

n-dimensionalism: the insides of words

HS “The Problem of Representation”, HS 2018, Philipp Blum, November 26, 2018

The contingent a priori

Suppose you never had never taken a philosophy course, and I told you that the following sentence

- (1) Water is watery stuff.

is a priori, explaining that by “a priori” I mean “knowable without recourse to sense-experience” and by “watery stuff” some description true of the stuff around us that is potable, colour- and odorless, in the lakes and rivers and so on. Would you not scratch your head? It is, of course, evident, undoubted and non-controversial that water is, say, colourless, but is this really something you can come to know in the armchair?

If I then explained to you that you would not apply the word “water” to something that does not have these properties, or at least does not have most of them, would you not remind me of the fact that this does *nothing* to show that we know that the stuff we are calling “water” actually has these properties, that it does not *follow* from “we would not have called *x* “the *F*” unless it were such and such” that we a priori *know* that the *F* is such and such.

Suppose I then said that a famous philosopher persuasively argued that the concepts of the a priori and the concept of metaphysical necessity are importantly different and not coextensive; that there are things that are not a priori but necessary and things that are not necessary and a priori, that his example was “Water is H₂O” and that the only way to explain this were to say that this sentence has two different intensions, one which is contingent and a posteriori, namely that the watery stuff is H₂O and that the other one, that H₂O is H₂O, is necessary and a priori and that the ‘bridge principle’, that water is watery stuff, has to be a priori for the whole come out necessary a priori, would you not then say that I did not really explain what I wanted to but that I explained it away, for there is not, after disambiguation, some one thing which is both necessary and a posteriori.

This, I take it, would be the point where two-dimensional semantics would enter the stage. The aim of my talk is to find out whether its appearance on scene really makes a difference.

In the following, I want to explain some motives behind this first reaction and make it appear less harsh. But let me first add some provisos. I have nothing against A-intensions or two-dimensional semantics in general and I can get myself in a frame of mind where I can agree that “water” is (in some sense equivalent to) some rigidified definite description. What I do not understand, however, is in what way A-intensions or two-dimensional semantics might yield an interesting and useable notion of a priori knowledge. By “interesting” and “useable”, I mean “non-tautologous” which, in this context, may be taken to mean “contingent”. So the question becomes: in what sense can (1) be understood to provide us with a priori knowledge of contingent facts?

My point does not turn on whether “watery” abbreviates the right descriptive phrase, nor on whether or not A-intensions are linguistically expressible. If they are not expressible, there is no knowledge to be had. I do not deny neither that many words have known associations with properties: the yellow flag on the ship can only mean that there are cases of yellow fever on board if people know that it does so – but this

just means that words have no use if we do not know what they are to be used for, not that we can get non-tautologous a priori knowledge about squares just in virtue of our knowledge of whatever is expressed by “the expression “square” stands for the property of being square”.

Let us examine some ways in which the alleged apriority of “Water is watery stuff” might be analysed, corresponding to different ways in which the primary intension of “water” might be read off from our use of the term. I will consider two ways in which the terms “water” and “watery stuff” might be linked. Either the content of “watery stuff” is part of the content of “water” or it is not. If we assume that “watery stuff” expresses the descriptive material associated with the natural kind term, the interpretations correspond to the semantic and the metaseantic construals of this descriptive material.

If we assume that “watery stuff” is the reference-fixing description for “water”, these two ways of construing “water” correspond to two ways in which we may speak of ‘reference-fixing descriptions’. In one way, it denotes the description, whatever it is, that has or might have been *used* to fix the reference of a particular term within our language community. In another way, it means whatever description is reference-fixing from a God’s eye point of view, that is uniquely satisfied by the referent of the term. Reference-fixing from God’s point of view, however, is better called rigidification.

The illusion of contingency

“Water is H₂O”, Kripke told us, is necessary. At first, this seems counterintuitive. Are we not able to imagine possible worlds where it is false; is not Twin Earth, as described by Putnam (1975), a world in case? It is not, says Kripke, for by “the world where water is XYZ” we are *misdescribing* something which we should rather call a “way our world might have turned out”, but did not. *Given* that it did not, it could not have: “water is H₂O” is, *if* true, necessarily true. What we are de facto imagining when we imagine water having another chemical structure than it actually has is not a possible world, a way our world could have been, but an impossible world, a way our world cannot be and could not have turned out to be. Twin Earth, if it is to be possible, is not a world containing water, but a world in the oceans and lakes of which there is a clear, drinkable liquid that we *would have taken* to be water (and hence called “water”) if we lived there. As we live here not there, however, the liquid we referred to when we actually fixed the reference of “water” had the chemical structure H₂O and so “water”, being a ‘rigid designator’, *actually* refers in all possible worlds to H₂O, though it could have done otherwise. For Kripke (1972: 307, 103), the “might” in “water might not be H₂O” is *merely* epistemic and represents “a present state of ignorance or uncertainty”: given the evidence before the empirical investigation of the nature of water, it could have turned out that water is not H₂O – but in this case “water” would not mean what it actually means. The qualification “given the evidence” is hence crucial, for given *our* evidence, it is no longer true that water could have turned out not to be H₂O.¹

David Chalmers has turned Kripke’s account upside down. Instead of distinguishing between a correct and an incorrect way of conceiving of (what one believes to be) a *p*-world, Chalmers claims that any given sentence *p* expressing a thought has two intensions: the secondary or subjunctive intension of “water is H₂O”, [water is H₂O]₂, is the set of possible worlds where H₂O is H₂O, i.e. the set of all worlds.² Its

1. Kripke (1972: 141–142, 332) claims that “it might have turned out that *p*” entails “it could have been the case that *p*” It seems incorrect to me to say, as Chalmers (2002c: 17) does, that “Kripke allows that it might *turn out* that Hesperus is not Phosphorus”. Cf. also: “...it seems reasonable to say that if the XYZ world *turns out* it be actual, then it will *turn out* that water is XYZ.” (Chalmers 2002a: 612) Chalmers (1995: 3) only made this claim for one sense of the “turns out” locution. All Kripke allows for is that it *might have turned out* in a world in which we had not discovered that it could not.

2. Following Chalmers’ usage in *The Conscious Mind* (1996), I am speaking here and in the following of primary and secondary intensions of *linguistic items*. The notion of primary and secondary intensions of *concepts*, which Chalmers uses in the 1995 version of “The Components of Contents” and the one of epistemic and subjunctive intension of concepts in the 2002 version both seem derivative. Cf.: “A concept, in thought, is the analog of a term in language.” (Chalmers 1995: 3), “Concepts are mental tokens that are often expressed in language by terms.” (Chalmers 2002a: 609)

primary intension, $[\text{water is H}_2\text{O}]_1$, is the set of worlds where the predominant clear drinkable liquid is composed of two hydrogen and one oxygen atom. Our world belongs to this latter set, but Twin Earth does not. In order to evaluate the primary intension of an expression α at a possible world, we have to “consider it as actual”: we have to figure out what we would refer to by α if that world turned out to be the actual world. To determine the primary intension of α , then, we have to extract from it by purely a priori means the “way in which [its] referent depends on how the actual world turns out” (Chalmers 1995: 4). To do this, we have to make use of what might be called a “presentation” of α , something giving us the way its reference depends on where we are, “encapsulating what it takes for an entity in the actual world to qualify as the concept’s referent” (Chalmers 1995: 3) What might such a presentation be? Quite often, philosophers stay content with loose paraphrases.³ Let us abbreviate the definite description “the dominant clear, drinkable liquid...” by “watery stuff”. Instead of misdescribing Twin Earth as a world where water is not H_2O , according to Chalmers, we should rather call it a world where watery stuff has a different microphysical structure than it has on our world, thereby saying something which distinguishes our world from Twin Earth.

The illusion that “water is H_2O ” is contingent, is thus explained by Kripke and Chalmers in radically different ways: for Kripke, this is modal error, partially explained by the common, but mistaken, belief that necessary truths are a priori. This belief is based on the failure to distinguish metaphysical or logical and epistemic possibility and the failure to acknowledge the existence of *mere* epistemic possibility. Chalmers turns this the other way round. When we call “water is H_2O ” contingent, we are basically right, for its primary intension, expressed by “watery stuff is H_2O ” is contingent, and primary intensions are what governs the rational relations between thoughts (Chalmers 1995: 2). Only its secondary intension, expressed by “ H_2O is H_2O ” is necessary.

This difference explains why Kripke and Chalmers give very different arguments against physicalism. Chalmers’ two-dimensional argument against a posteriori physicalism (what he calls ‘type B materialism’) can be represented as follows:⁴

- (2) The physical facts do not a priori entail the phenomenal truths.
- (3) If S is a posteriori, S has a contingent primary intension.
- (4) If materialism is true, the physical facts necessitate all true propositions.

From (2) and (4) it follows that psychophysical conditionals $p \rightarrow q$, where p is the conjunction of all physical truths and q is a particular phenomenal truth, are necessary a posteriori, if true at all. (3) and the conceivability of zombies entail that the primary intension of such an identification is not necessary. So there is, it seems, a world where the primary intension of the physical antecedent is true while the primary intension of the phenomenal consequent is false. As primary and secondary intensions are evaluated over the same space of possible worlds, it follows that not all phenomenal facts are necessitated by physical facts and, by (4), materialism is false.

Chalmers 1999 acknowledges that there is a loophole in the argument, as it has to be assumed that primary intensions of physical truths are still physical truths. If we were to deny this, however, we would get something like panpsychism which is not very materialistic either. I want to argue that there is a second loophole in the argument: how can we justify the step from the claim (2) that a phenomenal truth q is not a priori entailed by the physical truths to the claim that these physical truths do not necessitate the primary intension $[q]_1$ of q ? It is based on a supposedly general principle, namely that q and $[q]_1$ are conceptually

3. Chalmers, e.g., says: “...it is not too far off the mark to say that the primary intension of my concept “water” picks out the dominant clear, drinkable liquid in the environment that stands in an appropriate causal relation to the center of a world, but this is far from perfect.” (1995: 6, see also 2002a: 611, 613), “I will use “watery stuff” as a term of art to encapsulate the primary intension, whatever it is.” (Chalmers 1996: 58)

4. I am broadly following Chalmers’ presentation of it in (Chalmers 1999).

equivalent, in the sense that any necessitation of $[q]_1$ gives us an a priori entailment of Q : if something necessitates $[q]_1$, it a priori entails q ; if p does not a priori entail q , the conditional fails in a world considered as actual, and so $p \rightarrow [q]_1$ fails in some possible world. This presupposes that we can conclude from

(5) There is a p -world considered as actual where q is false.

that

(6) There is a p -world where $[q]_1$ is false.

The step from (5) to (6) is the crucial step in the ‘conceivability implies possibility’ argument: if we can conceive (ideally, after enough rational reflection etc.) of a world falsifying q , there is a sense in which q is contingent.⁵ According to Chalmers, then, there are only illusions, but no hallucinations of possibility.

This crucial step is related to the illusion of contingency in the following way: To explain why a necessary truth may be falsely taken to be contingent is to show that it is a posteriori in the sense of having a contingent primary intension. In order for this to count as an explaining away of the illusion, we have to produce a conceptually equivalent statement which *is* contingent and to argue that it has been falsely taken for the original one. For theoretical identifications, stating that two rigid designators are coreferential,⁶ Chalmers suggests a general device for doing this: Whenever we have a necessary a posteriori truth “ $a = b$ ”, its secondary intension “ $[a]_2 = [b]_2$ ” is necessary, while its primary intension “ $[a]_1 = [b]_1$ ” is contingent. Whenever someone thinks “ $a = b$ ” is contingent, he really means that “ $[a]_1 = [b]_1$ ” is contingent – which differs in modal properties, but is cognitively equivalent to the former in that “ $a = [a]_1$ ” and “ $b = [b]_1$ ” are both *a priori*.

Chalmers two-dimensional argument against a priori physicalism (2) proceeds from the nonexistence of a functional specification “the ϕ ” of what it is to be in a phenomenal state ψ (the so-called ‘explanatory gap’) to the claim that the phenomenal state ψ cannot be a posteriori identical to any brain state fulfilling the ϕ role. In doing this, he presupposes that “the ϕ ” and ψ are, if they pick out the same thing, *a priori* coreferential: the statement “ ψ is the ϕ ” is supposed to be a priori in the same sense than “water is watery stuff” is. The claim that “ $a = [a]_1$ ” is a priori, therefore, plays a crucial role in both arguments. So let us ask in what sense “water is watery stuff” may be said to be a priori.

Recall that “watery stuff” has been introduced to account for the sense in which we seem to be able to imagine worlds where water is not H_2O . Kripke used it (or variants thereof) to tell us that there really is *no sense* in which water might not have been H_2O : before learning that water actually *is* H_2O , we might have thought that “watery stuff is not H_2O ” is possible – but actually it is not. Our judgment to the contrary expressed a state of ignorance about what watery stuff is. Any qualitatively identical epistemic situation we are able to imagine *now* is one in which *our* watery stuff, i.e. water, is H_2O .

Kripke’s argument for dualism rests on a disanalogy between “water” and “pain”, i.e. on the premiss that “pain” could not, in a qualitatively identical epistemic situation, turn out to refer to something which is not pain, i.e. which is unlike the phenomenal quality we experience when we are in pain. Kripke formulates this as a claim about the ways in which we, i.e. our linguistic community, pick out the reference of “pain”.⁷ This thesis is thus not entirely a priori, for we can imagine creatures that pick out (what is for them) pain in other ways. Kripke’s point is just that *we* cannot.

5. I will use “conceivable” as dual to “a priori” and as “positively, ideally and primarily conceivable”. I will ignore what Chalmers (2002b: 9) calls “secondary conceivability” and come back to the question of idealisation later.

6. Like Kripke (1972: 148, 337) in the case of “pain”, Chalmers is committed to the claim that “consciousness” is a rigid designator, presumably referring to a natural kind. If it were not, then “consciousness is brain-state s ” would be of same type than “watery stuff is H_2O ”, i.e. be contingent a posteriori if true.

7. “If any phenomenon is picked out in exactly the same way that we pick out pain, then that phenomenon *is* pain.” (Kripke 1972: 153, 340)

For Kripke, considering a possible world as actual is, in effect, imagining the meanings of our words to be different from what they actually are.⁸ If we know that p is a necessary a posteriori truth, “it might turn out that $\neg p$ ” may be true if p is not entailed a priori by our evidence for it, but “it might have turned out that $\neg p$ ” is false, given what we know, because we *did* conclude from our evidence that p (Kripke 1972: 143, n. 72, 354). The sense in which it is conceivable that water is not H_2O , then, is relative to times, persons and a certain body of evidence (Kripke 1972: 35, 261). “Stick s is one meter long” is a priori *for* someone who used s to fix the reference of “one meter” (Kripke 1972: 56, 274).

Chalmers, however, wants “watery stuff” to do more than this. He believes that it (or at least a relevantly similar refinement of it) captures the way in which “water is H_2O ” still is and ever will be a posteriori. In his view, the illusion of contingency consists in assimilating “water is H_2O ” to “watery stuff is H_2O ”. Kripke would have taken this to be a case of conceptual confusion, of ignorance about what our words actually mean.⁹ Chalmers, however, thinks that two sentences really are conceptually equivalent and differ only in that the first, but not the second, is metaphysically necessary. They are conceptually equivalent because they are connected by the following bridge principle, which is true (and known by us to be true) *a priori*:

(7) Water is watery stuff.

On the face of it, (1) seems a very strong claim. Do we really now a priori that water is a liquid, fills our lakes and oceans, is taste- and odourless and drinkable? This seems a lot of information pulled out just of our ordinary concept.

I want to challenge the apriority of (1), by exploring two ways in which it might be understood and showing that in neither of them it is a priori in an interesting sense, i.e. a sense in which “phenomenal state ψ is what has the functional role ϕ ” is not. Chalmers assimilates (1) to Kripke’s cases of the contingent a priori, i.e. cases like “ a is one meter long” where “ a ” denotes the standard meter in Paris, by pointing to which the reference of “one meter” in our language has been fixed. But whereas Kripke simply *stipulated* that he uses “a priori” in a way such that it applies to “ a is ϕ ” whenever ϕ is the description by which the reference of a has been fixed, Chalmers has a substantial notion of the a priori, one he wants to put to use in his argument against physicalism. I want to argue that no such substantial notion is to be had which makes (1) turn out a priori but contingent.

The dilemma I want to put up for Chalmers turns on the thesis that he is committed to the claim that primary intensions are – at least in principle – expressible.¹⁰ I have three arguments for this, the first *ad hominem*, an internal second one and the third programmatic. Here is the first: If primary intensions were not expressible, Chalmers could not use the existence of an explanatory gap, i.e. the non-existence of a priori truths like (1) linking phenomenal and functional concepts, to argue against physicalism.¹¹ The internal argument concerns our motivation to introduce primary intensions in the first place, i.e. to explain away the illusion of contingency. Diagonalization is something we do in order to make sense of beliefs we take to be necessarily false. It should therefore give us something we can *ascribe* to people to which we do

8. “...being put in a situation where we have exactly the same evidence, qualitatively speaking, it could have turned out that Hesperus was not Phosphorus; that is, in a counterfactual world in which ‘Hesperus’ and ‘Phosphorus’ were not used in the way that we use them, as names of this planet, but as names of some other objects, one could have had qualitatively identical evidence and concluded that ‘Hesperus’ and ‘Phosphorus’ named two different objects.” (Kripke 1972: 104, 308)

9. Polywater, allegedly discovered in the Soviet Union, is described by Kripke (1972: 129, 323) as water which is not watery.

10. Chalmers would presumably deny this: “The real narrow content is an abstract function from centered worlds into extensions, and is only imperfectly captured by a given description.” (Chalmers 2002a: 38) – it is enough for my purposes, however, if primary intensions are expressible *in principle*. So I do not agree with the claim that “the PI framework doesn’t need to take a stand on whether the PI can be captured by a description” (Chalmers, <http://www.u.arizona.edu/~chalmers/class/596b/week4.txt>). If primary intensions were not specifiable at all, then how could we have cognitive access to them?

11. If all the physicalist has to do is to come up with a function, his task is too easy. Recall that Chalmers does not just present a semantical framework, but puts it to use against the physicalist, claiming that the physicalist cannot *specify* in functional terms what it is to be in a phenomenal state.

not want to ascribe belief in the impossible proposition. The programmatic argument is that apriority is tied to knowledge and knowledge to the *contents* of sentences. If we say of someone that he knows a priori that p , we do not merely ascribe to him knowledge that there is a sense in which p is and even cannot fail to be true, but we also ascribe knowledge of the truth actually expressed by p . If a priori knowledge of p were knowledge that p expresses a truth whatever it means, any epistemological interest in it would be lost. This means that primary intensions, if they are to do the work Chalmers wants them to do, have to be graspable: what is grasped when someone grasps the primary intension of “water”, then, is what I will call “watery stuff”. “[water]₁” and “watery stuff” are not only supposed to be coreferential *in the actual world*. Instead, the idea is that “watery stuff” picks out *in every possible world* what is there the value of the primary intension of “water”.¹² For any world w , what water would be if w turned out to be actual is the watery stuff in w .

Take Oscar, a philosophically interested water chemist, who – having read *Naming and Necessity* – uses “water” and “H₂O” interchangeably, i.e. such that, for him, “water is H₂O” is a priori. He differs from us, if the two-dimensionalist picture is correct, in his counterfactual language use. How can this difference be brought out?

Let him consider Twin Earth, where the watery stuff (the stuff in the lakes etc.) is not H₂O but XYZ and ask him whether his Twin-Earthian counterpart Twoscar, using “water” in the same way than he does, would call the Twin-Earthian watery stuff “water”. His answer will be “No”. Given that we know that water is H₂O and that we believe that this is necessarily so, this will be our answer too. Our counterparts will only call XYZ “water” if they do not know (or pretend not to know) what water is. But then, Oscar, a thorough Kripkean, will object, they do not use the word in the same way than we do.

So let us try a different route. Oscar thinks, and we allegedly deny, that “water is watery stuff” is a posteriori. Being told that water is watery stuff, he gets information while we do not; he is astonished while we are not. As we both agree that it is contingent that water is watery stuff, we both take Twin Earth to be possible, a world we both describe as a place where the local watery stuff is not H₂O. We only disagree in that he, but not we, would describe it further as a world where water is not watery stuff. But, Oscar will reply, did we not all learn from Kripke that we would thereby be misdescribing Twin Earth?

I think that Oscar has a point: To explain why “water = H₂O” is *not* a priori, even given that we know it is true, we have to consider worlds as actual where we lack this knowledge, i.e. where we take something to be water that is not H₂O. To explain why “water = watery stuff” is a priori, on the other hand, we have to restrict the realm of actual world candidates to those worlds where we *use the language as we do now*, i.e. where our words refer to sufficiently similar things than they do in our world. In order to do this, we have to exclude worlds where we are ignorant about what our words refer to: given that “water” uniquely refers to H₂O and that it is the macrophysical properties of H₂O that determine what counts as “sufficiently similar”, we have to exclude worlds where we do not know that water is H₂O. For if we did not know that water is H₂O, we could consider a world as actual where water has quite different superficial properties than H₂O actually has, thereby falsifying the claim that it is a priori that water has the superficial properties it has. We assume that macrophysical properties are more closely tied to meanings of our words than microphysical properties. Doing this, we *assume* rather than explain that we can be wrong about microphysics in a way we cannot be wrong about macrophysics.

How then is I will try to show that on both options, we can show that “water is watery stuff” is a priori only if we already *presuppose* that it is true. In the first case, this trivialises the notion of “a priori” in a way that “water is H₂O” would also be ‘a priori’, so understood. In the second case, we can show that it is a priori in the sense of having a necessary primary intension, but only at the price of making substantive empirical presuppositions about what our actual world is like. It seems to me that, given these presuppositions, “water

¹². We are thus asked to understand “environment” in the quotation of footnote 3 as a (universally bound) variable ranging over possible worlds.

is watery stuff” would be a posteriori, at least in the sense in which “water is H₂O” is. The moral seems to be that on none of the two construals we get any substantive disagreement with Oscar. The distinction between us, then, is a distinction without a difference.

Here is another way to put the dilemma: In order to evaluate the truth of (1) in other possible worlds, we either consider these worlds as actual or as counterfactual. When we consider them as counterfactual, we are treating “watery stuff” as something like a demonstrative, a rigid designator referring to whatever has been pointed to when the reference of “water” was fixed. This is the *dthat*-construal of (1). Rigid designators refer in all possible worlds to the same object. So they are not epistemically transparent, for they distinguish between epistemic alternatives. If we consider these worlds as actual, however, we cannot, I will argue, pick out by “watery stuff” *its* primary intension and thus cannot capture the way in which we consider possible worlds ‘as actual’. In the first case, we have strong necessities, i.e. a posteriori necessities with metaphysically necessary primary intensions, at the heart of Chalmers’s theory. In the second, we get *n*-dimensionalism and lack any justification to stop at any particular *n*.

Let me make three preliminary remarks. My point is not just that primary intensions are ‘ineffable’ in natural languages in so far as that they are only capturable in terms of secondary intensions, a problem that might be met by introducing square brackets or similar devices stipulated to *uniquely* express the primary intension of their argument. The deeper problem is that *any* way of specifying what the primary intension of a term is gives rise to the dilemma I want to put up for Chalmers in the following. The second caveat is that I have nothing against primary intensions in general: they work fine for indexicals, for example. What I am questioning is rather their supposedly intimate link to conceivability and the claim that they are scrutable by purely a priori means.¹³ Third, it is clear that, whatever else it is taken to mean, “watery stuff” cannot stand for a *purely* phenomenal concept, for then a world that only differs from our in that H₂O is not watery would already count as a zombie world, i.e. a world physically, but not phenomenally exactly like the actual world. “H₂O is watery”, however, is considered a posteriori on all sides.

I will now discuss in detail the two horns of my dilemma for Chalmers: primary intensions as descriptions by which reference has or could have been fixed and primary intensions as descriptions that are or can be rigidified.

Apriority by reference-fixing?

Under the first, meta-semantic interpretation of the descriptive material associated with the natural-kind term, we can again distinguish three variants. “Watery stuff” serves to introduce “water” as a directly referential singular term, according to the first variant of the theory, as an indexical expression according to the second and as a descriptive name according to the third.

If we understand “watery stuff” as a reference-fixing description in the first sense, we take (1) to have the following form:

(8) **dthat**(watery stuff) is watery stuff.

where “(...)” is *not* a rigidifying operator, but a directly referential singular term. This is the role **Kaplan** (1989: 579) originally intended for “**dthat**”. The content of the accompanying description “the ϕ ” in “**dthat**(the ϕ)” is not part of the content of the singular term: the description merely serves as a “demonstrative surrogate” **Kaplan** (1989: 581), completing the character and not the content of the term. It is, so to say, “off the record”, and not part of the semantics, but of the metasemantics of the referring term “**dthat**”.

13. Cf. **Chalmers** (1996: 57): “The primary intension of a concept, unlike the secondary intension, is independent of empirical factors: the intension *specifies* how reference depends on the way the external world turns out, so it does not itself depend on the way the external world turns out.”

I think this is the way Kripke intended his theory of the contingent a priori. I want to argue that it does not give us a substantial enough notion of the a priori to do the work Chalmers wants it to do.

The problem with this first reading is that it does not ensure that (8) is actually true. It is well known that we may fix the reference using descriptions which do not apply to the thing reference to which is fixed (Donnellan 1966). We pick out what we want to name by whatever uniquely identifying belief about it is shared among our conversational partners. It is shared belief, not shared knowledge, that is doing the reference fixing. Kripke is very clear about this in the case of proper names:

Thesis 5 says that the statement “If X exists, then X has most of the ϕ 's”, is *a priori* true for A . Notice that even in a case where [“If most, or a weighted most, of the ϕ 's are satisfied by one unique object y , then y is the referent of ‘ X ’.”] and [“If the vote yields no unique object, ‘ X ’ does not refer.”] happen to be true, a typical speaker hardly knows *a priori* that they are, as required by the theory. I think that my belief about Gödel is in fact correct and that the ‘Schmidt’ story is just a fantasy. But the belief hardly constitutes *a priori* knowledge. (Kripke 1980: 87)

A second variant of the meta-semantic route is to construe “water” as equivalent to “the watery stuff around here”. We then get

(9) The local watery stuff is watery stuff.

“The local watery stuff”, in (9), seems to function like a complex demonstrative. The analysis of complex demonstratives is rather complex, but it seems it boils down to two options. Either we allow for the case where “this F ” can refer to something which is not F , which brings us back to the first case, or we construe it as “this, which is F ”, where we predicate F of the thing we refer to by “this”. This latter option incorporates F into the semantics of the singular term – competent use of it requires some prior grounds for attributing F to the thing in question; (9) becomes something like “Whoever is singing now is singing” – if this is to provide us with knowledge that there is someone who sings, it requires prior knowledge that someone is singing now.

Something must have gone wrong. The diagnosis, of course, is that “watery stuff” is not just any ordinary singular description uniquely true of water; it is supposed to have played a special role in the fixation of the reference of “water” that distinguishes it from “ H_2O ” which has been coined long after the inclusion of the word “water” into our language. The claim, then, is that (8) is a case involving some sort of a “descriptive name” in the sense of Evans (1979). Evans invited us to stipulate that “Julius” denotes the inventor of the zip. He then contends that

(10) Julius invented the zip.

is a priori and what he calls ‘superficially contingent’. But there are two problems here. The first, discussed by Evans, is that there might have been no or no unique inventor of the zip. So what is a priori, if anything, is not (10) but “If there was a unique inventor of the zip, then Julius invented the zip”. But there is a second problem too. Let us agree that (8) has the same logical form than (10), namely:

(11) ϕ (whoever is the ϕ)

and let us assume that “the ϕ ” picks out a unique person in our world. The problem is that we have no assurance whatsoever that this person is ϕ , i.e. that (11) is true.

We do not have to know anything about the inventor of the zip to stipulate that he is to one we intend to call “Julius”. If we understand “the inventor of the zip” in “Let “Julius” denote the inventor of the zip” in

this *referential* and not in the attributive sense, we cannot be sure that we named thereby the actual inventor of the zip, given that we do not know who the inventor of the zip is.

When we wonder whether (10) is a priori, we have to ask ourselves whether we can consider a world w as actual which is such that there is someone x called “Julius” in w and it is not true of x in w that he or she invented the zip. Whether or not it is a priori that Julius invented the zip, thus depends on whether or not it is a priori that “Julius” denotes the inventor of the zip, that is on whether or not we can conceive of a world as actual where “Julius” does not denote the inventor of the zip. In one sense, this is easy. It is perfectly possible to conceive of a world as actual where the person who invented the zip is not called “Julius” (the actual world, e.g., will do).

If we, however, *stipulate* that our reference-fixing description be true of the thing reference to which is fixed, we get an empty notion of the a priori. Here is why: The original argument for the apriority of (10) had the following form:

- (12) Let “Julius” denote whoever invented the zip. Then it is a priori that Julius
invented the zip.

The crucial question now is: what does this stipulation amount to? If it does not amount to excluding any worlds from our consideration where the person who invented the zip there is not called “Julius” there, it might amount to excluding any worlds where the person *we would have called* “Julius” if we lived there did not invent the zip there.¹⁴ Here we have a double modality. There are two ways to cash it out. Either we take the stipulation to fix a *two-dimensional* intension for “Julius”. This would mean that we thereby fix not only our actual language, but also the language we would speak in counterfactual circumstances. Alternatively, we suppose that the stipulation is made in any world which we consider as actual. What it means to “keep the language fixed” and to “retain the concept of the *real* actual world” Chalmers (1995: 4), then, is to stipulate that they use the same words with the same meanings to fix the reference of “Julius” than we do:

...for this representation [of (10) as a contingent a priori truth] to be right, we must add that the stipulation in question was made in each of the worlds i , j and k . One who did not know about the stipulation, or did not understand it, would not know that the statement was true.
(Stalnaker 1999: 15)

Keeping the language fixed amounts to excluding from consideration worlds w where “Julius” denotes somebody other than whoever invented the zip in w . We do not gain new knowledge of contingent facts, like the fact that water is a liquid, but we are rather making true something.

There is an important difference between the metre and Julius examples and other cases of reference-fixing by description. If we use stick S to fix the metric system, then indeed we know “automatically, without further investigation, that S is one meter long” (Kripke 1980: 56). But this is because the introduction of the metric system was a performative act, which created metres in the first place. I think this is why Kripke is very cautious about the generalisability of the metre example:¹⁵

If someone fixes a meter as ‘the length of stick S at t_0 ’, then in some sense he knows *a priori* that the length of stick S at t_0 is one meter, even though he uses this statement to express a

14. This corresponds to the ‘shallow’ notion of the a priori where a statement is a priori if it expresses a truth (as used) in every context (cf. Stalnaker 1978: 83).

15. He says that the ‘definition’ we give while baptising the substance by “Gold is the substance instantiated by the items over there, or at any rate, by almost all of them” “express[es] an *a priori* truth, in the same sense as (and with the same qualifications applied as) ‘1 meter = length of S ’: it *fixes a reference*” (Kripke 1980: 135).

contingent truth. But, merely by fixing a system of measurement, has he thereby *learned* some (contingent) *information* about the world, some new *fact* that he did not know before? It seems plausible that in some sense he did not, even though it is undeniably a contingent fact that *S* is one meter long. (Kripke 1980: 63)

I take this to show that the meta-semantic explication of our alleged a priori knowledge is hopeless. And I think that this was also Kripke's view:

Following Donnellan's remarks on definite descriptions, we should add that in some cases, an object may be identified, and the reference of a name fixed, using a description which may turn out to be false of its object. [...] In such cases, the description which fixes the reference clearly is in no sense known *a priori* to hold of the object, though a more cautious substitute may be. If such a more cautious substitute is available, it is really the substitute which fixes the reference in the sense intended in the text. (Kripke 1980: 80, n. 34)

The "if" here is important: in all but the metre case, it seems, a more cautious substitute is not available.

It is not clear, however, how we can be as cautious as Kripke asks us to be. If being cautious is knowing that the description used to fix the reference is true of the object, then we indeed know that the cautiously chosen description is true of it – but this knowledge is not substantial.

But there is clearly something special about predicating reference-fixing descriptions? Yes there is, but I do not think that caution has anything to do with it. Rather it is coordination. Coordination is coreference by stipulation. But it does not give us substantial knowledge; it rather is a precondition of acquiring knowledge in the first place. The facts of coordination are fixed from a God eye's point of view, or – in Fine's relationalist semantics – from the point of view of an omniscient semanticist, establishing the theory describing exhaustively and accurately.

Let us follow the first horn first and consider in what sense (1) may be called a priori if "watery stuff" is taken to be the description possibly *used* to fix the reference of "water".

Like "water is H_2O ",¹⁶ "water is watery stuff" allows readings under which it is false. Block & Stalnaker (1999: 10) have described cases in which "watery stuff" could fail to pick out a unique referent, due to nonphysical ghost water. We do not even have to consider far fetched possibilities, however. There is a chemical substance, heavy water or deuterium oxide, which is watery stuff, but not water, for it has the microstructure D_2O . It is odour- and tasteless and you could not tell it from water.¹⁷ So "water is watery stuff" is true only if "watery stuff" is taken to include a descriptive component excluding D_2O . As D_2O has been synthesized for the first time in 1933 and water has been discovered to be H_2O already in 1784, the description by which the reference of "water" has been fixed presumably did cover heavy water as well.

To this it might be replied that the actual reference-fixing description does not matter. What is important, however, is the disanalogy between a description mentioning only superficial properties that might have

16. As Avrum Stroll (1989) has reminded us, if ice is in the same way 'identical to' (i.e. composed of) H_2O than water, then water is ice. But even if ice and steam count, under a relaxed sense, as water, water is *not only* H_2O , but contains various minerals, mud etc. So only *pure* water is H_2O . Pure water, however, is not what we most often mean by "water". If I order a glass of water in a restaurant, I would not like to get pure water, because pure water is not healthy to drink. So "water is H_2O " is true only if it is taken to mean " H_2O is the dominant chemical substance water is composed of". "Dominant", however is vague. Consider the sentence "air is O_2 ". There are many uses of "air" in which "air" is taken to refer to oxygen. Take, e.g., "The people in the submarine died because there was not enough air left for them to breathe." Nevertheless, air consists only to 21% of oxygen and it seems false to say that science discovered that air *is* oxygen. Moreover, air and oxygen share some but not all of their macrophysical properties. Like air, oxygen is odourless, transparent and breathable, but smoking a cigarette in a room filled with pure oxygen is even more unhealthy than smoking it in a room filled with air. It seems that on some uses of "air", "air is O_2 " is necessary, but on others it is not – and it seems reasonable to say the same thing of "water is H_2O ".

17. Apparently, mice can tell them apart. This suits them well, for D_2O makes them infertile. We also know that small fish cannot live very well in heavy water.

been used to fix the reference of “water” and a description of its microphysical properties. So it is said that “water is H_2O ” is a posteriori because we are able to imagine *of* water, i.e. H_2O , that it has a different microphysical structure (and consequently is not H_2O). Imagining this, we are “ H_2O ” as a description. This can be seen from the fact that it will not, under this reading, be necessary that what *is* H_2O , i.e. water, is H_2O , i.e. has the microphysical structure it (actually) has. So we are not conceiving of one and the same thing that is both necessary and conceivably false.¹⁸

The obvious reply is that we have, in order to evaluate the claim that “water is H_2O ” is informative, to *bracket* our knowledge that it is true. But what does this mean? It clearly will not do to imagine a situation where I know everything about the microphysical structure of water except that it is *called* H_2O . So I have to imagine being ignorant either about water or about its microphysical structure. If water *is* H_2O , this amounts to the same thing. To explain the possible truth of “I do not know that water is H_2O ” we have to read it as “I do not know what water is”.

The problem with this line is that it seems equally to apply to “water is watery stuff”. Bracketing our actual knowledge of water, we seem perfectly able to imagine of H_2O , that is water, that it has other macrophysical properties than it actually has. It seems easy to imagine worlds where water has become sparse, where only small samples of it are kept in laboratories and the lakes are filled with, say, petrol, but where the word “water” still applies to what it applies here, namely to H_2O . Entertaining the hypothesis that water might become sparse is not envisaging a linguistic change.

So we are left with a choice: either we *presuppose* that we would have called “Julius” some other person if some other person would have invented the zip, thereby making (10) *trivially* a priori or we allow for the possibility that we may falsely believe that we have renamed W.L. Judson, who invented the zip in 1891, but in fact have named *b*. To know whether a world *w* is a world where “Julius invented the zip” is false, we would then have to know whether *b* invented the zip. The upshot of this, I think, is that it does not follow from “if *a* were not ϕ , we would not call it ‘*a*’ ” that “ $\phi(a)$ ” is a priori.¹⁹

The point carries over to “water is watery stuff”. Either you allow for the possibility that the reference of water has been fixed by a description which is not in fact true of it or you stipulate, on the other hand, that “water” denotes watery stuff whatever it is. In the first case, we might still, with Kripke (1972: 78, 290), *stipulate* to use “a priori” as a predicate that applies to all and only the sentences of the form *Fa*, where “the *F*” is the description whereby the reference *a* has been fixed for all, many or most of us. But there are good reasons not to accept such a stipulation. For we primarily use “a priori” to speak about certain kinds of knowledge and this use is undercut by the stipulation. If “the man who discovered the Peano axioms” fixes the reference of “Peano”, some of us both know a priori that Peano is the man who discovered the Peano axioms *and* know a posteriori that Dedekind discovered them. So they know that Peano is Dedekind, which is false. Similarly, if we *know* a priori that Gödel discovered the incompleteness of arithmetic, then we know that arithmetic is incomplete. If a mistake were found in the proof, this might turn out to be false.

In the second, you do not only keep your language fixed, but a lot of empirical knowledge as well, namely your *de re* knowledge of water that it is watery. So you are no longer considering worlds as “actual”,

¹⁸. I take this to be compatible with Kripke’s insistence, in *Naming and Necessity*, on the importance to distinguish between metaphysics and epistemology, between what is necessary and what is a priori. The mistake of those who try to refute the necessity of “water is H_2O ” by imagining a world where watery stuff has a different microphysical structure is that they do not imagine a world where what Kripke claims to be necessary is false. In this sense, they are misdescribing the world they imagine, if they take it to be a counterexample to the necessity thesis.

¹⁹. The reason for this, I suppose, is that we use proper names for all sorts of purposes. If I say to you “if I were Maradona, I would marry Britney Spears” and I believe that Maradona is the best football player in the world, it may well be that if I knew that Schmaradona, but not Maradona, is the best football player in the world, I would instead say “if I were Schmaradona, I would marry Britney Spears”. But from this it does not follow that it is a priori that Maradona is the best football player in the world. Note also that the converse of the above principle is false as well. Take, e.g., the supposedly contingent a priori truth “I am the speaker”. If I say to you in a conversation “if I were you, I would do *p*” I am not misdescribing a possibility that should be properly described by “if the speaker were the audience, the speaker would do *p*”. Thanks to Andri Töndury for pointing this out to me.

but you are considering them as counterfactual, as ways the actual world might be *given* that water is watery. You are then supposing that the primary intension of “water” is given by the *secondary* intension of “watery stuff”, i.e. that $[\text{watery stuff}]_2 = [\text{water}]_1$. This makes “watery stuff” behave very much like a rigid designator. In every possible world, an utterance of “water” would thus refer to what is there the predominant liquid in the lakes and the oceans. This would follow e.g. if “watery stuff” contained an indexical element, that it denotes the clear drinkable liquid in *my*, or *the speaker’s* environment.²⁰ Either the relevant features of the speaker’s environment can be specified in non-indexical language or they cannot. If they can, as Chalmers (2002a: 6n) seems to think, nothing crucial changes. If they cannot, “watery stuff” distinguishes between epistemic alternatives,²¹ so it is not what we could call “epistemically transparent”. It was precisely their epistemic transparency, however, that led Chalmers to introduce primary intensions in the first place. If “watery stuff” is not epistemically transparent, if it is possible that you are unable to tell whether its application conditions are fulfilled, “water is watery stuff” is a posteriori, while it has a metaphysically necessary primary intension. I think one of the reasons why the contingent a priori has puzzled many philosophers is that claims that something is contingent and a priori are *doxastically indefensible*. If I claim to know (and even to know a priori) that p and at the same time admit that p might be false I do not contradict myself, but I will hardly convince anyone that I am right. It seems especially weird if I just stipulated p to be true. In this case, it seems, I have the best possible evidence for my knowledge claim. But stipulation only gives rise to knowledge that p if it *suffices* to make p true. And this is not the case with the contingent a priori. Whether we fix the two-dimensional intension of “Julius” or whether we iterate the naming procedure in each world we consider as actual, what may properly be called “a priori” is not (10) but the following conditional:

(13) Let “Julius” denote the inventor of the zip. Then Julius invented the zip.

(13), in contrast to (10) is not only a priori but also necessary. There is no world, conceived of as actual or not, where there is a unique inventor of the zip, “Julius” is used as a name for him and “Julius invented the zip” is false. So (10) comes out a priori under an empty notion of “a priori”: something is a priori in this sense in virtue of being of the form “if ‘ p ’ is true, then p ”.

“Water is H₂O” is necessary, *if* it is true. Can we thus say that “water is watery stuff” is a priori *if* it is true? Doing this we would trivialize our notion of “a priori”. Note that there is a crucial asymmetry between the two cases:

(14) If p is true, p is necessarily true.

(15) If p is true, p is a priori.

Kripke (1972: 109, 31) argues that we know the truth of (25) a priori, by philosophical analysis. But what about (26)? Kripke admits that even when we have fixed the reference of F with respect to a sample, it is still an empirical question whether all or most of the items in the sample are F .²² What we have to presuppose in the case of “water is watery stuff” is not only that it is true but that it is *known* to be true: apriority does not, like necessity, qualify truth but knowledge.²³ *Given* that we know that it is true, it may be called “a priori”, but vacuously so, for every p is entailed by a body of evidence which includes p itself.

20. This is suggested by Chalmers’ remark (Chalmers 1995: 5) that we have to evaluate the primary intension at *centered* worlds, worlds with an index fixing the speaker, the time and location of the utterance in question etc. We are thus supposed to include into the ‘center’ a sample of water, e.g. the one that members of our linguistic community pointed to when they introduced “water” into our language.

21. Two worlds are epistemic alternatives if they cannot be told apart by their inhabitants, if there is nothing with respect to which their inhabitants could distinguish them.

22. “*A priori*, all we can say is that it is an empirical matter whether the characteristics originally associated with the kind apply to its members universally, or even ever, and whether they are in fact jointly sufficient for membership in the kind.” (Kripke 1972: 137, 329)

23. If p is true, but not known to be true, how could it then be a priori (on any reasonable sense of “a priori”)?

Apriority by rigidification?

On the second construal of the relation between “water” and “watery stuff”, (i) comes down to:

(16) \ddagger (watery stuff) is watery stuff.

where \ddagger is a rigidifying operator that takes a description to form a singular term denoting in all possible worlds whatever uniquely satisfies the description in the actual world.²⁴ “ \ddagger (the ϕ)” is, in Kaplan’s terms, “a rigid description which induces a complex ‘representation’ of the referent into the content” (Kaplan 1989: 580) and thus corresponds to the semantic construal of “watery stuff”.

Unlike “ ϕ ((the ϕ))” (ii), “ ϕ (\ddagger (the ϕ))” (“whatever is ϕ is ϕ ”) is a truth of logic and can thus plausibly be taken to be a priori. The problem now is not with the a priori, but with contingency: how could “ ϕ (\ddagger (the ϕ))” fail to be true? Given that “water” is a rigid designator, why does it matter whether it is a rigidification of “watery stuff” and not of “ H_2O ”?

(16) claims that “watery stuff” is world-independent, i.e. that its secondary intension does not depend on the world in which it is uttered. Twoscar, if he uses language in the same way than Oscar, will call “watery stuff” in other worlds whatever has the superficial properties of XYZ in his world. That his XYZ qualifies as watery stuff *by our standards*, does not mean that any standard based on *its* superficial properties will be the same as ours. If we allow for the possibility that Twoscar uses “watery stuff” differently than we do, however, the matrices of “watery stuff” and its projection on the horizontal, “ \ddagger (watery stuff)” will have to look differently. Whether or not they pick out the same stuff, then, will depend on the world in which they are uttered.²⁵

In Chalmers’ framework, the difference between the two rigidifications is that the definite descriptions used play different roles in the fixation of the reference of “water”: you use “watery stuff” to rigidify its primary, and “ H_2O ” to rigidify its secondary intension. So we get the following variant of (16):

(17) \dagger (water) is watery stuff.

where \dagger is Stalnaker’s dagger, a function mapping a singular term a to another singular term $\dagger a$ denoting in every possible world the semantic value a would have if used there, i.e. a function projecting its diagonal intension onto the horizontal (Stalnaker 1978: 82).

For “water is H_2O ”, we get

(18) \ddagger (water) is H_2O .

where \ddagger is Stalnaker’s ‘upside-down dagger’, i.e. a function mapping a singular term a to another singular term $\ddagger a$ denoting in every possible world the semantic value of a attributed to your actual use of the term, i.e. a function projecting the diagonal proposition onto the vertical (Stalnaker 1978: 83, n.).

Given that “water” has the primary and secondary intensions it is claimed to have, (17) and (18) are true in the actual world. But Chalmers claims more than this: he claims that (17) and (18) are not only true but a

24. \ddagger corresponds to the ‘upside-down dagger’ of Lewis (1973: 63–64) and Stalnaker (1978: 83, n.) and to the “fixedly actual” operator $\mathcal{F}\mathcal{A}$ of Davies & Humberstone (1980).

25. I take myself to be in agreement with Stalnaker on this point: “It is straightforward to define an abstract property of propositional concepts – the necessity of the diagonal – but before drawing any conclusion about the epistemic status of an utterance associated with a propositional concept that has this property, one has to consider what set of possibilities the concept is defined on, and how the utterance is related to the propositional concept. One might conclude that *if* the relevant set of worlds included all those compatible with an agent’s knowledge, then she would know that the utterance expressed a truth. But this knowledge would not deserve the label “a priori knowledge” unless the agent knew a priori that the actual world was among those in the relevant set, and the two-dimensional apparatus does not pretend to offer any account of that.” (Stalnaker 1999: 14).

priori and a posteriori respectively. To be a priori, it is claimed, is to have the same primary intension. But what are the primary intensions of “watery stuff” and “H₂O”?

Let there be three worlds, the actual world where water, the watery stuff, is H₂O, Twin Earth where water, the watery stuff, is XYZ, and a third world where water, the watery stuff, is ABC. “Watery stuff” is stipulated to express the diagonal intension of “water” in the actual world. So it seems that “water” and “watery stuff” have the following two-dimensional intensions:

“water”	H ₂ O	H ₂ O	H ₂ O	“watery stuff”	H ₂ O	XYZ	ABC
	XYZ	XYZ	XYZ		H ₂ O	XYZ	ABC
	ABC	ABC	ABC		H ₂ O	XYZ	ABC

Given these matrices, we indeed have water = \ddagger (water) and watery stuff = \dagger (water). But wait. How can we be sure to have written down the right matrix for “watery stuff”? We stipulated that “watery stuff” should express the diagonal intension of “water” in the actual world. This was the projection of the diagonal onto the horizontal. This is what is done by Davies’ and Humberstone’s “actually”-operator *A*. But then we did another step, namely we applied Davies’ and Humberstone’s “fixedly”-operator \mathcal{F} , the vertical analogue of \square , and replicated the upper row on the lower two. Thereby we assumed that it does not matter for the evaluation of “watery stuff” in which world it is uttered.

But perhaps it does. Let us assume that “watery stuff”, if uttered in Twin Earth, would express the diagonal intension of *their* (possibly different) concept of “water” and so it would in the third world. I use subscripts to indicate the secondary intension of the concept Twin-Earthians use to express the primary intension of their water concept (that *we* may express by “twatery stuff”). So we get the following matrices:

“water”	H ₂ O	H ₂ O	H ₂ O	“watery stuff”	H ₂ O	XYZ	ABC
	XYZ ₁	XYZ ₁	XYZ ₁		H ₂ O ₁	XYZ ₁	ABC ₁
	ABC ₂	ABC ₂	ABC ₂		H ₂ O ₂	XYZ ₂	ABC ₂

At first sight, this does not seem to change very much. Both concepts still have the same diagonal intension and so are a priori equivalent. If we assume that XYZ = XYZ₁ and ABC = ABC₂ (as I did for the new matrix of “water” on the left side), it is still true that water = \ddagger (watery stuff). What we lost, however, is the idempotency of \dagger , i.e. it is no longer true that \dagger (watery stuff) = watery stuff. The new matrix for \dagger (watery stuff), i.e. $\dagger\dagger$ (water) is the following, which is different from that of “watery stuff” on the right hand side:

“ $\dagger\dagger$ (water)”	H ₂ O	XYZ ₁	ABC ₂	“watery stuff”	H ₂ O	XYZ	ABC
	H ₂ O	XYZ ₁	ABC ₂		H ₂ O ₁	XYZ ₁	ABC ₁
	H ₂ O	XYZ ₁	ABC ₂		H ₂ O ₂	XYZ ₂	ABC ₂

There are two problems: First, we only get “ \dagger (watery stuff) = watery stuff” at the price of assuming XYZ=XYZ₁ and ABC=ABC₂. Second, we can do exactly the same with “water is H₂O”, abbreviating by “twatery stuff” and “swatery stuff” whatever macrophysical descriptions uniquely pick out XYZ in Twin Earth and ABC in the third world respectively:

“water”	watery	twatery	swatery	“H ₂ O”	watery	watery ₁	watery ₂
	watery	twatery	swatery		twatery	twatery ₁	twatery ₂
	watery	twatery	swatery		swatery	swatery ₁	swatery ₂

So if we had introduced “H₂O” as an abbreviation of the chemical structure of watery stuff, whatever it is, and assume that twatery stuff=twatery stuff₁ and swatery stuff = swatery stuff₂, we would get that (17) and (18) are on the same level.

Does this matter? I think it does. To see why, consider Davies’s and Humberstone’s rendering of the claim that “water is watery stuff” is a priori:

(19) $\mathcal{F}A$ (water is watery stuff)

where $\mathcal{F}\mathcal{A}$ (“fixedly actual”) is defined as $\models_y^x \mathcal{F}A\alpha :\Leftrightarrow \forall y \models_y^y \alpha$ (where “ $\models_y^x \alpha$ ” means “ α , as evaluated in y is true of x ”). $\mathcal{F}\mathcal{A}$ corresponds to $\Box\ddagger$, where \ddagger is a function mapping the diagonal intension of a propositional concept onto the horizontal but leaving the other rows intact.²⁶

In this more uniform notion, (19) amounts to the following:

$$(21) \quad \Box\ddagger(\text{water} = \ddagger(\text{water}))$$

From this, it plausibly follows:

$$(22) \quad \Box(\ddagger(\text{water}) = \ddagger\ddagger(\text{water}))$$

Here, however, we have a double occurrence of \ddagger . Though both “watery stuff” and “ $\ddagger(\text{watery stuff})$ ” have the same diagonal, the apriority claim only goes through if they do not differ in the other fields. What does it mean to assume that they do not there differ? It means that “watery stuff” is not only supposed to pick out the primary intension of “water” in every possible world *considered as counterfactual* but also in every possible world *considered as actual*. This is even stronger than the claim that their secondary intensions are the same. It amounts to the thesis that we cannot even conceive of a world *as actual* where the counterfactual behavior of “watery stuff” and the primary intension of “water” differ, where e.g. $\text{ABC}_2 \neq \text{ABC}$. It has to be a priori that their secondary intensions are the same.

This, I think, is an unreasonably strong contention. Whatever our specification of the primary intension of “water”, it seems that we can always sensibly ask whether we managed to capture it. Whether or not we did, is an empirical question: it depends on what *we*, competent speakers of our language, would say in counterfactual circumstances. Whether or not the meaning of “watery stuff” depends on the world *in the same way* than the meaning of “water” is a fact we know in virtue of being members of a certain linguistic community, of having learned the language we speak in one way but not in another. And this may well depend on how the world is.

IS THIS TRUE? Do (19) and (20) express the claim that (17) and (18) are respectively a priori and a posteriori? I think not. For (19) and (20) are true iff (17) and (18) are true respectively, and, both projection operators being idempotent, they are even true iff “watery stuff is watery stuff” and “ H_2O is H_2O ” are true. This, however, means that “water is watery stuff” is a priori only if it is true.²⁷ If we are ignorant about which of various actual world candidates is the real actual world, we cannot specify one world as actual and so the two-dimensional analysis does not even get from the ground.

To see the problem, consider the following sentences:

$$(23) \quad \text{If XYZ is watery, then XYZ is water.}$$

$$(24) \quad \text{If H}_2\text{O is not watery, then water is not watery.}$$

Chalmers argues from the truth of (23) to the claim that “water is H_2O ” is a posteriori and from there to the claim that the primary intension of “Water is not H_2O ” is non-empty. To evaluate the truth of (23), he uses the Ramsey test: accept hypothetically the antecedent and check whether you can conclude to the consequent (Chalmers 2002b: 13). So if “water is watery stuff” is a priori, (24) had better be false. But is it? I am not so sure. Assume hypothetically that the antecedent of (24) is true. Given that you know that water *is* H_2O , cannot you therefrom conclude that water is not watery?

26. The a posteriori necessity of “water is H_2O ” is formalised by them as:

$$(20) \quad \mathcal{F}\Box(\text{Water is H}_2\text{O})$$

where $\mathcal{F}\Box$ is defined as $\models_y^x \mathcal{F}\Box\alpha :\Leftrightarrow \forall w, z \models_z^w \alpha$. $\mathcal{F}\Box$ (or, equivalently, $\Box\mathcal{F}$) corresponds to $\Box\ddagger$, where \ddagger now is a function mapping the diagonal proposition of a propositional concept onto the vertical and leaving the other rows intact.

27. In Davies’ and Humberstone’s two-dimensional logic, we cannot conclude α from $\mathcal{F}A\alpha$. All we can conclude from it is $A\alpha$, which is true iff α is true in the actual world, whatever this is.

Suppose watery stuff in w is XYZ, but that the inhabitants of w nevertheless use “water” to refer to H_2O .²⁸ They certainly can consider a world v as actual (e.g. their own actual world w) where H_2O is not watery, in the same way as we can imagine a world where H_2O is not watery. So v is a world, considered as actual by the inhabitants of w , where the substance they call “water” is not watery. It is also a world where what we (actually) call “water” is not watery, but not a world where what we *would* call “water”, if we lived there, is watery, for what we would call “water” in w is not H_2O , but XYZ. The inhabitants of w differ from us in their concept of water, not because they mean by it something other than H_2O (for they do not), but because they use it to refer to something which is not watery *by our standards*, i.e. which does not fill their lakes etc. This difference in counterfactual language use, however, between us and the inhabitants of w only shows up if we *presuppose* that it is H_2O and not XYZ that fills *our* lakes etc. If we were wrong about this, we could, for all we know, be inhabitants of w .

Even if we are not wrong about H_2O 's filling our lakes etc., we *can* easily *imagine* ourselves to be, i.e. we can conceive of w as actual. We can equally consider a world as actual in which we *falsely* believe that watery stuff is H_2O and which is thus a world in which we falsely believe to be able to distinguish between what we then would take to be our actual world and w . If we conceive of a world as actual in which we are, but do not think we are, using “water” to refer to XYZ and where we imagine that XYZ is not watery, we conceive of a world as actual where “water is watery stuff” is not a priori. So it is not a priori that “water is watery stuff” is a priori.

It is an empirical issue whether or not “watery stuff” and (what we express as) “twatery stuff” have the same secondary intensions, i.e. pick out the same stuff in all possible worlds considered as counterfactual. Part of this question concerns counterfactual language use, i.e. what competent speakers of our language would say if such-and-such a world would turn out to be their actual world. Another part, however, concerns *actual* language use: whether my word “watery stuff” and what I express as “twatery stuff” apply to the same Twin-Earthian substance. “Twatery stuff”, by stipulation, applies to all and only to things having the same macrophysical properties than XYZ. But the only thing I know of the macrophysical properties of XYZ is that they are the same than those of H_2O . So to answer the question whether “watery stuff” and “twatery stuff” have the same secondary intensions, I have to rely on my knowledge that “watery stuff” describes all and only the macrophysical properties of H_2O , that is on my *de re* knowledge of water that it is watery.

The moral to draw is rather, I think, that in order to evaluate primary intensions at worlds considered as actual, we cannot but evaluate *what we believe are the primary intensions of our words*. In this belief, however, we can be wrong. If we are wrong, it is not just that we are conceptually confused: what our words mean, even in counterfactual circumstances, depends on how the world is.

Imagining having fixed the reference of “water” using “watery stuff”, we distinguish the primary from the secondary intension of “watery stuff”. So we get a regress, in many respects similar to the regress of Fregean senses. This has been remarked by Davies and Humberstone:

For the record, we should remark that the process which yielded 2-dimensional modal logic from the more familiar 1-dimensional kind can be iterated: truth can be triply relativized to a *real* actual world w_1 , a ‘floating’ actual world w_2 , and a floating reference world w_3 . (Davies & Humberstone 1980: 10)

Even if every metaphysically possible world can be considered as actual, not all of them can be so considered at the same time and by the same person. You have to keep one world fixed and evaluate possible world

²⁸ Imagine w to be future earth, where tradition, especially semantic tradition, is highly respected, but water has become sparse and is kept in laboratories, while another substance, say petrol, fills the lakes. It is not important whether you would describe this case as a case where the word “water” would have different properties, e.g. would have become Millian or would have changed from a descriptive to a proper name. What is important, however, is that it would still be the same word.

candidates *for this world*. You are considering other worlds always from the perspective of one possible world which you take to be your actual world. It might be that some of your actual world candidates have themselves more and/or different candidates than the world you keep fixed.²⁹ If one of these worlds really *is* your actual world, then you might be wrong about what candidates you have to consider. By considering one particular range of candidates, you commit yourself to the actuality of one world out of a subset of the actual world candidates. But you can, as we did in the case of the primary intension of “watery stuff”, keep track of your commitment. The price to pay is a regress.

This is how we got into the third dimension: we asked ourselves how we know that “watery stuff”, stipulated to express the primary intension of “water”, does what it was stipulated to do. We allowed for the possibility that we, if we *were* inhabitants of Twin Earth, could have another concept of water than the one we take ourselves to have. So we indexed the secondary intension of “watery stuff” and, by diagonalization, got *its* primary intension. With respect to the expression we used to express this primary intension, namely “†(watery stuff)”, we simply assumed that its secondary intension is constant across worlds considered as actual. We thus implicitly applied the rigidifying operator \mathcal{F} . But in this we could equally have been wrong. To keep track of this possibility, we would need a fourth dimension. And so it goes up the ladder.

But something has to be a priori!

How could A-intensions be not be a priori? Does not something have to ‘encapsulate’ the “way in which [its] referent depends on how the actual world turns out” (Chalmers 1995: 4). Does this not mean that our competence with such terms is *ipso facto* competence in applying it to (descriptions of) counterfactual circumstances? Chalmers seems to think so:

“Given that we have the ability to know what our concepts refer to when we know how the actual world turns out, then we have the ability to know what our concepts would refer to *if* the actual world turned out in various ways.” (Chalmers 1996: 59–60)³⁰

And does this not mean that primary intensions are independent of empirical factors?

The primary intension of a concept, unlike the secondary intension, is independent of empirical factors: the intension *specifies* how reference depends on the way the external world turns out, so it does not itself depend on the way the external world turns out. (Chalmers 1996: 57) Like an epistemic intension but unlike a subjunctive intension, a two-dimensional intension can be evaluated a priori. One needs no empirical information about the actual world, since all the relevant information is specified in the epistemic possibility. (Chalmers 2002c: 21)

It may be that, given that we know how the actual world turns out (e.g. that water is H₂O), we know what our words refer to (e.g. “water” to H₂O). Does it therefrom follow that we have the *further* knowledge that, for any possible way the actual world could turn out, our words would refer to such or such things? This strong thesis is explicitly made by Chalmers and Jackson:

Possession of a concept bestows a *conditional ability* to identify the concept’s extension, given information about hypothetical epistemic possibilities (in the broad sense of “epistemic possibility”, invoking hypotheses about the actual world that are not ruled out a priori). (Chalmers & Jackson 2001: 5)

29. Note that the claim is that you could, if one of these worlds turned out to be actual, consider more and different worlds *as actual*. This is something Chalmers and Jackson miss in their reply to the argument that “If watery stuff is H₂O then H₂O is water” is not a priori because it would be false if uttered on Twin Earth: “The latter [“If H₂O would be watery stuff, then it would not be water” – I changed the example] is a subjunctive claim about an explicitly counterfactual scenario, so at best its truth reflects negatively on the *necessity* of an entailment from [“H₂O is watery stuff”] to [“Water is H₂O”].” (Chalmers & Jackson 2001: 13)

30. Cf. also (Chalmers 2002a: 612).

Our ability may be conditional, i.e. dependent on how the actual world turned and turns out, without thereby being an ability to determine, for whatever way the actual world could turn out, what the extensions of our concepts would be. Even if our competence with expressions and concepts is criterial and our knowledge of them in that sense conditional, this does not mean that it is knowledge of a whole array of conditionals. It is true, of course, that sufficient descriptive knowledge about a world w allows you to identify the referents of your terms in w if anything does. But this does not mean that you know to what your terms would refer to if w turned out to be actual.

Linguistic knowledge is knowledge about one particular this-worldly language, not knowledge about a whole range of different and at the same time sufficiently similar languages.³¹ “Conditional knowledge”, as Chalmers calls it, is not knowledge of a whole array of conditionals. It is true, of course, that sufficient descriptive knowledge about a world w allows you to identify the referents of your terms in w if everything does. But this does not mean that you know to what your terms would refer to if w turned out to be actual.³² To know this, you need further information, not about w , but about which of the worlds you consider as actual is actual. This is so even if w is your actual world. Even then you need, over and above the totality of truths in w , the further information that w is your actual world, in order to be able to take the meaning of “watery stuff” at face value. It is not just indexical information that you need, where and who you are, what time is it and the like. It is indexical information of the form “This world is my world” and the reliance on such information is just what makes a knowledge claim empirical.

But what would be the alternative? Here is Jackson’s argument that if Quine is right we cannot say how things are at all:

Now suppose that it is impossible to effect a partition among the possibilities independently of how things actually are. [1] No mental state, no linguistic item, no diagram, no system of semaphore, divides the possibilities, except *relative* to how things actually are. Then we can never say, diagram, depict, semaphore, think, ...how things are. [2] All we can do is say (depict, think, etc.) how they are *if* ... We are always in the position of one who only ever tells you what to do if you have high blood pressure, never what to do *simpliciter*. We can say how things are *conditional* on ..., but [3] can never make an unconditional claim about how things are. We cannot detach. This is a very radical doctrine. It is not that we cannot say with complete precision how things are. We really cannot say how things are at all. (Jackson 1998: 53)

The problem with this argument is that it equivocates on [2]. Under one reading, “All we can do is (say (depict, think, etc.) how they are) *if* ...” follows from [1]. Only the different, and much stronger, reading “All we can do is say (depict, think, etc.) (how they are *if* ...)”, however, gives us [3]. The difference, again, is between having an ability conditional on ...and having the ability to (if ..., then ...).

I accept the weak, but not the strong reading. It is not that all we can say is conditional; we can, and do, detach – it’s just that we do so fallibly. Whatever we do, we do it to the best of our capacities. We do not need a priori knowledge for this.

He thinks this is compatible with linguistic indeterminacy:

31. Note that the point is not only that you cannot infer $K(\phi \rightarrow \psi)$ from $K(\phi) \rightarrow K(\psi)$, but that you cannot infer $\forall F K(\phi(F) \rightarrow \psi(F))$ from $\exists G (K(\phi(G)) \rightarrow K(\psi(G)))$. It may seem that I am advocating a quite general meaning scepticism in my claim that we can falsely take a primary intension to be non-empty. If the meaning in question is taken to be that of “water”, e.g., then it is true that I advocate a moderate scepticism in the way of Kripke. But everybody does, so this does not count against my point. If it is that of “watery stuff”, however, I indeed think that we have a clear grasp of its primary intension. As “watery stuff” is a technical philosophical term, however, this claim is in no way contrary to what anyone, excluding philosophers, believes.

32. Chalmers seems to think it does: “Like an epistemic intension but unlike a subjunctive intension, a two-dimensional intension can be evaluated a priori. One needs no empirical information about the actual world, since all the relevant information is specified in the epistemic possibility.” (SI: 21)

In practice you may not have made the decision about how you will use language that settles whether you'd use 'Cats are not animals' to pick out the possibilities to which probability should be moved should some *very* surprising discoveries be made, or whether you'd use 'There are some things that would count as cats except that they are not animals'. The nature of the possibility is the same in either case; what is unsettled is how you'd pick it out in language – and the latter may be unsettled simply because the possibility is so exotic there is no point in expending energy in deciding ahead of time which way to jump should the need arise. We can, that is, agree with the point emphasized by Hilary Putnam: in practice it is hard to find *sentences* in a natural language that are *determinately* a priori. (Jackson 1998: 54)

The point can be illustrated by considering a disanalogy between the following two claims:

- (25) If p is true, p is necessarily true.
 (26) If p is true, p is a priori.

Kripke argues that we know the truth of (25) a priori, by philosophical analysis (NN 109, 311). But what about (26)? Kripke admits that even when we have fixed the reference of F with respect to a sample, it is still an empirical question whether all or most of the items in the sample are F .³³ There is another difference: (25) makes necessity-with-respect-to- p automatically transitive. But apriority, even with respect to sentences like "water is watery stuff", is not. It does not follow from the fact that you can conceive that you can conceive that p that you can conceive that p . What you can conceive of, depends on your epistemic situation, your cognitive makeup and many other things (even what you can ideally conceive of does). You can conceive of these parameters being different, without, of course, thereby conceiving of what you can not conceive of.

It seems that we are left with two options, corresponding to the two horns of my dilemma. Either we concede that (1) is not a priori after all, in virtue of the fact we can conceive of a world as actual where it is false, and thus concede that apriority does not imply contingency of primary intensions. Or we stipulate that primary intensions capture that cognitive content of statements and allow for a regress, corresponding to the non-transitivity of apriority. In the last part of my talk I would like to explore this second option.

But it's just a function!

The dilemma I want to put up for Chalmers turns on the thesis that he is committed to the claim that primary intensions are - at least in principle - expressible. I have three arguments for this, the first *ad hominem*, a internal second one and the third programmatic. Here is the first: If primary intensions were not expressible, Chalmers could not use the existence of an explanatory gap, i.e. the non-existence of a priori truths like (1) linking phenomenal and functional concepts, to argue against physicalism.³⁴ The internal argument concerns our motivation to introduce primary intensions in the first place, i.e. to explain away the illusion of contingency. Diagonalization is something we do in order to make sense of beliefs we take to be necessarily false. It should therefore give us something we can *ascribe* to people to which we do not want to ascribe belief in the impossible proposition. The programmatic argument is that apriority is tied to knowledge and knowledge to the *contents* of sentences. If we say of someone that he knows a priori that p , we do not merely ascribe to him knowledge that there is a sense in which p is and even cannot fail to be true, but we also ascribe knowledge of the truth actually expressed by p . If a priori knowledge of p were knowledge that p expresses a truth whatever it means, any epistemological interest for it would be lost.

33. "A priori, all we can say is that it is an empirical matter whether the characteristics originally associated with the kind apply to its members universally, or even ever, and whether they are in fact jointly sufficient for membership in the kind." (NN 137, 329)

34. If all the physicalist has to do is to come up with a function, his task is too easy. Chalmers' challenge is to the physicalist is to come up with some X such that water/watery stuff = pain/ X .

Counterfactual language use

When we falsely take a primary intension to be nonempty we take a world verifying it to be an actual world candidate which is not such a candidate. This could be so because “being an actual world candidate” is a non-empty condition on worlds.³⁵ Chalmers agrees that we have, when evaluating apriority claims, to make sure that we are not misdescribing worlds we consider as actual.³⁶ But how can we do this? How can Chalmers avoid the objection that “the fact that p might, for all I know, be true in the actual world, is just irrelevant to the issue whether it is true in some possible world.” (Yablo 1993: 9)?

To rule out linguistic error, Chalmers would need an extremely strong notion of the a priori according to which p is a priori iff it is not possible to believe $\neg p$. The corresponding notion of conceivability as possible believability, however, is far too weak to imply possibility, for many impossibilities may and have been believed. That we find ourselves unable to rule out a possibility does not show that it is a possibility we think we cannot rule out rather than an impossibility.

Does idealization help? Chalmers (2002b) claims that ideal positive and primary conceivability of p , that is conceivability after ideal rational reflection of a scenario that would verify p if actual, entails that p is possible. So it could be replied that the regress only arises in the case of less than ideal conceivability. I do not think that is true. To see why, consider Yablo’s analogy of conceiving with perceiving:

“Just as to perceive that p is to be in a state that (i) is veridical only if p , and that (ii) moves you to believe that p , to find p conceivable is to be in a state which (i) is veridical only if possibly p , and (ii) moves you to believe that p is possible.” (Yablo 1993: 7)

There is no a priori guarantee that perceptual error, even in principle and under ideal perceptual circumstances, can be ruled out (because, say, every possible perceptual apparatus is inherently defective). Why should there be such a guarantee in the case of conceivability? But the case is worse than this: with regard to perception, we are not forced to skeptical conclusions from the absence of such an a priori guarantee: even if there is no guarantee that p is true, given that I perceive that p under ideal circumstances, it may still be that my perceiving that p gives strong, but defeasible, evidence that p . Chalmers, however, cannot retreat to the claim that conceivability (even conceivability under ideal conditions) is defeasible evidence for possibility. His analysis of conceivability in terms of primary intensions entails that in any case where I really conceive that p , the primary intension of p is non-empty. So Chalmers is committed to the claim that under certain circumstances, conceivability is an *infallible* guide to possibility.

Rather than justifying the step from conceivability to possibility, the two-dimensional framework deployed presupposes the legitimacy of the very step in question.

It seems that we are left with two options, corresponding to the two horns of my dilemma. Either we concede that (i) is not a priori after all, in virtue of the fact we can conceive of a world as actual where it is false, and thus concede that aposterioricity does not imply contingency of primary intensions. Or we stipulate that primary intensions capture that cognitive content of statements and allow for a regress, corresponding to the non-transitivity of apriority.

35. This is e.g the case in Kaplan’s *Logic of Demonstratives* where contexts of utterance are required to contain a speaker in order to make “I exist” come out as logically true. As Chalmers takes primary and secondary intensions to be (possibly partial) functions defined over the same space of possible worlds, I will not pursue this line further. Another reason for being wrong about a primary intension, however, is that we may falsely believe that something is an actual world candidate which is not even possible, but belongs to the so-called “outer sphere” of possibility.

36. “On this view of conceivability, the conceivability of a statement involves two things: first, the conceivability of a relevant world, and second, the truth of the statement in that world. It follows that in making conceivability judgements, one has to make sure that one describes the world that one is conceiving correctly, by properly evaluating the truth of a statement in the world. One might at first glance think it is conceivable that Goldbach’s conjecture is false, by conceiving of a world where mathematicians announce it to be so; but if in fact Goldbach’s conjecture is true, then one is *misdescribing* this world; it is really a world in which the conjecture is true and some mathematicians make a mistake.” (Chalmers 1996: 68)

The conditional fallacy

Ignorance about the actual world does not matter for knowledge about the A-extensions of words. For the A-extension of T at a world w is the extension of T at w *given* w is the actual world, and so does not depend on whether or not w is in fact the actual world. Or, in other words, knowledge of the A-intension of T does not require knowledge of the nature of the actual world. [...] What we can know independently of knowing what the actual world is like can properly be called a priori. (Jackson 1998: 50)

How could A-intensions not be a priori? Does not something have to ‘encapsulate’ the “way in which [its] referent depends on how the actual world turns out” (Chalmers 1995: 4). Does this not mean that our competence with such terms is *ipso facto* competence in applying it to (descriptions of) counterfactual circumstances? Chalmers seems to think so:

“Given that we have the ability to know what our concepts refer to when we know how the actual world turns out, then we have the ability to know what our concepts would refer to *if* the actual world turned out in various ways.” (Chalmers 1996: 59–60)³⁷

And does this not mean that primary intensions are independent of empirical factors?

The primary intension of a concept, unlike the secondary intension, is independent of empirical factors: the intension *specifies* how reference depends on the way the external world turns out, so it does not itself depend on the way the external world turns out. (Chalmers 1996: 57) Like an epistemic intension but unlike a subjunctive intension, a two-dimensional intension can be evaluated a priori. One needs no empirical information about the actual world, since all the relevant information is specified in the epistemic possibility. (Chalmers 2002c: 21)

It may be that, given that we know how the actual world turns out (e.g. that water is H₂O), we know what our words refer to (e.g. “water” to H₂O). Does it therefrom follow that we have the *further* knowledge that, for any possible way the actual world could turn out, our words would refer to such or such things? This strong thesis is explicitly made by Chalmers and Jackson:

Possession of a concept bestows a *conditional ability* to identify the concept’s extension, given information about hypothetical epistemic possibilities (in the broad sense of “epistemic possibility”, invoking hypotheses about the actual world that are not ruled out a priori). (Chalmers & Jackson 2001: 5)

Our ability may be conditional, i.e. dependent on how the actual world turned and turns out, without thereby being an ability to determine, for whatever way the actual world could turn out, what the extensions of our concepts would be. Even if our competence with expressions and concepts is criterial and our knowledge of them in that sense conditional, this does not mean that it is knowledge of a whole array of conditionals. It is true, of course, that sufficient descriptive knowledge about a world w allows you to identify the referents of your terms in w if anything does. But this does not mean that you know to what your terms would refer to if w turned out to be actual.

But what would be the alternative? Here is Jackson’s argument that if Quine is right we cannot say how things are at all:

³⁷. Cf. also (Chalmers 2002a: 612).

Now suppose that it is impossible to effect a partition among the possibilities independently of how things actually are. [1] No mental state, no linguistic item, no diagram, no system of semaphore, divides the possibilities, except *relative* to how things actually are. Then we can never say, diagram, depict, semaphore, think, ...how things are. [2] All we can do is say (depict, think, etc.) how they are *if* We are always in the position of one who only ever tells you what to do if you have high blood pressure, never what to do *simpliciter*. We can say how things are *conditional* on ..., but [3] can never make an unconditional claim about how things are. We cannot detach. This is a very radical doctrine. It is not that we cannot say with complete precision how things are. We really cannot say how things are at all. (Jackson 1998: 53)

The problem with this argument is that it equivocates on [2]. Under one reading, “All we can do is (say (depict, think, etc.) how they are) *if* ...” follows from [1]. Only the different, and much stronger, reading “All we can do is say (depict, think, etc.) (how they are *if* ...)”, however, gives us [3]. The difference, again, is between having an ability conditional on ...and having the ability to (if ..., then ...).

***n*-Dimensionalism**

Even if every metaphysically possible world can be considered as actual, not all of them can be so considered at the same time and by the same person. You have to keep one world fixed and evaluate possible world candidates *for this world*. It might be that some of these candidates have more and different candidates than the world you keep fixed.³⁸ But you can, as we did in the case of the primary intension of “watery stuff”, keep track of your commitment. The price to pay is a regress.

Recall how we got to the third dimension: we asked ourselves how we know that “watery stuff”, supposed to express the primary intension of “water”, does what it was supposed to be. Could we not, if we *were* inhabitants of Twin Earth, have another concept of water? So we indexed the secondary intension of “watery stuff” and, by diagonalization, got *its* primary intension. With respect to the expression we used to express this primary intension, namely “†(watery stuff)”, we simply assumed that its secondary intension is constant across worlds considered as actual. We thus implicitly applied the rigidifying operator \mathcal{F} . But in this we could have been wrong. To keep track of this possibility, we would need a fourth dimension. So we get some sort of a Fregean regress of higher-order senses.³⁹

The Fregean regress is due to the presence of a functor $[\dots]_1$ mapping expressions to primary intensions in the object language and the claim that every primary intension should, at least in principle, be expressible.

Here is another way to argue for the induction step: A statement p is conceivable (i.e. $\neg p$ is not a priori), Chalmers claims, iff its primary intension $[p]_1$ is possible. The primary intension is supposed to represent the a priori accessible content of p . What we are doing, then, when we conceive of a p -world as actual, is not just conceiving of p having a non-empty primary intension. Instead, it is conceiving of a world where some other phrase q we take to express the primary intension of p is true. We are conceiving of the truth not of p but of the conceptually equivalent “ $[p]_1$ ”. The “conceivability implies possibility” thesis then amounts to:

$$(27) \quad p \text{ is conceivable} \iff “[p]_1” \text{ is possible}$$

³⁸. Note that the claim is that you could, if one of these worlds turned out to be actual, consider more and different worlds *as actual*. This is something Chalmers and Jackson miss in their reply to the argument that “If watery stuff is H_2 then H_2O is water” is not a priori because it would be false if uttered on Twin Earth: “The latter [“If H_2O would be watery stuff, then it would not be water” - I changed the example] is a subjunctive claim about an explicitly counterfactual scenario, so at best its truth reflects negatively on the *necessity* of an entailment from [“ H_2O is watery stuff”] to [“Water is H_2O ”].” (CA: 13)

³⁹. This has been remarked by Davies and Humberstone: “For the record, we should remark that the process which yielded 2-dimensional modal logic from the more familiar 1-dimensional kind can be iterated: truth can be triply relativized to a *real* actual world w_1 , a ‘floating’ actual world w_2 , and a floating reference world w_3 .” (1980: 10)

By the definition of primary intensions, conceiving of “[p]₁” is conceiving of something that expresses *its* primary intension, “[[p]₁]₁”. So (37) gives us:

(28) p is conceivable \iff “[[p]₁]₁” is possible

To conceive that p , it is not enough to believe that p expresses some truth or other at some world. One has to imagine that the sentence, in a world where its relevant meaning is preserved, expresses a truth which is conceptually equivalent to the truth we wanted to conceive of. For this to count as a way of imagining the truth of what is expressed by the sentence, however, one has to know what the sentence means. So one has to know its primary intension. In order to know that our evidence for the possibility of p does not just depend on our ignorance of its primary intension, we have to know that what “[p]₁” expresses is possible. To know this, we have to conceive of a world where it is true. We have no other access to what is possible than via conceivability. To claim otherwise would be to claim that brute facts about the way we speak constrain the realm of possibilities. To conceive of a world where “[p]₁” is true, however, we have to presuppose that what “[[p]₁]₁” expresses is possible. And so we get the regress.

The analysis of conceivability in terms of non-empty primary intensions seems to force upon us a non-S₅ logic of metaphysical modality. For we conceive of a world as actual where water is not H₂O, so there is a world where watery stuff is not H₂O. We can equally conceive of a world as actual where science tells us that water is necessarily H₂O. So there is a world where watery stuff is necessarily H₂O. So $\Diamond\Box\alpha$ does not imply $\Box\alpha$. It has to be, however, transitive, at least for necessary a posteriori truths like “water is H₂O”, for it is supposed to accommodate for Kripke’s observation that such sentences are, if true, necessarily true. So in any world w I consider as actual, if “water is H₂O”, as understood by us, is true in w , then it is true in all worlds reachable from w . If it is possible that water is *possibly* H₂O, then science has not found out what it claims to have found out.

It is, e.g., possible to imagine a world where the clear drinkable liquid in the lakes and oceans is not what we (or members of our linguistic community or experts or whoever) pointed to when we fixed the reference of “water”. It is possible that, unbeknownst to us, the reference of “water” has been fixed in other ways than we think it is or was. It is tempting to reply that it is (merely) *linguistic* knowledge that we would be lacking in this case. There are two replies to this: First, linguistic knowledge is not in an interesting sense a priori; second, if it were a priori, it would give us a priori empirical knowledge. ‘Linguistic’ knowledge in this use of the word, is knowledge about how reference is fixed, where ‘fixing the reference’ is an *activity* of certain members of my linguistic community, an empirical fact which I can easily imagine to have been otherwise. Kripke, Donnellan and others have described and imagined cases where originally descriptive names have become proper. How do we realize this? It certainly depends on whether a major or the most important part of our linguistic community is prepared to accept sentences of the form “ a is not ϕ ” (where “the ϕ ” is the description by which the reference of “ a ” has been fixed) as possibly true. The converse, however, is not true: even if they regard such sentences as plainly false, this does not show that they take it to be of the form “the ϕ is not ϕ ”. The second point pertains to a retreat position of the advocate of the contingent a priori sketched by Donnellan (1978: 55). Cannot we at least say that the following is a priori?

(29) Provided the ϕ exists, “ a is the ϕ ” expresses a contingent truth.

(29) is supposed to speak about a particular language L and more specifically about the truth-theory for that language. We know a priori that any truth-theory must give us all instances of the T scheme and so we know, if L is our own language, that (29) is true, provided that the ϕ exists, only if a is the ϕ .

But what about Twin Earth, where H₂O is not watery, but what you would call “water” there, namely XYZ, is? Twin Earth is a world where you do not know that water is H₂O and where you therefore consider it possible that water is not H₂O but, say, XYZ.

(30) If XYZ would be watery, then XYZ would be water.

Given that (23) is true, you can reason from XYZ being watery to its being water. Consider now the following sentences:

- (31) If XYZ is called “water”, then XYZ is water.
- (32) If H₂O is not called “water”, then water is not called “water”.

Even if you do not know that water is H₂O, there seems nothing that *precludes* (32) from being true. For all you know, there might be such a connection between the names of H₂O and of water

This contrast corresponds to a difference between indicative and subjunctive conditionals. Of the following two sentences,

- (33) If “tail” would mean “leg”, then horses would have four tails.
- (34) If “tail” means “leg”, then horses have four tails.

(33) is clearly false: a change in our language cannot change the anatomy of horses. (34), however, seems true: if “tail” and “leg” are synonymous, horses have four tails iff they have four legs. The upshot of this, I think, is that indicative but not subjunctive conditionals can change the ‘material parts’ (but not the structure) of the language in which their consequent has to be interpreted. The reason for this is that indicative conditionals require only that the consequent is true of every world *in* which the antecedent is true, while subjunctive conditionals require that it is true of every world *of* which the antecedent is true.

Consider again the following two claims:

- (35) If p is true, p is necessarily true.
- (36) If p is true, p is a priori.

(35) makes necessity-with-respect-to- p automatically transitive. But apriority, even with respect to sentences like “water is watery stuff”, is not. It does not follow from the fact that you can conceive that you can conceive that p that you can conceive that p . What you can conceive of, depends on your epistemic situation, your cognitive make-up and many other things (even what you can ideally conceive of does). You can conceive of these parameters being different, without, of course, thereby conceiving of what you cannot conceive of.⁴⁰

The Fregean regress is due to the presence of a functor $[\dots]_1$ mapping expressions to primary intensions into the object language and the claim that every primary intension should, at least in principle, be expressible, is in many ways similar to the Fregean regress of higher-order senses.

Here is another argument for the induction step: A statement p is conceivable (i.e. $\neg p$ is not a priori), Chalmers claims, iff its primary intension $[p]_1$ is non-empty. The primary intension is supposed to represent the a priori accessible content of p . What we are doing, then, when we conceive of a p -world as actual, is not just conceiving of p having a non-empty primary intension. Instead, it is conceiving of a world where some other sentence q we take to express the primary intension of p is true. We are conceiving of the truth not of p but of the conceptually equivalent “ $[p]_1$ ”. The “conceivability implies possibility” thesis then amounts to:

- (37) p is conceivable \iff “ $[p]_1$ ” is possible

40. Here, finally, is the opportunity to consider the import of idealization. Can an idealized notion of conceivability (“positively conceivable after ideal rational reflection”) make apriority transitive? I think not, for *every* notion of conceivability which makes it a *guide* to possibility (and is not simply equated with it) has to be relativized somehow to what humans (even in principle) can do. Humans, however, are clearly capable to imagine being able to do more than they are.

By the definition of primary intensions, establishing the non-emptiness of “[p]₁” is conceiving of something that expresses *its* primary intension, “[[p]₁]₁”. So (37) gives us:

$$(38) \quad p \text{ is conceivable} \iff "[[p]_1]_1" \text{ is possible}$$

To conceive that p , it is not enough to believe that p expresses some truth or other at some world. One has to imagine that the sentence, in a world where the *relevant* parts of its meaning are preserved, expresses a truth which is conceptually equivalent to the truth we wanted to conceive of. For this to count as a way of imagining the truth of what is expressed by the original sentence, however, one has to know what the sentence means. And this, if Kripke is right, is a posteriori knowledge, depending on how the real actual world turns out. So one has to know the primary intension of the sentence in question. In order to know that our evidence for the possibility of p does not just depend on our ignorance of its primary intension, we have to know that what “[p]₁” expresses is possible.⁴¹ To know this, we have to conceive of a world where it is true. We have no other access to what is possible than via conceivability.⁴² To conceive of a world where “[p]₁” is true, however, we have to presuppose that what “[[p]₁]₁” expresses is possible. And so we get the regress.

A world of monsters

Here is an example: To answer the question whether “water is watery stuff” is a priori, we have to do the following: fix a world w we take to be the actual world, consider a world v as actual and ask ourselves whether what we would call “water” in v is what we call “watery stuff” in w , when both are evaluated in v . If what we *here* call “watery stuff” is not what we *there* would call “water”, then it is not a priori. In the preceding sentence, “here” and “there” function as operators on contexts of utterance, something Kaplan calls ‘monsters’. For we can vary the world we fix as actual and imagine the meaning of “watery stuff” to be different from what we take it to be.

Two days ago, we met an operator that could do that, namely “if I were you”. It seems to me that the fact that “If I were you then I would stay away from me” is sensible, but that “If I am you then I stay away from me” is not, is evidence that we can and do consider contexts of utterance as counterfactual. If this is true and if we accept Chalmers linguistic criterion for the ‘considering as actual’ vs. ‘considering as counterfactual’ distinction, then we can consider as counterfactual the possibility that the primary intensions of our words are different from what they in fact are.

Another monster is the operator studied by Davies & Humberstone (1980: 10):

$$(39) \quad (w_1, w_2, w_3) \models R\alpha \iff (w_1, w_1, w_3) \models \alpha$$

which gives us strong necessities in the two-dimensional framework: $RA\alpha \leftrightarrow A\alpha$ will be deeply contingent (i.e. its \mathcal{FA} -modalization is false and it thus has a contingent diagonal intension) but arguably a pr

We easily get other monsters. One is the operator studied by Davies and Humberstone ((1980: 10):

$$(40) \quad (w_1, w_2, w_3) \models R\alpha \iff (w_1, w_1, w_3) \models \alpha$$

which gives us strong necessities in the two-dimensional framework: $RA\alpha \leftrightarrow A\alpha$ will be deeply contingent (i.e. its \mathcal{FA} -modalization is false) but arguably a priori. But of course we can define another notion of a priori for the three-dimensional case.

41. Though Chalmers is not committed to the claim that $[p]_1 = [[p]_1]_1$, he is, I think, committed to (38). With reference to Kripke, (Chalmers 1996: 133) claims that the primary and secondary intensions of “consciousness” and of phenomenal concepts in general coincide: “What it takes for a state to be a conscious experience in the actual world is for it to have a phenomenal feel, and what it takes for something to be a conscious experience in a counterfactual world is for it to have a phenomenal feel.” (cf. also Chalmers 1995: 6).

42. To claim otherwise would be to claim that brute facts about the way we speak constrain the realm of possibilities.

Earlier in the talk we met another multidimensional operator, namely Stalnaker's dagger, without adjunction of the "fixedly"-operator \mathcal{F} , corresponding roughly to the locution "what is said in the utterance s is true" and projecting the n -dimensional diagonal onto the $n - 1$ -dimensional hyperplane. It seems to me that iterations of this operator make sense.

I think that n -ary intensions may be useful in the philosophy of language. For what I say when I specify what is said by your utterance of p does not only depend on what your utterance means but also what the words mean I use to specify the content of your utterance.

I think that n -ary intensions could be useful in epistemology in general. Think of the world where you are a brain in a vat and of Putnam's famous quick-and-dirty argument that if he were a brain in a vat, he could not consider the world as actual where he is a brain in a vat and so, necessarily, he is not a brain in a vat. This is hardly convincing, but why? Partly because he could reason *like this* (or at least utter the same words) if he *were* a brain in a vat. But even if he were a brain in a vat, he could say "what would be said by 'I am a brain in a vat' if I were not a brain in a vat is true", and what he would say by this utterance would be true. Or take fictions: what is fun about fictions is not only that I can consider the fictional world as actual, but that I can imagine it *to be* actual, e.g. that I can take the fictional "I" to refer to me. For what I say when I specify what is said by your utterance of p does not only depend on what your utterance means but also what the words mean I use to specify the content of your utterance.

I hope to have laid out some middle ground between two extreme views on the a priori. On the one hand, we have Chalmers defending a substantive notion of apriority, keeping the language fixed and varying nevertheless in *some* sense at least the meaning of "water". On the other hand, we have people who think that for p to have a necessary primary intension is for p to be true whatever it means. I think it is worth the effort to vary the meanings of words independently, to make distinctions between varying degrees of keeping the language fixed and thus to go along with more dimensions than just with 2.

References

- Block, Ned & Robert C. Stalnaker, 1999. Conceptual Analysis, Dualism, and the Explanatory Gap. *The Philosophical Review* 108(1): 1–46.
- Chalmers, David J., 1995. The Components of Content. Unpublished manuscript.
- Chalmers, David J., 1996. *The Conscious Mind: In Search of a Fundamental Theory*. Oxford: Oxford University Press.
- Chalmers, David J., 1999. Materialism and the Metaphysics of Modality. *Philosophy and Phenomenological Research* 59: 473–496.
- Chalmers, David J., 2002a. The Components of Content. In Chalmers, David J. (editor) *Philosophy of Mind: Classical and Contemporary Readings*, pp. 608–633. Oxford: Oxford University Press.
- Chalmers, David J., 2002b. Does Conceivability Entail Possibility? In Gendler, Tamar Szabó & John Hawthorne (editors) *Conceivability and Possibility*, pp. 145–200. Oxford: Oxford University Press.
- Chalmers, David J., 2002c. On Sense and Intension. In Tomberlin, James E. (editor) *Philosophical Perspectives 16: Language and Mind*, pp. 135–182. Oxford: Basil Blackwell Publishers.
- Chalmers, David J. & Frank Jackson, 2001. Conceptual Analysis and Reductive Explanation. *The Philosophical Review* 110(3): 315–361.
- Cole, Peter (editor), 1978. *Syntax and Semantics 9: Pragmatics*. New York: Academic Press.
- Davies, Martin Kinsey & I. Lloyd Humberstone, 1980. Two Notions of Necessity. *Philosophical Studies* 38(1): 1–30.
- Davis, Steven (editor), 1991. *Pragmatics: A Reader*. Oxford: Oxford University Press.
- Donnellan, Keith S., 1966. Reference and Definite Descriptions. *The Philosophical Review* 75: 281–304. Reprinted in [Steinberg & Jacobovits \(1971: 100–114\)](#) and in [Schwartz \(1977: 42–65\)](#).
- Donnellan, Keith S., 1978. Speaker Reference, Descriptions and Anaphora. In [Cole \(1978\)](#), pp. 47–68. Reprinted in [French et al. \(1979: 28–44\)](#).
- Evans, Gareth, 1979. Reference and Contingency. *The Monist* 62(1): 161–189. Reprinted in [Evans \(1985: 178–213\)](#).
- Evans, Gareth, 1985. *Collected Papers*. Oxford: Clarendon Press.
- French, Peter A., Theodore E. Uehling, Jr. & Howard K. Wettstein (editors), 1979. *Midwest Studies in Philosophy 2: Contemporary Perspectives in the Philosophy of Language*. 2 edition. Minneapolis, Minnesota: University of Minnesota Press. First edition: ?.
- Jackson, Frank, 1998. *From Metaphysics to Ethics: A Defense of Conceptual Analysis*. Oxford: Oxford University Press.
- Kaplan, David, 1989. Demonstratives. In Almog, Joseph, John R. Perry & Howard K. Wettstein (editors) *Themes from Kaplan*, pp. 481–563. Oxford: Oxford University Press. Widely circulated from 1977 on.
- Kripke, Saul A., 1972. Naming and Necessity. In Davidson, Donald & Gilbert H. Harman (editors) *Semantics of Natural Language*, pp. 253–355, 763–769. Number 40 in Synthese Library, Dordrecht: D. Reidel Publishing Co. Reprinted as [Kripke \(1980\)](#).
- Kripke, Saul A., 1980. *Naming and Necessity*. Oxford: Basil Blackwell Publishers.
- Lewis, David, 1973. *Counterfactuals*. Cambridge, Massachusetts: Harvard University Press. Cited after republication as [Lewis \(2001\)](#).
- Lewis, David, 2001. *Counterfactuals*. Oxford: Basil Blackwell Publishers. Republication of [Lewis \(1973\)](#).
- Putnam, Hilary, 1975. Introduction: Science as Approximation to Truth. In *Mathematics, Matter and Method. Philosophical Papers, Volume 1*, pp. vii–xiv. Cambridge: Cambridge University Press. Second edition: [Putnam \(1979\)](#).
- Putnam, Hilary, 1979. *Mathematics, Matter and Method. Philosophical Papers, Volume 1*. 2 edition. Cambridge: Cambridge University Press. First edition: ?.
- Schwartz, Stephen P. (editor), 1977. *Naming, Necessity, and Natural Kinds*. Ithaca, New York: Cornell University Press.
- Stalnaker, Robert C., 1978. Assertion. In [Cole \(1978\)](#), pp. 315–332. Reprinted in [Davis \(1991\)](#) and in [Stalnaker \(1999: 78–95\)](#).
- Stalnaker, Robert C., 1999. *Context and Content: Essays on Intensionality, Speech and Thought*. Oxford: Oxford University Press.
- Steinberg, Danny D. & Leon A. Jacobovits (editors), 1971. *Semantics: An Interdisciplinary Reader in Philosophy, Linguistics, and Psychology*. Cambridge: Cambridge University Press.
- Stroll, Avrum, 1989. What Water Is or Back to Thales. In French, Peter A., Theodore E. Uehling, Jr. & Howard K. Wettstein (editors) *Midwest Studies in Philosophy 14: Contemporary Perspectives in the Philosophy of Language II*, pp. 258–274. Notre Dame, Indiana: University of Notre Dame Press.
- Yablo, Stephen, 1993. Is Conceivability a Guide to Possibility? *Philosophy and Phenomenological Research* 53(1): 1–42.