



# How not to conclude possibility from conceivability

Pierre Saint Germier

## ► To cite this version:

Pierre Saint Germier. How not to conclude possibility from conceivability. European Society for Analytic Philosophy, Milan, Sep 2011, Italy. hal-00932192

**HAL Id: hal-00932192**

**<https://hal.science/hal-00932192>**

Submitted on 16 Jan 2014

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# How not to conclude possibility from conceivability

Pierre Saint-Germier

August 30, 2011

## 1 Introduction

There is a long tradition of philosophical arguments seeking to prove substantial metaphysical theses starting from epistemic premises. In particular, in order to prove that something is possible, in an interesting metaphysical sense, some philosophers start from what is conceivable (in some sense to be determined). The aim of this paper is to distinguish and evaluate the different ways one can reject the metaphysical conclusion of a conceivability argument while accepting its epistemic premise. To do so, I will focus on a particular conceivability argument, namely the zombie argument attributed to David Chalmers (1996). However, the conclusions are intended to be generalizable to any conceivability argument.

The zombie argument is intended to be a refutation of materialism understood in modal terms as a supervenience relation between the phenomenal and the physical. If we represent the conjunction of all the physical truths obtaining in the actual world by  $P$  and the proposition that someone is phenomenally conscious by  $Q$ , then materialism is false just in case it is possible that  $P$  and not  $Q$ . The zombie argument then goes something like this:

- Z1. it is conceivable that  $P$  and not  $Q$  (ass.)
- Z2. if it is conceivable that  $P$  and not  $Q$ , then it is possible that  $P$  and not  $Q$  (ass.)
- Z3. therefore it is possible that  $P$  and not  $Q$  (1,2)

Consider two characters  $A$  and  $B$ . Suppose that  $A$  submits this argument to  $B$ . Suppose that  $B$  accepts premise 1. Our question is: how can  $B$  consistently accept premise 1 and reject premise 3?

It is clear that most of this depends on the bridge principle represented by premise Z2. How shall we understand this principle? What does “conceivable” and “possible” mean in this context ?

## 2 Conceivability and possibility

The notion of possibility that is relevant here should be distinguished from the notion of physical possibility, or possibility according to the laws of nature. A proposition is physically possible iff it is true in at least one physically possible world. When we deal with physical possibility, our quantifiers are restricted to a particular class of worlds. But when we are concerned with the zombie argument, it should be clear that we don’t restrict our quantifiers to a particular class of worlds, but quantify over possible worlds *tout court*. It may be physically impossible that (P and not Q). In fact, Chalmers agrees that there is no physically possible world such that (P and not Q) is true there. The question is whether there is a possible world, without such a restriction, in which (P and not Q). Let us call those worlds that are possible *tout court*, metaphysically possible worlds. We say that it is metaphysically possible that  $\phi$  iff for some world  $w$ ,  $\phi$  is true at  $w$ .

Chalmers, following Kripke, gives an intuitive test for metaphysical possibility. A world is metaphysically possible if it could have been created by God. For example, the physical constants might have been different from what they are in the actual world since it is perfectly intelligible that God could have given different values to the physical constants. So the proponent of the zombie argument contends that God could have created a world in which (P and not Q) while the critic of the argument claims that once God has created the physical world, no more work is required to give phenomenality a place in the natural world: once the physical is given, the phenomenal is given too. So there is no possible world at which (P and not Q).

Presumably, what “conceivability” means in the context of premise Z2, is something like conceptual coherence. It seems that incoherence is where our conceiving capacities end. We cannot conceive things given by incoherent descriptions, such as a male vixen, or a married bachelor.

So we can safely state the following principle:

(C-C) If it is conceivable that  $\phi$ , then  $\phi$  is coherent

Let us say that a proposition is incoherent if it is an instance of an antilogy or if it comes down to an antilogy after we substitute synonyms for synonyms. For example

(1) There is a married bachelor

comes down to the following antilogy if we substitute “unmarried man” to “bachelor”:

(2)  $\exists x((Fx \wedge \neg Gx) \wedge Gx)$

Now in order to back up premise Z2, it is necessary and sufficient to introduce the following principle of plenitude :

(PP) If  $\phi$  is coherent, then there is some possible world  $w$  such that  $\phi$  is true at  $w$ .

of which Z2 is an instance.

The conjunction of (C-C) and (PP) implies the following bridge principle

(C-P) If is conceivable that  $\phi$ , then it is possible that  $\phi$

Both (C-C) and (P-P) seem acceptable. For (C-C) to be false one would have to show that incoherent things can be conceived, which seems not to be the case. For (PP) to be false, there will have to be coherent propositions true at no world. But this would seem arbitrary. On what grounds could we refuse possibility to a perfectly coherent proposition ?

So far, we have mostly spoken for A. Let us see now how B can react.

### 3 Argument 1 : Against (C-C)

Let us first introduce a distinction, borrowed from Chalmers (2002), concerning the notion of conceivability :

- (3)  $\phi$  is prima facie conceivable if  $\phi$  is coherent on first appearances.
- (4)  $\phi$  is ideally conceivable if there is a possible subject for whom  $\phi$  is prima facie conceivable, with justification that is undefeatable by better reasoning.

Now B could say that the notion used by A is an ideal notion of conceivability. For certain very complex statement (and clearly “P and not Q” is a quite complex statement), there is a gap between what I can conceive and what an ideal reasoner could conceive. For example, there might be in such complex statements incoherences that I’m not able to detect, due to some cognitive limitations of mine. So B could say: suppose I can conceive

prima facie that P and not Q. It doesn't mean that an ideal reasoning on "P and not Q" could not defeat this conceivability judgement of mine. So I cannot infer safely from my judgement of conceivability to the absence of incoherence which is necessary to apply the (PP).

In other words, A faces a dilemma. In order to get (C-C) A need to remove the fallibility associated with prima facie conceivability. But so doing, it makes conceivability too remote from what I can conceive. If A wants to convince B, she needs to appeal to what B can conceive, not to what an ideal reasoner can conceive. Alternatively, if A chooses a more accessible notion of conceivability, then it introduces a source of fallibility in the argument : the (prima facie) conceivability of "P and not Q" is no proof of its possibility.

## 4 Argument 2: Against (PP)

Even if we grant the truth of (C-C), A still has a problem with (PP). It is easy to see that if there are a posteriori necessities, then there are counterexamples to (PP), that is there are impossible coherent propositions.

Let  $\phi$  be a necessary a posteriori proposition. Let  $\psi$  be the negation of  $\phi$ . It follows that  $\psi$  is not a priori false, because  $\phi$  is a posteriori. And  $\psi$  is impossible because  $\phi$  is necessary. Now if  $\psi$  is not a priori false, then  $\psi$  is coherent. For if  $\psi$  were incoherent, then  $\psi$  would be a priori false. So  $\psi$  is both coherent, and impossible. Therefore (PP) is defeated.

Thus a second line of resistance to conceivability arguments would be to accept that the considered scenario is coherent, while rejecting (PP) on the grounds that it is incompatible with a posteriori necessities.

This line of resistance however requires that B be able to show that there are relevant a posteriori necessities. Kripke has shown that there are a posteriori necessities of certain types ("Hesperus is Phosphorus", "water is H<sub>2</sub>O"). If "P and not Q" is an a posteriori necessity, as some type-B materialist argue, then the zombie argument does not go through. For "P and not Q" will come out conceivable but impossible.

However if B appeals to Kripkean a posteriori necessities (KAPNs), she has to consider seriously Chalmers' Two-Dimensional Semantics, which can be interpreted as a way to reconcile KAPNs with some versions of the (PP).

## 5 Plenitude with a posteriori necessities (I)

Before we turn to the two-dimensionalist solution, let's consider a possible move for A.

A first way to reconcile (PP) with a posteriori necessities would be to say that (PP) concerns only logical possibility, whereas KAPNs concern metaphysical possibility. So a second version of (PP) would be :

(PP\*) If  $\phi$  is coherent, then there is a logically possible world  $w$  such that  $\phi$  is true at  $w$ .

Then, the conflict between (PP\*) and KAPNS disappears. There are logically possible worlds at which “water is H<sub>2</sub>O” is false but no metaphysically possible worlds at which it is false.

However the cost of this solution is to introduce a modal dualism, that is a dualism at the level of worlds between logical possibility and metaphysical possibility. And this is bad for A. For the conceivability of (P and not Q) only entails the existence of a logically possible world, but not of a metaphysically possible world. Moreover, if (if P then Q) is an a posteriori necessity, as some physicalists contend, then there will be no metaphysically possible world at which (P and not Q). So the epistemic premise does not support the metaphysical conclusion.

Consequently, A needs to find a way to reconcile (PP) and KAPNS while preserving a form of modal monism. This is precisely what Chalmers’ two-dimensionalism allows to do.

## 6 Plenitude with a posteriori necessities (II)

Chalmers’ idea is to distinguish two ways one can consider the truth of a statement at a possible world. One can consider the truth of a statement at a possible world considered as *actual* or at a possible world considered as *counterfactual*.

Consider for example the statement

(5) Water is XYZ

and consider two world  $w_1$  and  $w_2$ . Both worlds are identical except that in  $w_1$  the clear drinkable liquid in our environment is XYZ whereas it is H<sub>2</sub>O in  $w_2$ . For all I know a priori,  $w_1$  could be the actual world, so (5) is true at  $w_1$  considered as actual. But  $w_2$  is the actual world, so (5) is true at no world considered as counterfactual. For if “water is H<sub>2</sub>O” is true at the actual world, then it follows from the Kripkean considerations that it is true at every world considered as counterfactual.

Given this distinction, we can define two kinds of intensions. The primary intension of a statement is the set of worlds considered as actual at which the

statement is true. The secondary intension of a statement is the set of worlds considered as counterfactual at which the statement is true. We say that  $w$  verifies  $\phi$  if  $w$  is a member of  $\phi$ 's primary intension and that  $w$  satisfies  $\phi$  if  $w$  is a member of  $\phi$ 's secondary intension.

Within this framework, we can define a new version of the principle of plenitude:

(PP\*\*) If  $\phi$  is coherent, then there is a possible world  $w$  such that  $w$  verifies  $\phi$

Now the thesis that there are KAPNs can be formulated as follows:

(K) If a KAPNs is true in the actual world, then every possible world satisfies it.

Then we can define notions of logical and metaphysical possibilities at the level of statements, without any commitment to Modal Dualism. Let us say that  $\phi$  is logically possible if some world verifies it and that  $\phi$  is metaphysically possible if some world satisfies it.

Then "Water is XYZ" will come out logically possible though metaphysically impossible. But this distinction is a distinction at the level of statements, not at the level of worlds. There is only one space of possible worlds.

So in this framework, we can have a version of the principle of plenitude, compatible with both the existence of KAPNs and Modal Monism.

Modal Monism is important because it seems to allow some move from the logical possibility of "P and not Q" to the metaphysical possibility of "P and not Q", since both notions of possibility are defined over the same space of worlds. The logical possibility of "P and not Q" introduces a *world* verifying P but not Q. What A needs to do now is to show that this world also satisfies "P and not Q".

However, this is not as easy as it seems. It would require that the primary and secondary intensions of "P and not Q" coincide. But this is debatable. Although it is plausible that the primary and secondary intensions of "Q" coincide, this is not so for "P". It is plausible that the primary and secondary intensions of the fundamental physical terms (like "mass", "electron", etc.) diverge. If this is the case, then there is no straightforward way from the logical possibility of "P and not Q" to the metaphysical possibility of "P and not Q".

It is interesting to note that Chalmers (2010, chap. 6) himself recognizes this fact and claims that the logical possibility of "P and not Q" only entails that "P and not Q" is metaphysically possible *or* that Russellian monism is

true. Russellian monism is the doctrine that consciousness is necessitated not by the structural properties of the physics of the actual world, but by the intrinsic nature of the basic physical properties. It is not clear whether Russellian monism qualifies as a kind of physicalism. In any case, it is not generally accepted by physicalists, so to the extent that physicalists are not Russellian monists, Chalmers considers that his argument against physicalism is successful.

One could remark that the Two-Dimensional framework does not provide a *general* way to go from epistemic premises to metaphysical conclusions. This would require that the following principle be valid :

(LP-MP) If  $\phi$  is verified by some world, then  $\phi$  is satisfied by some world.

But it is clear that this principle is not universally valid. Indeed, since the point of the Two-Dimensional Framework is to distinguish primary and secondary intensions, it follows that there will be many  $\phi$ s verified at some worlds, but satisfied by no world. Consider for example the statement “water is XYZ”.

So the somewhat positive result of the two-dimensional version of the zombie argument is not generalizable to every conceivability argument. It works only because of the particular semantic properties of “P and not Q”.

## 7 Argument 3: Against Modal Monism

There is another way one could dispute the generality of the argumentative strategy developed so far. Any use of the Two-Dimensional Framework relies on the doctrine of modal monism. But Modal Monism is itself disputable.

It has been attacked recently by Anand Vaidya (2008), following an argument by Nathan Salmon (1989) concerning the essentiality of origins. Salmon’s argument is based on the intuition that although an object could not have had an origin totally distinct from its actual origin, it could nevertheless have had an origin slightly different. For example, if a table T originates from a portion P of a tree, then T could have originated from a slightly different portion P’ of the same tree. If this intuition is correct, there should be a portion P” such that

1. T could not have originated from P”
2. if T had originated in P’, then it could have originated in P”



It follows that it is necessary that  $T$  does not originate from  $P''$ , but it is not necessarily necessary that  $T$  does not originate from  $P''$ . For it is possibly possible that  $T$  originates from  $P''$ .

If this is correct, then we have a violation of the S4-principle that if it is necessary that  $\phi$ , then it is necessarily necessary that  $\phi$ . So the logical systems S4 and S5 do not characterize correctly the logic of the metaphysical modality since they license wrong inferences.

Since it is plausible that S5 characterizes correctly the logic of the logical modality, it follows that there is a distinction at the levels of worlds between the logical modality and the metaphysical modality. If all metaphysically possible worlds are logically possible, then given that the accessibility relation for the metaphysical modality is not transitive, and that the accessibility relation for the logical modality is, it follows that there will be logically possible worlds that are not metaphysically possible. So we have a distinction at the level of worlds between the metaphysical and the logical modality.

In order to clarify the picture, the distinction proposed by Salmon between *a way for things to be* and *ways things might have been* can be useful. The first member of the distinction is relative to the notion of world. A world is a maximam way for things to be. The second member of the distinction is related to the notion of metaphysical modality. We have no reason to identify the two notions, unless we have an independent argument to do so. The notion of world does not even imply that worlds should be logically possible. It is a way for things to be that  $\phi$  and not  $\phi$ . So there are logically impossible worlds. When we introduce the notion of a logically possible world, we make a first restriction on set of worlds. And if the last argument is correct, we need to make a further restriction to have the set of metaphysically possible worlds.

If we accept this modal dualism, then the Principle of Plenitude will give only access to logical possibility but not to metaphysical possibility, so that it becomes difficult in general to argue for metaphysical conclusions from epistemic premises.

## 8 Conclusion

There are three graded ways to refuse the modal conclusion of a conceivability argument, while accepting its epistemic premise:

1. argue for a gap between prima facie conceivability and ideal conceivability

2. accept ideal conceivability, but deny (PP) on the grounds that it conflicts with KAPNs
3. accept ideal conceivability, accept (PP\*), but argue for a gap between logical possibility and metaphysical possibility.