Count(q) versus the Pigeon-Hole Principle

Søren Riis*

June 1994
Revised February 1996

Abstract

For each $p \geq 2$ there exists a model $\mathbf{M}^*$ of $I\Delta_0(\alpha)$ which satisfies the Count(p) principle. Furthermore, if p contains all prime factors of q there exist $n, r \in \mathbf{M}^*$ and a bijective map $f \in \text{dom}(\mathbf{M}^*)$ mapping $\{1, 2, ..., n\}$ onto $\{1, 2, ..., n + q^r\}$.

A corollary is a complete classification of the Count(q) versus Count(p) problem. Another corollary shows that the pigeon-hole principle for injective maps does not follow from any of the Count(q) principles. This solves an open question [Ajtai 94].

1 Introduction

The most fundamental questions in the theory of the complexity of calculations are concerned with complexity classes in which 'counting' is only possible in a quite restricted sense. Thus it is not surprising that many elementary counting principles are unprovable in systems of Bounded Arithmetic. These are axiom systems where the induction axiom schema is restricted to predicates of low syntactic complexity. For a good basic reference see [Krajicek 95].

The status of the elementary counting principles (which normally all are proved by some explicit or implicit reference to cardinality) is in a non-trivial way linked to questions in complexity theory. Let me give a few examples:

(1) If there is a model of S_2 in which the Gödel sentence Con(S_2) holds, but where the elementary pigeon-hole principle fails, then there is a model of S_2 in which NP $\neq$ co-NP [PW 87], [Krajicek 95]. In general, there is a model of S_2^1 in which NP $\neq$ co-NP if and only if there is a model of S_2^1 in which P $\neq$ NP [Krajicek 95].

(2) If there are models of $I\Delta_0$ where the pigeon-hole principle fails in the sense that for some n there exists a bijection from $\{1, 2, \ldots, n\} \to \{1, 2, \ldots, n^2\}$, then $I\Delta_0$ is not finitely axiomatizable. And by a similar argument it can be shown that if there exist models of S_2^2 where this version of the pigeon-hole principle fails, then there are models of S_2^2 in which NP $\neq$ co-NP.

(3) If S_2^1 proves a version of the pigeon-hole principle, then S_2^1 actually proves Bertrand's prime number theorem (there is always a prime between n and $2n$). According to S.Buss's theorem [Buss 85] there would be a polynomial time algorithm which produces a prime number of a given

*e-mail: pmtsr@amsta.leeds.ac.uk

number of bits. This is only known to be the case under strong conditional assumptions like Riemann's Hypothesis or P=NP.

Open problems like P $\neq$ NP have been acknowledged by prominent mathematicians to be one of our times' most outstanding problems (see for example [Smale 92A] and [Smale 92B]). Progress concerning any of the statements in (1)-(3) above are likely to go together with progress in the P $\neq$ NP problem and other related questions. It is generally believed by researchers in Bounded Arithmetic that the status of the elementary counting principles in models of Bounded Arithmetic has fundamental importance. Unfortunately the most fundamental versions of the problems are beyond the current techniques.

It is possible to soften up most of these fundamental problems. One way to do this is to add a new function symbol to the underlying language. Another essentially equivalent approach is to replace the underlying first order logic with second order logic with a restricted comprehension axiom schema [Riis 93A], [Riis 93B].

In [PW 85] A.Wilkie and J.Paris showed that the non-provability of the pigeon-hole principle (expressed by adding a new function symbol to the underlying language) would follow if it could be shown that the pigeon-hole principle does not have bounded depth polynomial size Frege proofs. And later M.Ajtai showed the validity of the converse implication. Actually Ajtai settled this issue for $I\Delta_0$. Later in [Ajtai 90] Ajtai considered the q-matching principle (in the case $q = 2$) which is in some sense is stronger than any version of the pigeon-hole principle. More specially he showed that there are models of $I\Delta_0(R)$ in which the Count(2) principle fails while any Δ_0-version of the ordinary pigeon-hole principles holds. Ajtai's results was later improved in various directions [BIKPPW 92], [KPW 95], [Riis 93B] and [BP 93].

The status of the Count(q) and Count(p) principle (in the basic case of $I\Delta_0$) was raised by J. Paris and A. Wilkie in the early 80s. Later Ajtai conjectured (in connection with [Ajtai 88]) that for different primes q, p the principles Count(q) and Count(p) are independent principles. Later Ajtai showed that this indeed is the case [Ajtai 94]. Ajtai's proof uses a list of deep results from the modular representation theory of the symmetrical group. Ajtai's proof depends strongly on q being a prime number.

In [Riis 93B] and [Riis 94A] the Count(q) versus Count(p) problem (also allowing composite numbers q, p) was reduced to a purely combinatorial conjecture. I showed that the existence of so-called 'exceptional forests', and the existence of implications between Count(q) and Count(p) go together. In [BIKPP 94] P. Beame, R. Impagliazzo, J. Krajicek T. Pitassi, and P. Pudlak were able to show that this type of problem is related to that of finding lower bounds on the degrees of the witnessing polynomials in Hilbert's Nullstellensatz. They obtained such lower bounds by a very careful repeated use of Ramsey's Theorem. This way they managed to solve a sufficiently strong part of a technical conjecture from [Riis 93B] and thereby obtain a complete classification of the Count(q) versus Count(p) problem in the base case (i.e. over $I\Delta_0$).

Independently in [Riis 94A] I managed to obtain an asymptotic classification of the exceptional forests and thereby solving the Count(q) versus Count(p) problem. Like [BIKPP 94] the proof in [Riis 94A] also involved a very complicated and technical use of Ramsey's Theorem. In this paper I have eleminated the involved and tricky use of Ramseys Theorem, and replaced it by a construction more in the style of [BKPPRS 95].

Our aim in this paper is to prove the following theorem

Theorem 1 *Let $q \geq 2$ and assume that r is an increasing function such that $r(n) \in \omega(1) \cap o(\log(n))$. For any countable language L of arithmetic in which all terms have polynomial growth rate, and for any sound extension of $I\Delta_0(L)$ (which leaves at least one function symbol f undefined) there is a model* M *such that:*

(i) *The $\Delta_0(L)$-Count(q) axiom scheme is valid in* M

(ii) *There exists $n \in$ M such that the function f defines a bijection from $\{1, 2, \ldots, n\}$ onto $\{1, 2, \ldots, n + q^{r(n)}\}$.*

A $\Delta_0(L)$-Count(q) axiom scheme is a scheme (for $\Delta_0(L)$-formulas) which formalizes the elementary matching principle stating that *if $\{1, 2, ..., n\}$ is divided into disjoint p-element subsets, then p divides n.*

2 Applications

The Count(q) principle implies many versions of the pigeon-hole principle, so the theorem shows that the matching principle Count(q) so to speak has an interesting blind spot.

In the future we'll let $\mathrm{PHP}^*_{*+s}(\mathrm{bij})$ denote the elementary principle stating that *there does not exist n and a bijective map from $\{1, 2, ..., n\}$ onto $\{1, 2, ..., n + s\}$*. This principle can also be stated for all $\Delta_0(L)$-formulas as a Δ_0 scheme.

Corollary 2 (Settling conjecture by Ajtai)
For different primes q, p $\quad \mathrm{Count}(q) \nvdash \mathrm{Count}(p)$
[Ajtai 94]

Corollary 3 (Obtaining the complete classification)
For fixed $q, p \geq 2$ the following is equivalent

(a) *p divides a power of q*

(b) $\mathrm{Count}(q) \vdash \mathrm{Count}(p)$.

[BIKPP 94], [Riis 94A]

Proof: The implication (a) $\Rightarrow$ (b) can be shown either by producing constant degree witnessing polynomial for the coresponding system of equations ([BIKPP 94]) or by constructing exceptional forests ([Riis 93B], [Riis 94A]). The implication (b) $\Rightarrow$ (a) follows from Theorem 1. According to this theorem $\mathrm{Count}(p) \nvdash \mathrm{PHP}^*_{*+q^{r(*)}}(\mathrm{bij})$. If p contain a prime factor which does not appear in q then $\mathrm{Count}(p) \vdash \mathrm{PHP}^*_{*+q^{r(*)}}(\mathrm{bij})$ and thus $\mathrm{Count}(q) \vdash \mathrm{Count}(p)$. □

Corollary 4 (Solving the Count versus PHP problem)
Let $r(n) \in w(1) \cap o(\log(n))$. For each $q, p \geq 2$
$\mathrm{Count}(p) \nvdash \mathrm{PHP}^*_{*+q^{r(*)}}(\mathrm{bij})$ *if and only if p divides a power of q*
(*if and only if* $\mathrm{Count}(q) \vdash \mathrm{Count}(p)$)

Let $\mathrm{PHP}^{*+p}_*(\mathrm{inj})$ be the the statement that *there is no n and no injective map from $\{1, 2, ..., n+p\}$ into $\{1, 2, ..., n\}$* and let $\mathrm{PHP}^*_{*+p}(\mathrm{sur})$ be the statement that *there is no n and no surjective map from $\{1, 2, .., n\}$ onto $\{1, 2, ..., n + p\}$.*

Corollary 5 (Answering an open question by Ajtai) [Ajtai 94]
(a) $\mathrm{PHP}^{*}_{*+1}(\mathrm{bij}) \nvdash \mathrm{PHP}^{*+1}_{*}(\mathrm{inj})$.
(b) $\mathrm{PHP}^{*+1}_{*}(\mathrm{inj}) \dashv\vdash \mathrm{PHP}^{*}_{*+1}(\mathrm{sur})$.
(c) $\mathrm{Count}(q) \nvdash \mathrm{PHP}^{*+1}_{*}(\mathrm{inj})$.

Proof: (a) follows from (c), because $\mathrm{Count}(q) \vdash \mathrm{PHP}^{*}_{*+1}(\mathrm{bij})$. To show (c) notice that $\mathrm{PHP}^{*+1}_{*}(\mathrm{inj}) \vdash \mathrm{PHP}^{*}_{*+q^{r(*)}}(\mathrm{bij})$ for any r. But according to Theorem 1 $\mathrm{Count}(q) \nvdash \mathrm{PHP}^{*}_{*+q^{r(*)}}$ when $r \in w(1) \cap o(\log)$. The bi-implication in (b) is a simple exercise. □
This shows that the pigeon-hole principle for injective maps are efficiently stronger than the pigeon-hole principle for bijective maps. Actually it shows that:

Corollary 6 *There exists a model* $\mathbf{M}^*$ *of* $I\Delta_0(\alpha)$ *in which* $\mathrm{Count}(p)$ *holds for each* $p \in \mathbf{N}\backslash\{1\}$. *Yet, there exists* $n \in \mathbf{M}^*$ *and an injective map* $f \in \mathrm{dom}(\mathbf{M}^*)$ *mapping* $\{1, 2, ..., n+1\}$ *into* $\{1, 2, ..., n\}$.

Proof: By the completeness theorem it suffices to show that for each finite set $p_1, p_2, .., p_l$ of integers, the conjunction $\mathrm{Count}(p_1) \wedge ....\wedge \mathrm{Count}(p_l)$ does not imply $\mathrm{PHP}^{*+1}_{*}(\mathrm{inj})$. This follows by an argument similar to the one given for (c) in corollary 5. □

Corollary 7 *Let* Γ *denote any collection of* $\mathrm{Count}(q)$ *principles,* $q \in \mathbf{N}$. *Then* $\Gamma \nvdash \mathrm{PHP}^{*+1}_{*}(\mathrm{inj})$. *If* Γ *is any collection of* $\mathrm{Count}(q)$ *principles where each* q *is a prime* $\neq p$, *then* $\Gamma \nvdash \mathrm{Count}(p)$.

According to corollary 2, Count(7) neither proves Count(5) or Count(2). Does Count(7) prove Count(5) $\vee$ Count(2)? None of the methods in [Ajtai 94], [BIKPP 94] and [Riis 93B] which deals directly with the $\mathrm{Count}(q)$ versus $\mathrm{Count}(p)$ principle are sufficient to answer this question. However it follows directly from Theorem 1 that

Corollary 8 *Suppose* $p_1, p_2, \ldots, p_k$ *all contain a prime-factor which does not appear in* q. *Then* $\mathrm{Count}(q) \nvdash \mathrm{Count}(p_1) \vee \mathrm{Count}(p_2) \vee \ldots \vee \mathrm{Count}(p_k)$.

Proof: Notice that $\mathrm{Count}(p_1) \vee \mathrm{Count}(p_2) \vee \ldots \vee \mathrm{Count}(p_k)$ imply $\mathrm{PHP}^{*}_{*+q^{r(*)}}$ for any r. But according to Theorem 1 $\mathrm{Count}(q) \nvdash \mathrm{PHP}^{*}_{*+q^{r(*)}}$ for certain functions $r(*)$. □

There is a natural way of translating a first order relational formula ψ into a Boolean propositional formula ψ_n of a universe with n elements. If, for example, $\psi \equiv \forall i \exists j \forall k\ R(i,j) \wedge S(i,j,k)$ then ψ_n can be written as $\wedge_i \vee_j \wedge_k\ x_{i,j} \wedge y_{i,j,k}$. For any relational formula ψ we consider the propositional formulas ψ_n. Notice that ψ holds in all finite models if and only if ψ_n is a tautology for each n. The *substitution axiom schema based on* ψ consists of the formulas ψ_n where each variable in ψ_n can be replaced by any propositional formula. The natural first-order formulations of the $\mathrm{Count}(q)$-principles and the pigeon-hole principles can be translated into a substitution axiom schema. The boolean version of the $\mathrm{Count}(q)$-principle becomes (after having introduced a variable y_A for each q-element subset $A \subseteq J$ for some $|J| \neq 0$ modulo q) the substitution schema $\bigvee_{j \in J} \bigwedge_{A \ni j} \neg y_A \vee \bigvee_A \bigvee_{B \neq A,\ B \cap A \neq \emptyset} (y_A \wedge y_B)$. A first order deduction rule $\frac{\theta_1, \theta_2, \ldots, \theta_k}{\theta}$ where θ_i, $i = 1, 2, \ldots, k$ and θ are relational first order formulas can naturally (for each n) be translated into a deduction rule for propositional logic. A *first order proposition proof system* $\mathbf{P}$ consists of a finite number of substitution axiom schemas together with a finite number of first order deduction rules. A $\mathbf{P}$*-proof*

(in Hilbert style) of a proposition η is a sequence $\eta_1, \eta_2, \ldots, \eta_u = \eta$ of Boolean formulas, such that each η_j, $j = 1, 2, \ldots, u$ is either a substitution instance of a substitution axiom scheme, or there are $i_1, i_2, \ldots, i_k < j$ such that $\frac{\eta_{i_1}, \eta_{i_2}, \ldots, \eta_{i_k}}{\eta_j}$ is a substitution instance of a deduction rule.

Absolute tautologies ψ_n are tautologies for which ψ besides being valid in all finite models also holds in all infinite models. Similarly an *absolute deduction rule* $\frac{\theta_1, \theta_2, \ldots, \theta_k}{\theta}$ is a rule for which $\theta_1 \wedge \theta_2 \ldots \wedge \theta_k \Rightarrow \theta$ is an absolute tautology. An *absolute proof system* is first order propositional proof system where all axiom schemes and all deduction rules are absolute.

A *Frege propositional proof system* is a propositional proof system which consists of: (i) a finite number of substitution schemas, i.e. Boolean formulas θ with special substitution variables $y_1, y_2, \ldots, y_k$. (ii) A finite number of deduction rules $\frac{\theta_1, \theta_2, \ldots, \theta_k}{\theta}$ where θ_i, $i = 1, 2, \ldots, k$ and θ are substitutions schemes. We only consider propositional systems which are consistent and sound (i.e. prove the usual tautologies).

Notice that Frege's propositional proof systems are absolute proof systems where the underlying first-order formulas are quantifier-free. The pigeon-hole principle PHP_n is not an absolute tautology because it fails for infinite sets. Elementary tautologies like $\theta_1 \wedge \theta_2 \rightarrow \theta_1$ are absolute. Modus Ponens $\frac{\theta_1, \theta_1 \rightarrow \theta_2}{\theta_2}$ is an absolute deduction rule. It is well known [Ajtai 88], [BIKPP 94], and [PW 85] that there are close links between results concerning provability in systems of Bounded Arithmetic and the length of bounded depth Frege Proofs. Our method of non-standard models (introduced by Ajtai [Ajtai 88]) allows us in a very straight forward way to generalize these results to absolute proof systems.
We can express theorem 1 in terms of absolute proof systems.

Theorem 9 *Let* $\mathbf{P}$ *be a propositional proof system which besides a finite number of absolute axiom schemas and absolute deduction rules contains the* $\mathrm{Count}(q)$ *substitution axiom scheme. Then there are no polynomial size bounded depth* $\mathbf{P}$*-proofs of* $\mathrm{PHP}^*_{*+q^{w(1)}}(\mathrm{bij})$.

From this theorem it is easy to obtain variants of corollaries $1, 2, \ldots 7$ where provability in $I\Delta_0(\alpha)$ has been replaced by provability by 'polynomial size, bounded depth, absolute proofs'. Furthermore, Theorem 1 follows by standard arguments from Theorem 9 which thus can be considered as the main result of the paper.

Theorem 9 is strongest possible in the sense that:

Theorem 10 *The implication* $\mathrm{Count}(q) \vdash \mathrm{PHP}^*_{*+q^k}(\mathrm{bij})$ *is absolute for any fixed* k.

The Theorem states that in a non-standard model $\mathbf{M}$ of first order arithmetic there exists a substitution instance $\mathrm{Count}_s(q)$ of $\mathrm{Count}(q)$ such that the tautology
$\mathrm{Count}_s(q) \rightarrow \mathrm{PHP}^*_{*+q^k}(\mathrm{bij})$ remains valid even if we allow arbitrary (i.e. not only M-definable) truth-table evaluations.

3 Proofs based on equations

Consider the identity

(Eq 1)
$$\binom{n+3}{3} + \frac{n}{2}\binom{n}{2} = \binom{n}{3} + \frac{n}{2}\binom{n+3}{2} + 1.$$

Assume naively that there exists a bijection $f : \{1, 2, .., n\} \to \{1, 2, .., n+3\}$. This would induce a bijection g from the 3-element subsets of $\{1, 2, \ldots, n\}$ onto the 3-element subsets of $\{1, 2, \ldots, n+3\}$, as well as induce a bijection h from the pairs of $\{1, 2, \ldots, n+3\}$ onto the pairs of $\{1, 2, \ldots, n\}$. In the case of n is even patch together $\frac{n}{2}$ copies of h together with g to obtain a bijection from $\{1, 2, \ldots, \binom{n}{3} + \frac{n}{2}\binom{n+3}{2}\}$ onto $\{1, 2, \ldots, \binom{n+3}{3} + \frac{n}{2}\binom{n}{2}\}$. In the case of n is odd just extend f to a bijection $f : \{1, 2, \ldots, n+1\} \to \{1, 2, \ldots, (n+1)+1\}$ and precede with $n := n+1$. We have just proved (in a very roundabout way and by reference to the ordinary pigeon-hole principle) that there can be no bijection from $\{1, 2, \ldots, n\}$ onto $\{1, 2, \ldots, n+3\}$.

This argument can (unlike the traditional cardinality arguments) be translated into a bounded depth polynomial size Frege proof, which uses a substitution schema for the pigeon-hole principle. And for essentially the same reason it can be translated into a proof in Bounded Arithmetic. The (more trivial) implication $\mathrm{PHP}^*_{*+3}(\mathrm{bij}) \vdash \mathrm{PHP}^*_{*+1}(\mathrm{bij})$ follows by taking 3 copies of a supposed counter example to $\mathrm{PHP}^*_{*+1}(\mathrm{bij})$. Thus

Proposition: *The pigeon-hole principles* $\mathrm{PHP}^*_{*+3}(\mathrm{bij})$ *and* $\mathrm{PHP}^*_{*+1}(\mathrm{bij})$ *hold in the same models of Bounded Arithmetic.*

Before we go into a deeper analysis of the pigeon-hole principle let us try to understand which of such implications can be proved by the use of these ideas. The positive results in this direction were first obtained in [Riis 93B]. What about the negative results (i.e. the classification of the implications which not are supported by any binomial equation)?

To answer this we must consider identities over $\mathbf{Z}_q$ (in the case of $\mathrm{Count}(q)$) as well as $\mathbf{Z}_\infty := \mathbf{Z}$ (in the case of the PHP). We must also be prepared to consider identities which contain polynomial expressions in nested binomial coefficients like:

$$c_1\left(\binom{\binom{r+17}{3}}{7}^5 - \binom{\binom{r}{3}}{7}^5\right) - c_2\left(\binom{\binom{r+19}{3}\binom{r+17}{4}}{11}^3 - \binom{\binom{r+2}{3}\binom{r}{4}}{11}^3\right) + \ldots$$

The theory we develop below (and which goes far beyond just considering arguments based on binomial equations) allows us to prove the following polynomial equation:

Theorem 11 *For* $r \in \mathbf{C}$, *and* $k, m \in \mathbf{N}$

$$\text{(Eq 2)} \qquad \binom{r}{k}\binom{r}{m} = \sum_{j=\max\{k,m\}}^{k+m} \binom{j}{k+m-j}\binom{2j-k-m}{j-m}\binom{r}{j}$$

The point (and usefulness) of this equation is that it allows us to replace a product $\binom{r}{k}\binom{r}{m}$ with a linear expression in $\binom{r}{j}$ $j \leq k+m$. Thus for example $\left(\binom{r+17}{7}\right)^5 = \left(\frac{1}{7!}\binom{r+17}{3}\left(\binom{r+17}{3}-1\right)\ldots\left(\binom{r+17}{3}-6\right)\right)^5$ can first be expressed as a polynomial in $\binom{r+17}{3}$ (of degree 35). The theorem allows us to express this polynomial as a linear expression in $\binom{r+17}{j}$ where $j = 15, 16, \ldots, 104, 105$. It turns out that (Eq 2) actually can be proved in Bounded Arithmetic (when k and m are fixed standard numbers and r is considered as a free variable). Furthermore, the (elementary) identity $\binom{r+a+b}{c} - \binom{r+a}{c} = \sum_{j=c-a}^{c} \binom{a}{c-j}\left(\binom{r+b}{j} - \binom{r}{j}\right)$ is also provable in systems of Bounded Arithmetic (when a, b and c are

fixed numbers). Thus to understand which proofs can be based on binomial equations (in a similar fashion to the argument based on (Eq 1)), it suffices to consider equations of the form:

(Eq 3) $$\sum_{j=0}^{u} c_j(\binom{r+k}{j} - \binom{r}{j}) \neq 0 \text{ modulo } q$$

where c_j; $j = 0, 1, \ldots, u \in \mathbf{Z}_q$. In the case $q = \infty$, $c_1, c_2, \ldots, c_u$ might depend on r. If the Count(q) principle is available we can consider such equations modulo q. Thus we also need to consider binomial equations over $\mathbf{Z}_q$. In the case we work over $\mathbf{Z}$ the argument only has a chance to take place in models of Bounded Arithmetic if $c_1, c_2, \ldots, c_u$ are integers bound by a term in the underlying language. This follows by Parikh's theorem [Parikh 71].

In general both $u := u(r)$ and $k := k(r)$ can be functions of r. Constraint on their growth-rate is closely linked to the systems of Bounded Arithmetic we have fixed. For instance the argument has only a chance to take place in systems where that $\binom{r}{u}$ is bound by a term $t(r)$ in the underlying language (again because of Parikhs theorem [Parikh 71]). So in the case of $I\Delta_0(\alpha)$ where all terms are polynomials, we need only to consider arguments where $u \in O(1)$. Summarizing we only consider the question whether $c_1, c_2, \ldots, c_u$ can be chosen bounded by a fixed polynomial in r such that equation
(Eq 3) has solutions for infinitely many r. Now $\binom{r+q^{\omega(1)}}{j} - \binom{r}{j} = 0$ modulo q, so for any $j \in O(1)$ (Eq 3) has infinitely many solutions. On the other hand there exists integers, which actually can be expressed as rational functions (like the function $n \to \frac{n}{2}$ in (Eq 1)), such that (Eq 3) has infinitely many solutions when $k(r) \in O(1)$. Thus

Corollary 12

For any p there exists a binomial equation which together with $I\Delta_0(\alpha) + \mathrm{PHP}^*_{*+1}(\mathrm{bij})$ *supports a proof of* $\mathrm{PHP}^*_{*+p^{O(1)}}(\mathrm{bij})$.

For any p there exists a binomial equation which together with $I\Delta_0(\alpha) + \mathrm{Count}(p)$ *supports a proof of* $\mathrm{PHP}^*_{*+p^{O(1)}}(\mathrm{bij})$.

There is no binomial equation which together with $I\Delta_0(\alpha) + \mathrm{PHP}^*_{*+1}(\mathrm{bij})$ *supports a proof of* $\mathrm{PHP}^*_{*+p^{\omega(1)}}(\mathrm{bij})$.

There is no binomial equation which together with $I\Delta_0(\alpha) + \mathrm{Count}(p)$ *supports a proof of* $\mathrm{PHP}^*_{*+p^{\omega(1)}}(\mathrm{bij})$.

To obtain our general result we have to consider <u>all</u> proofs (not just proofs based on binomial equations).

4 Exceptional forests

4.1 Stratification of the notion of existence

It has been said that existence does not come in degrees. The poor has as much existence as the queen. Many independence proofs in logic can be viewed as tampering with the notion of existence. In this section I present a method by which the existence of finitistic objects can be stratified.

In [PB 94] P.Pudlak and S.Buss considered a game $G = G(\psi)$ played between a *prover* and an *adversary.* In the game the adversary tries to persuade the prover that a certain propositional formula ψ is false. The prover can ask questions (of a type specified as part of the rules of the game). The adversary (who claims $\neg\psi$) can make up the answers. However if the adversary is caught in an elementary contradiction (like claiming both η and $\neg\eta$) the prover wins. P. Pudlak and S. Buss [PB 94] have shown that there is a close link between this game and the length of propositional proofs. For instance any Frege proof system has a canonical translation to a prover-adversary game. Actually Pudlak and Buss showed that *the minimal number $\mu(G)$ of rounds in the game $G(\psi)$ needed to trap the adversary is proportional to the logarithm of the length (counted as the number of steps) of the shortest Frege propositional proof of ψ.*

We now show that any prover-adversary game has some other complexity measures which relate to $\mu(G)$ in a non-trivial fashion.

A strategy for the prover can be represented as a decision tree: At the root the first question is assigned. For each possible answer (by the adversary) we have an edge. Each answer leads to a new situation in which the prover might (or might not) ask another question. At the end of each leaf l the prover has gathered a specific piece of information. Later we will refer to this piece of information as a *(forcing) condition.* Normally *we only focus on trees where the prover stops long before the adversary is trapped in an elementary contradiction.*

In other words, we consider trees where each leaf is assigned a condition which represents some partial knowledge concerning the adversary's assignment. At the root of the tree we have no knowledge. Each node corresponds to a concrete question Q while the various edges from a node represent the possible answers. All conditions in the leafs are clearly incompatible because different leafs contain conflicting pieces of information. We always assume (mostly for cosmetic reasons) that a question is relevant (i.e. its answer cannot be deduced from the previous questions).

Now given $q, n \in \mathbf{N}$ where $q \geq 2$. Let $\mathcal{F}$ be a forests of labeled trees in which all trees have height $\leq h$. Suppose that each condition appears 0 modulo q times in $\mathcal{F}$. Does the forest contain 0 modulo q trees? If not, we say *$\mathcal{F}$ is an exceptional forest.* The question whether there exist exceptional forests depends on the proposition ψ, the class of allowed questions and how an elementary contradiction is defined.

4.2 Exceptional forests are proofs

Proposition 13 *If there is an q-exceptional forest (based on the game $G(\psi)$), then ψ is valid.*

Proof: Suppose that ψ is invalid so the adversary can avoid any contradiction even if presented with the collection of all possible questions. Let the adversary chose a fixed strategy S. Now each tree contains exactly one branch which represents the adversary's answers (according to S). Thus if each branch appears 0 modulo q times, then $|\mathcal{F}| = 0$ modulo q. □

The proposition shows that *we can consider q-exceptional forests as proofs.* Like most syntactical correct strings not are proofs, so are most forests not q-exceptional forests. The relationship between the shortest proof (in a fixed proof system), the minimum number of rounds $\mu(G)$ in interactive proofs and the complexity (number of trees/height of trees etc) in q-exceptional forests is related in an interesting and non-trivial fashion. As a by-product of our analysis we will show that for

certain classes of propositions ψ and proof systems *the length of the shortest proof of ψ and the minimal height of the trees in q-exceptional forests (based on $G(\psi)$) correspond to each other in a well defined one to one fashion.*
The fact that a q-exceptional forests based on a prover-adversary game $G(\psi)$ can be viewed as a 'proof' of ψ is reflected in various other ways. For example, all the basic properties of logical deductions also hold for q-exceptionalness. As an example, if there exists an q-exceptional forest for $G(\psi \wedge \psi')$ (of height $\leq h$) there exists q-exceptional forests for both $G(\psi)$ and $G(\psi')$ (of heights $\leq h$).
In this section we briefly indicate how exceptional forests can be translated into bounded depth polynomial size Frege propositional proofs. This is another reason we can view q-exceptional forests as proofs. We consider a adversary-prover game $G_{(D,R)}$ where the adversary claims that there exists a bijection f from D onto R. The prover is allowed to ask questions of the form $f(d) =?$ or $f^{-1}(r) =?$. In this game we always think of D and R as being two big finite sets where $|R| \geq |D|$. In the following section we only focus on the games $G_{(D,R)}$ even though many of the results and ideas hold for most prover-adversary games.

Example: *Consider a forest $\mathcal{F}_{(D,R)}$ of (D,R)-labeled trees (all of height 1) which contain the trees with root questions $r?$; $r \in R$ together with $(q-1)$ copies of the trees with root questions $d?$; $d \in D$. Each branch can be represented as a pair (d,r); $d \in D$, $r \in R$. If $|D| \neq |R|$ modulo q, this forest is q-exceptional.*

These type of forests are so trivial that we like in [Riis 93B] and [Riis 94A] in some contexts will ignore them and only reserve the term exceptional to less trivial examples.

Assume that for some $r \neq 0$ modulo q there exists n and a bijection $f : \{1,2,\ldots,n\} \rightarrow \{1,2,\ldots,n+r\}$. The existence of $\mathcal{F}_{(\{1,2,\ldots,n\},\{1,2,\ldots,n+r\})}$ can be expressed as Boolean tautology. Each substitution instance of these tautologies have bounded depth polynomial size general proofs. Now we can prove that a bijection $f : \{1,2,\ldots,n\} \rightarrow \{1,2,\ldots,n+r\}$, defines a partitioning of the trees in $\mathcal{F}_{((\{1,2,\ldots,n\},\{1,2,\ldots,n+r\})}$ into disjoint q-element subsets. This violates the Count(q) principle. The general case where r is a fixed power of q is treated in an essentially similar fashion even though (as it turns out) we have to consider forests of height $\geq 2^{l-1}+1$.

4.3 Some basic results

Proposition: *If $|D| = |R|$, there are no q-exceptional forests.*

Proof: The adversary's proposition 'there is a bijection from D onto R' is valid. □

Proposition: *Suppose that the prover is only allowed to ask questions of the form $f(d) =?$. Then there are no q-exceptional forests.*

Proof: Let $f : D \rightarrow R$ be any injective map. The adversary can in a global way answer any collection of questions consistent according to this map. Thus there can be no exceptional forest. □

A homogeneous tree $[d_1, d_2, \ldots, d_l; r_1, r_2, \ldots, r_m]$ is a tree which consists of the conditions α for which each $d \in \mathrm{Dom}(\alpha)$, has $d \in \{d_1, d_2, \ldots, d_l\}$ or $\alpha(d) \in \{r_1, r_2, \ldots, r_m\}$. Let $\mathcal{F}[[k,m]]$ denote the forest which consists of the trees $[d_1, d_2, \ldots, d_k; r_1, r_2, \ldots, r_m]$ where $d_1 < d_2 < \ldots < d_k$ and

$r_1 < r_2 < \ldots < r_m$. It turns out that arguments based on a binomial equations are in some sense isomorphic to arguments based on forest of the form $\cup_{k,m} \lambda_{k,m}\ \mathcal{F}[[k,m]]$, where $\lambda_{k,m}$ denotes the multiplicity of the forest $\mathcal{F}[[k,m]]$.

We have already seen that theorem 11 allows us to reduce arguments based on binomial equations to a special normal form. Here is the analogous result for homogeneous forests of homogeneous trees:

Lemma 14 *The forest $\mathcal{F}[[k,m]]$ contains the same conditions as the forest which contains $\binom{j}{k+m-j}\binom{2j-k-m}{j-m}$, copies of $\mathcal{F}[[j,0]]$. The forest $\mathcal{F}[[j,0]]$ contains exactly the conditions of length j.*

Proof: The condition $\{\{d_1,r_1\},\{d_2,r_2\},\ldots,\{d_j,r_j\}\}$ appears $\binom{j}{k+m-j}\binom{2j-k-m}{j-m}$ times in the forest $\mathcal{F}[[k,m]]$ when $j = \max\{k,m\}$, $\max\{k,m\}+1,\ldots,k+m$. It appear once in $\mathcal{F}[[j,0]]$. □

Proof of Theorem 11: Let $|D| = |R|$. It suffices to show that for for all integers $d = r \geq k+m$

$$\binom{d}{k}\binom{r}{m} = \sum_{j=\max\{k,m\}} \binom{j}{k+m-j}\binom{2j-k-m}{j-m}\binom{d}{j}.$$

The left hand side denotes the number of trees in $\mathcal{F}[[k,m]]$. The right hand side denotes the number of trees in $\mathcal{F}' := \cup_{j=\max\{k,m\}} \binom{j}{k+m-j}\binom{2j-k-m}{j-m}\mathcal{F}[[j,0]]$. According to lemma 14 each condition appears the same number of times in $\mathcal{F}[[l,m]]$ and $\mathcal{F}'$. Now choose a bijection $\rho : D \to R$. This bijection select exactly one condition from each tree in both $\mathcal{F}[[l,m]]$ and $\mathcal{F}'$. Thus $|\mathcal{F}[[l,m]]| = |\mathcal{F}'|$. □

Two conditions (branches) α and β are *incompatible* ($\alpha \perp \beta$) if *there exists* $d \in D : \alpha(d) \neq \beta(d)$ *or there exists* $r \in R : \alpha^{-1}(r) \neq \beta^{-1}(r)$. Two conditions (branches) α and β are *compatible* ($\alpha || \beta$) if they not are incompatible. Suppose that T is a (D,R)-labeled tree, and suppose $\rho : D \to R$. Then T^ρ denotes the tree which is obtained by first removing all edges representing answers incompatible to ρ and, second by contracting all edges $< d, r >$ where $\rho(d) = r$.

Lemma 15 *Let T be a (D,R)-labeled tree of height h, let $\rho : D \to R$, and let $D' := D \setminus \mathrm{dom}(\rho)$ and let $R' := R \setminus \mathrm{ran}(\rho)$. Suppose that $|D'|, |R'| \geq h+1$, then T^ρ is a (D',R')-labeled tree. The tree T^ρ might have height 0, but it is never empty. It have height at most h*

Proof: Induction after $|\rho|$. As the induction is downwards starting from on arbitary point it suffices to show the first step in the induction. So assume that $|\rho| = 1$, and that $|D|-1, |R|-1 \geq h+1$. We can write $\rho = \{< d, r >\}$. For any question $d'?$, $d' \neq d$ ($r'?$ $r' \neq r$) in T remove the edge and the subtree above the edge $< d', r >$ ($< d, r' >$). The number of leafs in T is at least 2 because $|D|, |R| \geq h+1$. Thus each question in this new tree has at least one legitimate answer and will be non-empty after this procedure. For any question $d?$ ($r?$) keep intact the edge $< d, r >$ while removing all other edges and there subtrees on top of these. Contract the edge $< d, r >$. This way we get a tree which still has height h or in certain special cases $h-1$. In the case $T := (d)$ or $T := (r)$ the tree T^ρ becomes the (unique) tree of height 0. □

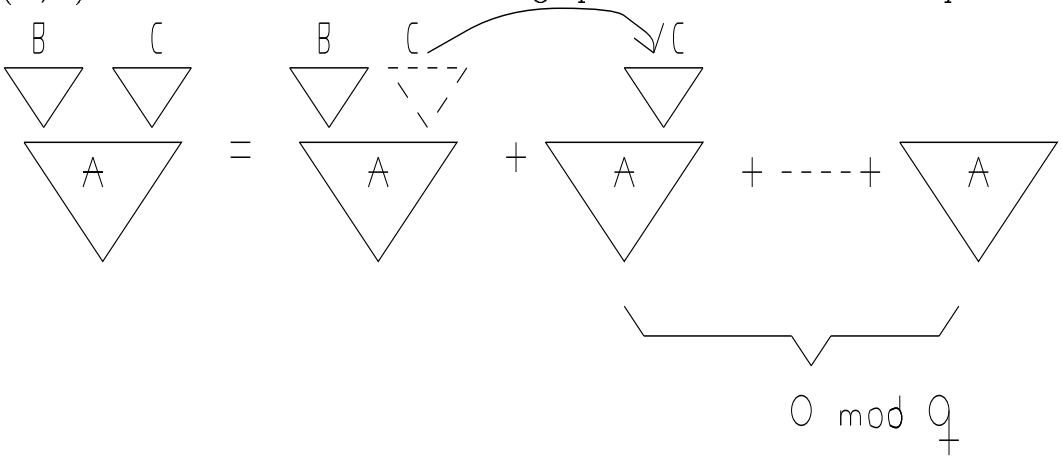
Corollary 16 *Let $\mathcal{F}$ be an q-exception forest of (D,R)-labeled trees of height $\leq h$. Let $\rho : D \to R$ be a partial map. Let $D' := D \setminus \mathrm{dom}(\rho)$ and let $R' := R \setminus \mathrm{ran}(\rho)$. Suppose that $|D'|, |R'| \geq h+1$. Then $\mathcal{F}^\rho$ is an q-exceptional forest of (D',R')-labeled trees.*

Proof: Suppose that $\mathcal{F} = \{T_1, T_2, \ldots, T_u\}$ is an q-exceptional forest. We have to show that $\mathcal{F}^\rho := \{T_1^\rho, T_2^\rho, \ldots, T_u^\rho)$ is a q-exceptional forest. We already know (lemma 15) that all restricted trees are non-empty, so $\mathcal{F}^\rho$ contains the same number as trees as $\mathcal{F}$. We have to show that each condition appears 0 modulo q times in $\mathcal{F}^\rho$ and that the number of trees of height 0 (=trees which contain only the empty condition) is 0 modulo q. It suffices to consider the case $|\rho| = 1$. Consider a branch $\{< d_1, r_1 >, < d_2, r_2 >, \ldots, < d_j, r_j >\}$ with elements in D', R'. Its total number of appearances is exactly the same as the number of appearances of the conditions $\{< d_1, r_1 >, < d_2, r_2 >, \ldots, < d_j, r_j >\}$ and $\{< d, r >, < d_1, r_1 >, < d_2, r_2 >, \ldots, < d_j, r_j >\}$. This number is 0 modulo q. The number of trees of height 0 is the same as the number of trees which contain the empty condition. This number is the same as the number appearances of the condition $\{< d, r >\}$ in $\mathcal{F}$ i.e. 0 modulo q. □

Lemma 17 *Let $q \geq 2$ and let $l \geq 1$. Let D, R be finite sets with $|R| - |D| = q^l$. Then there exists an exceptional forest $\mathcal{F}$ of (D,R)-labeled trees of height at most q^l.*

Proof: Let D, R be finite sets with $|R| - |D| = q^l$. First suppose that $\bar{d} := |D|$ is a power (at least $l+1$) of q. Consider the forest $\mathcal{F}_h$ which consists of $\mathcal{F}[[h,0]]$ together with $q-1$ copies of $\mathcal{F}[[0,h]]$. Each branch of height h appears exactly q times, while branches of all other lengths appears 0 times. The number of trees in $\mathcal{F}_h$ is $(q-1)\binom{\bar{d}}{h} + \binom{\bar{d}+q^l a}{h}$. If $h \geq q^l$ this is 1 modulo q, and $\mathcal{F}_h$ is an exceptional forest. In general (when there is no restriction on $\bar{d}$) chose $d' = q^u$ a big power of q such that $\bar{d}' \leq \bar{d}$. Construct an q-exceptional forest of (D',R')-labeled trees where $|D'| = \bar{d}'$ and $|R'| = \bar{d} + q^l$. Now apply corollary 16 to obtain an q-exceptional forest of (D,R)-labeled trees. □

At first sight many questions concerning specially labeled trees might seem hopeless. However in general we can break down trees and put them into a nice normal form. To see this let T be a (D,R)-labeled tree. Consider the following equation which holds modulo q



Notice that both sides of the equation contain 1 modulo q trees. Also that each condition appears the same number of times (modulo q) on each side of the equation.

Suppose that $\mathcal{F} := \{T_1, ..., T_u\}$ is any forest. Repeated application of the identity allows us to break down the trees in $\mathcal{F}$. Eventually each tree is brought on a normal form where at each level all but at most one node is a leaf. We call such trees *perfectly unbalanced* (=**PU**). Thus we have proved,

Lemma 18 *Fix $q \geq 2$, $q \in \mathbf{N}$. Let $\mathcal{F} := \{T_1, ..., T_u\}$ be any forest. There exists a forest $\mathcal{F}' := \{T'_1, T'_2, ..., T'_{u'}\}$ in which each condition counted modulo q appears the same number of times as in $\mathcal{F}$. Each tree in $\mathcal{F}'$ is a* **PU***-tree and furthermore $u' = u$ (modulo q).*

Notice that the **PU**-trees have a very simple representation. Each **PU**-tree can in a canonical fashion be represented by expressions of the form,
$(s_{11}?, s_{12})(s_{2,1}?, s_{2,2}), ..., (s_{j-1,1}?, s_{j-1,2})(s_j?)$, where $s_{i,1} \in D$ if and only if $s_{i,2} \in R$. Similar $s_{i,1} \in R$ if and only if $s_{i,2} \in D$.
It turns out that there are various useful identities between collections of **PU**-trees. For example $(1_D?, 1_R)(2_D?, 2_R)(3_R?) - (1_D?, 1_R)(2_R?, 2_D)(3_R?) = (1_D?, 1_R)(2_D?) - (1_D?, 1_R)(2_R?)$.
And $(1_D?, 1_R)(2_D?, 2_R)(3_R?) - (2_D?, 2_R)(1_D?, 1_R)(3_R?) = (1_D?, 1_R)(2_D?) - (2_D?, 2_R)(1_D?)$. The identities illustrate that *the difference between two trees which agree for all branches of length $\geq l$ can be expressed as the difference between two trees of height $l-1$.* These considerations show that we have a lot of flexibility below the top-level. From now we assume (without loss of generality) that we have brought all **PU**-trees to the form $(d_1?, r_1) \ldots (d_{l-1}?, r_{l-1})(u?)$ where u either belongs to D or to R.
Repeated use of the equations gives the following lemma.

Lemma 19 (Normal form) *Let $\mathcal{F}$ be an q-exceptional forest of (D, R)-labeled trees of height $\leq h$. Then there exists a q-exceptional forest of (D, R)-labeled* **PU***-trees of height $\leq h$. Furthermore, it is possible to ensure each tree is of the form:*
$(d_1?, r_1)(d_2, r_2) \ldots (d_{l-1}?, r_{l-1})(u?)$ *where $d_1 < d_2 < \ldots < d_{l-1}$, where $u \in D$ or $u \in R$, and where $l \leq h$.*

Here is the first class of (non-trivial) 2-exceptional forests I discovered. This happened during my doctoral work [Riis 93B]:
Example Consider a forest $\mathcal{F}_{(\bar{d},\bar{r})}$ which consists of the trees which contain all **PU**-trees of the form:

(1) $(d?, r_1)(r_2?)$ where $d \in \{1, 2, \ldots, \bar{d}\}$, $r_1, r_2 \in \{1, 2, \ldots, \bar{r}\}$, $r_1 > r_2$ and $r_1 - r_2$ is odd.

(2) $(d?, r_1)(r_2?)$ where $d \in \{1, 2, \ldots, \bar{d}\}$, $r_1, r_2 \in \{1, 2, \ldots, \bar{r}\}$, $r_1 < r_2$ and $r_2 - r_1$ is even.

(3) $(d_1?, r)(d_2?)$ where $d_1, d_2 \in \{1, 2, \ldots, \bar{d}\}$, and $r \in \{1, 2, \ldots, \bar{r}\}$.

Notice that $|\mathcal{F}_{(\bar{d},\bar{r})}| = \bar{d}\binom{\bar{r}}{2} + \bar{r}\binom{\bar{d}}{2}$. The forest $\mathcal{F}_{(\bar{d},\bar{r})}$ is 2-exceptional when $\bar{d}$ is odd and $\bar{r} - \bar{d} = 2$ modulo 4. Each branch appears an even number of times yet the forest contains an odd number of trees. The forest $\mathcal{F}_{(5,7)}$ contains 175 **PU**-trees. In all cases the forest $\mathcal{F}_{(\bar{d},\bar{r})}$ has height 2. ♣

4.4 Classification of the q-exceptional forests

For any property $\mathcal{P}$ of conditions we can define an equivalence relation '$\equiv'_{\mathcal{P}}$ by $\alpha \equiv_{\mathcal{P}} \beta$ if and only if $\forall \rho : \ \mathcal{P}(\rho) \Rightarrow (\alpha||\rho \Leftrightarrow \beta||\rho)$. The relation is clearly reflexive and symmetrical. For α, β, γ, for which $\alpha \equiv_{\mathcal{P}} \beta$, $\beta \equiv_{\mathcal{P}} \gamma$ for any ρ with $\mathcal{P}(\rho)$ we have $\alpha||\rho \Leftrightarrow \beta||\rho \Leftrightarrow \gamma||\rho$. Thus the relation $\equiv_{\mathcal{P}}$ is also transitive.

A property $\mathcal{P}$ of conditions is *transitive* if $\mathcal{P}(\alpha) \Rightarrow \mathcal{P}(\beta)$ when $\beta \subseteq \alpha$. Let T be a (D,R)-labeled tree. For a transitive property $\mathcal{P}$ of conditions we define $T(\mathcal{P})$ as the collection $\{\alpha \in T : \mathcal{P}(\alpha)\}$. A decision tree is *proper* if each question has at least one legitimate answer.

A condition α is k-extendable, if $\mathcal{P}(\alpha)$ and for each $d? \in D$ (and each $r? \in R$) each $r \in R$ ($d \in D$) with $\mathcal{P}((d,r) - \alpha)$ the condition $(d,r) - \alpha$ is $(k-1)$-extendable. A condition α is 1-extendable, if $\mathcal{P}(\alpha)$ and for each $d? \in D$ (and each $r? \in R$) there exists $r \in R$ ($d \in D$) such that $\mathcal{P}((d,r) - \alpha)$. A transitive property $\mathcal{P}$ is k-extendable if $\emptyset$ is k-extendable.

Lemma 20 *For any transitive property $T(\mathcal{P})$ can be organized into a decision tree. If, furthermore, $\mathcal{P}$ is h-extendable then $T(\mathcal{P})$ is a proper decision tree.*

Proof: First remove the conditions (branches) which do not satisfy $\mathcal{P}$ from T. The transitivity ensures that this is a tree. This tree T' might have top nodes in which a question where no legitimate answers can be produced. However if $\emptyset$ is h-extendable, a question on level $j \leq h$ must lead to an answer which is $\geq (h-j)$-extendable. □

Assume that T is a (D,R)-labeled tree of height $\leq h$. Let $\mathcal{P}$ be a transitive h-extendable property. Let $T^{\mathcal{P}}$ denote the tree which appears by contracting the edges with a single forced answer. We can (and will) view the tree $T^{\mathcal{P}}$ as being labeled by the equivalence classes defined by $\equiv_{\mathcal{P}}$.

A decition tree of $(D,R,\mathcal{P})$-condition is a decision tree for the game $G(\psi$ $(=: G(D,R,\mathcal{P})$ where $\psi \equiv$ "ρ defines a bijection $D \to R$ and $\mathcal{P}(\rho)$" and where a forced answer can be used in getting an elementary contradiction. From the definitions we get

Lemma 21 *Let $\mathcal{P}$ be a transitive h-extendable property. Then for any (D,R)-labeled decision tree T, the tree $T^{\mathcal{P}}$ is a decision tree for the game $G(D,R,\mathcal{P})$. If $\mathcal{F}$ is a q-exceptional forest of (D,R)-labeled trees of height $\leq h$, then $\mathcal{F}^{\mathcal{P}}$ is a q-exceptional forest of $(D,R,\mathcal{P})$-labeled trees.*

Our analysis towards the classification of exceptional forests is going to use various transitive properties.

(i) For $\rho_0 : D \to R$ let $\mathcal{P}_{\rho_0}$ be the property that $\rho || \rho_0$. This is a transitive property which is $(|D| - |\rho_0|)$-extendable. Notice that $(D,R,\mathcal{P}_{\rho_0})$-conditions are isomorphic to (D',R')-conditions where $D' := D \setminus \mathrm{dom}(\rho_0)$ and where $R' := R \setminus \mathrm{ran}(\rho_0)$.

(ii) Let $\mathcal{P}_{(D_1,D,R_1,R)}$ be the property that ρ maps D_1 into R_1 and maps $D \setminus D_1$ into $R \setminus R_1$. Notice that $\mathcal{P}_{(D_1,D,R_1,R)}$ is a transitive property and that it is $\mathbf{min}(\{|D_1|, |R_1|, |D| - |D_1|, |R| - |R_1|\})$-extendable.

(iii) Let Γ be a (consistent) collection of constrains of the form $\rho(d) = r \Leftrightarrow \rho(d') = r'$. Let $\mathcal{P}_\Gamma$ be the property that ρ satisfies all the constraints in Γ. Notice that $\mathcal{P}_\Gamma$ is transitive and is $\mathbf{min}(\{|D|, |R|\}) - |\Gamma|$-extendable.

(iv) Let D_{ij}, $j \in J_i$ (and R_{ij}, $j \in J'_i$) be partitions of D (and R) into disjoint sets. Assume that the sizes $|D_{ij}| = |R_{ij}| = c(i)$ only depend on i. Let $\mathcal{P}$ be the property that for each i and each j there exists k such that $\rho(D_{ij}) = R_{ik}$. Notice that $\mathcal{P}$ is a transitive property which is $\mathbf{min}_i(\{|J_i|, |J'_i|\})$-extendable.

(v) Let D_{ij} and R_{ij} be given as in (iv). Assume that each D_{ij} and R_{ij} are ordered. Let $\mathcal{P}'$ be the property that ρ besides satsifying $\mathcal{P}$ is order preserving. We notice that the property $\mathcal{P}'$ is transitive and is $\mathbf{min}_i(\{|J_i|, |J'_i|\})$-extendable.

Lemma 22 *Let $q \geq 2, l \geq 1$ be fixed integers. For any pair of sets (D, R) with $|D| + q^{l-1} + 2q^l = 0$ modulo $3q+2$ and $|R| = |D| + q^l$, there exists a partitioning $\mathcal{P}_D$ of D into $d' := \frac{|D|+q^{l-1}+2q^l}{3q+2}$ disjoint $(q+1)$-element subsets (forming a collection $\mathcal{P}_D^1$) and $d' - q^{l-1}$ disjoint $(2q+1)$-element subsets (forming a collection $\mathcal{P}_D^2 := \mathcal{P}_D \setminus \mathcal{P}_D^1$).*

Furthermore, there exists a partitioning $\mathcal{P}_R$ of R into d' disjoint $(2q+1)$-element subsets (forming a collection $\mathcal{P}_R^2$) and $d' - q^{l-1}$ disjoint $(q+1)$-element subsets (forming a collection $\mathcal{P}_R^1 := \mathcal{P}_R \setminus \mathcal{P}_R^2$).

Proof: Notice that $d'(q+1) + (d' - q^{l-1})(2q+1) = (3q+2)(d' - q^{l-1} - 2q^l) = |D|$ and that $d'(2q+1) + (d' - q^{l-1})(q+1) = (3q+2)(d' - q^{l-1} - 2q^l) + q^l = |D| + q^l = |R|$. □

Now fix a pairing $\tilde{\mathcal{P}}$ which: (i) pairs the members in P_D^1 (i.e. the selected $q+1$-element subsets of D) with the members of P_R^2 (i.e. the selected $2q+1$-element subsets of R) (ii) pairs the members in P_D^2 (i.e. the selected $2q+1$-element subsets of D) with the members of P_R^1 (i.e. the selected $q+1$-element subsets of R).

Fix a cyclic order on the elements in each selected $q+1$-element set (i.e. each member in $\mathcal{P}_D^1 \cup \mathcal{P}_R^1$). Also fix a cyclic order on the elements in each selected $2q+1$-element set (i.e. each member in $\mathcal{P}_D^2 \cup \mathcal{P}_R^2$).

In addition to these fixed choices we consider a selection S which chooses an emphasized point in each selected subset. Later we will run through all possible S. Notice that there are 1 modulo q possible selections S. Each emphasized point induces an order among the q (or $2q$) non-emphasized points in the same selected subset (by letting the selected point be the smallest point in the ordering).

Let $\rho'_S : D \to R$ be a map which maps the emphasized point in a $(q+1)$-element ($(2q+1)$-element subset) to the emphasized point in the corresponding (wrt. $\tilde{P}$) $(2q+1)$-element subset ($(q+1)$-element subset). According to (i) this property (which we denote $\mathcal{P}_S^{(1)}$) is transitive and $(|D| - 2d' + q^{l-1})$-extendable.

Let $\mathcal{P}_S^{(2a)}$ be the property that ρ maps the non-emphasized points in a given $(q+1)$-element subset ($\in \mathcal{P}_D^1$) onto the non-emphasized points in a $(q+1)$-element subset ($\in \mathcal{P}_R^1$). As noticed in (iv) this property is transitive and d'-extendable.

Let $\mathcal{P}_S^{(2)}$ be the property that ρ besides satisfying $\mathcal{P}_S^{(2a)}$, also maps the j^{th} element (after the emphasized point) to the j^{th} element (after the emphasized point) $j = 1, 2, \ldots q$. As noticed in (v) this property is transitive and d'-extendable.

Let $\mathcal{P}_S^{(3a)}$ be the property that ρ maps the non-emphasized points in a given $(2q+1)$-element subset ($\in \mathcal{P}_D^2$) onto the non-emphasized points in a $(2q+1)$-element subset ($\in \mathcal{P}_R^2$). This property is transitive and d'-extendable.

Let $\mathcal{P}_S^{(3)}$ be the property that ρ besides satisfying $\mathcal{P}_S^{(3a)}$, also maps the j^{th} element to the j^{th} element $j = 1, 2, \ldots 2q$.

Let $\mathcal{P}_S^{(**)} := \mathcal{P}_S^{(1)} \wedge \mathcal{P}_S^{(2)} \wedge \mathcal{P}_S^{(3)}$. This is a transitive property which is d'-extendable. We can also describe the $(D, R, \mathcal{P}_S^{(**)})$-labeling as arising from the game $G_{(D,R)}^{(S)}$ which is a modification of $G_{(D,R)}$. More specifically in $G_{(D,R)}^{(S)}$ the adversary has to ensure that the map defines a partial bijection from D into R. Furthermore, the adversary has to insure that the map:

(i) maps the emphasized point in a $q+1$-element subset ($2q+1$-element subset) to the emphasized point in corresponding (wrt. $\tilde{\mathcal{P}}$) $2q+1$-subset ($q+1$-element subset).
(ii) maps (in an order preserving fashion) the non-emphasized points in any given $q+1$-element subset ($\in \mathcal{P}_D^1$) onto the non-emphasized points in a $q+1$-element subset ($\in \mathcal{P}_R^1$).
(iii) maps (in an order preserving fashion) the non-emphasized points in any given $2q+1$-element subset ($\in \mathcal{P}_D^2$) onto the non-emphasized points in a $2q+1$-element subset ($\in \mathcal{P}_R^1$).
Now let $D' := \mathcal{P}_D^1$, $R' := \mathcal{P}_R^2$ and let $D'' := \mathcal{P}_D^2$, $R'' := \mathcal{P}_R^1$.
The partition $\tilde{P}$ induces an identification i_1 of elements in D' and R'' as well as an identification i_2 of elements in R' and D'' (actually any pair of identifications i_1, i_2 will do). Let $\mathcal{P}_S$ be the property that ρ (besides satisfying $\mathcal{P}_S^{(**)}$) induces maps $D' \to R'$ and $D'' \to R''$ such that for all $d \in D'$ $\rho^{-1}(i_2(\rho(d))) = i_1(d)$. The property $\mathcal{P}_S$ is transitive and $(d'-ql-1)$-extendable.
Thus for each selection S we have defined a property $\mathcal{P}_S$. The property $\mathcal{P}_S$ is designed such that

Lemma 23 *$(D, R, \mathcal{P}_S)$-conditions are isomorphic to (D', R')-conditions where $|D'| = d' - q^{l-1} = \frac{|D|+q^{l-1}+2q^l}{3q+1} - q^{l-1}$ and where $|R'| = |D'| + q^{l-1}$.*

Suppose that $T = (d_1, r_1)(d_2, r_2)\ldots(d_{h-1}, r_{h-1})(u)$ where $d_j \in D$, $r_j \in R$ and $u \in D \cup R$ is a **PU**-tree. Then $T^{(S)} := T^{\mathcal{P}_S}$ is a (D', R')-labeled **PU**-tree. For each **PU**-tree T we define a forest $\mathcal{F}(T)$ by letting $\mathcal{F}(T) := \cup_{S \in \tilde{S}} T^{(S)}$. As usual all sets are multi-sets. We now focus on the relationship between T and $\mathcal{F}(T)$. First we divide pairs (d, r) into 6 categories (which are independent of S):
(1) d belongs to a $(q+1)$-element subset $A \in P_D^1$ and r belongs to a $(2q+1)$-element subset $B \in P_R^2$ and $\tilde{P}(A, B)$.
(2) d belongs to a $(q+1)$-element subset $A \in P_D^1$ and r belongs to a $(2q+1)$-element subset $B \in P_R^2$ and $\neg\tilde{P}(A, B)$.
(3) d belongs to a $(2q+1)$-element subset $A \in P_D^2$ and r belongs to a $(q+1)$-element subset $B \in P_R^1$ and $\tilde{P}(A, B)$.
(4) d belongs to a $(2q+1)$-element subset $A \in P_D^1$ and r belongs to a $(q+1)$-element subset $B \in P_R^1$ and $\neg\tilde{P}(A, B)$.
(5) Both d and r belong to $(q+1)$-element subsets A and B (in respectively P_D^1 and P_R^1).
(6) Both d and r belong to $(2q+1)$-element subsets A and B (in respectively P_D^2 and P_R^2).
We say two pairs (d, r) and (d', r') *interact* if at least one of the following conditions is satisfied:
(i) Both d and d' belong to $A \in P_D^1 \cup P_D^2$
(ii) Both r and r' belong to $B \in P_R^1 \cup P_R^2$
(iii) The element d belongs to $A \in P_D^1 \cup P_D^2$, the element r' belongs to $B \in P_R^1 \cup P_R^2$ and $\tilde{P}(A, B)$.
(iv) The element d' belongs to $A \in P_D^1 \cup P_D^2$, the element r belongs to $B \in P_R^1 \cup P_R^2$ and $\tilde{P}(A, B)$.

Lemma 24 *Assume $T = (d_1, r_1)(d_2, r_2)\ldots(d_{h-1}, r_{h-1})(u_r)$ is a **PU**-tree in which (d_1, r_1) is a pair of type 5 or 6 which are not interacting with any other (d_j, r_j), $j > 1$. Then $\mathcal{F}(T)$ can be divided into q identical sub-forests $\mathcal{F}_1, \mathcal{F}_2, \ldots, \mathcal{F}_q$ together with 1 modulo q trees of height 1.*

Proof: Assume (d_1, r_1) is a pair of type 5 which does not interact with any (d_j, r_j), $j > 1$. Now fix all emphasized points except the emphasized point in $A \in P_D^1$ and the emphasized point in $B \in P_R^1$. We now want to fix the remaining emphasized points. There are $(q+1)^2$ choices. Of these q choices produce emphasized point in A and B such that the property $\mathcal{P}_S^{(2)}$ is satisfied. In the remaining $(q+1)^2 - q$ choices $\mathcal{P}_S^{(2)}$ is not valid. Thus no branch survives beoynd the first level so we get $(q+1)^2 - q = 1$ modulo q decision trees of height 1.

The case where (d_1, r_1) is a pair of type 6 is treated similarly except that instead of q choices there are $2q$ choices (where $\mathcal{P}_S^{(3)}$ holds). And instead of $(q+1)^2 - q$ trees of height 1 we get $(2q+1)^2 - 2q = 1$ modulo q trees. □

Lemma 25 *Assume that $T = (d_1, r_1)(d_2, r_2)\ldots(d_i, r_i)(d_{i+1}, r_{i+1})\ldots(u_h)$ is a* **PU***-tree where each (d_j, r_j) interacts with at least one other $(d_{j'}, r_{j'})$ $(j \neq j',\ j, j' \leq i)$. Assume that (d_j, r_j), $j > i$ are all of type $1, 2, 3$ or 4 which do not interact with any pair $(d_1, r_1), (d_2, r_2), \ldots, (d_i, r_i)$. Then $\mathcal{F}(T)$ consists of trees which all have height $\leq \lfloor i/2 \rfloor + 1$.*

Proof: Assume that $(d_{j_1}, r_{j_1}), (d_{j_2}, r_{j_2}), \ldots, (d_{j_k}, r_{j_k})$, $j_1 < j_2 < \ldots < j_k$ are all pairwise interacting. Now all (D', R')-labeled trees will only contain the edge corresponding to (d_{j_1}, r_{j_1}) because the edges $(d_{j'_k}, r_{j'_k})$, $j' > 1$ either represent redundant information or incompatible information. In the first case the edge gets contracted, while in the second case the edge and the part of the tree which is above it get removed. Thus the maximal number of pairs $(d_j, r_j), j \leq i$ which survives is $\lfloor i/2 \rfloor$. All pairs $(d_j, r_j), j > i$ get contracted or removed. The top node (u_h) is the only part of T above level i which survives. □

Lemma 26 *Assume that $T = (d_1, r_1)(d_2, r_2)\ldots(d_{h-1}, r_{h-1})(u_h)$ is a* **PU***-tree. Then there exists a forest $\mathcal{F}(T)'$ such that:*

(i) *$\mathcal{F}(T)$ and $\mathcal{F}(T)'$ contain the same conditions (counted modulo q) as $\mathcal{F}(T)$.*

(ii) *Both $\mathcal{F}(T)$ and $\mathcal{F}(T)'$ contain 1 modulo q trees.*

(iii) *The Forrest $\mathcal{F}(T)'$ can be divided into disjoint parts $\mathcal{F}(T)_1, \mathcal{F}(T)_2, \ldots \mathcal{F}(T)_p$ and $\mathcal{F}(T)_{\text{remainder}}$ such that $\mathcal{F}(T)_1, \ldots \mathcal{F}(T)_q$ are identical and all trees in $\mathcal{F}(T)_{\text{remainder}}$ have height $\leq \lfloor h/2 \rfloor + 1$.*

Proof: We have already seen that there are trees $T_1, T_2, ..., T_s$ for some $s = 1$ modulo q such that $T_1, T_2, \ldots, T_s$ contain the same conditions (modulo q) as T. These trees contain pairs $(d_1, r_1), \ldots, (d_{h-1}, r_{h-1})$ that can appear in any order we might wish. So without loss of generality we can assume that each tree T_j, $j \leq s$ is of a form so lemma 24 or lemma 25 is applicable. Let $\mathcal{F}(T)' := \cup_{j \leq s}\ \mathcal{F}(T_j)$. Thus with this notation we can assume that $\mathcal{F}(T)'$ satisfies (iii). Now T contains the same conditions as the forest $T_1, T_2, \ldots, T_s$ so $\mathcal{F}(T)$ and $\mathcal{F}(T)'$ also must contain the same conditions. Notice that the forest $\mathcal{F}(T)$ contains $|\tilde{S}| = 1$ modulo q trees. Finally notice that the forest $\mathcal{F}(T)'$ contains $s \cdot |\tilde{S}| = 1$ modulo q trees. □

Suppose $\mathcal{F}$ is a forest of (D, R)-labeled PU-trees. Then $\mathcal{F}^* := \cup_{T \in \mathcal{F}}\ \mathcal{F}(T)'$, where (of course) both F and the right hand side of the expression (as usual) are treated as multi-sets.

Now the forest $\mathcal{F}^*$ arises as a union of forests $\mathcal{F}(T)'$ which consists of trees with $(D, R, \mathcal{P}_S^{(**)})$-labelings. According to the identification in lemma 23 $\mathcal{F}^*$ consists of (D', R')-labeled trees where $|D'| = d' - q^{l-1} = \frac{|D| + q^{l-1} + 2q^l}{3q+1} - q^{l-1}$ and where $|R'| = |D'| + q^{l-1}$. Thus we have shown:

Lemma 27 *Let $q \geq 2, l \geq 1$ be fixed integers. Suppose that $|D| + q^{l-1} + 2q^l = 0$ modulo $3q+2$ and $|R| = |D| + q^l$. The procedure which transforms a forest $\mathcal{F}$ of (D,R)-labeled* **PU***-trees into a forest $\tilde{\mathcal{F}}$ of (D',R')-labeled* **PU***-trees has the following property:*

If $\mathcal{F}$ is a q-exceptional forest of trees of height $\leq h$ then $\tilde{\mathcal{F}}$ can be divided into two disjoint parts:
(i) *a q-exceptional forest of height $\leq \lfloor \frac{h}{2} \rfloor + 1$*
(ii) *a part of trees each appearing 0 modulo q times.*

We now apply the property $\mathcal{P}_{\rho_0}$ defined in (i)

Lemma 28 *Let D, R be finite sets with $|D|, |R| \geq h + 3q + 2$. Then there exists $\rho_0 : D \to R$ with $|\rho_0| \leq 3q+1$ such that $|D \setminus \text{dom}(\rho_0)| + q^{l-1} + 2q^l = 0$ modulo $3q+2$. Furthermore, if we let P_{ρ_0} denote the property that ρ_0 is compatible to ρ then a (D,R,P_{ρ_0})-labeling is in a canonical fashion isomorphic to a (D'',R'')-labeling where $D'' := D \setminus \text{dom}(\rho_0)$ and $R'' := R \setminus \text{ran}(\rho_0)$.*

Theorem 29 (Classification of exceptional forests) *Suppose that $q \geq 2, l \geq 1$ be fixed integers. Let D, R be finite sets with $|R| = |D| + q^l$ and suppose that $|D| \geq 4^{l+1}q^l$. Then:*
There is no q-exceptional forest of (D,R)-labeled trees of height $\leq 2^{l-1}$.
There are q-exceptional forests of (D,R)-labeled trees of height q^l.

Proof: The upper bound was already proved in lemma 17. The lower bound follows by repeated use of lemma 26 and lemma 28. The first application of lemma 28 let us pass from (D,R) to (D'',R'') where $|D''| \geq |D| - 3q - 1$. Now $|D'| \geq \frac{|D''|+q^{l-1}+2q^l}{3q+2}$ so $|D'| \geq |D|/(3q+2)$. To ensure that we can repeat this process satisfactorily it suffice to ensure that $|D| \geq 4^{l+1}q^l$. □

4.5 Bases

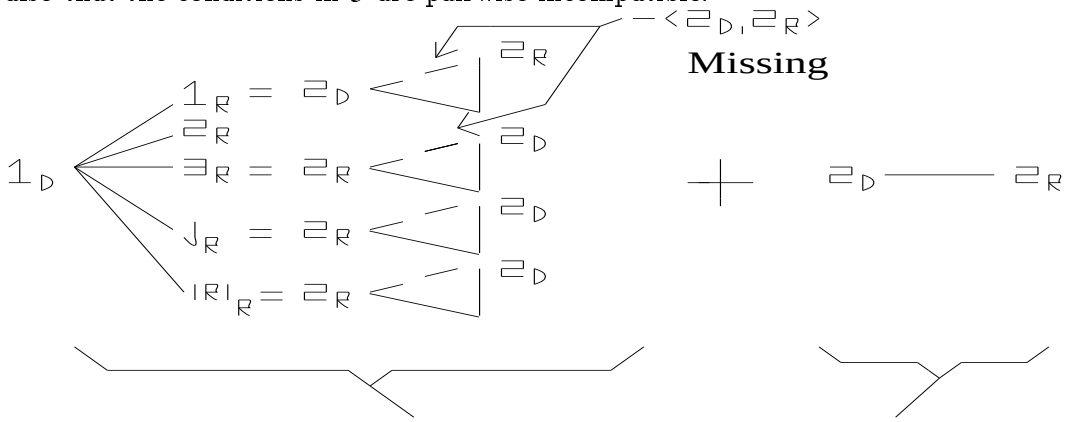
In our lower bounds we first lead to another and more general concept than that of a (D,R)-labeled decision tree. Let $\mathcal{C}$ be a collection of conditions (branches). If these can be organized into a (D,R)-labeled decision tree the conditions must be pairwise incompatible and further if $\rho : D \to R$ is a partial bijection (with $|\rho| \leq |D| - h$) then at least one condition α must be compatible to ρ.

A (l,D,R)-*basis* (or just l-basis when D and R are clear from the context) is a collection $\mathcal{C}$ of pairwise incompatible conditions which for each $\rho : D \to R$ with $|\text{dom}(\rho)| \leq l$ contain at least one branch α which is compatible with ρ. It is not hard to show that this notion is quite robust with respect to the choice of l. More specifically, $\mathcal{C}$ is an l-basis for some value $|D| - h(h+1)/2 \geq l \geq h(h+1)/2$ if and only if it is l-basis for any value $|D| - h(h+1)/2 \geq l \geq h(h+1)/2$. So when we let $l = |D|/2$ we could really have chosen many other values of l.

It is easy to show that the conditions in a (D,R)-labeled decision tree (of height $\leq |D|/2 - 1$) form a $|D|/2$-Basis. It would be convenient if the conditions in a (l,D,R)-basis always could be organized into a (D,R)-labeled tree. This is not (quite) the case:

Example (part 1): Let $\mathcal{C}$ consist of the conditions $\{< 1_D, 2_R >\}$,
$\{< 2_D, 2_R >\}, \{< 1_D, 1_R >, < 2_D, r_R >\}$ where $r_R \neq 2_R$ and
$\{< 1_D, r_R >, < d_D, 2_R >\}$ where $d_D \neq 2_D$ (and $r_R \neq 2_R$). If $|D| \geq 6$, $\mathcal{C}$ is an $|D|/2$-basis. To see this let $\rho : D \to R$ be given. If necessary extend ρ such that $\rho(1_D)$, $\rho(2_D)$ and $\rho^{-1}(2_R)$ are defined.

This is possible when $|\rho| + 3 \leq |D|$ (e.g. when $|\text{ran}(\rho)| \leq |D|/2$ and $|D| \geq 6$). Now notice that there must be exactly one condition in $\mathcal{C}$ which is extended by the map which extends ρ. Notice also that the conditions in $\mathcal{C}$ are pairwise incompatible.



The conditions in the (D, R)-system $\mathcal{C}$ cannot be organized into a (D, R)-labeled decision tree because none of the candidates for the question at the root ($1_D?, 2_D?$ or $2_R?$) appears in all conditions. ♣

In [Riis 93B] and [Riis 94A] I showed (for a different but similar labeling) that each (D, R)-system $\mathcal{C}$ can be 'refined' to a (D, R)-labeled tree. The price for this a blow-up in the length of the branches. In this paper I show a stronger result which avoids this. Consider again the example:

Example (part 2): If we make a dummy refinement of $\{< 2_D, 2R >\}$ and replace this condition by the conditions $\{< 1_D, r >, < 2_D, 2_R >\}_{r \in R \setminus \{2_R\}}$ we get a collection $\mathcal{C}'$ of conditions which can be organized as a decision tree. So the dummy refinement allowed us to glue the two connection components together. Now notice that $\mathcal{C}' = \mathcal{C} + (2_D?, 2_R)(1_D?) - (2_D?)$ so $\mathcal{C} = \mathcal{C}' - (2_D?, 2_R)(1_D?) + (2_D?)$. Now we can actually write $\mathcal{C}'$ as $[1_D; 2_R] - (1_D?, 1_R)(2_R?) + (1_D?, 1_R)(2_D?)$. Thus $\mathcal{C} = [1_D; 2_R] - (1_D?, 1_R)(2_R) + (1_D?, 1_R)(2_D?) - (2_D?, 2_R)(1_D?) + (2_D?)$. To check this directly notice that the condition $\{< 1_D, 2_R >\}$ appears in each tree in $[1_D; 2_R], (1_D?, 1_R)(2_R)$ and $(1_D?, 1_R)(2_D?)$. So counted with signs it appears 1 time. The condition $\{< 2_D, 2_R >\}$ appears only in $(2_D?)$. The conditions $\{< 1_D, 1_R >, < 2_D, r_R >\}$ $r_R \neq 2_R$ appears in $(1_D?, 1_R)(2_D?)$ but in no other trees. The condition $\{< 1_D, r_R >, < d_D, 2_R >\}$ $d_D \neq 2_D$, $(d_D \neq 1_D,\ r_R \neq 1_R, 2_R)$ appears only in $[1_D; 2_R]$. So all conditions in $\mathcal{C}$ appear 1 time in the linear combination. The condition $\{< 1_D, 1_R >, < d_D, 2_R >\}$ $d_D \neq 2_D$ do not appear in $\mathcal{C}$. It appears in $[1_D; 2_R]$ and $(1_D?, 1_R)(2_R)$ which appears with opposite sign. Finally notice (after having checked the remaining cases) that the right hand side contains a surplus of 1 tree. ♣

The example shows that a basis $\mathcal{C}$ need not consist of conditions which can be collected as a tree. On the other hand the example also shows that there is a linear combination $\Sigma_i\ (-1)^{\tau(i)} T_i$ of trees (with signs) such that $\Sigma_i\ (-1)^{\tau(i)} = 1$, and such that $\mathcal{C}$ and $\Sigma_i\ (-1)^{\tau(i)} T_i$ contain the same conditions (counted with sign). It turns out that this holds in general so actually any $|D|/2$-basis can be expressed as a linear combination of (D, R)-labeled trees.

Lemma 30 *Assume that $\mathcal{C}$ is a $|D|/2$-basis of conditions of length $\leq h$. Then there exists a linear combination $\Sigma_i\ \lambda_j\ T_j\ \ (\lambda_j \in \mathbf{Z})$ of trees $T_1, T_2, \ldots, T_u$ such that*
(i) $\Sigma_i\ \lambda_j = 1$
(ii) *The trees T_j have all height $\leq h\ (\leq h^2)$.*
(iii) *$\mathcal{F}_{\mathcal{C}}$ and $\mathcal{C}$ contain the same conditions (when these are counted with signs and multiplicity).*

Robustness: The lemma shows that the choice of $l = |D|/2$ is quite arbitrary. What really matters is that l do not get to close to 0 or $|D|$ in terms of square of heights of the trees.

A element $d \in D$ $(r \in R)$ is a *semi-root* for $\mathcal{C}$ if for each $r \in R$ $(d \in D)$ such that $(d, r) \in \alpha$. We say $d \in D$ $(r \in R)$ is a *root* if it is a semi-root, and it appears in all conditions in $\mathcal{C}$. Notice that a root can always serve (though it need not) as the root question in a decision tree.

Lemma 31 *Any $|D|/2$-basis $\mathcal{C}$ has a semi-root. Actually if $\alpha = \{< d_1, r_1 >, \ldots, < d_t, r_t >\} \in \mathcal{C}$, then there are semi-roots $u_1, u_2, \ldots, u_t$, where $u_i \in \{d_i, r_i\}$.*

Proof: Let $\alpha = \{< d_1, r_1 >, \ldots, < d_t, r_t >\} \in \mathcal{C}$. Assume that neither d_1 or r_1 are semi-roots. Then there exist r' and d' such that $\{< d_1, r' >\}$ and $\{< d', r_1 >\}$ do not appear in any $\beta \in \mathcal{C}$. Let α' consist of the pairs in α which are compatible to $\{< d_1, r' >, < d', r_1 >\}$. Let $\rho := \{< d_1, r' >, < d', r_1 >\} \cup \alpha'$. Notice that if ρ is compatible to $\beta \in \mathcal{C}$, then β and α' must be compatible. As β does not contain the pairs $\{< d_1, r' >\}$ and $\{< d', r_1 >\}$ none of the elements d_1, d', r_1 and r' belongs to pairs in β. But then β must be compatible to α which contradicts the assumption that all conditions in $\mathcal{C}$ are pairwise incompatible. □

In the example 1_D was a semi-root (and so was 2_D and 2_R). The conditions which did not contain 1_D (i.e. $\{< 2_d, 2_R >\}$) formed a connection component. We can view $\mathcal{C}$ to consist of 2 connection components. The dummy refinement leading to $T_{\mathcal{C}}$ 'glued' the two connection components together. In general a basis $\mathcal{C}$ of conditions of length $\leq h$ can contain arbitrarily many connection components. However the fact that all conditions are pairwise incompatible ensures that any sequence of dummy extensions (each of which reduces the number of connection components by one) will never extend any condition beyond a length of h^2.

First I show a weakened version of lemma 30. In this version the height h of the trees is only required to be bound by h^2. This version is actually sufficient to prove most of our general results.

Proof of weakened version of lemma 30: Notice that $\mathcal{C} = \mathcal{C} + \alpha - \alpha$. Suppose that we obtain $\mathcal{C}'$ from $\mathcal{C}$ by replacing $\alpha \in \mathcal{C}$ by 'dummy' extensions in a point u. If we chose a suitable u we reduce the number of connection components by one. If $u = d$ for some $d \in D$ we have $\alpha \cup \{< d, r >: r \in R \setminus \mathrm{ran}(\alpha)\}$. If $u = r$ for some $r \in R$ we have $\alpha \cup \{< d, r >: d \in D \setminus \mathrm{dom}(\alpha)\}$. In both cases we have the equation: $\mathcal{C}' = \mathcal{C} + (d_1, r_1)(d_2, r_2)\ldots(d_h, r_h)(u) - (d_1, r_1)(d_2, r_2)\ldots(d_{h-1}, r_{h-1})(d_h)$ where $\alpha = \{< d_1, r_1 >, < d_2, r_2 >, \ldots, < d_h, r_h >\}$. If $\mathcal{C}'$ only contains one connection component it is a decision tree and thus $\mathcal{C}$ can be written as a linear combination $T' - T_1 + T_2$ of decision trees (just let $T' := \mathcal{C}', T_1 := (d_1, r_1)(d_2, r_2)\ldots(d_h, r_h)(u)$ and $T_2 := (d_1, r_1)(d_2, r_2)\ldots(d_{h-1}, r_{h-1})(d_h)$).

Suppose that the conditions in $\mathcal{C}'$ cannot be organized into a decision tree. There must be $\beta \in \mathcal{C}'$ and a point $d \in D$ (or $r \in R$) such that if we replace β by all dummy extensions $\alpha \cup \{< d, r >\}$ where $r \in R \setminus \mathrm{ran}(\beta)$ (or $d \in D \setminus \mathrm{dom}(\beta)$) then the resulting collection $\mathcal{C}''$ contain one less connection component. Again by the same idea as before we can write $\mathcal{C}'' = \mathcal{C}' + T_3 - T_4$ where

$T_3 := (d'_1, r'_1)(d'_2, r'_2)\ldots(d'_h, r'_h)(u)$, where $T_4 := (d'_1, r'_1)(d'_2, r'_2)\ldots(d'_{h-1}, r'_{h-1})(d'_h)$ and where $\beta = \{< d'_1, r'_1 >, < d'_2, r'_2 >, \ldots, < d'_h, r'_h >\}$. In the case $\mathcal{C}''$ is a decision tree T'' we have $\mathcal{C} = T'' - T_1 + T_2 - T_3 + T_4$.

As I already pointed out, this procedure terminates before any condition gets length $h(h-1)/2$ (or just h^2). Thus eventually we construct a forest of trees $T_1, T_2, \ldots, T_u$ which (counted with sign) contains one tree and which contains the same conditions as $\mathcal{C}$. □

For the sake of completeness and to make sure later results appear in their full strength let me show the full version of lemma 30.

Example (part 3): Now consider $\mathcal{C}^*$ which contains the same conditions as $\mathcal{C}$ except that $\{< 2_D, 2_R >\}$ have been replaced by all conditions of the form $\{< 2_D, 2_R >, < 3_D, r_R >\}$ where $r_R \in R \setminus \{2_R\}$. Notice that $\mathcal{C}^*$ is a basis, and that the conditions in $\mathcal{C}^*$ can not be organised as a decision tree. If we follow the approach above and make dummy refinements trying to make 1_D into a root, we cannot keep all trees down below a height of 3. However it turns out that we can write $\mathcal{C}^* = [1_D; 2_R] + (2_R?, 2_D)(3_D?) - (2_R?, 2_D)(1_D?) + (1_D?, 1_R)(2_R?) - (1_D?, 1_R)(2_D?)$. All trees have height 2 so our refinement method is not optimal. ♣

The idea to solve the general case is to avoid making any refinements! It turns out that this can be achieved by repeatedly reorganising the tree during the construction.

Proof of lemma 30: First notice that any basis of height 2 can be written as a linear combination of trees T_j of height ≤ 2 such that the linear combination contains the surplus of exactly one tree. Assume (as part of the induction assumption) that there exists a basis $\mathcal{C}$ of conditions of length $\leq h$ which cannot be written as a linear combination of trees of height h. Clearly $\mathcal{C}$ must contain some conditions of length ≥ 3. We now embed $\mathcal{C}$ into a decition tree which might get height somewhat higher that h. Pick any semi-root $d \in D$ (or $r \in R$). For each $r \in R$ (or $d \in D$) consider the collection $\mathcal{C}^{\{<d,r>\}}$ of conditions in $\mathcal{C}$ which are compatible to $\{< d, r >\}$. Pich a semi-root for $\mathcal{C}^{\{<d,r>\}}$. At any stage we have constructed a tree where each node has associated a condition α. If there exists $\beta \in \mathcal{C}$ for which $\beta \not\subseteq \alpha$ we can choose a new semi-root for $\mathcal{C}^{\alpha}$. Notice that each condition $\alpha \in \mathcal{C}$ have one pair $< d', r' >$ assigned to an edge leading to a leaf. Now pick any δ which is assigned to a grandfather of a leaf. The basis $\mathcal{C}^{\delta}$ consists of conditions which have height ≤ 2. Thus we write $\mathcal{C}$ as a linear combination $\Sigma_j\ \lambda_j \mathcal{C}_j$ with $\Sigma_j\ \lambda_j = 1$ and where each $\mathcal{C}_j$ have one more grandfather node (than $\mathcal{C}$) for which $\mathcal{C}^{\delta}$ is a decition tree. Let $T(\mathcal{C})$ denote the maximal number of grandfather nodes for which $\mathcal{C}^{\delta}$ is a decition tree. Assume as part of the induction assumption that we chose our counter example $\mathcal{C}$ such that $T(\mathcal{C})$ takes the largest possible value among the counter examples of height $\leq h$. If $T(\mathcal{C})$ equals the number of 'leaf-grandfathers' we can get (by removing the dummy questions at the top) a basis $\mathcal{C}'$ of conditions of length $\leq h-1$ which cannot can be written as a linear combination of decition trees of height $\leq h-1$. This contradicts the induction assumption. □

A q-exceptional system $\mathcal{C}_1, \mathcal{C}_2, \ldots, \mathcal{C}_u$ (of height $\leq h$) is a collection of (D, R)-labeled $|D|/2$-bases (of heights $\leq h$) in which each condition appears 0 modulo q times while $u \neq 0$ modulo q.

Corollary 32 *There exists a q-exceptional system* $\mathcal{C}_1, \mathcal{C}_2, \ldots, \mathcal{C}_u$ *of* (D, R) *labeled* $|D|/2$*-bases of height* $\leq h$ *if and only if there exists a q-exceptional forests* $\mathcal{F}$ *of* (D, R)*-labeled trees of height* $\leq h$.

Proof: According to lemma 30, we can translate $\mathcal{C}$ into a forest $\mathcal{F}_{\mathcal{C}}$ which contains 1 modulo q

trees and which contains the same conditions (when the multiplicity is counted modulo q) as $\mathcal{C}$. All conditions in $\mathcal{F}_{\mathcal{C}}$ have length $\leq h$. Let $\mathcal{F} := \cup_j \mathcal{F}_{\mathcal{C}_j}$.

The converse implication follows trivially from the observation that the conditions in a (D,R)-labeled tree T form a (D,R)-labeled $|D|/2$-basis. □

4.6 Generic systems

Fix an integer $q \geq 2$. Let D, R and J be (big) finite sets. Assume that $|J| \neq 0$ modulo q. A (D,R,J,q)-*generic system* is an assignment $A \rightarrow \mathcal{C}_A$ which to each q-element subset $A \subseteq J$ assign a collection $\mathcal{C}_A$ of (D,R)-conditions such that $\forall j \in J \ \cup_{A \ni j} \mathcal{C}_A$ is a $|D|/2$-basis.

During my Doctoral work in Oxford I realized that the next technical lemma would 'give me everything'. To my friends I always referred to this lemma as "lemma 49". Originally I had thought "lemma 49" would be relatively easy to show (at least compared to the other parts of the proof), however the lemma surprised me in resisting any formal proofs. Eventually I decided to settle down for a smaller result for my thesis. But "lemma 49" still haunted me in my sleep. I often recalled the Danish Philosopher Piet Hein's wise words: *A problem worthy of attack bites back.*

"Lemma 49": *If $|R| - |D| = q^k$, then there are no (D,R,J,q)-generic systems in which all conditions have length $\leq k$.*

Proof: Assume that G is a (D,R,J,q)-generic system and assume that $|R| - |D| = q^k$. For each $j \in J$ let $\mathcal{C}_j := \cup_{A \ni j} \mathcal{C}_A$. As each $A \subseteq J$ have $|A| = q$ the collections $\mathcal{C}_1, \mathcal{C}_2, \ldots, \mathcal{C}_{|J|}$ must be a q-exceptional system. But then according to corollary 32 there must exist a q-exceptional forest of (D,R)-labeled trees of height $\leq k$. But according to Theorem 29 there are no q-exceptional forests of height k ($\leq 2^{k-1}$). □

We can actually improve "lemma 49" to the following equivalence!!

Theorem 33 *The following are equivalent:*

(a) *There exists J and a (D,R,J,q)-generic system of height $\leq h$*

(b) *There exists a q-exceptional (D,R)-system of height $\leq h$*

(c) *There exists a q-exceptional forest of (D,R)-labeled trees of height $\leq h$.*

Proof: (a) $\Rightarrow$ (b): For each $j \in J$ let $\mathcal{C}_j := \cup_{A \ni j} \mathcal{C}_A$. According to the definition each $\mathcal{C}_j$ is a $|D|/2$-basis. Thus if $|J| \neq 0$ modulo q the collection $\mathcal{C}_1, \mathcal{C}_2, \ldots, \mathcal{C}_{|J|}$ is a q-exceptional system.

(b) $\Rightarrow$ (c): Corollary 32.

(c) $\Rightarrow$ (b): Earlier observation.

(b) $\Rightarrow$ (a): Assume that $\mathcal{C}_j$, $j \in J$ is a q-exceptional (D,R)-system of height $\leq h$. Take all the conditions in $\mathcal{C}_1, \mathcal{C}_2, \ldots, \mathcal{C}_{|J|}$ and put them into disjoint classes each containing q identical conditions. Let $\mathcal{C}_A$ (where $A = \{j_1, j_2, \ldots, j_q\}$) consist of all the conditions α for which q-identical copies of α was chosen from $\mathcal{C}_{j_1}, \mathcal{C}_{j_2}, \ldots, \mathcal{C}_{j_q}$. We need to show that the map $A \rightarrow \mathcal{C}_A$ for each j has $\cup_{A \ni j} \mathcal{C}_A$ is a $|D|/2$-basis. But each condition in $\mathcal{C}_j$ appears exactly once in $\cup_{A \ni j} \mathcal{C}_A$. Thus $\cup_{A \ni j} \mathcal{C}_A = \mathcal{C}_j$ is a $|D|/2$-basis. □

4.7 Other combinatorial preparations

In this section we consider quite a different problem related to the sets D and R. We assume that $|D| \leq |R|$ and that $|R| - |D| \in |D|^{o(1)}$. Let $n := |D|$ and let $\mathbf{R}_p$ denote the collection of all partial bijections $\rho : D \to R$ with the probability measure which arises through the following procedure:

(i) Choose a set D' in D by picking each point with probability p

(ii) Choose randomly a set $R' \subseteq R$ such that $|D \setminus D'| = |R \setminus R'|$.

(iii) Choose randomly a partial bijection $\rho : D \to R$ with $\mathrm{dom}(\rho) = D \setminus D'$ and $\mathrm{ran}(\rho) = R \setminus R'$.

We follow [KPW 95] and will say set S of conditions *refines* a set H of conditions if for each $\alpha \in S$ either there exists $h \in H$ such that $\alpha \supseteq h$ or α is incompatible to all conditions in H. The following lemma is a minor modification of lemma 2D in [KPW 95].

Lemma 34 *Let $\mathcal{H}$ be a collection of conditions of length $\leq t$. Assume that $p < \frac{1}{100}$ and $pn \geq 40s$. Then for random $\rho \in \mathbf{R}_p$ with probability at least $1 - e(16p^4n^3t)^s - 2^{-O(1)pn}$ the following proposition holds:*

There exists a decision tree T_S of height $\leq 2s$ which refines all the conditions in $\mathcal{H}^\rho$. Furthermore, this remains true even if we add the requirement that $|\rho| \leq n - \frac{1}{2}pn$.

First a few comments. The above formulation of the lemma is slightly strengthened compared to lemma 2D in [KPW 95]. First, the conditions in the $2s$-complete system S of lemma 2D were only chosen to be a basis, while it was actually shown that S could be chosen as a decision tree. Second, the requirement that $|R| = |D| + 1$ has been weakened to the requirement that $|R| - |D| \leq n^{o(1)}$ where $o(1)$ denotes an infinitesimal non-standard rational.

For our application let $k > 5$ and t be finite numbers. Let $p = n^{1/k-1}$ and notice that the probability is at least $1 - e \cdot (1/n)^{s(1-4/k)}$. We can get rid of all negation by pushing these to the input gates (using the rules $\neg \wedge \neg \equiv \vee$ and $\neg \vee \neg \equiv \wedge$). And then using the fact the $\neg x_{d,r}$ can be 'expressed' as $\vee_{r' \neq r}\ x_{d,r'}$. Now by combining this with a standard switching lemma application we get

Lemma 35 *Assume that $\psi_1, \psi_2, \ldots, \psi_u$ is a collection of at most n^{k_1} circuits. Assume they all have size $s \leq n^{k_2}$ and depth $d \leq h$. Then there exists $s = s(k_1, k_2, d)$ and $\epsilon = \epsilon(k_1, k_2, d)$ such that if $\rho : D \to R$ is chosen randomly from $\mathbf{R}_p$ with positive probability all circuits can simultaneously be 'expressed' as disjunction of s-conjunctions.*

5 The model theoretical construction

5.1 Forcing setup

In this section we modify the construction in [Riis 93B] and [Riis 94A]. Let M be a countable non-standard model of $\mathrm{Th}(\mathbf{N})$ over a countable first order language L which extends the language of arithmetic. We have fixed sets $D := \{1, 2, ..., \bar{d}\}, R := \{1, 2, \ldots, \bar{r}\} \subseteq \mathrm{M}$, $\bar{d}, \bar{r} \in \mathrm{M} \setminus \omega$. Let L_{M} denote the language L extended with a constant c_m for each $m \in \mathrm{M}$. Let $L_{\mathrm{M}}(f)$ be L_{M} extended with an unspecified unary function symbol.
We say that $\rho : D \to R$ is a *partial bijection* if ρ maps its domain bijectively onto its range.

For $k \in \mathbf{N}$ let $\mathcal{P}_k(D,R) := \{\rho : D \to R \text{ and } (\bar{d} - |\mathrm{dom}(\rho)|)^k \geq \bar{d}\}$. We define $\mathcal{P}(D,R) := \cup_{k\in\mathbf{N}} \mathcal{P}_k(D,R)$. The elements in $\mathcal{P}(D,R)$ are ordered under inclusion. An element $\rho \in \mathcal{P}(D,R)$ is called a *(forcing) condition.* We use letters $\mathcal{A},\mathcal{B},\mathcal{C},...$ to denote subsets of $\mathcal{P}(D,R)$. When (D,R) is clear from the context we write $\mathcal{P}_k := \mathcal{P}_k(D,R)$ and let $\mathcal{P} := \mathcal{P}(D,R)$.

Notice that $\mathcal{P}_1 \subseteq \mathcal{P}_2 \subseteq .... \subseteq \mathcal{P}_r \subseteq .... \subseteq \mathcal{P}$, for each $r \in \omega$. The idea is to use $(\mathcal{P},\subseteq)$ as the set of forcing conditions.

We say that $\mathcal{D} \subseteq \mathcal{P}$ is *dense* if $\forall g \in \mathcal{P} \exists h \in \mathcal{D}\ h \supseteq g$.
We say that $\mathcal{D}$ is *quasi-definable* if there exists a formula $\theta(x) \in L_{\mathbf{M}} \cup \{R_\omega\}$ such that $\mathcal{D} := \{m \in \mathbf{M} :\ \mathbf{M} \models \theta(m)\}$ (the relation R_ω is defined by $R_\omega(a) \leftrightarrow a \in \omega$). Notice as an example that $\mathcal{P}$ is dense and quasi-definable (although $\mathcal{P}$ not is $L_{\mathbf{M}}$-definable). We say that $\rho_G \subseteq \mathcal{P}$ is a *generic filter* if (i) $\forall \alpha \in \rho_G \forall \beta \in \mathcal{P}\ \beta \subseteq \alpha \to \beta \in \rho_G$, (ii) $\forall \alpha,\beta \in \rho_G \exists \gamma \in \rho_G\ \gamma \supseteq \alpha \wedge \gamma \supseteq \beta$, and (iii) For $\mathcal{D} \subseteq \mathcal{P}$ dense and quasi-definable $\rho_G \cap \mathcal{D} \neq \emptyset$.

We use the abbreviation $\tilde{\rho}_G := \cup_{\alpha\in\rho_G} \alpha$.

Lemma 36 *If $\rho_G \subseteq \mathcal{P}$ is a generic filter, then $\tilde{\rho}_G : D \to R$ is a bijection.*

Proof: The only problem is to show $\mathrm{dom}(\tilde{\rho}_G) = D$ and $\mathrm{ran}(\tilde{\rho}_G) = R$. For an arbitrary $d \in D$ and $r \in R$ let $\mathcal{D}_{d,r} := \{\alpha \in \mathcal{P} :\ d \in \mathrm{dom}(\alpha) \wedge r \in \mathrm{ran}(\alpha)\}$. Notice that $\mathcal{D}_{d,r}$ is dense and quasi-definable so $\mathcal{D}_{d,r} \cap \rho_G \neq \emptyset$. Thus for each $d \in D$ and $r \in R$ there exists $\alpha_{d,r} \in \mathcal{D}_{d,r} \cap \rho_G$, and thus $d \in \mathrm{dom}(\tilde{\rho}_G)$ and $r \in \mathrm{ran}(\tilde{\rho}_G)$. □

Lemma 37 *For each $\rho_0 \in \mathcal{P}$ there exists a generic filter $\rho_G \subseteq \mathcal{P}$ such that $\rho_0 \in \rho_G$.*

Proof: Recall that both $\mathbf{M}$ and L are assumed to be countable, so there are only countably many quasi-definable dense sets. Let these be $\mathcal{D}_1,\mathcal{D}_2,....$ According to the definition of denseness there exists a sequence of conditions $\rho_1 \subseteq \rho_2 \subseteq .... \in \mathcal{P}$ with $\rho_j \in \mathcal{D}_j$, $j = 1,2,...$ and $\rho_1 \supseteq \rho_0$. Clearly $\rho_0 \in \rho_G := \{\rho :\ \rho \subseteq \rho_k \text{ for some } k \in \omega\}$ is a generic filter. □

For a sentence $\psi \in L_{\mathbf{M}}(P)$ we define the *forcing relation* $|\vdash$ by letting
$\rho| \vdash \psi$ iff $(\mathbf{M},\tilde{\rho}_G) \models \psi$ for all generic filters $\rho_G \ni \rho$.

Lemma 38 *If $(\mathbf{M},\tilde{\rho}_G) \models \psi$ for a generic filter ρ_G, then there exists $\rho_0 \in \rho_G \subseteq \mathcal{P}$ such that $\rho_0| \vdash \psi$.*

Proof: By use of induction on the logical complexity of a general formula $\psi(\vec{x})$, it is not hard to show that $\{(\vec{a},\rho) \in \mathbf{M}^r \times \mathcal{P} :\ \rho| \vdash \psi(c_{\vec{a}})\}$ is quasi-definable. Continuing this argument for each $L_{\mathbf{M}}(P)$-sentence ψ, we notice that $\mathcal{D} := \{\rho \in \mathcal{P} :\ \rho| \vdash \psi \vee \rho| \vdash \neg\psi\}$ is both quasi-definable and dense. For the required ρ_0 take any $\rho_0 \in \rho_G \cap \mathcal{D}$. □

Lemma 39 *For each bounded $\psi \in L_{\mathbf{M}}(f)$, $\rho| \vdash \psi$ iff $\rho| \vdash (\epsilon_\psi)^P = 1$.*

Proof: Induction on the number of logical constants in ψ. □
Recall that two conditions α and β are *incompatible* $(\alpha \perp \beta)$ if there exists $d \in D : \alpha(d) \neq \beta(d)$ or there exists $r \in R : \alpha^{-1}(r) \neq \beta^{-1}(r)$. Two conditions α and β are *compatible* $(\alpha||\beta)$ if they not are incompatible. A subset $\mathcal{B} \subseteq \mathcal{P}$ is *orthogonal* if $\forall \alpha,\beta \in \mathcal{B}\quad \alpha \neq \beta \to \alpha \perp \beta$ and is *complete* if

$\forall \rho \in \mathcal{P} \exists \alpha \in \mathcal{B}\ \rho || \alpha$. A *basis* is a collection $\mathcal{B} \subseteq \mathcal{P}$ which satisfies both these conditions (i.e. is both Orthogonal and Complete). Finally we let $||\mathcal{B}|| := \max_{\beta \in \mathcal{B}}(|\mathrm{dom}(\beta)|)$.
The next lemma allows us to repeat estimates in a scaled down version. More specifically it allows us to assume that $\emptyset| \vdash \psi$ in cases where $\rho_0| \vdash \psi$ for some $\rho_0 \in \mathcal{P}$. This is because the lemma allows us to replace (D, R) by (D', R') where $\bar{d}' := \bar{d} - |\mathrm{dom}(\rho_0)|$ and then smoothly pass from $\mathcal{P}(D, R)$ to $\mathcal{P}(D', R')$.

Lemma 40 *Fix $\rho \in \mathcal{P}$, let $D' := D \setminus \mathrm{dom}(\rho)$ and let $\bar{d}' := |D'|$. Define $\mathcal{P}_k(D', R') := \{\tilde{\rho} : \tilde{\rho}$ is a partial bijection of D' and $(\bar{d}' - |\mathrm{dom}(\tilde{\rho})|)^k \geq \bar{d}'\}$. Let $\mathcal{P}(D', R') := \cup_{k \in \omega}\ \mathcal{P}_k(D', R')$. Then $\mathcal{P}(D', R') = \mathcal{P}^\rho$, where $\mathcal{P}^\rho := \{\tilde{\rho} :\ \tilde{\rho} : D' \to R'$ is a partial bijection and $\tilde{\rho} \cup \rho \in \mathcal{P}\}$.*

Proof: First, we show $\mathcal{P}^\rho \subseteq \mathcal{P}(D', R')$. Suppose that $\tilde{\rho} \in \mathcal{P}^\rho$. There exists $k_0 \in \omega$ such that $\bar{d}' \leq \bar{d} \leq (\bar{d} - |\mathrm{dom}(\tilde{\rho} \cup \rho)|)^{k_0} = (\bar{d} - |\mathrm{dom}(\rho)| - |\mathrm{dom}(\tilde{\rho})|)^{k_0}$
$= (\bar{d}' - |\mathrm{dom}(\tilde{\rho})|)^{k_0}$. So $\tilde{\rho} \in \mathcal{P}_{k_0}(D', R') \subseteq \mathcal{P}(D', R')$.
Second, we show that $\mathcal{P}(D', R') \subseteq \mathcal{P}^\rho$. Suppose that $\tilde{\rho} \in \mathcal{P}(D', R')$. There exists $k \in \omega$ such that $\bar{d}' \leq (\bar{d}' - |\mathrm{dom}(\tilde{\rho})|)^k$. As $\rho \in \mathcal{P}$ there exists $l \in \omega$ such that $(\bar{d} - |\mathrm{dom}(\rho)|)^l \geq \bar{d}$. Thus we must have $(\bar{d} - |\mathrm{dom}(\tilde{\rho} \cup \rho)|)^{kl} = (\bar{d} - |\mathrm{dom}(\rho)| - |\mathrm{dom}(\tilde{\rho})|)^{kl} = (\bar{d}' - |\mathrm{dom}(\tilde{\rho})|)^{kl} \geq (\bar{d}')^l = (\bar{d} - |\mathrm{dom}(\rho)|)^l \geq n$.
Thus $\tilde{\rho} \cup \rho \in \mathcal{P}$ and $\tilde{\rho} \in \mathcal{P}^\rho$. □

Lemma 41 *Suppose that $\mathcal{B}$ is a basis for $\mathcal{P}$ and $\mathcal{H} \subseteq \mathcal{B}$. Suppose also that $||\mathcal{B}|| \in \omega$. Then*
(a) $\rho| \vdash (\vee_{h \in \mathcal{H}}\ h)^P = 1$ *iff ρ is incompatible with all conditions $h' \in \mathcal{B} \setminus \mathcal{H}$.*
(b) $\rho| \vdash (\neg \vee_{h \in \mathcal{H}}\ h)^P = 1$ *iff ρ is incompatible with all conditions $h' \in \mathcal{H}$.*

Proof: $(a) \Rightarrow$: Suppose that $\rho| \vdash (\vee_{h \in \mathcal{H}}\ h)^P = 1$, but ρ is compatible with $h' \in \mathcal{B} \setminus \mathcal{H}$. Now $\rho' := \rho \cup h' \in \mathcal{P}$ and as $\mathcal{B}$ is a basis h' must be incompatible to all conditions in $\mathcal{H}$. Clearly $\rho' \supseteq h'$ so ρ' is also incompatible with all conditions in $\mathcal{H}$. But then $(\vee_{h \in \mathcal{H}}\ h)^{\tilde{\rho}_G} = 0$ for each generic filter $\rho_G \ni \rho'$ (which exists by lemma 37). This contradicts $\rho| \vdash (\vee_{h \in H}\ h)^P = 1$.
$(a) \Leftarrow$: Assume that ρ is incompatible with all $h' \in \mathcal{B} \setminus \mathcal{H}$. Let $\rho_G \ni \rho$ be any generic filter (which exists by lemma 37).
Let $\mathcal{D} := \{\rho' \in \mathcal{P} :\ (\exists h' \in \mathcal{H}\ h' || \rho')$ or $\rho' \perp \rho\}$. Notice that $\mathcal{D} \subseteq \mathcal{P}$ is dense and quasi-definable. There exists $\alpha \in \mathcal{D} \cap \rho_G$, so there exists $h \in \mathcal{H}$ with $h \subseteq \alpha \subseteq \tilde{\rho}_G$.
$(b) \Rightarrow$ / $(b) \Leftarrow$ are proved by proofs very similar to $(a) \Rightarrow$ / $(a) \Leftarrow$. □

Lemma 42 *Let $\epsilon_1, \epsilon_2, ..., \epsilon_u\ \ u \in \mathbf{M}$, be an $\mathbf{M}$-definable sequence of Boolean circuits, each of the form $\epsilon_j :\equiv \vee_{h \in \mathcal{H}_j}\ h$. Let $\mathcal{B}_1, ..., \mathcal{B}_u$ be an $\mathbf{M}$-definable sequence and suppose that $t \in \omega$ such that:*
(a) *for each $j = 1, 2, ...u\ \ \mathcal{B}_j \subseteq \mathcal{P}$, is a basis for $\mathcal{P}$,*
(b) *for each $j = 1, 2, .., u\ \ ||\mathcal{B}_j|| < t$,*
(c) *for each $j = 1, 2, ..., u,\ \mathcal{H}_j \subseteq \mathcal{B}_j$.*

Then for every generic filter ρ_G either

(a) *for all $j \in \{1, 2, .., u\},\ \epsilon_j^{\tilde{\rho}_G} = 0$, or*
(b) *there exists $j_0 \leq u$ such that $\epsilon_{j_0}^{\tilde{\rho}_G} = 1$ and $\epsilon_j^{\tilde{\rho}_G} = 0$ for each $j < j_0$.*

Proof: Let $\mathcal{D} := \{\rho \in \mathcal{P} : (\exists j_0 \exists \beta \in \mathcal{H}_{j_0}\ \beta || \rho \wedge \forall \gamma \in \cup_{j<j_0} \mathcal{H}_j\ \rho \perp \gamma)$ or $(\forall \gamma \in \cup_{j \leq u} \mathcal{H}_j\ \rho \perp \gamma)\}$. Clearly $\mathcal{D}$ is quasi-definable. For each $\rho_0 \in \mathcal{P}$, if ρ_0 is compatible with some $\beta \in \cup_j \mathcal{H}_j$, then there must be a smallest j_0 such that ρ_0 is compatible with some $\beta \in \mathcal{H}_{j_0}$. Here we use the least number principle which is valid in M. Now $\rho := h \cup \rho_0 \in \mathcal{P}$, and thus $\rho \in \mathcal{D}$. So $\mathcal{D}$ is dense. Thus there exists $\rho \in \rho_G \cap \mathcal{D}$. This condition ρ is incompatible with all $h \in \mathcal{H}_j$, $j < j_0$. As $\tilde{\rho}_G \supseteq \rho \supseteq h \in \mathcal{H}_{j_0}$ clearly $(\vee_{h \in \mathcal{H}_{j_0}} h)^{\tilde{\rho}_G} = 1$. □

Recall that M is a countable non-standard model of Th(N) over a countable first order language L. As above we have fixed $D, R \subseteq \mathrm{M}$ with both $|D|$ and $|R|$ non-standard numbers. As above the set $\mathcal{P}$ of forcing conditions consists of partial bijections $\rho : D \to R$ with $|\mathrm{dom}(\rho)| \leq \bar{d} - \bar{d}^{\frac{1}{\omega}}$ for some $k \in \omega$. Now a direct application of lemma 35 gives

Lemma 43 (key lemma) *Let $\theta_1, \theta_2, ..., \theta_u$ be an* M*-definable sequence of depth $\leq d \in \omega$ circuits of size bounded by a fixed polynomial in n.*
Let $\rho_0 \in \mathcal{P}$. There exists $\rho \supseteq \rho_0$, $\rho \in \mathcal{P}$ and an M*-definable sequence $\epsilon_1, \epsilon_2, ..., \epsilon_u$ of circuits together with an* M*-definable sequence $\mathcal{B}_1, \mathcal{B}_2, .., \mathcal{B}_u$ and $s \in \omega$ such that*
(a) *for $j = 1, 2, .., u$ each $\mathcal{B}_j$, is a basis for $\mathcal{P}$,*
(b) *for $j = 1, 2, ...u$ each ϵ_j is of the form $\vee_{h \in \mathcal{H}_j} h$ for some $\mathcal{H}_j \subseteq \mathcal{B}_j$,*
(c) *for each $j = 1, 2, ..., u$, θ_j and ϵ_j holds in the same generic extensions of ρ.*
(d) *for each $j = 1, 2, ..., u$ $||\mathcal{B}_j|| \leq s$.*

If we combine the key lemma with lemma 42 we get:

Corollary 44 *If $\theta_1, \theta_2, ...\theta_u$ is an* M*-definable sequence of depth $d \in \omega$ circuits of size bounded in a fixed polynomial in n, then for any generic filter $\rho_G \subseteq \mathcal{P}$ either*
(a) *for all $j \leq u$ $\theta_j^{\tilde{\rho}_G} = 1$, or*
(b) *there exists $j_0 \leq u$, such that $\theta_{j_0}^{\tilde{\rho}_G} = 1$ and $\theta_j^{\tilde{\rho}_G} = 0$ for all $j < j_0$.*

Corollary 45 *If $A \to \theta_A$ is an* M*-definable map which maps every q-element subset $A \subseteq J$ into a depth d formula θ_A which is bounded by some fixed polynomial in $\bar{d}$. For each $\rho_0 \in \mathcal{P}$ there exists $\rho \supseteq \rho_0$, there exists $s \in \omega$, and a* M*-definable map $A \to \mathcal{C}_A$ which maps every q-element subset $A \subseteq J$ into a collection of conditions $\mathcal{C}_A$ of length s such that θ_A and $\vee_{h \in \mathcal{C}_A} h$ holds in the same generic extensions of ρ. In addition it is possible to ensure that the conditions in each $\mathcal{C}_A$ are pairwise incompatible.*

Proof: An elementary reformulation of lemma 43. □

Corollary 46 *Suppose that there exists an* M*-definable map $A \to \theta_A$ of depth d circuits which are all bounded by some polynomial in $\bar{d}$, such that $\rho_0 | \vdash$ "$A \to \theta_A$ defines a partitioning of J into disjoint q-element subsets and $|J| \neq 0$ modulo q". Then there exists a (D, R, q, J)-generic system.*

Proof: Suppose that there exists an M-definable map $A \to \theta_A$ of depth d circuits of polynomial size. According to corollary 45 there exists $s \in \omega$, and a M-definable map $A \to \mathcal{C}_A$ which maps every q-element subset $A \subseteq J$ into a collection $\mathcal{C}_A$ of conditions of length s such that θ_A and $\epsilon_A := \vee_{h \in \mathcal{C}_A} h$

hold in the same generic extensions of ρ. I claim that for each $j \in J$ $\mathcal{C}_J := \cup_{A \ni j} \mathcal{C}_A$ is a $|D|/2$-basis. To prove this claim first notice that all conditions in $\mathcal{C}_j$ must be pairwise incompatible. If $\mathcal{C}_j$ contained two (different) conditions which were compatible, there would be sets $A, B \ni j$ with $A \neq B$ such that some $\alpha \in \mathcal{C}_A$ are compatible to some $\beta \in \mathcal{C}_B$. But then $\alpha \cup \beta \cup \rho_0| \vdash$ "$A \to \theta_A$ defines a partitioning which contains both A and B" violating the assumption that ρ_0 forces the map $A \to \theta_A$ to define a partitoning of J into disjount q-element subsets. Second notice that it suffices to show that $\mathcal{C}_j$ is a $|D \setminus \mathrm{dom}(\rho_0)|/2$ basis, as this would imply that it is a $|D \setminus \mathrm{dom}(\rho_0)| - s^2$ basis (using the earlier robustness results). But suppose that there exists $\rho \in \mathcal{P}^{\rho_0}$ with $|\rho| \leq |D \setminus \mathrm{dom}(\rho_0)|/2$ which is incompatible to all conditions in $\mathcal{C}_j$. This is a contradiction because $\rho_0 \cup \rho \in \mathcal{P}$ and in no generic extension of $\rho_0 \cup \rho$ would j belong to any A in the partitioning. □

Notice that we have the converse in the sense that:

Lemma 47 *Assume that there exist a (D, R, q, J)-generic system $A \to \mathcal{C}_A$ where all conditions have length bounded by some standard number s.*
Let $\mathrm{Count}_{\mathrm{s}}(q) :\equiv \bigvee_{j \in J} \bigwedge_{A \ni j} \neg \epsilon_A \ \vee \ \bigvee_A \bigvee_{A \neq B, A \cap B \neq \emptyset} (\epsilon_A \wedge \epsilon_B)$ *where* $\epsilon_A := \vee_{h \in \mathcal{C}_A} h$. *Then any bijection $f : D \to R$ produces a truth-tabel evaluation violation which makes* $\mathrm{Count}_{\mathrm{s}}(q)$ *false.*

5.2 Proof of the main theorem

Proposition 48 *Let* **P** *be an absolute propositional proof system. Then there are no polynomial size bounded depth* **P***-proofs of* $\mathrm{PHP}^*_{*+q^{w(1)}}(\mathrm{bij})$.

Proof: Assume that there exists $d \in \omega$ such that for arbitarily large n there exists a sequence $\psi_1, \psi_2, \ldots, \psi_{u_n}$ of depth d formulas proving $\mathrm{PHP}^n_{n+q^{r(n)}}(\mathrm{bij})$. By the compactness theorem there would be a countable non-standard model M of first order arithmetic which for some non-standard number n there would be a sequence $\psi_1, \psi_2, \ldots, \psi_{u_n}$ of depth d formulas which proves $\mathrm{PHP}^n_{n+q^{r(n)}}(\mathrm{bij})$ (viewed within M). But then according to corollary 44 we must have $j_0 < u$ such that $\psi_{j_0}^{\tilde{\rho} G} = 1$ and $\psi_j^{\tilde{\rho} G} = 0$ for all $j < j_0$. All axioms are absolute so ψ_{j_0} cannot be an axiom. All deduction rules are absolute so ψ_{j_0} cannot be a consequence of any deduction. This is a contradiction. □

Theorem 49 *Let* M *be a countable non-standard model of Th*(N) *over a countable language which extends the language of arithmetic. Let $q \geq 2$ be a standard integer. Let $D := \{1, 2, \ldots, \bar{d}\}, R := \{1, 2, \ldots, \bar{r}\}$ be fixed initial segments of non-standard length. Assume that $\bar{r} - \bar{d}$ is a non-standard power of q and assume that $\bar{r} - \bar{d} < n^\delta$ for some non-standard infinitesimal δ. Let $\mathcal{P}$ the set of forcing conditions be defined as above, and let f be any generic bijection $f : D \to R$. There are no (D, R, q, J)-generic system if and only if* (M, f) *satisfy the Count(q) principle.*

Proof: Suppose that there exists a (D, R, q, J)-generic system i.e. suppose that there exists a M-definable map $A \to \mathcal{C}_A$ such that $\cup_{A \ni j} \mathcal{C}_A$ is a $|D|/2$-basis. Now consider the collection $\{A : (\mathrm{M}, f) \models$ 'f is compatible to some $\alpha \in \mathcal{C}'_A\}$. According to lemma 47 this defines a partitioning of J into disjount q-element subsets violating the $\mathrm{Count}(q)$ principle in (M, f).

Conversely suppose that there are no (D, R, q, J)-generic systems. But then according to lemma 45 and lemma 46 for no $d \in \omega$ does there exists a M-definable map $A \to \theta_A$ which assigns depth d

circuits of size bounded by a polynomial in $|D|$ which can be forced to violate the Count(q)-principle. □

Proof of Theorem 9: It suffices to show that the Count(q)-axiom sceme are always forced valid. The Count(q)-axiom sceme is not absolute so propostion 48 does not apply. However if Count(q) is forced false it follows from Theorem 49 that there exists a (D, R, J, q)-generic system where all the conditions have length bounded by some fixed standard number h. Yet according to Theorem 33 this is possible if and only if there exists a q-exceptional forest of (D, R)-labeled trees of height $\leq h$. But according to Theorem 29 (our classification of the q-exceptional forest) there are no q-exceptional forest of height $\leq 2^{l-1}$ when $|R| - |D| = q^l$. Thus if (and only if) l is a non-standard number each instance of the Count(q) axiom scheme always get forced true. □

Proof of Theorem 10: There exist a q-exceptional forest of (D, R)-labeled trees of height q^k (Theorem 29). Thus there exists of a (D, R, q, J)-generic system (Theorem 33) where all conditions have length bounded by q^k. This gives us a substitution instance $\text{count}_s(q)$ which comes out false for any truth-tabel evaluation induced by any bijection $f : D \to R$ (lemma 47). □

To end I would like to thank J. Krajicek and P. Pudlak, for their many useful comments as well as for their hospitality during my numerous visits in Prague. I would also like thank M. van Lambalgen and D. Zambella for inviting me to Amsterdam where this work was completed. Finally I would like to thank M. Ajtai. It has been a great pleasure to work on problems related to his work.

References

[Ajtai 88] M. Ajtai; On the complexity of the pigeonhole principle. 29^{th} Annual symp. on Found. Comp.Sci.(1988),pp 340-355.

[Ajtai 90] M. Ajtai; Parity and the pigeon-hole principle, in Feasible Mathematics Birkhauser, eds S.Buss and P.J. Scott (1990), pp 1-24.

[Ajtai 94] M. Ajtai; The independence of the modulo p counting principles, Proceedings 9^{th}-annual IEEE symposium on computer science (1994).

[BP 93] P. Beame, T. Pitassi; Exponential separation between the matching principle (1993) submitted

[BIKPPW 92] P. Beame, R. Impagliazzo, J. Kraijcek, T. Pitassi, P. Pudlak, and A. Woods; Exponential lower bounds for the pigeonhole principle, in: Proceedings of the 24th Annual ACM Symposium on Theory of Computing, ACM press (1992) pp 200-21.

[BIKPP 94] P. Beame, R. Impagliazzo, J. Krajicek, T. Pitassi, P. Pudlak; Lower bounds on Hilbert's Nullstellensatz and propositional proofs, submitted

[Buss 85] S.Buss; Bounded Arithmetic. Ph.D. dissertation, Princeton University, (1985). As book, Bibliopolis, Napoli (1986).

[BKPPRS 95] S. Buss, J. Krajicek, T. Pitassi, P. Pudlak, A. Razborov, J. Sergal; Polynomial bound on Nullstellensatz for counting principles (November 95).

[Ed: CK 93] P. Clote, J. Krajicek; Open problems, in: Arithmetic, Proof theory and computorial complexity, Oxford university press (1993) pp 1-19.

[Krajicek 95] J. Krajicek; Bounded Arithmetic, Propositional logic, and complexity theory, Encyclopedia of Mathematics 60, Cambridge University Press 1995.

[KPW 95] J. Krajicek, P. Pudlak, and A.Wood, Exponential lower bound to the size of bounded depth Frege proofs of the pigeonhole principle; Random Structures and Algorithms, Vol 7, No.1 (1995)

[KPW 91] J. Krajicek, P. Pudlak, and G. Takeuti; Bounded Arithmetic and the polynomial hierarchy, Annals of Pure and Applied Logic 52 (1991), pp 143-153.

[Macintyre 86] A. Macintyre; The strength of weak systems, Proceedings of the 11^{th} international Wittgenstein symposium (1986) pp 43-59.

[Parikh 71] R. Parikh; Existence and feasibility in arithmetic, J. Symbolic Logic, 36 (1971) pp 494-508.

[PWW 88] J. Paris, A. Wilkie, A. Woods; Provability of the pigeonhole principle and the existence of infinitely many primes. Journal of Symbolic Logic 53, (1988), pp 1231-1244.

[PW 85] J. Paris, A. Wilkie; Counting problems in Bounded Arithmetic, in: Methods in Mathematical Logic, LNM 1130, Springer (1985), pp 317-340.

[PW 87] J. Paris, A. Wilkie; On the scheme of induction for bounded arithmetic formulas, Annals of Pure and Applied Logic, 35 (1987) pp 261-302.

[PBI 91] T. Pitassi, P. Beame, and R. Impagliazzo; Exponential lower bounds for the pigeonhole principle, Computational complexity, 3, pp 297-308.

[PB 93] T. Pitassi, P. Beame; An Exponential separation between the Matching Principle and the pigeonhole principle. Proceedings 8^{th}-annual IEEE symposium on computer science (1993), pp 308-319

[PB 94] P. Pudlak, S. Buss; How to lie without being convicted In Feasible mathematics; Ed: Leivant (1995)

[Riis 93A] S. Riis; Making infinite structures finite in models of Second Order Bounded Arithmetic, in: Arithmetic, Proof theory and computorial complexity, Oxford university press (1993) pp 289-319.

[Riis 93B] S. Riis; Independence in Bounded Arithmetic; DPhil dissertation, Oxford University (1993)

[Riis 94A] S. Riis; Count(q) does not imply Count(p) Report Series, BRICS RS-94-21 Aarhus Denmark (1994).

[Riis 94B] S. Riis; Finitisation in Bounded Arithmetic; Report Series, BRICS RS-94-23 Aarhus Denmark (1994)

[Riis 94C] S. Riis; Count(q) versus the Pigeon-hole principle; Report Series, BRICS RS-94-25 Aarhus Denmark (1994)

[Smale 92A] S. Smale; Theory of computation, in: Mathematical Research Today and Tomorrow, eds. C. Casacuberta and M. Castellat, Berlin and Heidelberg, Springer-verlag (1992) pp 59-69

[Smale 92B] S. Smale; Round-Table discussion in: Mathematical Research Today and Tomorrow, eds. C. Casacuberta and M. Castellat, Berlin and Heidelberg, Springer-verlag (1992) pp 91