The quantifier complexity of NF

Richard Kaye

Abstract

Various issues concerning the quantifier complextity (i.e., number of alternations of like quantifiers) of Quine's theory NF, its axiomatizations, and some of its subtheories, are discussed.

1 Introduction

In this paper I shall investigate various issues concerning the quantifier complexity of NF. The main motivation for this work concerns the consistency problem for NF (see for example Boffa [1977]) and is as follows.

1. Kaye [1991] proved a generalization of a theorem of Specker's [1962] concerning the equiconsistency of NF and an extension of the theory of simple types, TST, the so-called *ambiguity axioms* Amb(ϕ). In particular, the author's modification shows Specker's theorem holds even if these ambiguity axioms are restricted to ϕ in a certain complexity class. The exact class of formulas here depends on the on the complexity of axiomatizations of NF. Roughly speaking, if we can find axiomatizations of NF of low complexity, the modification of Specker's theorem is more powerful and potentially more useful. Obtaining upper bounds on the complexity of axiomatizations of NF is the subject of section 1 below.

2. There is a great deal of interest in determining fragments of NF that are actually decidable. Hinnion [1972] has shown that (assuming NF to be consistent) the set of universal consequences of NF is decidable and complete in a certain sense. (See also Kreinovič and Oswald [1982].) Forster believes that a similar result for the stratified $\forall_2$ consequences of NF should hold, and although he has the best part of a

Received by the editors February 1995.
Communicated by P. Laubin.
1991 *Mathematics Subject Classification* : 03E70.
Key words and phrases : New foundations, quantifier complexity.

proof of this fact, he tells me that there is still a gap in his arguments to be filled in. Of course, by Gödel's results, if NF is consistent, it is undecidable. At what point does this undecidability occur? Similarly, we can ask whether all stratified formulas are equivalent in NF to stratified formulas of some bounded quantifier complexity. The answer turns out to be 'yes': this is the topic of investigation in section 2 below.

3. NF is known to be finitely axiomatizable (Hailperin [1944]). It is often said that this is a major stumbling block in attempts to prove NF consistent using compactness. In fact it is rather easy to obtain from various axiomatizations of NF natural subtheories of NF that are provably equiconsistent with NF, but not apparently finitely axiomatized, as I shall explain in section 3. The idea is to construct subtheories of NF over which NF is actually conservative for certain classes of sentences, just as Gödel–Bernays set theory (GB) is conservative over ZFC for sentences in the language of ZFC, but GB is finitely axiomatized and ZFC not.

This paper is selfcontained, except for references to the main result of Kaye [1991]. However, it may mean little to a reader lacking the background knowledge required to motivate the subject. For this I can recommend Boffa's paper [1977], or Forster's book [1992].

Part of this work was described in a talk presented at the ASL meeting in Granada in 1987 (see the abstract Kaye [1989]). The paper lay dormant for some time after that, at first due to other more pressing work on my desk, and then due to publication difficulties with the book for which it was refereed and accepted in 1993. I would like to acknowledge with thanks the friendly help and advice I have received from Maurice Boffa and Thomas Forster on the subject matter discussed here over the last few years.

2 The complexity of axiomatizations

$\mathcal{L}$ denotes the usual (one-sorted) language of set theory, $\{\in, =\}$. The collection of all open (i.e., quantifier-free) formulas of $\mathcal{L}$ is denoted O. $\exists_n$ and $\forall_n$ are defined as usual by:

$$\begin{aligned} \exists_0 = \forall_0 &= \mathrm{O}; \\ \exists_{n+1} &= \{\exists \vec{x} \gamma(\vec{x}, \vec{y}) : \gamma \in \forall_n\}; \end{aligned}$$

and

$$\forall_{n+1} = \{\forall \vec{x} \gamma(\vec{x}, \vec{y}) : \gamma \in \exists_n\}.$$

For $n \geq 0$ we define $\{x_1, x_2, \ldots, x_n\} = y$ to be the $\mathcal{L}$-formula

$$\forall z (z \in y \longleftrightarrow \bigvee_{i=1}^{n} (z = x_i)),$$

where the empty disjunction $\bigvee_{i=1}^{0}(z = x_i)$ takes truth-value *false*. O^+ is the smallest collection of $\mathcal{L}$-formulas containing $x = y$, $x \in y$ and $x = \{y, z\}$ for all possible choices of variables x, y, z and closed under $\wedge$, $\vee$ and $\neg$. $\exists_n^+$ and $\forall_n^+$ are then defined by:

$$\exists_0^+ = \forall_0^+ = \mathrm{O}^+;$$

$$\exists^+_{n+1} = \{\exists\vec{x}\gamma(\vec{x},\vec{y}\,) : \gamma \in \forall^+_n\};$$

and

$$\forall^+_{n+1} = \{\forall\vec{x}\gamma(\vec{x},\vec{y}\,) : \gamma \in \exists^+_n\}.$$

(My original definition of these classes allowed $\{x_1, x_2, \ldots, x_n\} = y$ to be present in O^+ for all n. I am indebted to Maurice Boffa for pointing out that the results presented below are true for the more restrictive classes as defined above.)

Notice that every O^+ formula is a boolean combination of $\forall_1$ formulas, so

$$\exists_{n+1} \supseteq \exists^+_n \supseteq \exists_n$$

and

$$\forall_{n+1} \supseteq \forall^+_n \supseteq \forall_n$$

for all $n \geq 1$. If Γ is a collection of formulas of $\mathcal{L}$, $\forall\Gamma$ denotes $\{\forall\vec{x}\gamma : \gamma \in \Gamma\}$, and similarly for $\forall\exists\Gamma$, etc.

The definitions of the complexity classes above also make sense in the many-sorted language $\mathcal{L}_{\mathrm{TST}}$ of the theory of simple types, and I shall occasionally employ the notations O, O^+, $\exists_n$, $\forall_n$, $\exists^+_n$, and $\forall^+_n$ in this context, without making explicit reference to the different language involved. If ϕ is a formula of $\mathcal{L}_{\mathrm{TST}}$ then ϕ^+ denotes the result of raising all type indices of variables in ϕ by 1, and $\phi^\#$ denotes the sentence of $\mathcal{L}$ obtained by omitting all reference to types in ϕ. The sentence $\mathrm{Amb}(\phi)$ is the $\mathcal{L}_{\mathrm{TST}}$-sentence $\phi \longleftrightarrow \phi^+$, and for a collection of formulas Γ, $\mathrm{Amb}(\Gamma)$ denotes the axiom scheme $\phi \longleftrightarrow \phi^+$ for all $\phi \in \Gamma$.

If Γ is a collection of $\mathcal{L}$-formulas, $str\Gamma$ denotes the collection of *stratified* formulas $\gamma \in \Gamma$, i.e., those $\gamma \in \Gamma$ for which the variables may be assigned natural numbers such that, if $x = y$ is a subformula then x and y are assigned the same number, and if $x \in y$ is a subformula then x is assigned a number one less than that assigned to y, or, put another way, the collection of all γ in Γ such that $\gamma = \phi^\#$ for some $\mathcal{L}_{\mathrm{TST}}$-formula ϕ.

The usual axiom of extensionality is an $\forall_2$ sentence, and will be denoted Ext. For a collection of formulas, γ, $\mathrm{NF}\Gamma$ denotes the theory axiomatized by Ext together with the scheme of *set-abstraction*,

$$\forall\vec{a}\exists x\forall y(y \in x \longleftrightarrow \gamma(y,\vec{a})) \qquad (*)$$

for all stratified formulas $\gamma \in \Gamma$. NF is Ext together with $(*)$ for all stratified γ. For example, it is not difficult to check (using induction on the complexity of γ in $(*)$) that NFO is equivalent to the finitely axiomatized theory consisting of Ext together with axioms stating that the following exist for all arguments x, y:

$$V = \{z : z = z\};$$

$$\Lambda = \{z : z \neq z\};$$

$$x \cap y = \{z : z \in x \wedge z \in y\};$$

$$x \cup y = \{z : z \in x \vee z \in y\};$$

$$x - y = \{z : z \in x \wedge z \notin y\};$$

$$\iota(x) = \{x\};$$

$$B(x) = \{z : x \in z\}; \text{ and}$$

$$-x = \{z : z \notin x\}.$$

This theory is known to be consistent, and has several pleasant model-theoretic properties—see Forster [1987] and [1992].

Using the complement function, $-x$, we see immediately that for all n,

$$\mathrm{NF}\exists_n = \mathrm{NF}\forall_n$$

and

$$\mathrm{NF}\exists_n^+ = \mathrm{NF}\forall_n^+.$$

THEOREM 2.1 $\mathrm{NF} = \mathrm{NF}\exists_2 = \mathrm{NF}\exists_1^+$.

Proof. (Maurice Boffa has pointed out to me that this follows directly from the Hailperin axiomatization of NF, Hailperin [1944]. I give here a direct proof.)

It suffices to show, for each $n \geq 1$, that $\mathrm{NF}\forall_n^+ \vdash \mathrm{NF}\exists_{n+1}^+$. Define

$$\langle x, y \rangle =_{\mathrm{def}} \{\{x\}, \{x, y\}\},$$

$$\langle x, y, z \rangle =_{\mathrm{def}} \Big\langle \{\{x\}\}, \langle y, z \rangle \Big\rangle,$$

$$\langle x, y, z, w \rangle =_{\mathrm{def}} \Big\langle \{\{\{\{x\}\}\}\}, \langle y, z, w \rangle \Big\rangle,$$

etc. (Notice that in an assignment of types to $y = \langle x_0, \ldots, x_n \rangle$ the x_is are given the same type, and y is given type $2n$ higher than this. This notion of an ordered tuple is due to Hailperin.) Suppose $\phi(x_0, x_1, \ldots, x_l)$ is $str\forall_n^+$ (possibly containing other parameters $\vec{a}$) where x_j is given type i_j in some stratification assignment of the variables in ϕ, and $i = \max\{i_j : j \leq l\}$. Then

$$a = \left\{ z : \forall x_0, \ldots, x_l \left(\begin{array}{r} z = \langle \iota^{i-i_0}(x_0), \ldots, \iota^{i-i_l}(x_l) \rangle \rightarrow \\ \phi(x_0, \ldots, x_l) \end{array} \right) \right\}$$

exists, by virtue of an axiom of $\mathrm{NF}\forall_n^+$; and

$$b = \{x_0 : \exists x_1, \ldots, x_l \, \phi(x_0, x_1, \ldots, x_l)\}$$

$$= \{x_0 : \exists x_1, \ldots, x_l, z \, (z = \langle \iota^{i-i_0}(x_0), \ldots, \iota^{i-i_l}(x_l) \rangle \wedge z \in a)\}$$

exists, by virtue of an axiom of $\mathrm{NF}\exists_1^+$. ■

Using the main theorem of Kaye [1991] we have

COROLLARY 2.2 NF is consistent if and only if $\mathrm{TST} + \mathrm{Amb}(\exists_4)$ is consistent if and only if $\mathrm{TST} + \mathrm{Amb}(\exists_3^+)$ is consistent. ■

For a collection Γ of formulas and an $\mathcal{L}$-theory T, ΓT denotes the collection of *sentences* $\gamma \in \Gamma$ provable in T. We make the convention that $str\Gamma T$ refers to $(str\Gamma)T$, the stratified Γ-consequences of T, and not to $str(\Gamma T)$, the stratified consequences of ΓT. Theorem 1.1 clearly implies $str\forall_4^+\mathrm{NF} = str\forall_5\mathrm{NF} = \mathrm{NF}$, i.e., NF has a $str\forall_4^+$-axiomatization. It can now be improved slightly to give...

Theorem 2.3 $\mathrm{NF} = str\exists_3^+\mathrm{NF} + str\forall_4\mathrm{NF}$. Indeed, NF is equivalent to NFO plus the parameter-free abstraction scheme,

$$\exists x \forall y (y \in x \longleftrightarrow \theta(y))$$

for all stratified $\theta \in \exists_1^+$ with only y free, plus the $str\forall_4$ axiom

$$\forall x \exists y \forall z (z \in y \longleftrightarrow \exists w (z \in w \wedge w \in x))$$

expressing the existence of the sum set $\bigcup x$ of x.

Proof. It suffices by the previous theorem to prove the existence of

$$\{x : \exists \vec{y}\, \phi(x, \vec{y}, \vec{a})\}$$

in the subtheory of NF indicated in the statement of the theorem, $\phi(x, \vec{y}, \vec{a})$ being an arbitrary stratified O^+ formula. Let $(y_0, y_1, \ldots, y_m)$ be the tuple of variables $\vec{y}$ and $(a_0, a_1, \ldots, a_n)$ be the tuple of parameters $\vec{a}$, where $y_0, y_1, \ldots, y_m$ are given types $i_0, i_1, \ldots, i_m$, $a_0, a_1, \ldots, a_n$ are given types $j_0, j_1, \ldots, j_n$, and x is given type k in some assignment of types to $\phi(x, \vec{y}, \vec{a})$.

Let l_x be the maximum of $\max(i_0, i_1, \ldots, i_m) + 2m$, $\max(j_0, j_1, \ldots, j_n) + 2n$, and k, and let $l_y = l_x - 2m$ and $l_z = l_x - 2n$. Notice that $l_x \geq k$, $l_y \geq \max(\vec{\imath})$, and $l_z \geq \max(\vec{\jmath})$. We write...

$$\langle\!\langle \vec{y} \rangle\!\rangle \text{ for } \langle \iota^{(l_y - i_0)}(y_0), \iota^{(l_y - i_1)}(y_1), \ldots, \iota^{(l_y - i_m)}(y_m) \rangle,$$

$$\langle\!\langle \vec{z} \rangle\!\rangle \text{ for } \langle \iota^{(l_z - j_0)}(z_0), \iota^{(l_z - j_1)}(z_1), \ldots, \iota^{(l_z - j_m)}(z_m) \rangle,$$

$$\langle\!\langle \vec{a} \rangle\!\rangle \text{ for } \langle \iota^{(l_z - j_0)}(a_0), \iota^{(l_z - j_1)}(a_1), \ldots, \iota^{(l_z - j_m)}(a_m) \rangle,$$

and

$$\langle\!\langle x \rangle\!\rangle \text{ for } \iota^{l_x - k}(x).$$

It is simple to check that, in the stratification assignment of the variables in $\phi(x, \vec{y}, \vec{z})$ given above, $\langle\!\langle x \rangle\!\rangle$, $\langle\!\langle \vec{y} \rangle\!\rangle$, and $\langle\!\langle \vec{z} \rangle\!\rangle$ are all given the same type, l_x. This means that, by parameter-free $\exists_1^+$-abstraction, the set

$$b = \Big\{ \Big\langle \Big\langle \langle\!\langle \vec{z} \rangle\!\rangle, \{\{\Lambda\}\} \Big\rangle, \Big\langle \langle\!\langle x \rangle\!\rangle, \Lambda \Big\rangle, \Big\langle \langle\!\langle \vec{y} \rangle\!\rangle, \{\Lambda\} \Big\rangle \Big\rangle : \phi(x, \vec{y}, \vec{z}) \Big\}$$

$$= \left\{ w : \exists x, \vec{y}, \vec{z}, r, s, t \begin{pmatrix} w = \Big\langle \Big\langle \langle\!\langle \vec{z} \rangle\!\rangle, t \Big\rangle, \Big\langle \langle\!\langle x \rangle\!\rangle, r \Big\rangle, \Big\langle \langle\!\langle \vec{y} \rangle\!\rangle, s \Big\rangle \Big\rangle \wedge \\ r = \Lambda \wedge s = \{\Lambda\} \wedge t = \{\{\Lambda\}\} \wedge \phi(x, \vec{y}, \vec{z}) \end{pmatrix} \right\}$$

exists, since $w = \Big\langle \Big\langle \langle\!\langle \vec{z} \rangle\!\rangle, t \Big\rangle, \Big\langle \langle\!\langle x \rangle\!\rangle, r \Big\rangle, \Big\langle \langle\!\langle \vec{y} \rangle\!\rangle, s \Big\rangle \Big\rangle$ is $\exists_1^+$ and ϕ is stratified.

Now let

$$c = \Big\langle \langle\!\langle \vec{a} \rangle\!\rangle, \{\{\Lambda\}\} \Big\rangle$$

(which exists in NFO). Notice that by the definition of the pairing function employed above,

$$\langle u, v, w\rangle = \Big\{ \big\{\{\{u\}\}\big\}, \big\{\{\{u\}\}, \langle v, w\rangle\big\} \Big\},$$

so

$$\begin{aligned} d &= b \cap B(\{\{\{c\}\}\}) \\ &= \Big\{ \Big\langle \Big\langle \langle\!\langle \vec{a} \rangle\!\rangle, \{\{\Lambda\}\} \Big\rangle, \Big\langle \langle\!\langle x \rangle\!\rangle, \Lambda \Big\rangle, \Big\langle \langle\!\langle \vec{y} \rangle\!\rangle, \{\Lambda\} \Big\rangle \Big\rangle : \phi(x, \vec{y}, \vec{a}) \Big\} \end{aligned}$$

exists by the NFO axioms. It is now straightforward to check that

$$\begin{aligned} e &= \Big(\bigcup{}^2 d\Big) \cap \Big\{ x : \exists x_1, x_2\, (x_1 \neq x_2 \wedge x_1 \in x \wedge x_2 \in x) \Big\} \\ &= \Big\{ \Big\langle \langle \langle\!\langle x \rangle\!\rangle, \Lambda\rangle, \langle \langle\!\langle \vec{y} \rangle\!\rangle, \{\Lambda\}\rangle \Big\rangle : \phi(x, \vec{y}, \vec{a}) \Big\}. \end{aligned}$$

Then

$$\begin{aligned} f &= \Big(\bigcup{}^2 e\Big) \cap \Big\{ \langle \langle\!\langle x \rangle\!\rangle, \Lambda\rangle : x \in V \Big\} \\ &= \Big\{ \langle \langle\!\langle x \rangle\!\rangle, \Lambda\rangle : \exists \vec{y}\, \phi(x, \vec{y}, \vec{a}) \Big\}, \end{aligned}$$

exists, and $\{x : \exists \vec{y}\, \phi(x, \vec{y}, \vec{a})\}$ equals $\bigcup^{(2+l_x)} f$ or $\bigcup^{(2+l_x)} f - \{\Lambda\}$. ■

I conclude this section with three open problems, to be solved under the assumption that NF is consistent. I conjecture that the answer is in the negative in each case.

PROBLEM 2.4 Does NF $=$ NF$\exists_1$?

PROBLEM 2.5 Does NF $= \forall_4$NF?

PROBLEM 2.6 Does NF $= \exists_3^+$NF?

All that I know on these lines is that, assuming NF to be consistent, $\exists_2$NF $\nvdash$ NF (for any sentence of $\exists_2$NF has a finite model, whereas $\exists_2$NF itself does not, so $\exists_2$NF is not finitely axiomatized), and $\forall_2$NF $\nvdash$ NF (see Forster and Kaye [1991], where the stronger result $\Pi_2^{\mathcal{P}}$NF $\nvdash \exists V \forall x\, x \in V$ is proved).

3 Quantifier complexity of terms, and decidability

An NF *term* is an expression of the form $\{x : \psi(x, \vec{y})\}$, where $\psi(x, \vec{y})$ is a stratified $\mathcal{L}$-formula. Thus the NF terms are exactly the natural expressions $t(\vec{y})$ for the sets whose existence is guaranteed by the stratified set abstraction axioms of NF. Similarly, we define NF$\exists_1^+$ terms (where the formula $\psi(x, \vec{y})$ must be stratified and of complexity no more than $\exists_1^+$), and *closed* terms (i.e., terms not containing free-variables $\vec{y}$), etc. It is important to realize that terms are only expressions for sets, and not themselves sets. Indeed two closed terms may turn out to be equal in one model of NF and different in a second one.

Forster [1987] has shown that there are natural definitions of $\in$ and $=$ on the closed NFO terms such that the collection of these terms with these relations forms a model of NFO, called *the term model for* NFO. At present though, this seems to be a special feature of NFO (and the theory NF$\forall$ considered also by Forster in his [1987] paper, where this is Ext together with the set abstraction scheme for stratified formulas with only one universal quantifier), and whether term models for the full theory NF exist is still open. However, the proofs of the two theorems in the last section, gives the following useful information about NF terms.

THEOREM 3.1 Let $t(\vec{y})$ be an NF term. Then there are $n \in \mathbb{N}$ and NF terms $t_0(\vec{y})$, $t_1(\vec{y}, z_0)$, $t_2(\vec{y}, z_0, z_1)$,... $t_n(\vec{y}, z_0, z_1, \ldots, z_{n-1})$, where $t_i(\vec{y}, z_0, z_1, \ldots, z_{i-1})$ is:

either	$z_j \cap z_k$	(some $j, k < i$)
or	$-z_j$	(some $j < i$)
or	$B(z_j)$	(some $j < i$)
or	$\bigcup z_j$	(some $j < i$)
or	y_l	(some $l < \mathrm{len}(\vec{y})$)
or	t	(some closed NF$\exists_1^+$ term t),

such that, in any model of NF and for any $\vec{a} \in M$, if

$$M \models b_i = t_i(\vec{a}, b_0, b_1, \ldots, b_{i-1}) \qquad (i = 1, \ldots, n)$$

then $M \models b_n = t(\vec{a})$.

Proof. By induction on the quantifier complexity of the term $t(\vec{y})$.

If $t(\vec{y})$ is $\{x : \psi(x, \vec{y})\}$ with ψ in $\exists_{n+1}^+$, use the argument in the proof of theorem 1.1. If ψ is $\forall_{n+1}^+$, find $t_n(\vec{y}, \vec{z}) = \{x : \neg\psi(x, \vec{y})\}$ using the induction hypothesis (since $\neg\psi$ is $\exists_{n+1}^+$), and let $t_{n+1}(\vec{y}, \vec{z}) = -t_n(\vec{y}, \vec{z})$. This reduces the proof to ψ in $\exists_1^+$. In this case, use the argument in theorem 1.3. ∎

COROLLARY 3.2 For all stratified $\mathcal{L}$-formulas $\phi(\vec{x})$ there is a stratified $\exists_3^+$ $\mathcal{L}$-formula $\theta(\vec{x})$ such that

$$\mathrm{NF} \vdash \forall\vec{x}\,(\phi(\vec{x}) \longleftrightarrow \theta(\vec{x}))$$

Proof. The idea is to let $\theta(\vec{x})$ be the formula

$$\exists z_0, \ldots, z_n \Big(\bigwedge_i z_i = t_i(z_0, \ldots, z_{i-1}) \wedge \langle \vec{x}\,\rangle \in z_n \Big)$$

where the terms t_0, $t_1(z_0)$,..., $t_n(z_0, \ldots, z_{n-1})$ are chosen by the previous theorem so that

$$t_n(t_0, t_1(t_0), \ldots, t_{n-1}(t_0, \ldots, t_{n-2}(\cdots))) = \{\langle \vec{x}\,\rangle : \phi(\vec{x})\}.$$

(Here the exact notion of coding tuples $\langle \vec{x}\,\rangle$ can be chosen using $\iota^{(i-i_j)}$—as in the proof of theorem 1.1—so that $\{\langle \vec{x}\,\rangle : \phi(\vec{x})\}$ is a valid NF term.) Although each expression of the form $z_i = t_i(z_0, \ldots, z_{i-1})$ is $\forall_2^+$ and stratified, there may be a problem in that the conjunction $\bigwedge_i z_i = t_i(z_0, \ldots, z_{i-1})$ is not stratified because the different component expressions require different type-assignments to the z_is. However, the solution to this problem is easily found: we simply create a different copy of each z_i at each different type that it is required at, using induction. ∎

Corollary 3.3 If NF is consistent, then neither of the following sets of Gödel numbers

$$A = \{\ulcorner\phi\urcorner : \phi \in str\exists_3^+ \text{ and } \mathrm{NF} \vdash \phi\}$$

and

$$B = \{\ulcorner\phi\urcorner : \phi \in str\exists_3^+ \text{ and } \mathrm{NF} \vdash \neg\phi\}$$

is recursive. Indeed A and B are disjoint recursively inseparable r.e. sets.

Proof. There is a well-known interpretation of PA in NF. If $\mathrm{Int}(\psi)$ is the interpretation of the $\mathcal{L}_{\mathrm{PA}}$-sentence ψ, then by the previous corollary we may assume $\mathrm{Int}(\psi)$ is $str\exists_3^+$, and clearly we may assume the mapping Int is recursive. The corollary now follows, for if $C \supseteq A$, $C \cap B = \emptyset$, and C is recursive, then

$$D = \{\ulcorner\psi\urcorner : \psi \in \mathcal{L}_{\mathrm{PA}} \text{ and } \mathrm{PA} \vdash \psi\}$$

and

$$E = \{\ulcorner\psi\urcorner : \psi \in \mathcal{L}_{\mathrm{PA}} \text{ and } \mathrm{PA} \vdash \neg\psi\}$$

could be separated by the recursive set

$$F = \{\ulcorner\psi\urcorner : \psi \in \mathcal{L}_{\mathrm{PA}} \text{ and } \mathrm{Int}(\psi) \in C\}$$

which is impossible. ■

It follows trivially from the last corollary that NF is not complete for stratified $\exists_3^+$ sentences, i.e., there is such a sentence ψ with neither $\mathrm{NF} \vdash \psi$ nor $\mathrm{NF} \vdash \neg\psi$. This is also true of unstratified $\exists_2$ sentences, for example: NF does not decide the sentence $\exists x \forall y (y \in x \longleftrightarrow y = x)$ stating 'there exists a Quine atom' (see for example Forster [1992] for a proof of this well-known fact). But on the other hand, Hinnion [1972] has shown that NF is complete for all $\exists_1$ sentences, and Forster conjectures that NF *does* decide all *stratified* $\exists_2$ sentences. Between these, there is still a lot to be known:

Problem 3.4 Is

$$\{\ulcorner\phi\urcorner : \phi \in str\exists_3 \text{ and } \mathrm{NF} \vdash \phi\}$$

recursive? What happens if we ask the same question for $\exists_2$, or for $str\exists_2^+$ in place of $str\exists_3$ here?

Problem 3.5 Does NF decide all $str\exists_3$ sentences? All $str\exists_2^+$ sentences?

4 Equiconsistent subtheories

The results already presented suggest several 'term model'-like constructions of models of NF, for which (we will see) the subtheory $str\exists_3^+\mathrm{NF}$ of NF plays an important role. Note also that the fact that every stratified expression in NF is equivalent to a stratified $\exists_3^+$ expression does not itself imply that $\mathrm{NF} = str\exists_3^+\mathrm{NF}$, and indeed there seems to be no *a priori* reason to expect that $str\exists_3^+\mathrm{NF}$ should be finitely axiomatizable. So perhaps this may in the future enable someone to construct models of NF from a ZF-like model, using the compactness theorem perhaps.

Now, it is a fact that, given a suitable quantifier complexity class Γ (such as $\exists_n$, $\forall_{n+1}$, $\exists_n^+$, or $\forall_{n+1}^+$ where $n \geq 1$, but unfortunately *not* $\forall_1$ or $\forall_1^+$) and a theory T, an axiomatization of ΓT can be written down directly. In general, this axiomatization may be extremely complicated and rather uninformative, but in particular cases (such as set theory or arithmetic) there are often several useful simplifications one can make. (Some examples from arithmetic are considered in the papers Kaye [1987] and Kaye, Paris and Dimitracopoulos [1988].) In the NF case, the results on terms presented above simplify the construction of models of NF from models of $\exists_3^+$NF, and these constructions can be used to verify that given axiom schemes do indeed axiomatize fragments such as $\exists_3^+$NF. Thus we will be able to find interesting subtheories of NF that are actually provably equiconsistent with it. The first result on these lines follows.

THEOREM 4.1 Let T_0 denote the theory $\mathrm{Ext} + \exists\Lambda\forall x x \notin \Lambda + \forall x, y \exists z\, z = \{x, y\}$, and let T denote T_0 together with all axioms of the form

$$\exists x_1, y_1, x_2, y_2, \ldots, x_n \begin{pmatrix} x_1 = \{x : \phi_1(x)\} \wedge \\ (x_1 \neq \Lambda \rightarrow y_1 \in x_1) \wedge \\ x_2 = \{x : \phi_2(x, x_1, y_1)\} \wedge \\ (x_2 \neq \Lambda \rightarrow y_2 \in x_2) \wedge \\ \ldots \\ (x_{n-1} \neq \Lambda \rightarrow y_{n-1} \in x_{n-1}) \wedge \\ x_n = \{x : \phi_n(x, x_1, y_1, \ldots, x_{n-1}, y_{n-1})\} \end{pmatrix}$$

where the $\phi_i(x, x_1, y_1, ..., x_{i-1}, y_{i-1})$ are $str\forall_1^+$ formulas of $\mathcal{L}$ with the free variables shown. Then $T = \Gamma$NF, where Γ is the class of sentences $\exists \vec{x}\, \phi(\vec{x})$ of $\mathcal{L}$ for ϕ a boolean combination of $str\forall_2^+$ formulas.

Proof. One direction is easy, for the theory T is clearly a consequence of NF, and the reader may easily check that all its axioms are sentences from the class Γ.

Now fix a sentence γ of Γ, and suppose that $T + \neg\gamma$ is consistent. We shall construct a model of NF $+ \neg\gamma$, thus showing γ is not a consequence of NF. Let $M \models T + \neg\gamma$ be $\aleph_1$-saturated and fix some enumeration

$$\phi_i(x, x_1, y_1, \ldots, x_{i-1}, y_{i-1})$$

of all $str\forall_1^+$ formulas with at most the free variables shown. By saturation and the axioms of T we may assume that M contains elements $a_1, b_1, \ldots, a_i, b_i, \ldots$ such that

$$M \models \bigwedge_{i \in \mathbb{N}} \begin{pmatrix} a_i = \{x : \phi_i(x, a_1, b_1, \ldots, a_{i-1}, b_{i-1})\} \wedge \\ (a_i \neq \Lambda \rightarrow b_i \in a_i) \wedge \\ (a_i = \Lambda \rightarrow b_i = \Lambda) \end{pmatrix}.$$

Let K be the substructure of M with domain $\{a_1, b_1, \ldots, a_i, b_i, \ldots\}$. We check that $K \models \mathrm{NF} + \neg\gamma$.

$K \models$ Ext. If $x, y \in K$ and $x \neq y$ then $x - y \neq \Lambda$ in M, by Ext in M. But then $a_i = x - y$ for some i, so $b_i \in a_i$.

$K \prec_{str\exists_2^+} M$. Suppose $M \models \exists \vec{x} \phi(\vec{x}, \vec{a}, \vec{b})$ where ϕ is $str\forall_1^+$ and $\vec{x} = (x_1, \ldots, x_l)$, x_j being given type i_j in a stratification assignment of ϕ, and $i = \max(\vec{\imath})$. Consider

$$a_r =_{\mathrm{def}} \{y : \forall \vec{x}\, (\langle \iota^{(i-i_1)}(x_1), \ldots, \iota^{(i-i_l)}(x_l) \rangle = y \rightarrow \phi(\vec{x}, \vec{a}, \vec{b}\,))\},$$

$$a_s =_{\text{def}} \{y : \forall \vec{x}\, (\langle \iota^{(i-i_1)}(x_1), \ldots, \iota^{(i-i_l)}(x_l) \rangle \neq y)\},$$

$$a_t =_{\text{def}} \{y : y \notin a_s\} = -a_s,$$

and

$$a_u =_{\text{def}} \{y : y \in a_r \wedge y \in a_t\} = a_r \cap a_t.$$

Since a_u is nonempty in M, we have $b_u \in a_u$, i.e.

$$b_u = \langle \iota^{(i-i_1)}(x_1), \ldots, \iota^{(i-i_l)}(x_l) \rangle$$

and

$$M \models \phi(\vec{x}, \vec{a}, \vec{b}\,)$$

for some $\vec{x} \in M$. Clearly, $\vec{x} \in K$, so

$$\exists \vec{x} \in K \quad M \models \phi(\vec{x}, \vec{a}, \vec{b}\,).$$

It follows from Tarski's test on elementary equivalence, and the fact that $str\exists_2^+$ is closed under subformulas, that $K \prec_{str\exists_2^+} M$.

In particular, from the last paragraph we have: for all sentences σ in Γ, $K \models \sigma$ implies $M \models \sigma$. Thus $K \models \neg\gamma$.

Finally, to check that $K \models$ NF, it suffices by theorem 1.1 and complements to show that $K \models \mathrm{NF}\forall_1^+$. But if $\phi(x, \vec{a}, \vec{b})$ is stratified and $\forall_1^+$ then

$$c =_{\text{def}} \{x : \phi(x, \vec{a}, \vec{b}\,)\} \in K$$

and also

$$\forall x\, (x \in c \longleftrightarrow \phi(x, \vec{a}, \vec{b}))$$

is $str\forall_2^+$, so absolute between M and K, so

$$K \models \forall x\, (x \in c \longleftrightarrow \phi(x, \vec{a}, \vec{b}))$$

as required. ■

Problem 4.2 Is the theory T in the statement of the last theorem finitely axiomatized?

Stronger results than the last theorem can be obtained by axiomatizing ΓNF for more restricted classes of formulas Γ. For example, the last theorem can be modified by (a) using one of the other axiomatizations in section 1 above, and/or (b) examining the quantifier structure of formulas such as $y = \{x : \phi(x, \vec{z})\}$ that appear in axioms of T above. It may perhaps be more interesting however to axiomatize *stratified* fragments of NF, and this is what I shall discuss now.

The most obvious approach to this type of problem is to use the ideas of the last theorem as follows: given a sufficiently saturated model M of the appropriate theory build a submodel $\langle K_0, K_1, \ldots \rangle$ of $\langle M, M, \ldots \rangle$ which is elementary for $\exists_2^+$ formulas, satisfies the $\forall_1^+$ set abstraction scheme of TST, and also satisfies a sufficiently strong ambiguity scheme.

To give an idea of the method, let us briefly discuss an axiomatization of the theory $str\exists_3^+$NF. Fix three families of variables $X = \{x_j^i : i, j \in \mathbb{N}\}$, $Y = \{y_j^i : i, j \in \mathbb{N}\}$, and $Z = \{z_j^i : i, j \in \mathbb{N}\}$, and suppose each $\phi_i(v, \vec{x}, \vec{y})$ is a stratified $\forall_1^+$ formula in free variables v, $\vec{x}$, and $\vec{y}$, where these have the property that no integer j greater than or equal to i can appear as a subscript in any x_n^m, y_n^m in $\vec{x}$, $\vec{y}$. If $\vec{u}$ is a tuple of variables from $X \cup Y \cup Z$, we write $\vec{u}^+$ for the result of increasing all superscripts of variables in $\vec{u}$ by one. Then, for any formulas ψ_j, the existential closure of the following formula is a consequence of NF:

$$\begin{array}{l}
x_1^1 = \{v : \phi_1(v)\} \wedge x_1^2 = \{v : \phi_1(v)\} \wedge \cdots x_1^k = \{v : \phi_1(v)\} \wedge \\
\bigwedge_i \forall z_1^i (z_1^i \in x_1^i \to y_1^i \in x_1^i) \wedge \\
x_2^1 = \{v : \phi_2(v, \vec{x}, \vec{y})\} \wedge \cdots x_2^k = \{v : \phi_2(v, \vec{x}, \vec{y})\} \wedge \\
\ldots \\
\bigwedge_i \forall z_{k-1}^i (z_{k-1}^i \in x_{k-1}^i \to y_{k-1}^i \in x_{k-1}^i) \wedge \\
x_k^1 = \{v : \phi_k(v, \vec{x}, \vec{y})\} \wedge \cdots x_k^k = \{v : \phi_k(v, \vec{x}, \vec{y})\} \wedge \\
\bigwedge_j \psi_j(\vec{x}, \vec{y}) \leftrightarrow \psi_j(\vec{x}^+, \vec{y}^+).
\end{array}$$

(Here, variables $\vec{x}$ and $\vec{y}$ range over X and Y respectively and k is some natural number.) The reason why this is provable in NF is that, in NF, we can take $x_i^1 = x_i^2 = \cdots x_i^k$, and similarly for the y_i^js. However, by the method of proof of the last theorem, and by the main result of Kaye [1991], the theory T_0 of theorem 3.1 together with the collection of *stratified* axioms of the above form, with the ψ_j all $\exists_2^+$ is an axiomatization of $str\exists_3^+$NF. I omit the proof here, but I have already indicted how it goes: one uses saturation in a ground model to build a model of the $\exists_1^+$ set abstraction scheme in the language of TST. Note in particular that we may assume without loss of generality that the type of a set realizing the variable x_j^{i+1} is one higher than a set realizing x_j^i (but the type of x_j^i may not be i.

In fact, a slight simplification can be made: it suffices that the formulas ψ_j above are all of the form $u \in v$ for variables $u, v \in X \cup Y$. This is because, in the constructed model where a_j^i realizes the variable x_j^i, and b_j^i realizes y_j^i, it suffices that the map $\tau \colon a_j^i \mapsto a_j^{i+1}$ and $b_j^i \mapsto b_j^{i+1}$ preserves $\in$. (It follows from this and extensionality that it will also be one-to-one.) This map τ may then take the place of a type-shifting automorphism of the many-sorted model (even though it is not necessarily onto) since we may get a one-sorted model with domain equal to the union of all the sorts modulo the equivalence relation generated by $u \sim v \Leftrightarrow \tau(u) = v$. (This trick was used in the last part of the proof of the main theorem on Kaye [1991], see the last paragraph on page 463 and the claim on page 464 for details.)

My main object here was to convince the reader that axiomatizations of 'reasonable' subtheories ΓNF can be obtained. Unfortunately, with the exception of that in theorem 3.1, none of the ones the author has managed to obtain is particularly elegant, and this is rather disappointing. With this in mind, I will close with one last problem which I thought about while writing this paper, but was unable to solve.

PROBLEM 4.3 Let S be the theory consisting of all stratified axioms of the theory T in theorem 3.1. Does S prove all stratified $\exists_3^+$ consequences of NF?

References

[1] Boffa, M. [1977] The consistency problem for NF. J.S.L., **42**, pp. 215-20.

[2] Forster, T.E. [1987] Term models for weak set theories with a universal set. J.S.L., **52**, pp 374–387.

[3] Forster, T.E. [1992] *Set Theory with a Universal Set.* Oxford Logic Guides **20**, O.U.P.

[4] Forster, T.E. and Kaye, R.W. [1991] End extensions preserving power set. J.S.L., **56**, pp. 323–9.

[5] Hailperin, T. [1944] A set of axioms for logic. J.S.L., **9**, pp. 1–19.

[6] Hinnion, R. [1972] Sur les modèles de NF. Comptes Rendus, Académie des Sciences de Paris, série A, **275**, p. 567.

[7] Kaye, R.W. [1987] Parameter-free induction in arithmetic. In: *Proceedings of the 5th Easter conference on model theory, April 20–25, 1987.* Sektion Mathematik der Humboldt-Universität zu Berlin. pp. 70–81.

[8] Kaye, R.W. [1989] A contribution to the consistency problem for NF. *(Abstract)* J.S.L., **54**, pp. 657–8.

[9] Kaye, R.W. [1991] A generalization of Specker's theorem on typical ambiguity. J.S.L., **56**, pp. 458–66.

[10] Kaye, R.W, Paris, J.B. and Dimitracopoulos, C. [1988] On parameter free induction schemas. J.S.L., **53**, pp. 1082–97.

[11] Kreinovič, V.Ja. and Oswald, U. [1982] A decision method for the universal theorems of Quine's New Foundations. Z.M.L., **28**, pp. 181–7.

[12] Specker, E.P. [1962] Typical ambiguity. In: *Logic Methodology and Philosophy of Science*, eds. E. Nagel *et al.* Stanford University Press, pp 116–24.

School of Mathematics and Statistics,
The University of Birmingham,
Edgbaston,
Birmingham,
U.K.
Email: R.W.Kaye@bham.ac.uk