Introduction

Propositional proof complexity studies the lengths of propositional proofs or equivalently the time complexity of non-deterministic algorithms accepting some coNP-complete set. The main problem is the NP versus coNP problem, a question whether the computational complexity class NP is closed under complementation. Central objects studied are propositional proof systems (non-deterministic algorithms accepting the set of propositional tautologies). Time lower bounds then correspond to lengths-of-proofs lower bounds.

Bounded arithmetic is a generic name for a collection of first-order and second-order theories of arithmetic linked to propositional proof systems (and to a variety of other computational complexity topics). The qualification *bounded* refers to the fact that the induction axiom is typically restricted to a subclass of bounded formulas.

The links between propositional proof systems and bounded arithmetic theories have many facets but informally one can view them as two sides of the same thing: the former is a non-uniform version of the latter. In particular, it is known that proving lengths-of-proofs lower bounds for propositional proof systems is very much related to proving independence results for bounded arithmetic theories. In fact, proving such lower bounds is *equivalent* to constructing non-elementary extensions of particular models of bounded arithmetic theories. This offers a very clean and coherent framework for thinking about lengths-of-proofs lower bounds, one that has been quite successful in the past (let us mention just Ajtai's [2] lower bound for the lengths of proofs of the pigeonhole principle in constant-depth Frege systems, see Chapter 21).

In this book we introduce a new method for constructing bounded arithmetic models, and hence for proving independence results and lengths-of-proofs lower bounds. A brief description could be *forcing with random variables* but

it also has features of non-standard analysis and of definable ultraproducts. The novelty lies neither in using forcing in bounded arithmetic or proof complexity (see 'Remarks on the literature' below) nor in forcing with random variables (that is well-established in set theory; see Scott [100] or Jech[49]), but rather in finding a way how to do this meaningfully in arithmetic, and further in using families of random variables that are sampled by algorithms restricted in a particular way (different from one application to another) rather than using the family of all random variables with a given sample space and range.

The models are built from random variables defined on a sample space Ω which is a non-standard finite set (often parameterized by a subset of $\{0,1\}^n$ with a non-standard n), and sampled by functions of some restricted computational complexity. This is considered inside an $\aleph_1$-saturated non-standard model of true arithmetic. One could equivalently work with sequences of bigger and bigger sample spaces and random variables defined on them, and consider their limit behavior using a suitable ultrafilter on $\mathbf{N}$, simulating indirectly the ultraproduct construction.[1] However, the use of a non-standard model from the beginning simplifies things considerably. This is analogous to the situation in non-standard analysis: while proofs using infinitesimals (and other features of non-standard analysis) can be translated into the $\epsilon - \delta$ formalization, the intuition or clarity of the original argument may be lost in the translation.

Random variables induce probabilistic distributions and probabilities of events. In particular, two random variables may be neither equal nor unequal; rather they may be equal with some probability. However, there is a fundamental problem: probabilities cannot be used as truth-values if classical logic is to be preserved. The (almost) right choice for the truth-value is the subset of the sample space consisting of those samples for which the two random variables are equal. At the heart of our construction is the realization that we can employ a bit of non-standard analysis (namely Loeb's measure: Loeb [80]) at this point: if one identifies two such truth-values (subsets of the sample space) if their symmetric difference has an infinitesimal measure one gets a *complete* Boolean algebra – this is the single most important feature of our method. Evaluation of first-order formulas in complete Boolean algebras is very natural and faithful (as it is well-known from Boole [10] for propositional logic and from Rasiowa and Sikorski [90] for predicate logic).

[1] This construction is explained in a way accessible to readers without a basic logic education in the Appendix.

The models we get are not classical but are Boolean-valued.[2] But that is perfectly sufficient for the purpose of independence results (and lengths-of-proofs lower bounds): in order to demonstrate that a sentence is not provable from a set of axioms it is enough to show that its truth value in some model is smaller than the truth value of any finite conjunction of the axioms.

Although some of the models appear interesting in their own right we interpret the construction primarily as

> *A method that reduces an independence result or a lengths-of-proofs lower bound to a combinatorial/complexity-theoretic statement about random variables.*

The combinatorial/complexity-theoretic statement we refer to here expresses that the truth-value of a particular sentence (in a particular model) is some particular value, typically $1_{\mathcal{B}}$ or $0_{\mathcal{B}}$. The validity of such a statement is a property of the particular family of random variables forming the model. For the families we consider it can often be formulated as a statement that an algorithm of a certain type can (or cannot) perform some computational task successfully for a high fraction of inputs.

Organization of the book

The book is divided into eight parts and an appendix. Part I (*Basics*) describes the general framework of the construction and develops a few basic properties of the method. This includes witnessing of quantifiers in the structures and linking the validity in the structures with the probability in the standard model. Part II (*Second-order structures*) extends the set-up to two sorted structures, with one sort for numbers and the other for bounded sets.

In Part III (AC^0 *world*) we construct two structures. The first one is a structure based on random variables computed by shallow decision trees. This is quite a rudimentary example and its basic properties are mirrored in several later models. The second structure is based on deep decision trees and it is a model of theory V^0_1. In Part IV ($AC^0(2)$ *world*) we construct an algebraic structure based on random variables defined by algebraic decision trees. This structure is a model of the theory $Q_2V^0_1$, extending V^0_1 by a bounded quantifier

[2] One can get classical models by applying a bit of logic: First apply the Löwenheim–Skolem theorem to the whole model-theoretic situation (i.e. not only the model but also the Boolean algebra and the truth valuation) to replace it by a countable one, and then apply the Rasiowa–Sikorski theorem [91] to collapse the Boolean algebra to the two-element Boolean algebra while preserving joins and meets used for defining the truth values, and hence collapsing the model to a classical one.

allowing us to count the parity of a bounded set. The key step in analyzing both the deep tree model and the algebraic model is bounded quantifier elimination. The combinatorial heart of these elimination procedures is provided by the Hastad's switching lemma and by the Razborov–Smolensky's approximation method respectively. In both Parts III and IV we use the models to derive anew a few known undefinability results, witnessing theorems and independence results for the theories. The purpose of including this material (as well as examples in Part VII) is to demonstrate that the method is a viable alternative to the usual proof-theoretic approach based on some form of a normalization of proofs.

Part V (*Towards proof complexity*) describes a general approach using the models for lengths-of-proofs lower bounds. This follows to a large extent Ajtai's method in [2], but with some important twists. In Part VI (*Proof complexity of F_d and $F_d(\oplus)$*) we use this approach to give a new proof of an exponential lower bound for PHP (the pigeonhole principle) proofs in constant-depth Frege systems. Then we discuss a long-standing open problem to prove the same lower bound also for constant-depth Frege systems with the parity gate. We do not manage to construct a model that would prove the elusive lower bound but we review some possibly relevant material about algebraic proof systems and, more importantly, we discuss in detail the issues that any construction of the desired structure has to tackle (in particular, the necessity of partially defined random variables). The models considered in this part are quite analogous to models in Parts III and IV.

The structures in Parts III–VI are second-order. In Part VII (*Polynomial-time and higher worlds*) we return to the first-order formalization and construct several models for theories like PV (polynomially verifiable), S^1_2 and T^1_2 and derive in this way some of the most important known witnessing theorems and conditional independence results in bounded arithmetic. In this part we also note a link between pseudorandom sets and a Löwenheim–Skolem phenomenon. Further, we define a model of PV naturally interpreting structural complexity results about random oracle.

In Part VIII (*Proof complexity of EF and beyond*) we first overview aims of proof complexity of strong proof systems and recall, in particular, background facts relevant to the Extended Frege proof system EF and to bounded arithmetic theories related to EF. We then expose in some detail the emerging theory of proof complexity generators aimed at constructing examples of hard tautologies. We also discuss several conjectures regarding these generators: on the hardness of proving circuit lower bounds, Razborov's conjecture about the Nisan–Wigderson generator and Extended Frege system, a conjecture about using random sparse Nisan–Wigderson generators as gadgets in gadget

generators, and related Rudich's demi-bit conjecture. Finally we construct a model relevant to some of these conjectures.

The main text is supplemented by an appendix in which we present in a self-contained and quite elementary way the construction of an ultrapower extension of the standard model of natural numbers. We also try to convey, using several examples, some mental picture about the model so that even a reader who is not familiar with non-standard methods can develop some intuition and follow the arguments in the main text.

Remarks on the literature

A form of forcing has been applied in bounded arithmetic earlier. Paris and Wilkie [86] and later Ajtai [2, 3, 4] and Riis [98] used a simple variant of Robinson's model-theoretic forcing (although combined with an involved combinatorial reduction in Ajtai's [2, 3, 4]). Wilkie[3] described a construction of Boolean-valued models of the theory S^1_2 and reproved using it a relation – known previously from Cook [28] and Buss [11] – between S^1_2 and the Extended Frege proof system EF. His construction has been further extended by Krajíček [55, 57, 60, 56] to a wider context. With a slight simplification one can describe the Boolean algebras involved in these constructions as Lindenbaum algebras but not based on provable equivalence of formulas (or circuits) as they are defined classically but rather on *feasibly provable* (i.e. with proofs of polynomially bounded length) equivalence. This works well in the sense that any valid lower bound can be proved, in principle, by such a forcing. But on the other hand the algebras are defined using the notion of a feasible proof about which we are supposed to say something by the construction in the first place, and so it is in a sense a vicious circle. Takeuti and Yasumoto [103, 104] changed the feasibly provable equivalence to simply 'true equivalence' – breaking this vicious circle – but it apparently did not help much as we know very little about the power of Boolean circuits of feasible size. Most importantly, the algebras used in all these constructions are not complete but are closed only under some definable unions. That makes it very hard to use them.

Background

This is an investigation in bounded arithmetic and in proof complexity and we expect that, ideally, the reader is familiar with established basic definitions,

[3] Unpublished lecture at the International Congress on Logic, Methodology and Philosophy of Science in Moscow, 1987.

facts, methods and aims of the field. The relevant background in bounded arithmetic and proof complexity can be found in Krajíček [56] but some reader may find useful shorter explanations of some basic points in Krajíček [53, 57, 61, 62] or in an excellent survey by Pudlák [89]. Nevertheless we always briefly review the relevant theories and propositional proof systems before they are studied, and thus a reader with at least a minimal logic background should be able to study the book. In addition Chapter 27 gives some very general proof complexity background. Recently Cook and Nguyen [30] offered an excellent exposition of basic theories and their relations to proof systems. Propositional proof systems and their complexity are also treated by Clote and Kranakis [25]. A reader looking for a background in model theory, and non-standard models in particular, may consult Chang and Keisler [21], Marker [83] (ultrapowers are there in Exercises 2.5.19 – 2.5.22 and 4.5.37) or Kaye [52].

Despite the natural character and simplicity of Boolean-valued models they were discovered only in the late 1960s by P. Vopěnka, and by D. Scott and R. Solovay as their versions of Cohen's forcing; the paper by Scott [100] is a beautiful exposition of the basic ideas aimed at non-logicians, the best to date (it also contains detailed bibliographical/historical comments[4]). The paper by Scott [100] as well as virtually all later expositions (e.g. in Takeuti and Zaring [105] or in Jech [49]) consider only Boolean-valued models of set theory. Mansfield [82] attempts a general theory but concentrates on model-theoretic properties of the class of such models, as opposed to properties of particular models, and gives a version that yields only elementary extensions and hence is unsuitable for independence results.

[4] Takeuti reports in [103] that Gödel recognized in Boolean-valued models a model-theoretic version of a reinterpretation of logical operations that he had developed earlier but had never used for independence results as it was too complicated.

PART I

Basics

1
The definition of the models

1.1 The ambient model of arithmetic

Let L_{all} be the language containing symbols for every relation and function on the natural numbers $\mathbf{N}$; each symbol from L_{all} has a canonical interpretation in $\mathbf{N}$. Let $\mathcal{M}$ be an $\aleph_1$-saturated model[1] of the true arithmetic in the language L_{all}. Such a model exists by general model-theoretic constructions; see Hodges [43]. Definable sets mean definable with parameters, unless specified otherwise.

The $\aleph_1$-saturation implies the following:

(1) If a_k, $k \in \mathbf{N}$, is a countable family of elements of $\mathcal{M}$ then there exists a non-standard $t \in \mathcal{M}$ and a sequence $(b_i)_{i<t} \in \mathcal{M}$ such that $b_k = a_k$ for all $k \in \mathbf{N}$.
We shall often denote this sequence of length t simply $(a_i)_{i<t}$.

For example, if all elements $\{a_k\}_{k\in\mathbf{N}}$ obey some definable property P then – by induction in $\mathcal{M}$ (aka overspill, see the Appendix) – also some b_s with a non-standard index $s < t$ will obey P. Such an element b_s will serve well as 'a limit' (interpreted here informally) of the sequence $\{a_k\}_{k\in\mathbf{N}}$.

Another property implied by the $\aleph_1$-saturation (and equal to it if we used a countable language) is the following:

(2) If A_k, $k \in \mathbf{N}$, is a countable family of definable subsets of $\mathcal{M}$ such that $\bigcap_{i<k} A_i \neq \emptyset$ for all $k \geq 1$, then $\bigcap_k A_k \neq \emptyset$.
However, the intersection $\bigcap_k A_k$ does not need to be definable in $\mathcal{M}$.

These two statements are essentially the only consequences of the $\aleph_1$-saturation that we will use.

[1] In the Appendix we give an elementary and self-contained construction (the so-called ultrapower) of such a model.

The ambient model $\mathcal{M}$ will suffice for our purposes everywhere in this book. However, in general we could take for $\mathcal{M}$ an $\aleph_1$-saturated elementary extension of $\mathbf{N}$ in a many sorted language having names not only for all elements of $\mathbf{N}$ and relations and functions on $\mathbf{N}$ as $L_{\rm all}$ has, but also names for all families of sets, families of families of sets, etc., for the whole so-called *superstructure* (this is commonly done in non-standard analysis and this terminology is used there). In such a rich model the properties above would hold also for sequences of sets, families, etc.

1.2 The Boolean algebras

Let $\Omega \in \mathcal{M}$ be an arbitrary infinite set called a **sample space**. As it is an element of $\mathcal{M}$, it is $\mathcal{M}$-finite. Let $N = |\Omega|$ be the size of Ω in the sense of $\mathcal{M}$. It is necessarily non-standard.

Let $\mathcal{A} := \{A \in \mathcal{M} \mid A \subseteq \Omega\}$. This is a Boolean algebra but not a σ-algebra as the class of definable sets is not closed under all countable unions (for example, while it contains all singletons it does not contain the countable set of those elements of Ω having only standardly many predecessors in Ω). The **counting measure** (i.e. the uniform probability) on $\mathcal{A}$ is defined by:

$$A \in \mathcal{A} \to \ |A|/N.$$

Its values are the $\mathcal{M}$-rationals. A positive $\mathcal{M}$-rational is called **infinitesimal** if it is smaller that all fractions $\frac{1}{k}$, $k \in \mathbf{N}$.

Define an ideal $\mathcal{I} \subseteq \mathcal{A}$ by:

$$A \in \mathcal{I} \ \text{ iff } \ |A|/N \ \text{ is infinitesimal.}$$

$\mathcal{I}$ is not definable in $\mathcal{M}$ (otherwise the set of natural numbers $\mathbf{N}$ would be definable, violating the overspill in $\mathcal{M}$). Using $\mathcal{I}$ define a Boolean algebra $\mathcal{B} := \mathcal{A}/\mathcal{I}$.

The induced measure on $\mathcal{B}$ (the so-called Loeb's measure) will be denoted μ. Hence $\mu(b)$ for $b \in \mathcal{B}$ is the standard part of $|B|/N$ (i.e. the unique standard real infinitesimally close to it) for any $B \in \mathcal{A}$ such that $B/\mathcal{I} = b$. It is a measure in the ordinary sense: The values of μ lie in the reals $\mathbf{R}$. It is σ-additive and a strict measure: $\mu(b) > 0$ if $b \neq 0_{\mathcal{B}}$.

The following key lemma is a combination of two well-known facts, one from non-standard analysis and one from measure theory.

Lemma 1.2.1 *$\mathcal{B}$ is a complete Boolean algebra.*

Proof: First, as in the construction of Loeb's measure Loeb [80], we use the $\aleph_1$-saturation to show that

Claim 1: *$\mathcal{B}$ is a σ-algebra and the measure μ is σ-additive.*

To establish the claim let $\{b_k\}_{k\in\mathbf{N}}$ be a countable subset of $\mathcal{B}$. Assume $b_k = B_k/\mathcal{I}$ for B_ks from $\mathcal{A}$. We may assume, without loss of generality, that $B_0 \subseteq B_1 \subseteq \ldots$. It is enough to find $C \in \mathcal{A}$ such that $B_k \subseteq C$ for all $k \in \mathbf{N}$ and $\mu(C) = \lim_{k\to\infty} \mu(B_k)$. It holds that for any $k \geq 1$ there is $n_k \geq 1$ such that for all $m > \ell \geq n_k$

$$\frac{|B_\ell|}{N} \leq \frac{|B_m|}{N} \leq \frac{|B_\ell|}{N} + \frac{1}{k}.$$

We may assume, by taking subsequence $\{B_{n_k}\}_{k\in\mathbf{N}}$, that $n_k = k$.

Take a non-standard extension $\{B_i\}_{i<t}$ of $\{B_k\}_{k\in\mathbf{N}}$ guaranteed to exist by $\aleph_1$-saturation (each B_k is an element of $\mathcal{M}$). Consider the following property P parameterized by s:

$$B_s \in \mathcal{A} \wedge \forall i \leq s; B_i \subseteq B_s \wedge \frac{|B_i|}{N} \leq \frac{|B_s|}{N} \leq \frac{|B_i|}{N} + \frac{1}{i}.$$

Property P is obeyed by all standard s and hence, by induction in $\mathcal{M}$, also by some non-standard $s_0 < t$.

It is easy to verify that $C := B_{s_0}$ has the required properties.

Claim 2: *$\mathcal{B}$ satisfies the ccc condition: any antichain is at most countable.*

This holds because the measure is strictly positive, i.e. $\mu(b) > 0$ for $b \neq O_{\mathcal{B}}$: Any antichain can contain only finitely many (non-zero) elements with measure in each interval $(1/(n+1), 1/n]$, $n \geq 1$.

As a consequence we get

Claim 3: *Any family of elements of $\mathcal{B}$ has the same set of upper bounds as one of its countable subfamilies.*

To see this, note that a family has the same set of upper bounds as the ideal it generates which in turn has the same set of upper bounds as any maximal antichain it contains – such antichains are countable by ccc (Claim 2), and each element of the antichain is majorized by a union of a finite number of elements of the family).