Stratification

Thomas Forster

June 4, 2025

Initially at least our language is $\mathcal{L}(\in, =)$. We might consider adding terms too – but later.

On the face of it nothing could be more straightforward than the idea of a stratifiable formula. It appeals to natural type-theoretic intuitions, but there are details that need to be carefully curated.

A *stratification* of a formula is in the first instance a function defined on the occurrences of variables in that formula, taking values in $\mathbb{Z}$.

It must give all occurrences of any one variable the same label. (We might relax this condition for *free* variables later; stay tuned.) A variable might be sitting next to a quantifier: any such occurrence must be decorated by the same integer as decorates all the occurrences of that variable that are bound by that quantifier.

We have to distinguish between *variables* and *letters*. This need arises because the wellformation rules of $\mathcal{L}(\in, =)$ admit formulæ such as $(\forall x)(x = x) \wedge (\forall x)(x \notin x)$. We want to say that although there is only one *letter* here, there are nevertheless two *variables*. Variables and letters have different identity criteria! Sameness-as-variables is a finer equivalence relation than sameness-as-letters. Two occurrences of a letter within a formula are occurences of the same variable if

(i) both occurrences are free; *or*
(ii) the two occurrences are bound by the one quantifier.

We need the subtleties in the preceding paragraph because of the need to define stratification even for formulæ that re-use bound variables. Quine saw all this and discusses it in section 3 of [3].

> $\mathcal{L}(\in, =)$ is a context-free language. In contrast we can use the pumping lemma to show that the fragment of $\mathcal{L}(\in, =)$ consisting of wffs that do *not* reuse bound variables is *not* context-free.

Then we say that if the formula contains a subformula $x \in y$ then the label given to 'x' by any stratification for that formula must be one less than the label given to 'y' and if it contains a subformula $x = y$ then the two variables must receive the same label. This condition extends seamlessly to formulæ containing terms (such as singular descriptions or – more specifically – set abstracts).

That completes the definition of a stratification.

A formula is *stratifiable* if it can be given a stratification.

> It is clear that composing a stratification on the left with a translation of $\mathbb{Z}$ results in another stratification, so in some circumstances it might be helpful to think of

stratifications *up to translation.* But formulæ with "disconnected" variables – for example $x \in y \wedge u \in v$ – obviously admit stratifications that are not equivalent up to translation. Generally these subtleties will not concern us.

A *stratified formula* is a formula decorated by a stratification.

This piece of nomenclature sounds pointless and anal-retentive, and it has to be admitted that it's only relatively recently that this distinction has arisen in the literature – hitherto the expression 'stratified formula' always covered both, and the word 'stratifiable' does not appear in the early literature. However the distinction does matter. We can think of a stratified formula as an expression in the language of strongly typed set theory, where each variable has an integer tattooed onto its midriff. The class of stratifi*able* formulæ is not a context-free language, whereas the class of stratif*ied* formulæ *is* context-free. See [7]. Also suppose we were interested in unification for stratifiable formulæ of $\mathcal{L}(\in, =)$ (a topic we are not going to touch on here!) one is not surprised to learn that two stratifiable formulæ can have a most general unifier ("mgu") that is not stratifiable (think $x \in y$ and $y \in x$) but one might expect that an mgu of two stratified formulæ should be stratified.

Thus we regard $(\exists y)(y \in x) \vee (\exists y)(x \in y)$ – which reuses the letter 'y' – as stratifiable, because when we give the two occurrences of 'x' the same label (as we must) – say 1 – we can give the first two occurrences of 'y' the label 0 and give the others the label 2. If we felt that the two 'y' variables were the same *variable* as well as the same *letter* we would regard $(\exists y)(y \in x) \vee (\exists y)(x \in y)$ as unstratifiable. This would clearly be the wrong way to go, since this formula is logically equivalent to – is an *alphabetic variant of*, indeed – the straightforwardly stratifiable formula $(\exists y)(y \in x) \vee (\exists z)(x \in z)$

0.0.1 Logical Equivalence does not preserve Stratification

This last example can also be used to illustrate how alphabetical variance preserves stratification but Logical equivalence doesn't. Although $(\exists y)(y \in x) \vee (\exists y)(x \in y)$ is stratified, it is logically equivalent to $(\exists y)(y \in x \vee x \in y)$ – which isn't. Dually for $\forall$.

Given that, we shouldn't be surprised that
(i) the rule of $\exists$-int can give a stratified conclusion from an unstratified premise, and
(ii) the rule of $\forall$-elim can give an unstratified conclusion from a stratified premise.

Ad (i) we can infer $\exists x\phi(x)$ from $\phi(t)$. The term t might be complex and unstratified and ϕ could be stratified. *Ad* (ii) we can infer $\phi(t)$ from $\forall x\phi(x)$ similarly.

1 Weak Stratification

The possibility of a concept of weak stratification was foreshadowed in Quine [4] p 78 and [3] section 4 and section 5, and discussed in Rosser [**?**]. However the idea was first isolated and its significance explained by Crabbé [1].

NF as set up in [4] has **stratifiable comprehension**:

$$(\forall \vec{x})(\exists y)(\forall z)(z \in y \longleftrightarrow \phi(z, \vec{x}))$$

is an axiom of NF as long as

(i) 'y' is not free in ϕ; and
(ii) '$(\forall z)(z \in y \longleftrightarrow \phi(z, \vec{x}))$' is stratifiable.

$\phi(z, \vec{x})$ is the *eigenformula* of the comprehension axiom and 'z' is the *eigenvariable*.

It was clear from early on that there are – shall we say – *subtleties* to do with proving the existence of $x \cup \{x\}$. Quine comments on this in [5].

If you have only stratifiable comprehension, you can use it to prove the existence of $x \cup \{y\}$, and you then do a $\forall$-elimination to instantiate the 'y' to 'x' – and then do a $\forall$-int to get the existence of all $x \cup \{x\}$. Thus stratified comprehension can prove some instances of unstratified comprehension. The trouble is that the proof has a cut in the middle – you do a $\forall$-int and then follow it immediately with a $\forall$-elim. Admittedly it's over-optimistic to expect NF to admit cut-elimination, but this particular cut-proof looks like one that one could do without.

The idea is to weaken our definition of *stratified* so that

$$(\forall x)(\exists y)(\forall z)(z \in y \longleftrightarrow z = x \vee z \in x)$$

is stratified in the new, weak, sense. So we say that a formula is *weakly stratifiable* if there is a stratification defined on its *bound variables*.

So nowadays we say that

$$(\forall \vec{x})(\exists y)(\forall z)(z \in y \longleftrightarrow \phi(z, \vec{x}))$$

is a (comprehension) axiom of NF as long as

(i) 'y' is not free in ϕ; and
(ii) '$(\forall z)(z \in y \longleftrightarrow \phi(z, \vec{x}))$' is weakly stratifiable.

This is *weakly stratifiable comprehension*.

Notice that condition (ii) is **not** that the eigenformula $\phi(z, \vec{x})$) should be weakly stratifiable – which is what *weakly stratified comprehension* might suggest to incautious readers; we have to ensure that the eigenvariable gets stratified, even tho' it is free in the eigenformula. (We do not have to stipulate that 'y' gets stratified: condition (i) ensures that there is only one occurrence of 'y' so there can be no type clash on 'y'.)

So if we have weakly stratifiable comprehension then

$$(\forall x)(\exists y)(\forall z)(z \in y \longleftrightarrow z = x \vee z \in x)$$

is an axiom.

Randall says that according to the definition of stratification in [3] this is an axiom of NF.

Thus it comes about that weakly stratifiable comprehension enables us to have normal proofs of existence of (e.g.) $x \cup \{x\}$. This treatment is due to Crabbé [1].

Brief reality check: stratifiable comprehension implies weakly stratifiable comprehension.

1.1 Weak stratification and the subformula relation

The class of weakly stratifiable formulæ is closed under the subformula relation. This is in sharp contrast to the class of stratifiable formulæ which isn't. '$x \in y$' is stratifiable but its substitution instance '$x \in x$' is not. And a substitution instance is a subformula.

There is quite a lot that can be said about this, but here-and-now might not be the place or time for saying it.

1.2 Odds and ends

Are the Croats saying anything helpful in [2]?
Globally stratifiable proofs?
k-stratifiable; stratification-mod-*k*.

References

[1] Marcel Crabbé, "Types Ambigus" C. R. Acad. Sc. Faris, t. 280 (6 janvier 1975) Série A pp 1–2

[2] Tin Adlešić and Vedran Čačić "A Modern Rigorous Approach to Stratification in NF/NFU" Logica Universalis **16**(3), pp. 451–468 (2022)

[3] Quine: "On the theory of Types" JSL **3** (1938) pp 125–139

[4] Quine New Foundations for Mathematical Logic

[5] Quine Set Theory and its logic

[6] J Barklay Rosser "The Burali-Forti Paradox" JSL

[7] Ryan-Smith, Calliope. "Stratifiable Formulae are not Context-Free" [2023] `https://arxiv.org/abs/2304.10291`