Math 117b - Homework 8

`Instructor:` Andrés Eduardo Caicedo

`Due:` March 13, 2007 at 12:00 pm.

This Homework is due either in the course box outside 253 Sloan or in my office by **Tuesday March** 13 **at noon**. Refer to the grading policy for additional requirements. This is the last homework set for this term. I start with an introduction to the problem, proceed to list the questions you need to solve, and close with a corollary.

1 Introduction

Our goal is to show that a purely combinatorial statement—call it $(*)$—is true but not provable in PA. In fact, let $\mathrm{Th}_{\Pi^0_1}(\mathbb{N}) = \{\phi \in \Pi^0_1 : \mathbb{N} \models \phi\}$. Then we will see that $\mathsf{PA} + \mathrm{Th}_{\Pi^0_1}(\mathbb{N})$ does not prove $(*)$.

Remark 1.1. This theory is definitely *not* r.e. But recall that PA proves that there is a truth predicate for Π^0_1-formulas. Using this, one can express the statement $\mathrm{Con}(\mathsf{PA} + \mathrm{Th}_{\Pi^0_1}(\mathbb{N}))$ as a single (Π^0_2) formula, and it is indeed the case that

$$\mathsf{PA} \vdash ((*) \longleftrightarrow \mathrm{Con}(\mathsf{PA} + \mathrm{Th}_{\Pi^0_1}(\mathbb{N}))),$$

although we won't show this.

Definition 1.2.

1. Given a set X and a number k, let

$$X^{[k]} = \{A \subset X : |A| = k\}.$$

   Given a set $A = \{x_1, \dots, x_k\} \subset \mathbb{N}$, we always assume $x_1 < \cdots < x_k$.

2. If $X \subseteq \mathbb{N}$ we say that a function $f : X^{[k]} \to \mathbb{N}$ is *regressive* iff for any $A \in X^{[k]}$, $f(A) < \min(A)$ whenever $\min(A) > 0$.

3. Given sets X and Y, a function $f : X^{[k]} \to Y$, and a subset $H \subset X$, we say that H is *homogeneous* for f iff for all $A, B \in H^{[k]}$, $f(A) = f(B)$.

4. Given $X \subseteq \mathbb{N}$, a function $f : X^{[k]} \to \mathbb{N}$, and a subset $H \subset X$, we say that H is *min-homogeneous* for f iff for all $A, B \in H^{[k]}$, if $\min(A) = \min(B)$ then $f(A) = f(B)$.

There are two versions of the classical Ramsey theorem, a finite and an infinite one. The finite one can be restated, using some (by now) easy coding, as a statement about numbers and is in fact a theorem of PA. The infinite one is not even expressible in PA (although it is also true).

The infinite version is as follows:

Theorem 1.3. *Let X be infinite, let C be finite, let $1 \leq k \in \mathbb{N}$ and let $f : X^{[k]} \to C$. Then there is an infinite subset H of X homogeneous for f.*

It is customary to refer to C as the set of *colors* and then one says that H is *monochromatic*.

There are several ways of proving this result. A popular one uses the following lemma:

Lemma 1.4 (König)**.** *Let T be an infinite finite branching tree. Then T has at least one infinite branch.* $\square$

That T is a tree means that $T \subseteq \mathbb{N}^{<\mathbb{N}}$ and that if $\tau \in T$ and σ is an initial segment of τ, then $\sigma \in T$. That T is finite branching means that any $\sigma \in T$ has only finitely many immediate successors. König's lemma is very easy to prove but it is not formalizable in PA, since it refers to infinite objects (there are more precise obstructions, in terms of the Turing complexity of the branches of T compared to the Turing complexity of T).

Given $f : \mathbb{N}^{[2]} \to C$ (we may as well assume $X = \mathbb{N}$), one builds a tree as follows: Nodes of the tree are numbers. Given n, the immediate successors of n in the tree are $m_1, \dots, m_s$ such that $f(\{n, m_1\}) \neq f(\{n, m_2\} \neq \cdots \neq f(\{n, m_s\})$; for any other $m > n$, $f(\{n, m\})$ equals one of the $f(\{n, m_i\})$, and the m_i are as small as possible. A branch $b_0 < b_1 < \dots$ through this tree (which exists, due to König's lemma) has the property that given any i and any $j, k > i$, $f(\{b_i, b_j\}) = f(\{b_i, b_k\})$. "Thinning this sequence out" produces then an infinite homogeneous set. And inductive argument easily gives the proof for $f : \mathbb{N}^{[k]} \to C$ with $k > 2$.

The finite version of Ramsey's theorem is as follows:

Theorem 1.5. *For all $n, k \in \mathbb{N}$, $k \geq 1$, and all finite C, there is $M \in \mathbb{N}$ such that whenever $|X| \geq M$ and $f : X^{[k]} \to C$, there is a homogeneous $H \subseteq X$ with $|H| \geq n$.*

This can be derived from the infinite version by another use of König's lemma: Clearly, we can restrict to sets X of the form $\mathbf{m} = \{0, 1, \dots, m\}$. Suppose that for some n, k, C there is no such M. Then for any m we have at least one function $f : \mathbf{m}^{[k]} \to C$ without homogeneous sets of size n (a "counterexample for m"). There are only finitely many such functions. Given $m_1 < m_2$, if $f_2 : \mathbf{m}_2^{[k]} \to C$ is a counterexample for m_2 then $f_2 \restriction \mathbf{m}_1$ is a counterexample for m_1. It follows that the counterexamples for different values of m are naturally organized on an infinite finite branching tree. This tree has a branch. This branch gives you a function $f : \mathbb{N}^{[k]} \to C$ without homogeneous sets of size n. This contradicts the infinite version of Ramsey's theorem.

Of course, once we know that M exists, it follows that the function $\pi : (n, k, l) \mapsto$ least M that works (for $|C| = l$) is recursive.

There are other, more efficient proofs of the finite version, and almost any book on combinatorics would contain such a proof. They give you explicit upper bounds on M. These proofs have an advantage over the proof I just sketched: The upper bounds are good enough that you actually obtain that the function π is in fact *primitive* recursive.

It may be useful to introduce some notation.

Definition 1.6 (Rado)**.** $M \to (n)^k_l$ means that M witnesses the finite version of Ramsey's theorem for n, k, l. In detail: Given any $f : \{1, \dots, M\}^{[k]} \to \{1, \dots, l\}$ there is $H \in \{1, \dots, M\}^{[n]}$ homogeneous for f.

The statement $(*)$ we are interested in, due to Kanamori and McAloon, is a modification of the finite version of Ramsey's theorem. We start with the infinite version.

Theorem 1.7. *If $X \subseteq \mathbb{N}$ is infinite, $1 \leq n \in \mathbb{N}$ and $f : X^{[n]} \to \mathbb{N}$ is* regressive*, then there is an infinite min-homogeneous subset of X.*

Before giving the proof, notice that we cannot in general expect to have an infinite *homogeneous* set. For example, consider the function $f(A) = \min(A)$.

Proof. Let $f : [X]^n \to \mathbb{N}$ be regressive.

We define a decreasing sequence of infinite subsets of X, $X \setminus \{0\} = H_0 \supset H_1 \supset H_2 \supset \dots$ such that letting $m_k = \min H_k$, then $(m_k)_{k \geq 0}$ is strictly increasing. Given H_k, let $\varphi : [H_k \setminus \{m_k\}]^{n-1} \to [0, m_k - 1]$ be the function $\varphi(s) = f(\{m_k\} \cup s)$. Let H_{k+1} be infinite and homogeneous for φ. (Using the infinite version of Ramsey's theorem.)

Then $\{ m_k : k \in \mathbb{N} \}$ is min-homogeneous for f. $\square$

Exactly as in the case of the finite version of Ramsey's theorem, this result gives us a finite version:

Theorem 1.8 $(*)$**.** *For all $n, k \in \mathbb{N}$, $n \geq 1$, there is M such that whenever $f : \{1, 2, \dots, M\}^{[n]} \to \mathbb{N}$ is regressive, there is a min-homogeneous subset of $\{1, \dots, M\}$ of size at least k.* $\square$

There is, however, a big difference between this theorem and the finite version of Ramsey's theorem. Already $(*)$ for $n = 2$ is significant. The function $g_2(k)$ =least M that witnesses $(*)$ for $n = 2$ and k, essentially coincides with the diagonal of Ackermann's function and is therefore not primitive recursive. (This is not too difficult to see, but we won't prove it.) The functions $g_2, g_3, g_4, \dots$ given by $g_n(k)$ =least M that witnesses $(*)$ for n and k, are each provably recursive in PA, but any provably recursive function is eventually dominated by one of them. This is what lies at the heart of the result we want to prove: That PA does not prove $(*)$.

Definition 1.9 (Kanamori, McAloon)**.** $M \to (k)^n_{reg}$ means that M witnesses $(*)$ for n, k. In detail, given any regressive $f : \{1, \dots, M\}^{[k]} \to \mathbb{N}$ there is $H \in \{1, \dots, M\}^{[n]}$ min-homogeneous for f.

We use similar notation in a few other cases, for example, $X \to (k)^n_{reg}$, X a subset of $\mathbb{N}$, means that for any regressive $f : X^{[k]} \to \mathbb{N}$ there is a min-homogeneous $H \in X^{[n]}$.

2 The proof

Throughout this argument you may assume that PA proves the finite version of Ramsey's theorem.

1. Assume $(*)$. Show that for any $e, k, n \in \mathbb{N}$ and any formulas $\psi_0, \dots, \psi_e$ in the language of arithmetic, each with at most $n+1$ free variables, there is a set $H \in \mathbb{N}^{[k]}$ of *diagonal indiscernibles* for these formulas. This means that given $c_0 < c_1 < \cdots < c_n$ and $c_0 < d_1 < \cdots < d_n$, all in H, and any $p < c_0$, then
$$\psi_i(p, c_1, \dots, c_n) \leftrightarrow \psi_i(p, d_1, \dots, d_n)$$
holds for each $i \leq e$.

What follows is an extended hint. You may as well assume that $k \geq 2n+1$. By Ramsey's theorem, there is w such that $w \to (k+n)^{2n+1}_{e+2}$. By $(*)$ there is m such that $m \to (w)^{2n+1}_{reg}$. Given $x_0 < \cdots < x_{2n} \leq m$ define $f(\{x_0, \dots, x_{2n}\})$ and $j(\{x_0, \dots, x_{2n}\})$ as follows: If there is an $i \leq e$ and a $p < x_0$ such that $\psi_i(p, x_1, \dots, x_n)$ and $\psi_i(p, x_{n+1}, \dots, x_{2n})$ have different truth values, then $f(\{x_0, \dots, x_{2n}\})$ is the least such p and $j(\{x_0, \dots, x_{2n}\})$ is the least such i. Otherwise, $f(\{x_0, \dots, x_{2n}\}) = 0$ and $j(\{x_0, \dots, x_{2n}\}) = e+1$.

Show that there is an H_1 of size $k+n$ that is min-homogeneous for f and homogeneous for j. Let $i \leq e+1$ be such that $j(\{x_0, \dots, x_{2n}\}) = i$ for all $\{x_0, \dots, x_{2n}\} \in H_1^{[2n+1]}$.

Show that if $i = e+1$ the set H consisting of the first k elements of H_1 is as we want.

Derive a contradiction from the assumption that $i < e+1$ (here it becomes useful that $k \geq 2n+1$).

2. Argue that PA suffices to prove the result of item 1. when we restrict the formulas ψ_i to be Σ^0_0.

You can be somewhat informal here, but explain how to even *state* the result in the language of PA. For this, it is useful to remember that PA proves that there is a truth-predicate for Σ^0_0 formulas.

[Be careful: It is not that PA proves "for any $e, k, \dots$" but rather that PA proves that $(*)$ implies "for any $e, k, \dots$".]

3. The *least number principle* states that if some number has a property then there is a least number that has the property. Show that the least number principle is equivalent to the induction principle.

Definition 2.1. Let $M \models \mathsf{PA}$. Let $a, b \in M$. By $[a, b]$ we denote the set $\{x \in M : a \leq x \leq b\}$ where $\leq$ is the usual order in M. If I is a proper initial segment of M, we write $I < M$. If $a \in I$ and $b \in M \setminus I$ we write $a < I < b$. We know that there is an initial segment of M isomorphic to $\mathbb{N}$. To simplify notation, we will identify this initial segment with $\mathbb{N}$ in what follows.

4. Let $M \models \mathsf{PA}$ be non-standard. Let $a, b, c \in M$, $c \in M \setminus \mathbb{N}$, and assume that $M \models [a, b] \to (2c)^c_{reg}$, where $[a, b] = \{x : a \leq x \leq b\}$. Show that there is $I < M$ such that $a < I < b$, I is closed under addition and multiplication, and $I \models \mathsf{PA}$.

Here is an elaborate hint: Let $\sigma(k)$ be the formula stating that there are k diagonal indiscernibles in the interval between a and b for the first k Σ^0_0 formulas (in some effective enumeration). Show that $M \models \sigma(n)$ for each $n \in \mathbb{N}$.

There are now two cases. Either $M \models \sigma(n)$ for each $n \in M$, in particular for some non-standard n, or else for some non-standard n, $M \models \neg\sigma(n)$. By the least number principle conclude that there is a least n such that $M \models \neg\sigma(n)$. Argue that this n is non-standard, so there is $m < n$, m non-standard, such that $M \models \sigma(m)$.

In either case, conclude that there is a set $\{c_i : i \in \mathbb{N}\}$ of elements of M, $c_0 < c_1 < \dots$ such that $a < c_i < b$ for each i and the c_i are diagonal indiscernible for all (standard) Σ^0_0 formulas.

Let $I = \{x \in M : \exists i (x < c_i)\}$, so I is a proper initial segment of M.

Suppose $i_0 < i_1 < i_2$. Assume that $p < c_{i_0}$ and $p + c_{i_1} = c_{i_2}$. Use indiscernibility to conclude that $p + c_{i_1} = c_j$ for all $j > i_2$. Conclude that $c_{i_0} + c_{i_1} \leq c_{i_2}$. Conclude that I is closed under addition.

Again, suppose that $i_0 < i_1 < i_2$. Assume that $p < c_{i_0}$ and $p \cdot c_{i_1} < c_{i_2} \leq (p+1) \cdot c_{i_1}$. Reach a contradiction (for example, by showing that indiscernibility implies from this that $(p+1) \cdot c_{i_1} < c_j$ for any $j > c_{i_2}$). Conclude that $c_{i_0} \cdot c_{i_1} \leq c_{i_2}$ and therefore I is closed under multiplication.

Explain why Σ^0_0 formulas are absolute between M and I (remember the argument for problem 2 of homework 6).

Given an arbitrary Σ^0_0 formula ψ and a number $n \in \mathbb{N}$, let $Q = \exists$ if n is odd and $Q = \forall$ if n is even. Let $p < c_{i_0}$ and show that

$$I \models \exists x_1 \forall x_2 \dots Q x_n \psi(p, x_1, \dots, x_n)$$

iff

$$\exists i_1 > i_0 \forall i_2 > i_1 \dots Q i_n > i_{n-1} M \models \exists x_1 < x_{i_1} \forall x_2 < c_{i_2} \dots Q x_n < c_{i_n} \psi(p, x_1, \dots, x_n)$$

iff

$$M \models \exists x_1 < c_{i_0+1} \forall x_2 < c_{i_0+2} \dots Q x_n < c_{i_0+n} \psi(p, x_1, \dots, x_n)$$

iff the same formula holds in I.

Now show that I satisfies all the instances of induction, by showing that it satisfies all the instances of the least number principle. For this, suppose $\phi(x_0, \dots, x_e, x_{e+1})$ is any formula, $p_0, \dots, p_e \in I$ and $I \models \exists x \phi(p_0, \dots, p_e, x)$. we must argue that there is a least such x. Use the reduction to Σ^0_0 formulas indicated in the paragraph above (and that M satisfies the least number principle for Σ^0_0 formulas) to conclude.

5. Conclude that $(*)$ is not provable in $\mathsf{PA} + \mathrm{Th}_{\Pi^0_1}(\mathbb{N})$.

For this, it is enough to show that there is a model of $\mathsf{PA} + \mathrm{Th}_{\Pi^0_1}(\mathbb{N})$ that is not a model of $(*)$. You may assume that there is a model $M \models \mathsf{PA} + \mathrm{Th}_{\Pi^0_1}(\mathbb{N})$, M non-standard. Let $a \in M \setminus \mathbb{N}$ and let b be least such that $M \models [a, b] \to (2a)^a_{reg}$ (if there is no such b, we are done). Let I be obtained from M as in the previous item. Show that $I \models \mathsf{PA} + \mathrm{Th}_{\Pi^0_1}(\mathbb{N})$ (recall problem 2 of homework 6).

3 Coda

Let me finish by explaining briefly why the function $\nu(n)$ =least m such that $[n, m] \to (2n)^n_{reg}$ eventually dominates all functions provably recursive in PA.

We argue by contradiction. Suppose that g is provably recursive in PA and infinitely often $g(n) \geq \nu(n)$.

Recall that the completeness theorem states that a theory is consistent iff it has a model. It is fairly easy to derive from this the *compactness theorem*: Let T be a theory. Suppose each finite subset of T has a model. Then T has a model.

This result allows us to construct non-standard models. In particular, we can use it to build a non-standard model $M \models \mathsf{PA}$ where there is an infinite element a such that $M \models \nu(a) \leq g(a)$.

[For example, we can work in the language of arithmetic augmented with a constant c, and let T be the union of PA with the statements that $c > n$ (for each $n \in \mathbb{N}$), that $\nu(c)$ exists, and that $\nu(c) \leq g(c)$. Any finite subset of T has a model, namely $\mathbb{N}$, with c interpreted as some sufficiently large integer m such that $\nu(m) \leq g(m)$. (This is possible since we are assuming that there are infinitely many such numbers m.) Then it follows from compactness that T has a model M. But this model is nonstandard since c^M is infinite.

A very similar argument can be used to show that there is a non-standard model of $\mathsf{PA} + \mathrm{Th}_{\Pi^0_1}(\mathbb{N})$ as required in item 5. This is a standard and very useful model-theoretic trick.]

Given such M with nonstandard a such that $M \models \nu(a) \leq g(a)$, consider I as in item 4., with $\nu(a)$ in the role of b. Then $a < I < \nu(a) \leq g(a)$, so $I \models \mathsf{PA}$ but g is not total in I, since $I \models$ "$g(a)$ does not exist". But this contradicts that PA proves that g is total.

References

[1] A. Kanamori, K. McAllon. *On Gödel incompleteness and finite combinatorics*, Annals of Pure and Applied Logic **33 (1)** (1987), 23–41.