RAMSEY THEORY AND ENORMOUS LOWER BOUNDS

Abstract

by

Harvey M. Friedman
Department of Mathematics
Ohio State University
friedman@math.ohio-state.edu
www.math.ohio-state.edu/~friedman/

April 5, 1997

Not everyone realizes that the classical Ramsey theorem's were stated and proved in order to solve a problem in mathematical logic. This can be seen by reading the title of the classical paper,

F.P. Ramsey, On a problem of formal logic, Proc. London Math. Soc. 30 (1930), 264-286.

The original logic problem was solved by Ramsey in that paper, and it was a decision problem. Although Ramsey's theorem is used in an essential way, this decision problem has a nondeterministic exponential time procedure, which is not an enormous amount of time by the standards of this talk, or by the standards of the numbers that appear in Ramsey's theorem.

The numbers that appear in the finite Ramsey theorem are quite unusual for a fundamental theorem of combinatorics. As we shall review, exponential stacks of arbitrary length appear in the upper and lower bounds.

We published a paper in the mid 80's which gave another related decision problem surrounding Ramsey's paper (the spectra comparison problem), where we show that the time complexity has upper and lower bounds involving exponential stacks of arbitrary length just as in Ramsey's theorem.

H. Friedman, On the Spectra of Universal Relational Sentences, Information and Control, vol. 62, nos. 2/3, August/September 1984, pp. 205-209.

We have another simple decision problem about finite structures with these same huge upper and lower bounds. We also have extended Ramsey's original logic problem to a more general context, and used Ramsey's theorem in a more exotic way to solve it. Again, the decision procedure is just nondeterministic exponential, but the numbers involved are far more gigantic. The Ackermann function is a mere speck of dust compared to the numbers involved here. This involves what is called the $<\epsilon_0$-hierarchy of numerical functions.

We then give a decision problem in this more general context, which is analogous to the one published in the mid 80's, and give upper and lower bounds involving the $<\epsilon_0$-hierarchy of numerical functions.

Thirdly, we give a decision problem in this more general context, which is analogous to the new decision problem about finite structures. We also establish upper and lower bounds involving the $<\epsilon_0$-hierarchy of numerical functions.

There is a related problem that we touch on. By the completeness theorem, every valid sentence in predicate calculus has a proof in predicate calculus. However, the proof may be very long. In fact, there may well be a reasonably short proof that a small sentence in predicate calculus is valid, but all proofs in predicate calculus are gigantic. We give some extreme examples of this related to exotic Ramsey theory.

We now review the classical Ramsey theorems. Here is a good reference for Ramsey Theory:

R.L.Graham, B.L.Rothschild, J.H.Spencer, Ramsey Theory, John Wiley & Sons, 1980.

We start with IRT = infinite Ramsey's theorem.

Let $S_k(N)$ be the family of all k element subsets of the natural numbers N = {1,2,…}. We use [n] = {1,…,n}.

IRT. Let $f:S_k(N) \to [r]$. Then f is constant on some infinite $S_k(E)$.

Proof: By induction on k. The case k = 1 is trivial. Suppose true for k. Let $f:S_{k+1}(N) \to [r]$. First find an infinite $E \subseteq N$ such that the values of f at k+1 element subsets of E depend only on their first k elements. This E is constructed by recursion starting with 1,2,…,k as the first k elements of E. Look at $f:S_{k+1}(E) \to [r]$. There is an obvious $g:S_k(E) \to [r]$ by ignoring top elements. Apply the induction hypothesis to g. (Strictly speaking, we can't apply the induction hypothesis to g because E isn't N. But E can be identified with N).

You may see that this proof is not very constructive or effective. There is no constructive or effective proof in light of the very well known

THEOREM 1. There is a computable $f:S_2(N) \to [2]$ of low computational complexity such that for no infinite recursive E is f constant on $S_2(E)$.

We now give the finite Ramsey theorem.

FRT. Let n >> k,r,p and $f:S_k[n] \to [r]$. Then f is constant on some $S_k(E)$, $|E| = p$.

We first give a slick proof from IRT that gives no information.

Fix k,r,p and suppose this is false. We can build a finitely branching tree with infinitely many nodes and apply what is called Konig's lemma to obtain an infinite path. The empty function is at the root. At the n-th level are the functions $f:S_k[n] \to [r]$ such that f is not constant on any $S_k(E)$, $|E| = p$. The children of a node (function) are just the nodes (functions) at the next level which extend it. By hypothesis, this tree is infinite. Now consider any infinite path. It provides an $f:S_k(N) \to [r]$ which is not constant on any $S_k(E)$, $|E| = p$, in direct contradiction to IRT.

Now for an informative proof. We again proceed by induction. It is convenient to let $R_k(r,p)$ be the least n such that the following holds.

For all r,p, every $f:S_k[R_k(r,p)] \to [r]$ is constant on some $S_k(E), |E| = p$.

For the induction hypothesis, we assume that R_k is everywhere defined. We want to show that the function R_{k+1} is everywhere defined.

Let r,p be given. We would like to choose n so large that for all $f:S_{k+1}[n] \to [r]$, there exists A of cardinality $R_k(r,p)$ such that the values of f at k+1 element subsets of A depend only on their first k elements.

Once we know that, we consider $f:S_{k+1}[A] \to [r]$, and apply the induction hypothesis to obtain the desired $E \subseteq A$, $|E| = p$.

The inductive construction needed to meet this requirement involves division of n roughly $R_k(r,p)$ times, where the divisions are roughly by $R_k(r,p)^k$. So it suffices to have, roughly, $n \geq (R_k(r,p)^k)^{Rk(r,p)}$. Therefore $R_{k+1}(r,p) \leq (Rk(r,p)^k)^{Rk(r,p)}$. And we can take $R_1(r,p) = rp$.

By playing around with these numbers we find that for $k \geq 3$ we see that $R_k(r,p) \leq$ an exponential stack of height k consisting of k-1 2's followed by krp on top. (Pretty crude, but good enough for here).

We now come to lower bounds for $R_k(r,p)$. In [GRS], a proof is given that even with r = 2, there is a lower bound for k-tuples, $k \geq 4$, with an exponential stack of k-2 2's with a constant factor times p^2 on top. They question whether this can be improved with k-1 2's. They credit Erdos and Hajnal.

Lower bounds are obtained by what are called stepping up lemmas. Here is the one used in [GRS]:

LEMMA. Suppose $k \geq 3$, and there exists $f:S_k[n] \to [2]$ which is not constant on any $S_k(E)$, $|E| = p$. Then there exists $g:S_{k+1}[2^n] \to [2]$ which is not constant on any $S_{k+1}(E)$, $|E| = 2p+k-4$.

From this, and essentially the same upper bound argument, [GRS] obtains:

THEOREM 2. For all $k \geq 4$ there is a constant c such that

$$2^{[k-1]}(cp^2) \leq R_k(2,p) \leq 2^{[k]}(cp).$$

Here $2^{[k]}x$ is a stack of k-1 2's with x on top. We write $2^{[k]} = 2^{[k]}(1)$.

To do the kind of complexity analysis we are concerned with, we actually need the following very crude estimate, which can be obtained by related but easier arguments:

THEOREM 3. There is an integer constant $c \geq 1$ such that
i) $2^{[k]}(p) \leq R_{ck}(2,c(k+p))$
ii) $R_k(r,p) \leq 2^{[ck]}(c(k+r+p))$.

The problem of logic solved by Ramsey concerns universal relational sentences. These are purely universal sentences in predicate calculus with identity and no constant or function symbols (thus only relation symbols). It begins with one or more universal quantifiers followed by a quantifier free part, which is a Boolean combination of statements of the form $R(y_1,\ldots,y_k)$ or of the form $z = w$, where $y_1,\ldots,y_k,z,w$ are variables.

Several different R of various arity ≥ 1 may appear. The results hold if constants symbols are allowed, but not if function symbols are allowed.

Ramsey asks: is there an algorithm for deciding whether a universal relational sentence has an infinite model?

Note that every restriction of every model of a urs is a model (to a nonempty subset of its domain). Using the famous compactness theorem for predicate calculus, we see that for a urs φ, there are two possibilities:

a) there are models of every nonzero cardinality;
b) there are models of every nonzero cardinality up to but not including an integer $p \geq 1$.

The spectrum of □ is the class of cardinalities of models of □. In case a), we regard this as , and in case b), we regard this as {1,…,p-1}.

The essence of what Ramsey did is to define what we call a sequence of atomic indiscernibles in a model (atomic SOI). This is a sequence (finite or infinite) of distinct elements $b_1,b_2,\ldots$ from the domain such that for each relation R of M and all indices $i_1,\ldots,i_k,j_1,\ldots,j_k$,

$$R(b_{i1},\ldots,b_{ik}) \;\square\; R(b_{j1},\ldots,b_{jk}),$$

provided $(i_1,\ldots,i_k)$ and $(j_1,\ldots,j_k)$ have same order type. He then proved the following:

THEOREM 4. Let □ be a urs. Then TFAE:

i) □ has an infinite model;

ii) □ has a model with an atomic SOI which exhausts all of the domain, and where the domain has cardinality the number of distinct variables used in □.

Proof: If □ has an infinite model then it has a countably infinite model. So by IRT, it has an infinite model with an atomic SOI which exhausts the model. Now just cut down to a subset of the required cardinality.

On the other hand, let M be a model with ii). Then we can stretch the SOI to length □ in the obvious way by imitating what is going on in M.

Let n be the number of distinct variables used in □. Then any sequence of length n from the stretching of M will look like a sequence of length n from M.

From the above equivalence, we can obviously read off a nondeterministic exponential time algorithm for deciding whether a urs □ has an infinite model. The nondeterministic exponential time completeness is due to H. Lewis, *Complexity of solvable cases of the decision problem for the predicate calculus*, in STOC 1978, pp. 35-46.

The hardness is easier to see if constant symbols are used with a 2n-ary predicate symbol.

One just has to use axioms about the behavior of the predicate symbol at the constant symbols in order to simulate a nondeterministic Turing machine computation of length 2^n. There will be a model if and only if there is an infinite model. The constant symbols can be eliminated by a trick.

We now consider the following decision problem:

DATA: two urs's □,□. QUESTION: do they have the same spectrum? (This is the spectra comparison problem for urs's).

THEOREM 5. Let □ be a urs. Then TFAE:

i) □ has an infinite model;

ii) □ has arbitrary large finite models;

iii) the spectrum of □ is ;

iv) □ has a model with an exhaustive atomic SOI of cardinality the number of distinct variables in □;

v) □ has a model with ≥ R_k(h(□),v(□)) elements, where k is the maximum arity of relations appearing in □, v(□) is the number of distinct variables in □, and h(□) is an at most exponential function.

The only thing new here is v) implies iv); i.e., the use of the Ramsey numbers to get atomic SOI’s.

Also, one can show that an iterated exponential function of □ is required in v).

From this Theorem, we can read off the following:

THEOREM 6. There is an algorithm for deciding the spectra comparison problem. It runs in time complexity $2^{[cx]}$ for some integer constant c ≥ 1.

We have shown the completeness of this decision problem in this complexity class (with respect to poly time reductions).

We will be content to give the crucial construction needed for the lower bound result.

Recall from the crude estimate on Ramsey’s theorem that $R_{ck}(2,ck) > 2^{[k]}$.

KEY CONSTRUCTION. There is a urs □ using binary <,=, binary S, unary P,Q, k+1-ary U, k-ary Y,Z, such that the following holds:

i) the spectrum of □ is 1,2,…,$R_{ck}(2,ck)$;

ii) in all largest models of □, < is a strict linear ordering, P(x) iff x is the least element, Q(x) iff x is the greatest element, S(x,y) iff y is the immediate successor of x in <.

To do this, we say that

a) < is a strict linear ordering of the universe;

b) P,Q hold of at most the least/greatest elements;

c) S(x,y) □ y is the immediate successor of x in <;

d) Y,Z are counterexamples to "the universe restricted to □P (or □Q) has the $R_{ck}(2,ck)$ Ramsey property", guaranteeing that the universe has cardinality □ $R_{ck}(2,ck)$; and if = $R_{ck}(2,ck)$ then P,Q both hold at at least one element;

e) for any x not greatest, U(x,_) is a counterexample to "the universe restricted to {y:□S(x,y)} has the $R_{ck}(2,ck)$ Ramsey property", guaranteeing that if the universe has cardinality $R_{ck}(2,ck)$, then for all x not greatest, there exists y such that S(x,y);

f) if the universe has cardinality $R_{ck}(2,ck)$ then P or Q hold of exactly the least or greatest element, and S is the successor relation.

We can do this so that □ is poly time computable in the unary expansion of k.

With this much control over models of cardinality $R_{ck}(2,ck)$, we can axiomatize the action of a Turing machine, where we cut off computation after $2^{[ck-1]}$ steps. We can compare the spectrum of the above □ with the spectrum of □' which includes an axiom saying that a given TM halts after $2^{[ck-1]}$ steps. The spectra will be equal if and only if the TM halts (before or right) after $2^{[ck-1]}$ steps.

To see this, note that if the spectra of □ and □′ are equal, then □′ must have a model of cardinality $R_{ck}(2,ck)$, and this model must correctly axiomatize the Turing machine action. So the TM halts after $2^{[ck-1]}$ steps.

On the other hand, suppose the TM halts after $2^{[ck-1]}$ steps. Then there is a model of □′ of cardinality $R_{ck}(2,ck)$, and no higher. Hence the spectra of □ and □′ are both $\{1,\ldots,R_{ck}(2,ck)\}$.

We now consider the following new decision problem:

DATA: two urs's □,□. QUESTION: do □,□ have a common largest model?

Here a largest model is one whose cardinality is largest; there may not be any.

Note that the closely related problem,

DATA: a urs □. QUESTION: does □ have a largest model?

is co-nondeterministic exponential time complete, since it is equivalent to "□ has no model or □ has an infinite model."

THEOREM 7. There is an algorithm for deciding whether two urs's have a common largest model which runs in time complexity $2^{[cx]}$ for some integer constant $c \geq 1$.

We have again shown the completeness of this decision problem in this complexity class (with respect to poly time reductions).

The earlier construction is modified as follows. We arrange that in the largest model of □, the full history of TM_1 running for $2^{[ck-1]}$ steps is encoded, and the same for □ with TM_2. Choose TM_1 and TM_2 so that the question of whether their respective histories (running for $2^{[ck-1]}$) steps are the same is appropriately computationally complete. The respective histories are the same if and only if □,□ have a common largest model.

What additional simple problems about urs's have this kind of computational complexity?

DATA: two urs's □,□. QUESTION: does □ have less, more, or the same number of models up to isomorphism as □?

We conjecture that this has the same computational complexity.

For a single urs, the best we have is the following:

DATA: a urs □. QUESTION: Does □ have a rigid largest model?

Here a rigid model is a model with no nontrivial automorphisms.

We can modify the construction to show that this has the same computational complexity.

Now consider

DATA: a urs □. QUESTION: does □ have a unique largest model?

I.e., is there a largest model of □ such that all largest models of □ are isomorphic to it?

We conjecture that this also has the same complexity.

We now move on to some real exotica. We consider universal sentences with constant, relation, and function symbols and equality.

It is well known that questions such as the existence of a model, a finite model, or an infinite model, are undecidable for universal sentences.

Here we work with models of a specific form. A 2-function is a function of the form

$$f:\{1,2,4,8,\ldots,2^n\}^k \;□\; N,$$

which is said to be of size n. We also allow the 2-functions

$$f:\{1,2,4,8,\ldots\}^k \;□\; N$$

of size .

We could work with functions

$$f:\{1,2,4,8,\ldots,2^n\}^k \; □ \; \{1,\ldots,2^n\},$$

getting the same complexity results.

We use universal sentences in one k-ary function symbol, where all terms must be unnested, and only = and < between terms is allowed.

In interpreting a us in a 2-function, one only requires that the sentence be true for all values of the variables that lie in $\{1,2,4,8,\ldots,2^n\}$; or in $\{1,2,4,8,\ldots\}$, respectively.

The 2-spectrum of a us □ is the set of all sizes of 2-functions in which it holds.

THEOREM 8. Let □ be a urs. The spectrum of □ must be one of the following:

i) a finite initial segment of N;
ii) N;
iii) N □ { }.

Furthermore, these attributes are nondeterministic exponential time complete.

Unlike the case of urs’s, ii) does not imply iii).

Problem ii) is shown to be equivalent to □ having a “small” 2-function model with an exhaustive set of “exotic” indiscernibles. The same holds for the problem of determining whether is in the spectrum of □.

THEOREM 9. Let □ be a us. TFAE:

i) the 2-spectrum of □ contains N;
ii) □ has a 2-function model of size that of an exotic Ramsey number involving v(□) and a number of colors exponential in □.

We can use Theorem 9 to establish a decision procedure for the following problem:

DATA: two us's □,□. QUESTION: do they have the same 2-spectrum?

THEOREM 10. The 2-spectra comparison problem for us's is decidable.

Proof: First use Theorem 8 to determine whether □,□ have 2-spectrum N or N □ { }, and identify which. If one or the other do then we are done. If neither do then by Theorem 9, the 2-spectra are bounded by an exotic Ramsey number which can be computed, and the 2-spectra compared.

THEOREM 11. The 2-spectra comparison problem is complete in a complexity class associated with the exotic Ramsey theorem. It is complete in the complexity class associated with the standard $\square_0$-recursive function that is not $<\square_0$-recursive.

DATA: two us's □,□. QUESTION: do □,□ have a common largest 2-function model?

We have shown that this has the same exotic computational complexity. We conjecture that

DATA: two us's □,□. QUESTION: does □ have less, more, or the same number of 2-function models up to isomorphism as □?

has the same exotic computational complexity.

We conjecture that

DATA: a us □.
QUESTION: does □ have a unique largest 2-function model?

has the same complexity.

An exotic atomic SOI for a k-ary 2-function f is a finite or infinite sequence of distinct powers of 2, $b_1,b_2,\ldots$, such that for each appropriate $i_1,\ldots,i_{2k}$ and $j_1,\ldots,j_{2k}$ of the same order type, and $1 \square p \square k$, we have

a) $f(b_{i1},\ldots,b_{ik}) < f(b_{ik+1},\ldots,b_{i2k}) \square f(b_{j1},\ldots,b_{jk}) < f(b_{jk+1},\ldots,b_{j2k})$;

b) $f(b_{i1},\ldots,b_{ik}) < b_{ip} \square f(b_{j1},\ldots,b_{jk}) < b_{jp}$;

c) $(b_{i1} = b_{j1}$ & … & $b_{ip} = b_{jp}$ & $f(b_{i1},\ldots,b_{ik}) \square b_{ip}) \square f(b_{i1},\ldots,b_{ik}) = f(b_{j1},\ldots,b_{jk})$;

d) $(b_{i1} = b_{j1}$ & … & $b_{ip-1} = b_{jp-1}$ & $f(b_{i1},\ldots,b_{ik}) \square b_{ip} \square b_{jp}) \square f(b_{i1},\ldots,b_{ik}) \square f(b_{j1},\ldots,b_{jk})$.

The existence of a finite 2-function with an exhaustive finite set of exotic indiscernibles as above (of length according to k) implies that the 2-spectrum contains N. Exotic indiscernibles actually have to be strengthened slightly to be equivalent: one needs to accommodate constants for the first k or so powers of 2.

For the 2-spectrum to contain , and hence be N □ { }, one adds the clause

e) $f(b_{i1},\ldots,b_{ik}) < b_q$,

where $q > i_1,\ldots,i_k$; again we must accommodate constants for the first k or so powers of 2.

The exotic Ramsey theorem used to derive the existence of such exotic indiscernibles is in the style of a principal lemma introduced by Paris and Harrington, and followed up by Kanamori and McAloon. It is somewhat sharper.

The bounds involved in the finite form of the appropriate exotic Ramsey theorem corresponds to the standard ϵ_0-recursive function that is not $<\epsilon_0$-recursive.