subobjects. The inner structure of objects and their relation to their subobjects are represented as bounded formulas.

6. The formal set theories we use in our representations are hierarchical, too (but need not to be constructive). They feature some form of the foundation principle. The identity of sets is determined by their elements (the identity of which is determined by their elements, and so on – but we don't get stuck in an infinite regression).

7. The extension of the formulas [5] we use to represent properties is upward invariant (as long as set theoretical universes are concerned): it depends only on those sets which correspond to the objects the properties are about. It doesn't matter how the higher levels of the hierarchy look like (or whether they exist at all)[6]

ticular, here we will consider V_ω a set theory, too. That said, the minimal set theory, and the unambiguous common part of all set theories.

[4]Intension, and extension, in the Fregeian sense. Consider planarity of graphs: the finite graph properties "G can be embedded to $\mathbb{R} \times \mathbb{R}$ with no intersecting edges" and "G has no topological subgraph isomorphic to K_5 and $K_{3,3}$", have the same extension, but not the same intension. This difference is also preserved in their set theoretic representation, as can be demonstrated by switching set theories: if we cut down ourselves to V_ω, the formula we have for emeddability to the plane sort of loses context and will be identically false (as there is no such thing as $\mathbb{R} \times \mathbb{R}$ in V_ω), while the formula with topological subgraphs will still make sense and will denote the same class of graphs as in "big" set theory.

[5]"Extension", again, in the Fregeian sense, not the result of extending something.

[6]... nevertheless, those higher levels should be worth for being addressed as "higher levels" (of a broader universe). I.e., a mathematical concept which matches this intuition is the extension of $\mathbb{N}$ by non-standard

The first thing we can spot that these declarations reflect a strong preconception – a bottom-up view of mathematics, which view extends to the non-formal part, too. Category theorists might object: their top-down view of the world is essentially different. However, two centuries of Foundations behind us have left a mark on our mathematical consciousness. Even category theorists sport some Pavlovian reflexes when foundational matters are on plate: we have already made a reference to Mac Lane's work on providing a set theoretic foundation for category theory.

Is the world really like this? Well… when your only tool is hammer, you will exhibit a tendency to treat everything as if it were a nail. No problem, you can live with it – just always keep in mind that your fingernails are not the kind of nails your tool should be applied to.

4.2 The Gimmick[7]

So then, kind reader, can we make a deal? Would you mind to sign it over there?… Yes, it is the Long Manifesto, there is no fine print, no strings attached[8]… No? Not yet? Hesitating? Really? Then please consider our very special offer, only for you, only now: if you take the Long Manifesto, then you can also take the Church-Turing Thesis, for free!

Recall finitary witnesses. We state the Church-Turing Thesis in terms of finitary witnesses: *A finitary property expressed in terms of finitary witnesses can be formalized within V_ω as a Σ_1 formula.* (We have sketched in the Introduction how it is a straightforward rephrasing of other forms of the Church-Turing Thesis – regardig the non-formal side of the "equation". With respect to the formal part, we can refer to recursion theory.) And we claim we can deduce it from the Long Manifesto. This deduction, of course, will not be a mathematical proof, as it's not dealing with mathematical entities. That said, we won't

Peano models, and not extensions by the means of arbitrary wild end extensions.

[7]We suggest the reader to revisit Section 1.3, as this section relies on that material

[8]This is, of course, a lie.

even pretend doing a mathematical proof: we will trace back our steps to certain points of the Manifesto only by and large. That is natural – a lot of subtle things are included implicitly in the Manifesto. One cannot expect a Manifesto to be bothered by details. We will speak in the *spirit* of the Manifesto.

Finitary objects are represented in an appropriate set theory by LM:3. Given that they are finitary, their representations will be hereditarily finite sets. Take a finitary property P, expressed in terms of finitary witnesses. By LM:4, it will be represented in set theory by some formula φ. By LM:7, the extension of φ – the class of hereditarily finite sets which satisfy φ – is the same in all set theories. So we can cut down our investigations to V_ω. The fact that P is expressed in terms of finitary witnesses means that it is worded along the following lines: a finitary object x has P iff there is a compound finitary object w which has x as its subobject, where the inner structure of w and its relation to x can be described by some condition D. Formalization preserves intension, as LM:4 says. Therefore on the formal side, φ will be of the form $\exists w\, \delta(w, x)$, where δ is the formalization of the description of the witness, D. By LM:5, δ will be a bounded formula, thus φ, the formal counterpart of P, is a Σ_1 formula.

4.3 Objects in the Rear View Mirror May Appear Closer than They Are

Proper journeys are the ones which end up somewhere, but not quite there where one have planned originally. This is what happened to me while getting here from the Introduction. Looking back, the reader can spot some differences between that what proposed in the Introduction and that what is written here. However, I did not go back and adjust the Introduction to match the actual exposition – the essence of the thoughts presented there is not affected, and this difference provides a handy opportunity for a review with some historical background (of the ideas expressed here). I also use this place to provide amendments to those parts of the Introduction which I see now differently.

This document now consists of two fairly independent parts, the mathematical part (Chapters 2, 3), an exposition of recursion theory in the context of hereditarily finite sets which culminate in the statement and the proof of the equivalence of decidability and definability with finitely determined formulas, and the metamathematical part (Chapters 1, Chapters 4), which discusses methodological issues related to the Church-Turing Thesis and tries to make a folklore paradigm of modern mathematics explicit. Both can be read as self-contained texts.

Originally I planned to have this Thesis more connected. The mathematical and metamathematical ideas I express have come to existence by the same train of thoughts. The argument I wanted to use for deducing the Church-Turing Thesis from the principle of set theoretic capture would have relied on the model theoretic definition of formal decidability. I thought that the model theoretic definition is self-explanatory and it is the "natural" definition (in some sense), and this is what can be stated of formalizations of algorithmic properties, once we accept it as a general principle that naive concepts can formalized in set theory. (By and large, I had the idea that evaluation of such formulas on appropriate finite initial segments of V_ω can is what you get when you formalize algorithmic definitions. This is the idea referred to at the end of 1.4.)

While possibility of converting algorithmic definitions to a finitely determined form might make us reassured that algorithmical problems have a strongly finitistic character, I realized that finitely determined formulas are far from being natural – you just won't dress in an armored diving suit for having an after-lunch seaside walk. The way we express mathematical ideas – including algorithms – does rely on a context, a universe of discourse which looks reasonably sane. This is fairly clearly demonstrated in 3.10.

On the other hand, I realized that Σ_1 formulas are indeed very natural constructs, and are free of the ad hocness of theoretic machine models and quasi-syntactic generative schemes. They clearly correspond to the naive idea of expressing properties in terms of finitary witnesses, which is a naive concept you can trade in for decision procedures without having to worry that you apply too much preconception upon doing so. And, while Σ_1 formulas also have their model theoretic characterization in V_ω – by means of the notion of positively finitely semideterminacy – it is not the model theoretic notion what corresponds to the naive concept of finitary witnesses, but the syntactic notion. Because it is intension is what matters, and intension is reflected by the structure of the formula. So yet again, fate was ironic with me and made fun from the title of my Thesis.

This is good news, in a way: this way my thoughts on Church-Turing Thesis became much more accessible, anyone has the necessary background who is familiar with the basic concepts of mathematical logic and theoretical computer science. It is not necessary to grok the non-established concepts of end extensions, fixation, positive finite semideterminacy and finite determinacy (end extensions are a standard concept but do not belong to the standard curriculum, either).

Finally, a few words about that what I dubbed as the "Long Manifesto", and the deduction of the Church-Turing Thesis from that. I have an idea which aspects of it are the ones which make my readers frown. One such thing might be the style of the deduction I use to get to the Church-Turing Thesis from the terms of the Long Manifesto. It *is* weird. No matter how much emphasis I put on the fact that it is not a mathematical proof, it does have a tone of a mathematical proof (even variables are used in it!). I can imagine that readers of a more technical vein will think it is quackery or as Hungarian speakers would put it, "a ferrule made of wood". I guess there was a time when such a tone was in fashion, when the basics of mathematical logic were laid down, back then not as a formal mathematical theory, but as part of an analytic exploration of the perceived true nature of mathematical concepts, as a kind of "homesteading of the Noosphere", when formal foundations were not yet laid down, or at least there was not an established consensus regarding them. It was a wild frontier where thinkers often had to throw entire books in the dustbin upon realizing that what they do leads to paradoxes. Since then mathematics is considered to be built upon a solid formal base, and similar reasonings either survived in a formal form as mathematical theorems or were regarded vague and ended up on the midden of the history of science. And regarding philosophy – I don't know, I am not familiar enough with that to make statements about contemporary trends, but I suspect that philosophers of these days do not operate via pseudo-formal reasonings. I am unsure, too, how this deduction qualifies methodologically. Nevertheless, while I am open to discussions about it (not just open, even curious!), at this point I do not really care about this issue – it is a sort of take-it-or-leave-it thing, I can do nothing about this. Regarding my personal stance: I feel it less of a quackery than those attempts to promote yet another arbitrary contrived abstract machine model as the Real Thing. I vote for a leaky deduction, rather than a pleasant induction, any day.

I am more concerned about the particular terms I used in that deduction. The reader might have the impression that I imposed an artificial pseudo-semantics on naive concepts, custom-made for my purposes, and then collected those premises which I needed to get where I was aiming, rolled it out under the aegis of the Long Manifesto, and in this puppet theatre I stage a drama entitled "The Deduction of the Church-Turing Thesis". What I admit is the fact that all frameworks are shaped by their applications. It is sure that the Long Manifesto would not consist of those particular terms it actually consists of now if I haven't had the motivation to use it as a set of premises for the Church-Turing Thesis. But I consider this a positive feedback. Proof-reading it, I haven't find a term in it which would suffer of being accidental, beyond the original accidentility of the set theoretic approach (which, of course, proves nothing but my awareness of the issue). For example, making mention of bounded formulas in the Manifesto were particularly suspicious for me – as they are so closely related to Σ_1 formulas, it might seem to be a qualified case of nepotism. But I thought into it yet again, and it seems to be sustainable that objects are defined in terms of bounded formulas. I could not recall a definition of something like an object which would not start with taking a base set and then would continue with cutting out its shape by means of formulas relativized to the base set.[9] Indeed it seems frighteningly plausible that objects are defined and described by means of bounded formulas...

Another dangerous bend shows up when I speak of objects and their subobjects. Maybe I could apply more successfully for being acknowledged as general if I spoke about their inner structure? Well, I just looked back to the Manifesto upon typing the previous sentence and I see did mention “inner stucture”, besides subobjects. So I think it is a harmless customization that I also mention subobjects. Still – when using elaborate expressions like “a compound finitary object w which has x as its subobject”, do I not go as far in arbitrariness as the perpetrators of abstract machines? After thinking it over again: I'm sticking with no. If you are not convinced, just rewind to procedures. Remember: we can turn a procedure p into a finitary witness for x by taking its run p_x on x. Even the notation used here suggests that x is a subobject of the run p_x! Read it again please: “we can turn a procedure p into a finitary witness for x by taking its run p_x on x” – do I have add any explanation to this sentence to have it make sense to you, or you just read it and grasp it? I can imagine an answer to this like “No problem with the sentence, as long as I am concerned; but one who is less familiar with the topic might have problems with objectifying the *run of the procedure*”. Even then, the concept of the run is organically attached to the concept of procedure, and getting in picture is just a matter of mastering the concepts. There is no such arbitrariness here as with Turing machines and friends.

At the end of the day, my conscience is clear with respect to calling for those kind of bids where it is made sure I am the only one who can win a pitch.

In a sense, all these concerns are of minor importance anyway, and any kind of defect

[9] Yet again, category theory. However different is the notion of an object in category theory, it won't have a word here in its own right until it compiles to the machine code of set theory. (And if some day it won't be like this anymore, then category theory shall be a separate paradigm which doesn't interfere with set theory.)

related to the deduction of the Church-Turing Thesis can be deemed accidental / technical. How so? Imagine that your mother serves roasted turkey for dinner, stuffed with apples. You see her as she processes the turkey before roasting it. Later on you meet again with the turkey by the table, roasted golden, waiting for getting sliced. What is inside the bird? Roasted apple. Do you have to wait for the moment when the turkey is opened up and you see it? No, you know it in advance. How so? It's so damn evident – a five year old can tell you that if you roast a turkey with an apple in it, then you get a roasted turkey with a roasted apple in it. Similarly, once you got reassured that you can formalize mathematical properties into first-order formulas in general, and took a look at finitary properties in terms finitary witnesses in particular, the proper question is not "what kind of formulas can we formalize them into?", but "how on earth could we get anything but Σ_1 formulas on V_ω?"

It is true that the turkey metaphor features a good deal of oversimplication. Nevertheless, when we match Church-Turing Thesis against the turkey and apple story, we can see which component of the kitchen life is missing on our side: roasting. We act like a sick gourmand who is told away from roasted food by his doctor, and then he creeps away by night to roast in secret.

This is the real problem. And the part of this Thesis which addresses this problem is the Short Manifesto.

Therefore probably the Short Manifesto is the greatest achievement of this work.

Chapter 5

Possible further directions

Here we just give a random burst of ideas and problems related to the topic of this Thesis, which could be subject of further research.

- Can fixation, semideterminacy, determinacy be applied to/in generlized recursion theory in a sensible way? What do get if apply these to classes of problems of various Turing degrees?

- Where do we get if we change the κ? That is, if we investigate κ-semideterminacy, κ-determinacy. Does it has any use in set theory? Are there non-trivial examples for κ-determinacy at all?

- The family of theorems which establish a correspondence between formula classes and model classes are called Birkhoff-style theorems. There is a trend of unifying such proofs in category theoretical frameworks, see eg. [NS82]. Could we reconstruct the Feferman-Marker Theorem as a simple application of such a framework?

- Relying on 3.11 we can construct a complexity measure based on finite determinacy: for a finitely determined formula $\varphi(x)$ and $u \in \omega$, let

$$\mathsf{cfix}_\varphi(u) = \min\{|I| : u \in I, I \text{ fixes } \varphi \text{ in } u\}.$$

Now if F is a class of $\mathbb{N} \longrightarrow \mathbb{N}$ functions, let $\mathsf{CFix}\text{-}F$ be the class of those problems (decidable sets) which can be defined by some finitely determined formula φ such that $\mathsf{cfix}_\varphi \in F$. For example, it can be shown that

$$\mathsf{NP} \cup \mathsf{coNP} \subseteq \mathsf{CFix\text{-}P} \subseteq \mathsf{PSpace}\,.$$

What else can we say about this complexity measure? And if me measure via the rank of I, rather than its cardinality?

- Why first-order logic? This is a very classic question, and has been discussed by many authors; and anyone who dips his or her finger into mathematical logic will have a feeling that it is a distinguished logical system. Still, a succinct, to-the-point answer to this question has not yet been born (or at least I know of none) – one which can help me out on the day when my child will come up to me, eyes open wide, to ask: "Dad, why first-order logic?"

Bibliography

[Ack37] W. Ackermann. Die widerspruchsfreiheit der allgemeinen mengenlhere. *Mathematische Annalen*, 114:305–315, 1937.

[DG08] Nachum Dershowitz and Yuri Gurevich. A natural axiomatization of computability and proof of Church's Thesis. *Bull. Symbolic Logic*, 14(3):299–350, 2008.

[Ena04a] Ali Enayat. Re: A new characterization of recursivity. `http://www.cs.nyu.edu/pipermail/fom/2004-June/008247.html`, 2 Jun 2004. Foundations of Mathematics mailing list.

[Ena04b] Ali Enayat. Re: A new characterization of recursivity. `http://www.cs.nyu.edu/pipermail/fom/2004-July/008305.html`, 1 Jul 2004. Foundations of Mathematics mailing list.

[Fef68] Solomon Feferman. Persistent and invariant formulas for outer extensions. *Compositio Math.*, 20:29–52 (1968), 1968.

[Fer07] José Ferreirós. *Labyrinth of thought*. Birkhäuser Verlag, Basel, second edition, 2007. A history of set theory and its role in modern mathematics.

[Fit07] Melvin Fitting. *Incompleteness in the land of sets*, volume 5 of *Studies in Logic (London)*. College Publications, London, 2007.

[Mar84] David Marker. A model theoretic proof of Feferman's preservation theorem. *Notre Dame J. Formal Logic*, 25(3):213–216, 1984.

[ML69] Saunders Mac Lane. One universe as a foundation for category theory. In *Reports of the Midwest Category Seminar. III*, pages 192–200. Springer, Berlin, 1969.

[NS82] István Németi and Ildikó Sain. Cone-implicational subcategories and some Birkhoff-type theorems. In *Universal algebra (Esztergom, 1977)*, volume 29 of *Colloq. Math. Soc. János Bolyai*, pages 535–578. North-Holland, Amsterdam, 1982.

[Sie02] Wilfried Sieg. Calculations by man and machine: conceptual analysis. In *Reflections on the foundations of mathematics (Stanford, CA, 1998)*, volume 15 of *Lect. Notes Log.*, pages 390–409. Assoc. Symbol. Logic, Urbana, IL, 2002.

[Świ03] S. Świerczkowski. Finite sets and Gödel's incompleteness theorems. *Dissertationes Math. (Rozprawy Mat.)*, 422:58, 2003.

[Syr08] Apostolos Syropoulos. *Hypercomputation*. Springer, New York, 2008. Computing beyond the Church-Turing barrier.