

## A Posteriori Necessities

**1. Introduction:** Recall that we distinguished between a priori knowledge and a posteriori knowledge:

**A Priori Knowledge:** Knowledge acquirable prior to experience; for instance, by merely considering the concepts and the relations between them (e.g., all bachelors are male). The denial of a priori truths result in a logical contradiction.

**A Posteriori Knowledge:** Knowledge acquirable only by experience (e.g., the Sun will rise tomorrow). The denial of a posteriori truths does NOT result in a logical contradiction.

Surely, there are a priori necessities. For instance, if <All bachelors are male> is true, then it is NECESSARILY true. That is, there is no possible world in which that proposition is false. In fact, plausibly, ALL of the a priori truths are also necessary truths.

Furthermore, all of the a posteriori truths that we've discussed so far have been contingent. That is, there are some possible worlds where they are true, and also some possible worlds where they are false; e.g., <The Sun will rise tomorrow>, <All ravens are black>, <Chad is a philosopher>, etc. all are possibly true, and possibly false. But, **are there any a posteriori necessities?** That is, are there any truths that are only knowable by observation that COULD NOT have been false? This is the question Kripke tackles. Prior to Kripke's time, many thought that ALL a posteriori truths are contingent, not necessary. Kripke supplies compelling evidence to the contrary.

**2. Hesperus and Phosphorus:** In the evening, a bright heavenly body is visible in the sky. The Ancient Greeks called this body **Hesperus**. There is also a bright heavenly body visible in the morning sky. The Ancient Greeks called this body **Phosphorus**. Much later, astronomers discovered that Hesperus and Phosphorus were really one and the same object; namely, the planet **Venus**. So, it turns out that **Hesperus IS Phosphorus**.

But, now ask: **Could Hesperus NOT have been Phosphorus?**

Note that we are NOT asking, Could it have been the case that there was one heavenly body in the evening sky that people CALLED Hesperus, and a DIFFERENT heavenly body in the morning sky that people CALLED Phosphorus. Certainly that scenario IS possible:

There certainly is a possible world in which a man should have seen a certain star at a certain position in the evening and called it 'Hesperus' and a certain star in the morning and called it 'Phosphorus'; and should have concluded—should have found

out by empirical investigation—that he names two different stars, or two different heavenly bodies. ... And so it's true that given the evidence that someone has antecedent to his empirical investigation, he can be placed in a sense in exactly the same situation, that is a qualitatively identical epistemic situation, and call two heavenly bodies 'Hesperus' and 'Phosphorus', without their being identical.

But, the example above only demonstrates that EPISTEMICALLY speaking, Hesperus and Phosphorus could have been two different objects. That is, given a certain kind of experience (namely, seeing a bright object in the evening, and seeing a bright object in the morning), there are some possible worlds where they turn out to be the same object, and other possible worlds where they turn out to be distinct.

However, Kripke is asking a METAPHYSICAL question. Namely:

**Could THAT object (pointing to Hesperus) have been numerically distinct from THAT object (pointing to Phosphorus)?**

Kripke thinks that the obvious answer to this question is: NO.

*[Note: This should not seem counter-intuitive. It is no different than asking, 'Could Mark Twain have been not identical to Samuel Clemens?' and answering 'No' for the simple reason that THAT object (pointing to Mark Twain) could not have been numerically distinct from THAT object (pointing to Samuel Clemens; i.e., the very same object).]*

But, then, it is not only the case that Hesperus is ACTUALLY Phosphorus (i.e., <Hesperus = Phosphorus> is true in the actual world). Rather, a much stronger claim turns out to be true: NECESSARILY, Hesperus is Phosphorus (i.e., <Hesperus = Phosphorus> is true in EVERY possible world. There is no possible world in which Hesperus is not Phosphorus):

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. But we, using the names as we do right now, can say in advance, that if Hesperus and Phosphorus are one and the same, then in no other possible world can they be different.

Note, though, that <Necessarily, Hesperus = Phosphorus> is not known a priori. Rather, this is something that is only known by EXPERIENCE. It was not until we made the discovery that we were pointing to the same object when we pointed at the evening star and then the morning star that we discovered this necessary truth. Therefore, Kripke concludes, there are **a posteriori necessities**.

**3. Rigid Designation:** Kripke's conclusion follows from his theory of naming. When we coined the name 'Hesperus', we pointed to a particular actual object. Similarly, when we coined the name 'Phosphorus', we pointed to a particular actual object. In short, we DESIGNATED those names as referring to particular objects. But, Kripke's claim is that, if we use a name to designate a particular object in the ACTUAL world, then that name designates that same object in every POSSIBLE world as well (i.e., designation is **rigid**).

**Rigid Designators:** If a name refers to X in the actual world, it refers to X in all possible worlds. Naming something "fixes the reference" of that name.

Once we discovered that 'Hesperus' and 'Phosphorus' are both names that rigidly designate ONE AND THE SAME OBJECT, we discovered that <Hesperus = Phosphorus> is true in the actual world—and therefore it is also true in EVERY possible world. In short, if there is a possible world where Hesperus exists, then that is ALSO a world where Phosphorus exists (as the very same object), and vice versa

*[Note: This is NOT to say that there is no possible world in which we could have CALLED Venus something else. Clearly, it is possible to have several names for one and the same thing. Heck, even in the ACTUAL world, we call the same planet by at least three names (Hesperus, Phosphorus, Venus, etc.). So, imagine the possibility that we could have decided to name Venus 'Flooglepuff'. In that possible world, the object CALLED Flooglepuff is just the same object as the one which 'Hesperus' (and 'Phosphorus' and 'Venus', etc.) rigidly designates in the actual world.]*

*Really, Kripke is just giving us an apparatus which allows us to "grab hold" of an individual and talk about what is possible for that individual. For instance, how should we make sense of the claim that <Chad could have been called 'Larry'>? Well, we consider a possible world where 'Chad' (i.e., the very same person who is writing these notes) exists, but is CALLED 'Larry'. All Kripke is saying is that we must still use the LABEL 'Chad' when considering this possibility—since the name 'Chad' rigidly designates the same object in every possible world—and we are just considering a scenario in which Chad calls himself by another name ('Larry' in this case).]*

If names are rigid designators, then identity statements are necessary, if true. Let  $R_1$  and  $R_2$  be two rigid designators. If  $\langle R_1 = R_2 \rangle$  is true, then  $\langle \text{Necessarily, } R_1 = R_2 \rangle$  is also true. This is not counter-intuitive at all. For instance, if I point to a guy and say, that is 'Mark Twain' ( $R_1$ ), and then I point to the SAME guy and say, that is 'Samuel Clemens' ( $R_2$ ), it turns out that  $\langle R_1 = R_2 \rangle$  is true—because they're the SAME THING. But, then,  $\langle \text{that guy} = \text{that guy} \rangle$  isn't just true in THIS world. It is NECESSARILY true. It could not be false.

Contrast this with NON-rigidly designating labels: If names were merely descriptions, then identity statements are contingent (that is, they could be true or false). Descriptions are NON-rigid, because the individuals that can fill that description could differ. For instance, let  $D_1$  and  $D_2$  be two NON-rigid designators. For instance, if I cite two descriptions and say let “the heavenly body that is seen in the evening” be ‘Hesperus’ ( $D_1$ ), and let “the heavenly body that is seen in the morning” be ‘Phosphorus’ ( $D_2$ ), then, although  $\langle D_1 = D_2 \rangle$  is true in the ACTUAL world (that is, in the actual world, those two objects are one and the same planet), this proposition won’t be true in EVERY possible world. For there are possible worlds where there is a heavenly body seen in the evening, and one seen in the morning, and they are NOT one and the same object.

$\langle$ The inventor of the bifocals is identical to the first postmaster general $\rangle$ . It turns out that this is TRUE in the actual world (Ben Franklin fits both of these descriptions). But, surely it is POSSIBLE for this claim to have been false. For instance, surely it is possible that Ben Franklin was the postmaster general, but NOT the inventor of the bifocals (e.g., if someone else had invented them before he did). In that possibility we are considering, the inventor of the bifocals is NOT identical to the first postmaster general.

**4. Water = H<sub>2</sub>O:** Kripke turns to scientific discovery. Originally, we pointed at some clear, tasteless liquid filling the rivers and lakes and used the name ‘water’ to designate it. Much later, scientists discovered that water is composed of H<sub>2</sub>O. In other words, we now know that water JUST IS H<sub>2</sub>O; i.e., we discovered that  $\langle$ Water = H<sub>2</sub>O $\rangle$  is true.

Now ask: **Could water have been composed of anything other than H<sub>2</sub>O?**

Kripke says: NO. For instance, if we discovered some clear, tasteless liquid that looked and felt like water but was NOT composed of H<sub>2</sub>O, we would conclude that we had found “fool’s water”, not water (similar to how pyrite is now considered to be “fool’s gold”, though it was once thought to be gold; or how nephrite is known to be distinct from jadeite, though they were once both thought to be the same thing: [jade](#)).

But, then, if  $\langle$ Water = H<sub>2</sub>O $\rangle$  is true, then it is NECESSARILY true—another a posteriori necessity! Now consider two more:

$\langle$ Light is a stream of photons $\rangle$

$\langle$ Heat is molecular motion $\rangle$

Kripke notes that light need not produce a visual sensation in order to exist. There are blind people, for instance. In their case, we say that they are not sensitive to light. In a world where ALL living things were blind, we would say that light existed, but that none of the creatures could detect it. Similarly with heat. Even if NOTHING could detect heat (e.g., if every sentient creature had something like [CIPA](#)), we would say that things were still HOT, but that none of the creatures could detect this.

So, 'heat' & 'light' designate things out there IN THE WORLD (not the sensations in us). The interesting result: Since heat JUST IS molecular motion and light JUST IS a stream of photons, it turns out that we have discovered two more a posteriori necessities:

<Necessarily, light is a stream of photons> <Necessarily, heat is molecular motion>

*[Note that we sometimes use terms like 'heat' ambiguously, to refer to the internal, mental SENSATION of heat; i.e., what it FEELS like to experience warmth. THAT isn't identical to molecular motion. So, what's this MENTAL thing identical to? Let's turn to that question.]*

**5. Implications for Philosophy of Mind:** Kripke has argued that, when scientists discover identity in nature (water= $H_2O$ , heat=molecular motion, etc.) they discover NECESSARY TRUTHS.

But: There is a recent trend among scientists and philosophers to identify certain mental states with certain brain states. These philosophers say that the mind (or, consciousness) is nothing over and above your body, or your brain. Rather, consciousness just IS (i.e., is IDENTICAL with) brain activity. This is a scientific identity claim. So, Kripke's thesis is importantly relevant here.

Suppose that neuroscientists have claimed that the sensation of pain is just C-fibers firing in the brain (we now know it's more complicated than this, but just suppose).

*[Note: the identity theorist is not just saying the the firing of C-fibers CAUSES pain. Rather, they are saying that **the firing of C-fibers IS pain**. (That's all pain IS! There's nothing more to it. Pain is nothing over and above this neural event.) Pain and the firing of C-fibers are identical in the same way as water and  $H_2O$ , or Sam Clemens and Mark Twain.]*

Let A and B be two rigid designators, which designate the following:

A = The sensation of pain you had when you touched a hot stove just now.

B = The brain state that corresponded to your pain (e.g., some C-fibers firing).

Now, if A=B (that is, if the sensation JUST IS the brain state), then it should turn out to be true that NECESSARILY, A=B. *[Remember, Kripke has established that, if <X=Y> is true, then <Necessarily, X=Y> is also true.]*

But, is that right? Is there a possible world where that brain event occurs (those C-fibers fire in your brain), but where you feel no sensation of pain? Kripke says, intuitively: Yes. Furthermore, Descartes argued (as we saw last time) that it ALSO seems possible for the sensation of pain to have occurred in the absence of that particular neurological event in the brain. Here is an argument:

1. For any two rigid designators,  $R_1$  and  $R_2$ , if  $\langle R_1=R_2 \rangle$ , then  $\langle \text{Necessarily, } R_1=R_2 \rangle$ .
2. Physicalist Assumption:  $\langle \text{Pain}=\text{C-fibers firing} \rangle$  (*where 'pain' rigidly designates the qualitative sensation of pain, and 'C-fibers firing' rigidly designates C-fibers firing*).
3. Therefore,  $\langle \text{Necessarily, Pain}=\text{C-fibers firing} \rangle$ .
4. But,  $\langle \text{Necessarily, pain}=\text{C-fibers firing} \rangle$  is false (*since there is a possible world where pain occurs but C-fibers do not fire, as well as a possible world where C-fibers fire, but pain does not occur; see below*).
5. Thus, pain is not identical to C-fibers firing; i.e., the physicalist assumption is false.

C-fibers Firing Without Pain: The crucial claim is P4. Here's why Kripke believes it is true: Imagine that there is a God, and that God is creating the world. It seems true that, in order to create water, all God would need to do is create  $\text{H}_2\text{O}$ . But, it does NOT seem that the mere creation of C-fibers firing in human brains is enough to create pain. Rather, God would still have some more work to do. He writes:

It would seem, though, that to make the C-fiber stimulation correspond to pain, or be felt as pain, God must do something in addition to the mere creation of the C-fiber stimulation; He must let the creatures feel the C-fiber stimulation as pain, and not as a tickle, or as warmth, or as nothing, as apparently would also have been within His powers.

Similarly, David Chalmers points out that what he calls '**philosophical zombies**' are metaphysically possible; i.e., there are possible worlds where there exist the sort of being that is "atom-for-atom identical to a conscious being such as you and me, but it is not conscious." This is another way of capturing Kripke's claim that, even once we have ALL of the physical facts in place, this alone does not guarantee consciousness. Consciousness is something OVER AND ABOVE the physical facts.

Pain Without C-fibers Firing: There also seem to be metaphysically possible worlds where conscious beings experience pain WITHOUT C-fibers firing. For instance, in one possible world, they experience pain while their XYZ fibers fire. In another possible world, they experience pain without brain activity at all! For, it seems metaphysically possible for there to be disembodied, non-physical souls in pain.

Conclusion: The argument above is generalizable to ANY mental phenomenon (not just pain). So, if Kripke is correct, then the identity thesis is false (i.e., consciousness is NOT identical to brain states, but is rather something OVER AND ABOVE your brain).

Stop here for a second. The power of this conclusion is astounding: If Kripke is right, then your mind is something OVER AND ABOVE your brain. Your consciousness is not MERELY some brain activity. Rather, it is something MORE than that! Whoa...

**6. The Illusion of Contingency (objection to premise 3):** Now, we might wonder how reliable conceivability is as a guide to metaphysical possibility. Most of the a posteriori necessities described above SEEMED contingent—that is, it seemed possible that they be false—but, Kripke explained that this intuition was mistaken as follows:

Consider the identity claim  $\langle \text{water} = \text{H}_2\text{O} \rangle$  again. If it SEEMS as if this could be false, your mistake is due to the fact that you are imagining someone being in exactly the same EPISTEMIC situation as we are. For instance, imagine someone on “Twin Earth”:

**Twin Earth (Water):** Imagine someone on a planet that looks qualitatively identical to Earth. On that planet, there is a clear, tasteless liquid, which fills the rivers and streams. They drink this liquid and even refer to it as ‘water’. However, it turns out that this substance is NOT composed of  $\text{H}_2\text{O}$ . Rather, it is composed of the substance, XYZ.

Supposedly, we think we are imagining a world where water exists in the absence of  $\text{H}_2\text{O}$ . But, when we imagine this supposed “counter-example” to the identity claim,  $\langle \text{Necessarily, water} = \text{H}_2\text{O} \rangle$ , we are making a mistake. This is not water in the absence of  $\text{H}_2\text{O}$ . Rather, it is some WATERY STUFF in the absence of  $\text{H}_2\text{O}$ . We may THINK we can imagine water in the absence of  $\text{H}_2\text{O}$ , but really we cannot. At best, we can imagine some OTHER substance (call it ‘twater’) in the absence of  $\text{H}_2\text{O}$ .

Thus, the apparent contingency of supposedly necessary truths is often illusory. Perhaps, when we say that we can conceive of scenarios where pain is NOT identical to the firing of C-fibers, we are getting confused in the same way. It may SEEM that the identity claim is false, but really it is still true.

**Reply:** Kripke disagrees that the mind-body identity theorist has access to this same strategy. Consider the possible world containing Twin Earth again:

**Twin Earth (Pain):** Imagine someone on a planet that looks qualitatively identical to Earth. On that planet, people feel painful sensations when they stub their toes, or put their hand on hot stoves, or stab each other, etc. However, it turns out that the C-fibers in their brains never fire when they have this sensation. Rather, some other part of the brain fires instead; namely, their XYZ fibers.

Supposedly, we are imagining a world where pain exists in the absence of C-fibers firing. Thus, supposedly we have a counter-example to the identity claim,  $\langle \text{Necessarily, pain} = \text{C-fibers firing} \rangle$  because here is a possible world where this identity does not hold.

The identity theorist must insist that we are **WRONG** about this. The strategy available to them is to insist that we are **NOT** imagining a world where pain exists in the absence of C-fibers firing. Rather, we are imagining a world where there is some **PAIN-Y STUFF** in the absence of C-fibers firing. We may **THINK** we can imagine pain in the absence of C-fibers firing, but really we cannot. At best, we can imagine some **OTHER** phenomenon (call it 'schpain') in the absence of C-fibers firing.

But, this response does not have the oomph that it did when raised against the Twin Earth (Water) case. The problem is that, while we **CAN** have the qualitative experience of interacting with water even in the absence of water, we **CANNOT** have the qualitative experience of being in pain even in the absence of pain. Rather, that experience is **ESSENTIAL** to pain; we cannot have the experience of pain unless we actually are in pain! In short, **not everything that looks like water is water, but everything that feels like pain is pain!**

To drive home the point: Compare the following two scenarios:

**Fool's Water:** Imagine that you are drinking some cool, clear liquid that you got out of a nearby lake, and it is quenching your thirst. Now imagine a chemist comes along and examines your beverage. "Nope," they tell you. "You're not drinking water. For, as it turns out, this substance is not composed of H<sub>2</sub>O at all. Rather, it's XYZ."

**Fool's Pain:** Imagine that you are writhing in bed, screaming and experiencing excruciating agony. Now imagine a neuroscientist comes along and examines your brain. "Nope," they tell you. "You're not in pain at all. For, as it turns out, none of the C-fibers in your brain are firing. Rather, your XYZ-fibers are firing."

You would probably accept the chemist's assessment, but reject the neuroscientist's assessment. "Oh weird. This isn't water!" you'd declare in the first case. In the second case: "Of course I'm in pain you moron! Why do you think I'm screaming!?"

[**Question:** Imagine a physicalist who abandons strict identity theory and instead posits that conscious states are "**multiply realizable**". That is, conscious states (such as pain) can be realized by any of several distinct kinds of brain states (e.g., C-fibers firing or XYZ-fibers firing). Thus, in one possible world, creatures whose C-fibers fire experience pain, and in another possible world, creatures whose XYZ-fibers fire experience pain. Can we make sense of this as a version of *physicalism*?]