Ramsey's theorem on Existential-Universal sentences of Predicate Calculus

Thomas Forster

June 21, 2010

The theorem we know as Ramsey's theorem—$\omega \to (\omega)^n_m$—was just a dry run for $(\forall n, m, k)(\exists x)(x \to (n)^m_k)$ which in turn was a dry run for $(\forall n, m, k)(\exists x)(x \to (n)^{<m}_k)$ which was merely a lemma *en route* to what he was *really* after. Ramsey's real theorem is that there is a method for deciding whether or not a $\forall^*$ sentence has arbitrarily large finite models, and the stronger result alluded to in the title is obtained from it by some elementary jiggery-pokery.[1]

1 Definitions

We fix a first-order language $\mathcal{L}$ with equality but without function letters. These two points are more important here than they usually are, because the result we are trying to prove—namely the decidability of the universal fragment—is trivial for a language without equality or function letters but impossible once one has both.

The variables of our first-order language are lower case Roman letters with subscripts. Calligraphic font letters denote models and the corresponding upper-case Roman letters the domain ("carrier set") of that model. A formula is **true** in a structure $\mathcal{M}$ iff all assignment functions satisfy it. An assignment function for a model $\mathfrak{M} = \langle M, \ldots \rangle$ ought to be a function sending variables of the language to elements of M but for technical reasons it is easier to take $\mathfrak{M}$-assignment functions to be functions taking *subscripts* of variables as their arguments, and values in M (the carrier set of $\mathfrak{M}$). Remember a formula is **valid** if it is true in all structures. Two formulæ are **logically equivalent** if they are true in the same structures.

[1]Historical point: this was before Church proved the undecidability of the predicate calculus. The question of the decidability of LPC in the air at that time.

The **Prenex Normal Form** theorem says that every formula of $\mathcal{L}$ is logically equivalent to one where all the quantifiers have been pulled to the front and all atomic subformulæ are within the scope of all quantifiers. By the prenex normal form theorem we may assume that our candidate formula Φ has had all its quantifiers pulled to the front and its matrix put into disjunctive normal form. (The **matrix** of a formula in prenex normal form is the stuff after the quantifiers.) We can be sure that for any finite length of quantifier prefix there can be only finitely many disjuncts that any $\forall^*$ sentence can contain. (This is because there are no function letters).

A universal sentence ($\forall^*$-sentence, Π_1-sentence etc. etc.) is one which, once coerced into prenex normal Form, has no existential quantifiers.

A universal sentence in a language without function letters or equality is satisfiable—if at all—in a universe with only one element.

Say something about function letters making life difficult.

2 Universal Sentences

Let us fix a universal sentence Φ, and suppose that the variables in Φ are 'x' with subscripts i for $i < p$. ('p' alludes to 'Φ'.) We manœuvre Φ into prenex normal form, with its matrix in disjunctive normal form. Let us follow Ramsey in calling these disjuncts **alternatives** rather than the more generic *disjuncts* (as they are customarily called nowadays) to remind ourselves that it is elements of precisely this family of disjuncts that we have in mind. A p-alternative—henceforth just an **alternative**—is a conjunction of atomics and negatomics in which every atomic formula in the variables x_i with $i \leq p$ appears precisely once. The $\forall^*$ sentence Φ that we are considering has only finitely many variables, and only finitely many predicate letters can appear in it. So there are only finitely many conjunctions of atomics and negatomics in the language of Φ. The set of disjuncts in the DNF of the matrix of Φ is merely a subset of this set. We assume wlog that all these disjuncts are *maximal* in the sense that each one contains every atomic formula as a conjunct, either negated or negated. This ensures that the alternatives of Φ are mutually exclusive and collectively exhaustive of all possibilities represented by Φ.

It might help the reader orient his/herself to think about how many alternatives there are. There is only one nontrivial language that is small enough for us to do anything by hand, and that is the language with equality and one binary relation. n variables give $\binom{n}{2}$ equations and n^2 things like $R(x_i, x_j)$ for a total of $2^{\binom{n}{2}+n^2}$ alternatives.

2.1 Bad Alternatives

So an alternative is a state-of-affairs that is consistent from the point of view of propositional logic. However, some alternatives—such as "$x_1 = x_2 \wedge x_2 = x_3 \wedge x_1 \neq x_3$"—violate the elementary theory of equality. Discarding them cuts the number of alternatives down to something more like $\Pi(n) \cdot 2^{n^2}$ where $\Pi(n)$ is the number of ways of partitioning an n-membered set.

Some violate principles which, despite being subtler, are nevertheless still logical—such as substitutivity of identity. '$F(x) \wedge \neg F(y) \wedge x = y$' is a case in point.

Do we want to consider these bad formulæ to be alternatives? We certainly want to discard them! So we either tighten up our definition of alternative so that such conjunctions of atomics and negatomics are not alternatives, or we can allow them to be alternatives, but then weed out from the PNF version of Φ any such bad alternatives that crept in during the normalisation process.

So let us consider that done, one way or another, so that the matrix of the PNF of Φ contains only sensible alternatives.

2.2 Orphans

There is a further simplification we can make, arising from the following consideration. The symmetric group S_p of all permutation of the natural numbers $\leq p$ (in other words the group of permutation of the variables in Φ) has an obvious action on the matrix of Φ (the set of alternatives). If A is an alternative and $\sigma \in S_p$ then $\sigma(A)$ is the result of rewriting the subscripts of the variables in A in accordance with σ.[2]

An **orphan** is an alternative A of Φ such that, for some $\sigma \in S_p$, $\sigma(A)$ is not in Φ.

LEMMA 1 Φ *is logically equivalent to (i.e., has the same models as) the result of deleting all orphans from it.*

Proof:

[2]There is an annoying subtlety here: Is there an order (lexicographic perhaps?) in which we write the atomics and negatomics in A? Or do we think of A as a conjunction of a set of atomics and negatomics instead of as a list? If the first, then the conjuncts in $\sigma(A)$ might be in the wrong order. So it might be best to think of alternatives in this second way. We certainly *don't* want to make the concept of alternative so fine that we regard $R(x, y) \wedge \neg R(y, x)$ as a different alternative from $\neg R(y, x) \wedge R(x, y)$. But in doing this we are already making a point about linearity of notation.

Let $\mathfrak{M} \models \Phi$ and suppose there is an assignment f which satisfies [the disjunction of alternatives that is the matrix of] Φ, and that f does this by satisfying an orphaned alternative A. Since A is an orphan, there is a permutation $\sigma \in S_p$ such that the alternative $\sigma(A)$ is not in [the disjunction of alternatives that is the matrix of] Φ. Then the assignment $\sigma \circ f$ will satisfy $\sigma(A)$. But then—since alternatives are mutually exclusive—$\sigma \circ f$ cannot satisfy any alternative in [the disjunction of alternatives that is the matrix of] Φ and therefore doesn't satisfy Φ, so Φ is not true in $\mathcal{M}$. ■

So we could rephrase lemma 1 as follows:

Lemma 2 *Every (universal closure of) a set of alternatives is logically equivalent to (the universal closure of) the union of the orbits included in it.*

Notice that this works for universal sentences, but not for formulæ of higher logical complexity in which different variables can be bound by quantifiers of different flavour ($\exists$ *versus* $\forall$.)

It could be said that the effect of the manipulations and simplifications inflicted on Φ have so far had the effect merely of correcting the mistake made by our syntax in requiring us—if we wish to think of a quantifier prefix $\forall x_1 \ldots \forall x_n$ as one quantifier not many—as a quantifer over *lists* not finite sets. We get into this situation because the tokens of symbols in a formula are linearly ordered by position. It is noteworthy that skolemisation involves deleting quantifier prefixes (on the assumption that naked variables are universally quantified) and this act destroys the order information. It shows that we never needed this information in the first place.

2.3 Conjunctions of Universal Formulæ

However there is another standard way of expressing universal sentences. Each alternative A in Φ is a conjunction of atomics and negatomics, and we can think of each alternative as a conjunction of two formulæ, both themselves conjunctions. The first conjunct is the conjunction of all the equations and inequations in Φ, and the second is the conjunction of all the remaining atomics and negatomics. What conjunctions of equations and inequations can possibly appear in alternatives in this way? Well, if Φ is to have any models at all, then any logical possibility—like $x_1 = x_2 \wedge x_2 \neq x_3 \ldots$—must be a subformula of at least one alternative, otherwise no assignment function f with $f(1) \neq f(2)$ will ever satisfy Φ and Φ will not be true in any model.

There are $\Pi(p)$ such logical possibilities, where $\Pi(p)$ is the number of partitions into nonempty pieces of a set of size p. This invites us to rephrase Φ along the following lines:

"If you pick up $x_1, x_2 \ldots x_p$ then either $x_1 = x_2 \wedge x_1 \neq \ldots$, in which case you are in one of these alternatives, or $x_1 = x_2 \wedge x_1 \neq \ldots$, in which case you are in one of these alternatives, or ..." with $\Pi(p)$ cases to consider.

That is to say, we can rephrase Φ as a conjunction:

$$\bigwedge_{j \leq \Pi(p)} (\forall x_1 \ldots x_p)(C_j \to \Psi) \tag{1}$$

where the C_j range over the $\Pi(p)$ consistent conjunctions of equations and inequations.

Notice that all the conjuncts in formula (1) are closed, so we can manipulate their variables independently. But that means we can now make a further simplification, as follows. The formula

$$(\forall x_1 x_2 x_3)(x_1 = x_2 \wedge x_2 \neq x_3 \to F(x_1, x_2, x_3))$$

is clearly logically equivalent to the formula

$$(\forall y_1 y_2)(y_1 \neq y_2 \to F(y_1/x_1, y_1/x_2, y_2/x_3))$$

($F(y_1/x_1, y_1/x_2, y_2/x_3)$ is of course the result of substituting 'y_1' for 'x_1', 'y_1' for 'x_2' and 'y_2' for 'x_3' in F.)

So we perform this simplification on all the conjuncts in formula (1) simultaneously, thereby enabling us to express Φ more specifically as a conjunction of expressions:

$$(\forall y_1 \ldots y_i)(\bigwedge_{j \neq k} y_j \neq y_k \to F_i) \tag{2}$$

one for each $i \leq p$, with the following rather special features.

1. The F_i are all disjunctions of conjunctions of atomics and negatomics from the *non-logical* vocabulary: all the equations and inequations were hived off earlier. These conjunctions of atomics and negatomics from the non-logical vocabulary are our new alternatives.

2. Lemma 1 tells us that we can remove orphans, so for each i we can take F_i to be the disjunction of all the formulæ in a union of S_i-orbits of (new-style) alternatives.

Ramsey writes out the set of formulæ whose generic representative is (2) as a matrix: one row for each i. I shall retain this nomenclature of 'rows' for the conjuncts like formula 2.

Definition 3 *For a formula like that in 2 a* **form** *is an orbit of alternatives.*

There is a certain amount of overloading/equivocation going on in that sometimes a form is a *set* of formulæ and sometimes a *disjunction* of formulæ.

Suppose we have a form F that uses variables $y_1 \ldots y_k$. For $I \subseteq \{1 \ldots k\}$ we can delete from the alternatives in F all occurrences of y_i for $i \in I$ and all the atomic subformulæ in which they appear. The result is something that is very like a form, except that the set of subscripts of the variables that appear in it might not be in an initial segment of $\mathbb{N}$. But this can be corrected by relettering: simply collapse the variables: $\{y_1, y_4, y_7\} \mapsto \{y_1, y_2, y_3\}$.

The form we obtain by doing this is a form that Ramsey speaks of as **involved** in the form we first thought of. I can see no reason for this involved[3] terminology and I will say "F_1 is a restriction of F_2" in these circumstances and write '$F_1 \prec F_2$'. This relation is clearly reflexive and transitive.

It matters because if we take a row of Φ, the kth for example, and delete some variables from it we obtain an allegation about a smaller number of variables, l, perhaps. This allegation had better not contradict the information about l variables contained in the lth row of Φ! Another way of putting it. A form with n variables in it in some sense contains a complete description of what can happen to and n-tuple. This information contains within it information about what can happen to n'-=tuples for $n' < n$. Thus a form with n variables is in danger of *overwriting* a form with fewer variables. This had better not happen!

We are now in a position to announce a partial result.

Lemma 4 *If $n \leq p$ then[4] Φ has a model of size n iff Φ contains a form F with n variables in it and contains every form $F' \prec F$.*

Proof:

$L \rightarrow R$ Suppose $\mathfrak{M} = \langle M, \ldots \rangle$ is an $\mathcal{L}$-structure with $|M| < p$ and is a model for some universal sentence Φ. We will try to recover what information we can about Φ given only this news.

[3] joke! Joke!!

[4] remember that p is the number of variables in Φ

The first thing to notice is that, for $k > n$, we can infer nothing about what might be in the kth row of Φ. This is because the kth row tells us what is true of a tuple of k distinct things—and M does not contain k distinct things. But this is all right, because we are looking for a form with n variables in it not $k > n$ variables. So Φ can have for its k row anything we please—or equally not have a kth row at all for any $k > n$.

Let f be an $\mathfrak{M}$-assignment function, a function sending the variables 'y_1' … 'y_k' (or, strictly speaking, their subscripts) to members of M.

From each f we will derive an alternative, A_f, as follows. $\mathfrak{M}$ comes equipped with interpretations (graphs) of each of the predicate letters appearing in Φ. If $\langle f(i), f(j)\rangle$ is in the $\mathfrak{M}$-interpretation of the two-place predicate letter 'R' then we put the formula '$R(y_i, y_j)$' into A_f; if $\langle f(i), f(j)\rangle$ is not in the $\mathfrak{M}$-interpretation of the predicate letter 'R' then we put the formula '$\neg R(y_i, y_j)$' into A_f. Clearly we have $A_{\sigma\circ f} = \sigma(A_f)$, so whenever we put an alternative A into Φ we also put in every alternative in the same orbit (form) as A.

What happens if $f(i) = f(j)$? By analogy with the foregoing we would want to put '$y_i = y_j$' into A_f. But alternatives do not contain equations, so this can't happen. We sublimate the urge to write down '$y_i = y_j$' by instead replacing all occurrences of 'y_j' by 'y_i' (assuming $i < j$). That way we get for A_f an alternative from a higher row, one with fewer variables in it.

(There is a slight difficulty here. What are we to do if, say, $p = 4$ and our assignment function f sends both 1 and 2 to $m \in M$ and sends both 3 and 4 to $m' \in M$? Clearly the message f has for us will concern row 2 rather than row 4. What is this message? To read it, we must "squash" 'y_1' and 'y_2' to 'y_1' and squash 'y_3' and 'y_4' to 'y_2', so that if $\langle m, m'\rangle$ belongs to the $\mathfrak{M}$-interpretation of R then we put '$R(y_1, y_2)$' into A_f. Q: Why not squash 'y_1' and 'y_2' to 'y_2' and 'y_2' and 'y_4' to 'y_1' instead? A: it doesn't make any difference! The result of squashing the variables in the second way will be put in by the assignment function f' that sends 1 and 2 to $m' \in M$ and sends 3 and 4 to $m \in M$? The kth row is a union of forms (orbits) and all this comes out in the wash.)

That is to say, the (ordinal number of) the row of formula (2) that A_f belongs to will depend on how many different values f takes. Let $|rn(f)|$ be the number of distinct values taken by f. Evidently A_f will belong to row $|rn(f)|$.

$R \to L$ Suppose Φ contains such a form—F, say—with n variables. We build an n-sized model $\mathfrak{M} = \langle M, \ldots \rangle$ of Φ. The kth rows, with $k > n$, all say "Whenever I pick up k distinct things, the following happens…" and so are vacuously true in any model with fewer than k things, and we don't have to worry about them.

Enumerate M somehow as $\{m_1 \ldots m_n\}$ (it doesn't matter how), and pick one alternative A from F (it doesn't matter which). Then consider the assignment $\lambda i.m_i$. We can now read off, from A, the interpretations of the predicate letters of $\mathfrak{M}$. If A says: $R(y_1, y_3)$, then we put $\langle m_1, m_3 \rangle$ into the $\mathfrak{M}$-interpretation of R. This gives a complete description of the interpretations of all the predicate letters for our model $\mathfrak{M}$.

So why do we need to worry about the restrictions of F to earlier rows? The point is that, for Φ to be true in $\mathfrak{M}$, it has to be satisfied by *all* assignments, including those that are not 1-1. Accordingly we have to ensure that the information given by earlier rows does not conflict, and that is why we need all the restrictions of F to be present in earlier rows.

Notice that the model $\mathfrak{M}$ we have constructed satisfies not only Φ but also the formula obtained from Φ by deleting all forms other than the one we chose (and of course its restrictions).

■

But what happens if there are more than p elements of our candidate model? Then we have some more work to do. First some more titivating. We notice that, the way Φ is expressed at the moment, a lot of the predicate letters can appear with the same variable in many positions. This makes a mess. We clear this up by inventing new predicate letters in a systematic way, so that—for example—instead of writing '$\psi(y_1, y_1, y_2)$' we write '$\psi_{1=2}(y_1, y_2)$' where '$\psi_{1=2}$' is a new predicate letter. Consider this done.

Say something about deleting from Φ the things not appearing in the chosen form. What is the logical status of this operation? They probably have the same spectrum…

Fix an alternative A (and remember a form is a disjunction of alternatives). Suppose A belongs to the ith row of the presentation of Φ in the style of formula 2 so that A has i free variables. Now let ψ be some (possibly newly coined) predicate letter with j argument places ($j \leq i$) which appears in A. A must, for each choice $\vec{z}$ of j variables from $y_1 \ldots y_i$, contain either $\psi(\vec{z})$ or $\neg\psi(\vec{z})$. Now $x_1 \ldots x_j$ give us $j!$ tuples of length j and A must make up its mind whether or not ψ holds for each of the $j!$ ordered j-tuples. *We are interested in the odd cases where this enables us to ascertain what A* *thinks happens to the remaining triples.* Let's fix some values for ease of illustration. Suppose $j = 3$ and $i = 24$, so that ψ has three argument places and there are 24 variables to be dealt with. There are $\binom{24}{3} \cdot 3!$ ways of fitting these variables into ψ and for each of these A must contain either it or its negation.

Given any of the $\binom{24}{3} \cdot 3!$ triples there is a canonical bijection between that triple and the triple $\{1, 2, 3\}$, since they are isomorphic as substructures of the linear order $\langle \mathbb{N}, \leq \rangle$. *Via* these canonical bijections, any ordered triple will be said to **correspond** to an ordered triple whose entries are $1, 2$ or 3. Thus $\langle 5, 7, 11 \rangle$ and $\langle 2, 3, 17 \rangle$ both correspond to $\langle 1, 2, 3 \rangle$; $\langle 4, 7, 2 \rangle$ corresponds to $\langle 2, 3, 1 \rangle$. In the case we are interested in, if $\vec{z}$ and $\vec{w}$ are two ordered triples that correspond to each other, then A thinks that $\psi(\vec{z})$ iff $\psi(\vec{w})$.

(Think of A as an authority on the truth value of ψ applied to triples of variables. We are interested in the case where A cares only about the order (increasing, decreasing, zig-zag ...) of the subscripts)

When this happens, Ramsey says that A **is serial in** ψ. If A is "serial" in all ψ then it is just plain **serial**. A form is **serial** iff one of its alternatives is serial. (Check for yourself: a form cannot have more than one serial alternative). Why 'serial'? Well, he's got to call it *something*, and the word 'serial' was used in those days for wellorderings, and he may have felt there was a connection of ideas.

We are now ready for some more definitions.

DEFINITION 5 *$\mathcal{I} = \langle I, \leq_{\mathcal{I}} \rangle$ is a* **set of indiscernibles** *(or* **s.o.i***) for $\mathcal{L}$ iff for all $\Phi \in \mathcal{L}$, if Φ is a formula, and n is the number of free variables in it then for all distinct n-tuples $\vec{x}$ and $\vec{y}$ from $\mathcal{I}$* **taken in increasing order** *we have $\Phi(\vec{x}) \longleftrightarrow \Phi(\vec{y})$.*

The idea of a set of indiscernibles first appears in the Ramsey paper under discussion, but the definition is not given explicitly.

It is standard in modern set theory that theorems like $(\forall n, m, k)(\exists x)(x \to (n)^m_k)$ imply the existence of sets of indiscernibles. Usually we are interested in infinite languages, and then we use the infinite version of Ramsey's theorem to show that infinite sets have infinite subsets that are s.o.i.s. Indeed it is usually understood that the $\mathcal{L}$ in the definition of s.o.i. is the full recursive datatype of all expressions that can be built up from a set of predicate letters and functions. However, in the case Ramsey is interested in, $\mathcal{L}$ is simply the finite set of predicate letters itself.

Let us illustrate with a simple case. Suppose we want an s.o.i. for the language containing only one ternary predicate letter F. Fix a total order of the universe, $<$. Partition the set of unordered triples $\{x, y, z\}$ into 2^6 pieces depending on whether or not F holds of the various six *ordered* triples one can make from this *un*ordered triple. By Ramsey's theorem if the Universe is large enough there will be a large monochromatic set. A monochromatic set is of course a set of indiscernibles.

The significance of this notion in the present context is that if A is a serial alternative and f is an assignment satisfying that serial alternative, then f enumerates a set of indiscernibles for the atomic formulæ of $\mathcal{L}$.

THEOREM 6

There is a finite number $m_{\mathcal{L}}$ such that, for every universal sentence Φ,[5]

1. *Either Φ contains a serial form F and all forms $F' \prec F$, in which case Φ has models of all sizes greater than $m_{\mathcal{L}}$; or*

2. *Φ contains no such form, in which case it has only finitely many models.*

Proof:

1. Suppose Φ contains a serial form F and all its restrictions. (In these circumstances we know already from lemma 4 that Φ has a model of size n, because Φ contains a form F and all its restrictions—whether that form be serial or not).

   This time we are given the extra information that the postulated form F is serial, but we are now required to furnish arbitrarily large models. Let M be a set of size $n > p$ and ask ourselves whether there is a way of finding subsets of M^n that interpret the predicate letters of $\mathcal{L}_\Phi$ (the set of predicate letters occurring in Φ) in such a way that Φ comes out true in the model $\mathfrak{M}$ expanded from M by adjoining those interpretations of the predicate letters. We start by fixing $\leq_M$, a wellordering of M. Let A be the serial alternative in F. Let us go back to our special case from earlier, where $\psi(\ )$ was a ternary relation. We can define the $\mathfrak{M}$-interpretation of ψ (a subset of M^3) as follows. We first fix a 1-1 assignment function f. We want to know whether to put $\langle m, m', m'' \rangle$ into the $\mathfrak{M}$-interpretation of ψ or not. m, m'' and m''' are ordered somehow by $\leq_M$. If $m \leq_M m' \leq_M m''$ then we put $\langle m, m', m'' \rangle$ into the $\mathfrak{M}$-interpretation of ψ iff A contains the conjunct $\psi(y_{f^{-1}(m)}, y_{f^{-1}(m')}, y_{f^{-1}(m'')})$. If $m' \leq_M m \leq_M$

[5] If $\mathcal{L}$ has a finite alphabet then $m_{\mathcal{L}}$ depends only on $\mathcal{L}$. Otherwise $m_{\mathcal{L}}$ depends on the predicate letters appearing in Φ.

m'' then we put $\langle m', m, m''\rangle$ into the $\mathfrak{M}$-interpretation of ψ iff A contains the conjunct $\psi(y_{f^{-1}(m)}, y_{f^{-1}(m')}, y_{f^{-1}(m'')})$, ... and so on, so that—in order to determine whether or not we want some given tuple $\langle m, m', m''\rangle$ to belong to the $\mathfrak{M}$-interpretation of ψ—we look at the vector of variables drawn from the set {'y_1', 'y_2', 'y_3'} to which the tuple $\langle y_{f^{-1}(m)}, y_{f^{-1}(m')}, y_{f^{-1}(m'')}\rangle$ corresponds to see whether or not A contains 'ψ' with that variable-vector inserted.

Since A is serial in all the predicate letters, this gives us an interpretation for all the predicate letters. We can think of the concept of 'serial' as being so designed that if A is serial this process never gives conflicting advice on whether or not to put any given tuple into the $\mathfrak{M}$-interpretation of any particular predicate.

This gives us a model $\mathfrak{M} \models \Phi$. In fact the way we have constructed this interpretation tells us a little more. Every tuple from M will satisfy something equivalent to the alternative A, so it satisfies the cut-down version of Φ that we alluded to at the end of the proof of leemma 4.

This construction of $\mathfrak{M}$ is parametrised by $\leq_M$ and A, so if we want to study this in more detail later it might be an idea to adopt a notation that makes this clear.

2. Assume there is a way of imposing interpretations on a set M of m elements to get a model $\mathfrak{M} \models \Phi$. We must show that Φ contains a serial form F and all forms $\prec F$.

   The idea is to show that if m is big enough then if $\mathfrak{M} \models \Phi$ then there is at least one p-tuple from M such that the true alternative for that tuple is a serial one. This is where we will need the finite version of Ramsey's theorem. Let $\mathfrak{M} \models \Phi$ be so big that by the finite version of Ramsey's theorem it has a p-sized set of indiscernibles $\vec{m} = \{m_1 \ldots m_p\}$ for the set of atomic formulæ. $|M|$ is of course the $m_{\mathcal{L}}$ of the statement of theorem 6.

   $\mathfrak{M} \models \Phi$, so every $\mathfrak{M}$-assignment function satisfies Φ. In particular, Φ will be satisfied by the rather special assignment function $f = \lambda i.m_i$ that sends the ith variable to the ith element of the set of indiscernibles. Like any ordinary assignment function, if f satisfies Φ at all it does so in virtue of satisfying a unique alternative, which we may as well call 'A_f' as before. The fact that the range of f is a set of indiscernibles is immensely informative. If A_f contains $\psi(y_1, y_3, y_5)$ (for example) this is because $\langle y_{f(1)}, y_{f(3)}, y_{f(5)}\rangle$ belongs to the $\mathfrak{M}$-interpretation of ψ. But $\langle f(1), f(3), f(5)\rangle = \langle m_1, m_3, m_5\rangle$ (by our careful choice of assignment f!) and—because $\vec{m}$ is a set of indiscernibles—$\langle y_{m_1}, y_{m_3}, y_{m_5}\rangle$ belongs to the $\mathfrak{M}$-interpretation of ψ iff $\langle y_{m_2}, y_{m_3}, y_{m_6}\rangle$ belongs to the $\mathfrak{M}$-interpretation of ψ iff ... and so on for all other increasing triples. So A_f contains $\psi(y_2, y_3, y_6)$ and so on, which means that it is serial in ψ. Indeed—since ψ was arbitrary—A_f is serial in everything else as well, making it serial *tout court*. The form to which A_f belongs is now the serial form whose coming was foretold.

■

Definition 7 *The spectrum of a sentence ψ is the set $\{n \in \mathbb{N} : \psi$ has a model of size $n\}$.*

Corollary 8 *The spectrum of any universal sentence is finite or cofinite.*

3 Existential-Universal Sentences

In the last two pages of the paper Ramsey "sketches" (his word) a way of extending this result to $\exists^*\forall^*$-sentences.

I shall follow the lead given by a witticism of Quine, who used to write existential-universal sentences in the style $(\exists\vec{x})(\forall\vec{y})(\ldots)$; the point being: $x$istential and $y$ouniversal!

Let us assume that our $\exists^*\forall^*$-sentence $(\exists x_1 \ldots x_n)(\forall y_1 \ldots y_k)\Psi$ is in a kind of PNF wher the matrix is in *con*junctive normal form. This enables us to import the universal quantifiers past the conjunctions so we can express it in the form:

$$(\exists\vec{x})([\bigvee_{i\in I} \Psi_i(\vec{x})] \wedge (\forall\vec{y})(\Theta(\vec{x}, \vec{y})))$$

where each Ψ_i is a conjunction of atomics and negatomics concerning $\vec{x}$ only, and Θ is quantifier-free.

Next we note that existential quantification commutes with disjunction to turn this into a disjunction of $\exists^*\forall^*$-sentences of this restricted kind. Then, by a device of relettering like that for which we used the operation `new` defined earlier, we can assume that in all the disjuncts all the x variables are deemed to denote distinct things. Finally, beco's of equivalences like that between $(\forall x)((x = a \wedge A(x)) \vee (x \neq a \wedge B(x)))$ and $(A(a) \wedge (\forall x \neq a)(B(x))$ we may assume that in all the disjuncts—which by now look like

$$(\exists\vec{x})(\Psi(\vec{x}) \wedge (\forall\vec{y})(\Theta(\vec{x}, \vec{y})))$$

the xs are all distinct from all the ys and from each other, and that Θ contains no atomic subformulæ featuring x variables only.

The intention is to prove that sentences of this kind obey theorem 6, for since any $\exists^*\forall^*$ sentence is a disjunction of finitely many of these sentences, that will be enough to ensure that all $\exists^*\forall^*$ sentences do. (If ψ is a disjunction of finitely many sentences each of which has finite or cofinite spectrum then so does ψ)

If a sentence of this new kind fails to be satisfiable we may assume that it is not on account of Ψ (which is after all a conjunction of atomics and negatomics). Next we banish all the existential quantifiers and replace the x variables by constants. Then we introduce a new suite of predicate letters which we will use to replace occurrences of old predicates which have constants in any of their argument places. Thus an old binary predicate letter 'F' occurring in contexts like '$F(a, y)$' and '$F(y, a)$' will give rise to new predicate letters in contexts like '$F_{a_1}(y)$' and '$F_{a_2}(y)$' and for each constant a. Finally atomic formulæ mentioning only 'x' variables get replaced by a propositional constant.

But what we now have is a $\forall^*$-formula, and we know how to test these for satisfiability!

References

[1] Ramsey, F. P. (1929) 'On a problem of formal logic'. Proc. London Math. Soc. **30** 264−−286.