The Limits of *The Limits of Abstraction*

A Review of Kit Fine's *The Limits of Abstraction*, 2002, Oxford, Oxford University Press.

Summary of Contents

Kit Fine's recent *The Limits of Abstraction*, an extended version of his [1998] paper, contains four chapters. The first two deal with philosophical aspects of abstraction, while the latter two provide the formal framework for a general theory of abstraction. Here we focus on the philosophical material, and, given the depth and complexity of the material, cannot attempt even a superficial summary of the entirety of the philosophical material. Thus, we content ourselves with sketching the contents. Two particular issues will then be examined in greater detail.

A theory of abstraction is, broadly construed, any account that reconstructs mathematical theories using second-order abstraction principles of the form:

$$\S xFx = \S xGx \leftrightarrow E(F, G)$$

(We ignore first-order abstraction principles such as Frege's direction abstraction.) The function ($\S$) occurring on the left-hand side of the biconditional provides objects corresponding to the concepts serving as the arguments of the function, and the identity conditions for the objects so introduced (the abstracts) are given by the equivalence relation E. The most notorious instance of a second-order abstraction principle is Frege's *Basic Law V*:

$$BLV: \quad \S xFx = \S xGx \leftrightarrow (\forall x)(Fx \leftrightarrow Gx)$$

which, infamously, is inconsistent.

There has been a resurgence of interest in abstractionism. Neo-Fregeans have abandoned *BLV* in favour of (presumably consistent) abstraction principles, each delivering a particular mathematical theory. For example, *Hume's Principle*:

$$Nx:Fx = Nx:Gx \leftrightarrow F \approx G$$

[$F \approx G$ abbreviates the second-order formula asserting the equinumerousity of F and G]

allows one to reconstruct second-order *Peano arithmetic*. Examining how far such a piecemeal approach to foundations can take us is currently an active and fruitful research project. Fine's approach is different. Instead of worrying over the acceptability of particular abstraction principles, or formulating principles that allow for the reconstruction of his favourite mathematical theories, Fine provides a general account of abstraction.

The first chapter, "Philosophical Introduction", is less an introduction than a dense examination of philosophical difficulties facing abstractionism, including worries about how to view the set theory used in the study of abstraction principles. Another issue discussed –

necessary and sufficient conditions for the identity of abstracts arising from distinct abstraction principles – is a variation of the traditional *Caesar Problem*. Fine also argues against the idea that abstraction principles can be understood in terms of reconceptualization, an argument we return to in the next section.

The second chapter (the principal addition to the original [1998] paper) concentrates on the use of Frege's *Context Principle* (*CP*) as the underlying justification of abstraction principles. Fine argues that *CP* is incapable of providing a "vindication" of abstraction principles thought of as contextual definitions and, furthermore, it conflicts with his proposed solution to the *Caesar Problem*. He also provides insightful discussions of interrelated issues, including referential determinacy and predicativity. Fine's interpretation of the *Context Principle* will be the topic of the third section of this review.

Next, Fine examines the proof theory and semantics of abstraction principles, thoroughly examining the conditions under which a particular abstraction principle is, or a collection of abstraction principles are, satisfiable, and determining the general conditions under which an abstraction principle is inflationary on a domain. Of particular interest is Fine's notion of hyperinflation – the idea that (with suitable constraints in place regarding the identity of abstracts arising from distinct abstraction principles) a collection of abstraction principles might together imply the existence of 'too many' objects even though no one of the principles does so individually.

In the final chapter, Fine formulates a general theory of abstraction, i.e. a formal theory describing the domain(s) containing all abstracts obtained through the application of acceptable abstraction principles. At the end he sketches how one can reconstruct arithmetic, analysis, and set theory within such a system and proposes some directions for future work.

Reconceptualization

In Chapter 1 Fine considers the idea that abstraction principles be understood as instances of reconceptualization. Unlike traditional implicit definition, and unlike applications of the *Context Principle* (on which more below), definition by reconceptualization centers, not on reference, but on sense – as Fine puts it, this approach "rests on the idea that new senses may emerge from the reanalysis of a given sense" (p. 35)

Ignoring problems with consistency, we can view *BLV* as stipulating that the sense of the (open) sentence on the right-hand side:

(A) $(\forall x)(Fx \leftrightarrow Gx)$

is identical to the sense of the (open) sentence on the left-hand side:

(B) $\S xFx = \S xGx$

The 'recarving' of content involved in moving from (A) to (B) is taken to endow the extension operator (§) with a sense. Fine flags a number of issues crucial to the success of definition by reconceptualization:

> Does the method work? Various questions should be distinguished here. Can we obtain a unique sense for the number operator in this way? Can we obtain at least one sense, even if not a unique sense? And, in each case, what is required by way of justifying the definition? (p. 38)

Fine argues that reconceptualization fails to deliver at least part of the goods, i.e. it fails to provide a unique sense for the operators in question. Since Fine's argument for this claim is difficult, it is worth providing in full:

> Let us adopt the following abbreviations:
>   (A) $(\forall x)(Fx \leftrightarrow Gx)$
>   (B) $\S xFx = \S xGx$
>   (C) $(\forall x)(x = \S xFx \leftrightarrow x = \S xGx)$
>   (D) $\S x(x = \S xFx) = \S x(x = \S xGx)$
>
> Then on the view under consideration, (B) should have the same sense as (A) and (D) the same sense as (C). Now presumably (C) should also have the same sense as (B). Indeed, it is hard to see how a criterion of synonymy could be so tolerant as to allow the synonymy of (B) to (A) (or of (D) to (C)) and yet not tolerant enough to forbid [sic] the synonymy of (C) to (B). If this is right, it follows that <u>(A)</u> has the same sense as (D). But then <u>if the sense of the operator that is given through reconceptualization is unique</u>, the sense of $\S xFx$ will be the same as the sense of $\S x(x = \S xFx)$ which is presumably not so. (p. 40, emphasis added)

We grant that "$\S xFx$" and "$\S x(x = \S xFx)$" have different senses, since (assuming that extensions behave like sets) the referent of the latter, but not the former, is guaranteed to have exactly one member. The question, however, is how Fine's argument is meant to reach the false conclusion that "$\S xFx$" and "$\S x(x = \S xFx)$" have the same sense.

    The argument comes in two parts. First, all of (A) through (D) have the same sense. Second, it follows from this that "$\S xFx$" and "$\S x(x = \S xFx)$" have the same sense. Neither claim is immune to doubt.

    First, although (B) and (A) have the same sense through reconceptualization (as do (D) and (C)), the claim that (B) and (C) have the same sense can be challenged. There are two reasons one might think that (B) and (C) have the same sense (neither of which is explicitly considered by Fine). First, one might think that logically equivalent formulae (such as (B) and (C)) have the same sense. Within higher-order logic this is tremendously dubious – second-order formulations of the *Axiom of Choice* and *Zorn's Lemma* are provably equivalent but, intuitively, have different senses. Second, one might argue that (B) and (C) have the same sense based on both their logical equivalence and their 'closeness' in terms of logical form (as

measured, perhaps, by the shortest proof of equivalence in some privileged deductive system). Notable among the problems with the 'closeness' approach is that sameness of sense becomes a vague matter.

Having noted the assumption, we grant, for the sake of argument, that (A) through (D) have the same sense. It remains to determine how Fine reaches the conclusion that "§xFx" and "§$x(x = $§$xFx)$" share a sense as well.

It is tempting to conclude that Fine's argument consists merely in moving from the sameness of sense of the identity statements (B) and (D) to the sameness of sense of the corresponding terms. The notion of reconceptualization makes such a move questionable at best. Reconceptualization allows formulae with significantly different logical structures to have the same sense. As a result, two identities (such as (B) and (D)) might have the same sense, but the explanation of this synonymy goes via an intermediate step (in this case (C)) with a different logical form. This intermediate step makes the similarity of logical form between (B) and (D), in a sense, accidental – after moving from (B) to (C) there is nothing in the structure or sense of (C) that *requires* us to (re-)reconceptualize so as to obtain (D). As a result, the sameness of sense of identities need not depend on the sameness of the sense of their components.

We can see that this simplistic reading is a mis-interpretation of Fine by attending to the two underlined parts of the argument. First, in adopting this interpretation we must read the underlined occurrence of "(A)" as a typo, since the problematic identity between senses is not between (A) and (D), but (B) and (D). Second, we must view as irrelevant to the argument what was meant to be its crucial premise – "the sense of the operator that is given through reconceptualization is unique" – yet the argument is meant to be a *reductio* on this claim. The simplistic interpretation requires only that each of four formulae involved have the same (unique) sense, it does not require that the operators occurring within these formulae do so as well.

Adopting this reading of Fine's position is unfortunately an easy mistake. The presentation of the argument is overly brief and, in addition, the wealth of typos elsewhere in the book provides some indirect support for the hypothesis that "(A)" is a typo here. We had ourselves adopted the simplistic interpretation (as had many of our friends and colleagues) until a opportune conversation with Fine himself, to whom thanks are due.

Taking the wording of the argument at face value, the identity of sense that causes the trouble is that between (A) and (D). Adopting the abbreviation:

$$@xFx \qquad = \qquad §x(x = §xFx)$$

we can rewrite (D) as:

(D*) $@xFx = @xGx$

Since (D*) is just the result of substitution of (explicitly defined) definiens for definiendum, (D) and (D*) presumably have the same sense.

The problem is now apparent: We have two operators, (§ and @) both of which result from reconceptualization of (A). Thus, if the sense of an operator given by reconceptualization is unique, "§xFx" and "@xFx" (i.e. "§$x(x = $§$xFx)$") must have the same sense (and thus the same reference).

We can avoid Fine's conclusion, however, by noting that "@", unlike "§", does not obtain its sense solely through reconceptualizition of (A). In reconceptualizing, we laid down *BLV* in order to provide a sense for §. If, in addition, we had adopted:

$$BLV^*: \ @xFx = @xGx \ \leftrightarrow (\forall x)(Fx \leftrightarrow Gx)$$

as a distinct instance of definition by reconceptualization, we would be forced to conclude that "§" and "@" have the same sense. Had we taken this route, with "@" as a primitive whose sense was provided directly by *BLV**, there would be no reason to think that @xFx is the singleton of §xFx (or more generally that "@xFx" and "§xFx" could not have the same sense).

Crucially, in Fine's argument "@" is <u>not</u> defined directly through reconceptualization, but is instead a complex operation <u>built up out of a simpler one</u>. "@" obtains its sense as a function of the sense of "§" in virtue of being explicitly defined in terms of "§". The fact that (D*) and (B) have the same sense as (A) should not fool us into thinking that the operator in (D*) obtained its sense in the same way as the operator in (B).

Fine's argument against reconceptualization, even if unsuccessful, does pinpoint certain crucial assumptions underlying the view. In particular, if we grant that logically equivalent statements have the same sense (and we need not do so), the equivalence relation expressed by "sameness of sense through reconceptualization" must be strictly stronger than that expressed by "sameness of sense" – the former carves up the linguistic world more finely than the latter. Although two formulae that have the same sense through reconceptualization will obviously have the same sense, crucially there will be formulae that have the same sense but which do not stand in the "sameness of sense through reconceptualization" relation. We do not see why this is all that bitter a pill, but will leave its defence for another time.

The Context Principle

Fine argues at length against the Neo-Fregean use of the *Context Principle* in the second chapter. He writes that:

> The 'context principle' can be regarded as an attempt to vindicate... contextual definitions. (p. 55)

After sketching the content of the *Context Principle*, Fine devotes a number of sections to demonstrating that justifying abstraction principles in terms of *CP* fails to solve a number of

deep philosophical problems. We shall have something to say regarding these objections below. Before doing so a short examination of Fine's taxonomy of definitions is in order.

Paraphrasing Fine, implicit definitions are definitions whose success depends on prerequisite knowledge of the truth of the corresponding Ramsey sentence – we have to know that the definition is satisfiable before we can successfully deploy it. This requirement, however, is not in place for *creative* definitions – no prior knowledge of the truth of the Ramsey sentence is demanded, rather the "objects are introduced into the discourse" in virtue of the definition.

It is important to distinguish an *epistemic* from an *ontological* understanding of creative definitions. The former entails merely that we don't have to <u>know</u> that the Ramsey sentence is true in order to successfully put forward the definition (it is still required that the Ramsey sentence is, in fact, true). The success of such a definition provides justification for believing in the truth of the relevant Ramsey sentence. The latter, ontological reading involves the idea that it is in virtue of a successful definition that the objects exist, i.e. it 'creates' the objects and 'adds' them to the domain. On this conception, prior to putting forth the definition the corresponding Ramsey sentence might be false, while the success of the definition guarantees its (subsequent) truth. This distinction is especially important since the epistemological reading is distinctive of the Neo-Fregean project.

Fine introduces a further genus, *contextual definitions* (p.60), which are bridge principles introducing new vocabulary by fixing their truth-conditions in terms of previously understood vocabulary. Abstraction principles are one species of contextual definition. As before, it is reasonable to distinguish *standard* or *implicit contextual* definitions, where success involves prior knowledge of the truth of the Ramsey sentence, from *creative contextual* definitions, where this constraint is not in place. Again, it is the latter (on its epistemic reading) that is of interest to the Neo-Fregean.

We now turn from definitions to the use of the *Context Principle* in their justification. The *Context Principle* traces back to Frege's [1884] *Grundlagen der Arithmetik*, where he observes that:

> ... one has to ask for the meaning of words in the context of sentences, not in their isolation. (*Grundlagen*, p. *x*).

The *Context Principle* is obtained by replacing "words" with "singular terms" and identifying meanings with referents. If what is distinctive about the occurrence of singular terms in sentences, as opposed to those same terms in isolation, is the truth value of the sentence, it follows that the reference of singular terms is tied to the truth values of sentences in which they occur.

Again, we can distinguish an *epistemic* from an *ontological* reading. The former understanding is akin to a minimalist conception of reference, spelling out constraints for successful reference. The latter is more substantial in that the objects come into existence in

virtue of *CP* (and the truth of appropriate sentences). Once again, it is the former, epistemic reading which is relevant to the Neo-Fregean project.

How, according to Fine, can the *Context Principle* be regarded as justifying creative/contextual definitions? One way of understanding Fine is that it is a consequence of *CP* that the usual order of explanation of identity statements is reversed:

> Instead of saying that the sentence '*t* = *t*' is true in virtue of the fact that the term *t* refers to a given object *x* and that object is identical with itself, we say that the term *t* refers in virtue of the fact that the sentence '*t* = *t*' is true – where this latter fact is itself explained without appeal to the reference of '*t*'. (p. 68)

This line of thought can be extrapolated to creative/contextual definitions. For example, instead of stating that *Hume's Principle* is true because numerical terms of the form "*Nx:Fx*" refer, and that the various referents stand in appropriate relations to one another, we say that the numerical terms refer, and that the referents (numbers) stand in certain relationships to each other, because particular instances of *HP* are true.

Fine further fleshes out the idea that *CP*, in conjunction with the truth of appropriate sentences, secures the reference of abstraction terms:

> ...the apparently linguistic behaviour of the terms endows them with the *potential* to refer, a potential that will be realized when the appropriate non-linguistic facts obtain. Thus referential potential is the counterpart, within the framework of *CP*, to an identifying sense within the standard Fregean model of reference. (p. 58)

Unfortunately, Fine nowhere explicitly tells us what these "non-linguistic facts" are.

At first glance one is tempted to identify the non-linguistic facts relative to *HP* with the commitments incurred by the corresponding Ramsey sentence. If the justification of creative/contextual definitions via *CP* required prior knowledge of the Ramsey sentence, however, then any creative definitions made on the basis of *CP* would collapse into standard implicit definitions (Fine does think that creative definitions justified via *CP* collapse into something like implicit definitions, but his reasons for thinking so are rather more subtle).

A different interpretation is that, given an instance of *HP*, the non-linguistic facts in question are the commitments of the (epistemologically) more basic right hand side of the biconditional, i.e. a one-to-one correspondence between concepts. The existence of such a correspondence guarantees, via the definition, the truth of the relevant identity claim regarding abstracts, which in turn entails, on the basis of *CP*, the reference of the number operator. Thus, *CP* only applies to definitions which are already successful (and thus true).

Once this is clarified, it becomes clear that much of Fine's critical examination of the *Context Principle* as a "vindication" of creative/contextual definitions misses the mark. *CP*, properly understood, is not intended to "vindicate" creative/contextual definitions in the sense of providing a <u>guarantee</u> that such definitions are successful – *Basic Law V* has taught us this much – nor does it account for the constraints on the success of individual creative/contextual

definitions (interestingly, the third chapter of Fine's book contains by far the most extensive investigation of this issue in the literature). Rather, *CP* is mobilized in order to secure the reference of numerical (and other) terms occurring in contextual definitions; the definitions themselves are justified by other means (perhaps through reconceptualization). The Neo-Fregean, however, does not (or at least should not) expect *CP* itself to do this justificatory work. Hence, most of the problems Fine subsequently raises for *CP* are unsurprising, at least in the sense that one would expect <u>some</u> <u>such</u> problems. Fine's identification and discussion of the particular difficulties facing creative/contextual definition is nonetheless immensely valuable. As a result we see no reason for abandoning *CP*, which seems rather benign in its epistemic version.

Fine, however, delivers a second threat to the *Context Principle*, by arguing that *CP* is incompatible with his solution to the *Caesar Problem*. This objection, if successful, would be more worrying, since presumably the solutions to various problems confronting creative/contextual definitions should be jointly consistent. Although limitations on space prevent a full discussion here, we feel that this second strategy also fails to be convincing (mainly on dialectial grounds, because there are other extant attempts at such a solution that are compatible with *CP*, and Fine gives no reason to reject those). Hence, we think that Fine fails to provide a decisive strategy against *CP,* and (on its epistemic reading) it emerges from his two-pronged attack alive and kicking.

Conclusions

Although, as noted above, Fine ultimately decides against the *Context Principle* as a possible means for justifying creative/contextual definitions, Fine is not averse to the possible success of such definitions in general. On the contrary, he explicitly accepts that creative definitions (understood ontologically) can provide a satisfactory foundation for mathematics. His view, *procedural postulationalism*, is based on the idea that:

> ... instead of stipulating that certain statements are to be true, one specifies certain procedures for extending the domain to one in which the statements will in fact be true. These procedures can be stated without invoking an abstract ontology; they achieve referential determinacy; and their legitimacy does not depend upon the prior knowledge that the objects which are to be introduced into the domain already exist. (p. 100)

Unfortunately, this positive view is only mentioned four times (pp. *v*, 36, 56, 100), with the remainder devoted to discussing and developing various problems concerning more traditional theories of abstraction. As a result, the book is somewhat disappointing, since one cannot but wonder how Fine's procedural postulationism will deal with the diverse and rather sophisticated problems raised in the text. Thus, the reader is likely to feel some frustration while waiting for the planned sequel, which promises to provide the conclusion to this abstractionist story.

Setting aside this minor disappointment, and notwithstanding our particular gripes surveyed above, the text is essential reading for anyone interested not only in abstractionist philosophies of mathematics, but the philosophy of mathematics and language in general. The philosophical chapters display a consistently high level of rigour and insight (with a corresponding level of difficulty) when discussing familiar topics, usually providing new perspectives on old problems. Additionally, the new philosophical problems raised are valuable and thought provoking, and promise to be the basis for much philosophical discussion to come. The later chapters on formal topics are equally essential, insofar as they drastically alter the landscape of formal work on abstraction principles and demonstrate the relative lack of serious mathematical work on this subject prior to Fine's book. A philosopher of our acquaintance recently noted that it is difficult to refer to a typical 'Fine' view without sounding like one is approving, but in this instance our overall assessment would have to be that this is, indeed, a 'Fine' book.

References:

Fine, K [1998], "The Limits of Abstraction", in *The Philosophy of Mathematics Today: Proceedings of an International Congress in Munich*, M. Schirn (ed.), The Mind Association, Oxford, Oxford University Press.

Frege, G. [1884], *Grundlagen der Arithmetik*, Verlag von Wilhelm Koebner, Breslau.

Roy T. Cook
rtc1@st-andrews.ac.uk

Philip Ebert
pae1@st-andrews.ac.uk

Arche: The AHRB Centre for the Philosophy of Logic, Language, Mathematics, and Mind
School of Philosophical and Anthropological Studies
University of St. Andrews
St. Andrews, Fife, Scotland