PROBLEMS FOR PROPOSITIONS

Samuel Z. Elgin[1]

Abstract

This paper consists of an investigation of three debates concerning propositional identity: the tension between structured propositions and higher-order logic, the principle Only Logical Circles, and Kaplan's Paradox. The literature at large has mistaken the consequences of each of these debates. Structuralists are not committed to the claim that identical properties have different extensions; rather, they are committed to existence monism. Only Logical Circles does not preclude the identification of green in terms of grue; some further (and, as of yet, undisclosed) principle is required for that result. And Kaplan's Paradox does not count against the possible-worlds conception of propositions; it applies to every account of propositions whatsoever.

There has recently been sustained interest in the nature and identity of propositions. This interest is perhaps attributable to the significance of the theoretical work that they are often taken to perform. Propositions are held to be the ultimate bearers of truth and falsity, and are the objects of attitudes like belief, hope, desire and fear. A sentence is said to be true just in case the proposition that it expresses is true—and sentences are said to be synonymous just in case they express the same proposition. Given the importance of these roles, it is incumbent for the metaphysician to ascertain what kinds of things propositions are.

Accounts fall on a spectrum of granularity. While coarse-grained accounts hold that all necessarily equivalent propositions are identical, fine-grained accounts differentiate between any that differ in their syntactic structure. In recent years, intermediate accounts—which distinguish between some, but not all, logically equivalent expressions—have received more attention then they have at any point in history. Unsurprisingly, given this context, a wide array of arguments have been advanced targeting the various accounts. Coarse-grained accounts struggle to accommodate the incongruous attitudes agents bear towards necessarily equivalent propositions; fine-grained accounts may be incompatible with orthodox principles of higher-order logic; and accounts of all stripes struggle with cardinality.

More surprisingly, the significance of many of these arguments has been widely—perhaps even universally—misunderstood. Some accounts do not have the implications philosophers have claimed that they have, and some arguments do not support the theories they have been taken to support. The aim of this paper is to correct these misunderstandings: a clean-up job, if you will. The result is that commitments that lie at the core of standard conceptions of propositions turn out to be deeply muddled. Propositions—if there are such

---

things—differ from how we have taken them to be.[2]

The first misunderstanding concerns structured propositions and higher-order logic. One of the central tenets of the structured-proposition view is that $Fa = Gb$ entails that $F = G$; if the proposition that John is a bachelor is identical to the proposition that John is an unmarried male, then the property of *being a bachelor* is identical to the property of *being an unmarried male*. Some claim that, when combined with the orthodox assumption that propositional identity is preserved through $\beta\eta$-conversion, this entails that properties are identical despite not being coextensive. Not so. It entails that the relevant properties *are* coextensive. This is because it has a much more interesting (and arguably more ominous) implication: monism is true. There exists but a single object to fall in the extension of any property.

A second misunderstanding concerns the principle Only Logical Circles—according to which logically circular identifications are innocuous, while nonlogical circularity is vicious. The primary motivation for this principle is that it allows for identifications of the form 'To be F is to be F,' while precluding identifications like 'To be green is to be grue and observed before time t or not so observed and bleen.' Only Logic Circles lacks this result. It only prevents the identification of green in terms of grue with the addition of other (and, as yet, undisclosed) principles.

A third misunderstanding concerns Kaplan's Paradox. It is often held that coarse-grained accounts have an untenable consequence: there are some propositions for which there exists no possible world in which they are uniquely considered (or transcribed, bet on, etc.). For this reason, coarse-grained accounts ought to be rejected in favor of an alternative. This is false. Every theory of propositions falls prey to Kaplan's Paradox, so the paradox provides no reason to reject one account in favor of another.

Before turning to these misunderstandings in detail, a quick note on the scope and limitations of this project. It is not my aim to fully rehabilitate theories that philosophers have previously rejected. It may be, for example, that metaphysicians find monism to be as untenable as the claim that properties with different extensions are identical. If so, then the fact that the theory of structured propositions entails that monism is true may be taken as a sufficient reason to reject that theory. Nevertheless, it is important, I think, to correctly understand the implications of our arguments—and to appreciate the reasons that we accept and reject the claims that we do. Relatedly, I do not address every problem that these theories face. There may well be other—perhaps even conclusive—reasons to reject various accounts of propositions. If I omit an argument, that should be taken to indicate that I have nothing to contribute to that debate—not to imply that the debate is unimportant. Lastly, while I am primarily concerned with accounts of propositions, some of the arguments have implications that philosophers in other areas may find of interest.

---

[2]My use of 'how' here is both deliberate and significant. I do not state 'propositions differ from *what* we have taken them to be' precisely because many of these debates are neutral on the ontological status of propositions. These arguments have no immediate implications about what sorts of things propositions are—but they do impact the logic of propositional identity.

In particular, I suspect that some may interpret the discussion the following section as a novel argument for monism—one that proceeds on entirely logical grounds.

It may also be helpful to briefly mention a topic that I do *not* address within this paper: the ontological status of propositions. There is an enduring and intractable debate over whether properties and propositions are 'real' (in some important sense of the word). Realists maintain that they are—that, when we write the fundamental book of the world, properties and propositions must somewhere be mentioned—while nominalists maintain that they are not. I write this paper in the context of a literature that is explicitly neutral on this debate. Many of the principles I discuss can be endorsed by realists and nominalists alike. For example, while a realist would naturally interpret '$\exists \lambda X.Xa$' as the claim that there exists some property of a, a nominalist might prefer an interpretation that makes no reference to properties.[3] But while the realist and nominalist diverge over how this sentence is to be interpreted, they may nevertheless agree with one another that it is true. As such, we can investigate the logic that governs these sentences without taking a stand on the debate between realism and nominalism.[4] It is notable that we can make significant progress on problems for propositions while remaining neutral on whether or not they are real.

Structure and $\beta\eta$-Conversion

Structuralists hold that propositions and properties are composed of worldly items in much the way that sentences and clauses are composed of words.[5] Propositions that are composed of different material are distinct—even if they have the same truth value in every possible world. For this reason, structured accounts allow for many distinctions that coarse-grained accounts do not. For example, the proposition that *grass is green or not green* is composed of material concerning grass, while the proposition that *the sky is blue or not blue* is composed of material concerning the sky.[6] Due to their differing compositions, the propositions are distinct—their necessarily identical truth-values notwithstanding. Structured accounts are thus appealing to those who would use propositions to make fine-grained distinctions.

One of the central tenets of structured propositions is the Principle of Singular Extraction (the PSE):

---

[3]The other alternative for the nominalist is to hold the interpretation of '$\lambda X.X$' fixed and to alter the interpretation of '$\exists$'.

[4]Even the possible-worlds conception of propositions (according to which a proposition is the set of worlds in which it is true) is available both to realists and to nominalists. While a realist may take possible worlds to be concrete objects like the one we occupy, a nominalist may take them to be maximally specific (and consistent) fictions.

[5]This sort of view is often attributed to Russell (1903)—though its roots can be traced back at least as far as Frege (1892). Over the years, proponents have become too numerous to discuss in any satisfactory manner here.

[6]On some versions, the components of these propositions are literally grass and the sky—rather than material concerning grass and material concerning the sky. This distinction is unimportant for what follows.

$$Fa = Gb \rightarrow F = G$$

If the proposition that Fa is identical to the proposition that Gb, then property F is identical to property G. This tenet can be seen as arising from the path-dependent nature of propositional construction. If F and G are distinct, then the propositions that they figure in are distinct, precisely because those propositions are 'built' out of different material. So, for example, if the proposition that *Jane is a sister* is identical to the proposition that *Jane is a female sibling*, then the property of *being a sister* is identical to the property of *being a female sibling*; had those properties been distinct, then the propositions they occur within would be distinct as well. In this way, the structured account allows for the extraction of the singular properties that occur within propositions.

Recently, Dorr (2016) has argued against the PSE. He holds that it violates a standard principle of higher-order logic concerning $\beta\eta$-conversion. Care is required in stating the conflict—as closely related principles can easily become confused. One principle is that of entailment; if ψ is the $\beta\eta$-reduction of ϕ, then ϕ entails that ψ. For example, because Fa is the β-reduction of $\lambda x.Fx(a)$, the latter entails the former. This inference is relatively uncontroversial (at least in contexts that lack opaque predicates like 'believes').[7] A more controversial—but nevertheless orthodox—assumption is that $\beta\eta$ equivalent expressions denote the same thing (a principle I dub 'β-identification'). That is, not only may one infer Fa from $\lambda x.Fx(a)$, but the two denote the very same proposition.[8] Dorr claims (falsely, in my view) that the combination of β-identification with the PSE has an untenable result: properties with different extensions are identical.

Generating the problem precisely requires some amount of formalism. Let us adopt a simple typed higher-order language. This language has two basic types—e and t—for entities and sentences respectively. Additionally, for any types τ_1 and $\tau_2 \neq e$, $(\tau_1 \rightarrow \tau_2)$ is a type. Nothing else is a type. Within this language, we can identify first-order monadic predicates as terms of type $(e \rightarrow t)$. They are functions from entities to sentences. In particular, their output is the sentence which predicates the relevant property of the input. So, for example, 'is tall' is to be treated as a function which generates sentences like 'Jones is tall,' 'Smith is tall,' etc. First-order dyadic predicates are to be identified with terms of type $(e \rightarrow (e \rightarrow t))$, second-order monadic predicates are to be identified with terms of type $((e \rightarrow t) \rightarrow t)$, etc. The $\neg$ operator is a term of type $(t \rightarrow t)$, and the binary connectives $\wedge, \vee, \rightarrow, \leftrightarrow$ are all of type $(t \rightarrow (t \rightarrow t))$. We allow for infinitely many variables of every type, as well as the corresponding λ-abstracts needed to bind them. We also introduce the quantifiers $\exists$ and $\forall$—of type $((\tau \rightarrow t) \rightarrow t)$ for every type τ. In first-order languages, these quantifiers perform dual functions. They serve both to express generality

---

[7]Perhaps 'John believes that Fa' is true while 'John believes that $\lambda x.Fx(a)$' is false because John is ignorant about λ abstraction. I note, however, that even systems with opaque predicates often preserve identity through $\beta\eta$-conversion. See, e.g., Caie, Goodman and Lederman (2020).

[8]A bit more precisely, Dorr maintains that propositional identity is preserved through *nonvacuous* $\beta\eta$-conversion—but this modification is unimportant for our purposes.

and to bind variables occurring within their scope. However, in higher-order contexts, these tasks are divided. While quantifiers serve to express generality, the task of variable binding is accomplished solely with λ-abstraction. Additionally, for every type τ we introduce a predicate $=$ of type $(\tau \to (\tau \to t))$ that is used to express identity. That is, the sentence $\ulcorner A^\tau =^{(\tau \to (\tau \to t))} B^\tau \urcorner$ is to be read as 'A is identical to B'.[9]

The putative conflict is generated in the following way. Select a binary relation R; for the sake of concreteness, let R be the relation of *being the same height as*; two people stand in this relation to one another just in case they are equally tall. β-identification entails:

$$\lambda x.Rxx(a) = \lambda x.Rxa(a)$$

Because the term on the right is the β-conversion of the term on the left, the two denote the same proposition. In natural language, this might be read as asserting that the proposition *Mary is the same height as herself* is identical to the proposition *Mary is the same height as Mary.* Given the PSE, the properties that figure within these propositions must also be the same. We may thus infer:

$$\lambda x.Rxx = \lambda x.Rxa$$

The property of *being the same height as oneself* is identical to the property of *being the same height as Mary.* But, intuitively, these properties are not even coextensive. After all, Jane is presumably the same height as herself—but she need not be the same height as Mary. For this reason, those who accept both the PSE and β-identification must hold that properties that differ in their extensions are identical: a truly untenable result.

So the standard argument goes. The mistake occurs near the end—in the assumption that Jane may fail to fall within the extension of *being the same height as Mary.* But Jane *must* fall within the extension of this property for one simple reason: Jane is identical to Mary, so the two bear all of the same properties.[10] This has nothing to do with the selection of Mary and Jane in particular; all objects fall within the extension of this property, because all objects are identical. There is only one object *to* fall in the extension of properties, and so no properties differ with respect to the objects that fall in their extensions.[11] Or, at least, this is what the structuralist ought to say.[12]

---

[9] Note that I shift freely between infix and prefix notation within this paper. In what follows, I occasionally omit the types of terms if it is contextually evident.

[10] This, of course, relies on the assumption that Leibniz's Law is true. I set aside issues arising from opacity for the purposes of this paper.

[11] Perhaps some worry that it is conceivable for properties to differ in their extension—in that F may have a within its extension while the extension of G may be empty. As I discuss below, the higher-order framework we are operating in precludes this possibility. Just as there is only one object, so too there is only one property: *being identical to a.* This property does indeed have a within its extension.

[12] The 'ought' here is meant in the following sense: s ought to say that p if s's commitments logically entail that p.

We can demonstrate that a unique object exists in the following way. β-identification entails:

$$\lambda x.(x = x)(a) = \lambda x.(x = a)(a)$$

As before, the term on the right is the β-conversion of the term on the left. In natural language, this might be read as the claim that the proposition *a is identical to itself* is identical to the proposition *a is identical to a*. Given the PSE, we may infer that the properties occurring within these propositions are identical; i.e.:

$$\lambda x.(x = x) = \lambda x.(x = a)$$

The property *is self identical* is identical to the property *is identical to a*. But, from classical logic, we have:

$$\forall \lambda x.(x = x)$$

All objects are identical to themselves; i.e., all objects fall within the extension of *is self identical*. Because all objects are self-identical, and because *is self identical* is, itself, identical to *is identical to a*, an application of Leibniz's Law results in:[13]

$$\forall \lambda x.(x = a)$$

All object are identical to a. Therefore, there exists only a single object, and monism is true.

Dorr's argument relied upon the existence of distinct objects. But accepting PSE and β-identification already forces the acceptance of monism. And because there is only one object *to* fall within the extension of properties, identical properties do not differ in their extensions; a falls within the extension of both $\lambda x.Rxx$ and $\lambda x.Rxa$—and there exists no other object that falls within the extension of one property but not the other.

As previously indicated, I do not take this point to resurrect the structured-proposition view. I have no doubt that many philosophers view monism with nearly as much incredulity

---

[13]A brief note on my appeal to Leibniz's Law. This principle is sometimes treated as object-restricted—i.e., as the claim that identical objects bear all of the same properties. This version of Leibniz's Law has no immediate implications for the derivation of monism, as I am concerned with the attributes of identical properties, rather than identical objects. In higher-order contexts like this, it is common to appeal too the following substitution principle instead (see, e.g., Caie, Goodman and Lederman (2020); Elgin (Forthcoming)):

$$\alpha = \beta \rightarrow (\phi \leftrightarrow \phi[\alpha/\beta])$$

That is, if α is identical to β then a sentence ϕ is true iff ϕ is true in which occurrences of α are substituted for occurrences of β. This guarantees that identical objects bear all of the same properties (because it permits the inference from $a = b$ to $Fa \leftrightarrow Fb$), but also stand in the same relations. Moreover, it is a version that applies to identical properties, as well as to identical objects. In particular, it permits the inference from $\lambda x.\phi = \lambda x.\psi$ to $\forall \lambda x.\phi \leftrightarrow \forall \lambda x.\psi$.

as the claim that identical properties have different extensions. It flies in the face of common sense to claim that a plurality does not exist. This being so, the fact that structuralism gives rise to monism may well be seen as a sufficient reason to reject the structured view. Nevertheless, if there is a reason to reject the structured account, it is because it entails monism—not because it entails that properties with different extensions are identical.

The argument is more-or-less complete. The remainder of this section expounds the variety of monism that the PSE and β-identification collectively entail. For a number of views have fallen under this label throughout the history of our discipline. There are substance monists—who count Thales and, on some interpretations, contemporary physicalists, among their ardent supporters. These monists hold that everything is composed of a single substance: water, for Thales—the physical, for physicalists. There are priority monists like Schaffer (2010), who hold that there is a unique fundamental object. Perhaps all objects ontologically depend upon (or are grounded in) a single object. But the most extreme form of monism is undoubtably existence monism. This neither holds that objects are composed of a single substance—nor that all things ontologically depend upon one—but rather that the world consists of a single, solitary object. It ought to be clear that existence (rather than substance or priority) monism is at issue.

Existence monism has had scattered support over the past few millennia; Parmenides, Spinoza and—more recently—Della Rocca (2021) have all given defenses. But while all existence monists agree that only a single object exists, they disagree over the nature of that object. Spinoza held that this object has different aspects or attributes. In some ways, we can conceive of it as having a giant, jello-like structure. It congeals, stretches and coalesces—so as to give rise to the manifest world that we observe. Parmenidean/Della Roccan monism, by contrast, is austere. Not only is there a single object, but it rests: constant, unchanging and uniform. Given that the PSE and β-identification collectively entail existence monism, we have yet to determine what the nature is of the unique object.

In many ways, this monism is closer in spirit to Parmenides'/Della Rocca's variety than it is to Spinoza's. For, just as it is possible to establish that there is but a single object, this higher-order framework enables us to establish that there is but a single property that this object bears: being identical to this object. From β-identification, we have:

$$\lambda X.(X = X)(\lambda x.x = a) = \lambda X.(X = \lambda x.(x = a))(\lambda x.x = a)$$

The proposition $\lambda x.x = a$ *is identical to itself* is, itself, identical to the proposition $\lambda x.x = a$ *is identical to* $\lambda x.x = a$. Given the PSE, this allows us to infer that the second-order properties occurring within these propositions are identical, i.e.:

$$\lambda X.(X = X) = \lambda X.(X = \lambda x.(x = a))$$

The second-order property of *is self-identical* is, itself, identical to the property of *is identical to* $\lambda x.x = a$. But, from classical logic, we have that:

$$\forall \lambda X.(X = X)$$

All properties are identical to themselves. This, when combined with Leibniz's Law, allows us to infer:

$$\forall \lambda X.(X = \lambda x.(x = a))$$

And so, not only is there but a single object, but there is also only a single property: being identical to the sole object. It should be clear that arguments of this structure could be generated for terms of arbitrary type.[14] There is only a single binary relation: identity. There is only a single monadic second-order property: being identical to being identical to a, etc. For this reason, the monism that results does not allow for variation in properties or attributes of the object which exists.

Restricting the PSE avoids this higher-order result. That is, if we take the principle to only license the inference from $Fa = Gb$ to $F = G$—and not from $A^{((e \to t) \to t)}F^{(e \to t)} = B^{((e \to t) \to t)}G^{(e \to t)}$ to $A^{((e \to t) \to t)} = B^{((e \to t) \to t)}$—then it remains possible to prove that a unique object exists, but it becomes impossible to prove that a unique property exists.[15] Those tempted by this restriction thus remain committed to existence monism, but need not accept the Parmenidean/Della Roccan variety of existence monism.

Arguably, the manner in which monism was established also suggests that we can comment on its modal status. If both the PSE and β-identification are principles of logic, then they may deserve the status of axioms within our system.[16] But if they are both to be treated as axioms, then it is provable from axioms that monism is true. The necessitation rule (which is permitted on the weakest standard modal logic, **K**) allows us to conclude that everything provable from axioms is necessarily true; i.e., $\vdash A$ entails $\vdash \Box A$. And so, if we treat PSE and β-identification as axioms, then monism is not only true, but necessarily true; we may infer $\vdash \Box \forall \lambda x.(x = a)$.[17]

I finish this section by discussing a potential objection to this derivation of monism.[18]

---

[14]Note that although it is possible to generate this argument for terms of arbitrary type, it is *not* possible to generalize, and thereby make a claim about terms of every type (at least within this language). Given how our language has been constructed, quantifiers can only make generalizations about terms lower on the hierarchy of types than themselves—so there exists no term that quantifies over all types.

[15]My thanks to Jonathan Schaffer for suggesting a restriction roughly along these lines.

[16]Of course, it remains an option to hold that these principles are both true—but are not principles of logic. Those who do so cannot establish the necessity of monism in this way.

[17]Strictly, we would need to expand the descriptive power of our language to include the modal operators $\Box$ and $\Diamond$ in order to express this sort of claim. I direct those interested in this type of expansion to Dorr, Hawthorne and Yli-Vakkuri (2021), who outline how higher-order modal logic is to be formalized. I also note that, in this text, the authors suggest that there is no interesting work for the category of 'logical truth' (as compared to truth). If nothing else, it strikes me that the category is useful in that it provides a path to demonstrating the necessity of claims along the lines described here. While the logical truths can be added to our axioms—and so the necessitation rule establishes the necessity of their entailments—this is not so for claims that are not logical truths.

[18]My thanks to Jonathan Schaffer for pressing me on this point.

Perhaps some interpret $\forall \lambda x.(x = a)$ as the claim that there is at *most* one object. But, arguably, this principle does nothing to guarantee that there is at least one object: if there were a possible world in which *nothing* existed, then all objects would—vacuously—be identical to object a. But an empty world is not one in which monism is true. So, perhaps this argument guarantees that there are either 0 or 1 objects, but it fails to guarantee that at least object exists.

Even if this were so, it would be of limited value to philosophers subservient to common sense. The number of objects would either be 1 or 0: the options monism or nihilism. This suggestion thus does not allow for the plethora of objects that we ordinarily countenance. But we need not settle for this unintuitive pair of alternatives—for there is an observation tracing at least as far back as Quine (1954) that establishes that classical logic is committed to the existence of at least one object. A standard way of axiomatizing first-order logic includes:

$$a = a$$

as an axiom. From this, one may immediately infer:

$$\exists \lambda x.x = a$$

There exists an object which is identical to a. So it is provable from the axioms that an object exists which is identical to a. For this reason, classical logic is not ontologically neutral; it is committed to the existence of at least one object. For this reason, while the PSE and β-identification collectively entail that there is at most one object, classical logic entails that there is at least one. Therefore, there is exactly one object.

This observation led Quine to reject classical logic. He held that logic ought to be neutral with respect to ontology. Precisely because first order logic is *not* neutral with respect to ontology (in that it precludes the possibility of an empty domain), we ought to reject the ordinary quantifiers, and adopt inclusive quantifiers (which are defined so as to be compatible with empty domains) in their place.

Quine would doubtlessly reject the assumptions that lead to existence monism. The logical assumptions at issue are far from neutral in a manner he would find satisfactory.[19] However, the interesting point is that those who accept classical logic can*not* reject the assumptions that lead to monism on the grounds that logic ought to be ontologically neutral.[20] By accepting classical logic, they *already* accept logic with implications regarding

---

[19]Indeed, the non-neutrality of higher-order logic led Quine to reject the very language in which this argument is presented. The fact that one may infer $\exists \lambda X.Xa$ from Fa takes a stand on the existence of properties. However, see Dorr (2016); Correia and Skiles (2019) for arguments that higher-order logic can be made compatible with nominalism.

[20]Conversationally, some have suggested that logic's commitment to at least one object is 'neutral enough,' while its commitment to at most one object is not. This suggests a notion of gradable neutrality, with some cutoff point regarding how neutral it is permissible for logic to be. I do not personally see how to precisify this or turn it into an argument concerning monism via logic.

ontology, so they lack the standing to object to a logic on those grounds. Moreover, in the post-Williamson (2013) era that we live in, there are many who maintain that what makes a system logic is its generality—not its neutrality.

The upshot is this: the PSE and β-identification collectively entail existence monism. Because of this, these principles do not entail that properties with different extensions are identical. Moreover, given the higher-order framework in which we operate, this monism holds for terms of arbitrary type. And if we are to treat these principles as axioms, then this monism is not contingent, but is necessarily true. Those who reject this form of monism ought to reject either the PSE or the claim that identity is preserved through $\beta\eta$-conversion.

Only Logical Circles

In a sense, all identifications are circular. In just the way that 'John is John' expresses something trivial yet true, so too 'To be just is to be just' expresses something trivial yet true. Identity is a reflexive relation—and reflexivity involves circularity. To be sure, there may also be a reading of sentences of form 'To be F is to be G' that is irreflexive.[21] Perhaps these sentences can be naturally read as definitions—rather than identifications— and definition is a directional (and irreflexive) relation. But interest in one reading does not preclude the existence of another. And it is uncontroversial that once we target the reading concerning propositional/property identity, the resulting sentences will be reflexive.

But there seems to be another sense in which these sentences ought not be circular. Quite plausibly, 'To be grue is to be green and observed before time t or not so observed and bleen' is true while 'To be green is to be grue and observed before time t or not so observed and bleen' is false. As Goodman (1955) noted, this identification of green is extensionally adequate; every green object is either grue and observed before time t or not so observed and bleen. Nevertheless, there in some sense in which the identification of green in terms of grue seems defective. Perhaps what it is to be green is to induce a certain phenomenal experience—or to reflect particular wavelengths of light, but it is not plausible that it is to be identified in terms of a property like grue. And if the reason that this identification is problematic is that it is circular (in that it allows grue to be identified in terms of green and green in terms of grue) then there must be an objectionable notion of circular identification.

To account for the distinction between vacuous and vicious circularity, Dorr defends the principle Only Logical Circles (OLC)—according to which logical circularity is innocuous, while nonlogical circularity is not. The precise principle is given as follows:

"$A \equiv_\tau B \to Logical_\Omega(C)$ where Ω and τ are any types, and A, B, and C are any terms, of type τ, τ and Ω respectively, such that B contains an occurrence

---

[21]See Cameron (2014) for a critique of the Rayo (2013) discussion of higher-order identification along the lines that it is not directional.

of A together with an occurrence of C that neither contains, nor is identical to, nor is contained by that occurrence of A, and none of the variables free in A or C are bound in either of these occurrences."—(Dorr, 2016, pg. 74)

This principle appears somewhat cumbersome; it can be represented more perspicuously as:

$$x \equiv_\tau \lambda v_1...v_n.y(z, x, v_1, ..., v_n) \rightarrow Logical_\Omega(z)$$

Only Logical Circles has broad implications for propositional identity. One of its most surprising is that no proposition is its own self-conjunction or self-disjunction. That is to say, p is a different proposition from $p \wedge p$ and $p \vee p$ (which may be described as rejecting idempotence). OLC requires that, in a circular identification (i.e., an identification where the same term appears on both sides of the identity sign), all other terms must be logical terms. Given β-identification, $p \wedge p$ is identical to $\lambda qr.(q \wedge r)(p, p)$; and so $p \equiv p \wedge p$ is equivalent to $p \equiv \lambda qr.(q \wedge r)(p, p)$. OLC would then entail that $Logical(p)$—that p is itself a logical term. But p need not be a logical term, and so idempotence is false.

The implications of OLC surpass the rejection of idempotence. It also forces the rejection of dissolution (the claim that p is identical to $p \wedge (q \vee \neg q)$), absorption (the claims that p is identical to $p \wedge (p \vee q)$ and that p is identical to $p \vee (p \wedge q)$), and the conjunction of the two distribution principles (the claims that $p \wedge (q \vee r)$ is identical to $(p \wedge q) \vee (p \wedge r)$ and that $p \vee (q \wedge r)$ is identical to $(p \vee q) \wedge (p \wedge r)$). Given the significance of these implications, it is incumbent to determine whether OLC performs the theoretical work that it ought to.

Dorr holds that the primary reason to endorse OLC is that allows for reflexive identity while precluding the identification of green in terms of grue. It is perfectly compatible with this principle that p is identical to p. But, he claims, "being green is strictly prior to being grue, since grue $\equiv \lambda x.(\lambda Fy.((Fy \wedge Oy) \vee (By \wedge \neg Oy))(green, x))$, and being green is not logical."(Dorr, 2016, pg. 80).[22]

But OLC does not itself establish the priority of green over grue. True, if it is the case that grue is so identified in terms of green, then OLC entails that green is not identified in terms of grue—precisely because that would allow nonlogical circularity. But it is equally true that if green is identified in terms of grue, then OLC entails that grue is not identified in terms of green. What the principle actually establishes is that the two are not mutually defined—it fails to establish which of the two is prior.[23] Some other principle is required.

There are two obvious ways to augment OLC to achieve the desired result: one suggested by the Lewis (1983) theory of relative naturalness, and the other by the Goodman

---

[22]I have omitted the types that Dorr provides—primarily because the typed framework he operates within is slightly different from the one I defined in the previous section. Additionally, I have corrected what I believe to be typos in the formalism. The actual description was "grue $\equiv (\lambda x.(\lambda f^{<e>} y^e.(fy \wedge Oy) \vee (By \wedge \neg Oy))(green, x)$"

[23]Note that this is a marked contrast from Goodman, who claimed that the two are mutually defined—presumably on the grounds that they are classically equivalent.

(1955)/Dasgupta (2018) theory of entrenchment. A predicate 'F' is said to be more natural than a predicate 'G' just in case it carves nature more closely at the joints than G does.[24] In contrast, a predicate 'F' is said to be more entrenched within a linguistic community than a predicate 'G' just in case it is more commonly used inferentially within that community. Let us say that a pair of predicates F, G are 'circularity-excluded' if OLC entails that they are not identified in terms of one another. 'Green' and 'grue,' on this account, are circularity-excluded. On the Lewisian approach, we might say that predicate F is 'prior' to predicate G just in case F and G are circularity-excluded and F is more natural than G is, while on the Goodmanean approach, we might say that F is prior to G just in case the two are circularity-excluded and F is more entrenched than G is. Given the assumptions that 'green' is both more natural and more entrenched then 'grue,' each will rule that green is strictly prior to grue. But what resources are there (if any) if we don't want to appeal either to entrenchment or naturalness?

We might make some progress if we accept certain principles about stipulative identification. Suppose that language L does not contain a predicate F (for simplicity's sake, we may let F be a monadic first-order predicate) that we would like to introduce. Permissivists hold that we may introduce any grammatically correct expression we would like. That is to say, if $\lambda x.\alpha$ is a grammatically correct expression in L with no free variables, then we are free to stipulate that $F := \lambda x.\alpha$.[25] After this stipulation, within L it holds that $F \equiv \lambda x.\alpha$; to be F is such to be an x such that α.

This might seem to permissive; we could restrict this account to cases where the variable x occurs somewhere within α—in order to avoid identifications like 'To be F is to be being such that grass is green' and in order to avoid complications arising from vacuous β-conversion.[26] Moreover, we may need to supplement this account with some sort of meta-theory of identification if this account of stipulative definition were to come into conflict with another account of identification.[27]

What Goodman did in *Fact, Fiction and Forecast* was to introduce a predicate that did not already exist within natural language; he provided a stipulative identification of 'grue'—which was a predicate which did not previously occur in English (or any other natural language). Given the permissive principle I am attracted to, this is perfectly fine;

---

[24]Naturalness, for Lewis, is a primitive relation that performs extensive explanatory work. I direct those interested in the details of this work to Lewis's original paper—and to Dorr and Hawthorne (2013), who argue (quite convincingly, in my view) that nothing could perform all of the work that Lewis claims.

[25]Note that x itself is bound by the λ-abstract, so any occurrences of x within α will not be free. Also note that this principle of stipulative identification is *not* the naïve schema: for any x, x has the property of being such that P is true iff P is true. This schema is quickly leads to inconsistency. Given the typed framework we are operating in—and the restriction to grammatically correct expressions within this language—the risk of contradiction for this sort of principle is minimal.

[26]See, again Dorr (2016) for a discussion of what some of these complications are.

[27]Here, I am thinking in particular about Lewis (1970)'s paper 'How to Define Theoretical Terms'—where he suggests identifying a theoretical predicate with the function it plays within that theory. In that paper, he seems to suggest that a predicate's stipulative identification 'trumps' its expanded postulate when the two conflict, which is one type of meta-theory.

'grue' means precisely what he stipulated it means. To be grue is to be green and observed before time t or not so observed and blue. But—given that this stipulative definition is correct—OLC then rules out the identification of green in terms of grue. The only way to *reject* this order of identification (while embracing OLC) is to claim that the principle of stipulative definition I endorsed is false. We may stipulatively introduce predicates into our language, but the identification of that predicate is not given by its stipulative definition. I find this hard to accept.

One concern is that my account of stipulative definition rules out the possibility that green is to be identified in terms of grue—but perhaps there is some other predicate F that is necessarily coextensive with 'grue' such that to be green is to be F and observed before time t or not so observed and G (where 'G' is an analogous predicate for 'being bleen'). Since F has not been introduced stipulatively into our language, my account of stipulative identification doesn't rule this possibility out—but surely it is just as unintuitive as identifying green in terms of grue. This is fair enough—but I would like to know what predicate F is. It cannot be the predicate 'grue' that Goodman introduced stipulatively for the reasons already discussed—it must be a predicate not stipulatively introduced in terms of green. I cannot think of what this predicate would be—nor is it one that anyone can stipulatively introduce.[28]

But the upshot is this: Only Logical Circles does not itself determine that green is prior to grue. Rather, it precludes the possibility that each is identified in terms of the other. In order to achieve the claim of relative priority, some further principle is required. Two possibilities, suggested by the literature, are given in terms of entrenchment and relative naturalness. However, I believe that an adequate account of stipulative identification will work as well. Furthermore, once we recognize that OLC must be supplemented with some other account—whether that be one in terms of relative naturalness, entrenchment, or stipulative identification—in order privilege green over grue, it's not entirely clear how much theoretical work Dorr's principle actually accomplishes. These theories were introduced precisely in order to secure the priority of green: and to do so without appeal to a principle like Only Logical Circles. These principles themselves do not have the widespread implications regarding the logic of identification that OLC does (like the rejection of idempotence, absorption, etc.). If appealing to them suffices to resolve the dilemma concerning priority, it is not obvious that we ought to accept OLC and the controversial implications that it has.

---

[28]Another possible response is that I am much closer to Goodman than I am to Lewis given that I rely upon the temporal order of stipulative definition in order to achieve the right result—which seems to depend upon sociological factors. This may be so—but I do not find it to be an objection (and if we think of a Goodman-theory as one which depends upon sociological factors and a Lewis-theory as one that does not, any theory will fall into one camp or the other).

Kaplan's Paradox

Kaplan (1995) raised an interesting puzzle for propositions—one that initially targets possible-worlds accounts.[29] These accounts maintain that propositions are to be identified with the sets of possible worlds in which they are true. So, for example, the proposition that *grass is green* is to be identified with the set of worlds in which grass is green, and the proposition *grass is not green* is to be identified with the set of worlds in which grass is not green. It is relatively straightforward to expand this sort of theory to accounts of properties and relations. Properties, for example, can be identified with functions from possible worlds to sets of objects—intuitively, with those objects that bear that property within that world. So, for example, *is red* is to be identified with a function that takes a possible world as its input, and has—as its output—the set of objects within that world that are red.

The paradox can be generated in the following way. Let us denote the set of possible worlds as P, and the cardinality of this set—whatever it may be—is C.[30] On the possible-worlds conception of propositions, every collection of possible worlds constitutes a proposition; that is to say, every subset of P can be identified with a proposition. For this reason, the set of propositions is the powerset of P. Select an arbitrary person at an arbitrary time. Intuitively, for each proposition, that person could have been thinking a thought whose content was that proposition (and no other). Because this is possible for each proposition, there exists a possible world where each proposition is uniquely thought of (by that person at that time). Therefore, it must be the case that there are at least as many possible worlds as there are propositions. So, the cardinality of P is greater than or equal to the cardinality of the powerset of P. But Cantor's Theorem entails that the cardinality of P is strictly lower than the cardinality of the powerset of P. Contradiction.

This argument is sometimes taken to count against the possible-worlds conception of propositions. The canonical response to this problem is given by Lewis (1986). Lewis denies the intuition that it is possible for each proposition to be uniquely considered (by a person at a time). This denial is not due to peculiarities of human consciousness—the fact that our human brains are incapable of thinking of about propositions concerning numbers of a certain size, for example. After all, it is readily possible for there to exist creatures with minds that surpass our own, and could be capable of considering numbers larger than we are capable of considering. It would be straightforward to generate the puzzle for these creatures as well, and so our response ought not be peculiar to human brains.

Lewis adopts a broadly functionalist conception of thought. What it is for a person

---

[29]This view is defended most famously by Lewis (1986). Kaplan's Paradox itself was first described in print by Davies (1981). Conversationally, some have expressed skepticism that this puzzle genuinely counts as a paradox—but rather indicates that the possible-worlds account of propositions has unintuitive implications. I myself am unsure of what criteria a puzzle need satisfy in order to count as a paradox, but I will continue to refer to it as 'Kaplan's Paradox' due to the use of this name in the literature.

[30]Lewis himself claims that beth-three is the lowest reasonable limit of this cardinality. I personally suspect that the cardinality is much higher than that.

(or creature with superhuman mental capacities) to entertain a proposition in thought is for them to be in a certain mental state which performs a particular functional role. The functional role at issue is given by the causal relation that that state has on the thinker's sensory input and behavioral output. So, for example, the reason why a person who entertains the proposition that *roses are red* differs from a person who entertains the proposition that *violets are blue* is due to the different causal roles that the states have for the respective thinkers. For example, the first may cause someone to assent to the question 'Are some flowers red?' while the latter may not perform this causal function.[31] Lewis holds that there are fewer functional roles than there are propositions. Of course, he cannot tell you which propositions there are which lack an appropriate functional role: but some or other such propositions exist. For this reason, we ought to accept that there are some propositions that cannot be the object of thought.

There is a reason to suspect that the metaphysics of thought is something of a red herring. Although Kaplan's Paradox is stated in terms of thinking of a unique proposition, structurally analogous problems can be given for any relation that humans (or, indeed, anything) stands in with respect to each proposition. For example, some might suspect that a rich enough language would have a different symbol for every proposition. Plausibly, there exists a possible world in which all propositions is transcribed: one where each of these symbols is written down. But although it seems possible for every proposition to be transcribed, there are some propositions for which it is not possible for them to be uniquely transcribed. For, if every proposition could have been uniquely written down, there would exist a mapping from every element of the powerset of propositions to a unique proposition. In this way, we arrive at a semantic version of Kaplan's Paradox—one where an appeal to functionalist conceptions of thought offer no obvious remedy.[32] More broadly, Kaplan's paradox shows that all instances of the following schema are false on the possible-worlds account of propositions:[33]

$$\forall p \Diamond \forall s (Fs \leftrightarrow p = s)$$

Because it is possible to witness this schema with no reference to thought, we might reasonably suspect that an adequate solution will not depend on particular conceptions of what it is to think of a proposition.

---

[31]Lewis largely sidesteps the debate over which version of functionalism is correct. So long as *some* version of functionalism is true—and there are fewer functions than there are propositions—then his move will succeed.

[32]A more general solution is given in Whittle (2009). Whittle argues that our language has sufficient descriptive power to describe its own semantics (in that it is capable of describing possible-worlds semantics). Any language with this expressive power will be incomplete; there will be sentences which grammatically appear to assert something yet are neither true nor false. Perhaps we ought to identify any sentences which give rise to Kaplan's paradox (like 'For any p, it is possible for agent S to be uniquely thinking of p') with these sentences.

[33]This formulation of the paradox follows that given in Anderson (2009).

This paradox is often held to count against the possible-worlds account of propositions (though perhaps not decisively). Minimally, it ensures that this account must reject intuitive judgments regarding our ability to think of unique propositions. Perhaps it motivates rejecting the possible-worlds account in favor of an alternative—one more faithful to our intuitions. It does not. Kaplan's Paradox provides no reason to adopt an alternative account of propositions for a simple reason: it applies to every theory whatsoever. There is no plausible account of propositions that is immune to Kaplan's paradox.

A bit roughly, the problem Kaplan's paradox raises is that there are too many propositions. In fact, there are *so* many propositions that we cannot map each proposition to a unique possible world—which is what prohibits it being possible for each proposition from being uniquely considered. But the possible-worlds conception of propositions forms a lower-bound on the number of propositions that acceptable accounts countenance. Every theory of propositions has at least that many (while some have more). For this reason, every account of propositions countenances sufficiently many that there are some propositions that cannot be uniquely considered. And because every account faces the problem raised by Kaplan's Paradox, the paradox provides no reason to reject one account in favor of another.

Take, for example, structured accounts of propositions. One potential virtue of these accounts is that they can distinguish between propositions that the possible-worlds approach cannot. Structuralists allow for p to be distinct from $p \vee (q \wedge \neg q)$, while coarse grained accounts maintain that they are identical (precisely because they are true in the same possible worlds). In this respect, structuralists countenance *more* propositions than coarse-grained accounts. But the problem Kaplan's paradox raised was that the possible worlds account already had too many propositions (in that there were so many that each could not be mapped to a unique possible world). Countenancing yet more will not alleviate this concern.

A bit more precisely, let us select an arbitrary account of propositions; one that countenances a set of possible worlds W and countenances a set of propositions P. If this account were to avoid Kaplan's Paradox, it must be such that $|P| \leq |W|$: the cardinality of propositions is less than or equal to the cardinality of worlds (as this is what is required for there to exist a possible world where each proposition is uniquely considered). Because this is so, it must be the case that $|P| < |\mathcal{P}(W)|$; the cardinality of propositions is less than the cardinality of the powerset of possible worlds. For this reason, there exists no mapping from each proposition to a unique collection of possible worlds.

Let us define a function *Representation* from propositions to sets of worlds. This function maps a proposition to the set of possible worlds in which that proposition is true.[34] Given the cardinality result, there must exist some sets of worlds that no proposition is represented

---

[34]Note that *representation* is agnostic with respect to the granularity of propositions. If propositions are fine-grained, then it maps distinct propositions to the same sets of worlds (e.g., $p \vee \neg p$ is presumably mapped to the same set of worlds as $q \vee \neg q$). However, if propositions are coarse-grained, then no to propositions are mapped to the same set of worlds.

by. That is, there must exist sets of worlds such that no proposition is true at all (and only) those worlds. Select an arbitrary such set of worlds, and denote this set with w.

The existence of w is untenable. There exists no proposition true at w and false for all other collections of worlds. So, for example, there is no way of describing what all of the elements of w have in common that the other worlds do not. If, for example, it were the case that grass is green in all worlds within w (but no others), then there exists no proposition that grass is green. In fact, it is impossible to make any claims about w that do not also make claims about other sets of worlds. For example, it cannot be that there exists a proposition which asserts that w is distinct from these other sets. But if w really is distinct from other sets of worlds, how could there be no proposition that asserts that w is so distinct?

The problem is that this sort of account has too few propositions. There are truths (such as the truth of the distinctness of w) that no proposition expresses, simply because there are not enough propositions to express all of the truths. At this point, it becomes doubtful that such an account really is an account *of propositions*. After all, one of the core theoretical roles that propositions were intended to perform was to be the ultimate bearers of truth and falsity. Propositions cannot play that role on this view, because there are more truths than propositions can express.

For this reason, every reasonable account of propositions must have more propositions than this: to hold that $|P| \geq |\mathcal{P}|W|$. To the best of my knowledge, all extant theories of propositions *do* hold that $|P| \geq |\mathcal{P}|W|$ (or take no explicit stand on either the cardinality of propositions or worlds). Therefore, all such accounts will face Kaplan's paradox—as they will all be unable to map each element of the set of propositions to a unique world. The result of the paradox is certainly unintuitive, but it provides no reason to abandon the possible worlds account in particular because every account faces a parallel result. Kaplan's Paradox may thus provide a reason to reject the notion of propositions entirely, but it does not provide a reason to adopt one account over another.

Conclusion

This paper has addressed three arguments concerning propositional identity. The first concerned structure and $\beta\eta$ conversion. It has been (falsely) held that these principles collectively entail that identical propositions have different extensions. This is not so— but they do entail that monism is true. The second argument concerned Only Logical Circles. Some have erroneously held that this principle ensures that green is strictly prior to grue. It accomplishes this only with the work of some additional principle—such as a view regarding stipulative identification. Lastly, some have incorrectly held that Kaplan's Paradox provides a reason to reject the possible-worlds account of propositions in favor of an alternative. But every plausible account of propositions faces Kaplan's paradox, so it provides no reason to reject one account in favor of another. But although the ramifications

of these debates have often become confused, they nevertheless raise interesting problems for propositions.

References

Anderson, Anthony. 2009. The Philosophy of David Kaplan. In *The Philosophy of David Kaplan*, ed. Joseph Almog and Paolo Leonardi. Oxford University Press pp. 85–92.

Caie, Michael, Jeremy Goodman and Harvey Lederman. 2020. "Classical Opacity." *Philosophy and Phenomenological Research* 101(3):524–66.

Cameron, Ross. 2014. "On the Lack of Direction in Rayo's *The Construction of Logical Space*." *Inquiry* 57(4):427–41.

Correia, Fabrice and Alexander Skiles. 2019. "Grounding, Essence and Identity." *Philosophy and Phenomenological Research* 3:642–70.

Dasgupta, Shamik. 2018. "Realism and the Absence of Value." *The Philosophical Review* 127(3):279–322.

Davies, Martin. 1981. *Meaning, Quantification, Necessity: Themes from Philosophical Logic*. Routledge and Kegan Paul.

Della Rocca, Michael. 2021. *The Parmenidean Ascent*. Oxford University Press.

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

Dorr, Cian and Johnathan Hawthorne. 2013. Naturalness. In *Oxford Studies in Metaphysics*, ed. Karen Bennett.

Dorr, Cian, Jonathan Hawthorne and Juani Yli-Vakkuri. 2021. *The Bounds of Possibility: Puzzles of Modal Variation*. Oxford University Press.

Elgin, Samuel. Forthcoming. "Definition.".

Frege, Gottlob. 1892. On Concept and Object. In *Translations from the Philosophical Writings of Gottlob Frege*, ed. Peter Geach and Max Black.

Goodman, Nelson. 1955. *Fact, Fiction and Forecast*. Harvard University Press.

Kaplan, David. 1995. A Problem in Possible-World Semantics. In *Modality, Morality and Belief: Essays in Honour of Ruth Barcan Marcus*, ed. Walter Sinnot-Armstrong. Cambridge University Press pp. 41–52.

Lewis, David. 1970. "How to Define Theoretical Terms." *Journal of Philosophy* 67(13):427–46.

Lewis, David. 1983. "New Work for a Theory of Universals." *Australasian Journal of Philosophy* 61(4):343–77.

Lewis, David. 1986. *On the Plurality of Worlds*. Oxford University Press.

Quine, W. V. O. 1954. "Quantification and the Empty Domaine." *The Journal of Symbolic Logic* 19(3):177–9.

Rayo, Augustín. 2013. *The Construction of Logical Space*. Oxford University Press.

Russell, Bertrand. 1903. *Principles of Mathematics*. Norton.

Schaffer, Jonathan. 2010. "Monism: The Priority of the Whole." *The Philosophical Review* 119:31–76.

Whittle, Bruno. 2009. "Epistemically Possible Worlds and Propositions." *Noûs* 43(2):265–85.

Williamson, Timothy. 2013. *Modal Logic as Metaphysics*. Oxford University Press.