ZFJ AND THE CONSISTENCY PROBLEM FOR NF

M. Boffa

Quine's system NF has turned fifty and its consistency relatively to a classical set theory (like ZF) is still an open problem. However, the basic works of Specker, Jensen and Grishin have led to a set of partial results showing that various subsystems of NF are consistent. These results are surveyed in [1] [2] [3], but I want to give here a new approach based on ZFJ (Zermelo-Fraenkel set theory with an automorphism J).

First, what is NF ? Formally speaking, it is the one-sorted theory generated by all proper axioms of the (many-sorted) theory of types TT based on axioms of extensionality and comprehension. In other words, NF is what TT becomes when only one domain plays the role of the different type-levels. The goal is to avoid the inconvenient that in TT the notions are duplicated infinitely many times. But this improvement does not work without some negative effects, for example : in NF, the universal set V (the set of all sets) is not cantorian (this means that it is not equipollent to the set of singletons of its elements). Moreover, Specker [11] showed that this set has no well-ordering, but this has the happy consequence that V is infinite (so that Frege's Arithmetic works in NF). Note that NF has the same language as ZF and strictly contains the language of TT (the stratified formulas). Now we have all we need about NF for our discussion (for people who are unfamiliar with NF, I recommend Quine's original paper [9], reprinted with additions in [10], and chapter III of [5]).

Let ZFJ be obtained from ZF (Zermelo-Fraenkel set theory without axiom of choice) by adding a unary function symbol J and the following axioms claiming that J is an automorphism :

$(\forall xy)(x=y \leftrightarrow Jx=Jy)$, $(\forall xy)(x\in y \leftrightarrow Jx\in Jy)$, $(\forall x)(\exists y)(x=Jy)$.

Since we will need J distinct from the identity, we have not added the comprehension and replacement axioms containing J (so J must be handled with care; for example : $\{x\in y \mid x\in Jy\}$ is a set, but $\{x\in y \mid x\in Jx\}$ is not). Let us first show in detail how to get a quick proof of Jensen's result [7] by using ZFJ (other simplified proofs already appeared in [1][4]). The result is that NFU is consistent (NFU is the one-sorted theory generated by TTU, the fragment of TT obtained by restricting the extensionality property to the nonempty sets). To each stratified formula $\varphi(x,y,\ldots)$ (of the language of ZF) let us associate two formulas $\varphi_1(u,x,y,\ldots)$, $\varphi_2(u,x,y,\ldots)$ (of the language of ZFJ) as follows (u is a variable not contained in φ) :

(1) φ_1 is obtained from φ by (simultaneously) replacing each quantifier (Qz) by $(Qz\in u)$ and each subformula of the form $(v\in w)$ by $(v\in Jw\subset u)$;

(2) φ_2 is obtained from φ by (simultaneously) replacing each quantifier (Qz) by $(Qz \in J^r u)$ where r is the type of z and each subformula of the form $(v \in w)$ by $(v \in w \subset J^s u)$ where s is the type of v.

A straightforward induction on φ shows that

$$\varphi_1(u,x,y,\ldots) \leftrightarrow \varphi_2(u,J^m x,J^n y,\ldots)$$

holds in ZFJ, where $m,n,\ldots$ are the types of $x,y,\ldots$.
In particular, $\sigma_1(u) \leftrightarrow \sigma_2(u)$ holds for each stratified sentence σ. Note that if we define $x \in_z y$ by $x \in y \subset z$, then $\sigma_1(u)$ means that σ is satisfied in the one-sorted "structure" $\langle u, \in_u J\rangle$ ("structure" because the graph of $\in_u J$ is not a set) and $\sigma_2(u)$ means that σ is satisfied in the many-sorted structure $\langle u, Ju, \ldots, J^n u; \in_u, \in_{Ju}, \ldots, \in_{J^{n-1}u}\rangle$ where n is the largest type appearing in σ. Now, if Pu(the power set of u)$\subset$Ju, then $P(Ju) \subset J^2u$, $P(J^2u) \subset J^3u, \ldots$, so that this many-sorted structure satisfies each axiom of TTU with types $\leqslant n$. So it is clear that $Pu \subset Ju \to \sigma_2(u)$ holds in ZFJ for each axiom σ of TTU, thus $Pu \subset Ju \to \sigma_1(u)$ holds for each axiom σ of NFU and so for each theorem σ of NFU (note that φ_1 can be defined for any φ, stratified or not). This gives an interpretation of NFU in ZFJ + $(\exists u)(Pu \subset Ju)$ i.e. in ZFJ + $(\exists \text{ ordinal } \alpha)(\alpha < J\alpha)$, which is consistent (relatively to ZF) by the Ehrenfeucht-Mostowski theorem (see [8]). And by choosing u finite, or infinite, or well-ordered, ..., we get the consistency of NFU + (V is finite), NFU + (V is infinite), NFU + (V is well-ordered), ... The consistency of NFU + axiom of counting (every finite set is cantorian), first proved in [6], simply follows from the consistency of ZFJ + $(\exists\alpha)(\alpha<J\alpha)$ + $(\forall n \in \omega)(Jn=n)$,

itself a consequence of the Ehrenfeucht-Mostowski theorem for ω-logic (see [8]).

A set u with $Pu=Ju$ would provide a consistency proof for NF; unfortunately such a set cannot exist, since $Pu=Ju$ entails $n+1=Jn$ (and so odd=even, an absurdity) where n is the largest natural number such that u is of the form $P^n v$. But the weaker assumption that Pu and Ju are equipollent sets (i.e. $Pu\cong Ju$) seems free of direct contradiction, and we will show how a bijective $h:Ju\to Pu$ induces an interpretation of NF. For each stratified formula $\varphi(x,y,\ldots)$ let us define $\varphi_1(u,h,x,y,\ldots)$ and $\varphi_2(u,x,y,\ldots)$ as follows :

(1) φ_1 is obtained from φ by replacing each quantifier (Qz) by $(Qz\in u)$ and each subformula of the form $(v\in w)$ by $(v\in hJw)$;

(2) φ_2 is obtained from φ by replacing each quantifier (Qz) by $(Qz\in P^r u)$ where r is the type of z.

By induction on φ, the following equivalence holds in ZFJ for any bijective map $h:Ju\to Pu$ and any $x,y,\ldots \in u$:

$$\varphi_1(u,h,x,y,\ldots) \leftrightarrow \varphi_2(u,h_m J\ldots h_2 J h_1 Jx, h_n J\ldots h_2 J h_1 Jy,\ldots)$$

where $m,n,\ldots$ are the types of $x,y,\ldots$, $h_1=h$ and $h_{i+1}(v) = \{h_i(w)\,|\,w\in v\}$. This implies that $\sigma_1(u,h)$ holds for each axiom σ of NF. So we have an interpretation of NF in $ZFJ + (\exists u)(\exists \text{ bijection } h:Ju\to Pu)$ i.e. $ZFJ + (\exists \text{ cardinal } c)(Jc=2^c)$ and I like to conjecture that this system is consistent.

Note that a c satisfying $Jc=2^c$ must be similar to the cardinal of the universal set V of NF; for example, it must be infinite but not an aleph.

I hope to have shown that the automorphisms of ZF provide a useful method for the problem of the consistency of NF.

References

[1] Boffa, *The consistency problem for NF*, J.S.L. 42 (1977), 215-220.

[2] ———, *La théorie des types et NF*, Bull. Soc. Math. Belg. A33 (1981), 21-31.

[3] ———, *The point on Quine's NF (with a Bibliography)*, Teoria IV/2 (1984), 3-13.

[4] Crabbé, *NF with Urelemente* (abstract), Tagungsbericht "New Foundations", Mathematisches Forschungsinstitut Oberwolfach, march 1987.

[5] Fraenkel, Bar-Hillel & Levy, *Foundations of set theory*, North-Holland (1973).

[6] Grishin, *An investigation of some versions of Quine's system*, Nauchno-Tekhnicheskaya Informatsiya (ser.2) 5 (1973), 34-37.

[7] Jensen, *On the consistency of a slight (?) modification of Quine's "New Foundations"*, Synthese 19 (1968/69), 250-263.

[8] Morley, *Homogenous sets* , Handbook of Mathematical Logic (Barwise, ed.), North-Holland (1977), 181-196.

[9] Quine, *New foundations for mathematical logic*, Amer. Math. Monthly 44 (1937), 70-80.

[10] ———, *From a logical point of view*, Harvard Univ. Press (1953).

[11] Specker, *The axiom of choice in Quine's New Foundations for mathematical logic*, Proc. Nat. Acad. Sc. U.S.A. 39 (1953), 972-975.

Faculté des Sciences
15, Avenue Maistriau
7000 MONS
(Belgium)