HAECCEITISM, ANTI-HAECCEITISM AND POSSIBLE WORLDS

By Bradford Skow

Possible-worlds talk obscures the debate about haecceitism, rather than clarifying it. I distinguish haecceitism and anti-haecceitism from other doctrines which sometimes go under those names. Then I defend the claim that any definition of 'haecceitism' using possible-worlds talk depends for its correctness on a substantive theory of the nature of possible worlds. This explains why using possible-worlds talk when discussing haecceitism causes confusion: the term will mean different things to parties who depend on different presuppositions about possible worlds.

I. POSSIBLE WORLDS AT WORK

Possible-worlds talk has proved its usefulness many times over. Even philosophers who deny that there are any such things as possible worlds find it hard to avoid talking as if there were. Some propositions are easier to express, some distinctions easier to capture and some debates easier to conduct using possible-worlds talk. Debates about supervenience are paradigm cases. It is much easier to grasp the differences between the many definitions of 'supervenes' when those definitions are given in possible-worlds talk.

But sometimes possible-worlds talk inhibits communication rather than improving it. Here the debate about haecceitism is the paradigm. Often when parties to this debate begin using possible-worlds talk, they end either by talking past each other, or by arguing about the truth of a different thesis altogether.

In this paper I aim to do two things. First, I shall distinguish haecceitism and anti-haecceitism from other doctrines that sometimes go under those names. Secondly, I shall defend the claim that there are no non-tendentious definitions of 'haecceitism' and 'anti-haecceitism' using possible-worlds talk. That is, any definition of 'haecceitism' using possible-worlds talk depends for its correctness on a substantive theory of the nature of possible worlds. This explains why using possible-worlds talk when discussing haecceitism causes

confusion: if the parties to the discussion presuppose different theories of the nature of possible worlds, then they will mean different things by 'haecceitism'.

II. HAECCEITISM AND ANTI-HAECCEITISM DEFINED

Could my brother and I have switched qualitative roles? (My qualitative role is just the conjunction of all of my qualitative properties, including my relational qualitative properties.) Could things be just as they actually are, except that I do not exist while someone who does not actually exist plays the qualitative role I actually play? Someone who answers 'Yes' to either of these questions is a haecceitist. Someone who answers 'No' to both, and to other questions like these, is an anti-haecceitist.

These are not yet definitions of 'haecceitism' and 'anti-haecceitism'. I have not said what 'other questions like these' means. It is easiest to define 'anti-haecceitism' first, and then define 'haecceitism' as its denial. What, then, is the definition of 'anti-haecceitism'? I shall approach this question obliquely. First I shall discuss an important consequence of anti-haecceitism. This consequence is a generalization of one answer to the questions I asked in the previous paragraph. Anti-haecceitism entails

1.   It is not possible for things to be just as they actually are, qualitatively, but differ in some non-qualitative way.

I can bring out what (1) means in this way. Imagine a language so powerful that it contains complete descriptions of the universe. The part of this language that can express only qualitative propositions (that is, contains no proper names or equivalent semantic devices) will then contain *complete purely qualitative* descriptions of the universe. Such a purely qualitative description will say everything there is to be said about how many things there are and what qualitative properties and relations those things instantiate, without saying just which individuals they are. (I shall sometimes use 'qualitative sentence' for sentences that express purely qualitative propositions.) If we imagine such a language, and if ϕ is a complete qualitative description of the universe in the qualitative part of the language, and ψ is a complete description of the universe in that language (ψ is not purely qualitative), then a more precise way to state (1) is

2.   Necessarily, if ϕ, then ψ.

As I have said, anti-haecceitism entails (1) and (2). But anti-haecceitism is stronger than (1). It is a kind of generalization of (1). So this is my official

definition: anti-haecceitism is the claim that (1) would still express a truth, no matter what the universe were like.

There are other more formal ways available of stating anti-haecceitism, which make use of some non-standard modal operators. Here are two such ways.

First, suppose we understand the actuality operator in the way Hodes does, as an operator that temporarily undoes the effect of just the most recent modal operator governing it, rather than the effect of all modal operators governing it.[1] Then anti-haecceitism can be stated as

3.   Necessarily, it is not possible for things to be just as they actually are, qualitatively, but differ in some non-qualitative way.

Secondly, using Forbes' indexed modal operators,[2] anti-haecceitism can be stated as

4.   Necessarily$_1$, it is not possible$_2$ for things to be just as they actually$_1$ are, qualitatively, but differ in some non-qualitative way.

Here the 'actually$_1$' undoes the effect of the modal operators governing it which have a different subscript – in this case, the single occurrence of 'possible$_2$'.

This completes my presentation of definitions of 'haecceitism' and 'anti-haecceitism'. But before I move on, two parenthetical comments.

First, I have actually defined 'weak anti-haecceitism' and 'strong haecceitism'. Strong anti-haecceitism is the view that for any possible individual b there is a qualitative property P such that necessarily, for all x, x has P iff x is identical with b. Strong anti-haecceitism entails weak anti-haecceitism. It also entails the principle of the identity of indiscernibles:

PII.   Necessarily, if x and y share all qualitative properties, then $x = y$.

In my view, this makes strong anti-haecceitism an implausible and so uninteresting doctrine. I shall confine my attention to weak anti-haecceitism, which does not entail (PII).

Second comment: David Lewis proposes a definition of 'haecceitism' in *On the Plurality of Worlds*. I discuss Lewis' definition below in §IV. But to prevent confusion, I should say here that my definition is different from Lewis'. I also think that my definition fits better the way the term was used before Lewis wrote his book; I address this topic in §IV as well.

---

[1] H. Hodes, 'On Modal Logics which Enrich First-Order S5', *Journal of Philosophical Logic*, 13 (1984), pp. 423–54.
[2] G. Forbes, *Languages of Possibility* (Oxford: Blackwell, 1989).

III. POSSIBLE WORLDS

I have struggled in the attempt to define 'anti-haecceitism' without using possible-worlds talk. It might seem that it would be easier to provide a definition if I gave in and used possible-worlds talk.

What might a definition couched in possible-worlds talk look like? Here is a route to one such definition. Anti-haecceitism looks like a supervenience claim, the claim that, roughly speaking, the non-qualitative facts supervene on the qualitative facts. In an unguarded moment, someone might express this supervenience claim in possible-worlds talk as

5. It is not true that there are distinct but qualitatively indiscernible possible worlds.

Why is (5) a tempting way to express anti-haecceitism? The reasoning goes like this: suppose haecceitism is true. Then it is possible for things to differ in some non-qualitative way, while being just as they actually are, qualitatively. So there is a possible world w that is distinct from the actual world; and since 'Things are just as they actually are, qualitatively' is true at w, w is qualitatively indiscernible from the actual world. So if haecceitism is true, then there is a possible world that is distinct from the actual world, but qualitatively indiscernible from it. The argument for the converse is similar, and (5) is a generalization of the conclusion.

This is a tempting argument for someone who is not being too serious about the nature of possible worlds, and is merely using the possible-worlds talk that appears in (5) as a shorthand way to express anti-haecceitism. But it is a very bad argument to address to philosophers who *are* serious about the nature of possible worlds. (5) is a tendentious definition of 'anti-haecceitism': whether it is equivalent to anti-haecceitism depends on which theory of possible worlds is true.

This claim requires support. The argument I gave to show that (5) is equivalent to anti-haecceitism had two premises:

6. Something that is actually false is possibly true if and only if there is a possible world distinct from the actual world
7. If 'Things are just as they actually are, qualitatively' is true at a possible world, then that possible world is qualitatively indiscernible from the actual world.

But these premises are controversial. There are theories of possible worlds according to which at least one of them is false. According to David Lewis'

version of modal realism, (6) is false, though (7) is true; I shall come to Lewis' theory later. For the moment I shall focus on theories according to which (7) is false.

Suppose possible worlds are *sets*, maximally consistent sets of sentences (sentence types) in some language. Suppose also that the language in question is capable of expressing non-qualitative facts, so that it contains proper names. This is one account of the nature of possible worlds.

Is (7) true according to this theory? If possible-worlds talk is being taken literally, not as a convenient fiction, then what is (7) saying? Presumably, if possible worlds are sets of sentences, then 'Things are just as they actually are, qualitatively' is true at w iff w and the actual world contain all the same qualitative sentences. So if this theory is true, (7) says that if two sets (which are possible worlds) contain all the same qualitative sentences, then those two sets are themselves qualitatively indiscernible. So the consequent of (7), read literally, says something about the qualitative features of certain *sets*. (Since in many contexts possible-worlds talk is not being taken literally, it can be hard to hear this reading of (7).)

This is what (7) says; and what (7) says is false according to this theory of possible worlds. Two sets that contain all the same qualitative sentences may fail to be qualitatively indiscernible, because there may be some non-qualitative sentence ('FF exists', perhaps) that is a member of one but not the other. Surely sets that differ over whether they contain some particular sentence differ qualitatively.

To be clear about this, my last assertion follows from the claim

8. For any sentence type, the property of having that sentence type as a member is a qualitative property (of sets).

Using (8), I argued that (7) is false. Of course, (7) is just a premise in an argument to establish that (5) is equivalent to anti-haecceitism; showing that a premise is false does not show that the conclusion is false. But in fact the conclusion *is* false. Given (8), the theory of possible worlds I am discussing *entails* (5). (Here is the argument: if two possible worlds – that is, according to this theory, two sets of a certain kind – are distinct, then (by the axiom of extensionality) they do not have all the same sentence types as members. By (8), sets that do not have all the same sentence types as members are qualitatively different. So distinct sets (of this kind) differ in their qualitative properties, which is what (5) says.) But the theory that possible worlds are sets of sentences does not entail anti-haecceitism. In fact, this theory is a good home for haecceitists. Take the set that (according to this theory) is the actual world; replace all occurrences of my name in each sentence that is a member of the actual world with my brother's name, and *vice versa*; if this

new set of sentences is a possible world, then haecceitism is true, even though (5) is also true. So, given this theory of possible worlds, (5) is not equivalent to anti-haecceitism.

(I say 'if this new set of sentences is a possible world', because two philosophers might agree that possible worlds are sets of sentences but disagree about just which sets of sentences are possible worlds. In fact, an anti-haecceitist who thinks that possible worlds are sets will deny that two sets that differ only over a permutation of names are both possible worlds.)

I have not yet said anything in defence of (8). (8) follows from the claim that distinct sentence types are qualitatively distinct and the claim that the membership relation is a qualitative relation. Is the membership relation a qualitative relation? This is not an easy question to answer. Sometimes we can tell whether a relation is qualitative by looking at its analysis. If its analysis mentions a particular individual, then it is not qualitative. That is how we know that the relation *being in the same room as Bob Dylan* is not qualitative. But the membership relation is a fundamental relation; it has no analysis. So we cannot use this procedure to determine whether it is qualitative. I think the best defence of (8) is the claim that all fundamental relations are qualitative. At any rate, the point I want to make can be made even if (8) is false. A philosopher who thinks that possible worlds are sets of sentences and who thinks that (8) is true should think that (5) is true; but he is not thereby an anti-haecceitist.

So (5) is not equivalent to anti-haecceitism, and the argument I presented to show that it is fails, because one of its premises, (7), is false if possible worlds are sets of sentences. Now that I have dealt with a particular case, I can express some general scepticism about the possibility of giving a non-tendentious definition of 'anti-haecceitism' using possible-worlds talk. A definition of 'anti-haecceitism' in possible-worlds talk must translate my definition of haecceitism, which is couched in ordinary modal discourse, into possible-worlds talk. But there is huge variation in what different theories say about the nature of possible worlds and about which sentences in ordinary modal discourse are equivalent to which sentences couched in possible-worlds talk. Why think that there is a definition of 'anti-haecceitism' in possible-worlds talk which is insensitive to these huge differences?

Here I expect to encounter complaints. Someone might say

True, (5) is not equivalent to anti-haecceitism. But that does not mean that there is *no* non-tendentious definition of 'anti-haecceitism' using possible-worlds talk. (5) is just carelessly worded (or taken in too literal-minded a way), and that was why it led to trouble. Possible worlds are *representations*. All theories of possible worlds agree on this. But (5), taken

literally, is not (directly) about the representational properties of possible worlds. This is why it is a bad definition. It is correct to point out that two possible worlds can agree on the qualitative features they represent the universe as having, even though those worlds themselves are not qualitatively indiscernible. One just needs to formulate anti-haecceitism directly in terms of the representational properties of possible worlds.

The idea is that if we formulate anti-haecceitism in terms of the representational properties of possible worlds, then the definition will be insensitive to the disagreements about the nature of possible worlds. Philosophers who disagree about whether possible worlds are abstract or concrete, whether they are sets of sentences or represent by magic, can still agree on which possibilities are represented. In the next section I examine an attempt to define 'anti-haecceitism' in this way.

IV. LEWIS' DEFINITION

Talk about what a possible world represents is talk about which sentences are true according to that possible world. As before, I take for granted a distinction between qualitative and non-qualitative sentences. Then we might try to express anti-haecceitism as

9.   For any possible worlds w_1 and w_2, if all the same qualitative sentences are true according to w_1 and w_2, then all the same non-qualitative sentences are true according to w_1 and w_2.

Why is (9) a tempting way to express anti-haecceitism? The reasoning goes like this: suppose haecceitism is true. Then it is possible for things to differ in some non-qualitative way, while being just as they actually are, qualitatively. But on any theory of possible worlds this is going to be equivalent to the claim that there is a possible world w such that (i) all the same purely qualitative sentences are true according to the actual world and w; but (ii) some non-qualitative sentence is true according to w but false according to the actual world. Since, in general, if some sentence is true at w_1 but false at w_2, then w_1 and w_2 are distinct, it follows from (ii) that w is distinct from the actual world. This appears to establish that haecceitism entails the denial of (9). Similar reasoning appears to establish the converse.

   This argument that (9) is equivalent to anti-haecceitism is better than the argument in §III that (5) is equivalent to anti-haecceitism. But it fails too. It is not true that every theory of possible worlds entails principle (10) below:

10. If s is true according to w_1 and false according to w_2 then w_1 and w_2 are distinct.

For example, (10) is false on David Lewis' version of modal realism. This is because he treats *de re* modal claims using counterpart theory, and he allows things to have more than one counterpart at a possible world. So (on a very permissive counterpart relation, where all it takes to be my counterpart is to be human), my brother is a counterpart of me at the actual world. That means that 'BAS lives in Los Angeles' is true according to the actual world, even though I do not live in Los Angeles. In fact, Lewis says that 'It is possible that my brother and I switch qualitative roles' is true on his theory; so he is a haecceitist, even though (9) is true on his theory.

(It sounds odd to say ' "BAS lives in Los Angeles" is true according to the actual world, even though I do not live in Los Angeles'. It may sound less odd if one bears in mind that according to Lewis' theory, the truth-value of a non-qualitative sentence depends on the world of evaluation *and* the contextually determined counterpart relation. We are rarely in contexts in which the counterpart relation is as permissive as this.)

(9), then, like (5) before it, is a tendentious definition of 'anti-haecceitism'. It is not equivalent to anti-haecceitism, no matter what theory of possible worlds is true. Since I can think of no better way to define anti-haecceitism using possible-worlds talk, I conclude that anti-haecceitism has no non-tendentious definition in possible-worlds talk.

Lewis, nevertheless, proposes (9) as his definition of anti-haecceitism.[3] For convenience I shall use 'Lewisian anti-haecceitism' as another name for (9). Then Lewis does not treat claims like

11. It is possible for my brother and me to switch qualitative roles

as propositions which *entail* haecceitism and (so) are incompatible with anti-haecceitism, but as propositions which merely *raise a problem* for (Lewisian) anti-haecceitism. The problem they raise is this: the easiest way to construct a theory of possible worlds on which sentences like (11) are true is to construct a theory for which (10) is true and Lewisian anti-haecceitism (9) is false. But Lewis has many independent arguments against these theories (they are either theories according to which possible worlds are abstract, or versions of modal realism other than his own). So he thinks he has good reason to accept (9), but also admits that (11) has considerable plausibility; the challenge is to reconcile them. He does so by abandoning (10) and making counterpart theory more flexible, in the way I have mentioned.

---

[3] Lewis' definition appears in §4.4 of *On the Plurality of Worlds* (Oxford: Blackwell, 1986). It differs a little in wording from (9).

Lewis reports that he has managed to defend anti-haecceitism by render-
ing it compatible with the intuitions that support haecceitism. In the current
context, this is false. What Lewis has done is show that Lewisian anti-
haecceitism is compatible with haecceitism; but this is not something haec-
ceitists should be upset about.

I do not deny that the debate about whether Lewisian anti-haecceitism is
true is an interesting and important one, at least for modal realists. In that
debate Lewisian anti-haecceitism is certainly the more attractive position. If
facts about which non-qualitative sentences are true according to a possible
world did not supervene on the qualitative features of that world (and all
other worlds too), then these facts would be deeply mysterious.[4] Better to
find a way to avoid having to say that these are brute facts.

Lewis seems to have thought that Lewisian anti-haecceitism is what
people had been arguing about all along. (He presents his definition as an
interpretation of the doctrine Kaplan called 'anti-haecceitism'.[5]) But he
is wrong. Interesting as the debate about Lewisian anti-haecceitism is, it is
different from the debate over whether sentences like (11) are true. That
debate too needs a name; and it had a name before Lewis published his
book – it was called 'the haecceitism debate'. That debate was alive and well
before the widespread use of possible-worlds talk in contemporary meta-
physics;[6] and it continues to be alive and well today, despite Robert Adams'
powerful arguments against anti-haecceitism.[7]

Here is an example from the philosophy of space and time. Leibniz
argued that if space exists, then it is possible for each thing to be located one
foot to the left of where it actually is, at each time.[8] Leibniz also claimed that
this is not in fact possible, and concluded that space does not exist. Some
substantivalists now think that the best way to respond to this kind of modal
argument is to accept some version of anti-haecceitism and reject the first

---

[4] Assuming, that is, that some version of modal realism with overlap is false. According to
modal realism with overlap, I do not just exist according to many possible worlds, I exist *in*
many possible worlds. So an individual at another possible world represents me if and only if
he *is* me. There is nothing mysterious about this account of representation *de re*, even though it
does not analyse representation *de re* in qualitative terms. Kris McDaniel defends a version of
modal realism with overlap in 'Modal Realism with Overlap', *Australasian Journal of Philosophy*,
82 (2004), pp. 137–52.

[5] D. Kaplan, 'How to Russell a Frege–Church', *Journal of Philosophy*, 72 (1975), pp. 716–29.

[6] See, e.g., N.L. Wilson, 'Substances without Substrata', *Review of Metaphysics*, 12 (1959),
pp. 521–39; A.N. Prior, 'Identifiable Individuals', *Review of Metaphysics*, 13 (1960), pp. 684–96. I
am no historian of the debate, but I am confident it existed in mediaeval philosophy.

[7] R. Adams, 'Primitive Thisness and Primitive Identity', *Journal of Philosophy*, 76 (1979),
pp. 5–26.

[8] G.W. Leibniz and S. Clarke, *Correspondence*, ed. R. Ariew (Indianapolis: Hackett, 2000).
The argument occurs in Leibniz's third letter. Leibniz actually talks of 'changing east into
west' rather than moving each body to the left.

premise of Leibniz's argument.[9] But accepting *Lewisian* anti-haecceitism is no help in responding to this argument, because Lewisian anti-haecceitism is compatible with Leibniz's first premise.

There is a distinct, but similar, argument against substantivalism: if space exists, then there is a possible world according to which each thing is located one foot to the left; but there is no such possible world; so space does not exist. Perhaps this argument, not the one in the previous paragraph, is the one Leibniz was interested in. If so, then Lewisian anti-haecceitism is helpful when responding to it.[10]

The first argument both makes better sense of things Leibniz says than the second one does and is more interesting in its own right. Leibniz offers the principle of sufficient reason as his reason for accepting the second premise of his argument. The principle says that all contingent facts have explanations – 'nothing happens without a reason why it should be so rather than otherwise' (Leibniz and Clarke, *Correspondence*, p. 7). Substantivalists who are Lewisian anti-haecceitists still accept unexplained contingencies – they still admit that it is possible for each thing to be located one foot to the left of where it actually is, at each time – no matter what they say about possible worlds. This would not have satisfied Leibniz.

I have one final remark about Lewisian haecceitism. The haecceitism debate and the Lewisian haecceitism debate take place at different levels. The Lewisian debate is a debate in the metaphysics of modality: assuming that one believes in possible worlds, and analyses modal operators as quantifiers over possible worlds, what is it that determines which non-qualitative sentences are true according to a given possible world? The haecceitism debate, by contrast, is more 'first-order'. It is a debate not about the form of the correct theory of the truth-conditions of modal sentences, but about which modal sentences are true. Clearly these debates are logically independent. Possible-worlds talk, useful as it is in philosophy, has obscured this distinction.

University of Massachusetts