Classless

Sam Roberts

1 Introduction

Classes are a kind of collection. Typically, they are too large to be sets. For example, according to standard theories of classes, there are classes containing absolutely all sets even though there is no set of all sets.[1] But what are classes, if not sets? As Boolos puts it:

> Wait a minute! I thought that set theory was supposed to be a theory about all, 'absolutely' all, the collections that there were and that 'set' was synonymous with 'collection'.[2] (Boolos, 1998, p.36)

It turns out that when our theory of classes is relatively weak, this question can be avoided. In particular, it is well-known that von-Neuman-Bernays-Gödel class theory (NBG) is *conservative* over the standard axioms of set theory (namely, those of Zermelo-Fraenkel set theory with the axiom of Choice (ZFC)):[3] anything NGB proves about the sets is already provable in ZFC. Thus, if all we assume about classes is that they satisfy the axioms of NBG, they can be treated as a convenient but dispensable fiction. For stronger class theories, however, conservativity can fail and it looks like classes have to be taken seriously.[4]

In this note I will prove a new conservativity result for a broad range of class theories (the Main theorem). It tells us that as long as our set theory T contains an independently well-motivated *reflection principle*, anything provable about the sets in *any* reasonable class theory extending T is already provable in T itself. Thus, assuming the reflection principle is true, classes can be treated as a convenient but dispensable fiction in a much broader range of cases than was previously thought.

2 Reflection and the cumulative hierarchy

It is well-known that the sets are organised into a cumulative hierarchy of levels, with one level for each ordinal. The first level contains no sets whatsoever. It is just the empty set, and we denote it "V_0". Then, given a level V_α, the very next level $V_{\alpha+1}$ contains all and only the subsets of V_α. In other words, $V_{\alpha+1}$ is the *powerset* of V_α. Formally: $V_{\alpha+1} = \mathcal{P}(V_\alpha)$. Finally, when λ is a limit ordinal – that is, an ordinal with no immediate predecessor – the corresponding level V_λ collects together all the sets from previous levels. In other words, V_λ is the *union* of previous levels. Formally, $V_\lambda = \bigcup_{\alpha<\lambda} V_\alpha$. The universe of absolutely all sets, V, is then the union of all levels, of all the V_αs.

[1] If there were a set of all sets, there would be a set of all sets that are not members of themselves by the Separation axiom. But that is impossible: such a set would have to both be a member of itself and also not a member of itself.

[2] See (Hellman, 1989, pp.44-45) for some further worries about classes, and Horsten (2018) section 5.1 for general discussion.

[3] See Kunen (2011) for definitions and details of any unexplained set theoretic notions used in this note.

[4] Usually, such theories prove the consistency statement for ZFC, which is unprovable in ZFC itself by Gödel's second incompleteness theorem. This is true, for example, for Morse-Kelley class theory (discussed below).

Some levels resemble the universe more closely than others. For example, if λ is a limit ordinal, V_λ is a natural model of all the axioms of ZFC minus the axiom schema of Replacement. And when α is a so-called *inaccessible cardinal*, V_α is a natural model of *all* the axioms of ZFC.[5] Call such V_α *inaccessible levels*.

How far does the cumulative hierarchy extend? *Reflection principles* say it extends so far that there are levels which resemble (or reflect) the whole hierarchy in various ways. For example, perhaps the simplest such principle says that any claim true in the universe is also true in some level.[6] Formally:

$$(\mathsf{PR}_S) \qquad \varphi \rightarrow \exists \alpha \varphi^{V_\alpha}$$

where φ is a sentence in the language of first-order set theory, $\mathcal{L}_\in$,[7] and φ^{V_α} is the result of restricting the quantifiers in φ to V_α.[8] Many take reflection principles to provide a way of justifying some of the more non-trivial axioms of ZFC (like Infinity and Replacement), and of justifying some new axioms (in particular, large cardinal axioms).[9]

Given that inaccessible levels are natural models of ZFC and ZFC is true in the universe of sets, a natural strengthening of PR_S says that the hierarchy extends so far that any claim true in the whole universe is also true in some inaccessible level. Formally:

$$(\mathsf{IPR}_S) \qquad \varphi \rightarrow \exists \alpha (\mathsf{In}(\alpha) \wedge \varphi^{V_\alpha})$$

where $\mathsf{In}(\alpha)$ expresses that α is an inaccessible cardinal and φ and φ^{V_α} are as before.

Inaccessible levels also provide natural models for class theory. The language of class theory, $\mathcal{L}^2_\in$, extends the first-order language of set theory, $\mathcal{L}_\in$, with a stock of second-order variables $X, Y, Z, ...$ etc (intended to range over classes) and takes "$x \in X$" and "$X = Y$" to be well-formed. We read "$x \in X$" as "x is an element/member of X", just as we do with sets. Typically, class theories start with the axioms of ZFC and swap its schemas of Replacement and Separation for their class-theoretic versions.[10] They then add an axiom of Extensionality – which says that classes with the same members are identical – and a principle of Global Well-ordering – which says that there is a class that codes a well-order of the universe of sets. They distinguish themselves primarily by which instances of the following *comprehension* schema they adopt. It says, for a particular condition, that there is class of all and only the things satisfying that condition. Formally:

$$(\mathsf{comp}) \qquad \exists X \forall x (x \in X \leftrightarrow \varphi)$$

[5] See, for example, (Maddy, 1988, p. 504) and (Drake, 1974, p.110).

[6] See the appendix for a discussion of the relation between this and other reflection principles.

[7] $\mathcal{L}_\in$ has, in addition to the usual resources of first-order logic, a single non-logical relation $\in$, intended to express membership. So, "$x \in y$" is read "x is a member/element of y".

[8] More precisely, φ^{V_α} is the result of replacing occurrences of "$\exists x$" in φ with "$\exists x \in V_\alpha$".

[9] See Koellner (2009) and (Maddy, 1988, p.503-504) for discussion.

[10] In particular, the axiom schema of Separation is swapped for:

$$(\text{Separation}_2) \qquad \forall X \forall x \exists y \forall z (z \in y \leftrightarrow z \in x \wedge z \in X)$$

and Replacement for:

$$(\text{Replacement}_2) \qquad \forall X (Fun(X) \rightarrow \forall x \exists y \forall z (z \in y \leftrightarrow \exists w \in x (\langle w, z \rangle \in X)))$$

where $Fun(X)$ abbreviates the claim that X is a class coding a function.

where φ is a formula in the language of class theory, $\mathcal{L}^2_\in$, without "X" free. When we add instances of comp for formulas without class quantifiers to the above principles, we get von-Neuman-Bernays-Gödel class theory (NBG). This theory is conservative over ZFC: whatever NBG proves in $\mathcal{L}_\in$, ZFC already proves (Felgner (1971)). As we add more and more instances of comp, we get stronger and stronger theories. For example, as soon as we add all instances for formulas with one class quantifier, we get a class theory that proves the consistency of ZFC and thus of NBG (Mostowski (1951)). When we add *all* instances of comp for formulas in $\mathcal{L}^2_\in$, we get Morse-Kelley class theory (MK). In this sense, MK is the strongest class theory.[11]

It is a standard result that inaccessible levels satisfy the theorems of MK when classes are interpreted as ranging over their subsets.[12] This fact can be used to prove the following conservativity result (see the appendix for the proof).

Main theorem. *Let* T *be any theory in* $\mathcal{L}_\in$ *which includes* ZFC + IPR_S*, and let* MK + T *be the theory in* $\mathcal{L}^2_\in$ *consisting of the axioms of* MK *together with those of* T*. Then* MK + T *is conservative over* T *for* $\mathcal{L}_\in$*: any* $\mathcal{L}_\in$ *sentence provable in* MK + T *is already provable in* T*. Thus,* MK + T *is consistent just in case* T *is.*

This means that if we are already committed to T – crucially, if we are already committed to the reflection principle IPR_S – then we can avail ourselves of MK without incurring a commitment to classes: they can be treated as a convenient but dispensable fiction.

3 Justifying IPR_S

The main theorem can only be used to eliminate classes if there is good reason to believe the reflection principle IPR_S. And, as I've mentioned, many take reflection principles like IPR_S to be well-motivated additions to the axioms of ZFC. Nonetheless, I want to end this note by briefly outlining one potential worry for IPR_S and a possible response.

Above, I mentioned that inaccessible levels of the cumulative hierarchy are natural models of ZFC. But it turns out that inaccessible levels are not the only levels of the cumulative hierarchy satisfying ZFC. Indeed, it is a standard result that below *any* inaccessible level, there are arbitrarily large levels satisfying ZFC (Drake, 1974, p.113). Moreover, the only levels satisfying MK are inaccessible levels (Drake, 1974, p.112). It follows that many levels satisfying ZFC do not satisfy MK. Consequently, the Main theorem does not go through when we replace IPR_S with the following weaker principle.[13]

(IPR^*_S) $$\varphi \to \exists\alpha(V_\alpha \vDash \mathsf{ZFC}+\varphi)$$

[11] Recently, class choice principles that go beyond MK have been investigated. See, for example, Hamkins et al. (2016). Since these principles are also true in any inaccessible level, the Main theorem covers class theories containing them.

[12] More precisely, ZFC proves:

$$\forall\alpha(\mathsf{In}(\alpha) \to \varphi^{V_\alpha})$$

where φ is a theorem of MK and φ^{V_α} is the result of replacing occurrences of "$\exists x$" in φ with "$\exists x \in V_\alpha$" and occurrences of "$\exists X\psi(X)$" with "$\exists y \subseteq V_\alpha\psi(y)$". See (Drake, 1974, p.112).

[13] In particular, MK proves the consistency statement for ZFC + IRP^*_S and is thus not conservative over that theory. The reason is that MK proves both the existence of a satisfaction class for the language of set theory and that relative to this class there are arbitrarily large levels of the cumulative hierarchy that are *elementary substructures* of the whole hierarchy. See Mostowski (1951). It follows that every such level satisfies the axioms of ZFC together with IPR^*_S.

The worry is that it looks like the motivation I gave earlier for IPR_S only extends to the weaker IPR^*_S.

Let me outline one promising way to bridge the gap between these two principles. Some prominent mathematicians and philosophers have argued that we should think of the axiom schemas of Separation and Replacement in an *open-ended* way: that they hold not only for formulas in the language of set theory, but also for formulas in any extension of that language.[14] Thought of in this way, it looks like the only levels which satisfy the axioms of Separation and Replacement are those levels satisfying their class-theoretic versions, Separation$_2$ and Replacement$_2$. (see footnote 10). To see this, suppose that Replacement$_2$ is false in V_α for some function f over V_α. Then, we can expand $\mathcal{L}_\in$ with a predicate F for f, and the corresponding instance of Replacement in this language would be false. Thus, Replacement would not hold in an open-ended way in V_α. In other words, if Replacement is open-endedly true in V_α, then Replacement$_2$ is true in V_α simpliciter. Similarly, for Separation. Thus, if the axioms of ZFC are open-endedly true in V_α, then V_α will be an inaccessible level.

Putting all of this together, the thought would be that if the axioms of ZFC are open-endedly true, then a natural strengthening of PR_S says that the hierarchy extends so far that any claim true in the whole universe is also true in some level in which the axioms of ZFC are open-endedly true. They are precisely the levels in which the axioms of ZFC together with Separation$_2$ and Replacement$_2$ are true. And those, in turn, are precisely the inaccessible levels (Drake, 1974, p.112). That gives us IPR_S.

4 Technical appendix

4.1 Proof of the main theorem

Proof. Suppose $\mathsf{MK} + \mathsf{T}$ proves φ from $\psi_0, ..., \psi_n \in \mathsf{MK}$ and $\chi_0, ..., \chi_m \in \mathsf{T}$. Since T extends ZFC, it will prove that every inaccessible level satisfies $\psi_0, ..., \psi_n$. Moreover, it will prove that any level satisfying $\psi_0, ..., \psi_n, \chi_0, ..., \chi_m$ also satisfies φ: in other words, it proves that levels preserve logic. So, T will prove that every inaccessible level satisfies:

(*) $$(\chi_0 \wedge ... \wedge \chi_m) \to \varphi$$

More precisely, T proves:

(**) $$\forall\alpha(In(\alpha) \to [(\chi_0 \wedge ... \wedge \chi_m) \to \varphi]^{V_\alpha}).$$

Since $\chi_0, .., \chi_m, \varphi \in \mathcal{L}_\in$, one instance of IPR_S will say that if (*) were false, it would be false in some inaccessible level. Formally:

(***) $$\neg[(\chi_0 \wedge ... \wedge \chi_m) \to \varphi] \to \exists\alpha(In(\alpha) \wedge (\neg[(\chi_0 \wedge ... \wedge \chi_m) \to \varphi])^{V_\alpha})$$

But we've just shown that (*) is *true* in *all* inaccessible levels. So, by (**) and (***), the antecedent of (***) must be false: that is, (*) must be true.[15] Thus, since T proves (**) and

[14] See Martin (2001), McGee (1997), Lavine (2006), and Parsons (1990).

[15] In general, it is easy to see that the reflection principle is equivalent to the principle which says that whenever φ is true in all inaccessible levels, it is true simpliciter. Formally:

$$\forall\alpha(In(\alpha) \to \varphi^{V_\alpha}) \to \varphi$$

(***), it proves (*). And since (*)'s antecedent is a conjunction of theorems of T, it follows that T also proves φ. □

It is worth noting that for the main theorem to go through, we actually only need that T includes IPR_S plus the resources to prove that inaccessible levels satisfy MK. Indeed, it is easy to see that the theorem generalises to arbitrary set theories T and class theories C on the assumption that T proves (for some Φ):

(1) $$\forall x(\Phi(x) \to \psi^x)$$

where ψ is an axiom of C, and:

(2) $$\chi \to \exists x(\Phi(x) \wedge \chi^x)$$

where χ is a sentence of $\mathcal{L}_\in$.

4.2 How do PR_S and IPR_S compare with other reflection principles?

The principle PR_S is the restriction to sentences of what is called the *partial reflection principle*. Formally:

(PR) $$\varphi \to \exists \alpha \varphi^{V_\alpha}$$

where φ is a formula in $\mathcal{L}_\in$ and φ^{V_α} is as before. Both PR_S and PR are extremely weak. In particular, let Z (for Zermelo set theory) be ZFC minus the axiom schema of Replacement. Then, $\mathsf{Z} + \mathsf{PR}$ proves that there is a V_α in which $\mathsf{Z} + \mathsf{PR}_S$ is true. Similarly, ZFC proves both PR and that there are V_α in which $\mathsf{Z} + \mathsf{PR}$ is true.

These partial reflection principles are to be contrasted with the *complete reflection principle*, which says that there are arbitrarily large levels of the cumulative hierarchy that are indistinguishable from the whole. Formally:

(CR) $$\forall \alpha \exists \beta > \alpha \forall \vec{x} \in V_\beta(\varphi \leftrightarrow \varphi^{V_\beta})$$

where φ is a formula in $\mathcal{L}_\in$ with free variables among $\vec{x}$. This principle is much stronger than PR. Indeed, it is equivalent to the axioms of Replacement and Infinity over the other axioms of ZFC.

We can similarly distinguish IPR_S from the version for arbitrary formulas in $\mathcal{L}_\in$ (IPR) and from the version of CR where the reflecting level is inaccessible (ICR). As before, each of these principles increases in strength. In particular, $\mathsf{Z} + \mathsf{IPR}$ proves that there is a V_α satisfying $\mathsf{Z} + \mathsf{IPR}_S$ and $\mathsf{Z} + \mathsf{ICR}$ proves that there is a V_α satisfying $\mathsf{Z} + \mathsf{IPR}$.[16] Since it didn't matter for my purposes which of these principles I used, I went for the simplest and weakest.

4.3 Model theoretic conservativity

The Main theorem is a *proof-theoretic* conservativity result: it tells us that what's provable in $\mathsf{MK}+\mathsf{T}$ is already provable in T. But there are stronger, *model-theoretic*, conservativity results. For example, it is well-known that any model of ZFC can be extended to a model of

[16]See Lévy (1960a), Lévy and Vaught (1961), and Lévy (1960b).

NBG minus Global Well-ordering with the very same sets.[17] It is natural to wonder if this stronger form of conservativity holds for MK + T and T when we drop Global Well-ordering.[18] It turns out that it does not. Indeed, the least V_α satisfying *any* theory T in the language of set theory cannot be extended to a model of MK minus Global Well-ordering with the same sets. In fact, it cannot be so extended to a model of MK minus Global Well-ordering and minus all instances of comp except those for formulas with one class quantifier (the so-called Π^1_1 instances) (Marek and Mostowski (1975)).[19]

References

Boolos, G. (1998). *Logic, Logic, and Logic*. Cambridge, MA: Harvard University Press. Edited by Richard Jeffrey.

Drake, F. R. (1974). *Set Theory: An Introduction to Large Cardinals*, Volume 76 of *Studies in Logic and the Foundations of Mathematics*. Amsterdam: North-Holland.

Felgner, U. (1971). Comparison of the axioms of local and universal choice. *Fundamenta Mathematicae 71*(1), 43–62.

Hamkins, J., V. Gitman, and T. Johnstone (Accessed 24th June 2016). Kelley-morse set theory and choice principles for classes. http://boolesrings.org/victoriagitman/files/2015/01/kelleymorse2.pdf.

Hellman, G. (1989). *Mathematics without Numbers: Towards a Modal-Structural Interpretation*. Oxford: Clarendon.

Horsten, L. (2018). Philosophy of mathematics. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Spring 2018 ed.). Metaphysics Research Lab, Stanford University.

Koellner, P. (2009). On Reflection Principles. *Annals of Pure and Applied Logic 157*(2-3), 206–219.

Kunen, K. (2011). *Set Theory*. College Publications.

Lavine, S. (2006). Something about everything: Universal quantification in the universal sense of universal quantification. In A. Rayo and G. Uzquiano (Eds.), *Absolute Generality*, pp. 98–148. Oxford University Press.

Lévy, A. (1960a). Axiom Schemata of Strong Infinity in Axiomatic Set Theory. *Pacific Journal of Mathematics 10*, 223–238.

Lévy, A. (1960b). Principles of Reflection in Axiomatic Set Theory. *Fundamenta Mathematicae 49*, 1–10.

Lévy, A. and R. Vaught (1961). Principles of Partial Reflection in the Set Theories of Zermelo and Ackermann. *Pacific Journal of Mathematics 11*, 1045–1062.

Maddy, P. (1988). Believing the Axioms. I. *The Journal of Symbolic Logic 53*(2), 481–511.

[17] Dropping Global Well-ordering is crucial here: some models of ZFC cannot be so extended to models of NBG. See Williams (ms) section 2.1 for discussion.

[18] Thanks to a referee at Analysis for raising this question.

[19] Thanks to Øystein Linnebo, Philip Welch, Kameryn Williams, and anonymous referees at Analysis for helpful comments and discussion.

Marek, W. and A. Mostowski (1975). On extendability of models of zf set theory to the models of kelley-morse theory of classes. In G. H. Müller, A. Oberschelp, and K. Potthoff (Eds.), *?ISILC Logic Conference*, Berlin, Heidelberg, pp. 460–542. Springer Berlin Heidelberg.

Martin, D. A. (2001). Multiple universes of sets and indeterminate truth values. *Topoi 20*(1), 5–16.

McGee, V. (1997). How We Learn Mathematical Language. *Philosophical Review*, 35–68.

Mostowski, A. (1951). Some impredicative definitions in the axiomatic set-theory. *Journal of Symbolic Logic 16*(4), 274–275.

Parsons, C. (1990). The uniqueness of the natural numbers. *Iyyun: The Jerusalem Philosophical Quarterly 39*, 13–44.

Williams, K. (ms). Least models of second-order set theories. manuscript under review.