Facts and Free Logic

R. M. Sainsbury

Facts are structures which are the case, and they are what true sentences affirm. It is a fact that Fido barks. It is easy to list some of its components, Fido and the property of barking. It is hard to say what the structure is (the glue is a notoriously tricky element), but happily this is not relevant for the present purpose. Intuitively, any sentence which refers to just these components, attributing barking to Fido, will affirm the same fact. By some standards, if Fido is the smartest dog on Elm Street, a sentence like "the smartest dog on Elm Street barks" meets this condition, and so also affirms the fact that Fido barks. This immediately provokes some uneasiness. Does not Elm Street have a claim to be a component of any fact which the more complex sentence affirms? More generally, should we qualify the intuition, so that it claims only that a sentence which refers just to Fido and the property of barking, *and to nothing else*, will affirm the fact that Fido barks? This question touches one central strand of Stephen Neale's subtle monograph. In a form which comes a little closer to the generality at which Neale's argument operates, a challenge to the very idea of facts can be made as follows: if a definite description "the *F*" refers at all, then it refers to the same thing as "the thing which is both *F* and *p*", for any true sentence "*p*". There is thus no upper limit to the additional material, up to a complete description of the world, that can be incorporated into a definite description. The conflicting intuitions are: (i) that all that matters to what fact is affirmed is what the definite description refers to; and (ii) that the extraneous material cannot be guaranteed to have no effect on which fact is affirmed. In light of the challenge, it is not surprising that Neale should applaud those fact theorists who, like Russell, denied that definite descriptions are referring expressions, thus cutting off the challenge before it can be so much as issued.

As his title indicates, Neale's primary focus is on facts but, as he makes plain, his considerations have a range of other bearings, for example on the nature of reference. A simple thought is that all referring expressions have essentially the same inferential properties, perhaps those captured, for singular referring expressions, by his principle PSST. A slingshot argument might be used to show, on the basis of this assumption, that definite descriptions are not referring expressions. This is one aspect of his work I wish to explore (see §3). The other aspect is how free logic bears on the issue. Here Neale's distinction between proofs dependent on a general principle of logical equivalence (PSLE) and proofs not so dependent has additional use. The free logician will deny that the supposed logical equivalences which have historically been offered are really equivalences. On the other hand, I agree with Neale that a free logician (of the kind I favour) can reconstruct the Gödelian proofs so that they are valid in his system. A free logician, as much as a classical one, will applaud the restrictions on theories of facts which Neale advocates, and I will shortly spell that out (see §2). There is also a more informal point to be made. The motivation behind free logic, or more precisely behind negative free logic, is at odds with usual views about facts. I begin by explaining why this is so.

1

According to negative free logic (NFL), some intelligible referring expressions have no referent. Any simple sentence (that is, one consisting of an n-ary predicate together with n referring expressions)[1] containing a referring expression with no referent is false. This follows from the composition axiom: a simple sentence $R(t_1, \ldots t_n)$ is true iff each t_i has a referent, a_i, and $\langle a_1, \ldots a_n \rangle$ satisfies R; otherwise it is false. Negation toggles truth and falsehood, just as in classical logic, so there are truths containing non-referring referring expressions. Quantification is restricted: existential generalization holds if the inputs are simple sentences, but for arbitrary sentences we need an additional premise affirming, for any referring expression t, the truth of "t exists". The same additional premise is required

[1] Simple sentences with complex referring expressions are not atomic, though all atomic sentences are simple.

if t is to be introduced into a conclusion of universal instantiation. A definite description ιxFx refers to something iff that thing uniquely satisfies F. The law of identity is that every object is self-identical, $\forall x\, x=x$. Not every instance of "$x=x$" is true. "The King of France is the King of France" is false so its negation is true, but that is no counterexample to the claim that every object is self-identical (the King of France is not an object).

Negative facts have always been problematic for theorists of facts. Should one say that "Socrates is not alive" affirms the negative fact that Socrates is not alive, which would presumably contain negation as a component? Or should one say that the sentence affirms just a positive fact, say the fact that Socrates is dead, or at least affirms that there is such a positive fact? However this is best resolved, it is classically not doubted that we can count on Socrates as a component. In negative free logic, by contrast, there are negative sentences which are true because they contain a non-referring referring expression, for example "Pegasus is not a horse" (how could he be? — he doesn't even exist)[2]. The assumption that such a statement affirms a fact is hard to reconcile with the absence of any object which could serve as a component corresponding to "Pegasus". However happy we are to allow negation and the property of being a horse as components, we simply lack the resources to construct a suitable fact.

This feature of NFL links with another: it is a logic for which standard kinds of model theory are inappropriate, at least as revealing the character of an appropriate semantics for the language. Certainly, some model theoretic approaches to free logic interpret non-referring referring expressions by associating them with a special entity, one lying outside the domain used for interpreting non-empty referring expressions and quantifiers. It would even be possible to associate semantically different non-referring referring expressions with distinct such entities. But it would be a mistake to suppose that this would be an appropriate description of how such expressions are actually used. No such special objects guide or are supposed to guide speakers in their use of such expressions. By the same token, NFL sees no value in the notion of a Russellian proposition, a

[2] The sentence is false in the myth, but we are concerned only with factual truth.

sequence of objects and properties, as a semantic representation. There are two reasons. The first is that NFL welcomes complex referring expressions, and it is plain that full justice to their semantics cannot be done simply in terms of their referent (supposing there is one); attention must also be paid to the complex means whereby the expression has this referent. The second reason is that in the case of non-referring referring expressions, there is no object to play the appropriate role in a supposed Russellian proposition. Enthusiasm for facts and enthusiasm for Russellian propositions are connected. One of Russell's early discussions (1912) introduces them in tandem. His account of belief factors out its various objects, those which later commentators regard as the constituents of Russellian propositions. The account of truth makes the truth of a belief consist in the existence of a complex, a fact, whose constituents are the objects of the belief, suitably "united". NFL is at odds with both ways of thinking. In place of model theory or Russellian propositions, truth theory is the natural semantics for NFL, and truths do all the work one might have looked to facts to do.

2

One would expect this antipathy to facts and to Russellian propositions to be reflected in inferential principles. In particular, one would expect that NFL would put no fewer restrictions on fact theories than those placed by a powerful slingshot.

The proof on Neale's p. 173 requires that φ and $(\alpha{=}\iota x(x{=}\alpha \mathbin{\&} \varphi) \vee (x{=}b \mathbin{\&} \neg\varphi))$ be logically equivalent, but according to NFL they are not. Suppose φ is true and α lacks a referent. Then the identity is false and equivalence fails. The proof could be revised so as to deal with arbitrary simple sentences *Fa* and *Gb* in place of φ and ψ. This blocks the counterexample to equivalence just given, since the truth of *Fa* requires that *a* have a referent; and there are no other counterexamples. The upshot is a slight weakening of the force of this slingshot, though it would remain strong enough to provide a detailed reason to encourage a fact theorist's natural reluctance to think that a connective like "the fact that *Fa* is the fact that …" permits substitution of logical equivalents salva veritate.

Neale implies, I think rightly, that one could reconstruct Gödel's slingshot within NFL. It cannot be used as it stands, since within NFL, ι-INTR is not valid:

$$\frac{T[\Sigma(x/\alpha)]}{T[\alpha=\iota x(x=\alpha \ \& \ \Sigma x)]}$$

Here is a counterexample. Put "is not a horse" for "Σ" and "Pegasus" for "α". "T" is null. The premise then formalizes "Pegasus is not a horse" which is true according to NFL, whereas the conclusion formalizes "Pegasus is identical to the thing which is identical to Pegasus and which is not a horse", and this, according to NFL, is false. It is a simple sentence composed of the binary identity predicate and two referring expressions, neither of which refers.

I have assumed that in replacing "Σ" by "is not a horse" the target English sentence is one in which the negation sign has wide scope relative to "Pegasus". If it has narrow scope, then the sentence attempts to predicate not being horse of Pegasus and there is no counterexample since, thus understood, the premise is false (according to NFL). If the sign for negation is introduced as a replacement for "T", then it retains its wide scope in the conclusion, and so the conclusion is true (according to NFL). One might simply stipulate that negation has wide scope in the target sentence.[3] It is up to Neale to say whether "is not a horse", with wide scope negation, makes for a legitimate replacement for "Σ". One can infer that he has no objection, since he allows "≠b" as such a replacement (with, of course, negation having what the NFL theorist will see as wide scope) in applying ι-CONV in the crucial derivation (pp. 183-4).

So ι-INTR, and hence ι-CONV, are not NFL-valid. Does this mean that the slingshot is irrelevant for NFL? Surely not. As far as I know, a difficulty arises only when ι-CONV is applied to a sentence with a non-referring referring expression. By adding an existence

[3] The formal language of NFL has the resources to attribute semantically significant scope properties to negation relative to names. Occasions on which natural language requires this are, however, rare. In particular, it is not easy to hear "Pegasus doesn't exist" with narrow scope negation (and so as false).

assumption to the premises, the rule can be rendered immune to this obstacle to its NFL-validity:

$$\frac{T[\Sigma(x/\alpha)],\ \exists x\ x=\alpha}{T[\alpha=\iota x(x=\alpha\ \&\ \Sigma x)].}$$

(The rule of ι-ELIM requires no change to meet NFL scruples.) This modification of ι-INTRO does not incapacitate the rule for use in the central proof (p. 183-4), for we already have a simple sentence containing the relevant name as a premise, and this guarantees that, if the premises are true, the name has a referent, and the conditions for the application of the NFL-version of ι-CONV are met.

It is one thing to protect ι-CONV from an obstacle to its validity and another to affirm that the result is valid. Neale displays no doubts, but with no restrictions on what can replace T, some doubts are in order, independently of free logic.[4] On the face of it, John might have said that Fido barks, without having said that Fido is identical to the thing which is Fido and barks. If "John said" is a sentence connective which can replace T in ι-CONV, then the principle is not generally valid, and the same would apply to its NFL version.

Even so, adopting NFL has only a rather minor impact on the force of slingshots. But the tables may be turned: may not slingshots prove too much, and thereby undermine views which a free logician would wish to adopt? In particular, do not slingshots, in conjunction with irresistible further assumptions, show that we cannot treat definite descriptions as referring expressions?

3

This relates to the question whether it makes sense to deny that coreferring expressions are everywhere substitutable salva veritate. Since plural referring expressions will not be

---

[4] As the example in the text shows, "T" must be restricted to exclude operators which are hyperintensional, and further restrictions may be in order. This is no doubt implicit in Neale's text.

at issue, we can narrow the question: does it make sense to suppose that PSST sometimes fails? If it never does, then treating (singular) definite descriptions as singular terms, as in NFL, would seem to help fuel the argument (as e.g. mentioned on p. 187) whose conclusion is that there are no non-truth functional sentence connectives, a conclusion I regard as a reductio of something that leads to it. Neale apparently thinks PSST must hold without exception, for he considers that a "treatment of definite descriptions according to which they are singular terms" is "hence subject to PSST" (p. 187). By contrast, I will suggest that slingshot arguments add to the considerations which favour the view that PSST should be restricted.

Despite the efforts of theorists of direct reference, the majority (I would guess) still think that proper names (those uncontroversial examples of singular terms) cannot be substituted salva veritate within contexts of propositional attitudes. "The ancient astronomers believed that Hesperus is Hesperus" may be true, even though "The ancient astronomers believed that Hesperus is Phosphorus" is false. "Frege said that it is one thing to know that Hesperus is Hesperus and another to know that Hesperus is Phosphorus" is true (setting aside textual nuances) whereas "Frege said that it is one thing to know that Hesperus is Hesperus and another to know that Hesperus is Hesperus" is false. More controversially, some think that even in more ordinary contexts such substitution may not be truth preserving: allegedly, relative to some event in the story, "Clark Kent entered the phone box and Superman came out" is true though "Clark Kent entered the phone box and Clark Kent came out" is false (cf Saul 1997). Our question is whether we should restrict PSST in modal contexts as a response to a slingshot argument whose conclusion is that there are no genuinely modal (and so non-truth functional) sentence connectives.

Such a slingshot, using Neale's devices, and in particular the presentation on p. 183–4, might run on the following lines. Starting from $\Box Fa$ we conclude that $\Box Gb$, for some arbitrary contingent truth concerning a distinct object, *Gb*. Regarding this conclusion as paradoxical, we look for something to reject in principles of inference the argument exploits.

1	[1]	$\Box Fa$	Premise
2	[2]	Gb	Premise
3	[3]	$a \neq b$	Premise
1	[4]	Fa	1, $\Box$-ELIM
1	[5]	$\exists x\ x=a$	4, $\exists$-INTRO
1	[6]	$\Box(a=\iota x(x=a \ \& \ Fx))$	1,5 ι-CONV
1	[7]	$a=\iota x(x=a \ \& \ Fx$	4,5 ι-CONV
3	[8]	$a=\iota x(x=a \ \& \ x \neq b)$	3,5 ι-CONV
1,3	[9]	$\iota x(x=a \ \& \ Fx) = \iota x(x=a \ \& \ x \neq b)$	7,8 =-TRAN
1,3	[10]	$\Box(a=\iota x(x=a \ \& \ x \neq b))$	6,9 PSST
2	[11]	$\exists x\ x=b$	2, $\exists$-INTRO
1,3	[12]	$\Box(a \neq b)$	10, ι-CONV
1,2,3	[13]	$\Box(b=\iota x(x=b \ \& \ a \neq x)$	11,12, ι-CONV
2	[14]	$b = \iota x(x=b \ \& \ Gx)$	2,11 ι-CONV
2,3	[15]	$b = \iota x(x=b \ \& \ a \neq x)$	3,11 ι-CONV
2,3	[16]	$\iota x(x=b \ \& \ Gx) = \iota x(x=b \ \& \ a \neq x)$	14,15 =-TRAN
1,2,3	[17]	$\Box(b = \iota x(x=b \ \& \ Gx))$	13,16 PSST
1,2,3	[18]	$\Box(Gb)$	17, ι-CONV

On any reasonable view of modality, the premises cannot genuinely entail the conclusion, since *Gb* could be a contingent sentence. □-ELIM (which need not have been used, and is not used in Neale's more general version) and ∃-INTRO cannot be challenged, so the inference rules to be examined are ι-CONV and PSST.

An additional possible culprit, one which Neale does not mention, is the assumption that necessity is properly expressed by a sentence connective. A familiar alternative view is that necessity is properly expressed by an expression which forms a closed sentence from an open one, that is, a quantifier . On this view, "necessarily *p*" is misleading: "necessarily" is better written as "for all worlds, *w*", and "*p*" would be better rewritten to show the variable "*w*".

While recognizing that this is the right form of response for some expressions (e.g. it gives a good explanation for why we should not think of "Nero fiddled while …" as a sentence connective, else we might infer "Nero fiddled while Bush invaded" from the premises "Nero fiddled while Rome burned", "Bush invaded" and "Nero≠Bush"), I will assume that it is not right for modality, so that attention focuses on the principles of inference.

A free logician who treats definite descriptions as referring expressions may, like anyone else, have doubts about ι-CONV in its full generality (we have mentioned doubts for the case in which T is replaced by an expression like "John said that"), but there is little room for doubt about the specific applications in the proof above. The unembedded applications seem sure to be truth preserving. The bidirectional inference between Fa and $a=(\iota x\ x=a\ \&\ Fx)$ is guaranteed by the logic of identity together with the guiding principle that a definite description ιxGx has as referent, if anything, the unique satisfier of Gx. The truth of Fa ensures that a has a referent, and this cannot diverge from the referent of $\iota x\ x=a$; adding the conjunct Fx, given that Fa, cannot disturb the referent. The reasoning runs as well in the reverse direction. Everything I have just said could be prefixed by "necessarily" without loss of truth, so it seems likely that a free logician could not resist a limited version of ι-CONV for which T is replaced by and only by □. The upshot is that this theorist needs to have something to say about the validity in these contexts of PSST.

If the extension of F varies from world to world, so will the extension of ιxFx. The definite description will be non-rigid. One ought not to expect a non-rigid designator to be substitutable salva veritate for a coreferring singular term in a modal context. Yet in effect this is what the proof does at line [17]. By hypothesis, Gb is contingent. So there are worlds at which not-Gb, worlds at which G has an extension other than its actual extension. For example, it may have the null extension. With respect to such a world, $\iota x(x=b\ \&\ Gx)$ has no referent. By contrast, the description which this one replaces at line [17] is rigid, assuming that a and b are rigid and that identity and distinctness are rigid relations (at least they should relate any pairs they actually relate at every world at which these pairs exist). This is enough to explain why such a replacement in a modal context may fail to preserve truth.

I conclude that it is prima facie acceptable to restrict PSST to singular terms which designate rigidly. There is a compelling explanation of why the restriction should be enforced, and (I am assuming) a compelling case in which some such restriction is required even for proper names. If one thought, with Russell, that PSST was a constitutive principle for singular terms, one would instead have to say that arbitrary definite descriptions cannot be treated as singular terms. One would also have to say, as Russell did, that ordinary proper names are not singular terms, in virtue of their failures of substitution in, for example, contexts of attributing beliefs and sayings. Very few would nowadays accept this extreme position. Restricting PSST is a more conservative response. Once it is restricted for proper names in attitude contexts, there is no general reason not to restrict it also for complex singular terms in modal contexts. Within free logic, this allows for the natural view that many definite descriptions are, as they seem, referring expressions.

Philosophy Department, University of Texas at Austin.
Philosophy Department, King's College London

REFERENCES

Russell, B. (1912). *Problems of Philosophy*. Oxford, Oxford University Press (1959).
Saul, J. (1997). "Substitution and Simple Sentences." *Analysis* 57: 102-108.