Contemporary Mathematics
Volume **106**, 1990

RAMSEY INTERPRETED:
A PARAMETRIC VERSION OF RAMSEY'S THEOREM

Gianluigi Bellin

to Michela

ABSTRACT. The No Counterexample Interpretation (NCI) and other proof-theoretic techniques are applied to a proof of the Infinite Ramsey Theorem. A parametric form of Ramsey's Theorem is obtained, that implies the Infinite, the Finite and the Paris-Harrington versions of Ramsey's Theorem. Applications of the proof theory of fragments of **PA** and of Linear Logic are suggested. The work was a preliminary experiment for the implementation of the NCI in Ketonen's Proof Checker *EKL*.

1. Motivation.

The project of applying basic results of Proof Theory to suitable mathematical problems to obtain results of general mathematical interest and computational significance has been pursued and advocated by G.Kreisel for over 30 years (Kreisel [1989]). Significant new results have been obtained (consider, for instance, Girard [1987 a] pp.237-251 and 484-496, Luckhardt [1989]). The aim of extracting information implicit in the proof of an infinitary theorem – such as bounds on the growth of functions – may be obtained by making the "constructive content" of the proof explicit, or, using a different jargon, by a synthesis of a program from the proof. Two basic techniques of Proof Theory can have computational significance, namely Gentzen's *Cut-Elimination* and *Functional Interpretations* – of the latter the Herbrand–Kreisel "No Counterexample Interpretation" (NCI) seems most suitable for practical applications. The following type of application is of interest to us. Suppose we have a proof of a statement of the form $\forall x.\exists y.\phi(x,y)$, expressing the fact that a certain set H is unbounded. Then the above techniques allow us to define

1980 *Mathematics Subject Classification* (1985 *Revision*). 03F10, 03F30, 05C55, 68T15

I wish to thank S. Feferman, J. Ketonen, G. Kreisel, J. McCarthy, P. Scowcroft, N. Shankar, W. Sieg and C. Talcott for their suggestions and support. This work is a part of the *EKL* project in Stanford, directed by J. Ketonen and J. McCarthy. Thanks to K. Hanson for proofreading the manuscript. This research has been supported by the NSF grant CCR-8718605 and Darpa contract N00039-84-C-0211.

This paper is in final form, and no version of it will be submitted for publication elsewhere.

a function(al) F such that

$$\forall x.\phi(x, F(x)) \tag{1}$$

and at the same time provide information about its complexity.

Fruitful use of this method in practice depends on the conditions of the proof. In the simplest and best–known case ϕ is quantifier–free, and (1) is (inductively) derived from universal axioms. What happens in the case of more complex axioms? Consider a formal deduction of a sentence of the form (1) from assumptions of the form

$$\forall u.\exists v.\forall z.\theta(u, v, z) \tag{2}$$

and suppose that (2) contains the definition of the set H. The NCI functional interpretation of (2) $\Rightarrow$ (1) yields functionals U, Z and F – U and Z functionals of $\mathbf{v}$, x – such that

$$\forall \mathbf{v} x.[\theta(U, \mathbf{v}[U], Z) \supset \phi(x, F(\mathbf{v}, x))]$$

The (set of) parameter(s) $\mathbf{v}$ may be crucial in determining the set H' which is defined by $\theta(U, \mathbf{v}[U], Z)$. If we let $\mathbf{v}$ satisfy Skolem axioms of the form $\forall u.\exists v.\forall z.\theta(u, v, z) \supset \forall uz.\theta(u, \mathbf{v}[u], z)$, then H is H'. However the choice of different parameters gives us the opportunity to search for *interesting* sets H'.

In this paper we analyze the Infinite Ramsey Theorem (see Graham, et al. [1980]) and present applications that are relevant to some finite versions of Ramsey's Theorem. [1] Let $[\mathrm{N}]^k$ be the set of all k-element subsets of N. Given any c-coloring $\chi : [\mathrm{N}]^k \to c$, a set A is *χ-homogeneous* if all k-element subsets of A are monochromatic.

Infinite Ramsey Theorem. (IRT) *For every c, k and any c-coloring $\chi : [\mathrm{N}]^k \to c$, there is an infinite χ-homogeneous set H.*

Let $[n_1, n_2]^k$ be the set of k-element subsets of $\{n_1, \ldots, n_2 - 1\}$, $[n]^k = [0, n]^k$. Say that A is *large* if $|A| > min(A)$.

Finite Ramsey Theorem. (FRT) *For every c, k, l there is a number $n = R(c, k, l)$ such that for any c-coloring $\chi : [n]^k \to c$, there exist a χ-homogeneous subset A of $[n]$, with $|A| = l$.*

Ramsey–Paris–Harrington Theorem. (RPH) *For every c, k and n_1, there exists $n_2 = LR(c, k, n_1)$ such that for every coloring $\chi : [n_1, n_2]^k \to c$ there is a large χ-homogeneous set.*

We consider the case of $c = k = 2$. The Infinite Ramsey Theorem can be formally proved in a system of second order arithmetic. Since k is fixed, in this paper we use instead first order Peano Arithmetic, with additional function and predicate symbols — we need a binary function symbol for χ and a unary function symbol for the auxiliary f^{stab}; a unary predicate for H, a unary predicate for the auxiliary set H_1 and a binary predicate for the sequence of sets $S(n)$, where f^{stab}, H, H_1, $S(n)$ are defined in Section (3.1) below. Given

[1] Consideration of this example was suggested in [KREISEL 1977].

an interpretation $\mathcal{I}$ for the extended language, $H_{\mathcal{I}}$ (in the standard model) is the desired $\chi_{\mathcal{I}}$-homogeneous set, assuming that the defining properties of f^{stab}, H, H_1 and $S(n)$ are satisfied in $\mathcal{I}$.

The proof can be divided into two parts: the first part shows that a set H is unbounded and the second that H is χ-homogeneous. Part one has the form (2) $\Longrightarrow$ (1), where (1) expresses that H is unbounded, and (2) contains the formal definition of H and of the auxiliary sets, together with the assumption that χ is a 2-coloring.

The main result is the following. We give a formula expressing restrictions for a "meaningful" choice of the parameters $\mathbf{v}$, called *Non-Triviality Condition* (NTC). This must be satisfied either globally, for all n, or locally, over the segment $[x, y]$, in which case we write (NTC)$[x, y]$. In the statement of the main theorem $F^{(l)}(\chi, \mathbf{v}, n) = F(\chi, \mathbf{v}, F^{(l-1)}(\chi, \mathbf{v}, n))$ is the l-th iteration of F.

Parametrized Ramsey Theorem (PRT): *There is a functional F, primitive recursive* [2] *in χ, $\mathbf{v}$, S_F and H_1, such that for every coloring $\chi : [\mathrm{N}]^2 \to [2]$ and every choice of $\mathbf{v}$,*

(i) *if $\mathbf{v}$ satisfies the* (NTC) *and $p = F^{(l)}(\chi, \mathbf{v}, 1)$, then $[p]$ contains a χ-homogeneous set of cardinality l;*

(ii) *if $\mathbf{v}$ satisfies the* (NTC)$[x, y]$ *and $p_0, \ldots p_l \in [x, y]$, where $p_i = F^{(i)}(\chi, \mathbf{v}, p_0)$, then $[x, y]$ contains a χ-homogeneous set of cardinality l.*

Moreover, for all χ and l, there exists $\mathbf{v}$ satisfying the NTC*; in particular:*

(1) *for any fixed χ, some $\mathbf{v}$ satisfies the* NTC *for all l;*

(2) *for any fixed l, there exist p and $\mathbf{v}$ satisfying the* NTC$[0, p]$ *for all χ;*

(3) *for any fixed n_1, there exist n_2 and $\mathbf{v}$ satisfying the* NTC$[n_1, n_2]$ *for all χ, with $l = F(\chi, \mathbf{v}, n_1)$.*

Corollary 1: *The Parametrized Ramsey Theorem and Compactness*[3] *imply the Infinite Ramsey Theorem* for $c = k = 2$.

Corollary 2: *The Parametrized Ramsey Theorem implies the Finite Ramsey Theorem* for $c = k = 2$.

Corollary 3: *The Parametrized Ramsey Theorem implies the Erdös-Mills version of the Ramsey-Paris-Harrington Theorem* for $c = k = 2$.

The PRT can be regarded as a generalization of *both* the Infinite and the Finite Ramsey Theorem as shown by the corollaries.[4] This theorem provides a precise mathematical content to the following "phenomenological" remark. In the proofs of the Ramsey Theorems under consideration, two components can be distinguished:

[2] In fact, F is in $\mathcal{E}_2(\chi, S_F, H_1, \mathbf{v})$, the third class in the Grzegorczyk hierarchy relativized to χ, $\mathbf{v}$, S_F and H_1.

[3] Since we consider only countable colorings, when we speak of "Compactness" we mean an application of (Weak) König's Lemma.

[4] Remember that we cannot derive the Infinite Ramsey theorem from the Finite version by König's Lemma (at least not in a conservative extension of PA, since IRT implies RPH).

(i) the common frame of the infinite *and* the finite versions of the theorem is the structure of their recursive definitions and inductive arguments;[5]

(ii) the properties of the set H — being χ-homogeneous, or of a certain cardinality, or "large" — depend on the choice of certain set and function parameters, different parameters yielding different versions of the theorem.

Component (i) is represented by the functional $\lambda\chi\ \mathbf{v}.F$; the choices in (ii) are represented by the parameters S and $\mathbf{v}$. The NTC provides constraints for the choices in (ii) to be meaningful. The computational complexity of the proof is determined by the choices in (ii) and by the verification of the NTC, while the computational complexity of F itself is very low.

Although we have not checked it in detail, it seems clear that one can prove the PRT for arbitrary exponents k, following the pattern of the proof given below for $k = 2$.

2. Proof Theoretic Tools.

In this section we review some classical proof-theoretic results. The proofs are well-known. In sections (2.1) and (2.2) we consider some refinements of those results for fragments of **PA**.

The following proposition, together with the Cut-Elimination Theorem for Predicate Calculus, yields *Herbrand's Theorem.* Let $\mathcal{L}$ be a first order language.

Proposition. (i) *Let* Γ *be a set of* Π^0_1 *sentences and let* ψ *be a quantifier-free formula of* $\mathcal{L}$. *If* $\Gamma \vdash \forall x.\exists y.\psi(x,y)$ *is provable in* **LK** *without Cut, and* a *is a parameter not used in the derivation, then there are terms* $t_1,\ldots,t_n \in \mathcal{L}$ *such that* $\Gamma \vdash \psi(a,t_1),\ldots,\psi(a,t_n)$ *is provable in* **LK** *without Cut.*

(ii) *Let* S *be any sequent in* $\mathcal{L}$ *and let* $\mathcal{L}'$ *be an extension of* $\mathcal{L}$ *containing Herbrand functions for the essentially universal quantifiers* [6] *of* S. *In* $\mathcal{L}'$ *there are terms* $\vec{\mathbf{t}}$ *and a quantifier-free sequent* $\mathcal{E}_{\vec{\mathbf{t}}}(S)$ (the Herbrand expansion of S) *such that* S *has a normal derivation if and only if* $\mathcal{E}_{\vec{\mathbf{t}}}(S)$ *has one.*

Remark. If conditional terms `if...then...else...` can be expressed in $\mathcal{L}$, then in part (i) we can take $n = 1$ and in part (ii) there are terms $\vec{\mathbf{t}} \in \mathcal{L}'$ *(functionals of predicate logic)* such that $\mathcal{E}_{\vec{\mathbf{t}}}(S)$ is obtained from S by replacing the essentially universal variables $\vec{y}$ by Herbrand functions $\vec{\mathbf{y}}$ and the essentially existential variables $\vec{x}$ with $\vec{\mathbf{t}}$. We write S_H for $S[\vec{x}/\vec{\mathbf{t}}, \vec{y}/\vec{\mathbf{y}}]$, when the $\vec{\mathbf{t}}$ are functionals of predicate logic.

[5] Thus an alternative proof of the Finite Ramsey theorem by double induction (see Graham et al. [1980], p.3) does not belong to this family of proofs.

[6] A quantifier and its bound variable are *essentially universal* if they are universal in a positive context or existential in a negative context; otherwise, they are *essentially existential.* We will use the boldface $\mathbf{y}$ for the Herbrand function as a formal symbol of the extended language $\mathcal{L}'$, where y is associated with an essentially universal quantifier $\forall y$.

The Herbrand theorem fails if **LK** is extended by the induction rule: we need a functional interpretation, the No Counterexample Interpretation (NCI). Let $\mathcal{L}_{\mathbf{PA}}$ be the first order language of Peano Arithmetic. If $S : \Gamma \vdash \Delta$ is any sequent in $\mathcal{L}_{\mathbf{PA}}$ and $\mathcal{L}'$ is as above, the language $\mathcal{L}_F$ of the NCI is $\mathcal{L}'$ extended by functionals of type ≤ 2: $\lambda \mathbf{y}_1 \ldots \mathbf{y}_m.X(\mathbf{y}_1,\ldots,\mathbf{y}_m)$, for the essentially existential quantifiers of S. The functionals are defined by α-recursion, for $\alpha < \epsilon_0$ (*extended Grzegorczyk hierarchy*). [7] Consider the extended Ackermann function

$$\begin{aligned} f_1(x) &= 2x \\ f_{\alpha+1}(x) &= f_\alpha^{(x)}(x), \\ f_\alpha(x) &= f_{\alpha(x)}(x), \qquad \text{for } \alpha \text{ limit,} \end{aligned}$$

where $f^{(x)}$ denotes the x-th iteration of f and where $\alpha(x)$ is the x-th element in the "natural sequence" for α. A function f *dominates* g if there is a n_0 such that $f(n) > g(n)$ for all $n > n_0$. Remember that the closure of Grzegorczyk's class $\mathcal{E}_m$ under one application of the scheme of primitive recursion is $\mathcal{E}_{m+1}$ and that the Ackermann function $f_x(x)$ dominates all primitive recursive functions (see Rose [1984], Chapter 2). In general, the functions of Grzegorczyk's class $\mathcal{E}_\beta$ are dominated by f_α, with $\beta < \alpha$.

No Counterexample Interpretation. (Kreisel [1952], see also Tait [1965 a] and [1965 b]) *Let* $\phi : \forall y_1 \exists x_1 \ldots \forall y_n \exists x_n.\psi(y_1,x_1,y_2,\ldots y_n,x_n)$ *be a theorem of* **PA**, *with* ψ *quantifier-free.*[8] *Then there exist functionals* $X_1, \ldots, X_n$, *which are* α*-recursive in* $\vec{\mathbf{y}} = \mathbf{y}_1,\ldots,\mathbf{y}_n$ *such that if* ϕ_F *is*

$$\phi_F : \qquad \psi(\mathbf{y}_1, X_1(\vec{\mathbf{y}}), \mathbf{y}_2[X_1(\vec{\mathbf{y}})], \ldots, \mathbf{y}_n[X_1(\vec{\mathbf{y}}), \ldots, X_{n-1}(\vec{\mathbf{y}})], X_n(\vec{\mathbf{y}}))$$

then $\phi_F[\vec{\mathbf{y}}/\vec{f}]$ *is true for every choice of (appropriate) numerical functions* $\vec{f}$ *for* $\vec{\mathbf{y}}$.

We may choose the Herbrand functions using the least number principle: $(\forall y.\phi(y)) \equiv \phi(\mu y.\neg\phi(y))$. Then

Corollary. *Let* ϕ *be* $\forall x.\exists y.\psi(x,y)$ *and let* F *be a function(al) such that* $F(a)$ *is the least* y *such that* $\psi(a,y)$ *holds. If* F *dominates* f_α *for all* $\alpha < \epsilon_0$, *then* **PA** *cannot prove* ϕ.

Finally, as G. Kreisel has often pointed out, the following fact is very useful to sharpen the estimation of the bounds. To prove it, use the fact that the functional interpretation of a Π^0_1 lemma $\forall x.A(x)$ is $A(\mathbf{x})$, where $\mathbf{x}$ is 0-ary.

Proposition. *The functional* F *such that* $\forall x.\phi(x, F(x))$, *constructed by interpreting a given proof of* $\forall x.\exists y.\phi(x,y)$, *does not depend on the proofs of* Π^0_1 *Lemmas.*

[7] See Rose [1984]. A short summary of the notions needed for Ketonen and Solovay's proof of the Paris Harrington result can be found in Graham et al. [1980], pp.150-154.

[8] For expository purpose it is convenient to consider prenex formulas.

2.1. Functionals in fragments of PA.

We can control the complexity of the functionals, if we succeed in working within fragments of **PA**. We are interested mainly in $(\Pi^0_2 - \mathbf{IR})$, in which the induction rule is restricted to induction formulas of the form $\forall\vec{x}.\exists\vec{y}.\psi$, with ψ quantifier-free. For this fragment, formalized in classical sequent calculus, we can define the NCI directly, using explicit definitions in correspondence with the logical and structural rules and *primitive recursion* in correspondence with the induction rule. A direct definition of the interpretation for the Cut rule is possible if this is restricted to Π^0_2-formulas. Let us call the NCI for the fragment $(\Pi^0_2 - \mathbf{IR})$ *primitive interpretation*, when it is defined in the form just described.

Proposition. *In* $(\Pi^0_2 - \mathbf{IR})$ *any functional* $X(\vec{y})$ *of the* (NCI) *is primitive recursive in* $\vec{y}$.

The proposition is a corollary of classical results of the proof theory of fragments of arithmetic and weak subsystem of analysis, established by Parsons and Minc and refined by Sieg [1985] and [1988]. These results provide a general method to obtain sharp upper bounds on the complexity of the functionals.

Let $(QF - \mathbf{IA})$ be the fragment of first order arithmetic **PA** in which first order quantification is allowed, but the induction formula in the induction axiom **IA** is quantifier-free; $(QF - \mathbf{IA})$ is a conservative extension of primitive recursive arithmetic **PRA**. Let $\mathcal{F}$ be any set of functions that includes the functions of Grzegorczyk's class $\mathcal{E}_3$ and let $\mathcal{F}_k$ be the k-th class of Grzegorczyk's hierarchy, relativized to $\mathcal{F}$. Let $(QF(\mathcal{F}_k) - \mathbf{IA})$ be the subsystem of $(QF - \mathbf{IA})$ where only function symbols and defining axioms for the functions of the class $\mathcal{F}_k$ are allowed. We obtain a hierarchy of proper subsystems of $(QF - \mathbf{IA})$; indeed, the reflection principle for $(QF(\mathcal{F}_m) - \mathbf{IA})$ can be proved in $(QF_{m+2} - \mathbf{IA})$ (Sieg [1988]).

Restriction of the number of applications of $(\Pi^0_2 - \mathbf{IR})$ yields the same hierarchy. Let $(\Pi^0_2(\mathcal{F}) - \mathbf{IR})_k$ be the fragment that allows at most k applications of the Π^0_2-induction rule and, in addition, only function symbols and defining axioms for the functions of $\mathcal{F}$.

Theorem (Sieg [1988]). $(\Pi^0_2(\mathcal{F}) - \mathbf{IR})_k$ *and* $(\mathrm{QF}(\mathcal{F}_k) - \mathbf{IA})$ *prove the same sentences in the language of* $(\Pi^0_2(\mathcal{F}) - \mathbf{IR})$*; this holds also for extensions of the theories by* Π^0_1 *-sentences. In particular,* $(\Pi^0_2 - \mathbf{IR})$ *is conservative over* **PRA**.

Now the corollary of the Cut-Elimination Theorem (section 2) gives the following information.

Corollary (Sieg [1988]). *If* $(\Pi^0_2(\mathcal{F}) - \mathbf{IR})_k$ *proves a* Π^0_2*-sentence* $\forall x.\exists y.\phi(x,y)$*, then there is a function(al)* $F \in \mathcal{F}_k$*, such that* $(\mathrm{QF}(\mathcal{F}_k) - \mathbf{IA})$ *proves* $\phi(a, F(a))$.

2.2. Functionals in Linear Fragments.

It is common practice in everyday mathematics to break down a long proof into several Lemmas, in order to make the verification easier. In choosing Lemmas, we tend to

consider propositions that have an independent interest or a simple formulation; a Lemma must be easy to grasp and to remember. In producing a functional interpretation, it is desirable that the definition of the functionals correspond closely to the main steps of the proof. Thus we may define functionals corresponding to a Lemma and then apply them to specific arguments provided in the rest of the proof. However, the values obtained by such a computation may be different from the values of functionals obtained from a "more direct" proof. It is therefore significant that if we work in the fragment of classical predicate logic called *Multiplicative Linear Logic* (Girard [1987 b]) the definition and evaluation of "primitive" functionals is independent of Cut-elimination.

Let **LML1** be a sequent calculus where sequents are *multisets*, the only structural rule is Exchange and propositional rules have a *multiplicative interpretation*, e.g.,

$$\frac{\Gamma \vdash \Delta, \phi \qquad \Pi \vdash \Lambda, \psi}{\Gamma, \Pi \vdash \Delta, \Lambda, \phi \wedge \psi} \qquad\qquad \frac{\Gamma, \phi, \psi \vdash \Delta}{\Gamma, \phi \wedge \psi \vdash \Delta}$$

to which we add the usual rules for first order quantifiers. The system **LD** for *Direct Logic* (Ketonen and Weyhrauch [1984], Bellin and Ketonen [1989]) is obtained from **LML1** by adding the Weakening rules.

Herbrand's Theorem (Proposition (ii) in section 2) has a particularly simple form in Direct and Multiplicative Linear Logic: the expansion $\mathcal{E}_{\vec{\mathbf{t}}}(S)$ has minimal complexity. Let $\mathcal{L}$, $\mathbf{y}[x]$, $\mathcal{L}'$ be as above.

Herbrand's Theorem: *Let* $\phi : \exists x_1.\forall y_1.\ldots\exists x_n.\forall y_n.\psi$ *be in prenex form.* [9] *There are terms* $t_1, \ldots, t_n \in \mathcal{L}'$ *such that* ϕ *is provable in First Order Direct Logic if and only if* $\phi_H : \psi(t_1, \mathbf{y}_1[t_1], \ldots, t_n, \mathbf{y}_n[t_1, \ldots, t_n])$ *is provable in Propositional Direct Logic.*

We have formulated Herbrand's Theorem as a corollary of the Cut-elimination Theorem. However we can easily define the Herbrand expansion for the Cut rule if this is restricted to Π^0_2 Cut-formulas. The following fact is practically useful:

Lemma. *Let* $\mathcal{D}$ *be a derivation of* S *in* **LML1** *in which the only application of Cut is the last inference and such that the Cut-formula is* Π^0_2*. Let* $S_H(\vec{\mathbf{s}})$ *be the Herbrand expansion of the conclusion, defined by induction on* $\mathcal{D}$*. If* $\mathcal{D}^*$ *is the result of the Cut-elimination procedure and* $S_H(\vec{\mathbf{t}})$ *is the Herbrand expansion of the conclusion defined by induction on* $\mathcal{D}^*$*, then* $\vec{\mathbf{s}} = \vec{\mathbf{t}}$*.*[10]

$(\Pi^0_2 - \mathbf{IR} - \mathbf{LML1})$ is Multiplicative Linear Logic with the Π^0_2 induction rule.

[9] Here the restriction is essential, as shown by the counterexample $\forall x.(\phi(x) \wedge \psi) \not\vdash (\forall y.\phi(y)) \wedge \psi$, $\phi(\mathbf{y}) \wedge \psi \vdash \phi(\mathbf{y}) \wedge \psi$.

[10] This is not true in general for Direct Logic: if the Cut-formula is introduced by Weakening, say, in the derivation of the left premise, then in $\mathcal{D}^*$ all the side formulas of the right premise are introduced by Weakening, thus some terms in $\vec{\mathbf{t}}$ may be simpler than some terms in $\vec{\mathbf{s}}$.

Proposition. (Π_2^0 – **IR** – **LML1**) *is the largest fragment in which the value of every functional defined from the "primitive" functional interpretation of a derivation $\mathcal{D}$ is always independent of the occurrence of Π_2^0-Cuts in $\mathcal{D}$.*

3. Proofs of Ramsey's Theorems.

In the following informal arguments we stress the structural similarity of the proofs of the Infinite Ramsey Theorem (IRT), the Finite Ramsey Theorem (FRT) and the Ramsey-Paris-Harrington Theorem (RPH). We consider only $c = k = 2$, namely, 2–colorings of graphs. It is enough to work with ordered pairs (x, y), where $x < y$. We write $r(l)$ for the Ramsey function $R(2,2,l)$ and $lr(n_1)$ for the Ramsey-Paris-Harrington function $LR(2,2,n_1)$.

3.1. Proof of the Infinite Ramsey Theorem.

As explained in Section 1, we distinguish two parts. The first part consists of three steps. Let $\chi : [\mathrm{N}]^2 \to 2$ be given.

Part 1, step 1.[11] For $i \in N$ define a chain of subsets $S(i)$ of N as follows. Let $S(-1) = \mathrm{N}$; let *Green* $(n) = \{y : \chi(n,y) = 0\}$, and *Red* $(n) = \{y : \chi(n,y) = 1\}$. Finally, let $T^0(n)$ be *Green* $(n) \cap S(n-1)$ and let $T^1(n)$ be *Red* $(n) \cap S(n-1)$.

$$S(n) = \begin{cases} S(n-1), & \text{if } n \notin S(n-1); \\ T^0(n), & \text{if } n \in S(n-1) \text{ and } |T^0(n)| \geq |T^1(n)|; \\ T^1(n) & \text{otherwise.} \end{cases}$$

By induction on n, one can prove that each $S(n)$ is unbounded.

($*$) Clearly, each $S(j)$ is a subset of $S(i)$, for all $i < j$.

Step 2. Define the "diagonal intersection" H_1 of the sets $S(n)$ by course-of-value recursion:

$$H_1 = \{x : \forall d \in H_1 \cap [x].x \in S(d)\}$$

We prove that H_1 is unbounded. If b is a bound for H_1, consider $S(b)$ and let c be an arbitrary element of $S(b)$. Let d be an arbitrary element of H_1. Since b is a bound for H_1, $d < b$; since $c \in S(b)$, $c \in S(d)$ by ($*$). Since d was arbitrary, c satisfies the definition of H_1 and thus $c < b$. But c was arbitrary, thus $S(b)$ is bounded by b, contradicting step 1.

Step 3. Define an *induced coloring* $f^{stab} : \mathrm{N} \to 2$ by

$$f^{stab}(n) = \begin{cases} 0, & \text{if } n \in S(n-1) \text{ and } T^0(n) \text{ is unbounded;} \\ 1, & \text{otherwise.} \end{cases}$$

($**$) Clearly, $f^{stab}(n) = 0 \vee f^{stab}(n) = 1$, for all n.

[11] Steps 1 and 2 are usually combined.

Finally, let $U^j = \{x \in H_1 : f^{stab}(x) = j\}$ for $j = 0, 1$. Define

$$H = \begin{cases} U^0, & \text{if } |U^0| \geq |U^1|; \\ U^1 & \text{otherwise.} \end{cases}$$

Since H_1 is unbounded and f^{stab} is a 2-coloring of H_1, H must be unbounded, too.

Part 2. It remains to verify that H is χ-homogeneous.

($***$) Clearly, if $x \in S(n)$ then $\chi(n,x) = f^{stab}(n)$ – by the definitions of $S(n)$ and f^{stab}.

We need to show that there exists v such that for all $n, m \in H$ with $n < m$, $\chi(n,m) = v$; we can choose $v = 0$ if U^0 is unbounded, $v = 1$ otherwise.

Suppose U^0 is unbounded and let $n \in H$, $m \in H$ and $n < m$. It follows from the definition of H_1 that $m \in S(n)$. By definition of H, $f^{stab}(n) = 0$, thus $h(n,m) = 0$ by ($***$). Since n and m are arbitrary, we are done. If U^0 is bounded, the argument is similar. ∎

3.2. Proof of the Finite Ramsey Theorem.

Let $n = 2^{2l-1} - 1$. We claim $r(l) \leq n$. Let $\chi : [n]^2 \to 2$ be given. If we take $S(-1) = [n]$ and we restrict all definitions to $[n]$, then the sets $S(i)$, H_1, F^{stab} and H are defined as in the proof of IRT.

Steps 1 and 2 are combined. We need to show that

$$|H_1| \geq 2l - 1$$

and we prove, by induction on p, that for $1 \leq p < 2l - 1$, there is $q_p < n$ such that

$$|H_1 \cap [q_p + 1]| = p \quad \text{and} \quad |S(q_p)| \geq 2^{2l-p-1)}.$$

The computation is easy — let q_1 be $0 \in H_1$, let q_{p+1} be the least element of $S(q_p)$ (remember that $q_p < inf\ S(q_p)$). Finally, since $|S(q_{2l-2})| > 0$, we can pick $q_{2l-1} \leq n$. In Step 3, the *pigeon-hole principle* is invoked to conclude that, since $|H_1| \geq 2l - 1$ and f^{stab} is a 2-coloring of H_1, we must have $|H| = l$. ∎

3.3. Proofs of the Ramsey-Paris-Harrington Theorem.

The following is an adaptation of the argument by Erdös and Mills [1981] to our context. Given n_1 we need to find n_2 such that for any $\chi : [n_1, n_2]^2 \to 2$ there is a large χ-homogeneous $H \subset [n_1, n_2]$. Let $r(n) = 2^{2n-1} - 1$. Define lr by $lr(0) = n_1 + 1$ and $lr(n+1) = lr(n) + r(lr(n))$. We claim that we can take $n_2 = lr(2n_1 - 2)$.

We use the following notation. Let $Green(x) = \{y : \chi(x,y) = 0\}$ and $Red(x) = \{y : \chi(x,y) = 1\}$. Let

$$\begin{aligned}
a_0 &= b_0 = n_1\\
a_{i+1} &= \mu x.x \in \bigcap_{j\leq i} Green(a_j) = \mu x.x > a_i \wedge \{a_0,\dots,a_i,x\} \text{ is } \chi\text{-homogeneous green.}\\
b_{i+1} &= \mu x.x \in \bigcap_{j\leq i} Red(b_j) = \mu x.x > b_i \wedge \{b_0,\dots,b_i,x\} \text{ is } \chi\text{-homogeneous red.}\\
A_{i+1} &= \{x : a_0,\dots,a_i,x\} \text{ is } \chi\text{-homogeneous green, but } \chi(a_{i+1},x) = 1 \text{ [= red].}\\
B_{i+1} &= \{x : b_0,\dots,b_i,x\} \text{ is } \chi\text{-homogeneous red, but } \chi(b_{i+1},x) = 0 \text{ [= green].}
\end{aligned}$$

Finally, let $C_{i,j} = \bigcap_{l\leq i} Green(a_l) \cup \bigcap_{l\leq j} Red(b_l)$. Then

$$\begin{aligned}
[n_1,n_2] &= \{a_0\} \cup C_{0,0}\\
&= \{a_0,\dots,a_{i-1}\} \cup \{b_0,\dots,b_{j-1}\} \cup \bigcup_{l\leq i} A_l \cup \bigcup_{l\leq j} B_l \cup C_{i,j}
\end{aligned}$$

for $i \leq p$, $j \leq q$, $p+q \leq 2n_1 - 2$. The picture is

$$\begin{tikzcd}[column sep=small]
 & & & & & a_0 = n_1 = b_0 \arrow[dll] \arrow[drr] & & & & & \\
 & & & a_1 \arrow[dl] \arrow[dr] & & & & b_1 \arrow[dl] \arrow[dr] & & & \\
 & & a_2 \arrow[dl] \arrow[dr] & & A_1 & & B_1 & & b_2 \arrow[dl] \arrow[dr] & & \\
 & \arrow[dl] & & A_2 & & & & B_2 & & \arrow[dr] & \\
 a_p \arrow[dr] & & & & & & & & & & b_q \arrow[dl] \\
 & A_p & & & & & & & & B_q &
\end{tikzcd}$$

We give two variants of the proof.

Proof 1. By induction on n, for $n \leq 2n_1 - 2$, we show that for some i, j with $i + j = n$
(i) either for some $l \leq i$, $|A_l| = r(a_l)$;
(ii) or for some $l \leq j$, $|B_l| = r(b_l)$;
(iii) or $\{a_0,\dots,a_i\}$ is χ-homogeneous green, $\{b_0,\dots,b_j\}$ is χ-homogeneous red and $|C_{i,j}| > n_2 - lr(n)$.
Clearly, if (i), then either there is a χ-homogeneous red set $D \subset A_l$ of cardinality a_l, and $\{a_l\} \cup D$ is χ-homogeneous and large, or there is a χ-homogeneous green set $D \subset A_l$ of cardinality a_l, and $\{a_0,\dots,a_{l-1}\} \cup D$ is χ-homogeneous and certainly large. Similarly, if (ii). Finally, if (iii) with $i + j = 2n_1 - 2$, then $|C_{i,j}| > 0$ and either $\{a_0,\dots,a_i,c\}$ or $\{b_0,\dots,b_j,c\}$ is χ-homogeneous and large, where c is the least element of $C_{i,j}$.

Proof 2. Let $S(n_1) = [n_1,n_2]$. For $n \geq n_1$, we define sets $S(n)$ as follows

$$S(n) = \begin{cases}
S(n-1), & \text{if } n \notin S(n-1);\\
Green(n) \cap S(n-1), & \text{if } n \in S(n-1) \text{ and}\\
 & \quad \text{either } p = n_1;\\
 & \quad \text{or } p < n_1, \text{ but } \exists j.|B_j| \geq b_j \wedge n = b_j;\\
 & \quad \text{or } p < n_1, \text{ but } \exists j.|B_j| \geq b_j \wedge n > b_j \wedge\\
 & \qquad \wedge |Green(n) \cap S(n-1)| \geq |Red(n) \cap S(n-1)|;\\
 & \quad \text{or } q < n_1, \text{ but } \exists j.|A_j| \geq a_j \wedge n > a_j \wedge\\
 & \qquad \wedge |Green(n) \cap S(n-1)| \geq |Red(n) \cap S(n-1)|;\\
Red(n) \cap S(n-1) & \text{otherwise.}
\end{cases}$$

Now we define the "diagonal intersection" H_1 as above and show:
(i) either $p = n_1$ and $|H_1| \geq n_1$;
(ii) or $q = n_1$ and $|H_1| \geq n_1$;
(iii) or $p, q < n_1$ and $|H_1| \geq l + 2c_l$, where $c_l = a_l$ or b_l, $l < n_1$.
This replaces steps (1) and (2) in the proof of FRT.

Step (3). We define $f^{stab}(n)$ in accordance with the new definition of $S(n)$. We define H as above. Then, arguing as in Proof 1, we show that $|H_1| = |H| \geq n_1$ if (i) or (2); if (iii), we show that $|H| \geq n_1$ if $n_1 \in H$ and that $|H| \geq c_l$ if $c_l \in H$. We verify that

$(***)$ if $x \in S(n)$ then $\chi(n,x) = f^{stab}(n)$

with the new definitions of $S(n)$ and f^{stab}. We prove that H is χ-homogeneous as before. ∎

4. Remarks and Questions.

Remarks. (*i*) The RPH Theorem, as formulated in section (1) can be formulated in $\mathcal{L}_{\mathbf{PA}}$, but is not provable in **PA**. The proof in Ketonen and Solovay [1981] (see also Graham et al. [1980]) is a remarkable application of the Corollary in Section (2), since it shows that $LR(c,k,n)$ dominates every f_α. For any fixed k, however, the result *is* provable in **PA**.

(*ii*) Let $k = 2$. Erdös and Mills [1981] show that $LR(n,2,n)$ is the Ackermann function $f_n(n)$. In particular, a coloring $\chi : [n, f_c(n) - 1]^2 \to c$ is exhibited such that every χ-homogeneous set has cardinality less than n. Namely, $\chi(x,y)$ is defined to be the least $i \leq c$ such that for some j

$$x, y \in [f_i^{(j)}(n), f_i^{(j+1)}(n) - 1]$$

(see also Graham et al. [1980], p.151). For a fixed c, however, $LR(c,2,n)$ is primitive recursive.

(*iii*) The FRT and the RPH Theorem follow by a compactness argument from the IRT – see section (6). Since we consider only countable colorings, (Weak) König's Lemma suffices.

(*iv*) From (*i*) and (*iii*) it follows that the IRT is not provable in a conservative extension of **PA**. For fixed k, the IRT is provable in a conservative extension of **PA**.

Notice that the compactness argument asserts the existence of an n, for which FRT holds, but does not explicitly give it as a function of c, k and l.

Question 1: *Can we actually obtain a bound for the function* $R(c,k,l)$ *by the functional interpretation of the compactness argument for the Finite Ramsey Theorem from the infinite version?*

Question 2: *The above proofs of the Infinite, the Finite and the Ramsey–Paris–Harrington theorems have different conclusions and computational complexities, but similar structure. Can we formulate the common features in the form of a more general lemma?*

5. Analysis of a First Order Proof of the Infinite Ramsey Theorem.

We postpone consideration of the compactness argument and of *Question 1* to the section (6). We focus on a formal proof of the Infinite Ramsey Theorem, and we restrict ourselves to the proof of IRT, exponent 2. Moreover, only Part 1 of the proof in section (3.1) is computationally significant, namely, the part in which we prove that the set H is unbounded. The language is that of first order **PA** with additional predicate and function symbols – namely χ, f^{stab}, S, H_1, H – as described in section 1.

Write a formal derivation in Σ^0_3-**IR** [12] of

Step 1. $\forall n.def\, S(n),\ \forall nq.\chi(n,q)=0 \vee \chi(n,q)=1, \quad \vdash \quad \forall n\, k.\exists i.S(n,i)\wedge i\geq k$

Here $def\, S(n)$ is the formal definition of the sets $S(n)$:

$$\begin{aligned}\forall x.S(0,x) \equiv & [\text{“}x\in T^0(0)\wedge T^0(0) \text{ is unbounded”}]\vee\\ & [\text{“}x\in T^1(0)\wedge T^0(0) \text{ is bounded”}],\\ \forall x.S(n+1,x)\equiv & [\neg S(n,n+1)\wedge S(n,x)]\vee\\ & [S(n,n+1)\wedge \text{“}x\in T^0(n+1)\wedge T^0(n+1) \text{ is unbounded”}]\vee\\ & [S(n,n+1)\wedge \text{“}x\in T^1(n+1)\wedge T^0(n+1) \text{ is bounded”}].\end{aligned}$$

Since in the first part of the proof only one direction of the equivalence is used, we may let $def\, S(n)$ be

$$\begin{aligned}&\forall x.[(\forall e.\exists d.\chi(0,d)=0\wedge d\geq e)\wedge\chi(0,x)=0]\vee\\ &\quad\vee[(\exists b.\forall m.\chi(0,m)=0\supset m<b)\wedge\chi(0,x)=1]\supset S(0,x),\\ &\forall x.[\neg S(n,n+1)\wedge S(n,x)]\vee\\ &\quad\vee[S(n,n+1)\wedge(\forall e.\exists d.S(n,d)\wedge\chi(n+1,d)=0\wedge d\geq e)\wedge S(n,x)\wedge\chi(n+1,x)=0]\\ &\quad\vee[S(n,n+1)\wedge(\exists b.\forall m.S(n,m)\wedge.\chi(n+1,m)=0\supset m<b)\wedge\\ &\qquad\qquad S(n,x)\wedge\chi(n+1,x)=1]\supset S(n+1,x).\end{aligned}$$

[12] Π^0_2-**IR** is not enough here, because the side formula $def\, S(n)$ in the antecedent increases the logical complexity of the *sequent* (see W.Sieg [1985], p.40, footnote 6).

Next, formalize the argument of section (3.1) to provide a derivation [13] of

Step 2. $def\, S,\ def\, H_1,\ \forall k.\exists i.S(k,i)\wedge i\geq k \quad\vdash\quad \forall b.\exists c.H_1(c)\wedge c\geq b$

$def\, H_1$ is a course-of-value inductive definition of the set H_1: $\forall x.H_1(x)\equiv(\forall d.d<x\wedge H_1(d)\supset S(d,x))$. In part 1 we use only the direction

$$\forall x.[\forall d.d<x\wedge H_1(d)\supset S(d,x)]\supset H_1(x).$$

Finally, write a formal derivation [13] of

Step 3.

$def\, H,\ \forall n.f^{stab}(n)=0\vee f^{stab}(n)=1,\ \forall b.\exists c.H_1(c)\wedge c\geq b \quad\vdash\quad \forall j.\exists m.H(m)\wedge m\geq j$

$def\, H$ is the explicit definition of the set H:

$$\forall x.H(x)\equiv[\text{“}x\in U^0\wedge U^0 \text{ is unbounded”}]\vee[\text{“}x\in U^1\wedge U^0 \text{ is bounded”}]$$

In part 1 we use only the direction

$$\begin{aligned}\forall x.\ &[(\forall p.\exists q.H_1(q)\wedge f^{stab}(q)=0\wedge q\geq p)\wedge H_1(x)\wedge f^{stab}(x)]\vee\\ &\vee[(\exists r.\forall u.H_1(u)\wedge f^{stab}(u)=0\supset u<r)\wedge H_1(x)\wedge f^{stab}(x)]\supset H(x)\end{aligned}$$

The only property of f^{stab} relevant here is $f^{stab}:[\mathrm{N}]\rightarrow 2$.

5.1. Estimation of the complexity.

Let "*def' S(n)*" be the Skolemization of the definition of $S(n)$ in Step 1 and "*def' f^{stab}*" be the Skolemization of the definition of f^{stab} in Step 3. [14] Assume we have a formal derivation of Part 1, using *def'* $S(n)$ and *def'* f^{stab}: *now every application of the induction rule is an instance of* Π^0_2-**IR**. The new derivation contains only one [15] application of Π^0_2-**IR** relevant to the complexity of the functionals. Let **v** be the set of Skolem functions in question. By applying the Corollary in section 2.1, we conclude

Proposition. *The functional F such that*

$$\Gamma\ \vdash\ \forall j.H(F(j))\wedge F(j)\geq j,$$

obtained from the interpretation of the derivation of part 1, is in $\mathcal{F}_2$, *where* $\mathcal{F}$ *is* $\{\chi,S,H_1,H,\mathbf{v}\}$, *namely the third class of Grzegorczyk's hierarchy relativized to* $\mathcal{F}$. ∎

[13] In Steps 2 and 3 induction is only used to prove the facts (∗) and (∗∗). These are Π^0_1 lemmas, thus they are irrelevant for the computation of the bounds, according to the Proposition of section (2). However, by the previous footnote, these inductions are formally instances of the Σ^0_3-**IR**.

[14] Notice that a Skolem function for a formula becomes a Herbrand function, when the formula is written in the antecedent of a sequent.

[15] See footnote [13] of the previous section.

Remark. In Section (5.3) we show that given $\chi : [\mathrm{N}]^2 \to 2$, *by iteration of F* we can generate a χ-homogenerous set $H_*^{(l)}$ of cardinality l for certain choices of the parameters and a large χ-homogeneous set $H_*^{[n_1,n_2]}$ for other choices. We also show that $H_*^{(l)}$ and $H_*^{[n_1,n_2]}$ are bounded by the usual functions $R(2,2,l)$ and $LR(2,2,n_1)$, respectively, for any χ.

By using auxiliary colorings [16] we may obtain a proof of the IRT with exponent $k = 2$ and an arbitrary number c of colors, and then construct a functional F_c which will be in $\mathcal{F}'_{c+2}$, for an appropriate set $\mathcal{F}'$ of parameters. Then χ-homogeneous sets of given cardinality and large χ-homogeneous sets can be obtained by iteration of F_c, for arbitrary c and $\chi : [\mathrm{N}]^2 \to c$ and for appropriate choices of parameters. Now $R(c,2,l)$ is primitive recursive but $LR(c,2,c)$ is the Ackermann function. *It is clear that the computational complexity of $H_*^{(l)}$ and $H_*^{[n_1,n_2]}$ ultimately depends on the parameters.*

5.2. Functional Interpretation of the Infinite Ramsey Theorem.

We give now the proof of the Parametrized Ramsey Theorem, stated in section (1). We consider derivations of **Step 1** – **Step 3** (section 5) and produce their interpretations **Step 1**$_F$ – **Step 3**$_F$. By the correctness of the NCI we know that for any choice of numerical functions for the parameters, the functionals will produce values making the *sequents* true. [17] It is reasonable to assume the following

Convention. The interpretations of the predicate letters $S(n,y)$, $H_1(x)$, $H(x)$ and of the function letter $f^{stab}(x)$ are sets $S_F(n)$, H_{1F}, H_F and a function f_F^{stab} such that the sentences *def* S_F, *def* H_{1F} and *def* H_F in the sequents **Step 1**$_F$ – **Step 3**$_F$ are true. [18]

[16] Let $\chi : [\mathrm{N}]^2 \to c$ for some c. Define auxiliary colorings χ_i for $i < c$ by letting $\chi_i(x,y) = 0$ if $\chi(x,y) = i-1$, $\chi_i(x,y) = 1$ otherwise. Then Part 1 of the proof of the IRT, with exponent $k = 2$ and $c+1$ colors, can be obtained by repeating c times Part 1 of the proof of the IRT. By the argument for the Proposition, we can construct a derivation with only c significant applications of the Π_2^0-**IR**.

[17] We may assume $\chi : [\mathrm{N}]^2 \to 2$ to be true; we may restrict ourselves to terms $\chi(x,y)$ where $x < y$. These assumptions are part of the *data* of our problem: there is no point here in considering deviant interpretations of those symbols.

[18] This last condition would be guaranteed if we used the comprehension axiom in second order logic to define the sets $S(n)$, H_1 and H, instead of introducing additional predicates. Consider, for instance, the base case **Step 1**$_0$ in the argument by induction that proves Step 1:

$$\forall p.\chi(0,p) = 0 \vee \chi(0,p) = 1 \quad \vdash \quad \forall k.\exists i.i \in \{x : \theta(x)\} \wedge i \geq k.$$

For some θ_{jF} and some j_0, the interpretation is

$$\chi(0,t) = 0 \vee \chi(0,t) = 1 \quad \vdash \quad \bigvee_{j \leq j_0}(I_j \in \{x : \theta_{jF}(x)\} \wedge I_j \geq k)$$

and, by the correctness of the interpretation, $\bigcup_{j \leq j_0}\{x : \theta_{jF}(x)\}$ is nonempty (assuming $\chi(0,t) = 0 \vee \chi(0,t) = 1$).

Interpretation of Step 1. Finally, we produce the definitions of the functionals. The most interesting part is the proof of **Step 1**. Write

Step 1$_0$ $\quad \forall n.def\ S(n)$, "$\chi : [\mathrm{N}]^2 \to 2$" $\ \vdash\ \forall k.\exists i.S(0,i) \wedge i > k$

for the base case, and

Step 1$_{n+1}^n$ $\quad \forall n.def\ S(n)$, "$\chi : [\mathrm{N}]^2 \to 2$", $\forall c.\exists a.S(n,a) \wedge a > c \ \vdash\ \forall k.\exists i.S(n+1,i) \wedge i > k$

for the induction step. Since the proofs of **Step1**$_0$ and of **Step1**$_{n+1}^n$ do not require applications of the induction rule, it is convenient to look at their Herbrand expansion.[19]

In the following table

$$\begin{array}{ll} \mathbf{e}_1 = \mathbf{e}[0,r] & \mathbf{m}_1 = \mathbf{m}[0,r,\mathbf{k}] \\ \mathbf{e}_2 = \mathbf{e}[0,\mathbf{m}_1] & \mathbf{m}_2 = \mathbf{m}[0,r,\mathbf{e}_2] \\ \mathbf{e}_3 = \mathbf{e}[0,r] & \mathbf{m}_3 = \mathbf{m}[0,t,\mathbf{e}_3], \\ t = max(\mathbf{e}_1,\mathbf{k}) & \end{array}$$

and we may take $\mathbf{k}$ for r. [20]

Herbrand Expansion of **Step 1**$_0$

$$\chi(0,r) = 0 \vee \chi(0,r) = 1, \qquad \chi(0,t) = 0 \vee \chi(0,t) = 1,$$

(1)
$$[((\chi(0,t) = 0 \wedge t \geq \mathbf{e}_1) \vee (\chi(0,\mathbf{m}_3) = 0 \wedge \mathbf{m}_3 \geq \mathbf{e}_3)) \wedge \chi(0,r) = 0] \vee$$
$$[(\chi(0,\mathbf{m}_2) = 0 \supset \mathbf{m}_2 < \mathbf{e}_2) \vee (\chi(0,\mathbf{m}_1) = 0 \supset \mathbf{m}_1 < \mathbf{k}) \wedge \chi(0,r) = 1] \supset S(0,r),$$

(2)
$$[(\chi(0,\mathbf{m}_2) = 0 \wedge \mathbf{m}_2 \geq \mathbf{e}_2) \wedge \chi(0,\mathbf{m}_1) = 0] \supset S(0;\mathbf{m}_1),$$

(3)
$$[(\chi(0,\mathbf{m}_3) = 0 \supset \mathbf{m}_3 < \mathbf{e}_3) \wedge \chi(0,t) = 1] \supset S(0,t)$$

$$\vdash$$

$$S(0,r) \wedge r \geq \mathbf{k}, \qquad S(0,\mathbf{m}_1) \wedge \mathbf{m}_1 \geq \mathbf{k}, \qquad S(0,t) \wedge t \geq \mathbf{k}$$

Now we define functionals B' and D' from (1):

$$B'(0,\mathbf{m},\mathbf{e},\mathbf{k}) = \begin{cases} \mathbf{e}_2, & \text{if } \chi(0,\mathbf{m}_2) = 0 \supset \mathbf{m}_2 < \mathbf{e}_2 \text{ is true;} \\ \mathbf{k} & \text{otherwise;} \end{cases}$$

$$D'(0,\mathbf{m},\mathbf{e},\mathbf{k}) = \begin{cases} t & \text{if } \chi(0,t) = 0 \wedge t \geq \mathbf{e}_1 \text{ (i.e., } \chi(0,t) = 0) \text{ is true;} \\ \mathbf{m}_3, & \text{otherwise.} \end{cases}$$

[19] Here we work in Direct Logic, using the result of Section (2.2). For details, see Bellin [1989].

[20] We may consider only $\mathbf{k} > 0$ and assume $\neg\chi(0,0) = 0 \wedge \neg\chi(0,0) = 1$.

Finally, on the basis of the above Convention, we define the functional $I_0(\mathbf{m},\mathbf{e},\mathbf{k})$ as follows

$$I_0(\mathbf{m},\mathbf{e},\mathbf{k}) = \begin{cases} r, & \text{if } S_F(0,r) \wedge r \geq \mathbf{k} \text{ is true;} \\ \mathbf{m}_1 & \text{if } S_F(0,\mathbf{m}_1) \wedge \mathbf{m}_1 \geq \mathbf{k}; \\ t & \text{otherwise.} \end{cases}$$

We proceed similarly in the induction step, and we define the functional $I_{ind}(\mathbf{a},\mathbf{m},\mathbf{e},\mathbf{k},n)$ as

$$I_{ind}(\mathbf{m},\mathbf{e},\mathbf{k}) = \begin{cases} r, & \text{if } S_F(n+1,r) \wedge r \geq k; \\ \mathbf{m}_1, & \text{if } S_F(n+1,\mathbf{m}_1) \wedge \mathbf{m}_1 \geq \mathbf{k}; \\ \mathbf{a}_1 & \text{otherwise.} \end{cases}$$

Here $\mathbf{a}_1$ is $\mathbf{a}[n,t]$; $\mathbf{a}[n,x]$ is a parameter deriving from the interpretation of the inductive hypothesis $\forall c.\exists a.S(n,a) \wedge a \geq c$ – namely, $S(n,\mathbf{a}_1) \wedge \mathbf{a}_1 \geq t$ –. Finally, we define the functionals B, D, X and I corresponding to the conclusion of the Induction Rule, **Step 1**. For instance, define I to be

$$I(0,\mathbf{m},\mathbf{e},\mathbf{k}) = I_0(\mathbf{m},\mathbf{e},\mathbf{k})$$
$$I(n+1,\mathbf{m},\mathbf{e},\mathbf{k}) = I_{ind}(\lambda k.I(n,\mathbf{m},\mathbf{e},k),\mathbf{m},\mathbf{e},\mathbf{k},n)$$

by primitive recursion. This concludes the interpretation of Step 1. The interpretation of Steps 2 and 3 is omitted.

5.3. Choices for the "Hidden Parameters".

Before giving a formal characterization of the Non-Triviality Condition, we consider some examples of evaluation of the functionals. Here let $H_* = \{F(1), F^2(1), \ldots\} \subset H_F$ be the result of iterating the functional $\lambda n.F(\chi,\mathbf{m},\mathbf{e},\mathbf{q},\mathbf{u},n)$ extracted from the proof of the Infinite Ramsey Theorem, for the given values of the parameters **m**,**e**,**q**,**u**, n; write $H_*^{[x,y]}$ for $H_* \cap [x,y]$ and let $H_*^{(n)}$ be the χ-homogeneous set given by n iterations of F.

Example 0. Let $m[n,x,y] = max(n+1,y)$, $e[n,x] = 0$. Then in the Herbrand expansion of **Step 1_0** we have $e_1 = e_2 = e_3 = 0$, $m_1 = t = k$, $m_2 = m_3 = 1$ and the defining conditions for S_F are tautological, thus S_F is N. Similarly for the expansion of **Step 1^n_{n+1}**. Thus $S_F(n) = \mathrm{N}$ for all n, and so $H_{1F} = \mathrm{N}$, too. Not surprisingly, the iteration of F in this case grows linearly.

Example 1: As usual, let $S(-1) = \mathrm{N}$, *Green* $(n) = \{z : \chi(n,z) = 0\}$ and *Red* $(n) = \{z : \chi(n,z) = 1\}$ (by our assumption, for all z in *Green* (n) or *Red* (n), $z > n$). Let

$$e[n,x] = \begin{cases} 0, & \text{if } \mathit{Green}\ (n) \cap S(n-1) \text{ is unbounded;} \\ max(\mathit{Green}\ (n) \cap S(n-1)) & \text{otherwise.} \end{cases}$$

$$m[n,x,y] = \begin{cases} \mu z.z \in \mathit{Green}\ (n) \cap S(n-1) \wedge z > y & \\ & \text{if } \mathit{Green}\ (n) \cap S(n-1) \text{ is unbounded;} \\ 0 & \text{otherwise.} \end{cases}$$

Functions m and e satisfying the above conditions are in fact Skolem functions for *def* $S(n)$, the defining conditions of $S(n)$, as well as for *def* f^{stab} and so "preserve the meaning" of those definitions. There is nothing to verify here.

Notice that for this m and e we can choose colorings that make $F^{(l)}$ – the l-th iteration of F – grow as fast as desired, i.e., this choice gives no bound for $R(2,2,l)$. Also notice that $H_*^{(n+1)}$ is an extension of $H_*^{(n)}$; thus from the set H_* by compactness we obtain again the Infinite Ramsey Theorem, as expected.

Example 2: Let $S(-1) = [p]$, *Green* (n) and *Red* (n) as before.

$$e[n,x] = \begin{cases} 0, & \text{if } |\mathit{Green}\ (n) \cap S(n-1)| \geq |\mathit{Red}\ (n) \cap S(n-1)|; \\ p & \text{otherwise.} \end{cases}$$

$$m[n,x,y] = \begin{cases} \mu z < p.z \in \mathit{Green}\ (n) \cap S(n-1) \wedge z > y & \\ & \text{if } |\mathit{Green}\ (n) \cap S(n-1)| \geq |\mathit{Red}\ (n) \cap S(n-1)|; \\ \mu z < p.z \in \mathit{Red}(n) \cap S(n-1), & \\ & \text{otherwise.} \end{cases}$$

If $p \geq 2^{2l-1}$, the above choice for **m** and **e** makes $S_F(n) \cap [p]$ satisfy the definition of $S(n)$ in the Finite Ramsey Theorem and *as k varies over* $[p]$, $I(n,m,e,k)$ generates the elements of $S(n)$ — notice that for $y \geq p$, $m[n,x,y] = 0$: considering the Herbrand expansion of **Step 1_0** we see that (1) becomes a tautology for $k \geq p$; thus $\mathrm{N} \setminus [p-1] \subset S_H(0)$ —. We may conclude that with this choice of parameters (and with suitable choices for **q** and **u**) by iterating the functional F l times we generate a χ-homogeneous set $H_*^{(l)}$. Following the proof of the Finite Ramsey Theorem (section 3.2), we can check that $H_*^{(l)} \subset [p]$ for any χ, i.e., the familiar bound is preserved.

Example 3: Let $[n_1,n_2]$, $a_0,\ldots,a_p$, $b_0,\ldots,b_q$, $A_1,\ldots,A_p$, $B_1,\ldots,B_q$ be as in section (3.3) and consider the definition by cases of $S(n)$ given in Proof 2, section (3.3).

$$e[n,x] = \begin{cases} 0 & \text{if } S(n) = \mathit{Green}(n) \cap S(n-1); \\ n_2 & \text{otherwise.} \end{cases}$$

$$m[n,x,y] = \begin{cases} \mu z < n_2.z \in \mathit{Green}\ (n) \cap S(n-1) \wedge z > y & \\ & \text{if } S(n) = \mathit{Green}(n) \cap S(n-1); \\ \mu z < n_2.z \in \mathit{Red}\ (n) \cap S(n-1) & \\ & \text{otherwise.} \end{cases}$$

As before, we check that with this choice of parameters (and with suitable choices for **q** and **u**) $S_F(n) \cap [n_1,n_2]$ satisfies the definition of $S(n)$ in the RPH theorem and as k varies over $[n_1,n_2]$ $I(n,m,e,k)$ generates the elements of $S(n)$ and F generates a χ-homogeneous large set $H_*^{[x,y]}$ for any χ, i.e., the familiar bound is preserved.

From the examples we see that our computation will produce one of the following outcomes, depending on the choice of the parameters and **m**, **e**:

(i) as k varies, the functional $I(n,m,e,k)$ gives values both in *Green* (n) and in *Red* (n) and the set H_* is not χ-homogeneous;

(ii) as k varies, the functional $I(n,m,e,k)$ gives values only in *Green* (n) or in *Red* (n), and thus the set H_* is χ-homogeneous;

(iii) as k varies within a segment $[x,y]$ of N, the functional $I(n,m,e,k)$ gives values only in *Green* (n) or in *Red* (n), thus the set H_* is χ-homogeneous in the segment $[x,y]$.

Finally, information about the cardinality of a certain segment of H_0 requires additional proof. In example 2 we needed to know that $p = r(l)$ in order to argue that $|H_*^{[0,p]}| \geq l$ independenly of χ and in example 3 we used the fact that $n_2 = lr(n_1)$ in arguing that $H_*^{[n_1,n_2]}$ is large, for any χ.

The condition for the "non-triviality" of the computation is easily spelled out. Given a certain choice of $\mathbf{m}$ and $\mathbf{e}$, let

$$\begin{aligned}
\rho(0,k) =& [(\chi(0,\mathbf{m}[0,X,B]) = 0 \wedge \mathbf{m}[0,X,B] \geq B) \wedge \chi(0,X) = 0] \\
\rho(n+1,k) =& [(S(N,\mathbf{m}[n+1,X,B]) \wedge \chi(n+1,\mathbf{m}[n+1,X,B]) = 0 \wedge \\
& \wedge \mathbf{m}[n+1,X,B] \geq B) \wedge S(n,X) \wedge \chi(n+1,X) = 0] \\
\sigma(0,k) =& [(\chi(0,D) = 0 \supset D < \mathbf{e}[0,X]) \wedge \chi(0,X) = 1] \\
\sigma(n+1,k) =& [(S(n,D) \wedge \chi(n+1,D) = 0 \supset D < \mathbf{e}[n+1,X]) \wedge \\
& \wedge S(n,X) \wedge \chi(n+1,X) = 1]
\end{aligned}$$

(here B, D and X depend on n and k, and by definition of X the value $X = k$ is tested). For the "non-triviality" of F (over a certain segment $[x,y]$) we need

(NTC): for all n, if $S(n-1,n)$, then $(\forall k.\rho(n,k)) \vee (\forall k.\sigma(n,k)$ and $\forall k.\rho(n,k) \equiv \neg\sigma(n,k)$.

(NTC)$[x,y]$: for all $n \in [x,y]$, if $S(n-1,n)$, then $(\forall k_{x\leq k\leq y}.\rho(n,k)) \vee (\forall k_{x\leq k\leq y}.\sigma(n,k)$ and $\forall k_{x\leq k\leq y}.\rho(n,k) \equiv \neg\sigma(n,k)$.

The Parametrized Ramsey Theorem (section 1) is now proved. ∎

6. Sketch of an interpretation of the Compactness Argument.

In the last section we have made choices of parameters $\mathbf{m}$, $\mathbf{e}$, $\mathbf{q}$ and $\mathbf{u}$, relying on well-known facts of finite combinatorics and we have shown that for those choices the desired sets are generated by iteration of the functionals within known bounds. It is conceivable that if we work with a compactness argument, there may be implicit in the part of the proof to which the IRT is applied (1) a particular choice of a coloring χ_0 and (2) specific numeric functions for $\mathbf{m}$, $\mathbf{e}$, $\mathbf{q}$ and $\mathbf{u}$, such that they may be mechanically uncovered by some proof-theoretic transformation. We might hope that (1) the coloring χ_0 would give information about a "worst case coloring" or (2) the numeric functions for $\mathbf{m}$, $\mathbf{e}$, $\mathbf{q}$ and $\mathbf{u}$ would be choice functions operating on an initial fragment of χ_0. In these cases the answer to *Question 1* (section 4) would be positive.[21] Part of the problem is to decide what the compactness argument is – different mathematical techniques may produce different proofs and different bounds. An obvious choice is the use of König's Lemma.

[21] Although a detailed analysis is beyond the scope of this paper, the following remarks provide evidence for our belief that the results presented in the previous sections are optimal for the techniques under consideration.

Compactness Argument. Assume the negation of the FRT, namely for given c, k, l, $\neg$FRT$^{(k)}$: *for each* n *there is a coloring* $\xi_n : [n]^k \to c$ *such that every set* $H \subset [n]$ *of cardinality* l *is not* ξ_n*-homogeneous.*

The set of finite colorings ξ_n falsifying the FRT, ordered by the relation "ξ_i *is extended by* ξ_j", forms a tree, which by our assumption is infinite. At each node $\xi_n : [n]^k \to c$ there are at most $c^{\binom{n}{k-1}}$ extensions $\xi' : [n+1]^k \to c$ of ξ_n. By König's Lemma there is an infinite path through our tree, i.e., a coloring $\chi : [\mathrm{N}]^k \to c$ such that every set H of cardinality l is not χ-homogeneous. This contradicts IRT$^{(k)}$, the IRT with exponent k. ∎

The above argument may be formalized by a derivation $\mathcal{D}$ in $\mathbf{PA}^2$, the sequent calculus for second order $\mathbf{PA}$ as follows. Let KL be a formalization of König's Lemma. [22] Let $\mathcal{D}_1^k$ and $\mathcal{D}_2$ be derivations in $\mathbf{PA}^2$ of IRT$^{(k)}$ and KL, respectively; let $\mathcal{D}_0$ be a derivation of

$$(\diamond) \qquad \mathrm{IRT}^{(k)}, \mathrm{KL}, \neg\mathrm{FRT}^{(k)} \vdash .$$

in $\mathbf{LK}^2$, the sequent calculus for pure second order logic. Now $\mathcal{D}$ results from $\mathcal{D}_1^k$, $\mathcal{D}_2$ and $\mathcal{D}_0$ by two applications of Cut. We are interested in the following transformation: apply Cut-Elimination to $\mathcal{D}$ and then the No Counterexample Interpretation to the resulting derivation $\mathcal{D}'$.

Notice that FRT$^{(k)}$ has the form $\exists n.\forall \xi_n.\phi(\xi_n, n)$ and that the NCI of $\mathcal{D}'$, yields a functional N such that $\phi_F(\xi[N], N)$ is true for all choice of numerical functions for the parameters. Here ξ is a new $k+1$-ary function parameter, representing an infinite sequence of finite colorings $\{\xi_n : [n]^k \to c\}_n$, i.e., any attempted counterexample to the FRT. Moreover, ξ may be regarded as an infinite coloring $\xi : [\mathrm{N}]^{k+1} \to c$ by

$$(*) \qquad \xi(i_0, \ldots, i_k) = \xi_{i_k}(i_0, \ldots, i_{k-1})$$

(we assume $i_0 < \ldots < i_k$).

We have not checked this in detail, but we conjecture a negative answer to *Question 1* on the basis of the following experiment. The Finite Ramsey Theorem with exponent k follows from the Infinite Ramsey Theorem with exponent $k+1$. Let $\mathcal{D}_1^{k+1}$ be a derivation of IRT$^{(k+1)}$ and let $\mathcal{D}_*$ be a derivation in $\mathbf{LK}^2$ of

$$(\triangle) \qquad \mathrm{IRT}^{(k+1)} \vdash \forall\zeta.\exists n.\phi(\zeta(n), n).$$

If $\mathcal{D}$ is the derivation ending with an application of Cut to the conclusions of $\mathcal{D}_1^{k+1}$ and $\mathcal{D}_*$, then it is easy to see that after the elimination of this Cut and application of the NCI, the functional N still depends on the coloring ζ — which is arbitrary — and on the *parameters* $\mathbf{m}$ and $\mathbf{e}$ — rather than on specific numeric functions. [23]

[22] We write IRT$^{(k)}$ and FRT$^{(k)}$ for their formalization, too.

[23] The significance of this experiment for an analysis of the compactness argument follows from

7. Conclusion.

What conclusions may we draw from our experiment? Our application of Proof Theory to Combinatorics has produced the Parametrized Ramsey Theorem, a general logical frame for "Ramsey-type" results. Additional computational instructions, in the form of choice-function parameters, yield the Infinite, the Finite and the Paris-Harrington versions of Ramsey's Theorem (see *Question 2*).

From this exercise we have not obtained new bounds for the Finite Ramsey Theorem. [24] It is an open question whether one can provide *new* function parameters for which the interpretation is non-trivial on a certain segment $[x, y]$ and such that $\{F(x), \ldots, F^{(l)}(x)\} \subset [x, y]$ for any coloring χ. [25]

We have considered fragments of arithmetics in which the No Counterexample Interpretation has a simple definition and the functionals are primitive recursive in the function parameters. In particular, using results by Sieg we have obtained a straigtforward estimation of the complexity of the functional interpretation of the Infinite Ramsey Theorem (section 5.1).

We have shown that Ketonen's Direct Logic and Girard's Linear Logic can be used to keep cumbersome computations under control – in these logics Herbrand's Theorem has minimal complexity and for proofs in Π^0_2-**IR** the result of evaluating the functionals is invariant under elimination of Π^0_2-Cuts.

The interest of this approach is increased by the fact that Direct Logic is the logic of the decision procedure of the proof-checker *EKL* (see Ketonen [1983]); thus implementation in *EKL* of the functional interpretation will make extensive experimentation more accessible.

8. Bibliography.

G. Bellin [1989]: *Mechanizing Proof Theory: Computer Aided Formalization and Unwinding of Proofs,* Dissertation, Stanford 1989

the following. (1) *Ceteris paribus*, the NCI of a derivation of $\vdash \exists n.\forall \xi_n.\phi(\xi_n, n)$ has the same form as the the NCI of a derivation of $\vdash \forall \zeta.\exists n.\phi(\zeta(n), n)$. (2) Following Kunen [1977], the argument in the proof of the Infinite Ramsey Theorem for the existence of an infinite pre-homogeneous set for $\xi : [\mathrm{N}]^{k+1} \to c$ is in fact an application of König's Lemma. In our terminology and in the case of $k = 2$, this is the argument that establishes the existence of the infinite set H_1 – corresponding to Steps 1 and 2.

[24] The functional obtained from the Infinite Ramsey Theorem depends on the given coloring $\chi : [\mathrm{N}]^2 \to 2$. By choosing suitable function parameters we made the functional depend only on an initial segment of χ, but the bound was explicitly provided with the parameters rather than extracted from the proof. On the other hand, the application of the functional interpretation to the compactness argument from the Infinite to the Finite Ramsey Theorem introduces a new parameter ξ that may be regarded as an arbitrary coloring $\chi : [\mathrm{N}]^3 \to 2$. We have been unable to mechanically extract choice-functions depending on an initial segment of ξ only.

[25] On the other hand, significant improvement on these bounds are obtained using a completely different proof idea for the Finite Theorem that has no analogue in the Infinite Theorem (see the alternative proof by double induction in Graham et al.[1980], p.3).

G. Bellin and J. A. Ketonen [1989]: *A Decision Procedure Revisited: Notes on Direct Logic, Linear Logic and its Implementation* (manuscript, submitted to Theoretical Computer Science)

P. Erdös and G. Mills [1981]: *Some bounds for the Ramsey-Paris-Harrington Numbers,* Journal of Combinatorial Theory, A 30, pp. 53-70, 1981

J-Y. Girard [1987 a]: *Proof Theory and Logical Complexity,* Vol.I, Bibliopolis, 1987

J-Y. Girard [1987 b]: *Linear Logic,* Theoretical Computer Science 50 pp. 1-102, 1987

R. L. Graham, B. L. Rothschild, J. H. Spencer [1980]: *Ramsey Theory,* Wiley Interscience, 1980

J. A. Ketonen and R. M. Solovay [1981]: *Rapidly growing Ramsey functions,* Annals of Mathematics 113, pp.267-314, 1981

J. A. Ketonen [1983]: EKL—*A Mathematically Oriented Proof Checker,* "Proceedings of the Seventh Conference on Automated Deduction", (Shostak ed.), Lecture Notes in Computer Science 170, pp.65-79, 1984

J. Ketonen and R. Weyhrauch [1984]: *A Decidable Fragment of Predicate Calculus,* Theoretical Computer Science 32, pp. 297-307, 1984

G. Kreisel [1952]: *On the Interpretation of non-finitist Proofs II,* Journal of Symbolic Logic 17, pp. 43-58, 1952

G. Kreisel [1977]: *From Foundations to science: Justifying and unwinding Proofs* in Proceedings from the Symposium "Set Theory. Foundations of Mathematics", Belgrad, *Recueil des Travaux de l'Institut Mathématique, Belgrade, Nouvelle série,* tome 2 (10), pp.63-72, 1977

G. Kreisel [1989]: *Logical Aspect of Computation: Contributions and Distractions,* preprint, Mathematical Institute, Oxford

K. Kunen [1977] *Combinatorics,* in *Handbook of Mathematical Logic,* J.Barwise (ed.), North Holland, 1977

H. Luckhardt [1989]: *Herbrand-analysen zweier Beweise des Satzes von Roth: polynomiale Anzahlschranken,* Journal of Symbolic Logic 54(1), pp.234-263, 1989

H. E. Rose [1984]: *Subrecursion. Functions and Hierarchies,* Oxford Logic Guides 9, 1984 Clarendon Press

W. Sieg [1985]: *Fragments of Arithmetic,* Annals of Pure and Applied Logic 28, pp.33-71, 1985

W. Sieg [1988]: *Derivations as Computations,* manuscript., 1988

W. W. Tait [1965 a]: *Functionals defined by transfinite recursion,* Journal of Symbolic Logic 30, pp. 155 -174, 1965,

W. W. Tait [1965 b]: *The substitution method,* Journal of Symbolic Logic 30, pp. 175 -192, 1965

G.BELLIN,
DEPARTMENT OF PHILOSOPHY,
STANFORD UNIVERSITY,
STANFORD CA 94305