

MODALITY, MOOD, AND DESCRIPTIONS

KAI FREDERICK WEHMEIER

§1. Introduction. By means of what semantic features is a proper name tied to its bearer? This is a puzzling question indeed: proper names — like “Aristotle” or “Paris” — are syntactically simple, and it therefore does not seem possible to reduce their meanings, by means of a principle of compositionality, to the meanings of more basic, and hence perhaps more tractable, linguistic elements.

It used to be widely held that the syntactical simplicity of proper names is misleading, that names are just abbreviations, in some sense or other, for certain other expressions, whose referential mechanism is clearer. The most prominent such theory, sometimes attributed to Frege and Russell, holds that the expressions for which proper names go proxy are definite descriptions, like “the man who taught Alexander the Great” or “the capital of France”. While not without problems of its own, this description theory of names has a great intuitive appeal, and yields a beautifully unified account of the phenomenon of singular reference.

Thirty years ago, however, an influential argument against the description theory was put forward by Saul Kripke. Appealing to the referential behaviour of names and descriptions, respectively, in contexts of metaphysical possibility and necessity, it has come to be known as the “modal argument”. According to Kripke, names are “rigid designators”, referring to the same individual no matter what actual or counterfactual situation we may be speaking about, whereas definite descriptions usually designate, if at all, non-rigidly, shifting their referents from possible world to possible world. The argument runs roughly as follows:

Consider the proper name “Aristotle” and some candidate for a description it might abbreviate — “the teacher of Alexander”, say. If the name abbreviates the description, the two expressions should be synonymous, and hence one should be substitutable for the other in all (but quotational) contexts *salva veritate*. Now

- (1) Under certain counterfactual circumstances, Aristotle would not have taught Alexander

is true: After all, Philip II of Macedonia might have decided to educate his son all by himself, in which case Philip, but not Aristotle, would have taught Alexander. However, the result of replacing “Aristotle” by “the teacher of Alexander” in (1),

- (2) Under certain counterfactual circumstances, the teacher of Alexander would not have taught Alexander,

is false¹. The reason for this divergence in truth value is, according to Kripke, that “Aristotle” rigidly designates Aristotle even in discourse about counterfactual situations in which he would not have done some of the things he did do, whereas “the teacher of Alexander” refers, with respect to any counterfactual situation, to the individual (if any) who would, in that situation, have taught Alexander. In any case, the assumption of synonymy is refuted by the difference in truth-value between (1) and (2). And the example can easily be generalised: most properties we might use to identify an individual descriptively are possessed by that individual only contingently, so that an argument analogous to the one given above can be constructed.

In this paper, I shall argue that the modal argument is fallacious. The observation on which my criticism is based is this: If we read “the teacher of Alexander” in (2) as “the man who would have taught Alexander”, the sentence is indeed false. But if we read it as “the man who taught Alexander” instead, it is true, so that the latter description appears to be substitutable for “Aristotle” *salva veritate*. The modal argument thus seems to depend on an ambiguity concerning verb mood.²

Concerning any such claims of ambiguity, Kripke (1980, pp. 59–60, footnote 22) has remarked:

Some philosophers have thought that descriptions, in English, are ambiguous, that sometimes they non-rigidly designate, in each world, the object (if any) satisfying the description, while sometimes they *rigidly* designate the object actually satisfying the description. (...) I find any such alleged ambiguities dubious. I know of no clear evidence for them which cannot be handled either by Russell’s notion of scope or by the considerations alluded to in footnote 3, p. 25.

¹Or rather, as Kripke points out, it has a reading which makes it false, namely the one according the description narrow scope with respect to the modal operator: when “ x teaches Alexander” is symbolised as “ Tx ” and (1) is rendered as “ $\diamond\neg Ta$ ”, the reading under which (2) is false is “ $\diamond[(\iota x)(Tx)]\neg T(\iota x)(Tx)$ ”, which, according to Russell’s theory of descriptions, expands to “ $\diamond(\exists!x Tx \wedge \exists x (Tx \wedge \neg Tx))$ ”.

²For a critical discussion of two other popular strategies against the modal argument, viz. the so-called “wide scope theory” and the “actualised descriptions” approach, see Soames (1998).

The main task in justifying my objection to the modal argument is, then, to show that the distinction of moods cannot be explained away by appeal to Russell's notion of scope for definite descriptions.³ To this end, it will be necessary to examine in some detail the main tool of logical analysis for ordinary modal discourse, viz., standard modal predicate logic. The results of this investigation seem to be relevant beyond the question of the modal argument's soundness, for they indicate that a modal logic adequately reflecting ordinary discourse should be extensional.

The paper is organised as follows: I explain my objection in more detail in §2. The (hypothetical) Kripkean reply, relying on Russell's notion of scope, is outlined in §3: according to it, verb mood simply indicates whether the respective occurrences of a predicate lie within or without the scope of a modal operator. This reply is found to be insufficient in §4, where it is shown that there are innocuous pieces of ordinary modal discourse whose formalisation in modal logic is precluded by such an explanation of mood. After discussing, in §6, another objection (outlined in §5) based on the consideration of a special kind of modal context, I develop, in §7, a formal system $S5^*$ of modal predicate logic that allows for the distinction of moods, and is thereby capable of formalising the aforementioned statements inexpressible in ordinary quantified $S5$. Within this framework, the informal objection to the modal argument presented in §2 is analysed formally in §8. Some remarkable properties of $S5^*$, like extensionality and absence of non-rigidity, are discussed in §9. §10 explores the relation of $S5^*$ and ordinary $S5$ (with and without actuality operator) to the two-sorted predicate calculus with explicit quantification over possible worlds. Finally, I discuss some consequences of my analysis for the philosophy of modality and the philosophy of language in §11.

§2. Descriptions in the mood. Let us try to repeat the argument of §1, this time by using “the man who taught Alexander” instead of the more elliptical phrase “the teacher of Alexander”. Sentence (2) now corresponds to

- (3) Under certain counterfactual circumstances, the man who would have taught Alexander would not have taught Alexander.

Clearly, (3), like (2), is false, and hence the non-synonymy of “Aristotle” and “the man who taught Alexander” follows, just as before.

There is a problem, however: The description “the man who taught Alexander” does not occur in (3). What we have in fact shown is that “the man who *would have taught* Alexander” is not synonymous with “Aristotle”. Let us try

³Kripke's footnote 3 refers to the distinction between speaker's reference and semantic reference, as used in his discussion of Donnellan-type ambiguities between referential and attributive uses of descriptions; it does not appear to be relevant in the present case.

again:

- (4) Under certain counterfactual circumstances, the man who taught Alexander would not have taught Alexander.

This time, we got the substitution right. But — sentence (4) is not false. Clearly, it might have been the case that the man who *taught* Alexander (viz., Aristotle) *would not have taught* him.

The situation seems to be this: If we substitute the expression whose non-synonymy with the name we are trying to establish — the indicative description “the man who taught Alexander” —, no change in truth-value is forthcoming. Such a truth-value change occurs only if we use the corresponding subjunctive⁴ expression “the man who would have taught Alexander” instead — but no-one would claim that “Aristotle” should be synonymous with it in the first place.⁵

Why did the original argument go through so smoothly? Only because “the teacher of Alexander” is homonymous between the indicative “the man who taught Alexander” — this *is* a designator and hence a plausible synonym

⁴Strictly speaking, “would have taught” is actually not a subjunctive, but a “conditional mood” form. A word of explanation for the apparently idiosyncratic use of the term “subjunctive” throughout this paper is thus in order. Before I justify the terminology with respect to the English language, let me mention that, unlike English and most modern Romance languages, both Latin and German employ the subjunctive (more precisely, the *coniunctivus irrealis* and the *Konjunktiv II*, respectively) in both the protasis and the apodosis of counterfactual conditionals (witness “*Si tacuisses, philosophus mansisses*” and “*Wenn du geschwiegen hättest, wärest du ein Philosoph geblieben*”, respectively), and hence also in modal constructions such as those used above. With regard to these languages, my terminology would therefore be perfectly adequate. As far as I am aware, English, French and the other Romance languages also used to work this way, and a separate “conditional mood” verb form only occurred in later stages of their development. Even in contemporary English, where the subjunctive proper seems to be almost extinct, remnants of the earlier practice can still be observed, as e.g. in the phrase “Were that true, there were no more to say” (Fowler 1983, 596; Fowler notes that the conditional form “would be” is the modern equivalent for the subjunctive “were” of the apodosis). From a systematic, semantic point of view it seems clear that the conditional and the subjunctive are to be grouped together, for they both refer not to how things stand in the real world, but to counterfactual, hypothetical situations. It does, therefore, not seem too much of an injustice to official grammatical classifications to apply the term “subjunctive” also to these conditional forms, especially in the light of Fowler’s (quoting the Oxford English dictionary) explication that *subjunctive* “is taken to mean the use of a verb-form different from that of the indicative mood in order to ‘denote an action or a state of affairs as conceived (and not as a fact), and [expressing] . . . a contingent, hypothetical, or prospective event’” (Fowler 1983, 595; the parentheses and square brackets are Fowler’s, the ellipsis is mine).

⁵The subjunctive phrase is, after all, not even a designator. For who is the man who would have taught Alexander? The question does not make sense, unless a counterfactual situation has been specified with respect to which the subjunctive is to be interpreted. That is, the subjunctive phrase is at best an incomplete singular term, much like “his father” when no value has been assigned to the variable “he”. Just as one would naturally ask, in reaction to an abrupt utterance of “His father is rich”: “Whose father?”, one would reply to “The man who would have taught Alexander was rich” by asking “The man who would have taught Alexander *if what?*”.

candidate for “Aristotle” — and the subjunctive “the man who would have taught Alexander”, as which it must be read in order to make the result of the substitution come out false. In other words: the modal argument as presented in §1 rests on an equivocation. Any defence of the modal argument will have to explain the indicative-subjunctive distinction away.

§3. Objection I: Verb mood and modal scope. Is the analysis of the preceding section not much too simplistic? Does it not rest on a quite elementary logical error? After all, the Kripkean can tell the following story in defence against the objection from grammatical mood:

“Consider modal predicate logic. Sentence (1), for example, can there be expressed as

$$(5) \quad \diamond \neg T a.$$

And the sentence

$$(6) \quad \text{Aristotle taught Alexander}$$

is formalised as

$$(7) \quad T a.$$

Note how both “would have taught Alexander” and “taught Alexander” are expressed by the same predicate symbol “T”. What distinguishes the occurrence of “T” in (5) from that in (7) is that the former, but not the latter, lies within the scope of a modal operator. In terms of Kripke’s semantics for modal logic, (5) holds true at the actual world if and only if there is a possible world v such that “ $T a$ ” is not true at v . This means that the occurrence of “T” in (5) refers not to how things are, but to how things might have been — which is the function of the subjunctive mood. On the other hand, the occurrence of “T” in (7) obviously refers to how things stand in the actual world, and that is what the indicative mood is for.

What these formal considerations show is that, for a predicate, being subjunctive simply means being within the scope of some modal operator, whereas being indicative means not lying within such a scope. There is thus no need to view indicative and subjunctive versions of a predicate as syntactically distinct — the only distinction needed is that between an occurrence of a predicate lying within or without the scope of a modal operator. The reason why natural language exhibits the redundancy of having, in addition to modal operators (with associated scopes), also syntactically distinct expressions for indicative and subjunctive predicates, is that it is not as precisely regimented as the formal idiom of modal logic: sometimes it would just be difficult to find out what the scope of a natural language modal operator is. Consider, for instance, the

sentence

- (8) Under certain counterfactual circumstances, some people who are unhappy would have been happy.

Were there no distinction in terms of mood between the predicates, we would have no choice but to parse the sentence according to its sequential structure, reading it as something like

$$(9) \quad \diamond \exists x (\neg Hx \wedge Hx),$$

which, being logically false, is certainly not what (8) expresses. But the indicative “is unhappy” in (8) makes it clear that this predicate occurrence is not within the scope of the modal operator. And the subjunctive mood of “would have been happy” conversely indicates that this predicate occurrence does lie within the scope of the possibility operator. Hence, we must render (8) as

$$(10) \quad \exists x (\neg Hx \wedge \diamond Hx).$$

The grammatical moods are thus nothing but parsing aids for the translation of ordinary modal discourse into the canonical notation of modal predicate logic.

Accordingly, the analysis of §2 misses the mark: By Russell’s theory of descriptions, sentence (4) is short for

- (11) Under certain counterfactual circumstances, precisely one man taught Alexander, and some man taught Alexander and would not have taught Alexander.

The grammatical moods of the predicates enable us to see that the true logical structure of (11) is

- (12) Precisely one man taught Alexander, and some man taught Alexander but would, under certain counterfactual circumstances, not have taught Alexander,

or, formally,

$$(13) \quad \exists!x Tx \wedge \exists x (Tx \wedge \diamond \neg Tx).$$

This is because we have two occurrences of the indicative “taught” in (11), which must therefore lie outside the scope of the modal operator, and one occurrence of the subjunctive “would have taught”, which accordingly must lie within the scope of the possibility operator.

Now (13) is just longhand for

$$(14) \quad [(\exists x)(Tx)] \diamond \neg T(\exists x)(Tx),$$

that is, for the result of substituting, in (5), the description “ $(\iota x)(Tx)$ ” with wide scope over the modal operator. But Kripke (1980, p. 13) acknowledges that his argument will not work if the description is given wide scope over the possibility operator, and points out that it is the reading with narrow scope which is relevant for his purposes. That is, Kripke is concerned with the sentence

$$(15) \quad \diamond [(\iota x)(Tx)] \neg T(\iota x)(Tx),$$

which expands to

$$(16) \quad \diamond (\exists!x Tx \wedge \exists x (Tx \wedge \neg Tx)).$$

The translation of this sentence into the vernacular, however, reads

- (17) Under certain counterfactual circumstances, precisely one man would have taught Alexander, and some man would have taught Alexander and would not have taught Alexander.

which, given Russell’s theory of descriptions, is just (3) (“under certain counterfactual circumstances, the man who would have taught Alexander would not have taught Alexander”). But, just as Kripke claims, (17), and hence (3), is indeed false. The logical analysis given here shows that no equivocation is involved at all: (15) arises from (5) through the substitution of the description “ $(\iota x)(Tx)$ ” for “ a ” with narrow scope, just as (14) is the result of substituting *the very same description* for “ a ” with wide scope. “The man who taught Alexander” and “the man who would have taught Alexander” are of identical logical form — the second is simply the form which the first assumes, in surface grammar, when inserted into the scope of a modal operator. The objection put forward in §2 is therefore without any force.”

This appears to be a powerful argument against the objection from grammatical mood. Nevertheless, it is itself fallacious, as I shall now attempt to show.

§4. Objection I rebutted. It is a tacit presupposition of objection I that standard modal predicate logic is an adequate tool for the logical analysis of the kind of modal discourse that we are concerned with. This presupposition is, however, not fulfilled.

Take the following piece of modal talk from Kripke’s *Naming and Necessity*:

Consider a counterfactual situation in which, let us say, fool’s gold or iron pyrites was actually found in various mountains in the United States, or in areas of South Africa and the Soviet Union.

Suppose that all the areas which actually contain gold now, contained pyrites instead (. . .) (Kripke 1980, p. 124).⁶

In short, what Kripke is saying here might be expressed thus:

- (18) It might have been the case that every area which contains gold would have contained iron pyrites.

To take another example:

- (19) Under certain counterfactual circumstances, everyone who has flown to the moon would not have flown to the moon.

These are perfectly innocent, intelligible, and arguably true statements. The careful distinctions of mood do not arise from mere pedantry: While (19) is probably true (imagine a course of history under which everyone who in fact became an astronaut would have become a race driver instead), its relative

- (20) Under certain counterfactual circumstances, everyone who has flown to the moon would have been female

is likely to be false, since there do not seem to be possible circumstances under which, say, Neil Armstrong would have been a woman. Nevertheless, the sentence

- (21) Under certain counterfactual circumstances, everyone who would have flown to the moon would have been female

is probably true, and the difference in truth value between (20) and (21) can only be due to the differences in mood. Granted, we often use additional devices to indicate whether a predicate is used *realiter*, that is, with reference to the actual world, or *irrealiter* (with respect to some counterfactual situation). For instance, (19) might be expressed by the more verbose

- (22) Under certain counterfactual circumstances, everyone who has in fact flown to the moon would not have flown to the moon.

But the crucial semantic feature is still the distinction of moods: For in

- (23) Under certain counterfactual circumstances, everyone who would in fact have flown to the moon would not have flown to the moon,

⁶Interestingly, the adverb “actually”, on its first occurrence in the quotation, does not force evaluation of the relevant subsentence with respect to the actual world. The same is true for: “Under certain circumstances, no-one would believe in aliens, although there would actually be aliens.” David Lewis has spotted an ambiguity of “actually” here (1970, section IX, 1983, section B), but, as I argue below, there is a contextual feature determining which world will be invoked by this adverb: viz., the mood of the predicate to which it is attached.

the expression “in fact” does not have the power to make the subjunctive predicate “would have flown to the moon” refer to the actual world. The addition of such modifiers may thus sometimes *facilitate* the proper interpretation of modal statements, but the orientation towards either actual or counterfactual circumstances is effected, first and foremost, by the indicative and subjunctive moods.

So let us try to formalise (19) within modal logic. The predicate “has flown to the moon” being indicative, it should, according to the analysis suggested in the preceding section, lie outside the scope of the possibility operator. If that is so, then we have no choice but to place the universal quantifier “everyone” outside the modal scope, too.⁷ This leads to the following tentative paraphrase of (19):

$$(24) \quad \forall x (Fx \rightarrow \Diamond \neg Fx).$$

But this is not what (19) says. In terms of possible worlds, (19) requires there to be one world in which everyone who has in fact flown to the moon would not have done so — an $\exists\forall$ -combination. (24), on the other hand, says that, concerning each individual who has flown to the moon separately, there is a world, possibly depending on the individual at issue, in which that individual would not have flown to the moon — a $\forall\exists$ -combination.⁸

The trouble with (19) is that the predicate “has flown to the moon” is *both* indicative *and* within the scope of the modal operator. This shows immediately that something is amiss with the Kripkean defence: the subjunctive’s function is not *just* to indicate that the respective predicate lies within the scope of a modal operator, but rather that it is also “modally bound” to it, i.e., that it is semantically relativised to the possible worlds conjured up by the operator. Standard modal logic, however, automatically relativises *all* predicates in the scope of a modal operator to such possible worlds, due to the stipulation that, for a given world w and formula ϕ ,

$$w \models \Diamond\phi \text{ if and only if for some world } v, v \models \phi,$$

which is the reason for the inexpressibility of (19) in such a framework.

This problem is highly germane to our discussion, because it suggests that a formal system adequately capturing ordinary modal discourse would have to distinguish indicative and subjunctive predicates *typographically*, whereas standard modal logic only distinguishes indicative and subjunctive *occurrences* of a predicate symbol by means of the syntactical criterion of modal scope.

⁷In the framework of Hintikka’s IF-logic, there may be other choices. There, one could perhaps argue that “has flown to the moon” lies within the priority scope, but not within the binding scope, of the modal operator. See, for instance, Hintikka (1997).

⁸Indeed, no sentence of quantified S5 has the same truth conditions as (19). This was conjectured by Hazen (1976) and later proved by Hodes (1984c). For a simple direct proof, see Wehmeier (200x).

A suitably revised modal logic, such as the system S5* of Section 7, will therefore admit *indicative* descriptions having *narrow* scope with respect to modal operators. In any case, it seems clear that the Kripkean explanation of the indicative-subjunctive distinction is not successful as it stands.

§5. Objection II: Indicative predicates shifting extensions. “Granted,” the Kripkean might say, “in the formulation you have given it, the modal argument fails to deliver a contradiction. But we can easily change the formulation and produce a valid argument. Take

(25) Possibly, Aristotle did not teach Alexander,

or

(26) It is not necessary that Aristotle taught Alexander.

In contradistinction to what you have claimed about the roles of the grammatical moods, we here have indicative predicates shifting their extensions at other possible worlds. If we now perform the substitution you suggest, we obtain

(27) Possibly, the man who taught Alexander did not teach Alexander,

and

(28) It is not necessary that the man who taught Alexander taught Alexander.

If the modal operators “possibly” and “it is necessary that” make the indicative predicates shift their extensions in (25) and (26), then (27) and (28) are both false when the description is given narrow scope with respect to the modal operators. And there the argument is, again.”

§6. Objection II rebutted. Before discussing the objection itself, let me point out how far-reaching our observation concerning the logical significance of the indicative-subjunctive distinction is. It is, for example, highly germane to the analysis of counterfactual conditionals. Thus, when we say

(29) If everyone who is rich had been poor, then someone who is poor would have been rich,

we again cannot explain the semantics of the indicative predicates by claiming that they are not within the scope of the conditional. Perhaps it is even clearer in the case of counterfactuals than in the case of the modalities that it is the interplay between the conditional operator and the subjunctive that effects extension-shifting to other possible worlds.

Consider now the following counterfactual:

- (30) If not Oswald, but someone else had shot Kennedy, then Ruby would have killed the man who would have shot Kennedy, not the man who (in fact) shot Kennedy.

Since, as (29) shows, indicative predicates can occur within the scope of the conditional without shifting their extensions, we may give the indicative description narrow scope, without affecting the truth value of (30). It is then also clear that “the man who shot Kennedy” may be replaced *salva veritate* by the name “Oswald”.

Counterfactuals, however, seem to be the basis for our intuitions concerning metaphysical possibility and necessity. Thus, we infer

- (31) Under certain circumstances Nixon would have gotten Carswell through (Kripke 1971, p. 175)

from

- (32) If Nixon had bribed a certain Senator, he would have gotten Carswell through (Kripke 1971, p. 176).

Thus, we naturally arrive at a modal statement by abstracting from, or perhaps quantifying away, the protasis of a counterfactual conditional. This close connection between counterfactuals and modal statements is severed by formulations using “possibly”, “it is necessary that” and their cognates, with indicative predicates intended to shift extensions. Thus, even if objection II could be sustained, it would pertain to a quite singular phenomenon of limited evidential value.

But in fact, objection II can hardly be sustained. The main reason is that “possible” and “necessary”, in ordinary language (and even within philosophical jargon), have so many meanings that the semantics assumed by modal logicians for sentences like (25) and (26) must be regarded as stipulated (that is, stipulated to fit standard possible worlds semantics) rather than modelled upon natural language.

To elaborate a bit: Clearly, the most natural reading for (25) is

- (33) For all we know, it may well be that Aristotle did not teach Alexander,

taking the possibility to be epistemic in nature. In the case of necessity, the situation is even worse: In addition to Kripke’s list of philosophical meanings of “necessary” (“a priori”, “analytic”, “certain”), it can mean “obligatory” (“it is necessary that I pick up my daughter from school”), or “urgently desirable” (“it is necessary that you be here in time”).

Were the meanings of “possibly” and “necessarily” sufficiently clear, there would be no need to resort to subjunctive formulations in order to explain their meanings. But, for instance, Forbes (1985, p. 2), deems it necessary to explain what he means by “it is possible that *P*”:

As a rough elucidatory guide, “it is possible that *P*” in the broadly logical sense means that there are ways things might have gone, no matter how improbable they may be, as a result of which it would have come about that *P*.

Incidentally, I should like to point out that, in this natural language example, it is quite clearly inadmissible to substitute one and the same sentence for both occurrences of the predicate letter “*P*”: On its first occurrence, it must be replaced with an indicative clause, but on the second occurrence, with a subjunctive one. For clearly, the correct formulation is

(34) There are ways things might have gone, no matter how
 improbable they may be, as a result of which it would
 have come about that Jones *would have been* rich,

not

(35) There are ways things might have gone, no matter how
 improbable they may be, as a result of which it would
 have come about that Jones *is* rich.

The artificiality of formulations of type (25) and (26) emerges even more clearly when the expressiveness problem discussed in Section 4 is taken into account. For how should one express

(36) Under certain circumstances, everyone who has flown to the
 moon would not have flown to the moon,

when only “possibly”, “necessarily”, and indicative predicates are available? It has been suggested that the adverb “actually” could help; but how natural (and comprehensible) are the sentences

(37) Possibly, everyone who has actually flown to the moon has not
 flown to the moon

and

(38) Necessarily, some of the things that are actually red are shiny?

It seems to me that we do not use such sentences. Instead, we use phrases like (36) or

(39) No matter how things might have gone, some of the things that
 are red would have been shiny.

I maintain, therefore, that objection II is based on contrived philosophers' jargon whose semantics has been stipulated to fit standard possible worlds semantics, whereas our initial formulation is well attested in ordinary language.

§7. A modal logic of indicativity and subjunctivity. The considerations of Section 4 suggest that an adequate logical analysis of the modal argument requires a revised system of modal predicate logic. In particular, it must be possible to express (19) in such a system. Given that this sentence is of $\exists\forall$ -form, a translation into logical notation should start with " $\diamond\forall x$ ". What comes next? Something similar to " $Fx \rightarrow \neg Fx$ ", just that, as was argued in Section 4, the two occurrences of the predicate "F" must be typographically distinguished so as to mark one as indicative, and the other as subjunctive.

An obvious constraint on any language of modal predicate logic is that its non-modal part should simply be the language of ordinary predicate logic. Therefore, indicative predicates are to be expressed by the ordinary predicate symbols of non-modal predicate logic (which, after all, formalises ordinary, indicative discourse), and it is the subjunctive for which we need to introduce a new notation. This constraint rules out the standard solution to the expressiveness problem, viz. the introduction of an "actually" operator " \mathcal{A} ".⁹ For consider the following example:

- (40) Someone has flown to the moon, but under certain counterfactual circumstances, everyone who has flown to the moon would not have flown to the moon.

With an actuality operator, it would have to be formalised as¹⁰

$$(41) \quad \exists x Fx \wedge \diamond\forall x (\mathcal{A}(Fx) \rightarrow \neg Fx),$$

but this is clearly not a faithful representation of the ordinary language sentence (40): In (40), there are two occurrences of the indicative predicate "has flown to the moon", and on both of these occurrences, the predicate has exactly the same semantic function, viz., to refer to how things stand in the real world. There is one occurrence of "would have flown to the moon", which is syntactically distinguished by being subjunctive, and semantically distinguished by referring not to the real, but to some counterfactual situation. In (41), however, the first occurrence of "has flown to the moon" in (40) is modelled by " Fx ", and so is the occurrence of "would have flown to the moon", whereas

⁹Logics with actuality operators have been studied rather extensively in the literature. See, for instance, Segerberg (1973), Hazen (1976), Crossley and Humberstone (1977), Peacocke (1978), Davies and Humberstone (1980), Hodes (1984c, 1984a, 1984b), and Forbes (1985, 1989).

¹⁰The sentence $\exists x \mathcal{A}(Fx) \wedge \diamond\forall x (\mathcal{A}(Fx) \rightarrow \neg Fx)$, while having the same truth conditions as (41), is ruled out by the constraint that the non-modal first conjunct of (40) should be expressed in modal logic by its formalisation in ordinary predicate logic. Thanks to Tom van der Beek for indicating the need for clarification here.

the second occurrence of “has flown to the moon” corresponds to “ $A(Fx)$ ”. This does not appear to be a transparent logical analysis — why should the two occurrences of the predicate “has flown to the moon”, which function in precisely the same way semantically, be modelled in *two* typographically distinct ways, as “ Fx ” and “ $A(Fx)$ ”? And why is the subjunctive predicate “would have flown to the moon”, referring to some counterfactual world, represented in exactly the same way as the indicative predicate “has flown to the moon” (as it occurs first in (40)), which refers to the actual world?¹¹

The idea suggests itself to introduce a typographical marker for the subjunctive. Let us symbolise the subjunctive version of the predicate “ F ” by “ F^* ”. We may then formalise (40) as

$$(42) \quad \exists x Fx \wedge \diamond \forall x (Fx \rightarrow \neg F^*x),$$

and this represents adequately the syntactic relations between the various occurrences of predicates in (40). Also, the intuition that the modal operators act on or bind those predicates in its scope which are subjunctive, but leave the others untouched¹², is brought out well by (42).

A system of this kind has in fact been introduced by Humberstone (1982)¹³. He employs a sentential subjunctivity operator “ S ”, writing “ $S(Fx)$ ” instead of our “ F^*x ”. I should like to propose a different solution here, essentially for the following reasons: Sentential operators can usually be iterated, but Humberstone’s semantics cannot accommodate iterations of “ S ” (and consequently, he does not count expressions with iterated subjunctivity operators as well-formed). From the point of view of natural language, this is of course no loss, since there is no such thing as iterated subjunctivity; still, one wonders whether, under such conditions, a sentential operator is the adequate formalisation. Similarly, his semantics cannot handle modal operators within the scope of “ S ”. Again, the exclusion of such expressions from the class of formulae is not objectionable as such, but the necessity of this measure suggests that subjunctivity does not work like a sentential operator. I therefore

¹¹Of course, the Kripkean already faces trouble when the expressiveness problem is solved by adding an “actually” operator, as pointed out e.g. by Stanley (1997, section V). But for the descriptivist there remains the curious presence of two very similar designating expressions, “ $(Ix)(Tx)$ ” and “ $(Ix)(A(Tx))$ ”, both ordinarily referring to the man who taught Alexander, only one of which (and in fact the more complicated one) providing the name “Aristotle” with the right meaning (see also the discussion in Soames (1998)). Such a duplication of terms is avoided by the logic $S5^*$ introduced below.

¹²The alternative, having modal operators act on *all* predicates within their scope indiscriminately, and introducing an operator-inhibiting device such as an actuality operator to undo this default binding, seems to be less intuitive. See also §10 below.

¹³Humberstone’s insightful paper is unfortunately mutilated by some grave misprints. A corrected version can be found in chapter 1 of Humberstone (2000), together with a number of “updates and afterthoughts”. These should be consulted for more on the relation between Humberstone’s original system and our $S5^*$, in particular with regard to the importance of introducing a subjunctive quantifier.

propose to stick with the notation introduced above: distinguish subjunctive predicates¹⁴ from their indicative relatives by flagging them with an asterisk “*”. But now for the formal details.¹⁵

Fix, for the remainder of this paper, a set P of predicate symbols (each of fixed finite arity) and a set C of individual constants. The corresponding language $L(P, C)$, or L for short, of $S5^*$ has as primitive symbols, besides the elements of P and C , the equality symbol $=$, some complete set of propositional connectives (\neg and \wedge , say), infinitely many individual variables x_1, x_2, \dots , the existential quantifier \exists , the subjunctive existential quantifier \exists^* (read: “there would have been”), the possibility operator \diamond , and, for each element P of P , the subjunctive P^* of P .¹⁶ A term is an individual variable or an individual constant from C . The formulae of L are defined inductively by the following clauses:

- If P is an n -ary predicate symbol from P and t_1, \dots, t_n are terms, then $Pt_1 \dots t_n$ and $P^*t_1 \dots t_n$ are formulae.
- If s and t are terms, then $(s = t)$ is a formula.
- If F and G are formulae, then so are $\neg F$ and $(F \wedge G)$.
- If F is a formula and x is an individual variable, then $\exists xF$ and $\exists^* xF$ are formulae.
- If F is a formula, then so is $\diamond F$.

An occurrence of a subjunctive predicate P^* or a subjunctive quantifier \exists^* in some formula is said to be free if it is not within the scope of any diamond. An L -formula F is subjunctively closed (s-closed for short) if it contains no free occurrences of subjunctive predicates (including subjunctive quantifiers). F is closed with respect to individual variables (i-closed) if no individual variable occurs free in F (in the usual sense of freedom). An L -sentence is an L -formula which is both s-closed and i-closed.

A Kripke structure for $S5^*$ is just an ordinary $S5$ model¹⁷ (with possibly varying domains). More precisely, a Kripke structure K is a tuple

$$\left(K, @, (D_w)_{w \in K}, (P_w)_{w \in K}^{P \in P}, (c_K)_{c \in C} \right),$$

consisting of a set K of possible worlds, a distinguished element $@$ of K (called the “actual” world), a K -indexed family $(D_w)_{w \in K}$ of sets, at least one of which

¹⁴As will be noticed in the formal definition below, the quantifiers are here, following Frege, also regarded as (second-order) predicates.

¹⁵Logical symbols will often be used autonomously in the more technical sections of this paper, and concatenation is indicated by juxtaposition. No Quinean quasi-quotation is used. This should not cause any confusion.

¹⁶Since the equality symbol will always be interpreted by the true identity relation, there is no need to introduce the subjunctive equality predicate $=^*$.

¹⁷ $S5$ appears to be the logic of choice for the kind of modality Kripke investigates in his (1980). There is no reason why the present approach should not be extended to other systems of modal logic.

is non-empty (so D , the union of all D_w , will be non-empty), a doubly indexed family $(P_w)_{w \in K}^{P \in \mathcal{P}}$ of relations, where for $P \in \mathcal{P}$ n -ary and arbitrary $w \in K$, P_w is a subset of D^n , and a \mathcal{C} -indexed family $(c_K)_{c \in \mathcal{C}}$ of elements of D .¹⁸

A variable assignment for K is a mapping σ from the set of individual variables into D . Given an individual variable x , we also write $K(x, \sigma)$ for $\sigma(x)$. For individual constants c from \mathcal{C} , we also write $K(c, \sigma)$ for c_K . We now define, by recursion, a relation $K \models_w F[\sigma]$, holding between Kripke structures K , worlds w of K (this w shall be called the “subjunctive world”), variable assignments σ for K and L-formulae F , as follows:

- $K \models_w Pt_1 \dots t_n[\sigma]$ iff $(K(t_1, \sigma), \dots, K(t_n, \sigma)) \in P_w$
- $K \models_w P^*t_1 \dots t_n[\sigma]$ iff $(K(t_1, \sigma), \dots, K(t_n, \sigma)) \in P_w$
- $K \models_w (s = t)[\sigma]$ iff $K(s, \sigma) = K(t, \sigma)$
- $K \models_w \neg F[\sigma]$ iff not $K \models_w F[\sigma]$
- $K \models_w (F \wedge G)[\sigma]$ iff $K \models_w F[\sigma]$ and $K \models_w G[\sigma]$
- $K \models_w \exists x F[\sigma]$ iff for some $e \in D_w$, $K \models_w F[\sigma\{x := e\}]$ ¹⁹
- $K \models_w \exists^* x F[\sigma]$ iff for some $e \in D_w$, $K \models_w F[\sigma\{x := e\}]$
- $K \models_w \diamond F[\sigma]$ iff for some $u \in K$, $K \models_u F[\sigma]$

The evaluation of an i-closed formula F is obviously independent of the choice of a variable assignment; we therefore define $K \models_w F$, for i-closed F , to mean $K \models_w F[\sigma]$ for every assignment σ . Likewise, the evaluation of an s-closed formula F is independent of the choice of the subjunctive world, hence, for s-closed F , we let $K \models F[\sigma]$ mean: for all $w \in K$, $K \models_w F[\sigma]$. For sentences F , we define $K \models F$ to mean $K \models_w F[\sigma]$ for all worlds $w \in K$ and all assignments σ . As in ordinary predicate logic, the meaningful statements of $S5^*$ are given by the sentences — just like “is greater than 0”, or perhaps “it is greater than 0” is not meaningful in itself, neither is “Aristotle would have been fond of dogs” (in both cases, we need metalinguistic devices to bestow a truth value on the expressions: a variable assignment in the first, and a subjunctive world in the second case). And while it seems quite natural, in the case of an actuality operator, to assume the actual world as its default referent, no such intuition would support a stipulation according to which the actual world should serve, by default, as *subjunctive* world.

The way this semantics works shall be illustrated by formalising sentence (19). It can be expressed as

$$(43) \quad \diamond \forall x (Fx \rightarrow \neg F^*x).$$

Let us see what it means for (43) to be valid in a Kripke structure K . We have:

$$K \models \diamond \forall x (Fx \rightarrow \neg F^*x)$$

¹⁸It could be argued that P_w should always be a subset of $(D_w)^n$. I do not wish to take a stand on this issue here and simply follow what seems to be common practice.

¹⁹ $\sigma\{x := e\}$ is the function which is like σ except that it maps the variable x to e .

iff for all $w \in K$, all assignments σ ,

$$K \models_w \diamond \forall x (Fx \rightarrow \neg F^*x)[\sigma]$$

iff for all $w \in K$ and all σ there is some $u \in K$ such that

$$K \models_u \forall x (Fx \rightarrow \neg F^*x)[\sigma]$$

iff for all σ there is some u such that for all $e \in D_{@}$,

$$K \models_u (Fx \rightarrow \neg F^*x)[\sigma\{x := e\}]$$

iff for all σ there is some u such that for all $e \in D_{@}$,

$$\text{not } K \models_u Fx[\sigma\{x := e\}] \text{ or not } K \models_u F^*x[\sigma\{x := e\}]$$

iff for all σ there is some u such that for all $e \in D_{@}$,

$$e \notin F_{@} \text{ or } e \notin F_u$$

iff there is some $u \in K$ such that for every e in $D_{@} \cap F_{@}$, e is not in F_u .

This is exactly what is required of a formalisation of (19). Let us now examine what form the modal argument takes in the setting of $S5^*$.

§8. The modal argument in $S5^*$. In $S5^*$, sentence (1) (“under certain counterfactual circumstances, Aristotle would not have taught Alexander”) is formalised as

$$(44) \quad \diamond \neg T^*a.$$

Now we are to substitute the description “ $(Ix)(Tx)$ ” for the name “ a ”. Kripke insists on giving the description narrow Russellian scope with respect to the diamond:

$$(45) \quad \diamond [(Ix)(Tx)] \neg T^*(Ix)(Tx).$$

According to Russell’s theory of descriptions, this may be expanded to

$$(46) \quad \diamond (\exists!x Tx \wedge \exists x (Tx \wedge \neg T^*x)).$$

Let us see what it means for (46) to hold in a structure K :

$$K \models \diamond (\exists!x Tx \wedge \exists x (Tx \wedge \neg T^*x))$$

iff for all w and σ ,

$$K \models_w \diamond (\exists!x Tx \wedge \exists x (Tx \wedge \neg T^*x))[\sigma]$$

iff for all w and σ there is some u such that

$$K \models_u (\exists!x Tx \wedge \exists x (Tx \wedge \neg T^*x))[\sigma]$$

iff there is some u such that $\text{card}(T_{@} \cap D_{@}) = 1$ and for some $e \in D_{@}$, $e \in T_{@}$ and $e \notin T_u$ iff $\text{card}(T_{@} \cap D_{@}) = 1$ and for some $e \in D_{@}$: $e \in T_{@}$ and there is some u with $e \notin T_u$ iff for all w and σ ,

$$K \models_w \exists!x Tx[\sigma] \text{ and for some } e \in D_{@}, K \models_w Tx \wedge \diamond \neg T^*x[\sigma\{x := e\}]$$

iff for all w and σ ,

$$K \models_w \exists!x \top x \wedge \exists x (\top x \wedge \diamond \neg \top^* x)[\sigma]$$

iff

$$K \models \exists!x \top x \wedge \exists x (\top x \wedge \diamond \neg \top^* x).$$

This last sentence is nothing but the result of substituting “ $(\iota x)(\top x)$ ” for “ a ” in (44) *with wide scope* over the modal operator:

$$(47) \quad [(\iota x)(\top x)] \diamond \neg \top^*(\iota x)(\top x).$$

Now if, as is the case when $@$ is our world, “ $a = (\iota x)(\top x)$ ” and (44) hold, then (47), and hence (46) and (45), are true: the substitution leading from (44) to (45) is *salva veritate* (which is what was to be expected from the informal objection to the modal argument in §2), whether the description is given wide or narrow scope over the modal operator.

Sentence (45) corresponds to our natural language formulation (4) (“Under certain counterfactual circumstances, the man who *taught* Alexander would not have taught Alexander”) of §2. What about (3) (“Under certain counterfactual circumstances, the man who *would have taught* Alexander would not have taught Alexander”)? It can be formalised as

$$(48) \quad \diamond [(i^*x)(\top^*x)] \neg \top^*(i^*x)(\top^*x),$$

which I shall use as shorthand for

$$(49) \quad \diamond (\exists^*x \forall^*y (\top^*y \equiv x = y) \wedge \exists^*x (\top^*x \wedge \neg \top^*x)).$$

This, just like (3), is evidently false. Does this fact provide a way out for the Kripkean, in that it at least proves “ $(i^*x)(\top^*x)$ ” to be a non-rigid designator non-synonymous with “ a ”? No: “ $(i^*x)(\top^*x)$ ” is not a designator, since “ $a = (i^*x)(\top^*x)$ ”, not being s-closed, has no determinate truth conditions, unless a subjunctive world has been chosen in advance. Informally speaking, this manifests itself in the incoherence of the question: “Who is the man who would have taught Alexander?”, when no possible situation has been determined with respect to which the subjunctive could be interpreted. By way of analogy, consider the variable “ x ” and the formula “ $a = x$ ”. It, too, has no determinate truth conditions, unless a variable assignment has been specified in advance, and accordingly, we do not call the free variable (as such, i.e., in the absence of a variable assignment) a “designator”.

The analysis of the modal argument in terms of S5* therefore seems to vindicate the informal objection outlined in §2 — indicative descriptions are inter-substitutable with coreferential names.

§9. The disappearance of non-extensionality and non-rigidity. The irrelevance of Russellian scope with respect to the possibility operator that was observed in the preceding section is of course just an instance of a more general phenomenon:

LEMMA 1. *Let F be an s-closed $S5^*$ -formula, and A any $S5^*$ -formula. Then for any structure K we have:*

- (a) $K \models \Box[(\iota x)F]A_x[(\iota x)F] \equiv [(\iota x)F]\Box A_x[(\iota x)F]$ and
- (b) $K \models \Diamond[(\iota x)F]A_x[(\iota x)F] \equiv [(\iota x)F]\Diamond A_x[(\iota x)F]$.

PROOF. We consider only the case of the box. Let K be a structure, $w \in K$ and σ any assignment for K . Then all of the following are equivalent:

- $K \models_w \Box \exists x (\forall y (F_x[y] \equiv x = y) \wedge A)[\sigma]$
- $(\forall u \in K) (\exists e \in D_{@}) (K \models_u \forall y (F_x[y] \equiv x = y)[\sigma\{x := e\}]$ and $K \models_u A[\sigma\{x := e\}]$
- $(\forall u \in K) (\exists e \in D_{@}) (\{f \in D_{@} : K \models_u F[\sigma\{x := f\}]\} = \{e\})$ and $K \models_u A[\sigma\{x := e\}]$
- $(\forall u \in K) (\exists e \in D_{@}) (\{f \in D_{@} : K \models_w F[\sigma\{x := f\}]\} = \{e\})$ and $K \models_u A[\sigma\{x := e\}]$ (because $K \models_u F[\sigma\{x := f\}]$, F being s-closed, does not depend on u)
- $(\exists e \in D_{@}) [(\{f \in D_{@} : K \models_w F[\sigma\{x := f\}]\} = \{e\})$ and $(\forall u \in K) K \models_u A[\sigma\{x := e\}]$
- $K \models_w \exists x (\forall y (F_x[y] \equiv x = y) \wedge \Box A)[\sigma]$

□

Three closely related corollaries may be drawn from the lemma.

First, it is well known for non-modal predicate logic that, under the hypothesis that $\exists!x F$, the scope of the description $(\iota x)F$ in any given formula A is semantically irrelevant. By the lemma, it follows that, given $\exists!x F$ for s-closed F , it does not matter how the description $(\iota x)F$ is scoped within $S5^*$ formulae $A_x[(\iota x)F]$. The second corollary is that $S5^*$ is extensional, as far as s-closed terms and formulae are concerned. We have, for example, that the schema

$$[(\iota x)F](y = (\iota x)F) \longrightarrow G_x[y] \equiv G_x[(\iota x)F]$$

is valid in all $S5^*$ models, where F is s-closed and the scope for $(\iota x)F$ in $G_x[(\iota x)F]$ is arbitrary. This follows by giving the description $(\iota x)F$ in $G_x[(\iota x)F]$ maximal scope, which, by the first corollary, does not affect the truth value of $G_x[(\iota x)F]$.

The third corollary, finally, is that there is no non-rigid designation in $S5^*$. Descriptions that are not s-closed do not qualify as *designators* at all (since their semantic values depend on the subjunctive world chosen), and s-closed descriptions designate *rigidly*: It is clearly $S5^*$ -valid, for s-closed F , that

$$[(\iota x)F](y = (\iota x)F) \longrightarrow [(\iota x)F]\Box y = (\iota x)F;$$

by the first corollary we have that

$$[(\iota x)F] \Box y = (\iota x)F$$

is equivalent to

$$\Box [(\iota x)F] y = (\iota x)F,$$

and hence

$$[(\iota x)F](y = (\iota x)F) \longrightarrow \Box [(\iota x)F] y = (\iota x)F.$$

§10. S5* and ordinary predicate logic. There is an ongoing debate concerning the relation of the modal idiom to languages with explicit quantification over possible worlds — see e.g. Lewis (1968), Hazen (1976), Forbes (1985, 1989), Cresswell (1990), Melia (1992), Forbes (1992), Chihara (1998) —, so it may be instructive to see how S5* relates to such a language. Let me first introduce some notation.

$L_2(P, C)$, or L_2 for short, is the two-sorted language of ordinary first-order logic given by

- two sorts of first-order variables: x, y, z , possibly indexed, of individual type, and $\alpha_0, \alpha_1, \alpha_2, \dots$, of world type;
- corresponding to each n -ary predicate symbol $P \in P$, an $n + 1$ -ary predicate symbol P° , whose first argument place is of world type and whose remaining argument places are of individual type;
- the individual constants from C ;
- the binary equality predicate $=$, both of whose argument places are of individual type; and
- a fresh binary predicate symbol I , whose first argument place is of individual type and whose second place is of world type.

Now we may associate, with each Kripke structure K as defined in §7, a classical structure \mathfrak{M}_K for L_2 by letting the world variables range over the set K of worlds of K , the individual variables over the union D of the D_w of K , interpreting every $c \in C$ by the element c_K of D , stipulating that eIw hold iff e is an element of D_w , and letting P° hold of (w, e_1, \dots, e_n) iff P_w holds of (e_1, \dots, e_n) in K . \mathfrak{M}_K will then satisfy the sentence $(\forall x)(\exists \alpha_1)(xI\alpha_1)$; and in fact every classical structure for L_2 which validates this sentence can be viewed as a Kripke structure in the obvious way.

We can now translate the L-formulae F of S5* into L_2 -formulae $(F)_2$ as follows:

- $(Pt_1 \dots t_n)_2 := P^\circ \alpha_0 t_1 \dots t_n$;
- $(P^* t_1 \dots t_n)_2 := P^\circ \alpha_1 t_1 \dots t_n$;
- $(s = t)_2 := (s = t)$;
- $(\neg F)_2 := \neg (F)_2$;
- $(F \wedge G)_2 := ((F)_2 \wedge (G)_2)$;
- $(\exists x F)_2 := \exists x (xI\alpha_0 \wedge (F)_2)$;

- $(\exists^* xF)_2 := \exists x (xI\alpha_1 \wedge (F)_2)$;
- $(\diamond F)_2 := \exists \alpha_1 (F)_2$.

The image of the translation $()_2$ is the fragment of L_2 consisting of the formulae which contain as world variables at most α_0 and α_1 (the former occurring free only) and within which every individual quantifier $\exists x$ occurs only bounded, that is, in the form $\exists x (xI\alpha_0 \wedge \dots)$ or $\exists x (xI\alpha_1 \wedge \dots)$. The translation of an s-closed formula of $S5^*$ is an L_2 -formula containing no world variable other than α_0 free.

Calling $()_2$ a translation is justified by the following fact:

Given a Kripke structure K , a world w of K , a variable assignment σ for K and a formula F of $S5^*$, $K \models_w F[\sigma]$ holds if and only if $\mathfrak{M}_K \models (F)_2[\sigma \cup \tau]$, where τ is any assignment of worlds to world variables mapping α_0 to $@$ and α_1 to w .

$S5^*$ is thus a mere notational variant of a certain fragment of L_2 .²⁰ The semantics are isomorphic: the subjunctive world can be seen as the variable assignment for α_1 , which is all we need to evaluate formulae in the image of $()_2$ (α_0 , functioning rather like a world constant, always being assigned $@$ as its value).

The crucial difference between ordinary $S5$ and $S5^*$ is that, technically speaking, the modal operators of $S5$, when interpreted as world quantifiers, cannot *selectively* quantify into just *some* of the argument positions, since these are all marked by the same world variable. This is somewhat disguised by the usual, roundabout way of translating $S5$ into L_2 (of first defining the relativisation of an $S5$ formula to an arbitrary world variable, and then defining the translation of a formula as its relativisation to the actual world, as e.g. in Forbes 1989, p. 10); but one can give a completely straightforward translation $()^\circ$ of standard $S5$ (in the language given by P and C as above) into L_2 as follows:

- $(Pt_1, \dots, t_n)^\circ := P^\circ \alpha_0 t_1, \dots, t_n$
- $(s = t)^\circ := (s = t)$
- $(\neg F)^\circ := \neg F^\circ$
- $(F \wedge G)^\circ := (F^\circ \wedge G^\circ)$
- $(\exists x F)^\circ := \exists x (xI\alpha_0 \wedge F^\circ)$
- $(\diamond F)^\circ := \exists \alpha_0 F^\circ$.

We clearly have that a Kripke structure K (defined as before) models (in the sense of $S5$) an $S5$ -formula F under the assignment σ if and only if the associated classical structure \mathfrak{M}_K satisfies F° under the assignment $\sigma \cup \tau$, where τ is any assignment of worlds to world variables mapping α_0 to $@$.

²⁰In fact, $S5^*$ can be extended to a logic $S5^{oo}$ intertranslatable with the *full* language L_2 , in a way somewhat analogous to the extension of “actually” logics by numerically indexed actuality and modal operators (see Peacocke (1978) and Forbes (1989)).

Given this translation, it is easy to see that the equivocation involved in the modal argument corresponds to an overloading of the world variable α_0 . First of all, α_0 plays the role of a name for the actual world (because it is standardly assigned @ as its value). But second, it is also used as a bound variable over possible worlds. This is bound to lead to fallacy; let us examine the modal argument, as it presents itself in S5 and L_2 , respectively:

The substitution, in the context of S5, of “ $(\iota x)(\top x)$ ” for “ a ” (narrowly scoped over the modal operator) in

$$(50) \quad \diamond \neg \top a$$

translates into L_2 as the substitution of

$$(51) \quad (\iota x)(xI\alpha_0 \wedge \top^\circ \alpha_0 x),$$

again with narrow scope over the existential quantifier, for “ a ” in

$$(52) \quad \exists \alpha_0 \neg \top^\circ \alpha_0 a.$$

In (51), α_0 occurs free only, that is, exclusively as a name of the actual world. In (52), however, the same variable becomes bound, thereby losing the individuality it enjoyed in (51). And it is entirely obvious that the substitution of “ $(\iota x)(xI\alpha_0 \wedge \top^\circ \alpha_0 x)$ ” for “ a ” in (52) (with narrow scope) is not legitimate, for elementary reasons — the variable α_0 , free in “ $(\iota x)(xI\alpha_0 \wedge \top^\circ \alpha_0 x)$ ”, having to become bound in the result of such a substitution.

Non-rigid designation, then, arises from considering the variable α_0 in the description $(\iota x)(xI\alpha_0 \wedge \top^\circ \alpha_0 x)$ once as being assigned the standard value @, and once as being a bound variable. It exploits the apparent connection between α_0 as a free and as a bound variable; but the connection really arises only through the bad strategy of using *one* variable to serve all purposes. There simply is no semantic link between α_0 as free and as bound variable; bound variables are, as it were, devoid of all individuality.

It may be instructive to revisit the “actually” operator \mathcal{A} at this point. Let $S5\mathcal{A}$ be S5 extended by the actuality operator \mathcal{A} and the actuality quantifier \exists^a , as in Hazen (1990). Then $S5\mathcal{A}$ can also be translated into L_2 ; but to bring out the peculiarity of actuality, I should like to present a translation into a slightly different language first.

Let L_3 be the language L_2 augmented with a new sentential operator sub . The language $L_{\mathcal{A}}$ is the fragment of L_3 inductively characterised by the clauses:

- If P is an n -ary relation symbol from \mathcal{P} and t_1, \dots, t_n are terms, then $P^\circ \alpha_0 t_1 \dots t_n$ is in $L_{\mathcal{A}}$.
- If s and t are terms (of individual type), then $(s = t)$ is in $L_{\mathcal{A}}$.
- If F and G are in $L_{\mathcal{A}}$, then so are $\neg F$ and $(F \wedge G)$.
- If F is in $L_{\mathcal{A}}$ and x is an individual variable, then $\exists x (xI\alpha_0 \wedge F)$ and $\exists x (\text{sub}(xI\alpha_0) \wedge F)$ are in $L_{\mathcal{A}}$.

- If F is in L_A , then so is $\exists\alpha_0 F$.
- If F is in L_A , then so is $\text{sub}(F)$.

The semantics of sub is given by the stipulation

$$\mathfrak{M} \models \text{sub}(F)[\sigma \cup \tau] \quad \text{iff} \quad \mathfrak{M} \models F[\sigma \cup \tau\{\alpha_0 := @\}],$$

where σ is an assignment for the individual variables and τ an assignment for the world variables.

Clearly, L_A is just a notational variant of the language of $S5A$. Given an $S5A$ -formula F , its translation $(F)_A$ into L_A is defined recursively by the clauses

- $(Pt_1 \dots t_n)_A := P^\circ \alpha_0 t_1 \dots t_n$
- $(s = t)_A := (s = t)$
- $(\neg F)_A := \neg(F)_A$ and $(F \wedge G)_A := ((F)_A \wedge (G)_A)$
- $(\exists x F)_A := \exists x (xI\alpha_0 \wedge (F)_A)$
- $(\exists^a x F)_A := \exists x (\text{sub}(xI\alpha_0) \wedge (F)_A)$
- $(\diamond F)_A := \exists\alpha_0 (F)_A$
- $(A(F))_A := \text{sub}(F)_A$.

Let us look at the translation of, say,

$$(53) \quad \diamond \forall^a x (A(Fx) \rightarrow \neg Fx)$$

into L_A , which is

$$(54) \quad \exists\alpha_0 \forall x (\text{sub}(xI\alpha_0) \rightarrow (\text{sub}(F^\circ \alpha_0 x) \rightarrow \neg F^\circ \alpha_0 x)).$$

As is obvious from the semantics of sub , this operator shields the variable α_0 off from quantification: Although, within the subformula occurrence of, e.g., $\text{sub}(F^\circ \alpha_0 x)$, the variable occurs within the scope of the existential quantifier $\exists\alpha_0$, it is not bound by it. Now in the canonical notation of L_2 , such a shielding operator is not available, and we must therefore take recourse to replacing the variable α_0 by a fresh one, α_1 , say, which is not bound by any quantifier; still, we must assign this fresh variable *the same value* as α_0 , since it is to simulate a *free* occurrence of α_0 . In other words: α_1 is nothing but the free variable α_0 disguised from the quantifiers.

More precisely, to extend the translation $()^\circ$ (of $S5$ into L_2) to $S5A$, add to the clauses for $S5$ given above the clauses

- $(A(F))^\circ := F^\circ[\alpha_0/\alpha_1]$,
- $(\exists^a x F)^\circ := \exists x (xI\alpha_1 \wedge (F)^\circ)$

where, for any L_2 -formula G , $G[\alpha_0/\alpha_1]$ is the result of replacing every free occurrence of α_0 in G by α_1 . Then the following holds:

A Kripke structure K models an $S5A$ -formula F under the assignment σ , if and only if the associated classical structure \mathfrak{M}_K satisfies $(F)^\circ$ under the assignment $\sigma \cup \tau$, where τ is any assignment of worlds to world variables mapping both α_0 and α_1 to $@$.

The introduction of a second world variable is clearly an improvement over the case of simple S5. But the roles of free and bound variables are still not kept strictly apart: On its *free* occurrences, the variable α_0 is identified with the invariably free variable α_1 (referring to the actual world) through the constraint on the variable assignment τ . And this must be so; for consider the S5A-formula

$$(55) \quad \diamond [\forall^a x (A(Fx) \rightarrow \neg Fx)] \wedge Fc$$

and its translation into L_2 ,

$$(56) \quad \exists \alpha_0 [\forall x (xI\alpha_1 \rightarrow (F^\circ \alpha_1 x \rightarrow \neg F^\circ \alpha_0 x))] \wedge F^\circ \alpha_0 c.$$

Since the first and third occurrences of “ F ” in (55) are to be evaluated with respect to the actual world, *both* α_0 and α_1 must be assigned @ to make (56) an adequate translation of (55). This strange handling of world variables reflects the misgivings expressed in §7 concerning the representation of the indicative predicate “has flown to the moon” by *two* constructions in (41): “ $A(Fx)$ ” when occurring within the scope of a modal operator, and “ Fx ” otherwise. To sum up, it seems preferable to use S5* instead of S5A not only from an informal, but also from a technical point of view.²¹

§11. Conclusion. What do these results mean for the philosophy of language? There seem to be at least three issues that deserve discussion. First, if the modal argument fails, what is the status of the description theory of proper names? Second, why is it so widely believed that metaphysically modal contexts should be non-extensional? And finally, are there further philosophical arguments that become dubious when due attention is paid to the indicative-subjunctive distinction?

The first point is rather simple. The conclusion to be drawn from the failure of the modal argument is just that the description theory is back in the race again. There are, to be sure, two other popular arguments against it — the semantic and epistemic arguments —, but the first does not seem to be very strong, and the second can easily be evaded by moving from the crude form of the description theory considered here (every name being backed up by one fixed definite description) to the so-called cluster theory, postulating that,

²¹It should be noted that S5A is interpretable in S5*. A given formula F of S5A is translated into S5* as follows: First delete all occurrences of “ A ” which are not within the scope of any modal operator. Say that a predicate letter or quantifier is directly within the scope of an “actually” operator if it is not separated from it by a modal operator. Attach an asterisk to every predicate letter and quantifier which lies in the scope of a modal operator, but not directly within the scope of any of the remaining “actually” operators. Finally, erase all remaining occurrences of “ A ”. This translation is clearly not one-one: Both “ Ta ” and “ $A(Ta)$ ” are translated into the single S5*-formula “ Ta ”.

associated with every name, there is some more or less vague *set* of descriptions giving the sense of the name.²²

So let me turn to the second point, concerning the non-intensionality of modal contexts. It might seem that our results flatly contradict the analytic tradition. Thus Quine, in whose writings the topic figures prominently, has argued as follows (see Quine 1953, p. 143): While, for instance, “9 is necessarily greater than 7” is true, replacement of “9” by the coreferential expression “the number of planets” yields the falsehood “the number of planets is necessarily greater than 7”.

Now, if it is metaphysical necessity that is at stake here, we can easily show that there is an equivocation between an indicative and a subjunctive description involved: From

(57) No matter how things might have gone, 9 would have been
greater than seven

and

(58) The x which numbers the planets is 9,

all that follows is the true assertion

(59) No matter how things might have gone, the x which numbers
the planets would have been greater than seven,

and not the false statement

(60) No matter how things might have gone, the x which would
have numbered the planets would have been greater than 7.

But, as a matter of fact, it is not metaphysical modality that Quine is aiming at, and he is very explicit about the kind of modality involved:

The general idea of strict modalities is based on the putative notion of analyticity as follows: a statement of the form “Necessarily . . .” is true if and only if the component statement which “necessarily” governs is analytic, and a statement of the form “Possibly . . .” is false if and only if the negation of the component statement which “possibly” governs is analytic. (Quine 1953, p. 143)

Quine is thus concerned with analyticity. But analyticity and metaphysical necessity are distinct notions: The former is truth in virtue of meaning, whereas the latter is concerned with ways the *world* might have been. Clearly, if a sentence is true solely in virtue of the meanings of its constituents, there is no guarantee that the replacement of some constituent by a coreferential one should preserve this property of analyticity. Metaphysical possibility

²²That this move suffices to evade the epistemic argument has been argued convincingly by Dummett (1973, pp. 135–137).

and necessity, however, are not in any obvious way concerned with meanings, and so there is little reason to believe that such contexts should be non-extensional. Quine's argument is, therefore, apparently correct, but does not reveal anything about contexts of metaphysical possibility and necessity. The widespread belief that the latter should not be extensional appears to be grounded in a confusion of analyticity (and perhaps other notions) and metaphysical necessity.

Concerning other arguments relying on modal claims, I should like to discuss two philosophical theses whose credibility appears to be weakened when due attention is paid to the indicative-subjunctive distinction. The first is Kaplan's "compulsive talker" argument against Reichenbach's token-reflexive analysis of the indexical "I". Kaplan writes:

But is it true, for example, that
 (A) "I" means the same as "the person who utters this token"?
 It is certainly true that
 I am the person who utters this token.
 But if (A) correctly asserted a synonymy, then it would be true that
 (B) If no one were to utter this token, I would not exist.
 Beliefs such as (B) could make one a compulsive talker. (Kaplan
 1989, pp. 519–520; I have altered Kaplan's number labels to letter
 labels).

What, exactly, is going on here? Presumably, the argument is meant roughly like this: I am the person who utters this token. Hence, if I exist, then the person who utters this token exists. By Russell's theory of descriptions, if I exist, then there is someone who utters this token. So, if no-one utters this token, then I do not exist. Now it is supposed to follow from the assumption of synonymy that if no-one were to utter this token, then I would not exist, i.e., (B). This consequence would clearly be unacceptable.

But is it really (A) that is responsible for this problem? Consider the following variant of Kaplan's argument: "The person who utters this token" means the same as "the person who utters this token". Hence, if the person who utters this token exists, then the person who utters this token exists. By Russell's theory of descriptions, if the person who utters this token exists, then there is someone who utters this token. So if no-one utters this token, then the person who utters this token does not exist. But now what? It simply does not seem to follow that, if no-one *were* to utter this token, then the person who *utters* this token would not exist. At best, it follows that, if no-one *were* to utter this token, then the person who *would utter* this token would not exist. But if "I" means "the person who *utters* this token", then it does *not* mean "the person who *would utter* this token", and so the last inference in Kaplan's original argument seems to be unwarranted.

The second philosophical thesis that, under the analysis proposed here, appears to be problematic, is the claim that there are necessities knowable only *a posteriori*. Take the example of Hesperus and Phosphorus. If Hesperus is Phosphorus, then they could not possibly have been distinct planets: One planet cannot be two. And the identity can certainly only be known *a posteriori* — without recourse to experience, it would have been impossible to find out that Hesperus is Phosphorus.

Let me digress for a second. “Hesperus is Phosphorus” is, by common consent, a singular statement. It is not a general truth, like “For every prime number, there is a greater prime number”. Yet, “Hesperus is Phosphorus” is trivially logically equivalent to the sentence

(61) “Everything is such that Hesperus is Phosphorus”,

where a redundant quantifier has been prefixed. This, of course, makes neither the identity statement itself nor the statement prefixed with the vacuous quantifier a general truth. Adding the quantifier is a cheap trick, because it does not bind any variable in its scope.

Now consider the statement

(62) No matter how things might have gone, Hesperus would always have been Phosphorus.

Here, the necessity operator is vacuous: there are no predicates which it could bind — except, perhaps, the equality predicate, but there seems to be a clear intuition that this is a rigid predicate, having true identity as its extension at every possible world (an intuition explicitly endorsed by Kripke). But apart from “would have been (identical to)”, there are no predicates in the embedded sentence, *a fortiori* no subjunctive predicates. Just as we do not call (61) a universal statement just because it has an initial (vacuous) universal quantifier, we should not call (62) a statement of necessity.

This vacuity of the necessity operator in (62) is somewhat disguised in S5, where the box is treated as a purely sentential operator. This makes genuine necessity statements take the same form as vacuous ones like (62). To take an example, consider

(63) No matter how things might have gone, Hesperus would have been the planet which would have appeared brightest in the evening sky.

This is a genuine, if false, necessity statement: The predicate “would have appeared brightest in the evening sky” is bound to the modal operator. In S5*, this is symbolised by

$$(64) \quad \Box h = (i^*x)(F^*x).^{23}$$

In ordinary S5, however, this is rendered simply as

$$(65) \quad \Box h = (ix)(Fx),$$

which assimilates it to the S5 sentence

$$(66) \quad \Box h = p,$$

only that, instead of the name p , (65) contains the description $(ix)(Fx)$. The fact that the box actually binds the predicate “ F ” in (65), but not the name “ p ” in (66), is concealed in this notation.²⁴

It is not exactly clear to me what we are to conclude from this observation concerning the (vacuous) necessity of identity. In any case, the argument for the existence of *a posteriori* necessities does seem to be in need of supplementation, if it can be sustained at all.

Acknowledgements. Concerning the intuitive objection to the modal argument set forth in §2, my main intellectual debt is to Ulrich Pardey. From his ordinary language investigations (1994, pp. 141–48, especially 142–45), I learned about the significance of the grammatical moods for logical analysis and about the systematic neglect of this distinction in mainstream analytic philosophy of language. It was he who encouraged me to take the intuitive objection seriously, and who helped me get clear about it in numerous conversations and discussions. Helge Rückert, Matthias Schirn, Göran Sundholm and Albert Visser read earlier versions of this paper and provided many helpful comments. Special thanks to Lloyd Humberstone and to Richard Holton for constructive criticism, discussion, and encouragement. Finally, I wish to thank the organizers of the Munich *Intensionality* conference, specifically Reinhard Kahle, for putting together such an inspiring event; thanks also to an anonymous referee for this paper who suggested valuable improvements here and there.

REFERENCES

- [1] CHARLES S. CHIHARA, *The worlds of possibility — modal realism and the semantics of modal logic*, Clarendon Press, Oxford, 1998.
- [2] MAX J. CRESSWELL, *Entities and indices*, Kluwer, Dordrecht, 1990.
- [3] JOHN N. CROSSLEY and LLOYD HUMBERSTONE, *The logic of “actually”*, *Reports on Mathematical Logic*, vol. 8 (1977), pp. 11–29.

²³Here and in what follows, I omit the descriptions’ scope indicator. The descriptions are always to be thought of as having narrow scope with respect to the modal operator.

²⁴Note that the sentence “ $\Box h = p$ ”, taken as an expression of S5 and of S5*, respectively, translates into the language L_2 as “ $\forall\alpha_0 h = p$ ” and “ $\forall\alpha_1 h = p$ ”, respectively, in each case revealing the presence of an empty quantifier.

- [4] MARTIN DAVIES and LLOYD HUMBERSTONE, *Two notions of necessity*, *Philosophical Studies*, vol. 38 (1980), pp. 1–30.
- [5] MICHAEL DUMMETT, *Frege — philosophy of language*, Duckworth, London, 1973.
- [6] GRAEME FORBES, *The metaphysics of modality*, Oxford University Press, Oxford, 1985.
- [7] ———, *Languages of possibility*, Basil Blackwell, Oxford, 1989.
- [8] ———, *Melia on modalism*, *Philosophical Studies*, vol. 68 (1992), pp. 57–63.
- [9] HENRY W. FOWLER, *A dictionary of modern English usage*, 2nd ed., Oxford University Press, Oxford, 1983, revised by Sir Ernest Gowers.
- [10] ALLEN HAZAN, *Actuality and quantification*, *Notre Dame Journal of Formal Logic*, vol. 31 (1990), pp. 498–508.
- [11] ALLEN HAZEN, *Expressive completeness in modal language*, *Journal of Philosophical Logic*, vol. 5 (1976), pp. 25–46.
- [12] JAAKKO HINTIKKA, *No scope for scope?*, *Linguistics and Philosophy*, vol. 20 (1997), pp. 515–44.
- [13] HAROLD T. HODES, *Axioms for actuality*, *Journal of Philosophical Logic*, vol. 13 (1984), pp. 27–34.
- [14] ———, *On modal logics which enrich first-order S5*, *Journal of Philosophical Logic*, vol. 13 (1984), pp. 423–54.
- [15] ———, *Some theorems on the expressive limitations of modal languages*, *Journal of Philosophical Logic*, vol. 13 (1984), pp. 13–26.
- [16] LLOYD HUMBERSTONE, *Scope and subjunctivity*, *Philosophia*, vol. 12 (1982), pp. 99–126.
- [17] ———, *Propositional attitudes: Some logical issues*, Typescript, 2000.
- [18] DAVID KAPLAN, *Demonstratives*, *Themes from Kaplan* (Joseph Almog, John Perry, and Howard Wettstein, editors), Oxford University Press, New York and Oxford, 1989, pp. 481–563.
- [19] SAUL A. KRIPKE, *Identity and necessity*, *Identity and individuation* (Milton K. Munitz, editor), New York University Press, New York, 1971, Reprinted in A. W. Moore (ed.), *Meaning and Reference*, Oxford University Press, New York and Oxford, 1993, pp. 162–91. Page references in the text are to the latter edition., pp. 135–64.
- [20] ———, *Naming and necessity*, Harvard University Press, Cambridge, Mass., 1980, Originally published (without the preface) in D. Davidson and G. Harman (eds.), *Semantics of Natural Language*, North Holland, Dordrecht, 1972, pp. 253–355.
- [21] DAVID LEWIS, *Counterpart theory and quantified modal logic*, *Journal of Philosophy*, vol. 65 (1968), pp. 113–26, Reprinted in his *Philosophical Papers*, Volume I, Oxford University Press, Oxford, 1983, pp. 26–39.
- [22] ———, *Anselm and actuality*, *Noûs*, vol. 4 (1970), pp. 175–88, Reprinted in his *Philosophical Papers*, Volume I, Oxford University Press, Oxford, 1983, pp. 10–20.
- [23] ———, *Postscripts to “Anselm and actuality”*, *Philosophical papers*, vol. I, Oxford University Press, Oxford, 1983, pp. 21–25.
- [24] JOSEPH MELIA, *Against modalism*, *Philosophical Studies*, vol. 68 (1992), pp. 35–56.
- [25] STEPHEN NEALE, *Descriptions*, The MIT Press, Cambridge, Mass., 1990.
- [26] ULRICH PARDEY, *Existenz und Reflexivität*, Beltz Athenäum, Weinheim, 1994.
- [27] CHRISTOPHER PEACOCKE, *Necessity and truth theories*, *Journal of Philosophical Logic*, vol. 7 (1978), pp. 473–500.
- [28] WILLARD V. QUINE, *Reference and modality*, *From a logical point of view*, Harvard University Press, Cambridge, Mass., 1953, pp. 139–59.
- [29] KRISTER SEGERBERG, *Two-dimensional modal logic*, *Journal of Philosophical Logic*, vol. 2 (1973), pp. 77–96.
- [30] ARTHUR F. SMULLYAN, *Modality and description*, *The Journal of Symbolic Logic*, vol. 13 (1948), pp. 31–37.
- [31] SCOTT SOAMES, *The modal argument: Wide scope and rigidified descriptions*, *Noûs*, vol. 32 (1998), pp. 1–22.

[32] JASON STANLEY, *Names and rigid designation, A companion to the philosophy of language* (Bob Hale and Crispin Wright, editors), Blackwell, Oxford, 1997, pp. 555–85.

[33] KAI F. WEHMEIER, *World travelling and mood swings, Foundations of the formal sciences II: Applications of mathematical logic in philosophy and linguistics* (Benedikt Löwe, Wolfgang Malzkorn, and Thoralf Räscher, editors), 200x, Forthcoming.

DEPARTMENT OF LOGIC AND PHILOSOPHY OF SCIENCE
SCHOOL OF SOCIAL SCIENCES
3151 SOCIAL SCIENCE PLAZA A
UNIVERSITY OF CALIFORNIA, IRVINE
IRVINE, CA 92697–5100, USA
E-mail: wehmeier@uci.edu