Resolution by Proxy

Samuel Z. Elgin[1]

In a series of recent papers, I presented a puzzle and theory of definition.[2] I did not, however, indicate how the theory resolves the puzzle. This was an oversight, on my part, and one I hope to correct. My aim here is to provide that resolution: to demonstrate that my theory can consistently embrace the principles I prove to be inconsistent. To the best of my knowledge, this theory is the only one capable of this embrace—which marks yet another advantage it has over competitors.

Both the puzzle and theory are expressed in the language of higher-order logic. I will assume broad familiarity with this sort of language here; those unfamiliar with this formalism would do well to read the original papers—or else some introductory text on the subject.[3] Suffice it to say that I will operate with a higher-order language with two basic types—e and t—the types of entities and sentences respectively. For every types τ_1 and τ_2, $(\tau_1 \rightarrow \tau_2)$ is a type, and nothing else is a type. There are infinitely many variables of every type—as well as the λ-abstracts needed to bind them. Predicates, quantifiers, signs for identity, and the logical connectives $\neg, \wedge$ and $\vee$ are all identified with terms of types in the standard way. I also introduce the predicate *Def* of type $(\tau \rightarrow (\tau \rightarrow t))$ which is used to express definitions. That is, '$Def(F, G)$' is to be read as 'F is, by definition, G.' The only additional formalism (introduced solely to reduce the otherwise-unwieldy length of types) is τ^2, which is shorthand for $(\tau \rightarrow (\tau \rightarrow t))$ for an arbitrary type τ.

The puzzle concerns the following five principles governing the logic of real definition:

COEXTENSIONALITY:	$Def^{t^2}(P^t, Q^t) \rightarrow (P^t \leftrightarrow Q^t)$
IRREFLEXIVITY:	$\neg\exists\lambda X^\tau . Def^{\tau^2}(X, X)$
CASE CONGRUENCE:	$Def^{(\tau \rightarrow t)^2}(F^{\tau \rightarrow t}, G^{\tau \rightarrow t}) \rightarrow Def^{t^2}(F^{\tau \rightarrow t}(a^\tau), G^{\tau \rightarrow t}(a^\tau))$
EXPANSION:	$(Def^{\tau^2}(F^\tau, G^\tau) \wedge Def^{\alpha^2}(H^\alpha, I^\alpha)) \rightarrow Def^{\tau^2}(F^\tau, G^{\tau\ [\mathrm{I/H}]})$
DEFINABILITY:	$\exists\lambda X^{\tau^2} . Def^{\tau^2 \rightarrow (\tau^2 \rightarrow t)}(Def^{\tau^2}, X)$

Coextensionality asserts that if p is, by definition q, then p and q have the same truth-value. For example, if the proposition that water is wet is, by definition, the proposition that H_2O is wet, then water is wet iff H_2O is wet. *Irreflexivity* precludes reflexive definitions; it is not the case that knowledge is, by definition, knowledge or that virtue is, by definition, virtue. *Case congruence* asserts that definitions apply to their cases. If *being a brother* is, by definition, *being a male sibling*, then the proposition that John is a brother is, by

[1] My thanks to Cian Dorr, who has substantially shaped my views in this area—and who first suggested resolving the puzzle of definition in this manner.

[2] See Elgin (2022, Forthcoming)

[3] See Bacon (2022) for one such introduction.

definition, the proposition that John is a male sibling. *Expansion*, for its part, allows for the substitution of terms for their own definitions within the contents of others. For example, if *being hydrogen* is, by definition, *being the element with a single proton*—and *being a proton* is, by definition, *being the subatomic particle consisting of two up-quarks and a down-quark*, then *expansion* entails that *being hydrogen* is, by definition, *being the element with a single subatomic particle consisting of two up-quarks and a down-quark*. *Definability*, lastly, asserts that definition is itself defined: it does not rank among the primitive relations.

Each of these principles is independently plausible and—as I have argued—many have independent support. However, they also seem to be incompatible. The inconsistency can be brought out in the following way (allowing D to represent the content of the definition of definition—whatever that content might be):

i)	$Def^{\tau^2}(F^\tau, G^\tau)$	Supposition
ii)	$Def^{\tau^2 \to (\tau^2 \to t)}(Def^{\tau^2}, D^{\tau^2})$	Definability
iii)	$Def^{\mathrm{t}^2}(Def^{\tau^2}(F^\tau, G^\tau), D^{\tau^2}(F^\tau, G^\tau))$	*ii*, Case Congruence
iv)	$Def^{\mathrm{t}^2}(Def^{\tau^2}(F^\tau, G^\tau), D^{\tau^2}(G^\tau, G^\tau))$	*i*, *iii*, Expansion
v)	$Def^{\tau^2}(F^\tau, G^\tau) \leftrightarrow D^{\tau^2}(G^\tau, G^\tau)$	*iv*, Coextensionality
vi)	$D^{\tau^2}(G^\tau, G^\tau)$	*i*, *v*, Classical Logic
vii)	$Def^{t^2}(Def^{\tau^2}(G^\tau, G^\tau), D^{\tau^2}(G^\tau, G^\tau))$	*ii*, Case Congruence
viii)	$Def^{\tau^2}(G^\tau, G^\tau) \leftrightarrow D^{\tau^2}(G^\tau, G^\tau)$	*vii*, Coextensionality
ix)	$Def^{\tau^2}(G^\tau, G^\tau)$	*vi*, *viii*, Classical Logic
x)	$\exists \lambda X^\tau . Def^{\tau^2}(X, X)$	*ix*, Classical Logic
xi)	$\perp$	*x*, Irreflexivity

Due to this conflict, various philosophers have suggested rejecting one or more of the principles in conflict. Werner (2022) argues that we ought to reject *expansion* in hyperintensional contexts (and that definition itself generates a hyperintensional context), while Jeremy Goodman has (conversationally) suggested rejecting *case congruence* on the grounds that only logical simples admit of definition—and all applications of *case congruence* concern terms that are not logical simples.

These abandonments are costly. I myself prefer neither of these approaches; I hold that all five principles are true. The task, then, is not to determine which principle to reject, but rather to construct a theory of definition on which this contradiction cannot be derived. The account in *Definition by Proxy* is such a theory.

Note, at the outset, a tacit assumption within this puzzle—that 'definition' is of type $(\tau \to (\tau \to t))$: it takes sequences of two terms *of the very same type* as its inputs. I take it that this is an extremely natural assumption, but it is one that I ultimately reject.[4] Definition, on my view, is not a relation between two terms of type τ, but rather a relation

[4] This rejection is not theoretically idle. I rely upon the different types in order to prove the irreflexivity of definition.

between a term of type τ *and a higher-order proxy for a term of type* τ. Thus, 'To be a bachelor is, by definition, to be an unmarried male' does not ultimately relate two terms of type $(e \rightarrow t)$, but rather a relation between one term of type $(e \rightarrow t)$ to a proxy for a term of type $(e \rightarrow t)$. The shift from terms to proxies for terms will ultimately allow us to consistently embrace these five principles.

Proxy-theory was first developed by Fritz (2021); higher-order proxies serve to make extraordinarily fine-grained distinctions while avoiding the problems that typically plague fine-grained accounts.[5] Rather than appealing directly to the proposition that Fa, proxy-theory makes use of the bihaecceity $\lambda X.\lambda x.(X = F \wedge x = a)$: the relation that property F stands in to object a, and that no other property stands into any other object. Unlike the proposition Fa (from which it is impossible to extract the unique property that occurs within it), it is possible to extract property F and object a from this proxy. We can provide a function that generates this sort of proxy (i.e., that generates proxies for the propositions Gb, $\forall xHx$, etc.) as follows:

$$\delta := \lambda X^{\tau_1 \rightarrow \tau_2}.\lambda x^{\tau_1}.\lambda Y^{\tau_1 \rightarrow \tau_2}.\lambda y^{\tau_1}.(X = Y \wedge x = y)$$

δ takes pairs of terms as its input and has—as its output—the relation that only the first input stands into the second. The only restriction for δ is that the second term be the functional input of the first. We can provide a function without this restriction as follows:

$$\gamma := \lambda X^{\tau_1}.\lambda x^{\tau_2}.\lambda Y^{\tau_1}.\lambda y^{\tau_2}.(X = Y \wedge x = y)$$

With the γ function in place, we can represent the proxy for the proposition that Fa as $\gamma(F, a)$. In order to simplify this notation still further, I represent the result of γ with [ ] notation (so that $\gamma(F, a)$ is to be represented as $[F, a]$).

We can also generate a recovery function that takes, as its input, instances of δ and has, as its output, the coarse-grained term it is a proxy for as follows:

$$Rec(\delta(\alpha, \beta)) = \alpha(\beta)^6$$

[5] In Fritz's case, to resolve puzzles of ground. The problems I allude to are the Russell-Myhill problem—which has recently generated an explosive literature—as well as the suggestion that property F can be extracted from the proposition that Fa. See Dorr (2016) for a discussion of the problems this problem generates.

[6] Note that the restriction to δ, rather than γ, is needed to ensure that $\alpha(\beta)$ is grammatical within our language. Had we defined *Rec* over γ, there would be no guarantee that the second term was the functional input of the first—in which case '$\alpha(\beta)$' would not occur within our language. Note, also, that this account of *Rec* may appear to be a restriction—rather than a full definition. (In that it ensures that $< \delta(\alpha, \beta), \alpha(\beta) >$ occurs within its extension—but says nothing about what else might be in this relation's extension). The intended interpretation is that *Rec* is undefined for any input that is not an instance of δ. Given some fine-grained conceptions of relations, there may be multiple relations that satisfy this restriction (if there are distinct, logically equivalent relations). If so, select an arbitrary relation that satisfies this description.

δ and γ allow us to express proxies of terms of arbitrary type: not only for propositions, but for properties, relations, etc. However, they are limited in that they are only sensitive to the outermost syntactic structure of a term. That is to say, while can use $[\neg, Fa]$ to denote the relation negation stands in to the proposition that Fa (and thereby express a proxy for the proposition $\neg Fa$), this is not a term from which we can extract either property F or object a. The obvious way to describe internal syntactic structure, in this higher-order framework, is recursively. To that end, we may define a family of relations of Decomposition—and say that one term may be decomposed into another. Decomposition is the smallest family of relations satisfying the following:[7]

1. If, $\alpha = Rec(\delta(\beta, \psi))$, then $Dec(\alpha, [\beta, \psi])$
2. If $Dec(\alpha, [\beta, \psi])$ and $Dec(\beta, [\eta, \epsilon])$ then $Dec(\alpha, [[\eta, \epsilon], \psi])$
3. If $Dec(\alpha, [\beta, \psi])$ and $Dec(\psi, [\eta, \epsilon])$ then $Dec(\alpha, [\beta, [\eta, \epsilon]])$

With the notion of decomposition in place, it is possible to represent two distinct proxies for the proposition $\neg Fa$. We might, first, represent it with $[\neg, Fa]$—or alternatively with $[\neg, [F, a]]$. While the first of these denotes the relation that negation stands in to the proposition Fa, the second denotes the relation that negations stands in to the relation F stands in to a. There is thus a spectrum of granularity of proxies where, at one end of the spectrum, a proxy is sensitive only to the outermost syntactic structure of a term while, at the other end, a proxy reveals the entirety of a term's syntactic structure. The latter is of special interest—so it is useful to introduce additional formalism describing that instance of decomposition. This can be accomplished (recursively) in the following way:

$$
\begin{array}{ll}
\lceil AB \rceil = [A, B] & \text{for constants } A^{\tau_1 \to \tau_2}, B^{\tau_1} \\
\lceil A\beta \rceil = [A, \lceil \beta \rceil] & \text{for constant } A^{\tau_1 \to \tau_2} \text{ and non-constant } \beta \\
\lceil \alpha B \rceil = [\lceil \alpha \rceil, B] & \text{for non-constant } \alpha \text{ and constant } B^{\tau_1} \\
\lceil \alpha\beta \rceil = [\lceil \alpha \rceil, \lceil \beta \rceil] & \text{for non-constants } \alpha, \beta
\end{array}
$$

Note that the $\lceil\ \rceil$ notation is only defined over grammatically correct sequences of two-or-more constants, and that the types of these terms depend not only on the types of the terms occurring within them, but on their syntactic structure as well. While $\lceil Fa \rceil$ strongly resembles $\lceil \neg\neg Fa \rceil$, the two are of entirely different types.

We can also define a function $\lfloor\ \rfloor$ that takes, as its input, a proxy of arbitrary type and has, as its output, the term it is a proxy for in the following way:

$$\lfloor \alpha \rfloor = \beta \text{ iff } Dec(\beta, \alpha)$$

Elsewhere, I have established that each proxy is a proxy for a unique term—so we can indeed be confident that $\lfloor\ \rfloor$ is a function. With these formalisms in place, I suggested the

[7] As with the definition of Rec, if there is no one smallest family, then select an arbitrary smallest family.

following condition on definition (which I dubbed Definition by Proxy—DBP):

$$Def^{*}(\alpha, \beta) \to Dec(\alpha, \beta)$$

If α is, by definition, β, then α can be decomposed into β. Note that, by necessity, Def^{*} is strictly a different relation than the former relation Def. Def was introduced of type $(\tau \to (\tau \to t))$, while Def^{*} takes terms of different types as its inputs.

DBP was sufficient to resolve the puzzles I then discussed (puzzles concerning the paradox of analysis, the granularity of definition, and the putative conflict between Leibniz's Law, the irreflexivity of definition, and the claim that definitions entail their corresponding identities). However, for my purposes now, I will consider the somewhat stronger claim:

$$Def^{*}(\alpha, \beta) \to Dec(\alpha, \lceil \eta \rceil) \text{ (where } \beta = \lceil \eta \rceil)$$

The difference between these is that the former principle is compatible with terms being defined by proxies of any granularity, while the latter requires that terms are defined by their most fine-grained proxy. That is to say, the former principle is compatible with $Def(\neg Fa, [\neg, Fa])$, the latter is not. Because $[\neg, Fa]$ is not an ultimately fine-grained proxy ($[\neg, [F, a]]$ being more fine-grained than it), it cannot serve to define the negation.

The five principles mentioned at the outset can then be restated for $Definition^{*}$—appropriately altering the types of the terms involved:

COEXTENSIONALITY*:	$Def^{t \to (\tau \to t)}(P^t, Q^\tau) \to (P^t \leftrightarrow \lfloor Q^\tau \rfloor)$
IRREFLEXIVITY*:	$\neg \exists \lambda X^\tau . Def^{\tau^2}(X, X)$
CASE CONGRUENCE*:	$Def^{(\tau_1 \to t) \to (\tau_2 \to t)}(F^{\tau_1 \to t}, \lceil G^{\tau_1 \to t} \rceil) \to Def^{t \to (\tau_3 \to t)}(F^{\tau_1 \to t}(\alpha^{\tau_1}), \lceil G^{\tau_1 \to t}(\alpha^{\tau_1}) \rceil)$
EXPANSION*:	$(Def^{\tau_1 \to (\tau_2 \to t)}(F^{\tau_1}, \lceil G^{\tau_1} \rceil) \wedge Def^{\tau_3 \to (\tau_4 \to t)}(H^{\tau_3}, \lceil I^{\tau_3} \rceil)) \to Def^{\tau_1 \to (\tau_5 \to t)}(F^{\tau_1}, \lceil G^{\tau_1} \rceil^{[H/I]})$
DEFINABILITY*:	$\exists \lambda X^{\tau_3} . Def^{(\tau_1 \to (\tau_2 \to t)) \to (\tau_3 \to t)}(Def^{\tau_1 \to (\tau_2 \to t)}, X)$

DBP is not agnostic with respect to some of these claims; I have previously established that definition is irreflexive on this account (on the grounds that for any terms A and B such that A is, by definition, B, B will fall higher on the hierarchy of types than A). It is also committed to coextensionality. Having established that $Def(P, Q) \to P = \lfloor Q \rfloor$, an application of Leibniz's Law ensures that P and Q have the same truth-values. This is thus committed both to *coextensionality** and *irreflexivity**.

Enough background. What is it that prevents us from carrying out the previous derivation—and generating an inconsistency for these five principles in precisely the manner we did for the original five?

We being in the same way. Select an arbitrary F and G such that

$$Def^{*\tau_1 \to (\tau_2 \to t)}(F^{\tau_1}, \lceil G^{\tau_1} \rceil) \tag{1}$$

As before, let D witness whatever it is that definition is defined in terms of. Given *definability**, we thus have:

$$Def^{*(\tau_1 \rightarrow (\tau_2 \rightarrow t)) \rightarrow (\tau_3 \rightarrow t)}(Def^{*\tau_1 \rightarrow (\tau_2 \rightarrow t)}, \ulcorner D^{\tau_1 \rightarrow (\tau_2 \rightarrow t)} \urcorner) \tag{2}$$

(Note, here, that while we can be confident that D is of the same type as *Def**, there is no way to determine the type of $\ulcorner D \urcorner$ without information about its syntactic structure—hence the appeal to type τ_3 in line 2). An application of *case congruence** on 2 then results in:

$$Def^{*t \rightarrow (\tau_4 \rightarrow t)}(Def^{*\tau_1 \rightarrow (\tau_2 \rightarrow t)}(F^{\tau_1}, \ulcorner G^{\tau_1} \urcorner), \ulcorner D^{\tau_1 \rightarrow (\tau_2 \rightarrow t)}(F^{\tau_1}, \ulcorner G^{\tau_1} \urcorner) \urcorner) \tag{3}$$

Appealing to 1, 3 and *expansion** then results in:

$$Def^{*t \rightarrow (\tau_4 \rightarrow t)}(Def^{*\tau_1 \rightarrow (\tau_2 \rightarrow t)}(F^{\tau_1}, \ulcorner G^{\tau_1} \urcorner), \ulcorner D^{\tau_1 \rightarrow (\tau_2 \rightarrow t)}(G^{\tau_1}, \ulcorner G^{\tau_1} \urcorner) \urcorner) \tag{4}$$

Given *coextensionality** this then entails:

$$Def^{*\tau_1 \rightarrow (\tau_2 \rightarrow t)}(F^{\tau_1}, \ulcorner G^{\tau_1} \urcorner) \leftrightarrow D^{\tau_1 \rightarrow (\tau_2 \rightarrow t)}(G^{\tau_1}, \ulcorner G^{\tau_1} \urcorner) \tag{5}$$

1, 5 and classical logic then give us:

$$D^{\tau_1 \rightarrow (\tau_2 \rightarrow t)}(G^{\tau_1}, \ulcorner G^{\tau_1} \urcorner) \tag{6}$$

We can then apply *case congruence** to 2 directly to conclude:

$$Def^{*t \rightarrow (\tau_4 \rightarrow t)}(Def^{*\tau_1 \rightarrow (\tau_2 \rightarrow t)}(G^{\tau_1}, \ulcorner G^{\tau_1} \urcorner), \ulcorner D^{\tau_1 \rightarrow (\tau_2 \rightarrow t)}(G^{\tau_1}, \ulcorner G^{\tau_1} \urcorner) \urcorner) \tag{7}$$

Given *coextensionality**, this then entails:

$$Def^{*\tau_1 \rightarrow (\tau_2 \rightarrow t)}(G^{\tau_1}, \ulcorner G^{\tau_1} \urcorner) \leftrightarrow D^{\tau_1 \rightarrow (\tau_2 \rightarrow t)}(G^{\tau_1}, \ulcorner G^{\tau_1} \urcorner) \tag{8}$$

6, 8 and classical logic then entail:

$$Def^{*\tau_1 \rightarrow (\tau_2 \rightarrow t)}(G^{\tau_1}, \ulcorner G^{\tau_1} \urcorner) \tag{9}$$

Up to this point, the proof has proceeded unhindered in just the manner of the original. (In particular, there were no problems with type-shifts that would be responsible for the proof's failure). However, a crucial difference now arises. In the original derivation, the 9th line read $Def(G, G)$—but the present case takes the form $Def(G, \ulcorner G \urcorner)$. In the first derivation we arrived at a reflexive definition: one were G is defined in terms of itself. But on the present theory G is defined in terms of a *proxy* for G—rather than in terms of G directly. We thus fail to violate *irreflexivity* in this way on the present theory of real definition.

The shift from terms to proxies for terms serves our purposes once again. Not only may we distinguish between terms in a manner as fine-grained as our syntactic structure—not only may we consistently embrace Leibniz's Law, *irreflexivity*, and the claim that definitions entail their corresponding identifications—not only do we resolve the paradox of analysis by providing an informational difference between a definiendum and its definiens—but we prevent the derivation of a contradiction between *coextensionality, irreflexivity, case congruence, expansion and definability.*

References

Bacon, Andrew. 2022. *A Philosophical Introduction to Higher Order Logic*. Routledge.

Dorr, Cian. 2016. "To be F is to be G." *Philosophical Perspectives* 30(1):39–134.

Elgin, Samuel. 2022. "Definition." *Philosophical Studies* .

Elgin, Samuel. Forthcoming. "Definition by Proxy.".

Fritz, Peter. 2021. "Structure by Proxy with an Application to Grounding." *Synthese* 198:6045–63.

Werner, Jonas. 2022. "Solving a Puzzle of Definition." *Inquiry* .