TWO THEORIES OF FORM

Abstract. Fine (2017) presents a puzzle of logical form and develops a solution based on his (1985) theory of arbitrary objects. This paper begins by reformulating Fine's puzzle as a "no-go theorem" in the context of abstract algebra. It then explores the space of theories of form in light of this puzzle. Two theories hold particular interest. One is naturally interpreted as a development of Fine's own theory. The other is novel. This paper argues that neither deserves the honorific 'the form of a formula' and that instead we should treat formulas of having many forms. The solution to Fine's puzzle is then achieved by disambiguation.

Introduction

According to one historically popular conception of logic, logic is the science of *logical form*.[1] Despite its historical prominence, explicit *metaphysical* theories of these objects, logical forms, are rarely offered. One recent exception is Fine (2017), who provides a theory of logical form based on his (1985) theory of arbitrary objects.[2]

Fine takes his theory to be motivated by a puzzle of logical form. There are three principles that, each taken on its own, intuitively governs logical forms, but when taken together, look to be unsatisfiable.

> *Existence*: For each formula there is an object that is the form of the formula.
> *Identity*: The forms of two formulas are the same if and only if the formulas are alphabetic variants.
> *Structural Similarity*: The form of a negative formula is the "negation" of a form. . . and the form of a conjunctive formula is the "conjunction" of two forms.

[1]Arguably this view traces back to Kant (see MacFarlane (2000).) It is of course more common to characterize logic as the science of *logical consequence*. But it is also not uncommon to say that logical consequence is consequence that holds in virtue of the form of the argument. In turn the form of an argument is specified in terms of the forms of its constituents.

[2]See also Fine (1998) for relevant background to Fine (2017).

(Fine 2017, p. 515)

Let p and q be propositional variables. From the principle *Existence* it follows that they have forms. And from the principle *Identity* it follows that their forms are the same, since any two propositional variables are alphabetic variants. The formulas $p \wedge p$ and $p \wedge q$ on the other hand are not alphabetic variants. From the principles of *Existence* and *Identity* it follows that they have distinct forms. The problem is now that from the principle *Structural Similarity* it follows that $p \wedge p$ and $p \wedge q$ are each the conjunction of two forms. And there appears to be only one option in each case. The form of $p \wedge p$ is the form of p conjoined with the form p. The form of $p \wedge q$ is the form of p conjoined with the form of q. But the form of p *is* the form of q. So the form of p conjoined with the form of p *is* the form of p conjoined with the form of q. Therefore the form of $p \wedge p$ is the form of $p \wedge q$. We have reached a contradiction.

Fine responds that that the form of $p \wedge p$ is a conjunction of forms, but it is not the form of p conjoined with the form of p. In order to make sense of this proposal he develops a theory of logical forms according to which they are *arbitrary formula*. An arbitrary formula is, roughly, a formula defined by a "let-clause." There is an arbitrary formula that could be any formula: it is the formula defined by the clause "let x be a formula." This is an example of an *independent* arbitrary formula. It is defined without reference to any other arbitrary formulas. Another example of independent arbitrary formula is the arbitrary conjunctive formula, the formula defined by the let-clause "let x be a formula such for for some formulas φ and ψ, $x = \varphi \wedge \psi$." But there are also dependent arbitrary formulas. Where x is the arbitrary conjunctive formula, a dependent arbitrary formula might be defined by the let clause "let y be a left-conjunct of x" or by the clause "let y be a right-conjunct of x." It is these dependent arbitrary formulas that Fine thinks serves as the conjuncts of conjunctive forms.

This paper develops a position on logical form according to which the principle *Identity* fails. We will see that there is a natural alternative to Fine's view that allows one to maintain *Structural Similarity* on its most plausible reading according to which the form of $p \wedge p$ is the form of p conjoined with itself. It may be that there are two distinct but perfectly intelligible notions of logical form. One obeys the principle *Identity* and the other the principle *Structural Similarity*. The resolution of the puzzle is then achieved by disambiguation. On no reading of 'logical form' are all of the principles true (at least on their most plausible interpretations), but each principle is true on some interpretation of 'logical form'.

I want to briefly say something about the context of this paper. Fine's (2017) focus on logical forms is partly to illustrate the role that a theory of arbitrary objects might play in an account of the structure of types more broadly.[3] While I will have less to say about this question, I too want to treat logical forms as a sort of case study by which we might draw some broader lessons. One of my overaraching goals, only gestured at in this paper, is to argue that the algebraic notion of a freely generated object can shed light on the somewhat obscure notion of an arbitrary object. An *arbitrary group*, for instance, might be thought of as a free group in the category of groups; and arbitrary Boolean algebra could be a free Boolean algebra in the category of Boolean algebras. On this conception of arbitrariness, the space of formulas themselves turn out to be a kind of arbitrary object. This raised the question of how the space of forms of formula should be thought to relate to the formulas themselves. The problem, however, is that it is hard to find any sort of natural characterization of the space of forms of formulas on Fine's view. This led me to consider an alternative conception of form that admits a more natural characterization. The contents of this paper lay out that notion of form.

In §1 we formulate Fine's puzzle as a sort of "no-go theorem" in the context of universal algebra. What this theorem shows is that logical forms, as Fine conceives them, cannot have certain features that we might have intuitively supposed them to have. The algebraic perspective on the puzzle leads naturally to an alternative conception of logical form that deviates from the principle of *Identity* just enough to secure *Structural Similarity*, and so provides a conception of form that has some of the features that fine's conception of forms lack. In §2, the alternative conception of form is developed and shown to have several nice

[3] This sort of approach to arbitrary object theory has some similarity with Leon Horsten's recent theory.

features. But it also has some features that are not so nice. In §3 an argument is provided that forms on this alternative conception lack certain features that intuitively forms ought to have. Forms, on Fine's conception, however, do not face this argument. I suggest in light of this that we abandon the idea there is any single notion of *the* form of a formula and instead hold that there are many forms of formulas.

1. Fine's Puzzle

There is a simple theorem underlying Fine's puzzle. We would like sameness of form to be a congruence relation on the space of formulas.[4] The problem is that the relation that two formulas stand in when they are alphabetic variants is not a congruence relation. Our intuitive conception of form is pushing us in two opposite directions. To bring out this perspective on the puzzle, we start with some background.

A *logical algebra* is a triple $(A, \wedge, \neg)$ such that A is a set, $\wedge$ is a binary operation on A and $\neg$ is a unary operation on A. There is no constraint whatsoever on what the elements of the set A are nor on what the operations $\wedge$ and $\neg$ are, other than that they have arities two and one respectively. The fact that we write the operations as $\wedge$ and $\neg$ of course indicates that we are going to be thinking of logical algebras as representing objects over which some notion of conjunction and negation are defined. But notice that there are many logical algebras for which this characterization may be misleading. For instance the algebra $(\mathbb{Z}, +, -)$, where $\mathbb{Z} = \{\ldots, -2, -1, 0, 1, 2, \ldots\}$ are the integers, $+$ maps two integers to their sum and $-$ maps each integer to its additive inverse, is a logical algebra.

Among the logical algebras there are certain distinguished ones—*term algebras*—that we can use to represent formulas of a given propositional language. We define these as follows.

Definition 1.1 (Term Algebra)**.** A *term algebra* is a logical algebra $(A, \wedge, \neg)$ admitting of a map $i : X \to A$ such that

- i, $\wedge$ and $\neg$ are all injections.
- The images of $i, \wedge$ and $\neg$ are all disjoint.

[4]Intuitively a congruence relation on formulas is an equivalence relation that respects the Boolean operations.

- A is the result of closing $i(X)$ under $\wedge$ and $\neg$.

The set X is said to *generate* the algebra $(A, \wedge, \neg)$.[5]

Let $(A, \wedge, \neg)$ be a term algebra generated by the set X. We can view the set X as a set of atomic formula, or propositional variables, and the map $i : X \to A$ as the inclusion of X into the set of formula generated by X. The assumption that $i, \wedge$ and $\neg$ be injections with disjoint images is an algebraic characterization of the familiar property of unique readability. In what follows

One of the questions motivating Fine's puzzle is: how similar are forms of objects, in structure, to the objects of which they are forms. Restricted to the special case of logical forms, we want to investigate the degree to which logical forms resemble their instances. As has become standard, we can study the structural relationships between objects by looking at the behavior of various structure preserving mappings in and out of those objects. Let $\mathcal{A} = (A, \wedge, \neg)$ and $\mathcal{B} = (B, \wedge', \neg')$ be two logical algebras. A *homomorphism* from $(A, \wedge, \neg)$ to $(B, \wedge', \neg')$ is a map $f : A \to B$ such that $f(x \wedge y) = f(x) \wedge' f(y)$ and $f(\neg x) = \neg' f(x)$, for any $x, y \in A$. An *isomorphism* from $\mathcal{A} \to \mathcal{B}$ is a homomorphism $f : \mathcal{A} \to \mathcal{B}$ that is *invertible*: there exists a homomorphism $g : \mathcal{B} \to \mathcal{A}$ such that $g \circ f = 1_{\mathcal{A}}$ and $f \circ g = 1_{\mathcal{B}}$ (where $1_{\mathcal{A}}$ and $1_{\mathcal{B}}$ are the identity maps on $\mathcal{A}$ and $\mathcal{B}$ respectively). An *endomorphism* is a homomorphism from a term algebra to itself. An *automorphism* is an endomorphism that is in addition an isomorphism.

In what follows we will call a *formula* an element of the countably infinite term algebra $\mathcal{F} = (F, \neg, \wedge)$.[6] We will write V for the set that freely generates $\mathcal{F}$ and call elements of V *propositional variables* or simply *variables*. A *substitution* on formulas is an endomorphism $f : \mathcal{F} \to \mathcal{F}$. The following universal property of $\mathcal{F}$ will be useful to appeal to in what follows.

[5]See Bergman (2015, p. 391) for a more general definition and discussion. It is well known that such term algebras *exist*. The proof proceeds basically by constructing explicitly the language with X as a set of propositional variables. See Bergman (2015, p. 392) Theorem 9.3.3.

[6]The assumption of uniqueness is justified by the fact that any two term algebras of the same cardinality are isomorphic.

Proposition 1.2. Let $\mathcal{B} = (B, \neg^B, \wedge^B)$ be a logical algebra. Let $f : V \to B$ be a function. Then there is a unique homomorphism $\bar{f} : \mathcal{F} \to \mathcal{B}$ with the property that

$$\bar{f} \circ i_V = f$$

where $i_V : V \to F$ is the inclusion of V into F.[7]

In particular, given any function $f : V \to F$ there is exactly one substitution $\bar{f} : \mathcal{F} \to \mathcal{F}$ whose restriction to V agrees with f. Hence substitutions on $\mathcal{F}$ correspond one-to-one with functions $f : V \to F$. Thus in order to know how a given substitution acts on formulas, it suffices to know where it sends the propositional variables.footnoteFor further backgroun see Burris (1981) and Bergman (2015).

It is worth pointing out here a relationship between the metaphysical notion of an arbitrary object of a given sort, and the algebraic notion of a term algebra. What proposition 1.2 shows is that, in effect, term algebras are *arbitrary logical algebras* since they can be homomorphically mapped into any logical algebra. Since an arbitrary logical algebra doesn't really have any interesting features, this fact is perhaps not super interesting. But when it comes to more interesting categories of algebras (groups say), there are groups that stand to all other groups as term algebras stand to logical algebras (the free groups).This suggests that in certain special cases, the notion of an arbitrary F can be given a precise mathematical description. I will say a bit more about this point later on.

Following Fine, we will restrict our attention to the forms of formulas of propositional logic (i.e., elements of $\mathcal{F}$). Ultimately a theory of form should of course be extended to account for difference in form between, for instance, formulas with quantifiers and those without. Since Fine's puzzle already arises at the level of propositional logic, including quantifiers would be a needless distraction.

[7]This fact is well known, but for a specific proof see Bergman(2015, p. 391); this proposition is an immediate corollary of Lemma 9.3.2

1.1. **Alphabetic Variants.** Two formulas, Fine tells us, are "*alphabetic variants* if they differ merely in the identity of their sentence letters." (Fine 2017, p. 515) Say that a formula φ is *interpretable* in a formula ψ, written $\varphi \leq \psi$, if there is a substitution f on $\mathcal{F}$ with the property that $f(\varphi) = \psi$. Formulas φ and ψ are *bi-interpretable* if $\varphi \leq \psi$ and $\psi \leq \varphi$. When φ and ψ are bi-interpretable, we write $\varphi \sim \psi$.

Proposition 1.3. $\sim$ is an equivalence relation on F.

Proof. The relation $\sim$ is automatically symmetric. It is reflexive since the identity map is a substitution. And it is transitive since the composite of two substitutions is a substitution. $\square$

It is natural to define alphabetic variants as bi-interpretable formulas. We recognize the formulas $p \wedge q$ and $r \wedge s$ as alphabetic variants because each can be obtained from the other by substitution. The formula $p \wedge p$ and $p \wedge (r \wedge s)$ are not alphabetic variants since, while $p \wedge (r \wedge s)$ can be obtained from $p \wedge p$ by substitution, there is no substitution mapping $p \wedge (r \wedge s)$ to $p \wedge p$.

There is another natural definition of alphabetic variants that may be slightly more intuitive. Say that formula φ is *isomorphic* to a formula ψ if there exists an automorphism of $\mathcal{F}$ that maps φ to ψ. This relation is evidently an equivalence relation. And it is also a plausible account of alphabetic variants. To see this let $Aut(\mathcal{F})$ be the automorphism group of $\mathcal{F}$ and let $Sym(V)$ be the symmetric group of V.[8]

Proposition 1.4. $Aut(\mathcal{F})$ is isomorphic (as a group) to $Sym(V)$.

Proof. First we show that the restriction map

$$f \mapsto f|_V : Aut(\mathcal{F}) \to Sym(V)$$

is a group homomorphism (i.e., the map that sends each $f : \mathcal{F} \to \mathcal{F}$ to the map $f|_V : V \to V$ obtained by restricting f to V). To show that this is well defined, we must show that the

[8]That is, $Aut(\mathcal{F})$ is the set of automorphisms of $\mathcal{F}$ equipped with operation $\circ$ of function composition and $Sym(V)$ is the set of bijections from V to V also equipped with function composition.

range of $f|_V$ is identical to V whenever f is an automorphism. Let $f \in Aut(\mathcal{F})$ and let $p \in V$. Then there is some $\varphi \in F$ such that $f(\varphi) = p$ (since f is an automorphism). If $\varphi = \neg\psi$ or $\varphi = \psi \wedge \chi$ for any formulas ψ or χ, then $p = \neg f(\psi)$ or $p = f(\psi) \wedge f(\chi)$. This contradicts the fact that the images of $\neg$ and $\wedge$ are disjoint from V. It follows that $\varphi \in V$. So for each $p \in V$, there is some $q \in V$ such that $f(q) = p$. Since f is automatically injective on V, it follows that $f|_V \in Sym(V)$. It only remains to show that it respects composition. And this is easy since for any $f, g \in Aut(\mathcal{F})$ and $p \in V$ we have

$$(g \circ f)|_V(p) = (g \circ f)(p) = g(f(p)) = g_V(f|_V(p)) = g|_V \circ f|_V(p)$$

Hence $f \mapsto f|_V$ is a homomorphism of groups. To show that it is an isomorphism, we finds its inverse.

For each $g \in Sym(V)$ let $\bar{g} : \mathcal{F} \to \mathcal{F}$ be the unique homomorphism such that $\bar{g} \circ i = g$. We show that the map

$$g \mapsto \bar{g} : Sym(V) \to Aut(\mathcal{F})$$

is a group homomorphism. Let $p \in V$ and $f, g \in Sym(V)$. Then

$$\begin{aligned}
(\bar{g} \circ \bar{f}) \circ i(p) &= \bar{g} \circ \bar{f}(p) \\
&= (\bar{g} \circ i) \circ (\bar{f} \circ i)(p) \\
&= g \circ f(p)
\end{aligned}$$

So $\bar{f} \circ \bar{g} \circ i = g \circ f$. Since $\overline{g \circ f}$ is the unique substitution with that property, $\overline{g \circ f} = \bar{g} \circ \bar{f}$. For each $g \in Sym(V)$, $\bar{g}$ has an inverse given by $\overline{g^{-1}}$ since

$$\overline{g} \circ \overline{g^{-1}} = \overline{g \circ g^{-1}} = \overline{1_V} = 1_{\mathcal{F}} = \overline{1_V} = \overline{g^{-1} \circ g} = \overline{g^{-1}} \circ \overline{g}$$

So $g \mapsto \bar{g}$ is a homomorphism of groups.

To finish the proof we note that $f \mapsto f|_V$ and $g \mapsto \bar{g}$ are mutually inverse. For $f \in Sym(V)$ and $p \in V$, we have

$$(\bar{f})|_V(p) = \bar{f} \circ i(p) = f(p)$$

Hence $(\bar{f})|_V = f$. And for any $g \in Aut(g)$ and $p \in V$ we have

$$g \circ i(p) = g|_V(p)$$

Hence $g \circ i = g|_V$ and so $\overline{(g|_V)} = g$. □

From this proposition we see that automorphisms of $\mathcal{F}$ are essentially permutations of propositional variables. So if φ and ψ are isomorphic, one can get to ψ from φ by permuting the variables in φ. And this seems to be exactly what alphabetic variation amounts to.

This provides us with two ways of thinking of alphabetic variation: isomorphism and bi-interpretability. But there is only one relation corresponding to these two ways of thinking, since it can be shown that that two notions coincide.

Theorem 1.5. *$\varphi \sim \psi$ if and only if φ and ψ are isomorphic.*

To prove theorem 1.5 it is helpful to first prove a lemma concerning the relationship between substitutions and subformulas. Given that φ and ψ are bi-interpretable, we want to somehow construct an automorphism of formulas that takes φ to ψ. Say that a substitution f *fixes* a formula φ if $f(\varphi) = \varphi$. The key to proving the theorem is that in $\mathcal{F}$, a substitution fixes a formula if and only if it fixes all of its subformulas.

Lemma 1.6. *Let f be a substitution and φ a formula. Then f fixes φ if and only if it fixes every subformula of φ.*

Proof. We argue for the nontrivial direction by induction on φ. The base case $\varphi \in V$ is trivial. Let $\varphi = \neg\psi$ and let f be a substitution. Suppose that $f(\varphi) = \varphi$. Hence $\neg f(\psi) = f(\neg\psi) = \neg\psi$. Since $\neg$ is an injection $f(\psi) = \psi$. Since the subformulas of φ are the subformulas of ψ plus φ itself, it then follows by induction that $f(\chi) = \chi$ for each subforula χ of φ. The conjunctive case follows by analogous reasoning. □

With this lemma in place the theorem is proved as follows.

Proof of theorem. The right to left direction follows immediately from the fact that every automorphism has an inverse. To show the other direction let f and g be substitutions such that $f(\varphi) = \psi$ and $g(\psi) = \varphi$. Then $g \circ f(\varphi) = \varphi$ and so by lemma 1.6, $g \circ f(p) = p$ for each propositional variable p in φ. And so since p is not in the range of $\neg$ or $\wedge$, $f(p)$ must be a propositional variable. Moreover, whenever p and q are distinct propositional variables occurring in φ, $f(p)$ and $f(q)$ are distinct since $g \circ f(p) = p$ and $g \circ f(q) = q$. Define $h : V \to V$ to map p to $f(p)$ if p occurs in φ and to itself otherwise. Then h is a bijection and so from proposition 1.4 there is an automorphism $\bar{h}$ that agrees with h on propositional variables. Where $V(\varphi)$ are the propositional variables occurring in φ, it follows by a trivial induction that $\bar{h}$ and f agree on any formula all of whose propositional variables occur in $V(\varphi)$. In particular it follows that $\bar{h}(\varphi) = f(\varphi) = \psi$. $\square$

Two formulas are isomorphic if and only if they are bi-interpretable. This is evidence that the precise notion of bi-interpretability captures the informal notion of alphabetic variation since we've seen that the two intuitive ways of making alphabetic variation precise coincide.

1.2. **Structural Similarity.** Bi-interpretability provides a precise account alphabetic variants and with it the principle of *Identity*. *Structural Similarity* also has an algebraic formulation. First, we require that the class of forms admit notions of "negation" and "conjunction". In other words, there should be a *logical algebra*, $\mathcal{U} = \langle U, \neg, \wedge \rangle$, where U is the set of forms of formula in F. And second, this logical algebra ought to preserve the structure of the formulas of which they are forms. In other words, there should be a homomorphism $f : \mathcal{F} \to \mathcal{U}$ with the property that $f(\varphi)$ is the form of φ.

It is sometimes assumed that sameness of structure requires the existence of an isomorphism rather than simply a homomorphism. But it would be misguided in the present setting to demand that there be an isomorphism between $\mathcal{F}$ and $\mathcal{H}$ since it would be misguided to demand that distinct formulas get assigned distinct forms. Logical forms are forms shared by many formula. Thus the weaker requirement of a homomorphism seems more appropriate here.

It should be noted though that the requirement of a homomorphism between formulas and their forms goes beyond what *Structural Similarity* actually says. According to structural similarity, the form of, for instance, $\varphi \wedge \psi$ should itself be a conjunction. The requirement of homomorphism says that it must be the conjunction of the forms of φ and ψ. It is this extra condition, we'll see, that leads to inconsistency with other plausible conditions on forms and is dropped in Fine's theory. However it seems to me to be a plausible requirement insofar as *Structural Similarity* is a plausible requirement. We'll return to this point when discussing Fine's view.

So let's suppose that the the requirement of *Structural Similarity* demands a homomorphism from $\mathcal{F}$ to $\mathcal{H}$ and the requirement of identity demands that two formulas have the same form if and only if they are bi-interpretability. Putting these together results in the requirment that there be a homomorphism between $\mathcal{F}$ and $\mathcal{H}$ that sends bi-interpretable formulas to the same element of $\mathcal{H}$.

In general, given a homomorphism $f : \mathcal{A} \to \mathcal{B}$, we call the *kernel* of f, $\ker f$, that binary relation on A that relates a and a' if f maps them to the same element (i.e., $\ker f = \{(a, a') \in A \times A \mid f(a) = f(a')\}$). Stated with this terminology, the joint effect of *Identity* and *Structural Similarity* is that $\sim$ should be the kernel of the homomorphism mapping formulas to their forms.

1.3. **Fine's Puzzle.** Putting all of this together we get the following precise formulation of the Fine's requirements:

> Existence: There is a function $f : F \to H$ that associates each formula $\varphi \in F$ with its form $f(\varphi) \in H$;
>
> Structural Similarity: This function $f : F \to H$ is a homomorphism of logical algebras $f : \mathcal{F} \to \mathcal{H}$;
>
> Identity: The kernel of this function is the relation $\sim$.

The puzzle is then that while these constraints each seem plausible on their own they cannot be jointly satisfied:

Theorem 1.7. *There does not exist a logical algebra $\mathcal{B}$ and homomorphism $f : \mathcal{F} \to \mathcal{B}$ with the property that* $\sim = \ker f$.[9]

Proof. Suppose otherwise. Let φ and ψ be distinct formulas such that $\varphi \sim \psi$. Then $f(\varphi) = f(\psi)$ and so $f(\varphi) \wedge f(\varphi) = f(\varphi) \wedge f(\psi)$. By assumption f is a homomorphism. Therefore $f(\varphi \wedge \varphi) = f(\varphi \wedge \psi)$. Since $\sim = \ker f$, $\varphi \wedge \varphi \sim \varphi \wedge \psi$. By theorem 1.5 there is an automorphism $g \in Aut(\mathcal{F})$ with the property that $g(\varphi \wedge \varphi) = \varphi \wedge \psi$. So $g(\varphi) \wedge g(\varphi) = \varphi \wedge \psi$. Since $\wedge$ is an injection, $g(\varphi) = \varphi$ and $g(\varphi) = \psi$. Therefore $\varphi = \psi$. Contradiction. $\square$

Given theorem 1.7, there is simply no way to construct a consistent theory of logical form satisfying the requirements of Existence, Identity and Structural Similarity. If sameness of form coincides with bi-interpretability, then the space of forms is rather unlike the space of formulas of which they are forms.

But this is puzzling. We often speak as if forms had linguistic structure. We say that $p \wedge q$ has a conjunctive form whereas $\neg p$ does not. What is it to have a conjunctive form if not to be a form that is a conjunction? In the next section we'll explore views that reject *Identity* and preserve *Structural Similarity*, finding one such alternative particularly compelling as an account of form.

2. Denying Identity

Theorem 1.7 shows that if the forms of formula have a similar structure to the formulas themselves, then it is false that all and only bi-interpretable formulas have the same form. The challenge for the Identity denying theorist of logical form is to construct a theory of sameness of form that replaces bi-interpretability with a relation that allows for Structural Similarity to be satisfied.

[9]The proof of this is almost immediate given the well known fact that $\sim = \ker f$ only if $\sim$ is a congruence. I will introduce the concept of congruence in the next section.

There are three kinds of relations that might act as replacements for bi-interpretability. A *liberal* theory holds that all bi-interpretable formulas have the same form, but in addition some formulas have the same form without being bi-interpretable. A *conservative* theory holds that formulas have the same form only if they are bi-interpretable, but there are some bi-interpretable formulas that differ in form. Lastly an *orthogonal* theory holds that bi-interpretability is neither necessary nor sufficient for sameness of form.

Conservative theories can be shown to be non-starters. The only relation contained in $\sim$ that can consistently be combined with Structural Similarity is the identity relation (this is made precise and demonstrated below). However there is an interesting liberal theory that is far from trivial. The main goal of this section is to develop this theory.

In order to develop this view more precisely we start with the following definition. Let $\mathcal{A} = (A, \wedge, \neg)$ be a logical algebra. A *congruence* on $\mathcal{A}$ is an equivalence relation on A such that for any $a_1, a_2, b_1, b_2 \in A$, if $a_1 \equiv a_2$ and $b_1 \equiv b_2$ then $\neg a_1 \equiv \neg a_2$ and $a_1 \wedge b_1 \equiv a_2 \wedge b_2$. The following is well known.

Proposition 2.1. Let $\mathcal{A} = (A, \wedge, \neg)$ be a logical algebra and $\equiv$ an equivalence relation on A. Then $\equiv$ is a congruence if and only if it is the kernel of a homomorphism $f : \mathcal{A} \to \mathcal{B}$ for some logical algebra $\mathcal{B}$.[10]

A consequence of proposition 2.1 is that if Structural Similarity holds, the relation of sameness of form is a congruence relation. The goal is then to find a plausible conception of form according to which it is a congruence.

2.1. **Conservative theories of form.** A theory is conservative if sameness of form is included in $\sim$. In order for the conservative to maintain Structural Similarity they must find some congruence relation $\approx$ contained in $\sim$. Unfortunately there is no plausible candidate.

Theorem 2.2. *The only congruence contained in $\sim$ is the identity relation*

[10]See for instance Burris and Sankappanavar (1981).

Proof. Let $\approx$ be a congruence relation contained in $\sim$ and suppose that $\varphi \approx \psi$. Since $\approx$ is an equivalence relation, we have $\varphi \approx \varphi$. And so since $\approx$ is a congruence relation

$$(\varphi \wedge \varphi) \approx (\varphi \wedge \psi)$$

By assumption $\approx \subseteq \sim$. Therefore

$$(\varphi \wedge \varphi) \sim (\varphi \wedge \psi)$$

By theorem 2.5 there is an automorphism f such that

$$f(\varphi) \wedge f(\varphi) = f(\varphi \wedge \varphi) = \varphi \wedge \psi$$

Since $\wedge$ is injective, $\varphi = f(\varphi) = \psi$. So $\varphi = \psi$. $\square$

Clearly any conception of logical form according to which φ is identical to its own form is inadequate. If we want a more plausible *identity* denying theory, we have to look elsewhere.

2.2. **Liberal theories of form.** The conservative theory collapses form and identity and so can be safely ignored. However there is a plausible liberal theory that can consistently be combined with *Structural Similarity* on its most plausible reading while maintaining quite a bit of intuitive appeal. The basic idea behind this theory is that two formulas should be regarded as having the same form if, ignoring repeated occurrences of propositional variables, the formulas are isomorphic. So for instance, the formulas $p \wedge p$ and $p \wedge q$ will have the same form, according to this theory, because if we ignore the fact that $p \wedge p$ has repeated occurrences of the propositonal variable p, say by replacing one occurrence of p with an occurrence of a distinct propositional variable r to get the formula $p \wedge r$, the formulas are isomorphic.

We'll say that a formula is *differentiated* if no propositional variable occurs more than once within it.[11] So, for instance, whenever p and q are distinct propositional variables, $p \wedge q$ is

[11]For each propositional formula $p \in V$ we define a map $\#_p : \mathcal{F} \to \mathbb{N}$ that keeps track of p's occurrences in formulas by induction. In particular $\#_p$ is the smallest function such that

(i) $\#_p(p) = q$ and $\#_p(q) = 0$ (for $q \neq p$).
(ii) $\#_p(\neg\varphi) = \#_p(\varphi)$.
(iii) $\#_p(\varphi \wedge \psi) = \#_p\varphi + \#_p\psi$.

differentiated and $p \wedge p$ is not. The idea of "ignoring" repeated occurrences of propositional variables can be explained in terms of associating with each formula φ a differentiated formula φ^* for which there exists a proper sort of substitution taking φ^* to φ. Intuitively this should be a substitution that preserves all of the logical structure of φ^* except perhaps the number of occurrences of propositional variables. Let $f : \mathcal{F} \to \mathcal{F}$ be a substitution. Then f is *simple* if the image of V is contained in V (i.e., $f(V) \subseteq V$)). Simple substitutions map propositional variables to propositional variables. Every automorphism is automatically simple but the converse is false. The substitution that maps every propositional variable to one propositional variable is simple, but not invertible.

With all of this in place we come to the main definition. The intuitive description of the relation we are after suggests that two formula φ and ψ will stand in this relation if there are isomorphic differentiated formula φ^* and ψ^* together with simple substitutions taking φ^* and ψ^* to φ and ψ respectively. However note that if f is an automorphism taking φ^* to ψ^* and g is a simple substitution taking ψ^* to ψ, then $f \circ g$ is a simple substitution taking φ^* to ψ. Thus we can require the superficially stronger condition that φ^* and ψ^* actually be identical, not only isomorphic. This leads to the following definition.

Definition 2.3 (Relative Bi-Interpretability)**.** Formulas φ and ψ are *relatively bi-interpretable*, $\varphi \equiv \psi$, if there is a differentiated formula χ and a pair of simple substitutions f and g such that $f(\chi) = \varphi$ and $g(\chi) = \psi$.

We can picture the relative bi-interpretability of two formulas φ_1 and φ_2 as being witnessed by a wedge:

$$\begin{tikzcd} \varphi_1 & & \varphi_2 \\ & \varphi \arrow[ul, "f_1"] \arrow[ur, "f_2"'] & \end{tikzcd}$$

Here φ is differentiated and f_1 and f_2 are simple substitutions. It is fairly natural to think of such wedges as "generalized isomorphisms" between formula. Thought of this way, two

Thus φ is differentiated if if $\#_p\varphi = 1$ or $\#_p\varphi = 0$ for each variable $p \in V$

formulas φ_1 and φ_2 are relatively bi-interpretable if and only if there is a generalized isomorphism connecting φ_1 and φ_2. It is not hard to see that there is a generalized isomorphism connecting each formula to itself and that if there is a generalized isomorphism connecting φ_1 to φ_2 there is a generalized isomorphism connecting φ_2 to φ_1 (so $\equiv$ is both reflexive and symmetric).

It is a bit more difficult to verify the transitivity of $\equiv$. What needs to be shown is that we can compose these generalized isomorphisms in a natural way so that if we have a diagram like the following,

$$\begin{tikzcd}
\varphi_1 & & \varphi_2 & & \varphi_3 \\
& \varphi \arrow[ul, "f_1"] \arrow[ur, "f_2"'] & & \psi \arrow[ul, "g_1"] \arrow[ur, "g_2"'] &
\end{tikzcd}$$

we can find a generalized isomorphism taking φ_1 to φ_3. It suffices to show that whenever we have a situation like the above, we get an automorphism η taking φ to ψ. We then obtain a generalized isomorphism between φ_1 and φ_2 by composing the simple substitution g_2 with the isomorphism η:

$$\begin{tikzcd}
\varphi_1 & & \varphi_3 \\
& \varphi \arrow[ul, "f_1"] \arrow[ur, "g_2 \circ \eta"'] &
\end{tikzcd}$$

Explicitly, we'll prove the following lemma.

Lemma 2.4. *For any differentiated formulas φ and ψ and simple substitutions f and g, if $f(\varphi) = g(\psi)$ then there is an automorphism h such that $h(\varphi) = \psi$.*

Proof. If ψ is a propositional variable, $f(\varphi) = g(\psi)$ implies φ is a propositional variable and so $h(\varphi) = \psi$ is immediate. Fixing φ, we prove the lemma by induction on ψ.

The base case is already done. Let $\psi = \neg\psi'$ and suppose $f(\varphi) = g(\psi)$. Since f is simple, and the images of $i : V \to F$ and $\neg : F \to F$ are disjoint, φ is not a propositional variable. Since the images of $\wedge$ and $\neg$ a disjoint, $f(\varphi_1) \wedge f(\varphi_2) = f(\varphi_1 \wedge \varphi_2)$ is distinct from $\neg g(\psi') = g(\psi)$, for any formulas $\varphi_1 \wedge \varphi_2$. And since F is generated by V under $\neg$ and $\wedge$ we conclude that $\varphi = \neg\varphi'$, for some formula φ'. Therefore, $\neg f(\varphi') = f(\varphi) = g(\psi) = \neg g(\psi')$.

Since $\neg$ is an injection $f(\varphi') = g(\psi')$. From the induction hypothesis $h(\varphi') = \psi'$ for an automorphism h. Therefore $h(\varphi) = \neg h(\varphi') = \neg\psi' = \psi$.

Now let $\psi = (\psi_1 \wedge \psi_2)$ and $f(\varphi) = g(\psi)$. By analogous reasoning, $\varphi = (\varphi_1 \wedge \varphi_2)$ and so $f(\varphi_1) = g(\psi_1)$ and $f(\varphi_2) = g(\psi_2)$. Applying the induction hypothesis, there are automorphisms h_1 and h_2 mapping φ_1 to ψ_1 and φ_2 to ψ_2 respectively. By assumption φ is differentiated and so φ_1 and φ_2 do not share propositional variables. We can thus define an automorphism h mapping φ to ψ by

$$h(p) = \begin{cases} h_1(p) \text{ if } p \in V(\varphi_1) \\ h_2(p) \text{ if } p \in V(\varphi_2) \\ p \text{ otherwise} \end{cases}$$

$\square$

From this lemma we observe the following proposition.

Lemma 2.5. *The relation $\equiv$ of relative bi-interpretability is an equivalence relation.*

Being an equivalence relation, however, doesn't suffice to show that it provides us with a workable notion of sameness of form. In order to show that, we must show that $\equiv$ is a congruence relation.

Proposition 2.6. $\equiv$ is a congruence relation.

Proof. Given lemma 3.5, it suffices to show that $\equiv$ preserves $\wedge$ and $\neg$. Suppose that $\varphi_1 \equiv \varphi_2$. Thus there is a generalized isomorphism between φ_1 and φ_2:

$$\begin{tikzcd} \varphi_1 & & \varphi_2 \\ & \varphi \arrow[ul, "f_1"] \arrow[ur, "f_2"'] & \end{tikzcd}$$

Since $\neg\varphi$ is differentiated whenever φ is, and $f(\varphi) = \varphi_i$ if and only if $f(\neg\varphi) = \neg\varphi_i$, it follows that there is a generalized isomorphism between $\neg\varphi_1$ and $\neg\varphi_2$:

$$\begin{tikzcd}
\neg\varphi_1 & & \neg\varphi_2 \\
& \neg\varphi \arrow[ul, "f_1"'] \arrow[ur, "f_2"] &
\end{tikzcd}$$

For the conjunctive case, let $\varphi_1 \equiv \varphi_2$ and $\psi_1 \equiv \psi_2$. Choose any generalized isomorphism betweeen φ_1 and φ_2:

$$\begin{tikzcd}
\varphi_1 & & \varphi_2 \\
& \varphi \arrow[ul, "f_1"'] \arrow[ur, "f_2"] &
\end{tikzcd}$$

Without loss of generality we can choose a differentiated formula ψ that does not overlap in its propositional variables with φ and g_1 and g_2 simple substitutions that together constitute a generalized isomorphism between ψ_1 and ψ_2.

$$\begin{tikzcd}
\psi_1 & & \psi_2 \\
& \psi \arrow[ul, "g_1"'] \arrow[ur, "g_2"] &
\end{tikzcd}$$

Since φ and ψ are differentiated and do not overlap in their propositional variables, $\varphi \wedge \psi$ is differentiated. Define h_1 to be that substitution such that $h_1(p) = f_1(p)$ if p occurs in φ and $h_1(p) = g_1(p)$ otherwise. And define $h_2(p) = f_2(p)$ if p occurs in φ and $h_2(p) = g_2(p)$ otherwise. Then the following generalized isomorphism witnesses $(\varphi_1 \wedge \psi_1) \equiv (\varphi_2 \wedge \psi_2)$:

$$\begin{tikzcd}
\varphi_1 \wedge \psi_1 & & \varphi_2 \wedge \psi_2 \\
& \varphi \wedge \psi \arrow[ul, "h_1"'] \arrow[ur, "h_2"] &
\end{tikzcd}$$

$\square$

By replacing bi-interpretability with relative bi-interpretability in the principle of *Identity* we obtain the following principle.

> *Weak Identity:* The form of φ and ψ is the same if and only if $\varphi \equiv \psi$.

More generally, relative bi-interpretability gives rise to the following theory.

> EXISTENCE: There is a function $f : F \to H$ that associates each formula $\varphi \in F$ with its form $f(\varphi) \in H$;

Structural Similarity: This function $f : F \to H$ is a homomorphism of logical algebras $f : \mathcal{F} \to \mathcal{H}$;

Weak Identity: The kernel of this function is the relation $\equiv$.

What Proposition 3.6 shows is that these principles are consistent.

2.3. **Logical forms.** One promising feature of this account of logical form is that it allows us to simply and literally speak of a form H being the conjunction of forms I and J. We can in fact strengthen the theory with the principle that the space of forms $\mathcal{H}$ be a term algebra and so satisfy unique readability.

To see this it is helpful to first explicitly construct an algebra that verifies that Structural Similarity. We do this by constructing a quotient algebra.

Definition 2.7. Let $[\varphi] = \{\psi \mid \varphi \equiv \psi\}$. Let $F/_{\equiv} = \{[\varphi] \mid \varphi \equiv \psi\}$. We define the quotient algebra as follows. The quotient algebra $\mathcal{F}/_{\equiv}$ has domain $F/_{\equiv}$ and operations defined by

$$\neg'[\varphi] = [\neg\varphi].$$

$$[\varphi] \wedge' [\psi] = [\varphi \wedge \psi]$$

The map $\varphi \mapsto [\varphi]$ is a homomorphism from $\mathcal{F}$ to $\mathcal{F}/_{\equiv}$ whose kernel is $\equiv$. The consistency of Weak Identity with Structural Similarity and Existence can then be demonstrated by identifying the form of φ with its $\equiv$-equivalence class $[\varphi]$. By identifying forms with elements of $\mathcal{F}/_{\equiv}$ we obtain a strong theory of forms according to which forms themselves have linguistic structure.

For each set X let $F(X)$ be the term algebra generated by X (so $\mathcal{F} = F(V)$).

Proposition 2.8. Let 1 be a singleton. Then $\mathcal{F}/_{\equiv}$ is isomorphic to $F(1)$.

Proof. It suffices to show that $\mathcal{F}/_{\equiv}$ is a term algebra generated by $[p]$ for $p \in V$. Let us first show that the maps $\neg'$ and $\wedge'$ are injections. Let φ and ψ be formulas. Suppose that $\neg'[\varphi] = \neg'[\psi]$. So $[\neg\varphi] = [\neg\psi]$. Therefore $\neg\varphi \equiv \neg\psi$. Let χ be differentiated and f and g be simple substitutions with $f(\chi) = \neg\varphi)$ and $g(\chi) = \neg\psi$. Since simple substitutions map variables to variables, negations to negations and conjunctions to conjunctions, it follows that $\chi = \neg\chi'$, for some formula χ'. Since χ is differentiated, $\neg\chi'$ is differentiated. Then $\neg f(\chi') = f(\neg\chi') = \neg\varphi$ and $\neg g(\chi') = g(\neg\chi') = \neg\psi$. Since $\neg$ is an injection, $f(\chi') = \varphi$ and $g(\chi') = \psi$. Hence $\varphi \equiv \psi$ and so $[\varphi] = [\psi]$. This shows that $\neg'$ is an injection. The argument that $\wedge$ is an injection is the same and so is omitted.

The operations $\neg'$ and $\wedge'$ have disjoint images that do not contain $[p]$ since simple substitutions always map variables to variables, negations to negations, and conjunctions to conjunctions. To complete the proof we need only show that $\mathcal{F}_{\equiv}$ is generated by $[p]$. Let $\Omega([p])$ be the result of closing $\{[p]\}$ under the operations of $\neg'$ and $\wedge'$. Let $\varphi(q_1, \ldots, q_n)$ be an arbitrary formula whose propositional variables are exactly $\{q_1, \ldots, q_n\}$. Let φ' be a differentiated formula with a simple substitution $f : \varphi' \to \varphi(q_1, \ldots q_n)$. Finally let $\varphi(p, \ldots p)$ be the result of replacing q_i with p. Then there is a simple substitution from φ' to $\varphi(p, \ldots, p)$. Thus $\varphi(q_1, \ldots, q_n) \equiv \varphi(p, \ldots, p)$. Hence $\varphi(p, \ldots, p) \in [\varphi(q_1, \ldots, q_n)]$ and so $[\varphi(q_1, \ldots, q_n)] = [\varphi(p, \ldots, p)] = \varphi([p], \ldots, [p])$. Since $\varphi([p], \ldots, [p]) \in \Omega([p])$ we are done. $\square$

The picture is thus the following. There is a single simple form P that is the form of all propositional variables. Anything that can be freely obtained by conjoining and negating a form is a form. Hence we have forms like $P \wedge P$ and $\neg(P \wedge P)$ and $P \wedge (P \wedge \neg P)$ and so on.

Notice that it is not required that we *identify* forms with equivalence classes of formulas to get this result. Rather we just require that the space of forms be isomorphic to $\mathcal{F}/_{\equiv}$. This is enough to secure the fact that forms possess linguistic structure of some kind.

This shows that weak forms provide candidate forms that allow us to consistently maintain Structural Similarity. But might there be some other notion of form that is more closely aligned with the intuitive notion of bi-interpretability that also validates structural similarity? We can actually show that the answer to this question is negative since $\equiv$ can be shown to be the *smallest* congruence extending $\sim$. To show this it is helpful to first prove the following lemma.

Lemma 2.9. *For any congruence $\sim'$ extending $\sim$, $\varphi \sim' f(\varphi)$ for any formula φ and simple substitution f.*

Proof. We prove this by induction on φ. If φ is is a propositional variable, then $\varphi \sim f(\varphi)$ for any simple substitution f. Let $\varphi = \neg\psi$ and let f be any simple substitution. By induction $\psi \sim' f(\psi)$. So $\neg\psi \sim' \neg f(\psi)$ since $\sim'$ is a congruence. Similarly, if $\varphi = (\varphi_1 \wedge \varphi_2)$ and f is a simple substitution, then $\varphi_1 \sim' f(\varphi_1)$ and $\varphi_2 \sim' f(\varphi_2)$ by induction. Therefore $(\varphi_1 \wedge \varphi_2) \sim' f(\varphi_1) \wedge f(\varphi_2) = f(\varphi_1 \wedge \varphi_2)$ since $\sim'$ is a congruence. □

Using this lemma we now prove our claim:

Theorem 2.10. *The relation $\equiv$ is the smallest congruence relation containing $\sim$.*

Proof. Let χ be a differentiated formula and f and g simple substitutions mapping χ to φ and χ to ψ. Let $\sim'$ be any congruence extending $\sim$. By lemma 2.9, $\chi \sim' \varphi$ and $\chi \sim' \psi$. Since $\sim'$ is an equivalence relation, $\varphi \sim' \psi$. Therefore $\equiv$ is contained in $\sim'$, for any arbitrary congruence that contains $\sim$. □

Summing up, the relation $\equiv$ of relative bi-interpretability is the smallest congruence extending $\sim$. Thus a theory stating that $\equiv$ captures sameness of form can be seen as the theory that diverges from the standard account in the minimal way required in order to satisfy *Structural Similarity.*

2.4. **Meta-forms.** If forms have linguistic structure, shouldn't they themselves have forms? One might worry we are on our way to something like a third man argument, since it would now appear that p and the form of p, $[p]$, themselves share a form. In this section I'll argue that there is no regress since the form of any form of a formula is just the form of that formula.

For the moment we will suppose forms are formulas belonging to $F(p) = F(\{p\})$ for some $p \in V$. Note that the relation $\equiv |_{F(p)}$ is simply the singleton $\{(p, p)\}$ since the only differentiated formula in $F(p)$ is p. But this fact seems to me not very deep, for there is another relation that coincides with $\equiv$ on $\mathcal{F}$ but whose restriction to $F(p)$ is less trivial.

Say that a formula φ is *uniform* if there is exactly one propositional variable p that is a subformula of φ. Say that φ and ψ are *uniformly bi-interpretable* if for some uniform formula χ, and simple substitutions f and g, $f(\varphi) = \chi$ and $g(\psi) = \chi$.

Proposition 2.11. Two formulas φ and ψ are relatively bi-interpretable if and only if they are uniformly bi-interpretable.

Proof sketch. If φ and ψ are relatively bi-interpretable, then there is a differentiated formula χ with simple substitutions f and and g taking χ to φ and χ to ψ respectively. But then there will be simple substitutions taking φ and ψ to the result of substituting out all propositional variables in χ for a single propositional variable p. Similarly if φ and ψ are uniformly bi-interpretable, there is a uniform formula χ and simple substitutions mapping φ to χ and ψ to χ. Then if we change each variable in χ so that it is differentiated, there will be simple substitutions taking this transformed formula to φ and ψ respectively. $\square$

Thus in order to study the space of meta-forms, we can look at how the relation of uniform bi-interpretability behaves over $F(p)$. The main result is the following:

Theorem 2.12. *Let $\varphi, \psi \in F(p)$. Then the following are equivalent.*

(1) *φ and ψ are isomorphic with respect to $F(p)$*

(2) *φ and ψ are uniformly bi-interpretable with respect to $F(p)$.*

(3) *φ and ψ are identical.*

Proof. Obviously (3) entails (1). To show that (1) entails (2), suppose that φ and ψ are isomorphic. Then there is a simple substitution mapping φ to ψ. So since every formula in $F(p)$ is uniform, it immediately follows that φ and ψ are uniformly bi-interpretable with respect to $F(p)$. To complete the proof we need to show that (2) implies (3). We will show this by induction. For the base case, note that anything uniformly bi-interpretable with p is p, since p is the only formula in $F(p)$ that is not the result of applying $\neg$ or $\wedge$ to some other formulas. Suppose that anything uniformly bi-interpretable with φ is φ and suppose that ψ is uniformly bi-interpretable with $\neg\varphi$. Then since φ and ψ are both uniform, there is a simple substitution f such that $f(\neg\varphi) = \neg f(\varphi) = \psi$. Since φ is uniformly bi-interpretable with $f(\varphi)$, $\varphi = f(\varphi)$. Hence $\neg\varphi = \neg f(\varphi) = \psi$. The case of conjunction is proved similarly. $\square$

The result is that the space of meta-forms, on the current conception, is isomorphic to the space of forms. The natural development of the theory is then to hold that while the theory of forms of formula is non-trivial, since no form is identical to the formula of which it is a form, the theory of meta-forms is trivial, since every meta-form is identical to the form of which it is a form.

It's unclear to me whether this should be seen as a criticism of this conception of form or not. The question of whether the form of a form is that form itself strikes me as something we have little independent grip on. Indeed as indicated above, it might be a benefit. On this conception of forms, we are not committed to an ever increasing hierarchy of meta-forms.

In the final section we will return to the question of the nature of forms. There, an interpretation of Fine's theory of forms is provided which takes *arbitrary formulas* to be elements of a clone corresponding to $\mathcal{F}$ (equivalently, forms can be be thought of as arrows of a certain Lawvere theory corresponding to this clone). We'll see that the forms singled out by $\equiv$ are naturally identified as a kind of "basis" of the forms singled out by $\sim$. But before we do that, I want to first look at a reason why one might think the forms singled out by $\equiv$ are insufficient as an account of *the* logical form of a given formula.

3. An Objection

There is an important objection to the idea that relative bi-interpretability serves as a notion of sameness of form. We can formulate the objection as argument:

- P1 If φ and ψ have the same logical form, then φ is logically valid if and only if ψ is logically valid.
- P2 For any formula φ, there is some differentiated formula ψ such that $\varphi \equiv \psi$.
- P3 No differentiated formula is logically valid.
- P4 Some formula is logically valid.

C There are formulas φ and ψ such that $\varphi \equiv \psi$, but φ and ψ do not have the same form.

Premise 1 seems to be a natural requirement on sameness of form: sameness of form preserves validity. Premise 2 is easily verified from the definitions. Premise 3 strikes me as plausible, though the specific argument will depend a bit on how the term 'logical validity' is understood. If by 'logically valid formula' one means a theorem of classical logic, the premise becomes a corollary of the following proposition.

Proposition 3.1. No differentiated formula is a theorem of classical logic.

Proof. A formula φ is a theorem of classical logic if and only if φ is true in every model v (where a model is a homomorphism $v : \mathcal{F} \to 2 = (\{0,1\}, 1 - \cdot, \min\{\cdot, \cdot\})$). To prove the proposition we show by induction that for any differentiated formula φ, there is a pair of models (t, f) such that $t(\varphi) = 1$ and $f(\varphi) = 0$. The induction is mostly trivial. The conjunctive case can be shown as follows. If $(\varphi \wedge \psi)$ is a differentiated formula then φ and ψ are differentiated. By induction there are pairs (t_1, f_1) and (t_2, f_2) mapping φ and ψ, respectively, to truth and falsity, respectively. Note that since $f_1(\varphi) = 1$, $f_1(\varphi \wedge \psi) = \min\{f_1(\varphi), f_1(\psi)\} = 0$. Thus it suffices to show that $(\varphi \wedge \psi)$ is true under some interpretation. Since $(\varphi \wedge \psi)$ is differentiated, φ and ψ do not overlap in propositional variables. We can thus define a homomorphism t so that it agrees with both t_1 and t_2 on propositional variables in φ and ψ. Then

$$t(\varphi \wedge \psi) = \min(t(\varphi), t(\psi)) = \min(t_1(\varphi), t_2(\psi)) = 1$$

□

Similar sorts of arguments can be made for logics that are not classical. For instance, in intuitionistic logic the argument can be made by replacing homomorphisms into 2 with homomorphisms into the variety of Heyting algebras. How far exactly the argument extends is somewhat outside of the scope of this paper.

Given that premises 2-4 are on solid footing, the only option is to deny premise 1. But how can this be plausible at all? Here is one natural thought. When we look at the space of formulas $\mathcal{F}$ from above, we are viewing formulas independently of context. And it is not only form but context that account for validity. Automorphisms preserve both context and form. But the relation of relative bi-interpretability only preserves only form. Thus, premise 1 fails because it fails to take into account the context that the form is placed in.

I think there may be a way to develop that account into a plausible account. I want to suggest another lesson, however. It seems to me that what this argument and Fine's puzzle bring out is that it is not correct to suppose that there is one item that is *the* form of a formula. Instead, a formula has several forms. Some of these forms track validity; others tracks logical complexity. In particular, there is a kind of form corresponding to bi-interpretability and a kind of form corresponding to relative bi-interpretability. Some of our intuitions concerning logical form go with relative bi-interpretability, and some with bi-interpretability. The solution to Fine's puzzle, then, is to disambiguate: the principle *identity* is true of one of our notions of form, while the principle *Structural Similarity* is true of our other notion of form.

4. Conclusion

This paper represents only a preliminary investigation into the metaphysics of form. The goal was to identify two different notions of form in terms of their structure and use them to dissolve a puzzle of form. Further investigation is required if we are to figure out what the *nature* of these forms are, and how the theory might be generalized to provide an account of the forms of other objects, both abstract and concrete.

References

Bergman, G.M. 2015. *An Invitation to General Algebra and Universal Constructions*. Universitext, 2nd. ed. Springer.

Burris, Stanley and H.P. Sankappanavar. 1981. *A Course in Universal Algebra*, Springer.

Fine, Kit. 1985. "Reasoning with Arbitrary Objects." Blackwell.

Fine, Kit. 1998. "Cantorian Abstraction: A Reconstruction and Defense." *Journal of Philosophy* 95 (12): 599-634.

Fine, Kit. 2017. "Form." *Journal of Philosophy* 114 (10): 509-535.

Kalish, Donald. 1952. "Logical Form." *Mind*, Vol. 61. No. 241 pp. 57-71.

MacFarlane, John. 2000. "What Does It Mean To Say That Logic Is Formal?" Ph.D. diss, University of Pittsburgh.

Russell, Bertrand. 1929. *Our Knowledge of the External World*. Routledge.