HOW TO LIVE WITHOUT IDENTITY—AND WHY

Kai F. Wehmeier

Identity, we're told, is the binary relation that every object bears to itself, and to itself only. But how can a relation be binary if it never relates two objects? This puzzled Russell and led Wittgenstein to declare that identity is not a relation between objects. The now standard view is that Wittgenstein's position is untenable, and that worries regarding the relational status of identity are the result of confusion. I argue that the rejection of identity as a binary relation is perfectly tenable. To this end, I outline and defend a logical framework that is not committed to an objectual identity relation but nevertheless expressively equivalent to first-order logic with identity. After it has thus been shown that there is no indispensability argument for objectual identity, I argue that we have good reasons for doubting the existence of such a relation, and rebut a number of attempts at discrediting these reasons.

Keywords: Identity, relation, arity, coreference

1. Introduction

Identity, conventional wisdom has it, is the binary relation that every object bears to itself, and to itself only. I will refer to this (putative) relation as 'objectual identity' in order to distinguish it from two other notions: first, the relation that obtains between names just in case they have the same bearer, sometimes called 'metalinguistic identity' but more aptly described as the coreference relation; and second, self-identity, a property every object has.

In §64 of the *Principles of Mathematics*, Russell observes:

> The question whether identity is or is not a relation, and even whether there is such a concept at all, is not easy to answer. For, it may be said, identity cannot be a relation, since, where it is truly asserted, we have only one term, whereas two terms are required for a relation. And indeed identity, an objector may urge, cannot be anything at all: two terms plainly are not identical, and one term cannot be, for what is it identical with?
>
> [1903: 63]

Noting that he cannot see how to do mathematics without identity, he concludes:

> Thus identity must be admitted, and the difficulty as to the two terms of a relation must be met by a sheer denial that two different terms are necessary.
>
> [1903: 64]

In the *Tractatus*, Wittgenstein denies that there is an objectual identity relation: ‘That identity is not a relation between objects is obvious’ [5.5301]. He elaborates: ‘Roughly speaking: to say of *two* things that they are identical is nonsense, and to say of *one* thing that it is identical with itself is to say nothing’ [5.5303]. This clearly echoes the complaint of Russell’s ‘objector’, and at 5.531-5.5321, Wittgenstein attempts to disarm indispensability arguments for objectual identity by transcribing some Russellian formulas involving ‘=’ into an identity-free notation.

Russell’s and Wittgenstein’s scruples regarding the identity relation have not sat well with mainstream analytic philosophy. I will argue that denying the existence of a binary relation of objectual identity is both a tenable and a reasonable position, that charges of 'muddleheadedness' [Mates 1986: 136] against it are misguided, and that it doesn't involve a 'clearly recognizable' mistake [Quine 1960: 117].

In §2, I explain the Tractarian proposal to purge logical notation of the equality sign and argue that its success entails the eliminability of appeals to an objectual identity relation for all purposes of first-order logic. I then address the role of objectual identity in natural language semantics and show how the 'is' of identity can be accommodated without invoking such a relation.

Section 2 thus constitutes a *Tractatus*-inspired dispensability argument. But is it reasonable to go further and claim that identity 'is no relation between objects'? In §3, I lay out a simple argument against the relational status of objectual identity and show how several attempts to escape its conclusion fail to achieve their purpose.

The final section 4 identifies four notions that are, according to our analysis, illegitimately lumped together under the heading of objectual identity, and reviews the dialectic of our argument.

2. The Dispensability of Identity in Logic and Language.

First-order logic with identity ($FOL^{=}$) is distinguished from first-order logic *simpliciter* (FOL) by the presence of a particular binary relation symbol, the equality sign '=', whose semantic value is stipulated to be the objectual identity relation. $FOL^{=}$ exceeds FOL in expressive power, e.g. in that numerical quantifiers like 'there are at least two objects such that' are definable in the former but not in the latter. Absent a relation of objectual identity, the equality sign would not be able to play its intended role, $FOL^{=}$ would collapse into FOL, and the expressive resources of logic would be compromised. It is thus tempting to conclude, with Russell, that any qualms one might have regarding objectual identity should be 'met by a sheer denial'.

I show here that this line of thought is mistaken: One can enjoy the full expressive benefits of $\mathrm{FOL}^{=}$ without using any linguistic device whose semantic value is the objectual identity relation.

To begin with, I restrict attention to first-order languages that contain no function symbols. There is, of course, a school of thought that regards functions as special cases of relations. Readers sympathetic to such a conception of functions are invited to skip the next three paragraphs.

In $\mathrm{FOL}^{=}$ every formula is logically equivalent to one in which function symbols 'f' occur at most in subformulas of the form $fs_1\ldots s_n = t$, where $s_1, \ldots, s_n$ and t are either individual variables or constants (see e.g. [Boolos, Burgess and Jeffrey 2002: 256]). We may thus stipulate, without loss of expressive power, that the *only* well-formed formulas are those in which function symbols occur at most in this way. Now rewrite the equation $fs_1\ldots s_n = t$ as $\Rightarrow(f, s_1, \ldots, s_n, t)$, which may be read 'for arguments $s_1, \ldots, s_n$, function f *yields* value t', or 'at $s_1, \ldots, s_n$, function f *evaluates to* t'. These formulations make it clear that we needn't invoke objectual identity to interpret formulas $fs_1\ldots s_n = t$, as long as we are prepared to acknowledge this primitive relation of yielding.

To be clear, I'm proposing neither the elimination of function symbols from logical notation (they survive in subformulas $\Rightarrow(f, s_1, \ldots, s_n, t)$), nor the excision of functions from our ontology (first-order structures may still assign n-ary functions to n-ary function symbols). Indeed, we may even continue to use terms of the form fx, say, if we treat them as definite descriptions $(\iota y)(\Rightarrow(f, x, y))$ to which Russell's contextual definition applies.

It is true, however, that in the new notation, the grammatical status of function symbols is essentially that of predicate symbols, since we may reparse $\Rightarrow(f, s_1, \ldots, s_n, t)$ as

$\Rightarrow_f(s_1, \ldots, s_n, t)$, with an $(n+1)$-ary predicate symbol '$\Rightarrow_f$'.[1] Languages with function symbols, if they are regimented in this way, thus behave like languages containing only predicate symbols, and we are therefore justified in considering only languages of the latter kind.[2]

There are advantages to beginning with an investigation of even more austere languages, namely those that don't contain individual constants either. Following Quine [1960: 228], I'll call such languages 'canonical notations'. For canonical notations, the advantage of FOL$^=$ over FOL consists in the possibility of expressing that distinct variables are assigned the same value.

Even in canonical notations, not all uses of '=' are essential: Formulas $x=x$ can always be replaced with an arbitrary tautology. One can rewrite $\forall x\ (x=y \rightarrow \phi(x))$ as $\phi(y)$, and likewise for $\exists x\ (x=y\ \&\ \phi(x))$. Contexts of the form $\forall x\ (\phi(x) \rightarrow x=y)$ or $\exists x(\neg x=y\ \&\ \phi(x))$ are a different story; it's here that the usefulness of the equality sign lies. But as we'll see shortly, the apparent non-eliminability of '=' from these contexts is just a quirk of FOL$^=$: It's a consequence of the convention that the values of variables may vary independently of each other. If one insisted that distinct variables *never* be assigned the same value, $\forall x\ (\phi(x) \rightarrow$

---

[1] One would then have to stipulate the totality and right-uniqueness of $\Rightarrow_f$.

[2] It may be instructive to consider an example of the suggested rewriting strategy. Take $R(g(x, h(c)), h(h(h(y))))$, where '$g$' is a binary function symbol, 'h' a unary function symbol, and 'R' a binary relation symbol. We'd rewrite this formula into

$$\exists r \exists s \exists t \exists u \exists v\ (\Rightarrow(g, x, t, r)\ \&\ \Rightarrow(h, c, t)\ \&\ \Rightarrow(h, u, s)\ \&\ \Rightarrow(h, v, u)\ \&\ \Rightarrow(h, y, v)\ \&\ Rrs).$$

Clearly the original notation is more concise and better suited for practical purposes. But since we're asking whether objectual identity is eliminable from logic *in principle*, such practical considerations carry no weight.

$x=y$) would be equivalent to $\forall x \neg\phi(x)$, and $\exists x (\neg x=y \,\&\, \phi(x))$ to $\exists x\, \phi(x)$ (because $x=y$ would be logically false).

This, by and large, is Wittgenstein's proposal at *Tractatus* 5.53: 'Identity of the object I express by identity of the sign and not by means of a sign of identity. Difference of the objects by difference of the signs.'[3]

Wittgenstein illustrates the proposal by transcribing various Russellian formulas into his own, equality-free notation. He does not, however, provide any systematic account, such as a general translation algorithm. This lacuna was filled by Hintikka [1956], who showed that, as far as canonical notations are concerned, one can indeed base first-order logic on an equality-free language without jettisoning the expressive power of $\mathrm{FOL}^{=}$. A brief explanation how such a Wittgensteinian predicate logic (W-logic for short) works may be in order.[4]

W-logic's characteristic feature is that, from the range of a bound variable 'x', all objects are excluded that are values of variables occurring free within the scope of the quantifier that binds 'x'.[5] Consider $\forall x \exists y\, \mathrm{R}xy$. It contains no free variables, so the range of 'x' is unrestricted, and $\forall x \exists y\, \mathrm{R}xy$ is true in the sense of W-logic (W-true for short) just in case for every object a, $\exists y\, \mathrm{R}xy$ is W-true when 'x' has the value a. Now $\exists y\, \mathrm{R}xy$ contains a free occurrence of 'x', so $\exists y\, \mathrm{R}xy$ is W-true when 'x' has the value a if, and only if, for some b *other than* a, $\mathrm{R}xy$ is W-true when 'x' has the value a and 'y' has the value b. Thus $\forall x \exists y\, \mathrm{R}xy$ is W-true just in case for every a there is a b other than a such that a bears R to b. Similarly, $\exists x \exists y\, (\mathrm{P}x \,\&\, \mathrm{P}y)$ is W-true if and only if at least two objects fall under P.

---

[3] 'Sign' is to be read quite generally here so as to include not just names, but also variables.

[4] Readers interested in exegetical issues such as Wittgenstein's own motivation for his convention should consult chapter V, §6 of [Fogelin 1976] as well as [White 1977/78], which helpfully sets out Wittgenstein's views against a Fregean background. For technical details see my [2004] and [2008].

[5] In languages containing constants, which we consider below, the referents of constants occurring within the scope of the quantifier binding 'x' will also be excluded from the range of 'x'.

As it turns out, for every sentence of a canonical notation in $FOL^{=}$, there exists a canonical-notation sentence without '=' that has exactly the same first-order models, under the semantics of W-logic, as the original sentence does under standard semantics. Hence, with respect to canonical notations, W-logic makes the equality sign, and with it the appeal to objectual identity as its semantic value, superfluous.

It's natural at this point to wonder whether W-logic really does get by without *any* appeal to objectual identity. After all, in expounding the meanings of Wittgensteinian formulas, I've made use of a metalanguage that seems to be committed to identity. I've said, for example, that $\exists y$ Rxy is W-true of *a* if and only if for some object *b other than a*, Rxy is W-true of *a* and *b*. What else could 'some object *b* other than *a*' mean than 'some object *b* such that not *b*=*a*'? It might thus seem that we've simply pushed the appeal to objectual identity from the object- into the metalanguage.

Not so. To be sure, my formulation of the truth conditions for $\forall x \exists y$ Rxy is tailored to an audience trained to read English through the lens of $FOL^{=}$.[6] Within a community whose understanding of natural language quantification is the Wittgensteinian one, however, I could simply say that $\exists y$ Rxy is true of *a* if and only if for some object *b*, Rxy is true of *a* and *b*: According to the Wittgensteinian reading of the quantifier, the value *a* would already be excluded from the range of the (metalinguistically) bound variable '*b*', because it is mentioned inside the scope of the quantifier 'for some object *b*'.

Let's now lift a restriction imposed earlier, and consider languages containing constants. Recall that in canonical notations, the equality sign makes it possible to express the

[6] Anyone who has ever taught an introductory logic course knows that many beginning students must be drilled into accepting that '$\forall x \forall y$' does not mean 'for any *two* objects *x* and *y*', that is, that '*x*' and '*y*' may be instantiated to the same value. The viability of W-logic shows that this pedagogical maneuver is merely the enforcement of a convention, rather than the correction of a conceptual mistake.

coincidence of the values assigned to distinct variables. In the presence of constants, '=' assumes two additional functions, namely to express, first, that the value assigned to a variable coincides with the referent of a constant, and second, that the referents of two constants coincide.

It's easy to see that adoption of W-logic renders the first of these functions superfluous.[7] The second is trickier. In my [2008], I proved that it's impossible to express, in W-logic as considered so far, that distinct constants corefer. However, I also showed that one can restore the full expressive power of $\mathrm{FOL}^{=}$ to W-logic by introducing a new binary predicate '≡', with the intention that an atomic formula $c \equiv d$ be interpreted as 'the constants "*c*" and "*d*" corefer'.[8] In other words, the triple bar stands for the coreference relation, the relation obtaining between names precisely when they refer to the same object.

It's important to observe that coreference does *not* contain a covert appeal to objectual identity: Letting 'ref' stand for the reference relation, we can formalize 'the constants "*c*" and "*d*" corefer' as

$$\exists x\ (\mathrm{ref}(`c', x)\ \&\ \mathrm{ref}(`d', x))$$

[7] Recall that the range of a bound variable also excludes the referents of all constants occurring in the scope of the quantifier binding that variable.

[8] The triple bar symbol is, of course, a tip o' the hat to Frege's [1879] introduction of identity in the *Begriffsschrift*. It's crucial to insist that the triple bar can only be flanked by individual constants, never by variables, to avoid the use-mention trouble diagnosed in Frege by, among others, White [1977/78: 178], Heck [2003: 87], and Mendelsohn [2005: 60-61].

even in first-order logic *without* identity. Objectual identity plays no role here; the operative notion is rather that of existential quantification: there is some *one* thing to which both names refer.[9]

This concludes the *Tractatus*-inspired argument for the dispensability of objectual identity with respect to first-order logic. Given the general translatability of $\text{FOL}^=$ into W-logic, it also follows that mathematics, at least to the extent that it's formalizable in $\text{FOL}^=$, can be carried out without invoking an objectual identity relation.[10]

We turn now to the question whether natural language commits us to objectual identity. In a sense, our logical results already imply that W-logic can formalize every English statement amenable to analysis in $\text{FOL}^=$. It may nevertheless be instructive to see that formalization in W-logic is no more artificial than in $\text{FOL}^=$.

By way of example, consider statements formed by means of 'except', such as

(1) Everyone except Alfred is smart.

Strictly speaking, (1) is ambiguous: it may or may not imply Alfred's failure to be smart. We only consider the former reading. Its standard analysis is

(1s) $\forall x\ (x{\neq}a \leftrightarrow \mathrm{S}x)$.

---

[9] Thus McGinn [2000: 8] errs when he writes that 'if we say "for some x, x is F and x is G", we are making tacit appeal to the idea of identity in using "x" twice here: it has to be the *same* object that is both F and G for this formula to come out true.' For an extended discussion see [Humberstone and Townsend 1994].

[10] Since the translation easily extends to second- and higher-order logics, there's no reason to think that W-logic might not be able to accommodate any part of mathematics.

One can avoid the use of '=', and the concomitant appeal to objectual identity, in the presence of Wittgensteinian quantifiers: In W-logic, (1) can be formalized as

(1w) $\forall x\,(\neg \mathrm{S}a\ \&\ \mathrm{S}x)$.[11]

The occurrence of '*a*' within the scope of the universal quantifier in (1w) exempts Alfred from the scope of the bound variable '*x*', as required, and (1w) does not differ from the surface form of (1) any more than the standard formalization (1s).

This analysis presupposes the availability of Wittgensteinian quantifiers in English. There's indeed evidence that English makes use of Wittgensteinian quantifiers, and I provide an example below. It's important to note, though, that our analysis by no means requires that *all* English quantifiers behave like those of W-logic. For as long as Wittgensteinian quantifiers are available, the simultaneous presence of standard quantifiers cannot diminish expressive power.

Here's the promised example of a W-logical quantifier in English—a staple of introductory logic classes (see [Jeffrey 1991: 84]).

(2) Everybody loves my baby, but my baby don't love nobody but me.

According to received wisdom, (2) is to be formalized as

[11] Actually, (1w) is adequate only under the assumption that the domain contains at least one individual besides Alfred, which is certainly a conversational implicature of any use of (1). In a one-element domain, (1w) is vacuously W-true, regardless of Alfred's smartness. A formalization that is adequate even in one-element domains would be $\forall x\,(\neg \mathrm{S}a\ \&\ \mathrm{S}x)\ \&\ \neg \mathrm{S}a$.

(2') $\forall x\ Lxb\ \&\ \forall x\ (Lbx \rightarrow x{=}i)$,

where ‘*i*’ stands for the speaker and ‘*b*’ for her baby. But then (2) has the counterintuitive consequence that the speaker and her baby are the same person: Instantiating the universal quantifier in the first conjunct of (2') with ‘*b*’ yields L*bb*; instantiating the quantifier in the second conjunct with ‘*b*’ yields $Lbb \rightarrow b{=}i$, hence by *modus ponens*, $b{=}i$.

Against this, it seems to me the counterintuitiveness of the purported consequence is due to the fact that no such consequence obtains in English. Outside textbooks of formal logic, it simply doesn’t follow from (2) that the speaker and her baby are one person. Where, then, has the argument of the preceding paragraph gone wrong? The *modus ponens* inference appears unassailable, so the failure of consequence must be due to the inadmissibility of the universal instantiations. The most straightforward explanation for the inadmissibility of instantiating ‘*x*’ to ‘*b*’ in (2') is that the quantifiers involved in (2) are Wittgensteinian rather than standard. After all, in W-logic, $\forall x\ Lxb$ doesn’t entail L*bb*, because ‘*b*’ occurs within the scope of the quantifier. Examples like (2) therefore make it plausible that English contains quantifiers abiding by the semantics of W-logic.

Let’s finally turn to the so-called ‘*is* of identity’. We needn’t consider statements involving definite descriptions such as ‘Scott is the author of *Waverley*’ or ‘The author of *Ivanhoe* is the author of *Waverley*’, since Russell’s analysis of descriptions, applied within W-logic, will make any putative occurrence of ‘=’ disappear.[12] This leaves us with statements in which ‘is’ is flanked by two proper names, as in the stock example

(3) Hesperus is Phosphorus.

[12] That is, we can analyze ‘*a* is the F’ as $Fa\ \&\ \neg\exists x\exists y\ (Fx\ \&\ Fy)$ and ‘The F is the G’ as $\exists x\ (Fx\ \&\ Gx)\ \&\ \neg\exists x\exists y\ (Fx\ \&\ Fy)\ \&\ \neg\exists x\exists y\ (Gx\ \&\ Gy)$.

I submit that (3) is of the form

(3') $h \equiv p$,

where the triple bar is the familiar coreference predicate. In other words, I propose to analyze (3) as:

(3'') There is an object x such that 'Hesperus' refers to x and 'Phosphorus' refers to x.

Questions regarding the informativeness of identity statements so understood, as raised in [Frege 1892], deserve more discussion than I can provide here. I refer the reader to [Fiengo and May 2006], where it is argued on philosophical and linguistic grounds that typical uses of statements like (3) are best understood as coreference claims. Fiengo and May observe that 'what the ancient astronomer has *said* [by uttering (3)] is that "Hesperus" and "Phosphorus" corefer', and suggest that a suitable notion of informativeness for identity statements must be 'derivative from the information . . . that the expressions flanking the identity sign corefer' [116; italics in the original]. Their position is thus very close to the view defended in this paper, for they also note that 'there aren't very many circumstances at all for the use of [objectual] identity statements, not at least in ordinary discourse' [117]. What compels Fiengo and May to hold on to objectual identity at all is their respect for the intuition of the necessity of identity statements. It's arguable, however, that this intuition has nothing to do with objectual identity. We'll discuss this point in connection with Kripke's views in the next section.

3. 'To say of *two* things that they are identical is nonsense.'

I hope to have convinced the reader that we can make do, in logic and natural language, without appeal to a relation of objectual identity. It's not clear, however, that we should take eliminability to imply non-existence. Is there, then, any additional reason for scepticism with respect to the existence of the identity relation?

To be clear: We're asking whether there exists a *binary* relation that every object bears to itself, and to no others. It's uncontentious that there exists a property (that is, in an inclusive usage of 'relation', a *unary* relation) of self-identity, a property every object has.

This raises the question what it means for a relation to be binary, rather than unary; more generally, *n*-ary rather than *m*-ary. A natural answer would seem to be this: The arity of R is the maximal number of objects that can possibly be related by R. Let us call this the Wittgensteinian Arity Principle, or WAP.[13]

Assuming WAP, scepticism regarding the binary nature of objectual identity can be buttressed by rigorous proof: The maximal number of objects that can possibly be related by objectual identity is one, because no *two* things can ever be identical (otherwise they wouldn't be two). Therefore, by WAP, objectual identity is unary, i.e. a property, rather than a proper relation.

Now WAP is certainly a natural and common characterization of arity: for instance, two recent philosophical reference works provide just this explanation.[14] But popularity is hardly evidence of plausibility. So let's explore the merits of WAP.

[13] I don't wish to suggest that this principle has any historical claim to being associated with Wittgenstein. Its designation is only intended to signal that it plays a role in the defence of a broadly Wittgensteinian claim, not that he himself would have endorsed it.

[14] To wit, Blackwell's *Companion to Metaphysics* (cf. [Mulligan 1995]) and the *Routledge Encyclopedia of Philosophy* (cf. [Bacon, Detlefsen and McCarty 1999: 7]). Both neglect to include the modal element and define

I note, first, that WAP yields the correct result for all standard examples of relations, with the sole exception of the contentious case of identity: it makes greenness and richness unary; *loving*, *stabbing*, and *giving birth to* binary; and betweenness ternary. Thus, for all I can see, the only case where WAP's correctness might be challenged is objectual identity.

Further, I take it to be analytic that a relation's arity is an upper bound for the maximal number of objects it can relate: seventeen objects cannot be related by a ternary relation (though perhaps three objects can be related by some 17-ary relations).

Now if there is a relation R whose arity is not a lower bound for the maximal number *m* of objects it can relate, the question arises how much room we should countenance between *m* and the arity of R, and what determines the size of that room. Surely friends of objectual identity don't want to end up with identity being ternary, or 17-ary, so there would have to be some principled way of restricting how much R's arity can exceed *m*.

It won't do simply to demand that arity be characterized not only in terms of the number of objects a relation can *successfully* relate, but also in terms of the number of objects a relation can *fail* to relate, for identity certainly fails to relate Frege, Russell, and Wittgenstein, but nobody thinks that this makes it ternary. To obviate this problem, the friend of identity would have to insist that application of identity to two objects, though never successful, is in some sense 'well-formed', whereas it's ill-formed when three or more objects are concerned. But unless this is postulated as a brute metaphysical fact, some explanation for the distinction between well- and ill-formed applications of objectual identity must be given.

---

the arity of a relation simply as the maximal number of objects it relates. In a world of strict narcissists, this would make the amatory relation unary. WAP gives the correct result because it would still be *possible* for the amatory relation to hold between two individuals, even if it does not in fact do so.

A *prima facie* attractive strategy is to explicate the distinction in linguistic terms, that is, to justify the binarity of the identity *relation* by way of the arity of the equality *sign* (or the 'is' of identity). Thus Quine [1960: 116]: 'What makes identity a relation, and "=" a relative term, is that "=" goes between distinct occurrences of singular terms, same or distinct, and not that it relates distinct objects.'

Let's grant that the arity of a predicate is evidence for the arity of the relation expressed. Now Quine presents his argument on the basis of a formalization of English in $FOL^=$, his preferred logical notation. We are well within our rights, of course, to adopt W-logic (including '≡') as *our* preferred logical notation; after all, it's equi-expressive with $FOL^=$ and, as argued in §2, just as natural as a logical analysis of English. But whatever example from $FOL^=$ Quine might cite for '=' occurring between singular terms, when we translate it into our notation, such occurrences either disappear (if they were flanked by at least one variable), or are transformed into occurrences of the triple bar (if the original occurrence of '=' was flanked by constants). Hence, for us, Quine's examples only show that the coreference relation is binary, which is uncontentious. The argument from the arity of the equality sign is thus unsuccessful. The purported distinction between well- and ill-formed applications of identity cannot be based on linguistic features of the equality sign, or the 'is' of identity, since these predicates arguably express the coreference relation rather than objectual identity.[15]

[15] There are other moves the friend of identity could make in order to ground a well-formed – ill-formed distinction. One might take recourse to the view that relations contain 'positions' or 'slots' that serve as receptacles for their relata, and that the arity of a relation is its number of slots. But at least in its most intuitive form, where relational facts are quasi-spatial complexes consisting of the relation with its positions occupied by the relata, such a view can hardly be reconciled with the binarity of objectual identity: Suppose identity has the two slots α and β. Insert any object x into α. Surely x itself cannot now be plugged into β, since it already occupies α. Thus identity facts are impossible. Worse, the same consideration shows that reflexive relational

Another strategy to circumvent the argument from WAP consists in adopting a set-theoretic account of relations as sets of ordered *n*-tuples. One can then block the argument by espousing an alternative characterization of arity: On the set-theoretic construal, the arity of a relation (*qua* set) R is just the number *n* such that all members of R are *n*-tuples. Indeed, for any domain D, the set $\{\langle x, x\rangle \mid x \in \mathrm{D}\}$ is clearly a binary relation-*qua*-set in which every member of D stands to itself, and itself only.[16]

Regardless of its mathematical virtues, this treatment of relations is not, as it stands, a convincing strategy against identity sceptics. An obvious objection would be that relations aren't sets, as evidenced in set-theoretic notation itself by the fact that expressions for sets (individual variables, and constants like '$\varnothing$') cannot occur syntactically in positions appropriate for relational expressions (such as '$=$' and '$\in$'), and vice versa. Indeed, if '$=$' and '$\in$' stood for classes of objects, the language of set theory would consist exclusively of singular terms, and it's hard to see how one could ever make set-theoretic statements. As Williamson [1985: 255] puts it: 'unless "$\in$" in set theory stood for a relation independently of these constructions [of relations as sets of *n*-tuples], they would not make sense. Relations are what they are and not another kind of thing.'

---

facts are impossible. While there may be ways to side-step these problems (e.g. by allowing two positions to collapse into one in case of reflexive facts, or by filling the positions not with the actual relata but by copies or clones of them), the measures to which one is driven seem so desperate that rejection of a binary objectual identity relation appears much more reasonable.

[16] This is subtler than it looks: On the now standard Kuratowski definition of the ordered pair, '$\langle x, y\rangle$' stands for the set $\{\{x\}, \{x, y\}\}$, and hence '$\langle x, x\rangle$' stands for $\{\{x\}\}$. Now $\{\langle x, x\rangle : x \in \mathrm{D}\}$, that is, $\{\{\{x\}\} : x \in \mathrm{D}\}$, looks like a binary relation when regarded as a subset of D×D, but like a property when considered as a subset of $\{\{\{x\}\}: x \in \mathrm{D}\}$. This shows that the assignment of arities to sets is relative to a background domain. Another case in point is the empty set, which is trivially, for any *n*, a set consisting solely of *n*-tuples. Thus without specification of a background domain, the empty set, considered as a relation, has *every* arity.

But this rejection of the set-theoretic account may be too swift. After all, it cannot be denied that classes of n-tuples do, for many purposes, successfully simulate n-ary relations. Thus even if relations aren't literally classes of n-tuples, it's still conceivable that there should be a one-to-one correlation between n-ary relations and classes of n-tuples. If so, the existence of the class of pairs $<x, x>$ would constitute evidence for the existence of a binary objectual identity relation.

However, Williamson [1985] and Fine [2000] have presented arguments against the possibility of such a one-to-one correspondence. In a nutshell, these arguments begin with the observation that the set-theoretic account implies the distinctness of any non-symmetric relation from its converse. For example, the amatory relation L, obtaining between x and y just in case x loves y (corresponding to the class $\{<x, y> \mid x \text{ loves } y\}$) would have to be distinct from its converse L^c that obtains between x and y when x is loved by y, i.e. the relation corresponding to $\{<x, y> \mid x \text{ is loved by } y\}$). Now according to Williamson and Fine, there's only one relation where the set-theoretic account generates two, L and L^c. This one relation serves as the semantic value of both 'loves' and 'is loved by', albeit with different linguistic conventions as to the location of the lover's and the beloved's names. In order to argue for this thesis, both authors propose to consider languages of a nonstandard sort; e.g. Fine [2000: 6] suggests a language whose relational expressions are physical bodies, with designated places in which to write the names of the relata. The predicate 'loves' might be a heart-shaped body with a red side (on which one inscribes the lover's name) and a black side (for the beloved's name). With respect to such a language, it makes no sense to ask whether the heart-shaped body stands for the amatory relation or its converse. It can stand for the one just as well as for the other; by the Principle of Sufficient Reason, it stands for the one if and only if it stands for the other. But it must stand for a unique relation, hence the relation and its converse must be one.[17]

This observation undercuts the presumption that there should be a one-to-one correlation between *n*-ary relations and classes of ordered *n*-tuples. There are, in a sense, more sets of tuples than relations, as witnessed by the fact that the sets {<*x*, *y*> | *x* loves *y*} and {<*x*, *y*> | *x* is loved by *y*} correspond to the same relation; in other words, at least some sets of tuples are artefacts of the way we represent relations within set theory. But then it is entirely conceivable that the identity relation, the class of pairs <*x*, *x*>, should itself be such an artefact, and that it maps to the same relation as the class of 1-tuples <*x*>.[18]

Taken together, these considerations seem to provide sufficient grounds to reject arguments from set theory for the binarity of objectual identity.

Let's turn to another influential defence of objectual identity. Kripke [1980: 108] suggests that even if the 'is' of identity can be taken to stand for the coreference relation, one is free to introduce a new relation, schmidentity, by *stipulating* that it is to hold between every object and itself, and in no other cases. In the context of our discussion, this is, of

[17] Set theory takes, as it were, a contingent *notational* feature of our natural and symbolic languages, namely the use of order to indicate in which way a relation applies to its relata, and misleadingly turns it into an essential feature of the *relation*.

[18] That there is a distinction between <*x*, *x*> and <*x*> is, incidentally, not obvious, even within set theory. As noted, <*x*, *x*> is typically taken to be the set {{*x*}}. There's no generally accepted set-theoretic definition of the 1-tuple <*x*>, presumably because the notion is not particularly useful for mathematical purposes. However, given that the only characteristic required of a definition of ordered *n*-tuples is that the identity of <x_1, ..., x_n> and < y_1, ..., y_n > imply the identity of each x_i with the corresponding y_i, it's perfectly in order to define <*x*> as {{*x*}}, for from the identity of {{*x*}} with {{*y*}} the identity of *x* and *y* certainly follows. But then <*x*, *x*> just is <*x*>, and the arity of the set-theoretic identity relation becomes ambiguous between one and two.

course, begging the question, since the dispute is precisely over whether such a stipulation would succeed in introducing a binary relation.

In a footnote, Kripke acknowledges this and elaborates:

> Of course, the device will fail to convince a philosopher who wants to argue that an artificial language or concept of the supposed type is logically impossible. In the present case, some philosophers have thought that a relation, being essentially two-termed, cannot hold between a thing and itself. This position is plainly absurd. Someone can be his own worst enemy, his own severest critic and the like. Some relations are reflexive such as the relation 'no richer than'. Identity or schmidentity is nothing but the smallest reflexive relation.
>
> [ibid.]

We should note that acceptance of WAP doesn't commit us to the position Kripke regards as 'plainly absurd'. All of his examples—*x* is a critic of *y*, *x* is an enemy of *y*, *x* is no richer than *y*—correctly come out as binary according to WAP, because the *maximal* number of individuals *x* and *y* such that *x* can be a critic of *y*, an enemy of *y*, or no richer than *y*, is two, regardless of whether those relations sometimes apply to an object and itself.

The only substantive point we need consider, then, is Kripke's suggestion that identity is simply 'the smallest reflexive relation'. If we construe binary relations as sets of ordered pairs, this is reasonable enough, but we've just seen that the reconstruction of relations as sets engenders artefacts that may be misleading as to the nature of the relations themselves.

An obvious reaction to Kripke's proposal is to ask: Why should a *smallest* reflexive relation exist? After all, there are plenty of cases where a notion of 'smallest' is ill-defined in general (even when unproblematic in the finitary case). Take the ordering of the reals: Any finite set of real numbers has a smallest member; however, the set of positive reals does not.

It would be hasty to dismiss Kripke's point on this basis, for objectual identity can be obtained (or so it might seem) not only by intersecting *all* reflexive relations, but even by intersecting just two, namely any reflexive relation and its converse (for instance, ≤ and ≥, if we are interested only in identity on numbers).

This particular way of putting the point isn't entirely satisfactory, given the Williamson–Fine arguments discussed above, for if a relation and its converse are the same relation, their intersection will not produce anything new.[19] But we can put the issue in terms that are not affected by Williamson's and Fine's observations: According to WAP, the relation of diversity (which any object bears to any other object, but not to itself) is binary. Identity is the complement of diversity, thus if relations are closed under complementation, it follows that identity, too, is a binary relation.

The argument from complementation contains a lacuna, however. To see this, we note that complementation is typically relative to a background domain; the complement of a set S, for instance, consists of the elements of an antecedently given domain—usually a superset of S— that are not members of S. But the background domain is not, in general, uniquely determined by S. Consider the relation M that holds between a man *m* and a woman *w* just in case *m* is married to *w*. One natural choice for the complement of M is the relation that holds between a man *m* and a woman *w* just in case *m* is not married to *w*. But one might just as well take M's complement to be the relation that holds between adults *a* and *b* just in case *a* is not married to *b*. In the latter case, but not in the former, M's complement relates Julia Gillard to Angela Merkel.

[19] In fact, intersection is not even well-defined on relations. See [Williamson 1985: 258-60].

Returning to diversity, we may ask: Relative to which background domain are we to take the complement of the diversity relation? The Kripkean answer is: with respect to the set of all pairs <*x*, *y*>, where *x* may be different from or identical with *y*. This is a natural reply from the set-theoretical perspective; however, it's not the only possible or even plausible answer. Given that diversity is introduced as the set of *two*-element pairs <*x*, *y*>, and that its application to one-element pairs <*x*, *x*> is thereby excluded on definitional grounds, it seems just as natural to take its complement relative to the set of all two-element pairs. But then the complement of diversity is the empty relation. So it's not at all clear that the complement of diversity really is identity.

It helps to consider the situation in the setting of W-logic. Kripke's proposal to introduce schmidentity by stipulation amounts to the introduction into W-logic of a binary predicate 'I' via the axiom

(Ax=) $\forall x\ \mathrm{I}(x, x)\ \&\ \forall x \forall y\ \neg \mathrm{I}(x, y).$

After all, the semantic value of such a predicate would have to be the objectual identity relation, or so it seems: Being a binary predicate, it must stand for a binary relation; the first conjunct ensures that this relation obtains between every object and itself; and the second conjunct, that no object bears it to any other.

We note that (Ax=) does not add expressive power to W-logic: Assuming a canonical notation, the argument places of the I-predicate can be filled in one of two ways: either with two tokens of one variable in both argument places, as in $\mathrm{I}(x, x)$; or with tokens of distinct variables, as in $\mathrm{I}(x, y)$. Given the first conjunct of (Ax=), formulas $\mathrm{I}(x, x)$ are always true, and hence logically equivalent to any tautology $\mathrm{Taut}(x)$, say $\mathrm{P}(x) \vee \neg \mathrm{P}(x)$. Thus we may always replace any subformula $\mathrm{I}(x, x)$ by $\mathrm{Taut}(x)$ *salva veritate*. By the second conjunct of (Ax=),

$I(x, y)$ is always false and therefore equivalent to any contradiction $\mathrm{Cont}(x, y)$, say $R(x, y)$ & $\neg R(x, y)$. Thus we may always replace $I(x, y)$ by $\mathrm{Cont}(x, y)$. In this way we can transform any sentence in which the I-predicate occurs into an equivalent one in which it does not.

Our reasoning arguably shows more: $I(x, x)$ is just an awkward and superfluous alternative notation for $\mathrm{Taut}(x)$, and $I(x, y)$ for $\mathrm{Cont}(x, y)$. The stipulation that was intended to introduce *one* binary relation, identity, has in fact picked out *two* relations: the universal unary relation, represented by $\mathrm{Taut}(x)$, and the empty relation $\mathrm{Cont}(x, y)$. It's notational legerdemain, as it were, to use the same predicate letter in the two conjuncts of (Ax=); the principle ought really to look like this:

$$\forall x\ \mathrm{Taut}(x)\ \&\ \forall x \forall y\ \neg \mathrm{Cont}(x, y),$$

which is, of course, a W-logical truth.

There is, thus, no unified notion of objectual identity—just a forced marriage, under the name of equality, of the universal property $\mathrm{Taut}(x)$ and the empty relation $\mathrm{Cont}(x, y)$.

Before we conclude our discussion of Kripke, it remains to address another facet of his views on identity—to wit, the necessity of identity— that threatens our analysis of identity statements between proper names. Consider again

(3) Hesperus is Phosphorus.

On our interpretation of the 'is' of identity, (3) has the form $h \equiv p$ and is not necessary: Mars and Jupiter might have been *called* 'Hesperus' and 'Phosphorus'; hence 'Hesperus' and 'Phosphorus' might not have coreferred.

Yet we have a robust intuition, brought to the fore by Kripke, that, in *some* sense, (3) is necessary. It's thus incumbent on me to explain the necessity of identity statements involving proper names without invoking objectual identity.

As a first step, I want to suggest that there's an even more basic intuition at work, namely the intuition that it's impossible for one object to be two objects. Sentence (3) seems necessary because, since it's true, Hesperus and Phosphorus are one object. Hence they could not have been two. Therefore they're necessarily one, that is, Hesperus is necessarily Phosphorus.

Note that the argument of the preceding paragraph glosses sentences of the form 'A is B' as 'A and B are one object'. When viewed through the lens of $\mathrm{FOL}^{=}$, this makes reference to objectual identity; but in W-logic, all it involves is elementary cardinality talk that can easily be accommodated in the presence of a modicum of second-order logic.[20]

To be more precise, I propose that the logical form of 'Hesperus and Phosphorus are one' is the W-logical sentence

(4) $\exists \mathrm{F}\, (\mathrm{F}h \;\&\; \mathrm{F}p \;\&\; \neg\exists y \exists z\, (\mathrm{F}y \;\&\; \mathrm{F}z)),$

[20] As noted by an anonymous referee, on the standard construal of the first-order quantifiers, second-order logic allows for a definition of identity as indiscernibility: one can introduce a binary predicate 'I' via the Leibniz principle $\forall x \forall y\, (\mathrm{I}(x, y) \leftrightarrow \forall \mathrm{P}\, (\mathrm{P}(x) \leftrightarrow \mathrm{P}(y)))$, and the semantics of second-order logic then ensures that the interpretation of 'I' is objectual identity. However, from our perspective, this involves the same legerdemain as the principle (Ax=) considered above, for the Leibniz principle translates into W-logic as the conjunction of $\forall x\, (\mathrm{I}(x, x) \leftrightarrow \forall \mathrm{P}\, (\mathrm{P}(x) \leftrightarrow \mathrm{P}(x)))$ and $\forall x \forall y\, (\mathrm{I}(x, y) \leftrightarrow \forall \mathrm{P}\, (\mathrm{P}(x) \leftrightarrow \mathrm{P}(y)))$. The former reduces to $\forall x\, \mathrm{I}(x, x)$, and the latter (under the assumption, which we must make anyway, that no two things can share all their properties) to $\forall x \forall y\, \neg \mathrm{I}(x, y)$. Thus the Leibniz principle is equivalent to (Ax=) in W-logic, and our earlier comments apply.

which says that we can form a collection to which h and p belong, but which does not have more than one member.

Let's note, first, that (4) follows from $h \equiv p$, assuming only the second-order W-logical principle

(5) $\forall x \exists F\, (Fx \,\&\, \neg \exists y \exists z\, (Fy \,\&\, Fz))$.

This is because (5) implies the existence of a singleton property F with Fh, and $h \equiv p$ then guarantees that Fp also holds.

In fact, (3) not only implies that (4) is true in the actual world, but also that it's true at every world where h or p exists: Since (5) is a logical principle, it holds at every possible world, and so any world w at which h exists contains a singleton property F holding of h. Now 'h' and 'p' are rigid designators, so that whenever we use these names to speak about counterfactual circumstances, they refer to whatever they refer to in the actual world. Since, by (3), 'h' and 'p' corefer in the actual world, they still corefer even when we use them to speak about w. Therefore, since Fh is true in w, Fp also holds in w. But then (4) is true in w.

Finally, note that $h \equiv p$ is logically equivalent with (4): We have seen that $h \equiv p$ implies (4). For the other direction, assume (4) and consider the property F. If $h \equiv p$ were false, $\exists z\, (Fh \,\&\, Fz)$ would follow from $Fh \,\&\, Fp$, because p would not be excluded from the range of the bound variable 'z'. The further inference to $\exists y \exists z\, (Fy \,\&\, Fz)$ is trivial because the range of 'y' is unrestricted. But this contradicts the defining property $\neg \exists y \exists z\, (Fy \,\&\, Fz)$ of F.[21] This explains why (3) often seems necessary to us—we tend to read it as the logically equivalent (4). However, objectual identity plays no role in all this.

[21] That the contingent sentence (3) can be logically equivalent to the necessary sentence (4) should not be too surprising, given the role rigidity plays in the argument. See e.g. [Zalta 1988].

4. Conclusion.

I've attempted to show that the rejection of a binary relation that every object bears to itself and to no others does not cripple our expressive resources, and that it is in fact well motivated.

We've seen that all uses of the equality sign in first-order logic can be eliminated by adopting W-logic, without thereby giving up any of the expressive power of $FOL^{=}$. I further argued that appeals to objectual identity in natural language semantics can be dispensed with if one takes W-logic to provide the logical forms of English sentences. These observations forestall the kind of indispensability argument that compelled Russell to meet objections to the binarity of identity by 'sheer denial'.

In the course of our inquiry, we encountered four relevant notions of identity, none of which was a relation that every object bears to itself, and to itself only:

There is, first, the trivial notion of self-identity, which is a *property* that every object has. Second, we have the coreference relation, which figures in identity statements like 'Hesperus is Phosphorus'. While binary, coreference is by no means a relation that *every object* bears to itself (only names bear it to anything), nor is it a relation that an object can bear *only to itself* (witness the case of 'Hesperus' and 'Phosphorus'). Logically, though not necessarily, equivalent, is the third notion of 'being one object'. Finally, there is a notion of identity proprietary to function-talk, to wit, the one figuring in the expression of the coincidence of function values, as in $f(g(x, c), y) = h(g(c, x))$. In such contexts the equality sign, I suggested, should be regarded as contextually defined, as being part of a shorthand notation that is more concise and amenable to symbolic manipulation than their equality-free translations. The substantive notion that remains after elimination of the contextual definition is that of a function's *yielding* ('⇒') a value for certain arguments; again, this doesn't involve a relation that every objects bears to itself, and to itself only.

We've thus split the seemingly unitary notion of an objectual identity relation into four: self-identity, coreference, being one, and yielding. In this respect, our treatment of '=' is not unlike the one the verb 'to be' received in the hands of Frege and Russell, who split the seemingly unitary notion of *being* into the four notions of existence, predication, subsumption, and identity.[22]

The second component in my defence of the rejection of objectual identity consisted in showing how a natural construal of the notion of arity leads one to conclude that any 'relation' every object bears to itself, and only to itself, must be unary, and how various strategies to escape this conclusion face serious objections from the identity sceptic. Among the identity-friendly strategies considered were Quine's attempt to ground the arity of the identity *relation* in the arity of the equality *sign*; the attempt to reduce relations to classes of ordered tuples; and Kripke's suggestion that the existence of 'schmidentity' as the smallest reflexive relation could simply be postulated. We also saw that Kripke's insight regarding the necessity of identity statements does not discredit our analysis by means of the coreference relation.

[22] Our quadrisection of the notion of identity may be compared with Pardey's [1994: 134-35] analysis of identity statements into seven categories. His categories (II) and (V) essentially correspond to the cases of identity statements in which at least one definite description flanks the equality sign. Category (III) consists of statements of self-identity, categories (IV) and (VI) of coreference claims, and category (VII) of statements concerning the evaluation of functions. This leaves Pardey's category (I), for which he provides the example 'Scott and the author of Waverley are identical (these supposedly two persons are in reality only one person)'. The parenthetical paraphrase suggests that he is thinking of our notion of 'being one'.

I hope it has become clear, at least, that objections to identity's relational status, as voiced by Russell and Wittgenstein, cannot be rejected out of hand.[23]

University of California, Irvine

REFERENCES

Bacon, John B., Michael Detlefsen and David C. McCarty 1999. *Logic from A to Z: The Routledge Encyclopedia Glossary of Logical and Mathematical Terms*, London: Routledge.

Boolos, George S., John P. Burgess and Richard C. Jeffrey 2002. *Computability and Logic*, 4th edition, Cambridge: Cambridge University Press.

Fiengo, Robert and Robert May 2006. *De Lingua Belief*, Cambridge, Mass.: MIT Press.

Fine, Kit 2000. Neutral Relations, *The Philosophical Review* 109/1: 1–33.

Fogelin, Robert 1976. *Wittgenstein*, London: Routledge and Kegan Paul.

[23] Thanks to two anonymous referees for this journal, to Aldo Antonelli, Bill Demopoulos, Kit Fine, Robert May, Richard Mendelsohn, Ulrich Pardey, Helge Rückert, Ansgar Seide, Bartosz Wieckowski and especially Lloyd Humberstone and Tobias Rosefeldt for helpful comments. I am also grateful to audiences at the universities of Bristol, Glasgow, Konstanz, Münster and Nancy, as well as the participants in my Metaphysics of Relations seminar.

Frege, Gottlob 1879. *Begriffsschrift*, Halle: Louis Nebert. Translated by Stefan Bauer-Mengelberg in: Jean van Heijenoort, ed., 1967. *From Frege to Gödel*, Cambridge, Mass.: Harvard University Press.

Frege, Gottlob 1892. Über Sinn und Bedeutung, *Zeitschrift für Philosophie und philosophische Kritik* 100: 25–50. Translated by Max Black as 'On Sense and Reference' in: Peter Geach and Max Black, eds, 1960. *Translations from the Philosophical Writings of Gottlob Frege*, Oxford: Blackwell.

Heck, Richard G. 2003. Frege on Identity and Identity-Statements, *Canadian Journal of Philosophy* 33/1: 83–102.

Hintikka, K. Jaakko J. 1956. Identity, Variables, and Impredicative Definitions, *The Journal of Symbolic Logic* 21/3: 225–45.

Humberstone, Lloyd and Aubrey Townsend 1994. Co-Instantiation and Identity, *Philosophical Studies* 74/2: 243–72.

Jeffrey, Richard 1991. *Formal Logic*, 3rd edition, New York: McGraw-Hill.

Kripke, Saul A. 1980. *Naming and Necessity*, Cambridge, Mass.: Harvard University Press.

Mates, Benson 1986. *The Philosophy of Leibniz—Metaphysics & Language*, New York: Oxford University Press.

McGinn, Colin 2000. *Logical Properties*, Oxford: Oxford University Press.

Mendelsohn, Richard L. 2005. *The Philosophy of Gottlob Frege*, Cambridge: Cambridge University Press.

Mulligan, Kevin 1995. Relations, in *A Companion to Metaphysics*, ed. J. Kim and E. Sosa, Oxford: Blackwell: 445–46.

Pardey, Ulrich 1994. *Identität, Existenz und Reflexivität*, Weinheim: Beltz Athenäum.

Quine, W.V.O. 1960. *Word and Object*, Cambridge, Mass.: MIT Press.

Russell, Bertrand 1903. *The Principles of Mathematics*, Cambridge: Cambridge University Press.

Wehmeier, Kai F. 2004. Wittgensteinian Predicate Logic, *Notre Dame Journal of Formal Logic* 45/1: 1–11.

Wehmeier, Kai F. 2008. Wittgensteinian Tableaux, Identity, and Co-Denotation, *Erkenntnis* 69/3: 363–76.

White, Roger 1977/78. Wittgenstein on Identity, *Proceedings of the Aristotelian Society* 78: 157–74.

Williamson, Timothy 1985. Converse Relations, *The Philosophical Review* 94/2: 249–62.

Wittgenstein, Ludwig 1922. *Tractatus Logico-Philosophicus*, translated by C.K. Ogden, New York: Harcourt, Brace & Co.

Zalta, Edward N. 1988. Logical and analytic truths that are not necessary, *Journal of Philosophy* 85/2: 57–74.