BEING SCEPTICAL ABOUT KRIPKEAN A POSTERIORI NECESSITIES AND NATURAL KINDS

The article discusses Saul Kripke’s influential theories of a posteriori necessary truths and natural kinds. With respect to the statements of identity involving proper names, it is argued that although their truth is a posteriori and necessary in the specific sense of counterfactual invariance, this is of no significance for substantial philosophical issues beyond the philosophy of language, because this counterfactual invariance is a trivial consequence of the use of proper names as rigid designators. The case is made that the expansion of the realm of necessary a posteriori truths to the statements of theoretical identity that involve “natural kind terms”, as well as the Kripkean essentialist theory of natural kinds, have no weighty argumentative support and fit badly both with science and language practice. This sets the stage for the development of an appropriately sophisticated “descriptivist” account of meaning and reference that would be better suited for a widened range of Kripke–Putnam style thought experiments. The general outlines of such a descriptivist account are provided.

Keywords: a posteriori necessary truth, necessity, contingency, metaphysical possibility, natural kind.

One of the most celebrated discoveries in the analytic philosophy of the second half of XX century is that of the existence of necessary a posteriori truths, made by Saul Kripke. As Hilary Putnam (the second most influential defender and developer of the theory of necessary a posteriori truths) puts it, “[s]ince Kant there has been a big split between philosophers who thought that all necessary truths were analytic and philosophers who thought that some necessary truths were synthetic a priori. But none of these philosophers thought that a (metaphysically) necessary truth could fail to be a priori” [Putnam, 1975: p. 151]. Kripke was the philosopher who first influentially challenged the view that all necessary truths are a priori and, more-
Over, succeeded (together with Putnam) in convincing the majority of analytic philosophers that there are necessary a posteriori truths, that they cover a wide range of statements, and that this is of great philosophical importance. In particular, statements of identity, such as “Hesperus is Phosphorus is Venus”, “Water is H₂O”, “Heat is mean kinetic energy”, are claimed to be necessary a posteriori truths.

The Kripkean view became widely accepted and often uncritically taken for granted, sort of orthodoxy in the analytic philosophy after 1970s. The appeal to a posteriori necessities plays a considerable role in the discussions in a number of substantial areas, such as philosophy of mind (the issue of mind-brain identity) and philosophy of science. On the other hand, criticisms are not lacking, although not influential enough to overthrow the Kripkean orthodoxy. A number of philosophers argued that Kripke’s argument involve grave faults, such as begging the question, unjustified essentialist assumptions [Salmon, 1979; Rosenberg, 1994]¹, treating words as if their meanings are rigidly fixed rather that changeable with time (especially, due to scientific discoveries) [LaPorte, 2004; Vasiliev, 2014: pp. 75-80] and contexts. Alongside with such criticisms, ever wider acceptance is gained by an approach that emasculates the idea of a posteriori necessities, the two-dimensional semantics (elaborated by David Chalmers [1996; 2004; 2006], Frank Jackson [1998a; 1998b; 2004], and others), in which perspective, “all necessary a posteriori truths are, in Gareth Evans’s (1979) terminology, ‘superficially necessary’, but ‘deeply contingent’” [Mackie, 2002: p. 225].

An important part of Kripke’s theory is the semantic theory of natural kind terms based on the essentialist ontological theory of natural kinds. It is this theory that grounds the claim that statements of theoretical identities (“Water = H₂O”, “Heat = mean kinetic energy”, etc.) are necessary a posteriori truths, and without this claim, the discovery of necessary a posteriori truths would lose most of its apparent philosophical significance.

The essentialist theory of natural kinds is that there is sort of unique genuine division of things into kinds that is determined by their true natures, or essences. According to this theory, which comes back to Plato and Aristotle, things have something called “essence”, and things belong to the same natural kind in virtue of their having the same essence. Of course, there can be a lot of “artificial” classifications based on arbitrary criteria, so that the same thing can be assigned to a number of different “artificial” kinds, but only one of these divisions is faithful to the true nature of things and (using, again, Plato’s metaphor) “carves nature at its joints”².

Accordingly, Kripke’s semantic theory is that such general terms as “water”, “tiger”,

¹ Cf.: “Kripke’s essentialist modal theses struck me as quirky and implausible, and while his critique of a particular interpretation of “Descriptivism” — that is, the thesis that the referent of a proper name is determined by way of an associated (descriptive) sense — was plainly on target, the specific version of Descriptivism that he had singled out for criticism seemed both deliberately and unnecessarily naïve” [Rosenberg, 1994: p. xi].

² “The second principle is that of division into species according to the natural formation, where the joint is, not breaking any part as a bad carver might” [Plato, 1892: p. 474].
“heat”, etc. designate classes of entities (natural kinds) that are individuated by the essences of these entities, and these essences are constituted by some deep qualities empirically discoverable by scientists rather than superficial qualities by which we identify water, tigers, etc. in ordinary life. The essentialist theory of natural kinds met acute and comprehensive criticisms in discussions in the philosophy of science, which make its credentials look dubious, to say the least. In particular, essentialist views on natural kinds are usually held on metaphysical grounds and opposed by “authors who are mainly oriented toward actual scientific practice” [Brzović, 2018]; some of the most important examples are [Dupré, 1993], [LaPorte, 2004], [Hacking, 2007], [Kendig and others, 2015], [Ereshefsky, 2017]. However, this does not impede the wide acceptance in the major (“linguistic”) part of analytic philosophy (that is primarily involved with language, meaning, and reference rather than scientific practice) of the Kripkean claim that truths of the statements of theoretical identities that involve “natural kind terms” (which are supposed to be such common words as “water”, “gold”, “tiger”, “heat”) are necessary a posteriori. This situation suggests scrutiny of the considerations advanced by Kripke and Putnam in favour of this claim.  

This article purports to re-evaluate the significance of the Kripkean claim about necessary a posteriori truths of the identity statements that involve proper names, highlight the far-fetchedness of the expansion of this claim to the statements of theoretical identity that involve “natural kind terms”, as well as several grave problems with the Kripkean essentialist theory of natural kinds and the deficiency of its argumentative support. This sets the stage for the development of an appropriately sophisticated account of meaning and reference both for proper names and what is supposed to be natural kind terms that would qualify, in Kripke’s terms, as “descriptivism” but has sufficient resources to counter the criticisms against descriptivism raised by Kripke and his followers and is better suited for a widened range of Kripke-Putnam style thought experiments than the Kripkean account. This article gives general outlines of such a descriptivist account; the details and the confrontation with the Kripkean counterfactual scenarios await for the sequel articles.

1. The Starting Point: Proper Names as Rigid Designators

With Kripke, the claim that there are necessary a posteriori truths arises initially for proper names, such as “Hesperus” or “Aristotle”, and then is extended to so called “natural kind terms”, such as “water”, “gold”, “tiger”, or “heat”. It is important to understand first why and in what sense true statements that involve proper names are, on Kripke’s account, necessarily true. As far as I see it, the clarity about these points should have two unfavourable consequences for the Kripkean view: first, it makes the discovery of necessary a posteriori truths far less exciting because of its

---

3 In one of his latest papers, Putnam [1990] had revised his views and renounced the notion of “metaphysical necessity”, without which there is no place for the notion of a posteriori necessary truths (except physically necessary). However, this didn’t stop the Kripkean bandwagon.
purely linguistic character and irrelevance to any extralinguistic matters; second, it
gives pretty good reasons to be sceptical about the expansion of the claim about
necessary a posteriori truths from proper names to so called “natural kind terms”.

So, let us elucidate the Kripkean case for the claim that such truths as “Hesperus is Phosphorus” are necessary a posteriori. The claim that they are knowable only a posteriori seems unproblematic, at least prima facie⁴: some centuries ago, people did not know that Hesperus, the starlike heavenly body first seen in the evening (“the evening star” for a shortcut), is Phosphorus, the starlike heavenly body last seen in the morning (“the morning star” for a shortcut); and they could not get to know this by means of mere armchair speculation (conceptual analysis); they lacked the empirical knowledge that allowed later to establish the identity. However, why is the truth of this identity statement necessary? Prima facie, it seems contingent: it could turn out that the starlike heavenly body last seen in the morning, called “Hesperus”, and the starlike heavenly body last seen in the morning, called “Phosphorus”, are two different heavenly bodies.

Kripke’s case for the claim that the truth “Hesperus is Phosphorus” is necessary is based on two pillars.

The first pillar can be called the counterfactual modal framework; it is the
framework of modal notions (necessity/contingency, possibility/impossibility) de-
defined in terms of what could have been the case, as distinct from what actually is the
case. Note that in this framework, claims about possibility are often cast in terms of
metaphysically possible worlds, and that most contemporary analytical philosophers,
including Kripke, when talking about possible worlds, mean nothing but “global”
counterfactual possibility — the way the world could have been, as distinct from the
way it actually is. Most contemporary analytical philosophers (including Kripke),
when making a statement in the form “There is a possible world such that …”, do not
mean that such a world somehow really exists; they mean nothing but “The world
could have been such that …”.

The second pillar is the concept of rigid designator and the observation that
proper names are used as rigid designators. A rigid designator is a term that means
the same thing in all possible worlds. That is, if “N” is a rigid designator, then
whenever we talk counterfactually about N (using the term “N”) — what could
have been the case with N, as distinct from what is really the case with N — we refer
to the same thing to which we refer when we talk about N as it actually is. Let us
take, for example, the name “Aristotle”. For most of us, it is mostly associated with
the idea of a great ancient Greek philosopher. However, it would be mistaken to
think that the name “Aristotle” just means “the great ancient Greek philosopher …”
(where “…” is to be replaced with some list of further properties required to iden-
tify Aristotle uniquely, given that there were other great ancient Greek philosophers).
It makes perfect sense to say that in some other circumstances (counterfactually),

⁴In fact, a number of philosophers [Fitch, 1976; Salmon, 1986: pp. 133–142; Soames, 2002] argued that this claim is mistaken and that the truth of the proposition “Hesperus is Phosphorus” is knowable a priori.
Dmytro SEPETYI

Aristotle could have been not a great philosopher (or not a philosopher at all). And it seems that for any description associated with the name “Aristotle”, it could have been that Aristotle would not fit this description if some circumstances of his life were sufficiently different from what they really were. The name “Aristotle” is not synonymous with any description (in particular, any definite description that identifies Aristotle uniquely); it functions differently in language, — in particular, in counterfactual talk. Name “Aristotle” is a rigid designator — it refers to the same person in all possible worlds, whether or not this person satisfies this or that description in these possible worlds.

I think that the claim that proper names function in natural language as rigid designators is correct, with the reservation “usually” ⁵. (Another, more important reservation is that, pace Kripke, this is consistent with appropriately sophisticated “descriptivism” about proper names.) If so, it trivially follows that such identity statements as “Hesperus is Phosphorus”, if true in the actual world, are true in all possible worlds. And that is just what “necessarily true” means in the counterfactual modal framework. This does not reveal any substantial truth beyond the mere non-modal fact that (in the actual world) Hesperus is Phosphorus. It is merely a matter of the way we use words when we talk about what could have been the case, as distinct from what is really the case. Given that in fact (in the actual world) Hesperus is Phosphorus, whenever we talk about what could have been the case, using the same names, we talk about the same heavenly body, which cannot, of course, fail to be itself. Hesperus could have been not the first starlike heavenly body seen in the evening (there are possible worlds in which Hesperus is not the evening star), and Phosphorus could have been not the last starlike heavenly body seen in the morning (there are possible worlds in which Phosphorus is not the morning star); however, given that Hesperus is Phosphorus, whenever we talk about this heavenly body counterfactually, both names are used to refer to this heavenly body.

The purely linguistic nature of Kripkean a posteriori necessities is clear from the fact, explicitly admitted by Kripke, that the truth that the evening star is the morning star is not necessary but contingent: it could have been that the first starlike

⁵ They do not always function that way. It is clear from the example advanced by Paul Ziff:

“Is it true that no one other than Nixon might have been Nixon? What’s so unusual about him? Perhaps no one other than Fidel might have been Fidel. But Nixon has no special qualities. Is what I am saying intelligible? Of course it is. Consider such remarks as these: ‘Do you know what made Hitler such a monster? He was dropped on his head as a baby. Hitler wouldn’t have been Hitler if that hadn’t happened …’, ‘Hilbert wouldn’t have been Hilbert but for an altogether fortuitous combination of genes effected by Otto and Maria Hilbert’. And so ‘Nixon might not have been Nixon had someone given him adequate moral instruction in his youth. He would have been a different person’. Contrary to Kripke’s thesis that names are always rigid designators any given name is a nonrigid designator on Kripke’s ‘intuitive test’” [Ziff, 1977: p. 327].

On the other hand, as Putnam [1990: p. 58] points out, the definite descriptions that are not usually used as rigid designators, can sometimes be used so — as, for example, “the president” in the statement (implying the person of the present-day president) “The president would never have become president if ...”.
heavenly body seen in the evening was not the last starlike heavenly body seen in the morning (there are possible worlds in which the evening star is not the same heavenly body as the morning star). It is just incorrect (according to Kripke) to describe such counterfactually possible situation (possible worlds) as such in which Hesperus is Phosphorus: given that Hesperus and Phosphorus is the same heavenly body and that “Hesperus” and “Phosphorus” are used as rigid designators (both refer to the same thing in the talk about all counterfactually possible scenarios, or in all “possible worlds”), if (counterfactually) the first starlike heavenly body seen in the evening was not the last starlike heavenly body seen in the morning, then either the evening star, or the morning star, or both, would not be that heavenly body to which we refer by names “Hesperus” and “Phosphorus”.

It is also important to see that the sense in which the statement “Hesperus is Phosphorus” is necessarily true is such that it has nothing to do with the informative content or cognitive significance of this statement, with what one gets to know by getting to know that Hesperus is Phosphorus. Imagine a person who knows about there being the evening star called “Hesperus” and the morning star called “Phosphorus” but who does not know that Hesperus is Phosphorus (as it was with all people before the identity was empirically discovered). Now, she gets informed that Hesperus is Phosphorus. What is the new knowledge she has acquired? What she knows now that she didn’t know before? It is clear that the new knowledge is that the evening star is the morning star. However, this (as Kripke admits) is contingent, not necessary— it could have been that the evening star was not the morning star (there are possible worlds in which the evening star is not the morning star). So, as far as the informative content of the statement “Hesperus is Phosphorus” is concerned, its truth (the fact that it is true) is contingent after all!

Suppose we agree with Kripke and Putnam (as most contemporary analytic philosophers of language do) that theoretical identities cast in “natural kind terms” (such as “Water is H2O”, “Heat is the mean kinetic energy”) are also a posteriori necessary truths. In this case, we should say the same about this necessity: it has nothing to do with the necessity of what we get to know by getting to know that water is H2O, or that heat is the mean kinetic energy. Consider the water-H2O identity, for example. Recollect yourself at times when you, while having a pretty good idea of water (the liquid stuff that is, if not too polluted, transparent and, if not for admixtures, drinkable and tasteless, like that in your surrounding that flows from faucets, fills rivers, falls from the sky when it rains, — let us convene to use “the watery stuff” as a shortcut for this definite description), were ignorant of chemistry. Then you start studying chemistry, and learn that there are the atoms of hydrogen, H, and the atoms of oxygen, O, and that they combine into the molecules H2O. This is still not enough to know that water is H2O; it is not knowable a priori that water is H2O, even if you are perfectly competent with the ordinary concept of water and the chemical concept of the H2O molecules; an empirical research is required to find out that the watery stuff available for research consists, in by far the largest part, of the H2O molecules. Now you get informed that water is H2O.
(or, more precisely, that water consists of the H$_2$O molecules). Or better yet, imagine yourself in the role of the chemist who, as a result of her experiments, has first discovered that water consists of the H$_2$O molecules. What you know now that you didn’t know before? It is clear that the new knowledge is that the watery stuff consists of the H$_2$O molecules. However, this (as Kripke admits) is *contingent, not necessary* — it could have been that the watery stuff was something else, not H$_2$O. (That is exactly why it is not knowable *a priori* that water is H$_2$O.) So, as far as the informative content of the statement “Water is H$_2$O” is concerned, its truth (the fact that it is true) is contingent after all!

The lesson that we should draw from this is that there are two distinct senses of “necessity”, “contingency”, and “possibility” of the truth of statements — one that is relevant to their informative content, or cognitive significance, but not to their counterfactual truth-values, and another that has to do with the counterfactual truth-values of statements but has nothing to do with their informative content, or cognitive significance. Let us call the former 1-necessity, 1-contingency, 1-possibility, and the latter 2-necessity, 2-contingency, 2-possibility. This outcome is in line with the post-Kripkean neo-Fregean semantic theory called “generalised two-dimensional semantics”, which accommodates “Kripkean intuitions” about *a posteriori* necessities to basically Fregean framework.

In fact, contemporary analytic philosophers usually recognise this distinction in a way, by distinguishing *epistemic possibility* and *metaphysical possibility*. However, formulating the distinction in these terms is often misleading, because the adjective “metaphysical” is likely to raise expectations of relating to truths about reality, its deeper ontological aspect going beyond the purview of natural science. Among the followers of Kripke, there is a marked tendency to treat Kripkean “metaphysical possibility” as *real* possibility, and treat epistemic possibility as suspect in the respect of its relevance to ontology. However, the preceding discussion suggests that

---


7 This point was admitted by Hilary Putnam in one of his latest articles, in which he, reacting to A.J. Ayer’s [1983: pp. 265-270] criticism of his and Kripke’s theories, has re-explained his views so as to distance himself from Kripke’s essentialism: “In spite of the term ‘metaphysical possibility’, no real metaphysics is involved over and above what was already involved in taking physical possibility to be an objective notion” [Putnam, 1990: p. 61].

8 Jens Kipper’s formulation of one of the objections against two-dimensionalism is characteristi- cistic. The objection is that although there are *epistemic scenarios* that correspond to 1-possibility (in more standard terms of two-dimensionalism, “the scenarios in the first dimension, i.e. those which constitute the primary intension”), — for example, the scenario in which water is not H$_2$O but something else, say XYZ, — calling such scenarios “worlds” is misleading: “The lesson we should have learned from Putnam and Kripke is precisely that there are no worlds where water is XYZ”. (“There are no worlds” here implies possible worlds.) On this view, we should “construe the worlds in the first dimension as purely epistemic possibilities, bearing no deeper connection with metaphysical possibility”. However, “the aim of conceptual analy-
Being Sceptical about Kripkean A Posteriori Necessities and Natural Kinds

quite the contrary should be the case: if we are interested in truths about the real (actual) world, about what is — or what can be — really the case (rather then in how to talk correctly about what is not but could have been the case), then the relevant possibilities are those going under the name “epistemic” rather than counterfactual possibilities going under the name “metaphysical”.

We should note, however, that — as the generalised two-dimensionalist analysis advanced by Jackson, Chalmers, and Kipper shows — the difference between the two kinds of possibilities (and, accordingly, necessities) is itself merely apparent because for each epistemic possibility (whatever cannot be excluded a priori, by means of logics and conceptual analysis), there is a counterfactual (“metaphysical”) possibility, — and it is essentially the same possibility described differently depending on whether we consider it as actual, what can be the case (then it qualifies as epistemic), or as counterfactual, what could have been (but is not) the case (then it qualifies as metaphysical).

Take, for example, the possibility that the evening star is not the morning star. If we consider it as actual (suppose we don’t know that if fact, the evening star is the morning star), then it is to be rightly described as “Hesperus is not Phosphorus” (if it is not really the case that the evening star is the morning star, then Hesperus is not Phosphorus). If we consider it as counterfactual (given that in fact, in the actual world, the evening star is the morning star), then we should describe that possibility as one in which the evening star is not the morning star but not as one in which Hesperus is not Phosphorus (in that scenario, either Hesperus is not the evening star or Phosphorus is not the morning star).

If Kripke’s and Putnam’s claim that “Water is H2O” is a posteriori necessary truth is correct, then the same scheme (one possibility — two ways of its considering and two descriptions) applies for it. If, perchance, the watery stuff in our surrounding we are used to call “water” is not H20 (suppose we were deceived or scientists got systematically wrong), then this possibility (considered as actual) is to be described by the statement “Water is not H20”. However, when we consider the same possibility (of the watery stuff in our surrounding being not H20) counterfactually, then (given that in fact, in the actual world, water is H20), we should describe this possibility by the statement “The watery stuff in our surrounding could have been not H20” or “The watery stuff in our surrounding could have been not water” but not by the statement “Water could have been not H20”.

The outcome is, as David Chalmers [Chalmers 2010: p. 190] suggests, “a sort of modal monism, with a single primitive space of worlds, along with two sorts of evaluation of sentences over this space” (one, according to primary, Fregean intensions

sis is to gain insight into what is really, i.e. metaphysically, possible or necessary. If for instance conceptual analysis can only reveal that it is epistemically possible that water is not XYZ (which is just another way of saying that it is not a priori that water is not XYZ), then this hardly teaches us anything about water...” [Kipper, 2013: p. 26].

In this formulation, the following equation is question-beggingly taken for granted:

real possibility = metaphysical possibility = Kripkean counterfactual possibility.
of the terms involved, when we consider a possible world as actual, and another, according to secondary, Kripkean intensions, when we consider a possible world as counterfactual). This "suffices to account for all the modal phenomena that we have reason to believe in and that we might invoke possible worlds to explain. ... metaphysical necessity is simply subjunctive necessity over the space of logically possible worlds" [Chalmers, 2010: p. 187]. Kripkean "metaphysical possibility of statements is logical possibility with an a posteriori semantic twist" [Chalmers, 1996: p. 38]; “rather than ruling conceivable worlds 9 impossible, a posteriori constraints simply cause worlds to be redescribed” [Chalmers, 1997: p. 14].

Note also that from the evolutionary viewpoint the fact that human beings have developed the capability of counterfactual thinking is hardly understandable except as the by-effect of their developing the capability of thinking about actual possibilities, as a projection of the latter capability. It is our capability to think of what can be the case (given what we know) — rather than our capability to think of what could have been the case — that was evolutionary advantageous. Counterfactual thinking is understandable as the retrospective application of our capability to think of actual possibilities.

The general outcome of this discussion with respect to the statements of identity involving proper names on both sides (such as “Hesperus is Phosphorus”) is that although their truth is a posteriori and necessary in the specific sense of counterfactual invariance (their actual truth-value is retained whenever we talk of what could have been the case, given what is actually the case), this is of no significance for substantial philosophical issues beyond the philosophy of language, because this counterfactual invariance is a trivial consequence of the use of proper names as rigid designators. With respect to other purported Kripkean a posteriori necessities, such as “Water is H₂O”, “Heat is the mean kinetic energy”, etc., the situation would be similar (and manageable by means of the general two-dimensionalist account), if there were no strong disanalogy between such statements and statements like “Hesperus is Phosphorus”, as well as other reasons not to accept the Kripkean claim that the true identity statements involving natural kind terms are necessary a posteriori truths.

2. Natural kinds, essences, and theoretical identities

So far, we discussed the cases of theoretical identities involving general terms (such as “water”, “heat”, etc.) rather than proper names with the proviso “if the Kripkean claim that such identity statements, if true, are necessarily true is correct”. We concluded that if that is the case, then the generalised two-dimensionalist account of such necessary truths should be correct. By far the largest majority of contemporary analytical philosophers, including the most prominent two-dimensionalists, accept the Kripkean claim at issue. In what follows, I am going to challenge it.

The first point to be noted here is that what makes identities involving proper names counterfactually necessary does not hold for identities involving so called “natural kind terms”, such as “water”, “tiger”, “heat”, etc. Kripke and his followers

9 i.e., maximal scenarios that correspond to epistemic possibilities.
obscure this point by claiming that natural kind terms, like proper names, are *rigid designators*. However, how can that be? In the case of proper names, a rigid designator is the term that designates the same thing in all possible worlds (taken as counterfactual), or when we talk counterfactually about what could have been the case. “Hesperus” means in such talk the same heavenly body, whatever counterfactual possibilities we consider (including all those in which Hesperus is not the evening star). In more technical terms, the *extension* of “Hesperus” in all possible worlds (what “Hesperus” picks out, or refers to, in those worlds) is the same heavenly body. Now consider some paradigmatic “natural kind terms”. Let us begin with terms for countable things, such as “tiger”. The extension of “tiger” in the actual world are all tigers that actually exist. However, it could have been that more tigers existed; there are possible worlds in which there are tigers that do not exist in the actual world. So, the extension of “tiger” in these possible worlds is not the same as in the actual world; it includes extra tigers. “Tiger” is not a rigid designator in the same sense in which “Hesperus” or “Aristotle” is a rigid designator, because it refers to different (only partially interlapping) sets of items in different possible worlds. With “water”, the situation is a bit more complicated, because it is an uncountable noun. However, there certainly could have been more water in the world than there actually is; there could have been some extra rivers or seas, or extra H$_2$O molecules, etc. So, the extension of “water” is not the same in all possible worlds; in some possible worlds it picks out more stuff than in some others. Hence, “water” is not a rigid designator in the same sense in which proper names are rigid designators.

The Kripkean expansion of the claim about necessary *a posteriori* truths from proper names to so called natural kind terms is based on the equivocation with the term “rigid designator”. In the case of proper names, a rigid designator is a term that refers to the *same concrete object* in all counterfactual scenarios. In the case of “natural kind terms” it is taken to be a term that refers not to the *same concrete object* or the *same set of concrete objects*, but to the *same kind of objects*, where the sameness of the kind is determined according to a conception that involves an opaque *actual-world-indexed reference to some properties*, — that is, indirect reference to these properties as such that actually have certain second-order properties (Kripke’s “properties that are essential for the largest part of the portions of stuff to which the name ‘water’ was historically applied, whatever these properties are”, “properties that are essential for most animals to which the name ‘tiger’ was historically applied, whatever these properties are”, Putnam’s “chemical composition of water, whatever it is”, etc.). Because which properties $P_1$ do have these second-order properties $P_2$ is a contingent matter discoverable only *a posteriori* (empirically), the truths about natural kinds are also *a posteriori*. However, because the conception of a natural kind $K$ (is supposed to be such that) involves the reference to those properties $P_{1a}$ that *actually* have second-order properties $P_2$ (*actual-world-indexed reference*), in the counterfactual talk about possible scenarios the natural kind term “$K$” continues to refer to the objects in that scenario that have properties $P_{1a}$ rather than to the objects that have properties $P_{1p}$ that have second-order properties $P_2$ in those
scenarios (possible worlds); this ensures that if the statement “All Ks have properties \( P_{1a} \)” is in fact (actually) true, then this statement and all the statements that logically follow from it are true in the counterfactual talk about any possible scenario in which Ks exist, that is, they are counterfactually (“metaphysically”) necessary truths.

Kripke holds that natural kind terms designate distinct kinds of things that are things of that kind in virtue of their distinctive nature, or essence. This is basically Platonic-Aristotelian view that things are what they are (that is, belong to a certain kind) in virtue of sharing something called “Form” (εἶδος), or “essence”, or “nature” (this essentialist “nature” should not be confused with “nature” of naturalism). Those essences are taken to be real entities (universals) that exist independently of our minds and linguistic conventions. This view was dominant in the philosophy of ancient Greece and Middle Ages but was losing its influence from the 17th century\(^\text{10}\) to the moment of Kripke’s appearance on the philosophical scene, as an outdated view that is clearly at odds with the scientific outlook. Kripke had managed to revive essentialism and impart it with the appearance of scientific respectability by identifying “essences” with deep properties that are discovered by science, usually as a matter of “theoretical identities”. (For example, the scientific discovery that water consists of the \( \text{H}_2\text{O} \) molecules is assimilated into the claim about the essence of water — that water is essentially the stuff composed of the \( \text{H}_2\text{O} \) molecules. Accordingly, any stuff that is not composed of the \( \text{H}_2\text{O} \) molecules is claimed to be not water, even if it has all those properties — transparent, liquid, tasteless, drinkable and thirst-quenching, etc. — by which we ordinarily identify water.) The essences are taken to be what grounds natural — rather than conventional — division of things into kinds. Although we can divide things into kinds conventionally, by any arbitrarily selected properties, or sets of properties, or superficial similarities, there is sort of privileged division — the only “natural” division — one that, as Plato said, “carves nature at its joints”. “Natural kind terms”, such as “water”, “tiger”, “heat” are taken to function, in natural language, in accordance with this essentialist doctrine.

However, there are grave troubles with this view and its fit both with science and language practice.

To begin with, although “natural kind” is a special philosophical term, there is nothing even remotely approximating to consensus among philosophers about the way to draw the line between natural kinds and other kinds\(^\text{11}\). There is no tenable

---

\(^{10}\) The first influential attack on it was made by Descartes, who disparaged the notion of “substantial forms” as explanatorily useless and argued that everything in nature, except human minds, is to be explained in terms of “the attribute of extension”, that is, in terms of spatial relations and their changes with time.

\(^{11}\) Some other important kinds are functional kinds. Functional kinds (wings, computers, etc.) are distinguished in terms of roles in the network of relationships, and allow for multiple realisability (for example, wings, computers, etc. can be made of different stuffs, have different forms, structures etc.). Consider the question: why should we take “water” to refer to a natural kind identifiable in terms of a scientifically discoverable hidden essence rather than a functional kind that refers to whatever has those properties by which we usually identify water, insofar as they are significant for us, language-users, “play roles” in our lives?
account of what there is in common between all those Xs that are usually taken (in particular, by Kripke) to be natural kinds, as opposed to all those Ys that do not qualify as natural kinds. In the article with a telling title, “Natural Kinds: Rosy Dawn, Scholastic Twilight”, Ian Hacking draws attention to the heterogeneity of what is usually supposed to be natural kinds (standard examples from the philosophical discussions of the 1970s include tigers, lemons, water, gold, multiple sclerosis, atoms, heat and the colour yellow) and argues that “there is neither a well-defined class, nor any useful vague class, that collects together these heterogeneous examples in the ways that philosophers have hoped for” [Hacking, 2007: p. 206]: “…how could there be a class that fruitfully collects together such a wonderful array of interesting kinds of ... — kinds of what, anyway? Tiger names a kind of animal, lemon a kind of fruit, and also, in a related sense, a kind of tree. Gold names an element, and in another sense a metal, a substance that comes in lumps, dust and flakes. Beyond the animals, vegetables, and minerals, the standard examples do not seem to be kinds of anything. Of what is heat a kind? The sheer heterogeneity of the paradigms for natural kinds invites scepticism” [Hacking, 2007: pp. 206-207].

Moving from standard examples to theories of natural kinds, we can observe the “proliferation of incompatible views”, “a slew of distinct analyses directed at unrelated projects” [Hacking, 2007: p. 203]. Hacking argues that “there are so many radically incompatible theories of natural kinds now in circulation that the concept itself has self-destructed” [Hacking, 2007: p. 205].

Since 1970s, the essentialist theory of natural kinds was criticized in specialized philosophical discussions about natural kinds and scientific classifications, as inadequate to the practice of science and up-to-date scientific knowledge. The idea that things have essences and that nature is neatly divided into kinds according to these essences finds no scientific support and no scientific use. Science has no business with essences. It has to do with properties and relations of the objects it studies and with regularities, or laws of nature, that connect such properties and relations. Depending on properties and relations we are interested in, things can be classified in different ways, so that the same thing can belong to a number of different kinds; and all these kinds are real or “natural”, insofar as they are distinguished according to real properties to be found in nature. Such cross-cutting classifications abound in science (in particular, in chemistry and biology), and there is no reason to think that this is due to the imperfection of present-day science; that in perfect science

12 Within scientific realism, major alternative views on natural kinds are cluster kind approaches and promiscuous realism. According to the former, natural kinds are open sets of items that “share some subset of properties that tend to cluster together due to some underlying common causes” [Brzović, 2018]. The latter proposes to use “natural kind” as a convenient term to designate “a set of objects that are found to have a substantial number of shared properties” [Dupré, 1993: p. 83]. The main point of promiscuous realism, or pluralism [Dupré, 1981; 1993; 1996; 2001; Kitcher, 1984; Ereshefsky, 1992; 1998; Ereshefsky and Reydon, 2015], is that “there are countless legitimate, objectively grounded ways of classifying objects in the world. And these may often cross-classify one another in indefinitely complex ways” [Dupré, 1993: p. 18].
this would have no place. In other words, there is no reason to think that there is one unique way in which “nature is carved at its joints”.

Characteristically, Zdenka Brzović points out that “[b]iological species …, which were taken as standard examples of natural kinds, do not fulfill essentialist requirements” [Brzović, 2018], and Mark Ereshefsky writes of “the death of essentialism” in biology [Ereshefsky, 2017]. John Dupré points out, with respect to the conception of a natural kind as consisting “of all and only those things that share some essential property”, that “[i]t is widely agreed among biologists that no such property can be found to demarcate species, so that if an essential property is necessary for a natural kind, species are not natural kinds” [Dupré, 1993: p. 53]. More generally, “[t]here is no God-given, unique way to classify the innumerable and diverse products of the evolutionary process. There are many plausible and defensible ways of doing so, and the best way of doing so will depend on both the purposes of classification and the peculiarities of the organisms in question…” [Dupré, 1993: p. 57]. Likewise, Ian Hacking notes that “[b]iological species have long served as paradigms of natural kinds — they are (one feels) natural kinds if anything is” [Hacking, 2007: p. 234], and explains that the contemporary scientific views are not favourable for the concept of natural kinds even in biology:

“When Kripke first published, current scientific folklore held that in the next few years molecular biology would discover in the DNA of a species the necessary and sufficient conditions for being of that species. Hence one blithely talked of the essence of tigers and lemons. We would move from the phenotype to the genotype and on to the essence. Within twenty years the folklore had been superseded. There are no necessary and sufficient conditions for being a tiger or a lemon. No essence” [Hacking, 2007: pp. 233-234].

Generally, Hacking argues that “[a]lthough one may judge some classifications as more natural than others, there is neither a precise nor a vague class of classifications that may usefully be called the class of natural kinds. … In short, … there is no such thing as a natural kind” [Hacking, 2007: pp. 238-239]. At best, “natural kind” can be used as a loose “handy tag” that implies only “that in this or that context, some kinds are more — or less — deeply rooted than others”, and does not imply the existence of a distinct class of natural kinds.

It can be objected that this scepticism about natural kinds does not directly applies to Kripke’s and Putnam’s theories at issue, because these theories are not about the world (whether or not there is a distinct class of natural kinds) but about language — that there is (in natural language) a distinct class of general terms that function in the way Kripke and Putnam suggest. The explanation of such functioning may be that human beings are “inborn essentialists” (even if this essentialism is mistaken and not favoured by science) — that we have sort of innate predilection to essentialism, and this predilection played a crucial role in the historical formation of natural language and its individual acquisition with respect to such words as “water”, “tiger”, “heat”, etc. 13.

13 Ian Hacking mentions a psychological theory — developed and defended by Frank Keil [1979; 1989], Scott Atran [1990; 1998], and others — along these lines. According to this theory, human beings have innate mental modules, one of which is the natural-kind module that “en-
However, we should consider carefully the question: how was it established (or what reasons do we have to believe) that such words as “water”, “gold”, “tiger”, “heat” function in the way Kripke and Putnam suggest? The answer is that this is taken (by Kripke and his followers) to be established by mere consulting intuitions about how we should describe certain counterfactual scenario. Kripke and Putnam just propose to consider some such scenario and suggest that we should describe them the way that fits their theory rather that the way that fits the theory they claim to refute — the traditional view dubbed by Kripke as “descriptivism”.

“Descriptivism” is, roughly, the view that such general terms as “water”, “gold”, “tiger”, “heat”, etc. pick out their referents by means of their meanings (senses), of which we are normally aware (we know what we mean by “water”, “gold”, “tiger”, “heat”, etc.) and which can, in principle, be formulated as descriptions. For example, on the descriptivist account, “water” is sort of a shortcut for a description like “the liquid transparent drinkable stuff like that filling our rivers, lakes, etc.”. Kripke objects against this that water is essentially H2O and so if we conceive of a possible world in which the liquid transparent drinkable stuff that fills rivers, lakes, etc. is not H2O but, say, XYZ, then it would be incorrect to describe that world as a world in which water is XYZ; this world should be described as one in which the watery stuff is not water. Why so? There is no non-circular argument for this claim. As for circularity, Kripke’s explanations often look as if in order to support his claims about the correct description D of a counterfactual possibility P he appeals to Essential properties of the objects involved, whereas the only support he provides for the claim

ables children at an early age to begin to classify, to generalize over classes, and to pick up common names for the classes”. Besides, “[t]here is the additional thesis that children act as if the classes for which they are innately primed have essences”. It is also suggested that “Aristotle’s accounts of living things pretty well reflects what he calls folk-biological concepts, which are universal in the human race” [Hacking, 2007: pp. 232-233]. However, if our “innate essentialism” is like that of Aristotle, then it is not suitable to support the Kripkean view: Plato’s and Aristotle’s and Medieval philosophy’s essences have functional character; they are not hidden properties discoverable by the methods of empirical science; rather, we can discern them by our “mind’s eyes” (“armchair” speculation), if we carefully think through our familiar experiences and circumstances.

There is another reason why our “innate essentialism” and mental “natural-kind module” (even if there are such things) would not provide good reasons to accept the Kripkean view. As Jens Kipper points out, even if “it is a part of our ‘folk theory’ that water is a natural kind”, it is hardly the case that this theory is an inalienable part of our meaning of “water”:

“If it had turned out, say, that what flows in our rivers and lakes etc. is actually a mixture of many different chemical substances, would it really have turned out that there is no water? Hardly, I think. People once surely believed that air is a natural kind — like water, it was even considered as an element from Antiquity to early Modern history. But when it turned out that what we breathe is not at all a natural kind, but rather a mixture of many different gases, this did not imply that there is no air” [Kipper, 2012: p. 77].

14 Because the word “essential” is often used in a loose sense nearly synonymous with “important”, I will use “Essential” with the capital “E” for the stronger, essentialist sense. In this stronger sense, “Essential properties” are properties that belong to the essence of a thing, properties without which the thing would not be the thing it is.
that these properties are Essential is the invitation to conceive the counterfactual possibility P and the suggestion that its correct description should be D. Jay Rosenberg [Rosenberg, 1994: pp. 6-27] discusses a number of Kripke’s claims and what can be taken as arguments for them, and arrives at the conclusion: “...each of Kripke’s specific unconditional essentialist necessity claims depends upon an intuitive general conditional essentialist principle. In each case, the principle is instantiated in such a way that its antecedent becomes an empirical, a posteriori, truth, and its consequent is then detached by modus ponens as the desired unconditional claim of ‘metaphysical necessity’. If we search for textual or argumentative support for these general conditional essentialist principles themselves, however, we will be disappointed. In each instance, all we will find is that Kripke simply appeals to his prior convictions regarding such matters, to ‘how it seems to him’, or, to invoke a more fashionable phrase, to his ‘modal intuitions’” [Rosenberg, 1994: p. 22].

In other words, “the Kripke-Putnam arguments for essentialism are, in effect, direct appeals to intuition” [Bird and Tobin, 2017]. Are those intuitions universally shared, compelling, and unequivocal? By no means so. It seems that if we are interested in how certain terms function in natural language, and this functioning is a matter of communication between all sorts of people, we should have consulted the intuitions of a numerous representative sample, and we should have been careful that the counterfactual situations proposed to the consideration of the respondents were not selected and formulated in a way that favours some intuitions rather than counter-intuitions. Of course, nothing like that was done. Instead, there are Kripke’s and Putnam’s own intuitions (hardly independent of their philosophical views) and the fact that they have found the approval of the dominant part of analytic philosophers. The reasons of this dominant approval are, again, hardly just a matter of linguistic competence; rather, theoretical commitments of academic analytical philosophers and their prospects of professional advancement were also involved.\textsuperscript{15}

\textsuperscript{15} In particular, the commitment of the majority of contemporary analytic philosophers to materialism seems to play a considerable role in the wide acceptance of Kripke’s-Putnam’s view. First, the apparently irreducible intrinsic intentionality of thought (represented by Fregean senses) conflicts with materialism, and this provides a strong motivation for the development of causal and externalist theories of reference (congenial with Kripke’s and Putnam’s theories), which strive to provide the account of reference in terms of causal chains and linguistic behaviour and to do away with “meanings in the head”. (Putnam’s famous slogan “Meanings are not in the head” was well-timed.) Second, the notion of necessary a posteriori truths is used by a number of materialist philosophers as a defensive tool against the anti-materialist arguments (such as the zombie argument, or the knowledge argument) that proceed from the apparent and seemingly unbridgeable gap from the physical to consciousness (the explanatory gap, the conceivability of phenomenal zombies, etc.) to the falsity of materialism. (These philosophers use the notion of necessary a posteriori truths to establish the gap between the epistemic and the metaphysical realms, and use that gap to neutralize the explanatory gap from the physical to consciousness by relegating this gap to the epistemic realm.)

\textsuperscript{16} Something like Kuhnean model can be applied. Before Kripke, there was sort of “normal science” within the broadly Fregean framework, and the framework was mostly satiated. The “scientific revolution” — the shift from the broadly Fregean to the broadly Kripkean framework opened a wide space for the research of new vistas and, accordingly, for academic publi-
3. Looking for the satisfactory descriptivist account of meaning and reference

The Kripkean theory of *a posteriori* necessities coupled with the essentialist theory of natural kinds was meant to replace the kind of accounts of meaning and reference that was almost universally accepted before the publication of *Naming and Necessity* and associated, among others, with the names of John Stuart Mill, Gottlob Frege, and Rudolph Carnap. (Although some authors advanced accounts similar to that of Kripke earlier, those accounts had not drawn much attention). Kripke dubbed it “descriptivism”. Roughly, descriptivism is the view that linguistic terms pick out their referents by means of meanings that are “in the head”, of which the speaker is aware, and which correspond to certain descriptions, so that a term (either proper or general name) functions in natural language as a shortcut for such a description. “Water” is a shortcut for something like “the transparent tasteless drinkable liquid like that in our surrounding that fills rivers, lakes, etc.”; “Hesperus” is a shortcut for something like “the starlike heavenly body first seen in the evening”, and so on. John Stuart Mill accepted this view for general terms but not for proper names; he believed that proper names directly refer to particular things, and have no *connotation* (that is, no meanings expressible in descriptions). Kripke’s theory is sort of marriage of this Millean view with essentialism and its expansion beyond proper names — to natural kind terms.

Kripke’s and Putnam’s arguments against descriptivism have a character of thought experiments that are supposed to invoke our modal intuitions — intuitions about what is possible (usually counterfactually, sometimes epistemically). The reader is invited to imagine a number of scenarios, is suggested how these scenarios should be described correctly, and is expected to see that this is indeed the case; and it is shown that the correctness of these descriptions (or the incorrectness of alternative descriptions) is inconsistent with descriptivism. However, it is usually far from obvious that these scenarios should be described the way that fits Kripke’s theory rather that the way that fits the descriptivist account. Arguments to this point, if provided, turn out to be question-begging — they take for granted certain (in fact, very disputable) premises about essences, or Essential properties. Kripke’s and Putnam’s attack on descriptivism also fails to appreciate the available resources of descriptivism (if appropriately sophisticated) to accommodate many of the modal intuitions to which Kripke and Putnam appeal.

Take, for example, the water-\( H_2O \) identity. It is far from self-obvious that we should describe the counterfactual scenario in which the watery stuff in our surrounding is not \( H_2O \) as such in which the watery stuff is not water. There is a well-known example of the mineral called “jade” that was discovered to be of two chemical varieties: a combination of sodium and aluminium (the variety that has got the name “jadeite”) and a combination of calcium, magnesium, and iron (the variety...
that has got the name “nephrite”). Why should we hold that “water” is a “natural kind term” that functions in the way Kripke and Putnam suggest rather than a functional kind term like “jade” that allows for the possibility of multiple realizability? (Both Kripke and Putnam mentioned the case of jade, but neither provided the reason why “water” and “jade” should be treated in so different ways.)

Even if, nevertheless, the claim that in the envisaged scenario the watery stuff is not water seems plausible, we should ask: why it seems so? Is not it the case that if we tend to agree that in such a scenario the watery stuff is not water, we tend so because we know that in fact, water is H₂O, and this knowledge is assimilated into our meaning of “water”? Of course, if by “water” we mean H₂O or the conjunction of being H₂O with some other properties (perhaps, the liquid stuff that consists of H₂O and is transparent, drinkable, etc.), then water is necessarily H₂O; however, the truth of “Water is H₂O” is a priori. In that case, we have the dilemma both horns of which fit descriptivism and clash with the Kripkean account: if we use the word “water” in the sense that does not imply H₂O, then the truth of the statement “Water is H₂O” is knowable a posteriori and contingent; otherwise, if we use the word “water” in the sense that implies H₂O, then the truth of the statement “Water is H₂O” is necessary and knowable a priori; in no case it is knowable a posteriori and necessary. Kripke’s and Putnam’s arguments against descriptivism presume that meanings of words in natural language and “in the heads” of language users are fixed and never change, and they fail to take seriously the possibility that meanings of words in natural language, and even more so “in the heads” of language users, evolve and accommodate the newly acquired knowledge.

Besides this point about the mutability of word-meanings, the adequate account of the meaning and reference of words like “water”, “gold”, “tiger”, “heat” should take into consideration the indexical (relative to the speaker, or the community of speakers) character of the meanings of many words in natural language. For example, if we talk about water as “drinkable stuff”, we should take into account (as Putnam in his famous Twin Earth thought experiment did not) the fact that “drinkability” is relative to a drinker or a drinker community. So, given our organism’s constitution (the place of the H₂O molecules in our bodies) it may be logically impossible (inconceivable) that any other stuff but H₂O could replace water for our drinking purposes, and so we would not recognize any transparent liquid XYZ stuff (perhaps drinkable for some other intelligent humanlike looking creatures like water is drinkable for ourselves) as water. If so, then in the Twin Earth scenario the relevant question should be not “Whether we would call the transparent liquid XYZ stuff that is drinkable for Twin-Eartheans ‘water’?”, but “Whether we would call it ‘water’ if our organisms were so constituted that for all our important ordinary purposes, including drinking, XYZ would be a perfect replacement for H₂O?”.

17 Putnam [1975: pp. 141-142] considers the objection that perhaps the term “water” had not the same extension in 1750 and in 1950 (on both Earths), and answers this objection by a statement of his and Kripke’s theory of reference for the terms like “water” rather than by providing any argument.
I think that descriptivism about proper names is also tenable. It can meet the Kripkean objections by two main devices.

First, admitting that names function as rigid designators but pointing out that this can be accounted for by descriptions that contain explicit reference to the actual world. For example, the description that adequately expresses the meaning of the name “Hesperus” is not “the starlike heavenly body first seen in the evening” but “the starlike heavenly body actually first seen in the evening”. The difference is that in the description of a counterfactual scenario, the former expression (that does not contain the reference to the actual world) will pick out the starlike heavenly body first seen in the evening in that scenario (possible world), which can happen to be not Hesperus, whereas the latter expression (that contains the reference to the actual world) picks out the starlike heavenly body first seen in the evening in the actual world, which cannot fail to be Hesperus. (For the simplicity’s sake, we leave out of consideration the possibility that in the actual world some other starlike heavenly body can become first seen in the evening. To take account of this possibility, we should make our description a bit more complicated by adding to it the explicit reference to the relevant temporal period: the starlike heavenly body actually first seen in the evening for several centuries or millennia before 2020 CE”.)

Second, for many, if not for all, proper names the Kripkean account of how proper names pick out their referents should be assimilated into the descriptivist framework as the account of these names’ meanings. Kripke suggested that a proper name N used be a speaker S at a time T picks out as its referent the object that is connected to this particular use of “N” by an appropriate communicative chain that begins with the event of “baptism”, when this object was given this name. Kripke considered this account as an alternative to descriptivism; however, this (as well as any other) account of reference, insofar as it is arrived at by armchair speculation of what we would say about various counterfactual scenarios (rather than by empirical methods) can be assimilated by descriptivism. As a first approximation, a descriptivist can say that a proper name “N” used be a speaker S at a time T means the object that is actually connected to this particular use of “N” by an appropriate communicative chain that begins with the event of “baptism”. If we didn’t mean something of the sort, we could not arrive at the account of reference proposed by Kripke by mere means of conceptual analysis, as Kripke did. However, some inadequacies can be found in this view, and the way to correct them is likely to require the distinction between:

— on the one hand, personal meanings “in the heads” of the users of a proper name, which need not involve the whole communicative chain that begins with the baptism-event and ends with its particular use by a particular user; instead, it can involve only the last link of the chain, beginning with the event of this particular user being introduced to this name and acquainted with its referent); accordingly, these meanings “in the heads” of different users will be different but pick out the same referent, and belong to one “tree” of the communicative chains rooted in the baptism event;
— on the other hand, the generalisation that can be considered as the proper name’s meaning in language, and which should involve the whole “tree” of the communicative chains rooted in the baptism event rather than one particular chain ending in one particular use.

The details of this account require elaboration in separate publications. The careful re-examination of Kripke’s and Putnam’s thought experiments involving purported natural kind terms and countering their analysis with the descriptivist (or conceptualist) account along the lines suggested above is also forthcoming.

REFERENCES


Being Sceptical about Kripkean A Posteriori Necessities and Natural Kinds


Received 05.10.20