

'If I Were You'

In this paper I will present an analysis of a puzzling counterfactual, 'If I were you', in terms of an epistemic two-dimensional semantics based on "centered worlds" (Chalmers, 2004 & forthcoming). The hope is 'If I were you' could be a convenient short-hand for picking out centered worlds just like 'If JFK was not killed' picks out the possible worlds where JFK had not been killed. As we shall see, 'If I were you' resists an interpretation in terms of centered worlds, and the problems it poses for centered worlds revolve around the reference of indexicals - problems that the project of epistemic two-dimensionalism aims to solve. I argue that 'If I were you' designates but fails to specify a unique centered world. 'If I were you' fails to specify, it might be that we cannot specify centered worlds in such a way for everyday statements like we can for statements such as "If JFK were still alive."

Introduction

Think of the possible world's where all of your beliefs are true. Unless we are immodest about our epistemic powers, we do not believe that the actual world we are in is such a world. There is a way the world seems *as* actual, by our beliefs, while we know that what we think of *as* actual is unlikely in every case to be correct about what *is* actually actual. Take our friend Throckmorton. We also know that the actual world we are in is unlikely to be one of the world's where all of Throckmorton's beliefs hold true. There is a way the world seems *as* actual, by Throckmorton's beliefs. Throckmorton is also unlikely in every case to be correct about what *is* actually actual. Apart from all the stuff me and Throckmorton agree on, the reference of the word 'chair', the fact that water is H₂O,

there remains stuff which from our different perspectives, seems different. It seems there are possibilities about how the world could be which we disagree on. Of the possible worlds that could be our world we disagree on which of these it is. But sometimes when we have been arguing, it turns out we never disagreed on how the world is, we just misunderstood what each of us stood for, what each of us was meaning by our words. In such cases we did not disagree in an objective sense on how the world actually *is*, we disagreed in the way the world could be seen *as*, we disagreed in some cognitive sense.

I shall start by introducing epistemic two dimensionalism. I then follow by introducing the problem of how to interpret "If I were you". I then attempt an analysis based on conventional possible worlds semantics, and then on ways in which we assume the reference of indexicals to not be automatically assigned. I then show how "If I were you" is an instance of how two-dimensional semantics at work. I shall finish by highlighting some problems that 'If I were you' poses for centered worlds.

Epistemic Two Dimensionalism

One motivation for two-dimensional semantics is the need for a theory to explain why terms that differ in cognitive significance intuitively seem to have different truth values. Kripke's argument that "Hesperus is Phosphorus" is *aposteriori* necessary in Naming and Necessity (Kripke, 1981) is one such place where the natural intuition that 'Hesperus may not have been Phosphorus' is violated. "Hesperus is Phosphorus" being necessary is 'true at all possible worlds' but only discoverable empirically. One goal of epistemic two dimensional semantics is to ground the sense in which it seems epistemically possible that the term 'Hesperus' may not have picked out the same thing as 'phosphorus'.

Another motivation for two-dimensional semantics is to explain the relationship between indexicals and propositions. One such problem for the relationship between

indexicals and propositions is the problem of how indexicals refer. Take such a statement as "I am making a mess". If John Perry was making a mess and uttered the sentence "John Perry is making a mess", it could be believed by John Perry exclusive to the belief, "I am making a mess" (Perry, *The Problem of the Essential Indexical*, 1993). If indexicals, such as 'I' refer to the person talking, then 'I' is translatable into 'the name of the person talking' for the purposes of finding a truth-evaluable proposition. For John Perry to believe that he is making a mess, in addition to "John Perry is making a mess" he would also need to believe "...and I am John Perry". It seems then that while "I am making a mess" and "John Perry is making a mess" may have the same extension, and may pick out the same individual across worlds, that is, have the same intension, there is something missed out by translating indexicals such as 'I', 'here' and 'now' into intensions that pick out persons, locations and times.

The epistemic two-dimensional semantics of David Chalmers (Chalmers, *Two-Dimensional Semantics*, 2006) (Chalmers, *Epistemic Two Dimensional Semantics*, 2004) generalizes what is similar between "I am John Perry" and "Hesperus is Phosphorus". While 'Hesperus' and 'Phosphorus' pick out the same thing across possible worlds, that is, Venus, and in the case of John Perry's utterance, "I" and "John Perry" pick out the same things across possible worlds, there is nevertheless something that is required to be believed by subjects that is *aposteriori* for John Perry that "I am John Perry" and for an ancient astronomy student "Hesperus is Phosphorus". This 'extra thing to be believed', claim advocates of epistemic two-dimensional semantics (Chalmers, *The Nature of Epistemic Space*, forthcoming) is characterized by a primary intension of expression tokens that picks out the extension of those expression tokens in *ways the world could be considered as actual*, differentiated from their familiar intension that pick out *ways the world could be counterfactually*.

It could be considered a view held by someone of the world considered as actual that 'Hesperus' is a completely different star to 'Phosphorus'. Of course this view is horribly mistaken, so mistaken, that given the way the world is, such a person that continued to refer to 'Hesperus' could not fail to refer to 'Phosphorus' (Kripke, 1981). The point is, that once we consider a the way the world is as actual, the way the world is actually sets in stone what we are talking about when we are talking about ways the world could be. But, so claims Chalmers, just as we can choose which worlds to talk about, we can choose which worlds we can consider-as-actual in picking out what it is we are talking about in those worlds. The claim is there is an intension of 'Hesperus' which, amongst others, picks out possible cases in which the world could be considered actual & where 'Hesperus' is not 'Phosphorus'.

As way of an example, we could find out that the colorless-tasteless liquid in our oceans that we know of as "water" could turn out to be XYZ rather than H₂O. It turns out all along we are on 'twin earth' (Putnam, 1975). If this happened, we would have to consider the XYZ-World as actual, and cease to consider the H₂O-world as actual. 'Water' would then pick out XYZ across worlds. 'Water is XYZ' could not fail to be false. 'Water is XYZ' is necessary *aposteriori*.

Consider rows below represent ways *the world could be considered as actual* and the columns pick out what these terms would refer to in worlds *considered as counterfactual*.¹ One intension seems to rely heavily on ways the world could be considering metaphysical possibilities, the horizontal, 'metaphysical' intension, and the other intension relies heavily on ways the world could be considered *as actual* epistemic possibilities.

¹ Chart from (Chalmers, Two-Dimensional Semantics, 2006)

'Water' Worlds

	H ₂ O-world	XYZ-world	...
H ₂ O-world	H ₂ O	H ₂ O	...
XYZ-world	XYZ	XYZ	...
...

Of course the token of expression 'water', cannot fail to pick out H₂O. However, it is does so because we take the epistemic possibility that "water is H₂O" as actual. This corresponds to the first row of the table. It is also an epistemic possibility that we are centered in the twin-earthian XYZ world, and 'water' could pick out XYZ. We then associate water with an alternate *epistemic* intension and a *metaphysical* intension, and the token "water" would then pick out XYZ across worlds.

It is possible however to have the extension picked out by the *epistemic* intension and the metaphysical intension misaligned. So for instance, take a scenario where we are located in the H₂O world, it is actual, but we consider that we are in the XYZ world. The primary intension of our expression picks out XYZ, but the metaphysical intension picks out H₂O. Our instances that "Water is XYZ" are thereby false, unknown to us, we are actually talking about H₂O. The function from the two dimensional intension to an extension then fails. So for a term to have an extension, both dimensions of intensions must align. 'Water' refers to H₂O, iff we take the H₂O-world considered *as actual* and also that the H₂O-world *is actual*.

Epistemic Possibility

It is an open question on how to characterize *epistemic* possibility. As metaphysical intension of tokens picks out that token in possible worlds, the hope is for a similar account for epistemic intension, that we can pick things out in epistemic space. One proposal is to account for it in terms of 'centered worlds' (Chalmers, *The Nature of Epistemic Space*, forthcoming). A centered world, C , is specified by a possible world, w , with a time and location in which it is centered. We can write this $C(w,t,l)$. For instance, "I am making a mess" is true in C , if that centered world is John Perry in the supermarket. $C(@,t,supermarket-aisle-1)$. But it is false in the same possible world considered as actual, but centered at a different location, say the petrol station. We might think of centered worlds as being specified more easily by people. "I am Napoleon" is true at the world centered at Napoleon, but not true of the world centered at David Chalmers.

It seems from the above that we have a lucid way of designating such a world. Perhaps 'If I were you' could be a convenient short-hand for picking out centered worlds just like 'If JFK was not killed' picks out the possible worlds where JFK had not been killed. Perhaps "If I were John Perry" or "If I were Napoleon" easily picks out centered worlds. That is the hope.

In order to evaluate such statements as candidates for picking out epistemic possibility, we need to carefully contrast them with the fact that they can also pick out metaphysical possibility. For instance, "If I were John Perry, I would be making a mess" could be true in a metaphysical sense, and specify more than just a centered world. This seems part and parcel of the two-dimensional semantic picture, expression tokens pick out two intensions, related to two types of possibility. On one reading of the epistemic possibility, it seems that "If I were John Perry" innocently picks out how the world might seem to me "If I were John Perry" and perhaps it follows from that that "I am making mess". However, there is a reading of "If I were John Perry" in a metaphysical sense too.

There is something inherently odd about this statement as regards to metaphysical possibility, how on earth could "I" be "John Perry".

Being John Perry

Think of sentences of the form "If I were you..."? Are we asserting things about being in another's situation? Their situation plus their personality? Or a gerrymandered thingy, a 'me-you'? "If I were you..." is a common everyday statement that plays a role in social exchanges, no person has trouble interpreting it when it is mentioned in passing. Such sentences make good sense in conversation and can convey quite complex facts & advice. Are we really talking about two different things, like saying "If X's were Ys" or are we talking about something that is an hybrid object, a "me-you"?

We shall attempt an analysis based on the 'metaphysical intension'. This seems proper since a sincere asserter of "If I were you" generally intends to say something about what they would do/believe/thing/feel/see if they counterfactually were the person they utter it about. Imagine witnessing a conversation between two people in which one asserts something to the other that begins along the lines of "If I were you I would wear the hat ". Is the suggestion that 'they' could be 'that person'?² Or is the 'could' that makes "If I were you" a counterfactual the fact that 'they' could be in 'that persons' situation? Is there a difference?

Are they saying that if 'they' were in *that* situation they would wear the hat? The latter, *that* situation, is an assertion about the external context that go into making the situation, the former 'that persons' situation, is more ambiguous and could include aspects of that persons history/beliefs/tastes/emotions. Are we talking about one aspect of your situation, or your entire situation? If "If I were you" can be true, just what parts of a situation should we take into account? In order to explain "If I were you" I

² Here I will follow a convention of calling the person picked out by the "I" in "If i were you", "they" and I will call the person picked out by "you" in "If I were you", "that person".

think it is necessary for a two-dimensional account. In order to motivate such an approach, it would be useful to see how attempting to use one-dimensional semantics of intension such as David Lewis's (Lewis, 1983) 'counterpart theory' of possible-worlds is inadequately accounts for "If I were you".

Let's consider the following one-dimensional semantic interpretation of "If I were you";

Explanation 1. *There is a counterpart of me and you, a 'me-you', in a world*

W₁, that is equivalent to this world except "I am you" and "you are me".

So for instance, "If I were John Perry..." takes that world (this one) where John Perry is perhaps currently walking down a supermarket aisle and spilling sugar all over the floor, and takes me in the library and picks out 'our' counterpart in another world, W₁, so that I am currently walking down a supermarket aisle and spilling sugar all over the floor. So it follows from "If I were John Perry..." that "I would be making a mess".

The Second Indexical Problem: The second "I" picks out both 'John Perry' and 'me'. "If I were John Perry I would be making a mess". Who is uttering this second 'I'? If it is me who is stating the counterfactual claim, I am not John Perry, so I could not be making a mess. If it is John Perry, then it is no sense true that "I would be making a mess". If it is the me-you, the counterpart in W₁, then we have an interesting occurrence of "I". A statement such as "I am here now", is well documented occurrence of a statement that must be true uttered at a context (Predelli, 1998). "I am here now" may be false given a world, for instance, the world that picks out is the world where I am here, at the library, on a Wednesday. Counterfactually, I could be in the library on a Sunday, and the extension of the proposition picked out by "I am here now", uttered where I am, here, now, in the library, on a Wednesday, would be false. But I could not fail to utter "I am here now" in a context or scenario and for it to be false at that context or scenario. The reason this is, is because the reference of "I", "here" and "now" is automatically assigned, as opposed to 'intentionally' assigned as in the case of 'that thing' (Perry, Indexicals and

Demonstratives, 1997). But the second "I" of "If I were John Perry, I would be making a mess" seems to evade its reference being automatically assigned. The statement "If I were John Perry, I could utter 'I am here now' and it would be true", is true while "If I were John Perry, I would be here now", seems false, due to the ambiguity of the reference in the second occurrence of "I". It is evident something is lacking in a conventional semantic account on this front.

Another Second Indexical Problem: "If I were John Perry I would be making a mess" is equivalent to "If I were the me-you, I would be making a mess", which is equivalent to "If I were the me(me-you), I would be making a mess" and so on. The problem for picking out a counterpart, is that it leads to this infinite regress where each layer seems to give no more light on what is at play in "If I were you". "If I were John Perry I would be making a mess" is equivalent to "If I were the me-you, I would be making a mess", but there seems, if there could be such a thing as the me-you, that there are counterfactual worlds where the me-you would not be making a mess. In other worlds the me-you does not pick out the worlds where "I am making a mess" is true.

The One Counterpart Problem: The counterpart of me and you in W_1 , has as its counterpart in this world both me and you. In some sense, I must be a counterpart of you! But each thing in a world has one and only one counterpart, and nothing and its counterpart can be in the same world.³

I am John Perry

Given the problems with 'being john perry' counterfactually, perhaps there is a way to analyze "If I were you" indexically, so the indexicals of "I am making a mess" uttered by John Perry share a reference with "If I were John Perry, I would be making a mess".

Perhaps there is a sort of situational semantics in which utterances of the world "I" can be shared across individuals. If so, there might be a way in which we could make sense

³ See (Lewis, 1983)

of claims such as "If I were you I", in terms of a shared situation. Such an explanation might work as follows;

Explanation2. There is a situation that you are in, and if I were in your situation, I would x.

So for example, it could be as straight forward as "If I were John Perry, I would be making a mess" can be analyzed in terms of "If I were in John Perry's situation, I would be making a mess". While this is certainly true of some cases of "If I were you", it doesn't seem true of all cases⁴. Consider cases in which aspects of a person's personality are important, perhaps John Perry is dispositionally inclined to make a mess in supermarket aisles. "If I were John Perry, I would not be making a mess". But perhaps this is a cheap shot, perhaps it is an essential part of the situation of John Perry's, that a mess is being made. Perhaps "If I were John Perry (making a mess), I would be upset", or "... I wouldn't be philosophically reflective and inclined to write a paper about it!", but not "if I were John Perry (making a mess), I would not be making a mess".

The problem with such a clear-cut analysis is how to differentiate the utterance of "I" related to the situation of evaluation from the utterance of "I" related to the individual at that time. They seem perfectly intertwined. Perhaps there is a part of John Perry's

⁴ Here I am steering clear of cases that might involve normativity, they have their own unique issues that make it difficult to evaluate truth for. However, there is no doubt that normativity plays an important role in common assertions of "if I were you". It is possible to make assertions of "If I were you" which are not synonymous with "If I were in your situation". For example "If I were you I would learn the piano", might be true alongside "if I were in your situation I would learn the piano" but it is not necessary, I could for instance, believe that "if I were in your situation I would learn the trumbone", given my personal aesthetic etc, yet on consideration of your character traits, along side my judgement, I could still believe "If I were you I would learn the piano". But we are talking about judgements here. No doubt the difference is clear here, but perhaps when "If I were you" is different to "If I were in your situation", what we really mean by "If I were you" is something along the lines of "you should". Just as there are cases where it is possible to describe "If I were you" as equivalent to "If I were in your situation" there are cases where an assertion of "If I were you" is equivalent to "you should", after all if we are describing what you can or can't do, "should implies can". But it is not true that the fact that you can do something implies you should, so there are instances where "If I were you" is different to "If I were in your situation" and "You should", that correctly describe capabilities with no imperative power.

situation that involves being philosophically reflective in addition to making a mess. It also seems true that "If I were John Perry, I would be the philosopher who wrote 'The Problem of the Essential Indexical'" etc. This does not sit at ease with statements like "If I were John Perry, I would not be philosophically reflective about the mess I am making"!

So we need some charitable principle along the lines of "A situation talked about in 'If I were you' terms, is to be interpreted as close as possible to the situation of the 'you' specified". This would be analogous to the principle in possible world semantics that says worlds such as "the white t-shirt world" should be interpreted as close as possible to the actual world with the only difference being that everyone wears white t-shirts. So we should consider "being philosophically reflective" an essential part of John Perry's situation, unless specified otherwise.

We might consider such a situation as being specified by all of John Perry's I-statements; "I see a trail of sugar", "I am in a supermarket aisle", "I am in a philosophically reflective mood", and so on, and any specification of "If I were John Perry..." would pick out the difference in that list. But this raises the problem of the essential indexical (Perry, *The Problem of the Essential Indexical*, 1993), John Perry may be making a mess, indeed, John Perry may believe "John Perry is making a mess", but this does not mean that John Perry believes that "I am making a mess", despite the fact being that John Perry is making a mess. It would still be true that "If I were John Perry, I would be making a mess". So a situation cannot be specified by the bunch of 'I' statements John Perry or the 'you' in "If I were you", thinks it is epistemically warranted to believe.

Perhaps then, a situation is a combination of a world, with a list of statements a subject thinks it epistemically warranted to believe. Of the world in which John Perry is making a mess and John Perry doesn't believe "I am making a mess", it is still true that "If I were John Perry, I would be making a mess", then we have simply equated "I" with John Perry,

who factually we know is making a mess. Perhaps if a situation is a combination of a world with a list of statements a subject thinks it epistemically warranted to believe, "If I were John Perry, I would be making a mess, but I would not believe 'I am making a mess'" is true. There is something we have left out of the analysis, and that is the way in which "If I were you" statements are not just true at a situation, but also rely on the situation of the asserter, the first "I" in "If I were you". What we want to say is, despite John Perry's beliefs of whether he is or is not making a mess, a truer rendering of "If I were John Perry, I would be making a mess" is "If I were John Perry, I would believe I am making a mess".

The World Centered Around John Perry

So we have seen that attempting an analysis of 'If I were you' poses serious difficulties if we try to give a conventional explanation in terms of possible worlds, or if we perhaps allow a violation of the rule that says the reference of 'I' is automatic. It seems the analysis that 'If I were you' needs is in terms of two-dimensional semantics. Let's recall the diagram of 'Water' worlds earlier. Something similar occurs with indexicals. Recall that the rows represent ways the *world could be considered as actual* and the columns represent ways *the world could be counterfactually*.

"I" worlds

	JPerry-world	NotJP-world	...
JPerry-world	JPerry (1)	JPerry (2)	...
NotJP-world	Not-JP (3)	Not-JP (4)	...
...

- (1) Imagine that an individual considers the possible world where “I” designates John Perry. At the context of utterance, where “I” designates John Perry this is true. “John Perry is making a mess” = “I am making a mess”.
- (2) Imagine that an individual considers that “I” designates John Perry as actual, at a context of utterance where “I” does not designate John Perry. “John Perry is making a mess, therefore I am making a mess” is false.
- (3) Imagine that an individual considers ‘I’ does not designate ‘John Perry’ at a context of utterance when ‘I’ picks out ‘John Perry’. “John Perry is making a mess” without believing “I am making a mess”.
- (4) Imagine that an individual considers “I” does not designate John Perry at a case where “I” does not pick out John Perry. “John Perry is making a mess & I am not making a mess”.

Centered worlds are specifications of a world considered as actual within a context of utterance (a possible world). In the above diagram (1) and (3) pick out the centered-worlds that are John-Perry-centered. So, can we give an analysis of “If I were John Perry” in terms of picking out those centered worlds?

Let’s try the following formulation for “If I were you”:

There is one way that the world is considered as actual for-you, w_a , at a context/world W . If I considered w_a as actual, then given you centered at w_a , there are a set of ways, w_x , the world, w_a , could be considered as actual at-you.

The problem here is that we are attempting to specify “If I were you” in terms of “If I were the individual in (1)”, or “If I were the individual in (3)”. If this were the case, then the way centered worlds are specified would be ambiguous. A centered world picks out a way the world could be considered as actual (an epistemic intension), and context in a world. To evaluate a centered world we must consider it as actual.. But if we took to

evaluating this centered world, we would have to consider the specification of *the way the world considered as actual for-you* as the way the world is actually.

To stress the problem, let's consider "If I were John Perry". So

If I were John Perry I would be making a mess

By the diagram above, (1) & (3) pick out the worlds centered at John Perry. In order for me to evaluate a proposition at the centered worlds (1) and (3), I would have to be in a situation analogous to (2), in which I am using "I" to designate John Perry at a context of utterance (eg. "If I were John Perry I") where "I" does not designate John Perry. We are not then talking about John Perry at all. But to evaluate it as a centered world I would have to be in a situation where I consider that "I" actually designates John Perry, (1) or (3). But given in evaluating a centered world picked out by "If I were John Perry I would be making a mess" I could only ever be in situation (2) in evaluating claims about "If I were John Perry I..."

So the way to look at is not as "If I were John Perry" as "If I were in situation (1) or (3)", but rather "If I were John Perry" is synonymous with situation (1) and (3). Occasionally "If I were you" would assert something close to situation (2), where "If I were John Perry, I would not be philosophically inclined" is true more of the asserter than of John Perry.

Where to from here?

I think the problem that we have encountered concerning how to evaluate truth at a centered world is a product of a state where the asserter of "If I were you"s primary or epistemic intensions surrounding 'I' are misaligned with the 'you's primary intensions. When we consider something as an epistemic possibility for the purposes of evaluating a centered world, we do not consider things "X is Y" where Y is a class of objects or properties, as we do when we are considering an epistemic possibility for the world.

Instead we consider things "X as Y", *as* actual; 'as' rather than 'is'; analogy rather than identity or the sort of epistemic commitment that leads to metaphysical commitment.

Bibliography

Chalmers, D. (2004). Epistemic Two Dimensional Semantics. *Philosophical Studies*, 118, 153-226.

Chalmers, D. (forthcoming). *The Nature of Epistemic Space*. Retrieved June 17, 2010, from consc.net/papers/espace.html

Chalmers, D. (2006). Two-Dimensional Semantics. In E. & Lepore, *Oxford Handbook of the Philosophy of Language* (pp. 574-606). Oxford University Press.

Kripke, S. (1981). *Naming and Necessity*. Oxford: Blackwell Publishing.

Lewis, D. (1983). Counterpart Theory and Quantified Modal Logic. In D. Lewis, *Philosophical Papers Volume 1*. New York: Oxford University Press.

Perry, J. (1997). Indexicals and Demonstratives. In R. & Hale, *Companion to the Philosophy of Language* (pp. 586-612). Oxford: Blackwell Publishing.

Perry, J. (1993). *The Problem of the Essential Indexical*. Oxford: Oxford University Press.

Predelli, S. (1998). I Am Not Here Now. *Analysis*, 58, 107-112.

Putnam, H. (1975). The Meaning of Meaning. *Minnesota Studies in the Philosophy of Science*, 7, 131-193.