Lecture 3

Transworld Identity v. Counterparts

1. <u>Review</u>

Last time, we compared a couple different abstract modal realist positions with concrete modal realism. We focussed on issues that arose from these respective positions on the *nature* of **possible worlds**. We noted, though, that this is a distinct question from the nature of **individual possibilia**. We set the latter question aside last week. We'll turn to it now.

2. <u>Setting up the Question</u>

Let's begin by getting clear on what the question is here. We begin with the intuition that we can make true modal claims about ourselves and other individuals. So, for instance:

- (1) I could have been an astronaut.
- (2) I could have been two inches taller.
- (3) I might not have been born in Canada.

Now, just as we wondered what made our modal claims true, we can ask, what makes it true **of me** (i.e. a **de re** modal truth) that I could have been an astronaut (for instance)? It doesn't look like it can be because of any of the actual properties I have now. After all, I do not have the property of being an astronaut.

We will consider this question from within a *realist* perspective. If you're a realist about possible worlds, you're going to think that whatever makes this true of me, it has something to do with other possible worlds. But once you think this, the natural next question is: **what relation do I have to other possible worlds** such that it's true of me that I could have been an astronaut?

If you already accept other possible worlds (whether abstract or concrete), one thing you might think is that **we (at least in part) exist at other possible worlds**. And it is in virtue of my being an astronaut at some other world that it is true of me here that I *could have been* an astronaut. Very roughly, this is the **Transworld Identity Theory** (TWI).

Another thing you might think is that I don't exist or have any parts at other worlds. Instead, there are individuals at other worlds that stand in for me in some way. And since at some world, one such individual is an astronaut, it is true of me here at this world that I could have been an astronaut. Roughly, this is **Counterpart Theory** (CT).

NOTE: it is important that these two different accounts do not entail, and are not entailed by, one of either concrete modal realism or abstract modal realism! Lewis himself is a concrete modal realist and a counterpart theorist. And Plantinga (a prominent proponent of TWI) is an abstract modal realist and a transworld identity theorist. These views need not necessarily come together as a package. That is, **the different positions** *cut across* **one another**. TWI is perfectly consistent with concrete modal realism. CT is perfectly consistent with abstract modal realism. Of course, when combined with different accounts of possible worlds, the

views on individuals may have different *consequences*. But none of these consequences is logically impossible.

3. Transworld Identity Theory

TWI is the view that we, or par of us, exist at other worlds, and it is in virtue of this that we modal claims about us are true. So, let's consider the claim:

(C) Clara Schumann could have had only nine fingers.

According to TWI, (C) is true because, there is some world (call it w_1) where Clara Schumann herself exists (or existed) and has only nine fingers rather than ten. Now this view seems to raise an immediate problem: if Clara Schumann exists at w_1 and at $w_{@}$, then she has two inconsistent properties, namely the property of having only nine fingers and the property of having ten fingers. But surely nothing can have inconsistent properties!

Modal Stages

Proponents of TWI have different ways of avoiding this result. One of these we can call the **Modal Stages view**. On this view (from Yagisawa (2010) who, incidentally, is also a concrete modal realist), individuals do not wholly exist at any one world; they are made up of different parts or 'stages' that exist at different worlds. So, in the example, only part of Clara Schumann exists at $w_{@}$, and she also has a w_1 -stage among infinitely others. This resolves the problem of contradictory properties by making it the case that only distinct stages possess these properties. Clara Schumann's w_1 -stage has only nine fingers, and Clara Schumann's $w_{@}$ -stage has ten fingers. Moreover, Clara Schumann is neither nine-fingered nor ten-fingered; these are properties only possessed by some stage and not by their sum.

Compare (by way of analogy): there is no problem with ascribing inconsistent properties to different spatial parts of me. Let's say my forearm is 30cm long and my thumb is 5cm long. No inconsistency arises from attributing these properties to me, since they are only properties of distinct spatial parts of me. What is more, I am neither 30cm long nor 5cm long; it doesn't follow from some part of me having one of these properties that I—the whole—have that property.

This is an analogous view to the **perdurance** view of our persistence over time on which we do not exist wholly at any one time, but rather have different temporal parts. And these temporal parts can have conflicting properties like being 3ft tall and being 6ft tall.

World-Indexed Properties

On a different version of TWI (from McDaniels (2004) who is a concrete modal realist, and Plantinga (1974) who is an abstract modal realist), individuals exist wholly at each world where they exist, but have **world-indexed properties**. So, returning to the example again, Clara Schumann exists wholly at the actual world, and the property she bears such that (C) is true is the property of *having-nine-fingers-at-w*₁. This property is consistent with the property *having-ten-fingers-at-w*_@, and so the original problem is resolved.

Compare again the spatial case: this would be analogous to saying that I have the property of *being-30cm-long-in-the-forearm* and of *being-5cm-long-in-the-thumb*. So it is no longer the case that I have the conflicting properties of being 30cm long and being 5cm long. And this is so even though I—the whole—am the bearer of the relevant properties, and not just parts of me.

Compare the temporal case: this is analogous to the **endurance** view of our persistence over time on which we wholly exist at each moment in time. And, at each time, we have time-indexed properties like being 3ft-tall-at-t₁ and being 6ft-tall-at-t₂.

4. Problems for TWI

It is clear that the proponent of TWI can get around issues arising from the possession of (apparently) inconsistent properties. But some (e.g. Lewis) have argued that there are other more pressing problems for the view. Problems that give us reason to be counterpart theorists.

The Problem of Accidental Intrinsics

There are some properties of us that we think could have been different. These are sometimes referred to as "accidental" properties. (This is after the Aristotelian distinction between essential and accidental properties; a distinction later adopted by Scholastic thinkers.) Some of these accidental properties are intuitively intrinsic properties—i.e. they are properties their bearers possess in virtue of features internal to that bearer. They are not possessed in virtue of any external relation the bearer stands in to anything else.

Consider (C) again. The property of having a certain number of fingers looks like just such an **accidental intrinsic property**. How many fingers Clara Schumann has is plausibly a function of her physical features alone. The trouble is, on the version of TWI on which properties are world-indexed, the property of having ten fingers (say) is indexed to a world. It is rather the property of having-ten-fingers-at- $w_{@}$ (in the example). In other words, it is a **relational** property—an **extrinsic** property. Clara Schumann's actually having ten fingers is, on this view, a relation between Schumann and the actual world.

Indeed, it is going to turn out that *all accidental properties are relational* on this version of TWI. How come? Consider the following:

- P1. On TWI, for any object O that has any accidental intrinsic property P, there is some world w₁ where O has P and some world w₂ where O does not have P. (By the definition of an accidental intrinsic property, and TWI)
- P2. On the world-indexed properties view, P1 entails that O has the property Pat- w_1 . (By the world-indexed property view of TWI)
- P3. But P-at- w_1 is a relational property holding between its bearer, P, and w_1 . (By any plausible understanding of relational properties)
- C. Therefore, all (intuitively) accidental intrinsic properties are relational properties on the world-indexed property view of TWI.

Chisolm's Paradox

This objection concerns the need for individual essences on TWI. Consider the following case:

Take two distinct people: Elsa and Anna. Suppose Elsa at the actual world has the properties of being heir to the throne, having blonde hair, and having magic powers. Suppose Anna at the actual world has the properties of being second in line to the throne, having red hair, and having a preference for building snowmen. None of these properties is essential to either Elsa or Anna. Indeed these and a great many more of Elsa's and Anna's properties will be accidental. Given that this is (plausibly) the case, it follows that there is a world w_1 where Elsa is a bit more like Anna, and Anna is a bit more like Elsa. Indeed, if Elsa has the accidental properties (E_1 , ..., E_n) and Anna has the accidental properties (A_1 , ..., A_n), then there will be a series of worlds $(w_1, ..., w_n)$ such that at w_m (where $1 \le m \le n$) Elsa has the properties $(A_1, ..., A_n)$ A_m and E_{m+1} ... E_n), and Anna has the properties (E_1 ... E_m and A_{m+1} ... A_n). I.e. there will be a series of worlds where, at each subsequent world, Anna and Elsa swap one more of their accidental properties. Eventually, there will be a world where, intuitively Anna will be in the Elsa-role; she will be heir to the throne, have blonde hair, and have magic powers (among other properties Elsa has at the actual world). And Elsa will be in the Anna-role; she will be second in line to the throne, have red hair, and have a preference for building snowmen (among other properties Anna has at the actual world). But this is just to say that Anna could have been Elsa and vice versa.

What is more, nothing in the example hung on Anna or Elsa in particular, so for any two individuals I_1 and I_2 , I_1 could have been I_2 and vice versa. I could have been you. You could have been Catherine the Great. Catherine the Great could have been Clara Schumann.

At this point, TWI theorists have two options. **ONE**: say that we each have essential properties—i.e. that it is not the case that all of our properties are accidental. But just saying that we have some essential properties does not rule out the possibility of *something* else also possessing those essential properties. After all, essential properties are just those properties that I could not be without; they do not preclude others also possessing those properties. So, we need some essential property that no other individual does *or could* possess. If Elsa, say, possessed some property that is essential to her, and that could not be possessed by any one else (i.e. and that no other individual, possible or actual, possesses), then this would preclude the possibility that Anna could be Elsa. Such a property is sometimes referred to as an **individual essence**. **TWO**: accept that there is something like **bare identity** (identity independent of properties) across possible worlds in virtue of which Anna is the one in the Elsa-role at w_n .

But what could such a property be? It seems like we couldn't say anything more about such a property other than what has just been said. One concern on this objection, then, is that positing such individual essences is **ad hoc**—it is an entity posited just in order to avoid role-switching cases, and has no further justification. The second concern is that positing these essences allows for the possibility that there is a world that is exactly like the actual world, except the individual essences have been shuffled around. And this seems either incredible or gratuitous.

5. Counterpart Theory

The rival view to TWI is the view that we do not exist, either in our entirety or in part, at any other world. Instead, *de re* modal claims are true in virtue of **counterparts** actual individuals have at other possible worlds.

What is a counterpart? On Lewis' view, **counterparthood** is a (non-spatiotemporal!) similarity relation between objects at different worlds.

X is a **counterpart** of Y iff X is sufficiently similar to Y.

You'll immediately notice that this definition is very vague! But Lewis takes this to be a virtue of his account. Consider the intuition that, in some conversations it seems fair to say that "I could have been a tortoise" (e.g. imagine we're wondering what animal we would be if we weren't human); but then, in other contexts this seems like a claim I would reject (e.g. if we're wondering how my childhood could have been different). Counterpart Theory (CT) allows for this. Similarity is a vague notion that varies according to context. Thus, in some contexts, there be some Y such that Y is a tortoise and is sufficiently similar to me; and in other contexts that same Y will fail to be sufficiently similar to me.

Let's return to the example from earlier:

(C) Clara Schumann could have had only nine fingers.

On CT, this is true just in case there is some individual C* that is non-identical to Clara Schumann at another possible world, C* is sufficiently similar to Schumann, and C* has only nine fingers.

What about the Elsa and Anna case? Well, CT accepts that there is some world where a counterpart of Anna has all the actual properties of Elsa. Thus, CT accepts that Anna could have been Elsa. But there is no problem with individual essences or bare identities for the proponent of CT since Anna is not identical with anything that bears all of Elsa's actual properties.

6. Objections to Counterpart Theory

There are several different objections to CT. Here we'll consider three.

"The Humphrey Objection" (Kripke 1980)

Hubert Humphrey lost the 1968 US election to Richard Nixon. We can imagine that Humphrey might ruefully think "I could have won the election":

(H) Humphrey could have won the election.

Now, on CT, (H) is true in virtue of the fact that at some possible world, a non-identical but sufficiently similar individual to Humphrey—a counterpart of Humphrey—won the election. But, the objection goes, why should Humphrey care about that? He cares that *he*—himself!

No one else!—could have won the election. Not that someone an awful lot like him could have won the election. (H) is about Humphrey, and not about someone an awful lot like him.

Lewis' response to this objection is that it simply misunderstands the nature of CT. CT does not deny that (H) is about Humphrey. Here is Lewis in his own words:

Counterpart theory [says] that someone else - the victorious counterpart - enters into the story of how it is that another world represents Humphrey as winning, and thereby enters into the story of how it is that Humphrey might have won. Insofar as the intuitive complaint is that someone else gets into the act, the point is rightly taken. But I do not see why that is any objection, any more than it would be an objection against ersatzism that some abstract whatnot gets into the act. What matters is that the someone else, or the abstract whatnot, should not crowd out Humphrey himself. And there all is well. Thanks to the victorious counterpart, Humphrey himself has the requisite modal property: we can truly say that *he* might have won. (196, original italics, my bolding)

Self-Identity

Plantinga poses this objection to CT. He argues that CT has the unacceptable result that selfidentity and identity-to-self come apart in their modal profiles. Consider the proposition:

(F) Necessarily, Rosalind Franklin is identical to Rosalind Franklin.

This is ambiguous between two different propositions, corresponding to the *de re* and *de dicto* readings of the proposition:

(F-*de re*) Rosalind Franklin is necessarily identical to Rosalind Franklin. (F-*de dicto*) Necessarily, Rosalind Franklin is identical to Rosalind Franklin.

Plantinga argues that these two claims should both be true. The property of being selfidentical, and the property of being identical-with-Franklin "coincide on" Franklin. Or, as he defines it, "it is impossible that [s]he have either one without the other". But on CT, they do not coincide on Franklin.

On CT, these two propositions are equivalent to:

(F-*de re*-CT) RF is identical to her counterparts at all possible worlds. (F-*de dicto*-CT) At all possible worlds, if there is a counterpart of RF then that counterpart is identical to itself.

The first one of these is false since, on CT *no one* is identical with *any* of their counterparts. The second one, on the other hand, is true. In other words, on CT, Franklin necessarily has the property of being self-identical. But, Franklin does not necessarily have the property of being identical to herself (to Franklin)—this she only has contingently at the actual world.

Essential Properties

According to Plantinga, it follows from CT that, for instance:

(S) If I had had one more hair on my head, I would not have existed.

Or, more general, for every property P that any object O actually has, necessarily O has the property of having P or of being a different object to O. And so, O has all of its properties *essentially*.

Plantinga argues that this follows from the fact that all objects are non-identical with their counterparts. So, consider each possible world. At the world where O exists, O has the property P (*ex hypothesi*); at all other worlds where some counterpart of O (O*) exists, O* is non-identical with O (by CT). On this objection, it is because of this that (S) is meant to be true. Afterall, I do not exist at the world where my counterpart has one more hair on my head. I am worldbound.

Here, the response from the CT theorist is to say that (S) is ambiguous between the following:

(S*) If I had had one more hair on my head, I would have popped out of existence. (S**) The individual at the actual world with n-1 hairs and the individual at w_1 with n hairs are not identical.

The (S*) is false, but is arguably what drives the intuition about (S). (S**) on the other hand is true, but is just a statement of one part of CT (i.e. that we are not identical to our counterparts), and so is itself no objection.