THE REVIEW OF SYMBOLIC LOGIC
Volume 2, Number 1, March 2009

EXTERNALISM, INTERNALISM, AND LOGICAL TRUTH

CORINE BESSON
St Hugh's College, Oxford

Abstract. The aim of this paper is to show what sorts of logics are required by externalist and internalist accounts of the meanings of natural kind nouns. These logics give us a new perspective from which to evaluate the respective positions in the externalist–internalist debate about the meanings of such nouns. The two main claims of the paper are the following: first, that adequate logics for internalism and externalism about natural kind nouns are second-order logics; second, that an internalist second-order logic is a free logic—a second order logic free of existential commitments for natural kind nouns, while an externalist second-order logic is not free of existential commitments for natural kind nouns—it is existentially committed.

§1. Introduction. Externalists and internalists about the meanings of natural kind nouns disagree about whether the meanings of such nouns should be explained in terms of facts about the external physical environment of the speakers using them. Externalists claim that this external physical environment plays an essential role in individuating the meanings of natural kind nouns; internalists claim that it plays no role in individuating such meanings. This paper considers the consequences of these opposing views for logic. It shows what sorts of logics are required by externalist and internalist accounts of the meanings of natural kind nouns, and proceeds to develop and assess these logics. Considering the consequences of externalism and internalism for logic enables us to evaluate them from a new perspective. In particular, the paper argues for the following two claims:

1) The sorts of logics required by externalism and internalism about natural kind nouns are *second-order logics*.
2) An internalist second-order logic for natural kind nouns is a *free logic*—a logic free of existential commitments for such nouns; and an externalist second-order logic is not free—it is existentially committed.

The general picture of the relation between semantics and logic which explains the connection between the externalist–internalist debate about natural kind nouns and claims 1) and 2) is this: semantics determines logic. That is, one's account of linguistic meaning determines which logical relations—which relations of logical entailment—hold between the sentences of a language. Different accounts of meaning require different logics.[1] This is a natural picture to have: one role of logic is to tell us what it is for an inference to be

Received: October 30, 2007

[1] The claim that semantics determines logic is metaphysical, not epistemological. I take it that, if true, externalism and its logic apply to natural languages; but I make no claim as to whether or how speakers know the basic semantic and logical mechanism that operate in their language and no claim as to whether or how they represent such semantic and logical facts to themselves.

© 2009 Association for Symbolic Logic

doi:10.1017/S1755020309090091

https://doi.org/10.1017/S1755020309090091 Published online by Cambridge University Press

logically truth-preserving; which inferences are truth-preserving is determined by what it takes for the premises and the conclusion to be true; and this is determined by the way meaning assignments work for these sentences. That is, these assignments determine which patterns of inference ought to be valid in the language for which the semantic theory is given.[2] Thus, in the case in hand, externalists and internalists regard different sentences containing natural kind nouns as logically true and different sequences of sentences containing such nouns as logically valid because they provide different accounts of the meanings of natural kind nouns and of the sentences in which these nouns occur.

The remainder of this introduction explains the connection between the externalist–internalist debate about the meanings of natural kind nouns and claims 1) and 2).

The standard setting for the externalist–internalist debate about the contribution of the external physical environment to the meanings of natural kind nouns is Putnam's Twin-Earth thought-experiment: we consider the natural kind noun 'water' in two possible situations (Earth and Twin-Earth), where the external physical environments are different only with respect to water. While there is water (H_2O) on Earth, there is twater (XYZ) on Twin-Earth.[3] Externalists claim that 'water' means something different on Earth and Twin-Earth: the environment partly individuates the meanings of natural kind nouns, and so a difference in the environment with respect to water entails a difference in the meaning of 'water'. Internalists claim that 'water' means the same on Earth and Twin-Earth: the environment plays no role in individuating the meanings of natural kind nouns, and so a difference in the environment with respect to water on Earth and Twin-Earth does not entail a difference in the meaning of 'water'.[4]

One aspect of the externalist–internalist debate that is relevant to logic is that concerning the meanings of *empty natural kind nouns*, that is, natural kind nouns for which there is nothing in the external physical environment that could individuate their meanings. The issue of emptiness has been made salient in the externalist–internalist debate by using a thought-experiment similar to that of Twin-Earth, namely the so-called 'Dry-Earth' scenario.[5] In Putnam's Twin-Earth scenario, the issue is whether 'water' would have meant something different, if it had referred to a different liquid (on Twin-Earth) from that which it actually refers to (on Earth).[6] In a Dry-Earth scenario, we compare Earth, where 'water' refers to H_2O, and Dry-Earth, where 'water' refers to nothing whatsoever—where there is no natural kind that 'water' refers to. (I come back to this characterization shortly.) For instance, we can suppose that our near-intrinsic duplicates on Dry-Earth are the victims of

[2] A familiar example here is how antirealism about meaning serves to motivate intuitionistic logic (see Dummett, 1975). Another, which is relevant here, is of course how different views on the semantics of empty names generate different logics—classical or free (see Lehmann, 2002).

[3] See Putnam (1975). N.B. the externalist–internalist debate at issue here is that which concerns whether the natural kinds that are part of the external physical environment contribute to the meanings of natural kind nouns. There is another externalist–internalist debate about the meanings of natural kind nouns that concerns the extent to which speakers' understanding of natural kind nouns is determined by social interactions. It is assumed here that these two debates are independent, and only the former is at issue here.

[4] These claims are made precise in section 2, where the notion of meaning is made precise.

[5] See Boghossian (1997), who first brought up this scenario in the context of compatibilism—the issue of whether externalism is compatible with privileged access. See McLaughlin and Tye (1998), Stoneham (1999), Sawyer (2003), and Goldberg (2006) for externalist discussions of the Dry-Earth scenario.

[6] In this paper, natural kind nouns are taken to be *referring expressions*. This is explained further in section 2.

https://doi.org/10.1017/S1755020309090091 Published online by Cambridge University Press

the collective illusion that there is water on their Earth; however there is no such liquid. The issue here is whether 'water' has a different meaning on Dry-Earth and on Earth.[7]

Internalists who have discussed the Dry-Earth scenario (or a similar one) claim that 'water' means the same on Earth and Dry-Earth: if the external physical environment does not play a role in individuating meaning, the meaning of 'water' is the same. Intuitively, externalists should at least say that 'water' means something different on Earth and Dry-Earth: if the environment partly individuates meaning, a difference in the external physical environment with respect to water entails a difference in meaning. Internalists have in fact used this scenario to argue that externalists cannot give a satisfactory account of the meaning of 'water' on Dry-Earth, and, generally, of the meanings of empty natural kind nouns.[8] By contrast internalists can offer a straightforward account, which flows from a general account of the meanings of natural kind nouns: take whatever Earth you like—Earth, Twin-Earth, or Dry-Earth—'water' means the same on those Earths.

An externalist might reply that this charge is misguided, for although there might be something like an internalist line on empty natural kind nouns, there is nothing like an externalist one: externalists are interested in giving an account of the relation between the meanings of natural kind nouns and the natural kinds they refer to, and not one of natural kind nouns that are empty. They are only interested in natural kind nouns that actually refer. They may agree that it would be desirable to have an account of empty natural kind nouns, but that is just not what they are aiming to give, and so they need not have a tailor-made account for them.

Internalists will find this reply unsatisfactory. They will argue that externalists owe us an account of empty natural kind nouns and further that they cannot give one. For instance Gabriel Segal has argued that externalists are committed either to regard empty natural kind nouns as meaningless or to fall back on a form of internalism and say that 'water' means the same on Earth and Dry-Earth, despite the difference in the external physical environment with respect to water.[9] It is not possible for externalists to say that 'water' means the same on Earth and Dry-Earth. It would trivialize their account: if 'water' means the same on Earth and Dry-Earth, then presumably it means the same on Twin-Earth and Dry-Earth (from the standpoint of Dry-Earth, there is nothing special about Earth—Dry-Earth lacks both water and twater). But then by transitivity 'water' means the same on Earth and Twin-Earth, which is precisely what externalists deny. So the natural thing to say for externalists is that as far as externalism goes, empty natural kind nouns are meaningless. They could further argue that this is something that they can live with or that some other separate theory can be given for them.

In principle, there is another possible option that externalists could consider, which is to say that 'water' refers to an empty kind on Dry-Earth. In general, there are two ways in which the idea of an empty natural kind noun can be understood: either as a noun that fails to refer to anything whatsoever or as a noun that refers to an empty natural kind. So far I have treated emptiness as a case of reference-failure and it is worth seeing whether this other option would be a good externalist strategy to endow 'water' with a meaning on Dry-Earth. One problem with saying that 'water' refers to an empty natural kind is that it is

[7] I do not discuss the somewhat similar case of natural kind nouns that refer to motleys (e.g. 'jade'). It is standard for externalists to say that natural kind nouns refer uniquely, and for internalists to allow them to refer to motleys.

[8] See Boghossian (1997) and Segal (2000).

[9] See Segal (2000).

https://doi.org/10.1017/S1755020309090091 Published online by Cambridge University Press

unclear how such an empty kind could determine the meaning of 'water': that empty kind is not part of the external physical environments of speakers. Also, externalists generally appeal to a causal theory of reference to explain the relation between the meaning of a natural kind noun and the kind it refers to.[10] But empty natural kinds are causally inert. Perhaps there is another sense of 'environment' to which externalists can appeal here: perhaps an environment that is external but not causal. But the problem is that whatever sort of environment this is, it is going to be the same on Earth, Twin-Earth, and Dry-Earth, because, by stipulation, the external physical environment is the same on all of them except for the presence of samples of water. With respect to this sort of external environment, there is no difference between these Earths, and so no reason to count 'water' as having a different meaning on any of them. This notion of external environment will not enable the externalist to say consistently that 'water' means something different on Dry-Earth as opposed to Earth or Twin-Earth because there it refers to an empty kind.

A further problem for the externalist is that no empty kind is a natural candidate to be the referent of 'water' on Dry-Earth: there is no more reason to say that 'water' refers to the empty kind water (which is nonempty on Earth) than there is to say that it refers to twater (which is nonempty on Twin-Earth). Again, from the standpoint of Dry-Earth, there is nothing special about Earth as opposed to Twin-Earth. And of course, 'water' on Dry-Earth cannot refer to both the empty kinds water and twater: these kinds are necessarily distinct. Perhaps the externalist should say that 'water' on Dry-Earth refers to an empty kind which is neither water nor twater—some sui generis empty kind. But it is hard to see which kind that might be: there just is no obvious or natural candidate;[11] and it is ever more unclear how people on Dry-Earth could ever refer to that kind. It thus seems that any assignment of an empty kind to 'water' different from water (or twater, etc.) is quite arbitrary. But, as just argued, 'water' cannot refer to water (or twater, etc.) on Dry-Earth. So 'water' does not refer to anything whatsoever on Dry-Earth—not even an empty natural kind.[12]

Given that externalists cannot appeal to empty kinds to endow empty natural kind terms with meanings, and given that, for instance, 'water' on Dry-Earth cannot have the meaning it has on Earth, the most consistent way to go for externalists is to say that, as far as externalism is concerned, empty natural kind nouns are meaningless; semantic resources other than externalism will have to be invoked to account for them.[13] Concerning internalists, they need not appeal to empty natural kinds to endow empty natural kind nouns with meanings: just as they do not need to appeal to water on Earth or to twater on

[10] See Kripke (1972) and Putnam (1975).

[11] If natural kinds are represented using some set-theoretical apparatus, the empty set will be a candidate to be the referent of 'water' on Dry-Earth (as well as the referent of 'unicorn' and 'phlogiston'—if they are empty there). But we cannot just presuppose that kinds are some sorts of set-theoretic entities which necessarily exist; and it would still be unclear how reference to them could be achieved.

[12] Here it should be stressed that the claim concerns endowing a natural kind noun with a meaning or introducing the noun for the first time. For instance, externalists can obviously consistently say that natural kind nouns can meaningfully refer to extinct natural kinds (if there are such things)—that is, kinds that were exemplified at the time the noun was introduced but that no longer are.

[13] Not all externalist options have been considered here. See Sterelny (1983, p. 108) for the suggestion that empty natural kind nouns partially refer to instantiated natural kinds (e.g. 'phlogiston' partially refers to oxygen). See Bilgrami (1992, pp. 383–384) for an (externalist) descriptivist account of such nouns. None of these seem promising.

https://doi.org/10.1017/S1755020309090091 Published online by Cambridge University Press

Twin-Earth. Having said that, they could not simply say that empty natural kind nouns refer to empty kinds. This is again for the reason that there is no nonarbitrary referent for 'water' on Dry-Earth (water or twater or some sui generis empty natural kind, etc.).

In the rest of the paper, by an 'empty natural kind noun' I shall mean a natural kind noun that fails to refer. Again, because they are not part of the external physical environment, empty natural kinds play no role in explaining what is at stake between externalism and internalism about the meanings of natural kind nouns. And by a 'nonempty natural kind noun' I shall mean natural kind noun that refers to an exemplified, nonempty, natural kind.

The state of play now is that as far as externalism is concerned empty natural kind nouns are meaningless and as far as internalism is concerned they are meaningful. This aspect of the disagreement between externalists and internalists is relevant to logic—and in particular to claim 2) that an internalist second-order logic for natural kind nouns is free of existential commitments, and that an externalist such logic is not free. If natural kind nouns have to be nonempty to be meaningful, externalism requires an *existentially committed account* of the meanings of such nouns. A meaningful natural kind noun is a noun that refers to an exemplified natural kind. If so, an externalist logic for natural kind nouns is existentially committed in that it requires that such nouns refer to exemplified natural kinds. Internalists do not require natural kind nouns to be nonempty to be meaningful—indeed they do not have to require natural kind nouns to refer to any kind (exemplified or not) to be meaningful; and so they do not provide existentially committed accounts of natural kind nouns. Thus an internalist logic for natural kind nouns is free of existential commitments.[14]

Usually, a free logic is defined as a logic that is free of existential commitments with respect to *singular terms*. It is an alternative to first-order classical logic, which does not allow for empty singular terms. If empty singular terms are allowed in the language, classical principles of inference involving singular terms are invalid. Free logic is the result of modifying those principles to ones that are valid when empty singular terms are allowed in the language.[15] More precisely, the hallmark of free logic is the rejection of the classical principle universal instantiation (UI):

(UI) $\forall x \Phi(x) \rightarrow \Phi(t)$, where t is free for Φ, and t is a singular term.

In free logic, it is standard to have a principle weaker than (UI) is as follows, such as ($\mathrm{UI_F}$)—'F' for free:

($\mathrm{UI_F}$) $\forall x \Phi(x) \rightarrow (\exists x x = t \rightarrow \Phi(t))$, where t is free for Φ, and t is a singular term.

In ($\mathrm{UI_F}$), the clause $\exists x x = t$ guarantees that something in the domain is t. In section 4, where externalist and internalist logics for natural kind nouns are developed, principles relevantly similar to (UI) and ($\mathrm{UI_F}$) are given.

Let us turn to claim 1) that the logics for externalism and internalism are second-order. Again free logics are typically first-order. However it is natural kind *nouns* that are at issue

[14] Of course not every externalist or internalist will recognize herself in these characterizations. It is virtually impossible to find general principles on which all externalists or internalists would agree. But I hope that these characterizations are found plausible by most, if not all.

[15] Free logic is often associated with broadly Fregean semantics about singular terms, whereby either meaning (sense) determines reference or is independent of reference—to the effect that singular terms can be meaningful and yet empty. If there are empty meaningful singular terms in the language, we need a free logic. There are several ways of implementing these broadly Fregean semantics, many of which are obviously related to internalism. See Burge (1974), Larson and Segal (1995), Sainsbury (2001b), and Lehman (2002) for discussions.

https://doi.org/10.1017/S1755020309090091 Published online by Cambridge University Press

here. Prima facie, first-order languages are not adequate to reflect the different ways in which externalists and internalists understand the existential commitments of natural kind nouns; for prima facie, nouns are not singular terms. It is natural here to turn to second-order logic. In section 3, I argue that second-order logic is the right place to look for constructing logics for externalism and internalism about natural kind nouns. I show how issues about existential commitments can be framed in second-order logic.

The paper is organized as follows. In section 2, a general semantic and metasemantic framework is given in which the externalist–internalist debate about the meanings of natural kind nouns and the question of emptiness can be understood. Section 3 relates these characterizations to free logic and second-order logic. Section 4 proceeds to outline and discuss the important features of these externalist and internalist logics. In section 5, these logics are assessed. In an appendix, a systematic exposition of the logics is offered.

§2. Externalism and internalism about natural kind nouns.

2.1. Preliminaries. A few things will be presupposed in the discussion. First, the issue between externalists and internalists only concerns natural kind nouns, that is, *simple* or single-word expressions for natural kinds. Descriptions for natural kinds such as 'transparent liquid that falls as rain' or complex noun-phrases such as 'albino tiger' or 'yellow grass' will not be discussed here. On an externalist account, complex expressions are meaningful insofar as their constituent expressions are; and so externalism only holds indirectly for those complex expressions. Also, although by externalists' standards 'grass' is meaningless if empty, this is not the case for 'yellow grass', which may be meaningful if empty (provided that its constituent parts are meaningful). For instance in a world in which there is grass and the colour yellow but no yellow grass, 'yellow grass' is meaningful. There is no disagreement with internalists about this (although there is a disagreement about the way in which such constituent parts get to be meaningful). So the internalist–externalist debate is only about how to account for the meanings of single-word natural kind nouns.

Secondly, natural kind nouns are nouns, and not, for instance, singular terms; in particular they are not proper names (which are also sometimes called 'proper nouns'). The reasons for this are mainly *syntactic*: unlike proper names, nouns are words that can form noun-phrases, be preceded in English by (in)definite articles and attributive adjectives; semantically, unlike proper names, the same noun may be true of more than one thing.[16]

Thirdly, natural kind nouns are treated here as *referring expressions*, expressions whose function is to refer to natural kinds—their basic *semantic* function is the same as that of singular terms.[17] It is of course not mandatory to do so. Some might think it more natural

[16] Arguably there are uses of some natural kind nouns as singular terms, namely those that are mass terms (such as 'water'). For instance, some philosophers think that in 'water is wet', 'water' functions as a singular term. A discussion of whether mass nouns ever really function as singular terms would lead us too far afield. I will assume here that here they do not. But see Koslicki (1999) for an interesting discussion.

[17] Saying that natural kind nouns are referring expression is by no means failing to recognize that their syntactic role (mentioned above) is different from that of singular terms. For it is not mandatory to tie the semantic property of being a referring expression to specific syntactic properties.

https://doi.org/10.1017/S1755020309090091 Published online by Cambridge University Press

not to do so, and to regard them as expressions whose semantic function is merely to apply or be true of things (i.e. samples of kinds).[18] The reason I treat them as referring expressions is that it is standard to frame the externalist–internalist debate in this way: externalists and internalists agree that natural kind nouns are referring expressions but they disagree about the role reference to natural kinds in the external physical environment plays in individuating their meanings.[19] Nothing substantive in what follows rests on treating natural kind nouns as referring expressions, and logics similar to those given here can be developed in which natural kind nouns are not treated as referring expressions.[20]

Fourthly, I offer no metaphysical account of natural kinds, no explanation of what makes a natural kind a natural kind (e.g. which sorts of properties define it, which lawlike principles are connected to it, etc.), and make no attempt to characterize the sorts of entities that natural kinds might be (e.g. properties, universals, sets, tropes, sui generis kinds, or what have you). I leave it open what the persistence-conditions for natural kinds are, for example whether there are kinds that were never exemplified or whether kinds subsist once they are extinct. What these are partly depends on what sorts of entities natural kinds are. Externalists and internalists need not have a view on these matters: they are not making metaphysical claims, but semantic ones about the conditions under which natural kind nouns refer and are meaningful.

Finally, again, there are two ways in which the notion of an empty natural kind noun can be understood: either empty natural kind nouns are nouns that fail to refer to anything or they are nouns that fail to apply to anything, but refer to empty kinds. It is the former that is at issue between the externalist and the internalist about natural kind nouns.

2.2. A Kaplanian framework. I now outline a framework in which the disagreement between externalists and internalists about the meanings of natural kind nouns and its logical interface can be made precise. It is an extension of David Kaplan's framework, which he developed to account for the semantics and logic of indexicals and demonstratives.[21] It is standard to apply this framework to the externalist–internalist debate; it is also convenient to do so to highlight connections between semantics and logic.

Kaplan defines a *context of utterance* as the parameters (such as time, place, speaker, world, etc.) needed to generate the content of an expression from its character or meaning. The *character* of an expression is the rule or procedure that determines the content of an expression in a given context. It can be characterized as a function from an expression's possible contexts to its content. The *content* of an expression is its extension in a given context of utterance—for example for a sentence it is a proposition, for a proper name it is its referent. The character of a sentence is not part of what is said, and so is not evaluable as true or false, only the content is. A *circumstance of evaluation* is an actual or

[18] See Devitt (2005).

[19] Internalists need not deny that natural kind nouns are referring expressions (e.g. endorse a form of Russellian descriptivism), nor that kinds are really referred to, nor that the kinds referred to are really out there in the external physical environment. Internalists may appeal to descriptive notions to explain the meanings of natural kind nouns, but they do not have to. Many internalists endorse conceptual role semantics, whereby the meaning of an expression is determined by the cognitive role it plays in a speaker's psychology (see Block, 1986). See Fodor (1987) and Chalmers (2002) for other possibilities.

[20] See Besson (Manuscript).

[21] See Kaplan (1989a, 1989b). See in particular Kaplan (1989b, 568ff.) for the framework developed in this subsection.

https://doi.org/10.1017/S1755020309090091 Published online by Cambridge University Press

a counterfactual situation with respect to which we evaluate the contents of sentences as true or false (given the contents of their subsentential parts).

Kaplan uses this framework to distinguish between parametric and non-parametric expressions. Indexicals, for instance, are *parametric* in that they have unfixed or context-sensitive characters: characters that can determine different contents in different contexts. The rule that determines an indexical's referent in a given context is its *semantic character*; it is semantic because it is part of its meaning. For instance the rule for the indexical 'I' is roughly that 'I' always refers to the agent of the context. Although the character of 'I' is not part of the proposition expressed by a sentence in which it occurs, it gives 'I' its meaning. This accounts for the fact that 'I' has a constant meaning outside specific contexts of use—for the fact that it is not just ambiguous. Proper names by contrast are *nonparametric*, they do not have unfixed or context-sensitive characters: once fixed, their contents remain the same in all contexts. Insofar they have characters—insofar as there are rules or procedures (e.g. baptism) that determine their referents—those characters are not semantic, but *metasemantic*. They belong to the possibly nonsemantic facts that explain their basic semantic features. In this case, it is natural to say that the (semantic) character–content distinction collapses—that proper names have no semantic character. Thus a generic proper name such as 'David' has no meaning outside a specific context of use, when it has not been assigned a specific referent, and can be regarded as systematically ambiguous. Whenever it refers to different people, its metasemantic character or its content is different.

The externalist–internalist debate about the meanings of natural kind nouns can be framed in terms of this contrast between parametric and nonparametric expressions. Internalists claim that natural kind nouns have context-sensitive meanings (akin to unfixed semantic characters), while externalists claim that they do not.[22] More precisely, internalists extend Kaplan's account of indexicals to natural kind nouns and claim that just as the indexical 'I' can be used in different contexts with the same meaning to refer to different people, 'water' can be used in different contexts with the same meaning to refer to different natural kinds: 'water' can be used on Earth and Twin-Earth to refer with the same meaning to different liquids.[23] Externalists deny this: if 'water' does not refer to water, it does not have the meaning that it actually has—it has a different metasemantic character or a different content.[24]

2.3. Metasemantics and emptiness. To address the issue of empty natural kind nouns, I now distinguish between semantic and metasemantic contexts. A *metasemantic context* is a context at which an expression is introduced—a context at which its *character* (whether

[22] See Fodor (1987, 44ff.), Chalmers (1996, 56ff., 2002), and Stalnaker (2001, 2003). Putnam (1975, p. 234) once said that natural kind nouns have an unnoticed indexical element—by this he meant that indexicals are used in reference-fixing definitions for natural kind nouns ('this is water'), not that natural kind nouns are parametric in the sense defined here.

[23] Internalists sometimes identify this extended notion of character with so-called 'narrow meaning', and say that on Earth and Twin-Earth 'water' has the same narrow meaning, but different wide meanings (i.e., contents). Externalists tend to think that there is no such thing as a narrow meaning that is the same on Earth and Twin-Earth. See Burge (1979) and Stalnaker (1989, 1990).

[24] The distinctions unfixed versus fixed character and semantic versus metasemantic character are orthogonal. There are four possibilities: (i) unfixed semantic character: for example Kaplan's indexicals; (ii) unfixed metasemantic character: for example ambiguity account of indexicals; (iii) fixed semantic characters: for example causal descriptivism; (iv) fixed metasemantic character: for example Kaplan's account of proper names as directly referential.

https://doi.org/10.1017/S1755020309090091 Published online by Cambridge University Press

semantic or metasemantic) is determined.[25] A *semantic* context is a context at which the *content* of an expression is determined (given its character). For instance, the semantic character of an indexical is determined at a metasemantic context and its content is determined at a semantic context (given its semantic character);[26] the metasemantic character or content of a proper name is determined at a metasemantic context (supposing that proper names are not parametric) and so its content is the same at all semantic contexts. Given this distinction, we can take externalism to entail the claim that if Twin-Earth rather than Earth had been the actual metasemantic context, 'water' would not have meant what it actually does: it would not have had the metasemantic character or content that it actually has. By contrast, internalists claim that if Twin-Earth rather than Earth had been the actual metasemantic context, 'water' would have had the same meaning: it would have had the semantic character or meaning that it actually has. For externalists, 'water' has different meanings (content) at different metasemantic contexts and (given a specified meaning at the metasemantic context) always has the same content at different semantic contexts. For internalists, 'water' has the same meaning at all metasemantic contexts, but may have different contents at different semantic contexts.

It is metasemantic contexts, and not semantic contexts, that are relevant to the question of the meanings of empty natural kind nouns. In the introduction, I said that as far as externalism is concerned, empty natural kind nouns are meaningless. What this means in our framework is that a natural kind noun that is introduced at a metasemantic context at which it is empty is not meaningful: 'water' would not be meaningful if it were introduced at a metasemantic context at which water, or any other relevant natural kind, does not exist (e.g. on Dry-Earth). For internalism empty natural kind nouns are meaningful. What this means is that 'water' can be meaningfully introduced at a metasemantic context at which there is no water, or any other relevant natural kind.

To make it clear that it is metasemantic contexts that are relevant to the issue of emptiness in the externalist–internalist debate, consider the two following situations:

> (α) 'Water' is *introduced* at a *metasemantic context* at which there is no water or no sample of any relevant kind.
>
> (β) 'Water' is *uttered* with its actual meaning (as determined on Earth) at a *semantic context* at which there is no water or no samples of any relevant kind.

Consider the sentence 'There is water' in situation (β) where 'water' has its actual meaning—the situation for instance of an Earthian going for a weekend on Dry-Earth. Externalists and internalists agree that 'water', as it occurs in this sentence in this context, is meaningful. Moreover, they agree that the sentence is not true: for externalists what that sentence says is that there is water (H_2O). This is false. For internalists, what it says is that there is watery stuff (say). This is false.[27] However, externalists and internalists disagree about whether 'water' is meaningful in situation (α)—where the name 'water' is introduced. Externalists claim that 'There is water' is not meaningful because 'water'

[25] Talk of an expression being 'introduced' should be taken to allow for a gradual process, not just a single act of introduction (e.g. baptism or single reference-fixing definition).

[26] Semantic contexts are of course akin to Kaplan's contexts of utterance. I prefer using the notion of a semantic context because it gives a clearer contrast with metasemantic contexts.

[27] I come back to this issue of assignments of truth-value in section 4.3.

https://doi.org/10.1017/S1755020309090091 Published online by Cambridge University Press

is not meaningful: there is nothing in the external physical environment that determines a meaning for it. But internalists claim that 'water' and 'There is water' are meaningful in (α). So the relevant context to theorize about the different externalist and internalist treatments of the meanings of empty natural kind nouns is the metasemantic context.

2.4. Characterizations of externalism and internalism. Given this broadly Kaplanian framework, precise formulations of externalism and internalism can be given. These formulations are given in terms of semantic contexts since externalism and internalism are semantic claims. But it is what goes on at the metasemantic context that explains the nature of these semantic claims. Also, externalism and internalism about natural kind nouns are formulated as contrary positions—externalism, if true, is true of all natural kind nouns, and internalism, if true, is true of all natural kind nouns. This is a natural assumption to make.[28]

Externalism says that if a natural kind noun 'N_k' refers to a kind K at a semantic context, there are no semantic contexts at which 'N_k' means the same as in the original semantic context but does not refer to K. The following is thus a meaning-postulate for externalism:

$$\text{(Externalism)}\ \forall n_k\ \exists K\ \forall C\ \forall K^*(n_k \text{ refers to } K^* \text{ at C } \leftrightarrow K = K^*),$$

where 'n_k' ranges over natural kind nouns, 'K' and 'K^*' range over natural kinds, and 'C' ranges over possible semantic contexts of utterance.

(Externalism) entails that natural kind nouns are so-called *obstinately rigid designators*: a natural kind noun designates the same natural kind at all semantic contexts, even at contexts at which that kind does not exist.[29] Treating natural kind nouns as obstinately rigid enables us to reflect the difference between situation (α) where 'There is water' is not meaningful, and (β) where it is meaningful. Not every construal of rigid designation enables us to do so. For instance, suppose that natural kind nouns are not obstinately rigid, but *persistently rigid*: they refer to the same natural kinds at all semantic contexts at which those kinds exist but do not refer to anything at semantic contexts at which those kinds do not exist. With this account of rigidity, the claim that the sentence 'There is water' is meaningful in situation (α) but is not in situation (β) cannot be expressed adequately. Consider (β): if 'water' is persistently rigid and does not refer at semantic contexts at which there is no water, there will be no semantic context at which 'water' is meaningful and there is no water. At those contexts 'water' will fail to refer and to be meaningful; thus, 'There is water' will not be meaningful. That is, if 'water' is persistently rigid, the externalist cannot consider semantic contexts at which there is no water using 'water' with its actual meaning. But intuitively this is something that she should be able to do, and which she can do if natural kind nouns are obstinately rigid.

Internalism says that if 'N_k' refers to K at a given semantic context C, there are semantic contexts at which 'N_k' means the same as it does at C, but does not refer to K. That is, a suitable meaning-postulate for internalism is the following:

$$\text{(Internalism)}\ \forall n_k\ \forall K\ \exists C\ \exists K^* \neg(n_k \text{ refers to } K^* \text{ at } C\ \leftrightarrow K = K^*).$$

(Internalism) entails that natural kind nouns need not refer to unique natural kinds at all semantic contexts: either they refer to different natural kinds at different semantic contexts

[28] In principle they could both be false and there could be mixed positions; these are ignored here, and in effect externalism and internalism are treated as contradictories.

[29] See Salmon (2005, pp. 33–34) for the contrast between the different types of rigidity mentioned in this section.

https://doi.org/10.1017/S1755020309090091 Published online by Cambridge University Press

or they do not refer to anything at some semantic contexts. (Internalism) does not entail any form of rigidity.

§3. Logics.

3.1. Metasemantics and logical truth. The notion of logical truth is introduced as follows: count as a logical truth a sentence that is *true at all metasemantic contexts*. More precisely, a (meaningful) sentence containing a (meaningful) natural kind noun is a logical truth if it is true at all metasemantic contexts at which a meaning is determined for that natural kind noun (as well as for its other constituents). Let me explain. Consider the claim (Externalism), which says that a natural kind noun refers to the same unique natural kind at all semantic contexts. The rationale for this claim is that externalists require that once a natural kind noun refers to a given nonempty natural kind at the metasemantic context, it refers to it at all semantic contexts. Thus consider the following claims:

$$(I_{\mathrm{Nk}})\ \exists K\ K = N_k$$

$$(E_{\mathrm{Nk}})\ \forall K\ \exists x\ \ (x \text{ is } K)$$

(I_{Nk}) says that some kind K is N_k; (E_{Nk}) says that all natural kinds K are exemplified. For externalists (I_{Nk}) and (E_{Nk}) are true at all metasemantic contexts, and this is why instances of (I_{Nk}) and (E_{Nk}) are logical truths in the logic for (Externalism).

Consider (I_{Nk}) and the noun 'water'. Suppose that we are at the actual metasemantic context—Earth. At this context the instance of (I_{Nk}) 'Some kind is water' is true because there is a nonempty natural kind (water) which 'water' refers to. Now suppose that the actual metasemantic context is Twin-Earth where 'water' does not refer to water, but to twater. Given externalism, since 'water' refers to different natural kinds on Earth and Twin-Earth, it means something different on Earth and Twin-Earth. Still, 'water' is a meaningful natural kind noun on Twin-Earth because, as on Earth, there is a nonempty natural kind there (twater) to which it refers. The reason why 'water' is meaningful on Twin-Earth is that the relevant instance of (I_{Nk}) is true. Suppose finally that 'water' is introduced at a metasemantic context at which there is nothing which the noun refers to. At that context 'water' is not meaningful and 'Something is water' is not a meaningful sentence. Here, we say that it is *not an instance* of (I_{Nk}). Indeed, if 'Something is water' is not meaningful at that metasemantic context, that means that that context is not a metasemantic context at which a meaning is determined for 'water'. That is, it is not a metasemantic context that is quantified over in our account of logical truth as truth at all metasemantic contexts.[30]

Here, the direction of explanation between semantics and metasemantics is two-way. First, from metasemantics to semantics. The claim that 'water' means something different

[30] Saying that 'Something is water' is true at all metasemantic contexts, that is, true no matter what meaning is assigned to 'water' at that context, presupposes that the same natural kind noun (type) is reinterpretable. An externalist might think that natural kind nouns are not reinterpretable, and think of nouns as pairs of physical types and interpretations—as metasemantically individuated as it were. In this case she will not think of truth at all metasemantic contexts in terms of truth no matter which interpretation we give to a natural kind noun, but truth no matter what natural kind noun gets substituted for another. This view could be accommodated without requiring any substantive change to the way logical truth has been characterized here. But note that there might be problems with opting for substitutional account of logical truth. Here see Tarski (1936) for discussion.

https://doi.org/10.1017/S1755020309090091 Published online by Cambridge University Press

whether Earth and Twin-Earth are the metasemantic contexts *justifies* (the relevant instance of) the claim (Externalism) that 'water' refers to water at all semantic contexts. In a slogan: at all metasemantic contexts: 'N_k' refers to K at all semantic contexts. Secondly, from semantics to metasemantics. When we consider empty natural kind nouns, we have to come back to the metasemantic context given the way in which the semantic claim (Externalism) is framed: a natural kind term refers to a unique natural kind at all semantic contexts. Thus given (Externalism), 'water' is meaningless at metasemantic contexts at which it is empty.

Consider now internalism: (Internalism) allows natural kind nouns to be meaningfully introduced at metasemantic contexts at which they are empty. For example, there may be a metasemantic context (e.g. Dry-Earth) at which 'water' is meaningfully introduced, while no relevant natural kind exists. At that context, the relevant instance of (I_{Nk}) is false. Thus for internalists, instances of (I_{Nk}) need not be true at all metasemantic contexts: (I_{Nk}) is not logically true. Instances of (E_{Nk}) are not true at all metasemantic contexts either: 'water' may be meaningfully introduced on Dry-Earth.

We can thus give a precise content to the claim made in introduction that externalists provide existentially committed accounts of the meanings of natural kind nouns, and internalists do not: (Externalism) requires instances of (I_{Nk}) and (E_{Nk}) to be true at all metasemantic contexts, (Internalism) does not.

Concerning the status of logical truths as truths at all metasemantic contexts, note first that a sentence may be logically true while not true at all semantic contexts—for example for externalists (I_{Nk}) and (E_{Nk}) need not be true at all semantic contexts.[31] Secondly, logical truths need not be necessarily true: it is contingent that water exists, so it is contingent that something answers to the name 'water' at the metasemantic context at which it is introduced.

This second feature is reminiscent of Kaplan's account of logical truth. In Kaplan's logic for indexicals, the sentence 'I am here now' is logically true but not necessarily true. It is logically true because, given his semantics of indexicals, the meanings (semantic characters) of 'I', 'here', and 'now' (and 'am'), 'I am here now' is true at every semantic context (context of utterance).[32] There is a similarity between our reasons for counting a sentence as logically true and Kaplan's, but also a substantive difference. The similarity is that logical truth is conceived as truth given the sort of semantic and metasemantic principles that govern certain expressions in the language. Again, on this conception, semantics determine logic. The difference is this. The sentence 'I am here now' is logically true because it is *true at all semantic contexts*. But it is not true at all metasemantic contexts. For instance it is not true at a metasemantic context at which 'now' means what 'yesterday' means in English—that is at a metasemantic context at which the character of 'now' is different. An externalist logic regards the sentence 'Something is water' as logically true because it is *true at all metasemantic contexts*–at all contexts at which a meaning is determined for 'water'. But it is not true at all semantic contexts; there are semantic contexts at which nothing (relevant) is water (as in situation (β) of section 2.2)).

[31] For externalists (and internalists) instances of (I_{Nk}) and (E_{Nk}) need not be true at all semantic contexts—for instance at contexts at which there are no samples of a given kind. Whether instances of (I_{Nk}) are true at all semantic contexts depends on whether natural kinds exist at semantic contexts at which they are empty. Again (see section 2.1), externalists (and internalists) need no take a view on whether (I_{Nk}) is true at all semantic contexts. What matters for the logic for (Externalism) is that (I_{Nk}) and (E_{Nk}) are true at all metasemantic contexts.

[32] See Kaplan (1989a, 508ff.).

https://doi.org/10.1017/S1755020309090091 Published online by Cambridge University Press

So while Kaplan generalizes over semantic contexts, we generalize over metasemantic contexts: contexts at which natural kind nouns are endowed with meanings.[33]

3.2. Free logic. I now turn to *universal instantiation* and show how the logics for (Externalism) and (Internalism) respectively validate and invalidate a version of this principle for natural kind nouns. (Externalism) requires a logic that is existentially committed with respect to natural kind nouns, (Internalism) does not. These differences are crucial when we consider a version of universal instantiation which involves natural kind nouns; for it fails if meaningful natural kind nouns are allowed to be empty.

Consider the following principle of universal instantiation:

$(\mathrm{UI_{Nk}})\ \forall K\ \Phi(K) \rightarrow \Phi(N_k)$, where N_k is free for Φ and N_k is a natural kind noun.

($\mathrm{UI}_{\mathrm{N}k}$) says that if condition Φ applies to K, it applies to N_k. ($\mathrm{UI_{Nk}}$) is logically true in a logic for (Externalism): since 'N_k' is a natural kind noun, it refers to a natural kind, and since the universal quantifier ranges over all natural kinds, K is in the domain. Consider (2) and the actual metasemantic context:

(1) $\forall K \Phi(K) \rightarrow \Phi$ (phlogiston).

By externalist standards 'Φ(phlogiston)' is not meaningful because 'phlogiston' is not meaningful: (1) is not an instance of ($\mathrm{UI_{Nk}}$) in the logic for (Externalism).

As far as internalism is concerned empty natural kind nouns are meaningful, and so ($\mathrm{UI_{Nk}}$) is not logically true in the logic for (Internalism): the antecedent of ($\mathrm{UI_{Nk}}$) may be true but its consequent false. In particular, at the actual metasemantic context, the antecedent of (1) may be true but its consequent false—every kind in the domain of the quantifiers may satisfy Φ, while phlogiston does not satisfy Φ, because phlogiston is not in the domain.

Thus, a logic for (Externalism) regards universal instantiation as logically valid and a logic for (Internalism) does not. The failure of universal instantiation is the cornerstone of free logic: (Internalism) requires a free logic.

3.3. Second-order logic. (Externalism), (Internalism), (I_{Nk}), (E_{Nk}), and ($\mathrm{UI_{Nk}}$) all contain quantifiers that range over natural kinds. I will now argue that these quantifiers are second-order quantifiers.

Although second-order logics are not designed to be logics for natural kind nouns, they are intuitively appropriate here. First, questions concerning existential commitments are closely related to the truth of certain second-order principles (e.g. universal instantiation, existential generalization, comprehension principle), which are the sorts of principle in which we are interested. Secondly, a natural view of logic is that it aims to codify the logical behavior of natural languages' expressions given their meaning assignments. Thus a natural outlook on second-order logic is that it aims to codify the logical behavior of predicates in natural languages given their meaning assignments. And this is what has

[33] If logical truth were truth at all semantic contexts, (Externalism) would require a free logic. If (Externalism) treats natural kind nouns as obstinately rigid designators and does not require (I_{Nk}) to be true at all semantic contexts, that means that there may be contexts at which 'water' has a referent but that referent does not exist. That is, at those semantic contexts, the fact that 'water' refers does not entail that water exists. This would be an interesting if somewhat esoteric way of generating free logic.

https://doi.org/10.1017/S1755020309090091 Published online by Cambridge University Press

been done here for natural kind nouns: looking at their logical behavior, given the sorts of meanings they are assigned by different accounts of their meanings.[34]

There are two basic ways of characterizing second-order logic:

(i) Second-order logic is the *logic of (first-order) predicate position*: it is the result of adding quantifiers with respect to predicate position to the first-order quantifiers.[35] Thus quantification into the position of 'is a horse' in (2) is second-order quantification:

(2) Shergar is a horse.

The problem with (i) is that intuitively a logic for natural kind nouns should quantify into *noun position* rather than predicate position; by (i), that logic is not second-order. One possibility here would simply be for us to instead quantify into predicate position, where predicates are formed with nouns. And it would be easy to form predicates with natural kind nouns, since to every natural kind noun there corresponds a predicate. Also, focusing on predicates rather than nouns would allow us straightforwardly to apply the symbolism of second-order logic to them: issues about existential commitments as they are expressed in standard second-order languages could be applied to the special case of natural kind predicates.[36]

However, there are important differences between nouns and predicates, and focusing on predicates would raise further complicated issues for us. Consider (Externalism), where quantification is into natural kind noun position. It may not be appropriate for us to use quantifiers into predicate position, because it is unclear that predicates refer and, if they do, it is unclear that they refer to natural kinds.[37] Also some philosophers have argued that nouns cannot be assimilated to predicates on the grounds that, unlike predicates, nouns carry with them principles of (sortal) identity.[38] So taking natural kind predicates as the subject matter of our logics may look appealing because of the convenience of using the symbolism of second-order logic. But it may also make us change philosophical topic. It is better to stick to nouns.

---

[34] There already exists several logics for nouns that are first- or second-order. See for instance Gupta (1980), Roeper (2004), and Cocchiarella (2001, 2007). I cannot discuss them here.

[35] See Bostock (1997, p. 34) and Sainsbury (2001a, p. 236).

[36] It would also alleviate a possible worry about quantification into noun position which concerns negation: negation is often said to go with predicates and not with nouns (or constituent expressions of predicate). The worry here is that those expressions we quantify over (nouns) are not those expressions we negate (predicates). However, it is not obvious that we ought to say that we quantify over what we negate (i.e. that if we have quantification into noun position, we should have noun negation). In the first-order case, there is certainly no temptation to say that. It may be easier to deal unilaterally with predicates, but there is no real tension in not doing so. Also, it is not obviously true that negation goes with the predicate: in standard logic, negation is a sentential operator, and as far as natural language is concerned, things are not so clear. Suppose that we want to know what Shergar is. It seems that a good answer to the question 'What is Shergar?' could be 'not a man' or 'a horse, not a man', where the negation seems to be attached to the noun. Negation is complicated and I cannot do it justice here.

[37] For instance Wiggins (1984, p. 134) claims that predicates such as 'is a man' have no reference, because proper or substantial reference has already dropped out when we reach complex expressions such as predicates. Predicates have semantic values (make a contribution to truth-conditions) which are determined in terms of their constituent parts, but they do not have referents.

[38] See for instance Gupta (1980, 20ff.) and Dummett (1981, 212ff).

https://doi.org/10.1017/S1755020309090091 Published online by Cambridge University Press

Also, since the issue here is the way natural kind nouns work in *natural language*, we should focus on nouns because in natural language the relevant sort of quantification is into noun position rather than predicate position. For instance, from (2) above, we can infer (3) but not (4):

(3) Shergar is something.

(4) Shergar something.

(4) is ill-formed.[39] In (3), quantification is into noun position.[40] From this standpoint, quantification into predicate position is an artifact of the symbolism of second-order logic. So it would be wrong to insist that a logic has to quantify into predicate position, as opposed to, for instance, noun position to count as second-order. What seems to be the case is that talk of quantification into predicate position in natural language is just a convenient general way of talking about quantification into the position of what may be called 'predicative expressions', that is expressions that can typically be used to form first-order simple predicates (e.g. nouns, adjective, verbs, etc.). If that is so, quantification into noun (adjective, verb, etc.) position should after all count as second-order quantification.

Condition (i) provides a sufficient condition for second-order logic, but I consider a second condition that is often set as a requirement for second-order logic.

(ii) Second-order logic is obtained when *the range of the second-order variables is determined by that of the first-order ones*—when the first-order domain determines both the ranges of the first- and second-order quantifiers.

For instance, a 'standard' semantics for second-order logic uses only the first-order domain for the assignments of values to the first- and second-order variables—second-order variables range over subsets of the domain of the first-order variables.[41]

With quantification over kinds, there is a straightforward way in which the first-order domain determines the second-order one. For intuitively, what objects there are in the first-order domain—in particular the objects that are members of natural kinds—determines what kinds there are: if we have all the objects, we have all the natural kinds. Let us say that *the kind(s) of the domain D* or *the kind(s) of D*, is (are) the kind(s) determined by the objects in the first-order domain *D*. Now, suppose that we want to restrict our attention to the domain of horses. That domain is our first-order domain *D*; then, the kinds of *D* are the kinds horses belong to, for example the kind horse, mammal, animal, and so forth. If the domain *D* is of horses and lilies, the kinds of *D* are the kinds horse (etc.) and lily (etc.). In this way, the domain of the quantifiers that range over kinds is determined by that over which the first-order quantifiers range. Strictly speaking, a domain of kinds need not be specified.[42]

Given that quantification into natural kind noun position satisfies suitably understood versions of conditions (i) and (ii), the logics for (Externalism) and (Internalism) can be

[39] I do not consider the possibility of inferring 'Shergar something*s*' from (2). This really stretches English; and it is unclear that 'somethings' is not just elliptical for 'is something'.

[40] See again Wiggins (1984, p. 133).

[41] Shapiro (1991, p. 73) and Cocchiarella (1993, 456).

[42] This characterization of natural kinds only works for nonempty kinds. Empty kinds would not be determined by the first-order domain. Again, empty kinds are not directly relevant to our discussion of externalism and internalism and it is not required to have them in the domain of the quantifiers.

https://doi.org/10.1017/S1755020309090091 Published online by Cambridge University Press

regarded as second-order;[43] in particular, (UIN_k) of section 3.2) and similar principles given below are second-order principles.

§4. Elements of the logics for (externalism) and for (internalism). In this section, I outline the main elements of the logics for (Externalism) and for (Internalism).[44]

4.1. Second-order logic for (Externalism). In a logic for (Externalism),
(I_{Nk}) $\exists K\, K = N_k$ holds, and a second-order principle of universal instantiation is validated:

($\mathrm{UI}_{\mathrm{Nk}}$) $\forall K\ \Phi(K) \rightarrow \Phi(N_k)$, where N_k is free for Φ.

The logic also validates a second-order principle of existential generalization (EG) and a comprehension principle (CP):

($\mathrm{EG}_{\mathrm{Nk}}$) $\Phi(N_k) \rightarrow \exists K\ \Phi(K)$, where N_k is free for Φ

($\mathrm{CP}_{\mathrm{Nk}}$) $\exists K\ \forall x\ (x \text{ is } K \leftrightarrow x \text{ is } N_k)$.

($\mathrm{EG}_{\mathrm{Nk}}$) says that if the kind N_k is Φ, then some K kind is Φ. And ($\mathrm{CP}_{\mathrm{Nk}}$) says that every natural kind noun 'N_k' refers to a natural kind. From the obvious fact that $\forall\, x\ (x \text{ is } N_k \leftrightarrow x \text{ is } N_k)$ and ($\mathrm{EG}_{\mathrm{Nk}}$), we can get a principle that guarantees that there exists a kind of the domain D over which noun variables range:

(D_{Nk}) $\exists K\ \forall x\ (x \text{ is } K \leftrightarrow x is K)$.

Also, given ($\mathrm{UI}_{\mathrm{Nk}}$), ($\mathrm{EG}_{\mathrm{Nk}}$), by transitivity, we can derive:

($\mathrm{UE}_{\mathrm{Nk}}$) $\forall K\ \Phi(K) \rightarrow \exists K\ \Phi(K)$.

The logic for (Externalism) is *standard* or *classical* insofar as it validates all classical principles and inferences. However, the adjunction of (E_{Nk}) as a logical truth is a significant departure:

(E_{Nk}) $\forall K\ \exists x\ (x \text{ is } K)$.

(E_{Nk}) is the rejection of empty kinds and its presence ensures that all quantification over kinds and other reference to kinds is to nonempty kinds. It should be remembered here that the logic is only concerned with natural kinds as the values of (simple) natural kind nouns that can serve to form monadic predicates together with the copula. If we were considering syntactically complex nouns – noun-phrase – it would not be appropriate to require them to refer to exemplified natural kinds: for instance we surely want 'transparent liquid that falls as rain' or 'yellow grass' to be successfully introduced at metasemantic contexts at which they do not have samples.[45]

---

[43] I do not discuss the possibility of characterizing second-order logic by saying that the *range* of the second-order quantifiers is different from that of the first-order ones, which range over objects. Given that opinions differ widely as to what sorts of entities are quantified over in second-order logic (sets, properties, universals, concepts, tropes, pluralities, or what have you), a characterization of second-order logic in terms of any of these entities is bound to be too contentious to be helpful. Also, the standard interpretation of second-order logic has the second-order quantifiers ranging over sets, which are objects, and so would fail if we used this as a characterization of second-order logic. See Shapiro (1991, p. 13).

[44] See the Appendix for a systematic exposition of these logics.

[45] Externalism and internalism are claims about single word expressions, natural kind nouns (see again section 2.1), and their logics only contain natural kind nouns. This is why (E_{Nk}) holds for externalists. But for them the semantically complex noun-phrase 'transparent, colorless,

https://doi.org/10.1017/S1755020309090091 Published online by Cambridge University Press

Now, in the metalinguistic claims (Externalism) and (Internalism), and in principle (I_{N_k}), the identity sign is used.[46] And it is natural for the logics of (Externalism) to have versions of the principle of self-identity and of Leibniz's Law:

$$(\mathrm{Id.}_{\mathrm{Nk}})\ \forall K\, K = K$$

$$(\mathrm{LL}_{\mathrm{Nk}})\ N_{k1} = N_{k2} \rightarrow (\Phi(N_{k1}) \rightarrow \Phi(N_{k2})).$$

This also enables us to express (D_{N_k}) in the following way:

$$(D^{*}_{\mathrm{Nk}})\ \exists K\ K = K.$$

4.2. Second-order free logic for (Internalism). A logic for (Internalism) does not validate (I_{Nk}), and thus ($\mathrm{UI}_{\mathrm{Nk}}$) and ($\mathrm{EG}_{\mathrm{Nk}}$) are not validated either. This means that the logic for (Internalism) is *free*. Also, the logic does not require that there are natural kinds at all (nonempty or empty). Thus (D_{Nk}), which asserts the existence of natural kinds, is not validated. This means that the logic is a *universally free* logic—a free logic in which the domain of the quantifiers may be empty. If so, ($\mathrm{UE}_{\mathrm{Nk}}$) cannot be derived in it.

($\mathrm{UI}_{\mathrm{Nk}}$) has to be weakened in order to allow for cases in which 'N_k' is empty. In first-order free logic, the standard way of weakening universal instantiation consists in adding a clause ('$\exists x\, x = t$') which requires the existence of referents for singular terms. Here, the following naturally suggests itself:

$$(\mathrm{UI}_{NkF})\ \forall K\ \Phi(K) \rightarrow (\exists K\, K = N_k \rightarrow \Phi(N_k)),\ \text{where } N_k \text{ is free for } \Phi.$$

It is also natural to have the same sort of restriction for ($\mathrm{EG}_{\mathrm{Nk}}$) and ($\mathrm{CP}_{\mathrm{Nk}}$)—($\mathrm{CP}_{\mathrm{Nk}}$) says that for every natural kind noun, there is a natural kind and this need not be true from an internalist perspective:

$$(\mathrm{EG}_{NkF})\ \Phi(N_k) \rightarrow (\exists K\, K = N_k \rightarrow \exists K\ \Phi(K)),\ \text{where } N_k \text{ is free for } \Phi.$$

$$(\mathrm{CP}_{NkF})\ \exists K\, K = N_k \rightarrow (\exists K\ \forall x\ (x \text{ is } K \leftrightarrow x \text{ is } N_k)), \text{where } N_k \text{ is free for } \Phi.$$

The logic for (Internalism) validates the principle of identity ($\mathrm{Id}_{\mathrm{Nk}}$) and Leibniz's Law ($\mathrm{LL}_{\mathrm{Nk}}$). But given that the logic is universally free, it does not validate (D^{*}_{Nk}) and (I_{Nk}). Finally, concerning (E_{Nk}), the restricted internalist principles listed above require neither that natural kind nouns refer nor that there are objects in the first-order domain. And it is natural for an internalist logic to allow the first-order domain to be empty—and so (E_{Nk}) is not validated.

---

liquid …' could be successfully introduced at a metasemantic context at which there are no samples of water (provided that its constituent expressions are meaningful). Thus if our externalist logic also contained natural kind noun-phrases, (E_{Nk}) would not hold in them, and we would need different styles of natural kind variables (and values) indicating the sorts of expressions (e.g. simple natural kind noun versus complex noun-phrase) which they stand for: we would have to distinguish between a natural kind as referred to by a simple natural kind noun and one as referred to by a complex noun-phrase.

[46] One possible worry about having the identity-sign here is that for instance '(a) person is identical with (a) human' might not be grammatical (c.f. section 2.1). We should perhaps rather write 'the kind person is identical with the kind human', where we flank the expression for identity with expressions that have the syntax of singular terms. On the other hand, it seems that if it is appropriate to treat natural kind nouns as singularly referring to kinds, it must be in order to make identity statements involving such nouns. Also, it is desirable to be able to express things such as 'Water = H_2O' or 'Water ≠ twater' or '(The kind) horse = (the kind) *equus caballus*'.

https://doi.org/10.1017/S1755020309090091 Published online by Cambridge University Press

4.3. Semantics. The semantics for the second-order logic for (Externalism) is straightforward. If only nonempty natural kind nouns are allowed in the logic, sentences of the form:

(5) $\Psi(N_k)$

are true just in case condition 'Ψ' holds of the kind referred to by 'N_k'. For instance, the sentence 'Shergar is a horse' is true if 'Shergar is' holds of the referent of 'horse'. (For convenience, it is assumed here that the proper name 'Shergar' is meaningful only if it is nonempty.) Thus 'Shergar is a horse' is true just in case the referent of 'Shergar' is (the 'is' of predication) the referent of 'horse'. Accordingly, 'Shergar is a horse' is false just in case the referent of 'Shergar' is not the referent of 'horse'.

As for the semantics for the logic for (Internalism), there is a plethora of semantics we can choose between when it comes to free logic: actualist or nonactualist, negative, positive or neutral semantics, and so forth.[47] It is natural here to go for a negative semantics: a semantics in which no referents are assigned to empty natural kind nouns and in which atomic sentences containing empty natural kind nouns are false.[48]

A general reason in favor of this semantics is that it is a well-developed bivalent semantics for free logic. According to it, the sentence 'Something is water' is false on Dry-Earth—whether it is considered as a metasemantic or a semantic context. Its negation is true. There is a lot to say in favor of bivalent semantics—for instance, it is consistent with a Tarskian-style theory of truth.[49] That said, there is no intrinsic reason for an internalist about the meanings of natural kind nouns to adopt bivalence—this is not motivated by internalism itself, but by other theoretical preferences.

A specific reason is that (Internalism) does not require natural kind nouns to refer to kinds to be meaningful, and so there is no intrinsic reason to stipulate referents of sorts for empty natural kind nouns. So semantics for free logic that assign referents of sorts to empty natural kind nouns (e.g. nonexistent kinds, the empty set, or what have you) are not natural candidates here.

According to standard first-order negative free logic (8) is false:

(6) Pegasus is a horse.

Because Pegasus does not exist, the predicate 'is a horse' is not true of the referent of 'Pegasus'. Thus (6) is not true, and, given bivalence, it is false.

It is straightforward to extend this negative semantics to the logic for (Externalism): an atomic sentence that contains an empty natural kind noun is false. Thus for instance,

(7) Shergar is a unicorn,

is false because there are no unicorns. The relevant principle that connects the truth of an atomic sentence containing a natural kind noun with the existence of a natural kind that is the bearer of that noun can be formulated as follows:[50]

---

47 See Lehmann (2002) for a review of these semantics.

48 See Burge (1974).

49 Preserving bivalence may be crucial for the meaning-theorist. If she is working with a Tarskian theory of truth (or at least uses conventional disquotational truth-schemas) as many meaning-theorists (with different outlooks on the nature of meaning) are, denying bivalence might lead to inconsistencies. See Williamson (1992, 265–266) for this point.

50 See Burge (1974) for a first-order analogue.

https://doi.org/10.1017/S1755020309090091 Published online by Cambridge University Press

$$\text{(TC)}\ \Psi(N_{k1}\ldots N_{kn}) \rightarrow (\exists K_1\ K_1 = N_{k1} \& \ldots \& \exists K_n\ K_n = N_{kn}),$$

where Ψ is any atomic predicate, including identity.

(TC) connects the truth of an atomic sentence containing a natural kind noun with the existence of a kind that is the bearer of that noun. If no such kind exists, then, given bivalence, the sentence is false.

One application of (TC) yields that (8) is false because 'phlogiston' is empty:

(8) Phlogiston = oxygen.

§5. Discussion. This paper presents a new way of motivating free logic, namely internalism about natural kind nouns, and a new way of setting up free logic, within a second-order framework. I know of only one other development of a second-order free logic for natural kind nouns (or rather, simple and complex natural property predicates), namely Cocchiarella's logic for conceptual natural realism.[51] It is a second-order logic in that it contains quantifiers quantifying into predicate position; the predicate variables range over what he calls 'natural properties and relations', some of which are intended to be things like natural kinds. The logic is universally free in that the domain of the second-order quantifiers may be empty and it is free of existential presuppositions regarding predicate constants and variables.

Although there are many differences between Cocchiarella's logic for conceptual natural realism and that for (Internalism), mainly due to the fact that his system is much richer (e.g. it includes relations, and contains modal operators), crucially, neither of them regards universal instantiation as a valid principle, and both of them impose restrictions on that principle. Also, he characterizes his system as an ontological system, whereas the logics for externalism and internalism are determined by semantic theories.

Cocchiarella's reasons for thinking that a logic whose subject matter is natural properties and relations should be free are epistemological, and not semantic and metasemantic. Cocchiarella thinks that it is an 'empirical matter' whether there are any natural properties, and so a natural predicate (a predicate that purports to name a natural property) cannot just be assumed to have a natural property as value; it may fail to refer to anything. To be meaningful, it is not required that a natural property predicate has a natural property (or relation) as referent: the existence of a natural property is not the *semantic ground* for natural kind predication (although successful predication requires the existence of such a property). An account of what makes such predication meaningful is to be found somewhere else and Cocchiarella finds it in a form of conceptualism: natural property predicates express predicable concepts which may be vacuous. Those concepts are identified as the rule-following cognitive capacities that underlie our use of referential and predicable expressions.[52]

What Cocchiarella puts forward is a version of a standard argument that free logicians give against first-order classical logic: classical logic is at odds with the traditional idea

[51] See Cocchiarella (1986, 1993, 2007, esp. chapter 12). His logic for conceptual natural realism is a subsystem of a general system of conceptual realism: the general system is not free, only the subsystem is.

[52] See Cocchiarella (1983, p. 315, 2007, chapter 12).

https://doi.org/10.1017/S1755020309090091 Published online by Cambridge University Press

that logic is a priori.[53] Classical logic rests on the assumption that singular terms refer to things in the domain, and thus regards the following as a priori:

(9) $\forall n(n$ is a meaningful singular term $\rightarrow \exists x(t$ refers to $x))$.

In the case of natural object and properties, Cocchiarella's argument is that something like (10) is not a priori—(10) is false if empty natural kind nouns are allowed in the language:

(10) $\forall n_k(n_k$ is a meaningful natural kind noun $\rightarrow \exists K(n_k$ refers to $K))$.

If logic is a priori, and in particular logical relations such as validity are a priori, then it would seem that a logic relying on (10) demands that we know a priori that any given meaningful natural kind noun in the language refers to an exemplified kind. But intuitively this is not something that can be known a priori.

This is not the place to review in detail Cocchiarella's conceptualism and how it might connect with some of the things internalists claim about the nature of meaning. But it is worth stressing that Cocchiarella's argument for free logic is very similar to one sort of epistemological argument put forward by internalists. (Externalism) is, in Cocchiarella's terms, a theory in which the existence of a natural kind is the semantic ground for correct natural kind predication. That means that (10) is a truth of externalism and is thus a priori (if (Externalism) is); (Internalism) is a theory in which the existence of a natural kind is not the semantic ground for correct natural kind predication. And many philosophers have argued against externalism on the ground that (10) should not be regarded as true, let alone a priori. If (10) is a priori, externalists allegedly find themselves open to incompatibilist arguments that are designed to show the incompatibility of externalism and privileged access, such as:[54]

(i) 'water' is a meaningful natural kind noun $\rightarrow \exists K$ 'water' refers to K.
(ii) 'water' is a meaningful natural kind noun.
(iii) $\exists K$ 'water' refers to K.

(i) is an instance of (10)—it is a priori for externalists. (ii) is said to be a priori because of access: an expression's semantic properties are a priori. Thus, by detachment, (iii) ought to be a priori, but intuitively it is not. Accordingly, it seems that externalists should deny that (ii) is a priori. Maybe doing this is something that they can live with, and maybe there are other ways out of the problem.[55]

What is important to see here is that one epistemological problem facing externalism is the same one that faces classical logic: it is a problem of incompatibility between the existential presuppositions required by a semantic theory or a logic, and a priori knowledge. So philosophers who think that such epistemological problems are insuperable should turn to free logic.

Leaving aside issues concerning compatibilism and free logic, another question concerns how generally we should think of the status of some of the logical truths of the logic for (Externalism), in particular (I_{Nk}): $\exists K\, K = N_k$. It is of course surprising that (I_{Nk}) should be a logical truth, if we think in terms of the standard canons of logical truth: (I_{Nk}) does

[53] See Bostock (1997, 351ff.) and Sainsbury (2001a, p. 240) for discussion.
[54] See McKinsey (1991).
[55] See Brewer (2000), Davies (2000), and McLaughlin and Tye (1998) for externalist discussions.

https://doi.org/10.1017/S1755020309090091 Published online by Cambridge University Press

not have many of the properties that are usually taken to be essential to logical truth, such as those of a priority, necessity, and perhaps analyticity. So should we really take the idea that (E_{Nk}) is a logical truth seriously? As we have seen in section 3.1, there are other sorts of sentences that have a legitimate claim to be regarded as logical truths because of the way natural language works—for instance Kaplanian sentences such as 'I am here now'. And they too do not fit the canons of logical truth.

One might simply deny that externalist and Kaplanian logical truths are genuine logical truths. But to do that, we would need some principled reason for demarcating sentences such as (E_{Nk}) from standard logical truths; we cannot arbitrarily declare that (E_{Nk}) is not a genuine logical truth. Also, we could not do this without undermining externalism: if a logic that validates (E_{Nk}) should not count as logic, the semantics that generates such a logic cannot be right. Internalists could argue against externalists in this way—in the same way, that is, in which free logicians argue that classical logic cannot be the right logic because of its existential commitments, and that we should not endorse the semantics that generates such a logic.

Another possibility is to reassess our canons of logical truth, which might lead us to say that there are logical truths which do not have some of the traditional properties associated with logical truth—and (E_{Nk}), or a Kaplanian logical truth could be examples of those. Such an assessment was not the aim of this paper. But the hope is that it has been shown how, if (Externalism) is true, (E_{Nk}) is a genuine contender for logical truth, because it is motivated by very general, and plausible externalist semantic and metasemantic principles. If it is true that externalism is committed to the logical truth of (E_{Nk}), it is also true that externalism should make us rethink what our conception of logical truth should be in the same way as it made us rethink our conception of meaning.

§6. Acknowledgments. Thanks to Thomas Kroedel, Hemdat Lerman, Bruno Whittle, Timothy Williamson, Andrew Hudson, and two anonymous referees for very helpful comments on this paper.

Appendix: Second-order logics for (Externalism) and (Internalism)

(i) Language

Definition of a second-order language L_{Nk}, for a logic for natural kind nouns.

The logical vocabulary consists of the connectives: material implication $\rightarrow$ and negation $\neg$, the universal quantifier $\forall$, and a sign for the binary relation of identity $=$; as well as brackets (,) and variables:

$$x_1, x_2, \ldots, x_i, \ldots \text{ are object variables}$$

$$K_1, K_2, \ldots, K_i, \ldots \text{ are natural kind variables}$$

Definitions:

Given material implication $\rightarrow$ and negation $\neg$, the other connectives are introduced as abbreviations as follows:

$$\Phi \vee \Psi : \neg\Phi \rightarrow \Psi \qquad \text{(disjunction)}$$

$$\Phi \,\&\, \Psi : \neg(\Psi \rightarrow \neg\Psi) \qquad \text{(conjunction)}$$

$$\Phi \leftrightarrow \Psi : (\Phi \rightarrow \Psi)\&(\Psi \rightarrow \Phi) \qquad \text{(material equivalence)}$$

https://doi.org/10.1017/S1755020309090091 Published online by Cambridge University Press

The existential quantifiers are defined as follows:

$$\exists x \Phi : \neg \forall x \neg \Phi$$

$$\exists K \Phi : \neg \forall K \neg \Phi$$

Nonlogical vocabulary:

$n_1, n_2, \ldots, n_n$ are object names

$N_{k1}, N_{k2}, \ldots, N_{kn}$ are natural kind nouns

Metavariables:

Object names and variables are object terms: $t_1, t_2 \ldots, t_i, \ldots$

Natural kind nouns and variables are natural kind terms: $T_1, T_2, \ldots, T_i, \ldots$

The second-order quantifiers should be understood as quantifying into noun position. The following translation scheme between (quasi) natural language and the formal language is introduced:

Tt : t is T.

Formation rules:

F2. If T is a natural kind term and t is a object term, then Tt is an atomic formula.

F3. If t_1 and t_2 are object terms, $t_1 = t_2$ is an atomic formula.

F4. If Φ is a formula, $\neg \Phi$ is a formula.

F5. If Φ and Ψ are formulae, $(\Phi \rightarrow \Psi)$ is a formula.

F6. If x is an object variable and Φ is a formula, $\forall x \Phi$ is a formula.

F7. If T_1 and T_2 are natural kind terms, $T_1 = T_2$ is a formula.

F8. If K is a natural kind variable and Φ is a formula, $\forall K \Phi$ is a formula.

A sentence is a closed formula (which I do not define here).

(ii) Logic for (Externalism)

The logic is classical in that the logical constants and quantifiers are understood as in standard second-order logic 1: all object and natural kind terms take values in the domain of the quantifiers which is nonempty. The intended interpretation of the second-order quantifiers is as ranging over the kinds of the domain of objects. This interpretation is modified below. The holding of $(\mathrm{EN}_K) \forall K \exists x K(x)$ is a departure from classical logic. The resulting system is somewhat dissimilar from classical second-order logic in that it lacks the empty natural kind.

a) Model theory

It is straightforward to use a Tarskian model-theoretic account of logical consequence for the logics for (Externalism) and (Internalism). The idea of logical truth as truth at all metasemantic contexts connects with the model-theoretic account as follows: each such context generates an interpretation of natural kind nouns, and so in this respect, truth at all metasemantic contexts is like truth in all interpretations (of the nonlogical vocabulary). So truth at all metasemantic contexts connects with truth in all models via the notion of an interpretation.

Models are mathematical objects framed in set-theoretic terms. If we want to use model theory, we should not make assignments to kinds of the domain, but to sets—or to 'kind-sets', sets of objects that go proxy for kinds. There is a price to pay for adopting the

https://doi.org/10.1017/S1755020309090091 Published online by Cambridge University Press

standard set-theoretic understanding of a model: it prejudges the issues of what kinds are and that of their identity conditions. It is a fair price to pay for commonality and precision. Also, the model theories developed here validate the exact same principles as those given in section 4, and that is what really matters for our purposes.

I have introduced in section 3.3) the idea of a kind of the domain D, where that kind is determined by what is in D. Now quantification over kinds is replaced by quantification over kind-sets in such a way that the second-order domain is determined by the first-order one. To this effect, the existence of a function that maps the domain D to a kind-domain D^* is assumed:

$$f^* : D \rightarrow D^*.$$

f^* determines which subsets of D correspond to kinds. The objects in D^* are defined as follows: D^* is a set of ordered pairs, the first members of which are nonempty subsets of D and the second members of which are arbitrary objects. An arbitrary object is required to be the second member of these ordered-pairs if we are to have distinct but coextensive kinds. If the members of D^* were just nonempty subsets of D, as members of D^* all coextensive kinds would be identical.[56]

Thus a model for the language $L_{\text{Ex.Nk}}$ is a structure $M_{\text{Ex.Nk}} =< D, I >$, in which D is a nonempty set, and I is an interpretation function that assigns (a) objects in D to the object names of $L_{\text{Ex.Nk}}$, and (b) members of D^* to the natural kind nouns of $L_{\text{Ex.Nk}}$.

That is to say, the interpretation of the nonlogical vocabulary of $L_{\text{Ex.Nk}}$ is such that:

I1. If n is an object name, then $I(n) \in D$.

I2. If N_k is a natural kind noun, then $I(N_k) \in D^*$.

A variable assignment s for $L_{\text{Ex.Nk}}$ is a function that assigns a member $s(x)$ of D to each object variable x of the language and a member $s(K)$ of D^* to each natural kind variable K. Under I and s, object terms refer to objects in D and natural kind terms refer to members of D^*:

R1. Ref. of (x), $s = s(x)$

R2. If t is an object name, then ref. of (t), $s = I(t)$

R3. Ref. of (K), $s = s(K)$

R4. If T is a natural kind noun, then ref. of (T), $s = I(T)$

R5. Ext. of (K), $s =$ the first member of $s(K)$

R6. If T is a natural kind noun, then ext. of (T), $s =$ the first member of $I(T)$

The conditions under which a model $M_{\text{Ex.Nk}}$ and assignment s satisfy a formula Φ, that is the conditions under which $M_{\text{Ex.Nk}}, s \models \Phi$ are as follows:

S1. $M_{\text{Ex.Nk}}, s \models Tt$ iff ref. of (t), $s \in$ ext. of (T), s.

S2. $M_{\text{Ex.Nk}}, s \models t_1 = t_2$ iff ref. of (t_1), $s =$ ref. of (t_2), s.

S3. $M_{\text{Ex.Nk}}, s \models \neg\Phi$ iff it is not the case that $M_{\text{Ex.Nk}}, s \models \Phi$.

S4. $M_{\text{Ex.Nk}}, s \models \Phi \rightarrow \Psi$ iff it is not the case that ($M_{\text{Ex.Nk}}, s \models \Phi$ and not $M_{\text{Ex.Nk}}, s \models \Psi$).

[56] Another option would be to have two separate domains, D and D^* and not use 'f^*'. However we would not be able to capture the idea that D determines D^*.

https://doi.org/10.1017/S1755020309090091 Published online by Cambridge University Press

S5. $M_{\text{Ex.Nk}}, s \models \forall x \Phi$ iff $M_{\text{Ex.Nk}}, s[x/o] \models \Phi$ for every object o in D.[57]
S6. $M_{\text{Ex.Nk}}, s \models T_1 = T_2$ iff ref. of (T_1), s = ref. of (T_2), s.
S7. $M_{\text{Ex.Nk}}, s \models \forall \mathrm{K} \Phi$ iff $M_{\text{Ex.Nk}}, s[K/U] \models \Phi$ for every member U of D^*.

If $M, s \models \Phi$ for every assignment s, then we say that M *is a model of* Φ. It is easy to see that if s and s' are two assignments that agree on the free variables of Φ, then $M, s \models \Phi$ iff $M, s' \models \Phi$. A set of formulae Γ *is satisfiable* if there is a model M and an assignment s on M such that $M, s \models \Phi$ for every Φ in Γ. Φ is satisfiable, if $\{\Phi\}$ is satisfiable. A formula Φ *is a semantic consequence of* Γ ($\Gamma \models \Phi$), or $< \Gamma, \Phi >$ *is valid*, if the union of Γ with $\{\neg\Phi\}$ is not satisfiable. That is Φ is a semantic consequence of Γ if for every M and s on M, if $M, s \models \Psi$ for every Ψ in Γ, then $M, s \models \Phi$. And Φ is a *logical truth* if $< \emptyset, \Phi >$ is valid.

b) Deductive system

Let Γ be a set of formulae and Φ a single formula. A deduction of Φ from Γ is defined as a finite sequence $\Phi_1 \ldots \Phi_n$ such that Φ_n is Φ and, for each $i \leq n$, Φ_i is in Γ, Φ_i is an axiom, or Φ_i follows from previous formulae in the sequence by one of the rules.

Propositional fragment
Axiom schemas for the propositional fragment:

(Ax. 1) $\Phi \rightarrow (\Psi \rightarrow \Phi)$
(Ax. 2) $(\Phi \rightarrow (\Psi \rightarrow \Delta)) \rightarrow ((\Phi \rightarrow \Psi) \rightarrow (\Phi \rightarrow \Delta))$
(Ax. 3) $(\neg\Phi \rightarrow \neg\Psi) \rightarrow (\Psi \rightarrow \Phi)$

Rule for the propositional fragment:

(MP) From Φ and $\Phi \rightarrow \Psi$, infer Ψ

First-order fragment
Axiom schemas for the first-order-fragment:

(UI) $\forall x \Phi(x) \rightarrow \Phi(t)$, where t is free for Φ
(Id.) $\forall x\, x = x$
(LL) $t_1 = t_2 \rightarrow (\Phi(t_1) \rightarrow \Phi(t_2))$, where t_1 and t_2 are free for Φ

Rule for the first-order fragment:

(G) From $\Phi \rightarrow \Psi(x)$, infer $\Phi \rightarrow \forall x \Psi(x)$, provided that x does not occur free in Φ or in any member of Γ

Derived principles:

(EG) $\Phi(t) \rightarrow \exists x \Phi(x)$, where t is free for Φ
(D) $\exists x\, x = x$
(I) $\exists x\, x = t$
(QR) $\forall x \Phi(x) \rightarrow \exists x \Phi(x)$

[57] '$s[x/o]$' reads: 'the assignment that is just like s with the exception that x is sent to o'; '$s[K/U]$', below, reads accordingly.

https://doi.org/10.1017/S1755020309090091 Published online by Cambridge University Press

Second-order fragment
Axiom schemas for the second-order fragment:

$(\mathrm{UI_{Nk}})\forall K\,\Phi(K) \rightarrow \Phi(T)$, where T is free for Φ
$(\mathrm{E_{Nk}})\forall K \exists\, x\, K(x)$
$(\mathrm{Id._{Nk}})\forall K\, K = K$
$(\mathrm{LL_{Nk}})T_1 = T_2 \rightarrow (\Phi(T_1) \rightarrow \Phi(T_2))$, where T_1 and T_2 are free for Φ

Rule for the second-order fragment:

($\mathrm{G_{Nk}}$) From $\Phi \rightarrow \Psi(K)$, infer $\Phi \rightarrow \forall K\ \Psi(K)$, provided that K does not occur free in Φ or any member of Γ.

Derived principles:

$(\mathrm{CP_{Nk}})\exists K \forall\, x(K(x) \leftrightarrow T(x))$
$(\mathrm{EG_{Nk}})\Phi(T) \rightarrow \exists K\,\Phi(K)$, where T is free for Φ
$(\mathrm{D_{Nk}})\exists K \forall x(K(x) \leftrightarrow K(x))$
$(\mathrm{I_2})\exists K\, K = T$
$(\mathrm{QR_2})\forall K\,\Phi(K) \rightarrow \exists K\,\Phi(K)$

(ii) Logic for (Internalism)

The logical constants and quantifiers are understood as above. The logic is universally free in that object names and natural kind nouns are not assumed to take values in the domain of the quantifiers, which may be empty.

a) Model theory

As for the logic for (Externalism), the model theory of the logic for (Internalism) departs from its natural interpretation: the things in the second-order domain are not natural kinds but kind-sets. D^* is a kind-set, a set of ordered pairs of a subset of D and an arbitrary object. The difference however is that both D and D^* may be empty. To indicate this, the subscript 'F', for 'free' is used: thus we write D_{F} and D^*_{F}.

Thus, a model for a free second-order language $L_{\mathrm{In.Nk}}$ is a structure $M_{\mathrm{In.Nk}} =< D_{\mathrm{F}}, I >$, in which D_{F} is a set that may be empty, the domain of the model, and I is a partial interpretation function that assigns (a) objects in D to the defined object names of $L_{\mathrm{In.Nk}}$, and (b) members of D^*_{F} to the defined natural kind nouns of $L_{\mathrm{In.Nk}}$.

That is to say, the interpretation of the nonlogical vocabulary of $L_{\mathrm{In.Nk}}$ is such that:

$\mathrm{I1_F}$. If n is an object name and $I(n)$ is defined, then $I(n) \in D_{\mathrm{F}}$.
$\mathrm{I2_F}$. If N_k is a natural kind noun and $I(N_k)$ is defined, then $I(N_k) \in D^*_{\mathrm{F}}$.

A variable assignment s for $L_{\mathrm{In.Nk}}$ is a partial function that assigns members $s(x)$ of D_{F}, to object variables x of the language, and members $s(K)$ of D^*_{F} to natural kind variables K. Under I and s, object terms refer to objects in D_{F} and natural kind terms refer to members of D^*_{F}:

$\mathrm{R1_F}$. If $s(x)$ is defined, then ref. of (x), $s = s(x)$; otherwise x does not refer at s.
$\mathrm{R2_F}$. If t is an object name and $I(t)$ is defined, then ref. of (t), $s = I(t)$; otherwise t does not refer at s.

https://doi.org/10.1017/S1755020309090091 Published online by Cambridge University Press

$R3_F$. If $s(K)$ is defined, then ref. of (K), $s = s(K)$; otherwise K does not refer at s.

$R4_F$. If T is a natural kind noun and $I(T)$ is defined, then ref. of (T), $s = I(T)$; otherwise T does not refer at s.

$R5_F$. If $s(K)$ is defined, then ext. of (K), $s =$ the first member of $s(K)$; otherwise K has no extension at s.

$R6_F$. If T is a natural kind noun and $I(T)$ is defined, then ext. of (T), $s =$ the first member of $I(T)$; otherwise T has no extension at s.

The conditions under which a model $M_{\text{In.Nk}}$ and assignment s satisfy a formula Φ, that is the conditions under which $M_{\text{In.Nk}}, s|{=}\Phi$ are as follows:

$S1_F$. If t and T refer, then $M_{\text{In.Nk}}, s|{=}Tt$ iff ref. of (t), $s \in$ ext. of (T), s; otherwise it is not the case that $M_{\text{In.Nk}}, s|{=}Tt$.

$S2_F$. If t_1 and t_2 refer, then $M_{\text{In.Nk}}, s|{=}t_1 = t_2$ iff ref. of (t_1), s = ref. of (t_2), s; otherwise it is not the case that $M_{\text{In.Nk}}, s|{=}t_1 = t_2$.

S3. $M_{\text{In.Nk}}, s|{=}\neg\Phi$ iff it is not the case that $M_{\text{In.Nk}}, s|{=}\Phi$.

S4. $M_{\text{In.Nk}}, s|{=}\Phi \rightarrow \Psi$ iff it is not the case that ($M_{\text{In.Nk}}, s|{=}\Phi$ and not $M_{\text{In.Nk}}, s|{=}\Psi$).

S5. $M_{\text{In.Nk}}, s|{=}\forall x\Phi$ iff $M_{\text{In.Nk}}, s[x/o]|{=}\Phi$ for every object o in D_F.

$S6_F$. If T_1 and T_2 refer, then $M_{\text{In.Nk}}, s|{=} T_1 = T_2$ iff ref. of (T_1), s = ref. of (T_2), s; otherwise it is not the case that $M_{\text{In.Nk}}, s|{=}T_1 = T_2$.

S7. $M_{\text{In.Nk}}, s|{=}\forall K\Phi$ iff $M_{\text{In.Nk}}, s[K/U]|\ \Phi$ for every member of D^*_F.

The definitions of satisfaction, logical truth, and logical consequence are the same for the logic for (Externalism), so I do not repeat them.

b) Deductive system

The propositional fragment is the same as in the logic for (Externalism). The rules (MP), (G), and (G_2) are the same. What is different is the axiom schemas for the first- and second-order fragments:

Axiom schemas for the first-order-fragment:

(UI_F) $\forall x\Phi(x) \rightarrow (\exists x\, x = t \rightarrow \Phi(t))$, where t is free for Φ.

(Id.) $\forall x\, x = x$.

(LL) $t_1 = t_2 \rightarrow (\Phi(t_1) \rightarrow \Phi(t_2))$, where t_1 and t_2 are free for Φ.

Derived principle:

(EG_F) $\Phi(t) \rightarrow (\exists x\, x = t \rightarrow \exists x\Phi(x))$, where t is free for Φ.

Axiom schema for the second-order fragment:

(UI_{NkF}) $\forall K\Phi(K) \rightarrow (\exists K\, K = T \rightarrow \Phi(T))$, where T is free for Φ.

(Id.$_{Nk}$) $\forall K\, K = K$.

(LL_{Nk}) $T_1 = T_2 \rightarrow (\Phi(T_1) \rightarrow \Phi(T_2))$, where T_1 and T_2 are free for Φ.

Derived principles:

(E_{NkF}) $\Phi(T) \rightarrow (\exists K\, K = T \rightarrow \exists K\Phi\,(K))$, where T is free for Φ.

(CP_{NkF}) $\exists K\, K = T \rightarrow \exists K\forall x(K(x) \leftrightarrow T(x))$.

https://doi.org/10.1017/S1755020309090091 Published online by Cambridge University Press

BIBLIOGRAPHY

Besson, C. (Manuscript). Language and existence: On a new application of free logic. DPhil Thesis, University of Oxford.

Bilgrami, A. (1992). Can externalism be reconciled with self-knowledge? In Pessin, A., and Goldberg, S., editors. *The Twin Earth Chronicles: Twenty Years of Reflection on Hilary Putnam's 'The Meaning of 'Meaning'*. New York: M.E. Sharpe, pp. 362–393.

Block, N. (1986). Advertisements for a semantics for psychology. *Midwest Studies in Philosophy*, **10**, 615–678.

Boghossian, P. (1997). What the externalist can know *a priori*. In Smith, B., and McDonald, C., editors. *Knowing Our Own Minds: Essays on Self-Knowledge*. Oxford, UK: Oxford University Press, pp. 271–284.

Bostock, D. (1997). *Intermediate Logic*. Oxford, UK: Oxford University Press.

Brewer, B. (2000). Externalism and a priori knowledge. In Boghossian, P., and Peacocke, C., editors. *New Essays on the A Priori*. Oxford, UK: Oxford University Press, pp. 415–432.

Burge, T. (1974). Truth and singular terms. In Lamber, K., editor. *Philosophical Applications of Free Logic*. Oxford, UK: Oxford University Press, pp. 189–204.

Burge, T. (1979). Individualism and the mental. In French, P., *et al.*, editors. *Midwest Studies in Philosophy* 4: *Studies in Metaphysics*. Minneapolis, MN: University of Minnesota Press, pp. 72–121.

Chalmers, D. (1996). *The Conscious Mind: In Search of a Fundamental Theory*, Oxford, UK: Oxford University Press.

Chalmers, D. (2002). The components of content. In Chalmers, D., editor. *The Philosophy of Mind: Classical and Contemporary Readings*. Oxford, UK: Oxford University Press, pp. 608–633.

Cocchiarella, N. (1983). Philosophical perspectives on quantification in tense and modal logic. In Gabbay, D., and Guenthner, F., editors. *Handbook of Philosophical Logic*, Vol. IV. Dordrecht, The Netherlands: Kluwer Academic Publishers, pp. 309–353.

Cocchiarella, N. (1986). *Logical Investigations of Predication Theory and the Problem of Universals*. Napoli, Italy: Bibliopolis.

Cocchiarella, N. (1993). Book review: Stewart Shapiro, *Foundations Without Foundationalism: A Case for Second-Order Logic*. *Notre Dame Journal of Formal Logic*, **34**, 453–468.

Cocchiarella, N. (2001). A conceptualist interpretation of Lesniewski's ontology. *History and Philosophy of Logic*, **22**, 29–43.

Cocchiarella, N. (2007). *Formal Ontology and Conceptual Realism*. The Netherlands: Springer Synthèse Library. Dordrecht.

Davies, M. (2000). Externalism and armchair knowledge. In Boghossian, P., and Peacocke, C., editors. *New Essays on the A Priori*. Oxford, UK: Oxford University Press, pp. 384–414.

Devitt, M. (2005). Rigid application. *Philosophical Studies*, **125**, 139–165.

Dummett, M. (1975). The philosophical basis of intuitionistic logic. In Benacerraf, P., and Putnam, H., editors. *Philosophy of Mathematics: Selected Readings* (second edition). Cambridge, UK: Cambridge University Press, pp. 97–129.

Dummett, M. (1981). *Frege: Philosophy of Language* (second edition). London, UK: Duckworth.

Fodor, J. (1987). *Psychosemantics: The Problem of Meaning in the Philosophy of Mind*. Cambridge, MA: MIT Press.

https://doi.org/10.1017/S1755020309090091 Published online by Cambridge University Press

Goldberg, S. (2006). Anti-individualist semantics for natural kind terms. *Grazer Philosophische Studien*, **70**, 55–76.

Gupta, A. (1980). *The Logic of Common Nouns: An Investigation in Quantified Modal Logic*. New Haven, CT: Yale University Press.

Kaplan, D. (1989a). Demonstratives. In Almog, J., and Perry, J., editors. *Themes from Kaplan*. Oxford, UK: Oxford University Press, pp. 481–563.

Kaplan, D. (1989b). Afterthoughts. In Almog, J., and Perry, J., editors. *Themes from Kaplan*. Oxford, UK: Oxford University Press, pp. 565–614.

Koslicki, K. (1999). The semantics of mass predicates. *Noûs*, **33**, 46–91.

Kripke, S. (1972). *Naming and Necessity*. In Davidson, D., and Harman, G., editors. *Semantics for Natural Languages.* Dordrecht, Netherlands: Reidel. Reprinted in S. Kripke (1980). *Naming and Necessity.* Oxford, UK: Blackwell.

Larson, R., & Segal, G. (1995). *Knowledge of Meaning: An Introduction to Semantic Theory*. Cambridge, MA: MIT Press.

Lehmann, S. (2002). More free logic. In Gabbay, D., and Guenthner, F., editors. *Handbook of Philosophical Logic* (second edition), Vol. V. Dordrecht, The Netherlands: Kluwer Academic Publishers, pp. 197–259.

McKinsey, M. (1991). Anti-individualism and privileged access. *Analysis*, **51**, 9–16.

McLaughlin, B., & Tye, M. (1998). Externalism, Twin Earth, and self-knowledge. In Smith, B., and McDonald, C., editors. *Knowing Our Own Minds: Essays on Self-Knowledge*. Oxford: Oxford University Press, pp. 285–320.

Putnam, H. (1975). The meaning of 'meaning'. In *Mind, Language and Reality: Philosophical Papers*, Vol. 2. Cambridge, UK: Cambridge University Press, pp. 215–271.

Roeper, P. (2004). First- and second-order logic of mass terms. *The Journal of Philosophical Logic*, **33**, 261–297.

Sainsbury, M. (2001a). *Logical Forms: An Introduction to Philosophical Logic* (second edition). Oxford, UK: Blackwell.

Sainsbury, M. (2001b). Sense without reference. In *Departing From Frege: Essays in the Philosophy of Language*. London, UK: Routledge, pp. 205–223.

Salmon, N. (2005). *Reference and Essence*. Oxford, UK: Blackwell.

Sawyer, S. (2003). Sufficient absences. *Analysis*, **63**, 202–208.

Segal, G. (2000). *A Slim Book About Narrow Content*. Cambridge, MA: MIT Press.

Shapiro, S. (1991). *Foundations Without Foundationalism: A Case for Second-Order Logic*. Oxford, UK: Oxford University Press.

Stalnaker, R. (1989). On What's in the Head. In *Context and Content: Essays on Intentionality in Speech and Thought*. Oxford, UK: Oxford University Press, pp. 169–193.

Stalnaker, R. (1990). Narrow content. In *Context and Content: Essays on Intentionality in Speech and Thought*. Oxford, UK: Oxford University Press, pp. 194–209.

Stalnaker, R. (2001). On considering a possible world as actual. In *Ways a World Might Be. Metaphysical and Anti-Metaphysical Essays*. Oxford, UK: Oxford University Press, pp. 188–200.

Stalnaker, R. (2003). Conceptual truth and metaphysical necessity. In *Ways a World Might Be. Metaphysical and Anti-Metaphysical Essays*. Oxford, UK: Oxford University Press, pp. 201–215.

Sterelny, K. (1983). Natural kind terms. In Pessin, A., and Goldberg, S., editors. *The Twin Earth Chronicles: Twenty Years of Reflection on Hilary Putnam's 'The Meaning of 'Meaning"*. New York: M.E. Sharpe, pp. 98–114.

https://doi.org/10.1017/S1755020309090091 Published online by Cambridge University Press

Stoneham, T. (1999). Boghossian on empty natural kind concepts. *Proceedings of the Aristotelian Society*, **99**, 119–122.

Tarski, A. (1936). The concept of logical consequence. In Corcoran, J., editor. *Logic, Semantics and Metamathematics*. Indianapolis, IN: Hackett, pp. 409–420.

Wiggins, D. (1984). The sense and reference of predicates: A running repair to Frege's doctrine and a plea for the copula. In Wright, C., editor. *Frege: Tradition and Influence*. Oxford, UK: Blackwell, pp. 126–143.

Williamson, T. (1992). Vagueness and ignorance. In Keefe, R., and Smith, P., editors. *Vagueness: A Reader*. Cambridge, MA: The MIT Press, pp. 265–280.

ST HUGH'S COLLEGE, OXFORD
OX2 6LE, UK
E-mail: corine.besson@philosophy.ox.ac.uk

https://doi.org/10.1017/S1755020309090091 Published online by Cambridge University Press