Reflections on Saul Kripke's
Naming and Necessity (1980)

Jesper Jerkert

February 9, 2012

Introduction

Saul A. Kripke's *Naming and Necessity* was first published in 1972. A revised and enlarged edition appeared in 1980 and has been continually available in print since then. The book is essentially a transcription of three lectures delivered in January 1970 at Princeton University, with some amendments and a preface added.[1] This text will give an overview of Kripke's book and offer some comments on its philosophical content. Of course, I will not cover each and every argument in the book. In particular, I will comment on the second half of the third lecture only briefly.

Lecture I

The relation between proper names and definite descriptions is Kripke's first topic. By definite description is meant a phrase of the form "the x such that Fx", where F is one or several properties or characteristics. According to John Stuart Mill, names have denotation but no connotation; in other words, a name has no meaning apart from its giving a reference to the bearer. By contrast, many philosophers agree with Gottlob Frege that names do indeed have connotation (or "sense"). There must be some descriptive content in names, Kripke concurs (p. 28), or else it is very difficult to understand how people ever started using names to refer to anything. Another simple argument in favor of this view could be that it is possible to discover that two distinct names have the same referent. For example, Hesperus, the evening star, has been found to be identical to Phosphorus, the morning star.[2] If there were no meanings in the names Hesperus and Phosphorus except for their making reference to the planet Venus, we would not gain any new information from learning that 'Hesperus is Phosphorus'. But intuitively, we do learn something from this phrase.

In spite of crediting Frege (and others) for this insight, Kripke holds that Frege, Russell, Searle, Strawson, and others were wrong in their analyses leading to the usage of a cluster concept for the sense of names. According to this view, in a way names do correspond to definite descriptions, but not to any *particular* definite description. For it is

---

[1] Surprisingly, the lectures were given without any manuscript or even any notes, according to Kripke's footnote on p. 22!

[2] In Greek mythology, one could argue that Hesperus and Phosphorus are still distinct, though.

easy to see that it would be highly problematic to take names as synonymous to particular definite descriptions (Morris, 2007:74ff). For example, a person called John Smith may be equivalent to "the bald guy with the old Volvo living in Knoxville Street" to someone, but he may be described as "the math lecturer with the funny teeth" by someone else. Of course, both these descriptions may be correct, but it seems unfeasible to attach any significance to a particular description in a theory of the meaning of names. For one thing, if it is part of the meaning of John Smith that he owns a Volvo, then the statement "John Smith owns a Volvo" would be tautological. But intuitively, there is nothing tautological with such a statement; on the contrary we may find it quite informative. Therefore, the idea of a cluster concept, or perhaps of a "family resemblance", has emerged, most notably promoted by Searle (1958). According to this idea, whatever satisfies most of (or enough of) the family of descriptions attached to a name is the true bearer (the referent) of that same name (Kripke, 1980:31).

But some problems still present themselves in such a theory. First, it is not clear that the problem of propositions being tautological disappears completely. Secondly, as pointed out by Kripke (1980:45ff), there is a problem with defining the meaning of a name with qualitative descriptions, whether part of a "cluster" of descriptions or not. This problem – as far as I understand Kripke – appears when we ask what it would mean to say that, e.g., Richard Nixon might have lost the presidential election in 1968.[3] The most popular way of making sense out of such statements (at least among philosophers) is to discuss whether there are "possible worlds" in which Nixon lost the election in 1968. To identify the meaning of Nixon as a cluster of descriptions and then ask whether Nixon might have lost the election in 1968 would then amount to the question of whether there is a possible world in which there is a person sufficiently similar to a cluster of descriptions intended to define Nixon (e.g., holding such and such political views, having such and such appearances, talking in such and such ways, growing up in this or that place, etc.) but who did not win the election in 1968. This seems to be a rather difficult thing to assess: how do you go about judging the similarity between a description cluster and a referent, or between two description clusters? And even if it was a feasible thing to do, there is still something unintuitive about the procedure. Says Kripke:

> "We just say 'suppose this man had lost'. It is *given* that the possible world contains *this man*, and that in that world, he had lost. (…) But, if we have such an intuition about the possibility of *that* (*this man's* electoral loss), then it is about the possibility of *that*. It need not be identified with the possibility of a man looking like such and such, or holding such and such political views, or otherwise qualitatively described, having lost. We can point to the *man*, and ask what might have happened to *him*, had events been different." (Kripke:1980:45f)

This quote points to the intuition that the very man Nixon is not identical to (a cluster of) qualitative statements describing him, not even if the descriptions are highly accurate. Instead, when we talk about Nixon having lost in another possible world we take for granted that we are talking about Nixon and no one else – the very same man as in our actual world. I share Kripke's view (again: as I have understood it) that descriptions cannot be taken as definitions of names, neither their references nor their meanings. Kripke introduces (1980:48) the term "rigid designator" for something designating the

---

[3]For those unfamiliar with this era of American modern history, Nixon in fact won the election in 1968 and was inaugurated as President of the United States in January 1969.

same object in every possible world, and he claims that proper names are rigid designators. Note that the existence of a rigid designator, e.g. Richard Nixon, does not imply that the object does exist in every possible world. It merely implies that in those worlds where Nixon exists, he is the same object.

So far, I believe Kripke's position is wholly reasonable (though his style is not always as clear as I would have wished for). There is something intuitively unsatisfactory with equating names with descriptions (or with description clusters). I found the following quote especially persuasive:

> "What I do deny is that a particular is nothing but a 'bundle of qualities', whatever that may mean. If a quality is an abstract object, a bundle of qualities is an object of an even higher degree of abstraction, not a particular" (Kripke, 1980:52).

In other words, a categorical mistake seems to be committed by supporters of the description theory. Richard Nixon is a particular and can in no way be equated with a bundle of qualities or with a cluster of descriptions, for the latter are no particulars but abstract entities. According to this view, the problem of assessing whether the reference of a name (say, Richard Nixon) in a possible world not identical to our actual world is correct or not is judged to be a pseudo-problem and disappears. By 'Richard Nixon' we refer to the same Richard Nixon in all worlds, period. (But in some of these worlds he might not exist.)

Indeed, rather fundamental difficulties arise if we think that the reference of a name is determined by descriptions. For example, if we discuss whether Aristotle could have failed to be the teacher of Alexander, and maybe could have failed to take an interest in philosophy at all, people committed to a description theory will immediately object with an argument approximately like: "Well, if we are talking about a person that didn't even take an interest in philosophy, it's doubtful that we are still talking about Aristotle". But this seems to be a misdirected objection. Surely it could have happened in our actual world as well as in other possible worlds that *the man* Aristotle did not take an interest in philosophy. I cannot see anything unfeasible with such a line of thought. There might be, of course, names the referents of which we know practically nothing except for one or a few easily listed facts. But even if very little – say, one single fact – were known about a certain person, that does not imply that imagining this fact to be false would amount to talking about someone else. For example, let's assume that the only fact known about Homer is that he authored the *Iliad* and the *Odyssey*. It is entirely possible, in my view, to imagine a world in which the man Homer did not author the *Iliad* and the *Odyssey* (for example, a world in which the *Iliad* and the *Odyssey* do not exist, but where Homer still existed). Although in such a situation we might be unable to tell anything about Homer within that world, it seems to me wrong to say that Homer could not exist in that world.

In the final pages of the first lecture (pp. 64–70), Kripke lists and starts discussing what he takes to be six fundamental commitments or theses of the descriptive theory of names (of individuals). If the speaker is called *A*, the theses that the supporter of the descriptive theory is committed to are as follows (Kripke, 1980:64f).

(1) To every name or designating expression '*X*', there corresponds a cluster of properties, namely the family of properties φ such that *A* believes 'φX'.

(2) One of the properties, or some conjointly, are believed by *A* to pick out some individual uniquely.

(3) If most, or a weighted most, of the φ's are satisfied by one unique object *y*, then *y* is the referent of '*X*'.

(4) If the vote yields no unique object, '*X*' does not refer.

(5) The statement, 'If *X* exists, then *X* has most of the *φ*'s' is known *a priori* by the speaker.

(6) The statement, 'If *X* exists, then *X* has most of the *φ*'s' expresses a necessary truth (in the idiolect of the speaker).

The "vote" referred to in (4) is something that Kripke imagines must occur for (3) to be assessed, especially if different descriptions are weighted differently. Kripke states that he takes (1) to be true by definition but (2)–(6) to be false.[4] In addition to (1)–(6), he adds a principle of non-circularity that he takes to be reasonable to impose on (1)–(6) (and also on other sets of theses constituting a similar theory of descriptions):

(C) For any successful theory, the account must not be circular. The properties which are used in the vote must not themselves involve the notion of reference in such a way that it is ultimately impossible to eliminate.

Lecture II

At the outset of the second lecture, Kripke returns to an issue on which he commented briefly at the end of the first lecture, namely the theory that a name means nothing but 'the individual called <the name>'. For example, according to this theory 'Socrates' means 'the individual called "Socrates" '. There have been philosophers holding this to be a correct theory. Kripke rejects this view, citing condition (C). Although the theory stating that a name means nothing but 'the individual called <the name>' is not necessarily trivial (for there are cases when a name is currently used for a person for which that very name was most likely not used historically – the prophet Isaiah could be one example), it seems very difficult for such a theory to escape (C). Since I agree that (C) is a reasonable principle, I also agree that a theory stating that the meaning of 'Socrates' is 'the individual called "Socrates" ' is not good enough.

After this, Kripke again goes back to criticizing a theory according to which the referent of a name in other possible worlds that we might discuss, say the name 'Aristotle', must be identified by some process of judgment where descriptions are compared or evaluated. It may be true that the most distinctive attributes of Adolf Hitler was his hateful and murderous political deeds, but that would still not entail that we would cease talking about Hitler in a counterfactual scenario where he hadn't entered politics at all (p. 77). Kripke holds that this shows (6) to be false: it is not necessarily true that an individual must have most of certain properties. Thesis (5) is false as well, according to Kripke. He holds that (5) is derived from (1)–(4) in some way, although he doesn't give the details (p. 73).

Talking about necessities and possibilities (and in particular about possible worlds), one might try to distinguish different ways in which phrases like "It is possible that the *F* is *G*" or "It is possible that the *F* is not *G*" can be understood. Taking $\Diamond$ to be the possibility operator, one interpretation of "It is possible that the *G* is *F*" could be

$$\Diamond(\exists x(\forall y(Fy \leftrightarrow x = y) \land Gx)),$$

---

[4]As acknowledged by Kripke, (1)–(6) do not exhauste all theses that a supporter of the description theory might reasonably be taken to hold (pp. 65f).

whereas another one could be

$$\exists x \big(\forall y (Fy \leftrightarrow x = y) \wedge \Diamond (Gx) \big).$$

In the first case, the modal operator has a scope stretching over the whole expression, but in the second case the existential quantification has the widest scope. In a modal expression involving a negation, the number of interpretations may be even greater (depending on the scopes of negation as well as of modality).[5]

One could have guessed that Kripke's critique of descriptive theories that are using possible world semantics would hinge on such ambiguities of scope. But in the preface (pp. 5ff) he seems to deny that this is the case. He asks us to consider propositions such as these:

    (1) Aristotle was fond of dogs.
    (2) The last great philosopher of antiquity was fond of dogs.

Kripke takes Russell to be a philosopher who would regard (1) and (2) as equivalent, since their truth conditions would agree extensionally if it is indeed the case that Aristotle was the last great philosopher of antiquity (p. 7). But Kripke himself takes names to be rigid, and he therefore does not agree that (1) and (2) are equivalent. In Kripke's words: "But counterfactually, Russell's conditions can vary wildly from those supposed by the rigidity thesis" (ibid.). An important point is that (1) and (2) do not involve modality, so there can be no ambiguity arising from different scopes of modality operators here (p. 11) — nor, for that matter, from scopes of a negation operator. The question of rigidity is still there, so Kripke concludes that "the view that *reduces* rigidity to scope (…) is simply in error" (p. 12).

According to a descriptive theory of names, it would be natural to assume that names refer in virtue of their implicit descriptive content: a name refers because it is really a disguised description of an object that fits the description better than any other object. Reference is created by the similarity between reality and the description.[6] But since Kripke does not believe in descriptive theories, he will have to come up with some other theory of reference. He comes up with an outline of a causal theory of reference (pp. 91ff). The idea is that a name spreads as if by a chain through talk of various kinds between people. The chain starts when a baby gets its name and the parents start talking about the child using that name. At the other end there could be a someone who never met the person and certainly do not know the path taken by the name chain to reach him. For example, a person might refer to the famous American physicist Richard Feynman (1918–1988) by using the name Feynman

> "even though he can't remember from whom he first heard of Feynman or from whom he ever heard of Feynman. He knows that Feynman is a famous physicist. A certain passage of communication reaching ultimately to the man himself [Feynman] does reach the speaker. Then he is referring to Feynman even though he can't identify him uniquely. (…) [He could] have trouble distinguishing between Gell-Mann[7] and Feynman. So, he doesn't have to know these things, but, instead a chain of

---

[5]Cf. Glüer & Pagin (2006:509f) for formalizations in modal logic of different readings of a Kripkean example.

[6]I am not suggesting that this is the *only* available theory of reference for a descriptivist; only that it is a theory that would spring easily to mind to a descriptivist.

[7]Murray Gell-Mann (b. 1929) is another famous American physicist.

communication going back to Feynman himself has been established, by virtue of a membership in a community which passed the name on from link to link" (p. 91).

In Kripke's view, the important thing is that there is a link from the initial "baptism". But the details are unclear. For example, in Kripke's quote, the speaker "knows that Feynman is a famous physicist". Is this necessary? Couldn't the speaker be ignorant of this fact without creating trouble for Kripke's theory?

Generally speaking, there are several potential problems with a causal theory of reference. First, there seem to be non-referring names. For example, 'Father Christmas' does not refer to a real person (unless we hold that the name refers to Saint Nicholas, perhaps). Would it then be meaningless to talk about Father Christmas? This problem is mentioned by Kripke (p. 93), so one would guess that his theory could avoid the problem somehow.

Secondly, the chain may be broken somewhere, or there may be a change of reference along the chain. For example, imagine that the first person in the chain is connected to one follower only, but this follower misunderstands the first person's intentions, and uses the name learned from the first person for something else than what was intended by the first person. Actually, similar scenarios may be true for quite important names in real life. Gareth Evans has suggested that the name 'Madagascar' shifted its meaning with Marco Polo. Before Marco Polo, the name referred to part of mainland Africa, but due to a misunderstanding on Marco Polo's part, 'Madagascar' started being used as the name of an island still known as Madagascar today (Reimer, 2010). Would this matter to Kripke's theory? Well, he doesn't present a full-fledged theory, only an overall idea about a theory, so I cannot really say.

One could go on listing potential problems for a causal theory of reference. Thirdly, for example, a causal theory will have to explain carefully what exactly will constitute a "link" in the chain. Under what circumstances has a link been established, and under what circumstances has it failed to be established? However, since Kripke's theory is presented in such a rough form (he even states that it should *not* be called a theory; p. 97 and *passim*), there is perhaps little point in listing more questions one would like to ask him.

In the final pages of the second lecture, Kripke discusses identity statements (pp. 97–105), returning to the 'Hesperus is Phosphorus' example and to statements such as 'heat is the motion of molecules'. The question is whether the truth of such statements (i.e., 'Hesperus is Phosphorus' as well as 'heat is the motion of molecules') is necessary or contingent. Before Kripke, the received view (according to Kripke) was that these are contingent truths, for it is possible (or so people tended to think) that they could have been false. Kripke holds that they are necessary truths (p. 99). This seems to follow quite naturally from the rigidity thesis. If a name is a rigid designator it points to the same individual in all possible worlds. Equating two rigid designators and claiming that they could possibly point to different individuals would then amount to a self-contradiction. As for natural kinds such as 'heat' and 'gold', their nature is discussed in lecture III, below.

Lecture III

Kripke's third lecture is mainly about natural kind terms. His views are generally considered to be akin to Hilary Putnam's as exposed in the article "Meaning and Reference"

(1973).

Kripke repeats that if an identity statement between names is true, it is also necessarily true. However, it needs not be true *a priori*; we could learn empirically that it is true (pp. 109f). He continues discussing essential properties. For example, a given person might have had other properties than the ones she actually has, and might have experienced other things than the ones she actually experienced. But some properties cannot be removed: "How could a person originating from different parents, from a totally different sperm and egg, be *this very woman*?" (p. 113). He extends the reasoning to all material objects and suggests a principle (in embryonic form): "If a material object has its origin from a certain hunk of matter, it could not have had its origin in any other matter" (p. 114 n56, italics removed from original).

But this talk of "matter" is of course highly abstract and not necessarily of much help to sort out which properties are essential for a certain object. However, Kripke gets a little more concrete when he discusses the properties of gold (pp. 117–125). Gold is taken as a prime example of a natural kind. What properties of gold could turn out to be different from what we believe without making us no longer talk about gold? Kripke asks us to consider two questions. The first is: Could we discover that gold is not yellow? (p. 118). He answers in the affirmative. It is possible to imagine a situation where we discover that gold is in fact blue. We would then not stop using the word "gold" and invent a new word for what we formerly called gold. Instead, we would continue calling it gold and simply acknowledge that up until the discovery that gold is blue, we were wrong about the color of gold. So this seems to be true:

(3) It is not necessary that gold is yellow.

What about this question, then: Could gold turn out not to have atomic number 79 (which it has, as far as we know)? (p. 123). Here, Kripke's answer is again yes: the standard atomic theories could turn out to be wrong (ibid.). But then he continues discussing the question under the assumption that gold *does* have atomic number 79. In that case, could gold turn out not having atomic number 79? This time, the answer is no. To me, the reasoning here is a little unclear. My interpretation is that Kripke's message is *not*

(4) It is necessary that gold has atomic number 79.

but rather

(5) If gold has atomic number 79, then it is necessary that gold has atomic number 79.

Very well, given that gold is considered a natural kind term, and that such terms have certain rigidity in the Kripkean cross-worlds sense, (5) is reasonable. What makes me confused is that I interpret Kripke as contrasting the question on the color of gold with the question on gold's atomic number. They have different answers. As for the color of gold, the answer is (3), but for the atomic number question, the answer is (5). These answers do not have the same form. But couldn't they have more similar forms? It seems to me that the answer to the color question could be

(6) If gold is yellow, then it is necessary that gold is yellow.

But I take it that Kripke does not agree. But why not? Maybe because having atomic number 79 is core part of scientific theories, whereas yellowness is not? But can we

draw this line between (5) and (6) with certainty? Maybe it is true that yellowness is deducible from scientific laws concerning metals (I would believe this is the case, but I'm no chemical expert). Then perhaps the answer to the color question could be

> (7) If our standard theories on chemical elements (atomic numbers, etc.) and on visual perception are correct, then it is necessary that gold is yellow.

So even if (6) is not a good answer, maybe (7) is. I would have liked to know why (6) and (7) are not admissible answers to the color question, if this is indeed what Kripke holds.

In the remains of lecture 3, Kripke argues, as I read him, for the existence of essentials connected to kinds such as 'light' and 'heat', and he draws conclusions for the philosophy of mind (the latter is in pp. 134ff). It seems that he wants to condition assertions of the existence of essentials in ways similar to (5). For example, we could imagine a world in which heat is not molecular motion, but *if* heat is molecular motion then necessarily heat is molecular motion. (This is my interpretation of Kripke, pp. 129ff.) Even if there were creatures sensitive to, e.g., light "in such a way that they felt the same thing that we feel when we feel heat" (p. 132), we would not say that heat has become light or vice versa. But it is true, says Kripke, that *the reference* of heat may be given in terms if feelings of heat.

Final words

In insisting that names are not disguised descriptions, one could perhaps argue that Kripke not only contributed to the philosophy of language, but also launched an attack on the idea that analysis of language is the main (or only) method of philosophy. The primacy of analysis of language has been a widely held belief in analytic philosophy during the 20th century (indeed, it has sometimes been cited as the hallmark of analytic philosophy). It would be paradigmatic of a language analysis approach to hold that names are really descriptions in disguise, so that when you say 'Aristotle' you are really saying 'the x such that Fx and Gx and …'. But according to Kripke, this is not enough (and not even correct). Metaphysical considerations will have to enter the analysis, and we have to think about the notions of necessity and possibility. Of course, metaphysical arguments are standard ingredients of contemporary philosophy of language, but it still seems to me that Kripke's approach and criticisms of a more straight-forward linguistic analysis of the meaning of names constitute an eye-opening challenge to earlier beliefs.

For this and other reasons, it is not difficult to understand that *Naming and Necessity* has become very influential. According to one textbook, "Kripke's arguments against description theories of names inaugurated a revolution in the philosophy of language" (Morris, 2007:94). But I would like to raise one general criticism. On the first reading of *Naming and Necessity*, Kripke's prose seemed to flow effortlessly and I got hooked on his overall message (or, at least, on what I *perceived* as his overall message). The second time, however, the confusion increased considerably. In short, Kripke's lecture-style text could be clearer. As evidenced by large portions of Kripke's preface, there have been several misunderstandings of what he wanted to say. It is a pity he did not spend more time editing the manuscript into a more carefully argued and phrased text. The lecture script makes an easy read from one perspective, but partially it is a difficult read from another, namely in terms of clarity.

Of course, Kripke is aware of at least some of the deficiencies of his text. For example, when he presents his outline of a causal theory of reference he immediately recognizes difficulties that would need to be addressed; his excuse for not addressing them is that "first, I'm sort of too lazy at the moment; secondly, rather than giving a set of necessary and sufficient conditions which will work for a term like reference, I want to present just a *better picture* than the picture presented by the received views" (p. 93). Yes, maybe he has offered a better picture, though still a blurred one. On the other hand, maybe the blurriness has added to the fame of *Naming and Necessity*. Fellow philosophers may have found it an irresistible challenge to scrutinize and discuss a work open to various interpretations from a distinguished philsopher that prior to *Naming and Necessity* was known mainly for his work in logic.

References

Glüer, Kathrin & Pagin, Peter (2006). "Proper Names and Relational Modality", *Linguistics and Philosophy* 29(5), 507–535.

Kripke, Saul (1980). *Naming and Necessity*. Cambridge, Mass.: Harvard University Press, 2nd edition.

Morris, Michael (2007). *An Introduction to the Philosophy of Language*. Cambridge: Cambridge University Press.

Reimer, Marga (2010). "Reference". In: *Stanford Encyclopedia of Philosophy*, available at http://plato.stanford.edu/archives/spr2010/entries/reference/.

Searle, John R. (1958). "Proper names", *Mind* 67, 166–173.