NECESSARY EXISTENTS

Timothy Williamson

It seems obvious that I could have failed to exist. My parents could easily never have met, in which case I should never have been conceived and born. The like applies to everyone. More generally, it seems plausible that whatever exists in space and time could have failed to exist. Events could have taken an utterly different course. Our existence, like most other aspects of our lives, appears frighteningly contingent. It is therefore surprising that there is a proof of my necessary existence, a proof that generalizes to everything whatsoever. I will explain the proof and discuss what to make of it. A first reaction is that a 'proof' of such an outrageous conclusion must contain some dreadful fallacy. Yet the proof does not collapse under scrutiny. Further reflection suggests that, suitably interpreted, it may be sound. So interpreted, the conclusion is not outrageous, although it may not be the view you first thought of.

1. The proof rests on three main claims.[1] The first is that my nonexistence strictly implies the truth of the proposition which states my nonexistence:

(1)   Necessarily, if I do not exist then the proposition that I do not exist is true.

For that things are so-and-so is just what it takes for the proposition that they are so-and-so to

1

be true. The second main claim is that the truth of the proposition strictly implies its existence:

(2)   Necessarily, if the proposition that I do not exist is true then the proposition that I do not exist exists.

For if the proposition did not exist, there would be nothing to be true. The final main claim is that the existence of the proposition strictly implies my existence:

(3)   Necessarily, if the proposition that I do not exist exists then I exist.

For if I did not exist, there would be nothing for the proposition to state the nonexistence of.

Given (1)-(3), the rest of the proof is easy and more or less uncontroversial. For if p strictly implies q, q strictly implies r and r strictly implies s then p strictly implies s by the transitivity of strict implication. Thus (1), (2) and (3) entail (4):

(4)   Necessarily, if I do not exist then I exist.

Consequently, my nonexistence strictly implies a contradiction and is therefore impossible. Hence my existence is necessary; (4) entails (5):

(5)   Necessarily, I exist.

Of course, any thinker could go through (1)-(5) to prove their own necessary existence.

Indeed, nothing in the proof depends on the use of the first person 'I'; other names and demonstratives would do in its place. Indeed, we can generalize the proof by substituting a variable for 'I' to derive the result that for every x, necessarily x exists (a result which we might prefix with a further 'necessarily').

One can construct a parallel proof with the phrase 'at all times' in place of 'necessarily' throughout (1)-(5). It would prove eternal rather than necessary existence. That conclusion seems even more obviously wrong. Surely I did not exist as a soul before my conception. Moreover, a variant of the argument with 'this body' in place of 'I' would prove the eternal existence of this body. Surely this body did not exist as a foreshadow of itself before conception, and will not exist as a ghost after death. Arguably, if these are counterexamples to the proof of eternal existence, then they are also counterexamples to the proof of necessary existence, for the latter entails the former. Something exists necessarily only if it exists in all possible situations; but all past, present and future situations are possible, so it exists in all past, present and future situations, so it exists eternally. Therefore, if I did not exist before my conception or will not exist after my death, then my existence is not necessary. In any case, with or without this temporal corollary, (5) looks counterintuitive enough.

2. As already noted, (5) is a fairly uncontroversial consequence of (1)-(3). The obvious strategy is to examine the grounds for those three premises in the hope of finding a weak link. First, however, it will be useful to clarify the notion of a proposition used in all three premises.

We can refer to propositions with 'that' clauses, such as 'that it is snowing' or 'that I do not exist'. Propositions are bearers of truth and falsity. The proposition that it is snowing

is true if and only if it is snowing, and false if and only if it is not snowing. Propositions are the objects of propositional attitudes, such as believing, hoping and saying. One can believe that it is snowing, hope that it is snowing, or say that it is snowing. Propositions are expressed by the sentences in the corresponding 'that' clauses, although the same proposition can be expressed by different sentences and the same sentence can express different propositions in different contexts. I can express the proposition that I am tired by saying 'I am tired' in English or 'Ja sam umoran' in Serbian. But if you say 'I am tired' you say that you are tired, not that I am tired; the proposition which you express with that sentence is the proposition that you are tired (or something like it), not the proposition that I am tired. To express the proposition which you express with the sentence 'I am tired', I must use a sentence with a different linguistic meaning, such as 'You are tired'. A proposition can also be a premise or conclusion of an argument. For example, someone who says 'John is taller than James; therefore James is not taller than John' is expressing an argument in which the premise is the proposition that John is taller than James and the conclusion is the proposition that James is not taller than John. This preliminary explanation will be amplified below.

Now consider premise (1). It is one half of one instance of a quite general principle which characterizes truth for propositions:

(1+)   Necessarily, the proposition that P is true if and only if P.

Here 'P' may be replaced by any declarative sentence which says that something is the case, in the instance of (1) 'I do not exist'. The principle 'The proposition that P is true if and only if P' is the standard characterization of propositional truth; (1+) adds that it holds of necessity.[2]

To motivate (1+), consider a standard notion of a valid argument as one in which, necessarily, if the premises are true then the conclusion is also true. This is not the purely logical notion of validity, since it does not require the truth-preservation to be guaranteed by the logical form of the argument; the connection between premises and conclusion might be an informal one. It is nonetheless a useful notion to apply, particularly when we want to use arguments in order to draw out what follows from a counterfactual supposition. For example, although we know that space has at least three dimensions, we may still wonder what would have been the case if space had had only two dimensions; then we need to know what conclusions follow from the supposition that space has only two dimensions, in the sense that necessarily they are true if the supposition is. Now the argument from the premise that P to the conclusion that Q passes the test of necessarily preserving truth on condition that necessarily, if the proposition that P is true then the proposition that Q is true. But the usefulness of this test depends on the equivalence of that condition with the simpler condition that necessarily, if P then Q. For example, when we argue 'John is taller than James, therefore James is not taller than John', our interest is primarily in the comparative heights of John and James, and in the truth of the propositions only to the extent to which it correlates with our primary interest. We want to know whether necessarily, if John is taller than James then James is not taller than John. Our question is answered by the information that necessarily, if the proposition that John is taller than James is true then the proposition that James is not taller than John is true, *provided* that principle (1+) holds, for then, necessarily, the proposition that John is taller than James is true if and only if John is taller than James, and the proposition that James is not taller than John is true if and only if James is not taller than John. Without (1+), we have no obvious reason for using a notion of a valid argument as one in which the truth of the premises necessitates the truth of the conclusion. Thus our ordinary

5

way of thinking about the validity of arguments assumes the correctness of (1+).

Arguably, many other aspects of our use of the notion of propositional truth also depend on (1+). For example, what would it have been like had all Napoleon's (actual) hopes come true? He hoped that Russia would be conquered, so in the relevant circumstances Russia is conquered. But that assumes that, in the counterfactual circumstances in which all his hopes come true, the proposition that Russia is conquered is true if and only if Russia is conquered. Our confidence in that equivalence rests on (1+).

Schema (1+) does not commit us to any particular theory about the nature of propositions. For instance, it is neutral as to whether propositions are in some loose sense linguistic items. However, it does exclude some theories. For example, it is inconsistent with a theory on which, with respect to any possible circumstances, the phrase 'the proposition that dogs bark' simply denotes the string of letters 'Dogs bark', with whatever meaning it has in those circumstances, for the string could have meant what 'Cats philosophize' actually means, and been false even though dogs barked. Thus a theory on which propositions are linguistic items would need somehow to hold the relevant meanings fixed across such circumstances in order to accommodate (1+). That is not intuitively surprising.

None of this makes (1+) as unassailable. Some argue that it must be revised in the light of the semantic paradoxes such as the Liar. For example, suppose that by some self-referential device we can construct a proposition p to be the very proposition that p is not true. By an instance of (1+), the proposition that p is not true is true if and only if p is not true; but the proposition that p is not true just is p, so (since identical items have the same properties) p is true if and only if p is not true, which implies a contradiction. However, it is by no means clear that any proposition p can be the very proposition that p is not true. Perhaps sentences which appear to express such propositions do not really succeed in saying

6

that anything is the case, and therefore cannot be substituted for 'P' in (1+). The resolution of the semantic paradoxes remains a matter of considerable obscurity. It is not obvious what kind of qualification, if any, (1+) requires. At any rate, it should be qualified only where absolutely necessary. In premise (1) of our argument, the sentence 'I do not exist' replaces 'P'; this substitution has no special link with the semantic paradoxes. Appeal to them does not destroy the strong presumption in favour of (1). Let us therefore proceed on the revocable basis that premise (1) holds.

Premise (2) is an instance of the principle that existence is a precondition of truth:

(2+)   Necessarily, if the proposition that P is true then the proposition that P exists.

For consider a possible world in which the proposition that P does not exist. If that world had obtained, there would have been no proposition that P. *A fortiori*, there would have been no true proposition that P. Thus a counterexample to (2+) seems ruled out. Given that (2+) holds, so does (2).

It is sometimes said that a proposition can be true *of* a possible world without being true *in* that world.[3] We can express propositions in one world about another world. Thus a proposition might be true of a possible world without existing in that world. But this idea does not address the case for (2+), for (2+) does not say that the proposition exists in any possible world of which it is true. We could paraphrase (2+) thus: for any possible world w, if the proposition that P would have been true if w had obtained, then the proposition that P would have existed if w had obtained. We can abbreviate that by saying that for any possible world w, if the proposition is true in w then the proposition exists in w. The antecedent concerns truth in w, not truth of w, so the distinction poses no threat to (2+).

Does the distinction pose a threat to (1+)? Someone might suggest replacing (1+) by the schema: for any possible world w, the proposition that P is true of w if and only if, in w, P. In particular, the proposition that I do not exist is true of w if and only if, in w, I do not exist. That does not yield (1), which requires that if, in w, I do not exist then, *in w,* the proposition that I do not exist is true. Thus the argument for (4) and (5) would fail. Since (1+) was motivated by the use of the condition that necessarily the conclusion is true if the premises are as the standard for a valid argument, the replacement of (1+) would require a corresponding replacement of that standard by the condition that the conclusion is true *of* any possible world *of* which the premises are true.

But now a threat of circularity emerges. For the concept of a possible world is a technical one, itself in need of explanation. What is a possible world? A natural answer is that it is a consistent and complete class of propositions. A class X of propositions is consistent if and only if for every pair of contradictory propositions p and $\sim p$, there is not both a valid argument from X to p and a valid argument from X to $\sim p$. X is complete if and only if for every pair of contradictory propositions p and $\sim p$, there is either a valid argument from X to p or a valid argument from X to $\sim p$. Thus the concept of a possible world is explained in terms of the concept of validity. But, on the envisaged view, the concept of validity is explained in terms of the concept of a possible world!

That objection is not immediately decisive, for there are rival explanations of the concept of a possible world. But the distinction between truth in a world and truth of a world faces another problem. We say that the open sentence 'x is a capital city' is true of London and not of Oxford because London is a capital city and Oxford is not. The true-of relation between an open sentence and an object depends on the assignment of the object to a variable in the open sentence. Different propositions result from different assignments. The

proposition that London is a capital city is true, the proposition that Oxford is a capital city false. Can we apply this model to the postulated true-of relation between propositions and worlds? Consider the contingently true proposition that Blair was Prime Minister in 2000. It is supposed to be true of the actual world @ and false of some other possible world w. On the model, the sentence contains a tacit variable; if @ is assigned to the variable, a truth results, if w is assigned, a falsehood. But that does not make the resulting propositions contingent. There is genuine contingency in how things are only if, once values have been assigned to *all* variables, the resulting proposition could still have differed in truth-value. It is not contingent that Blair was Prime Minister in 2000 in @ and that he was not Prime Minister in 2000 in w. What is contingent is simply that Blair was Prime Minister in 2000. Its contingency requires it not to have a variable waiting to be assigned a world. The reply 'But contingency just is variation in truth-value with variation in the value of the world variable' betrays a failure to grasp what contingency is.

According to David Lewis's modal realism, contingency consists in differences between possible worlds, which are conceived as equally real, mutually disconnected spatiotemporal systems.[4] Consider the common sense claim 'It is contingent that there are no talking donkeys' ($\sim \exists x(Tx \ \& \ Dx) \ \& \ \Diamond \exists x(Tx \ \& \ Dx)$). If one interprets the quantifier as unrestricted, modal realism makes the claim false by making its first conjunct false: the modal realist holds that there really are talking donkeys, in spatiotemporal systems other than ours. For modal realism to make the claim true as uttered in the actual world, one must interpret the quantifier as implicitly restricted to the objects in a world. Our spatiotemporal system contains no talking donkeys but, on Lewis's account, other spatiotemporal systems do contain talking donkeys. The restricted quantifier is given an implicit argument place for a world. Intuitively, however, a difference between spatiotemporal systems in itself constitutes

no contingency at all. For all that has been said, it is necessary that another spatiotemporal system contains talking donkeys while this system does not, in which case the matters at issue are not contingent. A necessary difference between spatiotemporal systems constitutes no contingency. Even if there are mutually disconnected spatiotemporal systems such as Lewis postulates, they are not the distinctive subject matter of modal discourse. They are simply more of what there is, about which we can ask genuinely modal questions: for instance, whether there could have been more or fewer spatiotemporal systems than there actually are.

To put the point another way, the modal realist claims that one can fully specify how things are in an extensional language without modal operators, restricted quantifiers or other expressions indexed to worlds. Yet, still according to modal realism, nothing stated in that language is contingent. Thus the view implies that it is not genuinely contingent how things are. Of course, the view also implies that one may truly say 'It is contingent that there are no talking donkeys'; that shows that it is also wrong about the truth-conditions of modal statements. Lewis misidentifies contingency as a special kind of indexicality, just as Berkeley misidentified material objects as special groups of sense impressions.

The modal realist's postulation of an implicit argument place for worlds is not faithful to our understanding of modal vocabulary. Since there is no argument place for worlds of the required kind, the postulated true-of relation between propositions and worlds does not behave like the standard true-of relation. Absent some special explanation, the postulate rests on a false analogy.

We can grasp a distinction between truth in a world and truth of a world for utterances. An utterance of the sentence 'There are no utterances' in this world is true of a world in which there are no utterances. For the way the utterance says things to be is the way they are in that world. But that is just a notational variant of the point that the utterance

10

actually expresses a proposition which would be true if that world obtained; in other words, the proposition is true in that world. The utterance need not exist in that world in order to be true of it because the proposition which it expresses in this world exists in that one. We need not carry the utterance across from this world to that one precisely because we can carry the proposition across instead. There is the illusion of a distinction between truth in a world and truth of a world for propositions because we appear to be able to model such a distinction on a corresponding distinction for utterances, forgetting that the presence of the latter depends on the absence of the former. On critical reflection, both (1+) and (2+) withstand the threat from the purported distinction.

Finally, consider premise (3). It too is a special case of a more general principle, roughly, that a proposition about an item exists only if that item itself exists:

(3+)   Necessarily, if the proposition that P(o) exists then o exists.

Here 'o' is to be replaced by a referring singular term such as a simple demonstrative, indexical or ordinary proper name, whose function is to refer in a given context to a particular object (o) and enable us to say something about it; 'o' is not to be replaced by a definite description. 'P(o)' is to be replaced by a sentence which has that singular term as a constituent and says that something is the case. In (3), 'I' replaces 'o' and '... do[es] not exist' replaces 'P(...)'; 'I' is a referring singular term of the requisite kind, a non-descriptive indexical and a constituent of 'I do not exist'.

A simple defence of (3+) is based on the Russellian view that the proposition that P(o) is a structured entity of which one constituent is the object o. For example, the proposition that that dog is barking is a complex consisting of that dog and the property of barking. On

this view, the terms that may replace 'o' are *directly referential* in David Kaplan's sense; the contribution of such a term to the proposition expressed by a sentence in which it occurs is simply its referent.[5] If a structured object has a given constituent, then necessarily the former exists only if the latter is a constituent of it and therefore exists too. Since o is a constituent of the structured proposition that P(o), necessarily, the proposition that P(o) exists only if o exists. On Kaplan's view, 'I' is a paradigm of a directly referential term. It is being used as such rather than mentioned in the sentence 'I do not exist', so (3) is a genuine instance of (3+).

However, (3+) is plausible even independently of the direct reference view. For example, on a more Fregean view propositions ('Thoughts') are structured objects, but the constituent corresponding to the term 'o' is a mode of presentation of o rather than the object o itself, a sense of which o is the referent. Thus the sense of the demonstrative 'that dog' in the present context, but not that dog itself, is a constituent of the proposition expressed in the present context by the sentence 'That dog is barking'. Even so, how could something be the proposition that that dog is barking in circumstances in which that dog does not exist? For to be the proposition that that dog is barking is to have a certain relation to that dog, which requires there to be such an item as that dog to which to have the relation. The argument is quite general; it does not even require propositions to be structured objects. Necessarily, if o does not exist then there is no such item as o, so there is no such item as the proposition that P(o), so the proposition that P(o) does not exist. It is crucial to the argument that the function of the singular term 'o' is to refer to a particular object, and not merely to introduce a description, for otherwise the existence of the proposition might imply only the existence of the description, whether or not anything satisfied it. But the function of the indexical 'I' evidently is to refer to a particular object (in a given context) and not merely to introduce a

12

description. 'I' does not function like the definite description 'the actual producer of this utterance': one might be under the illusion 'I am not the actual producer of this utterance' without being under the illusion 'I am not me'.[6] Although some remarks in Frege suggest a purely descriptive conception of singular terms, more recent developments from his views acknowledge the kind of object-dependence which the present argument requires. Thus it is not only on a Russellian view that (3) is one of the instances of (3+) to which the argument applies. Necessarily, if the proposition that P(o) exists then o stands in some kind of relation to it (such as being a constituent or being the referent of a constituent), and therefore exists.

Nevertheless, a subtle objection might be made to (3+), and correspondingly to (3). The argument for (3+) assumes that when we use the phrase 'the proposition that P(o)' in speaking of a counterfactual situation (in the scope of 'necessarily'), we thereby refer to something which would have the corresponding property (of being a proposition to the effect that P(o)) *in the counterfactual situation*. But there is another possibility. Perhaps we are using the phrase 'the proposition that P(o)' to pick out the object which has that property *in the actual situation* and then talking about how things could have been with that very object in a counterfactual situation, whether or not it had the property in the counterfactual situation. In other words, the argument treats 'the proposition that P(o)' like the definite description 'the winner' in a typical utterance of the sentence 'The winner could have been someone else': with respect to a counterfactual situation the description denotes whoever won in that situation (the description has narrower scope than the possibility operator in 'could have'). An alternative is to treat 'the proposition that P(o)' like 'the winner' in a typical utterance of 'The winner could have lost': with respect to a counterfactual situation the description denotes whoever won in the actual situation (the description has wider scope than the possibility operator). Let the actual proposition that P(o) be *p*. Thus *p* actually has a relation

13

to o. Suppose, however, that in some counterfactual circumstances *p* lacks the property of being a proposition to the effect that P(o). In those circumstances, *p* might lack the relation to o, so we lose our reason for expecting the existence of *p* to necessitate the existence of o. But, on the alternative reading, (3+) says that, necessarily, if *p* exists then o exists; thus we lose our reason for accepting (3+), and with it our reason for accepting (3). Perhaps the object which actually has the property of being a proposition to the effect that I do not exist would have lacked that property if it had been true.

Although the objection may initially sound plausible, it is hard to substantiate. If the actual proposition that I do not exist would not have been a proposition to the effect that I do not exist if I had not existed, why should it have been true in those circumstances? What would its content have been? The point is general. Consider again the truth schema (1+). Let *p* be the proposition that P. On the reading which the objection requires, (1+) says that, necessarily, *p* is true if and only if P, even if *p* could have lacked the property of being a proposition to the effect that P. But in circumstances in which *p* is not to the effect that P, why should it be true if and only if P? That is, the motivation for (1+) requires that, with respect to counterfactual circumstances, the phrase 'the proposition that P' denotes something which *in those circumstances* would have the property of being a proposition to the effect that P, which is exactly what the objection treated as optional. The most natural view is that the proposition that P is essentially a proposition to that effect, so that the distinction between the two readings makes no difference to the argument. Alternatively, the phrase 'the proposition that P' might pick out a different object with respect to counterfactual circumstances, but one with the property in those circumstances of being a proposition to the effect that P. Either reading fits (3+) and (3) in fitting (1+) and (1).

It is quite unclear what could have been the proposition that I do not exist (if I had not

existed) other than the actual proposition that I do not exist. Moreover, on the natural reading of the standard for validity, 'Necessarily, if the premises are true then the conclusion is true', the descriptions 'the premises' and 'the conclusion' are understood rigidly, as denoting the same propositions with respect to all circumstances. Since that test for validity is well-motivated only if we can unpack the truth-conditions of premises and conclusion by means of something like (1+), the natural suggestion is that 'the proposition that P' is to be understood as rigidly designating something which is essentially a proposition to the effect that P.

We can make the same point in other terms. Consider the special case of (1+) in which 'P' is replaced by something of the form 'P(o)': necessarily, the proposition that P(o) is true if and only if P(o). For any possible circumstances, (1+) requires an item which in those circumstances is true if and only if P(o). That equivalence is guaranteed only if the item has a relation to o in those circumstances. If o did not exist in those circumstances, then there would be nothing for the item to have the relation to. Thus the motivation for (1+) underpins (3+) too. Since the objection to (3+) did nothing to show how (1+) could fail, the objection is not sustained.

So far, the argument for necessary existence has withstood scrutiny. Each of the three premises (1)-(3) has highly plausible grounds. Moreover, the grounds for different premises are mutually reinforcing; they do not pull in different directions in the way characteristic of sophistical arguments. That point has just been noted in one respect for (1) and (3), and will be reinforced below in another respect for (2) and (3). We should therefore take seriously the possibility that the argument is sound, its conclusion strange but true. How *can* its conclusion be true, though? What is supposed to be wrong with the apparently compelling grounds for regarding its conclusion as false?

3. We can make some progress by considering how the concept of existence was applied in the motivation for (2) and (3). In both cases, the argument was that if a given item had not existed, then there would have been no such item as it, and therefore nothing to have a property or a relation to something. Here 'property' and 'relation' are understood in a broad sense in which any predication ascribes a property or relation. If the proposition that P had not existed, there would have been no such item to be true. If the object o had not existed, there would have been no such item to be constitutively related to a proposition. Existing was taken as a necessary precondition of having any properties or relations whatsoever.

The motivation assumes that, necessarily, if x does not exist then there is no such item as x. By contraposition: necessarily, if there is such an item as x then x exists. The converse is scarcely controversial. Thus a necessary and sufficient condition for x to exist is that there be such an item as x. We can therefore symbolize 'x exists' by the familiar formula $\exists y\ x{=}y$, where the quantifier is not restricted to any particular kind of thing. In particular, it must not be restricted merely by definition to what has spatial or temporal location.[7] Call that the *logical sense* of 'exist'.

The motivation for the argument further assumes that a given object o could not have had a property or relation without existing in the logical sense, without there being such an item as o to have the property or relation. That point is sometimes challenged by appeal to past objects. For example, it is said, Trajan's Column in Rome is now a trace of the Emperor Trajan, and the name 'Trajan' refers to him, so various objects now stand in causal and semantic relations to Trajan. By the same token, Trajan now stands in causal and semantic relations to various objects. He still has relations, but does not still exist.

Such examples are not decisive. Doubtless, *in some sense* Trajan no longer exists. Specifically, he is no longer anywhere; he lacks spatial location. Although atoms which once

composed him may still be spatially located, he is not identical with those atoms. More generally, we may say that he is no longer *concrete*. But he still counts for one when we ask 'How many Emperors of Rome were there?'. Suppose that in fact there were n Emperors of Rome. The past tense formulation with 'were' of course does not mean that at some past time there were then n Emperors of Rome, for they were not all Emperor simultaneously. Rather, it means that the number of objects with the property of having been Emperor of Rome at some time or other is n. If there are m apples in the bowl, then the number of objects with the property of either being an apple in the bowl or having been Emperor of Rome is $m+n$. Whatever can be counted exists at least in the logical sense: there is such an item. Past objects are no counterexamples to the principle that having properties or relations entails existing in at least the minimal sense. 'Trajan does not exist' is true when 'exist' is used in the nonlogical sense of concreteness, not when it is used in the logical sense. Existence in the sense of concreteness is of crucial significance for metaphysics; for logic it is just one more property, which objects may have or lack.

Fictional objects threaten the argument still less. The question is whether some object could have had a relation without existing in the logical sense. A positive answer could not be supported by a claim such as 'Satan does not exist, but he has the relation of being-worshipped-by to Satanists'. For the name 'Satan' yields a verifying instance of the claim that something could have had a relation without existing in the logical sense only if the name refers to something, in which case 'Satan exists' is true in the logical sense of 'exists' in the very world in which the name putatively refers to something which has the relation of being-worshipped-by to Satanists: the actual world.

Nonexistence in the logical sense is a very radical matter indeed, for it entails having no properties or relations whatsoever. It is not obvious that I could have failed to exist *in the*

17

logical sense. The argument for (5) depends on reading 'exist' in the logical sense, for that is the one needed to make (2+) and (3+) plausible. Its conclusion is therefore to be interpreted as the claim that it is necessary that I am something. What is surely not necessary is that I 'exist' in the sense of being concrete:

(6)   Possibly, I am not concrete.

From (5) and (6) we can deduce by standard modal reasoning that existing in the logical sense does not necessitate being concrete:

(7)   Possibly, I exist and I am not concrete.

We should not assume that the only alternative to being concrete is being abstract. When Trajan died, he did not become an abstract object, although he ceased to be concrete. He did not become the value of some abstraction operator. He became something neither abstract nor concrete, but something that had once been concrete. Trajan is an ex-concrete object. Similarly, if my parents had never met, I would have been something neither abstract nor concrete, but something that could have been concrete.[8] I would have been a possible concrete object. I would not have been a physical object, but I would have been a possible physical object.

We must be clear what we mean by phrases of the form 'possible F', such as 'possible physical object'. They are sometimes given a *predicative* reading, on which '*x* is a possible F' is equivalent to the conjunction '*x* is possible and *x* is an F', just as '*x* is a spherical stone' is equivalent to the conjunction '*x* is spherical and *x* is a stone'. Then '*x* is possible' is in turn

read as something like '*x* could exist'. On the predicative reading, a possible physical object is a physical object, one which could exist. Thus if each physical object could exist, the possible physical objects are simply the physical objects. But the predicative reading is irrelevant to the preceding claims, for they imply that I could have been a possible physical object without being a physical object. The relevant reading here is an *attributive* one, on which '*x* is a possible F' is equivalent to 'it is possible that *x* is an F' ($\Diamond$F*x*), just as '*x* is a pretended cure' is equivalent to 'it is pretended that *x* is a cure', not to the scarcely intelligible '*x* is pretended and *x* is a cure'. On the attributive reading, a possible physical object need not be a physical object; it may qualify simply because it could have been a physical object. We may define a *merely* possible F in the attributive sense as a possible F that is not an F. For example, if you are not a government minister but could have been, then you are a merely possible government minister. If my parents had never met, I would have been a merely possible physical object. Since I am actually a physical object and actuality implies possibility, I am a possible physical object; but I am not a merely possible physical object.

Someone might still ask 'What *kind* of thing is a merely possible physical object?'. The answer that 'possible physical object' already demarcates a kind is liable to elicit the complaint 'I asked what it *is*, not what it *could have been*'. Presumably, the complainant wants an answer in non-modal terms. But what justifies the presumption that there should be such an answer? When we think of past physical objects, we are content to classify them in terms of what they were; we do not insist on a classification in terms of what they are now, without reference to the past. Why should possible physical objects be different?

One source of unease may be an inability to imagine what a merely possible physical object would be like. But what exactly is it that we cannot do? We can intellectually grasp the concept of a merely possible physical object; in effect it has just been defined by the open

19

sentence '*x* is not a physical object but *x* could have been a physical object'. Consequently, we can formulate the general existential thought that there are merely possible physical objects. We cannot perceptually imagine a merely possible physical object as such, just as we cannot imagine a number, but that has no more tendency to show that there are no merely possible physical objects than it has to show that there are no numbers. It is impossible to perceive numbers or merely possible physical objects, for they lack spatiotemporal location and causal relations.[9] Perception does not exhaust our contact with reality; we can think too. We have been given no reason to accept the empiricist prejudice that what cannot be perceptually imagined is thereby suspect. Of course, good questions arise about our ability to grasp and apply modal concepts; if something is not actually the case, how do we know whether it could have been the case? Since we do have such knowledge, those questions must have answers. It is still far from clear what those answers are. We should not assume that they will make knowledge of (1), (2) and (3) more problematic than knowledge of more familiar claims of necessity, or knowledge of (6) more problematic than knowledge of other claims of possibility. From those premises, and others like them, the relevant conclusions follow.

Even the claim that merely possible physical objects are unperceivable must be formulated with care. What is true is the *de dicto* claim that it is impossible that someone perceives some merely possible physical object. But the corresponding *de re* claim is false, that for some merely possible physical object it is impossible that someone perceives it. For a merely possible physical object could have been a physical object; in normal cases it could have been a perceived physical object. Merely possible physical objects are unperceivable only in the sense in which unperceived physical objects are unperceivable. Of course, for physical objects the difference between being perceived and being unperceived may be purely extrinsic, whereas the difference between being a physical object and being a merely possible

20

physical object is intrinsic in some sense. That suggests a different objection.

On the envisaged view, two very different states are possible for one object. It is capable of being an embodied person, knowing, feeling and acting in space and time. It is also capable of being a merely possible person, disembodied, spatiotemporally unlocated, knowing nothing, feeling nothing and doing nothing. Is so radical a difference in properties consistent with the identity of the object? But the two sets of properties are not wholly disparate. The person actualizes the potential to have properties characteristic of a person. The merely possible person has the unactualized potential to have such properties. What they share is the potential. Why should that not suffice?

Consider identity and distinctness for persons and merely possible persons. If person A is somewhere person B is not, then A is distinct from B. But if A is a merely possible person, then A is nowhere, and therefore does not satisfy that condition for distinctness from B. Nevertheless, A still *could have been* somewhere B was not; in those circumstances, A would have been distinct from B. By the necessity of identity, if A could have been distinct from B then A is distinct from B, for if A and B are identical and A could have been distinct from B then A could have been distinct from itself (by the indiscernibility of identicals), which is impossible. Thus the mere potential for A to be somewhere B is not suffices for the actual distinctness of A and B. Quite generally, suppose that, necessarily, Fs are identical if and only if they stand to each other in a relation R. Then, necessarily, possible Fs are identical if and only if they could both be F and stand to each other in R. For let A and B be possible Fs. If they are identical then, in possible circumstances in which A is an F, B is the same F and they stand to each other in R. Conversely, if A and B could both be F and stand to each other in R, then they could be identical, and therefore are identical, by the necessity of distinctness (if A and B are distinct, they could not have been identical).[10] To the extent to

21

which one can state identity conditions for Fs, one can state identity conditions in correspondingly modalized terms for possible Fs.

A different sort of complaint about the envisaged view is that it has a massively inflationary effect on our ontological commitments. Since any human sperm S and egg E could have united to result in a person, who would have existed necessarily; therefore, given the view, there actually is a possible person who could have resulted from S and E.[11] Arguments of this type yield an infinity of merely possible animals, vegetables and minerals. Is this an objectionable cluttering or crowding of our ontology? Of course, the spatial metaphor of clutter is misleading, for it is crucial to the new objects that they lack spatial location. The nonmetaphorical complaint is that the theory commits us to too many objects. At this point appeals may be made to Ockham's Razor: 'Do not multiply entities without necessity'. Of course, it is objectionable to postulate without reason that there are entities of some kind. It is also objectionable to postulate without reason that there are no entities of some kind; it is objectionable to make any postulate without reason. But merely possible animals, vegetables and minerals have not been postulated here without reason; the argument explained above for necessary existence gives a reason for postulating them. Underlying Ockham's Razor we can also discern the insight that simplicity of theory is a virtue. But the simplicity of a theory is not proportional to the size of its ontology. Zermelo-Fraenkel set theory postulates a high infinity of sets but is comparatively simple; with *ad hoc* modifications one could massively reduce the size of its commitments while massively increasing its complexity. The proposed conception of necessary existence effects a major simplification of both the proof theory and semantics of quantified modal logic. It simplifies the proof theory because it validates certain formulas (such as the Barcan Formula and its converse), which are derivable in the simplest axiomatizations; other views invalidate those

formulas, and must therefore complicate the proof theory in order to block their derivation. The conception simplifies the semantics because it obviates the need to associate each possible world with a domain of quantification containing just those objects which exist at that world. If Ockham's Razor amounts to a preference for simple theories, it tells strongly in favour of the proposed conception. As for any preference for theories that estimate numbers of entities as low rather than high as such, it seems to carry no independent weight. For example, if two cosmological theories of equal simplicity estimate the number of galaxies in the universe, and one estimate is twice the other, that by itself seems to be no reason at all for preferring the theory with the smaller estimate.[12]

There are few knockdown arguments in philosophy, and the foregoing argument for necessary existence is not one of them. Someone determined to reject its conclusion at all costs can surely reject one of its premises, perhaps by abjuring the very idea of a proposition.[13] The argument is directed to those with more open minds, who are willing to rethink the status of its superficially implausible conclusion in the light of the argument itself and of the proposed metaphysics. The cost of rejecting a premise may be higher than the cost of accepting the conclusion.

On the view defended here, an object is essentially a locus of potential. How far it actualizes its potential may be a radically contingent matter. But the existence of that object with that potential is wholly noncontingent. Logical properties and relations such as existence and identity are not subject to contingency.[14]

Notes

1   The argument reworks for another purpose material from A. N. Prior, *Past, Present and Future* (Oxford: Clarendon Press, 1967), 149-151. See also Kit Fine's 'Postscript' to A. N. Prior and K. Fine, *Worlds, Times and Selves* (London: Duckworth, 1977), 149-150, and his 'Plantinga on the Reduction of Possibilist Discourse', *Alvin Plantinga*, J. E. Tomberlin and P. van Inwagen (eds.) (Dordrecht: D. Reidel, 1985), 160-180, and A. Plantinga, 'On Existentialism', *Philosophical Studies* **44** (1983), 9-10, and 'Reply to Kit Fine', *Alvin Plantinga*, op. cit., 341-349. My interest in the argument was aroused by the work of my pupil, David Efird.

2   In his defence of the unnecessitated principle, Paul Horwich suggests that the necessitated version might be derivable from the assumption that the unnecessitated version is explanatorily fundamental: *Truth*, 2nd ed., (Oxford: Clarendon Press, 1998), 21.

3   Fine makes a similar distinction between outer and inner truth, 'Plantinga on the Reduction of Possibilist Discourse', op. cit., 163.

4   D. K. Lewis, *On the Plurality of Worlds* (Oxford: Blackwell, 1986).

5   D. Kaplan, 'Demonstratives: An Essay on the Semantics, Logic, Metaphysics, and Epistemology of Demonstratives and Other Indexicals', *Themes from Kaplan*, J.

Almog, J. Perry and H. Wettstein (eds.) (Oxford: Oxford University Press, 1989).

6   Even the description 'the producer of this utterance' contains the demonstrative 'this utterance', which is not purely descriptive, but 'I' does not refer to the utterance.

7   For a defence of unrestricted quantification see my 'Existence and Contingency', *Aristotelian Society* **100** (2000), 117-139.

8   The underlying assumption is that since I am concrete, it is necessarily possible that I am concrete (so even if I had not been concrete it would still have been the case that I could have been concrete). This is an instance of the so-called Brouwerian principle $p \supset \Box\Diamond p$ in modal logic, which is plausible when $\Box$ and $\Diamond$ stand respectively for metaphysical necessity and metaphysical possibility. The principle corresponds to the symmetry of the accessibility relation in possible worlds semantics. It is a theorem of the attractively simple modal system S5, a good candidate for the logic of those notions, but also of much weaker systems without the S4 principle $\Box p \supset \Box\Box p$, which corresponds to the transitivity of accessibility.

9   If the objects of perception are not all physical then the objection will need to be stated with more care, but that is the objector's problem.

10   On the necessity of identity and distinctness see my 'The Necessity and Determinacy of Distinctness', *Essays for David Wiggins: Identity, Truth and Value*, S. Lovibond and S. Williams (eds.) (Oxford: Blackwell, 1996).

11   The reasoning again depends on the Brouwerian principle.

12   The view defended in this paper is put forward in my 'Necessary Identity and Necessary Existence', *Wittgenstein — Towards a Re-evaluation: Proceedings of the 14th International Wittgenstein-Symposium*, vol. 1, R. Haller and J. Brandl (eds.) (Vienna: Holder-Pichler-Tempsky, 1990), and elaborated and given further support in my 'Bare Possibilia', *Erkenntnis* **48** (1998), 257-273, and 'The Necessary Framework of Objects', *Topoi* **19** (2000), 201-208. For a somewhat similar view see B. Linsky and E. Zalta, 'In Defense of the Simplest Quantified Modal Logic', *Philosophical Perspectives 8: Logic and Language*, J. Tomberlin (ed.) (Atascadero: Ridgeview, 1994) and 'In Defense of the Contingently Nonconcrete', *Philosophical Studies* **84** (1996): 283-294.

13   If the rejection of propositions nevertheless permitted some way of simulating them, the argument for necessary existence might still be simulated by a sound argument to the same conclusion. For example, quantification over propositions might be simulated by nonsubstitutional quantification into sentence position; for the latter see my 'Truthmakers and the converse Barcan formula', *Dialectica* **53** (1999), 253-270.

14   Thanks to audiences in Oxford and at the Royal Institute of Philosophy for comments on earlier versions of this paper.