Notre Dame Journal of Formal Logic
Volume 29, Number 4, Fall 1988

On the Methodology of Possible Worlds Semantics, I: Correspondence Theory

DAVID PEARCE and HEINRICH WANSING*

1 Motivation Though possible worlds semantics has long been established as the dominant research tradition in philosophical logic and its applications, its various theories, background assumptions, and norms have seldom been systematically investigated from a methodological point of view. By way of illustration, consider the matter of semantic *adequacy*. When a new or revised logical system is proposed, the first and often the only significant 'test' to which it is subjected is that of "completeness": Can the logic be shown to be complete with respect to a suitable semantics? Providing the system has a minimum of intrinsic interest, an affirmative answer to this question is virtually a ticket to 'official' recognition, while even incomplete systems of no intrinsic interest whatsoever may acquire, in virtue of their incompleteness, a kind of rarity value in the catalogue of logics. In short, (in)completeness proofs are the mainstay of many a journal article and provide the meat of many logic textbooks.

If completeness is genuinely to represent a criterion of internal adequacy, and not merely a logical nicety, we must ask ourselves exactly what cash-value a complete semantics possesses. This question leads naturally to a further problem. Given a well-defined model theory and appropriate rules of interpretation, the matter of completeness is a factual (or better a logical) one, to be settled by formal analysis. But if we are considering logical semantics in general, or one tradition like that of possible worlds in particular, the question is no longer purely logical: it contains a methodological component and can be answered only on the basis of adopting certain *conventions*. To prove completeness we may need recourse to some nonstandard interpretation of the logical constants, or to some alternative specification of the intended 'models'. Consequently, issues of the following sort arise: Within what limits are we free to modify the stan-

*Parts of an earlier version of this paper were presented at the 5th Finnish-Soviet Logic Symposium held in Tampere, Finland, May 1987. For helpful comments on the earlier draft we are grateful to Johan van Benthem and Ingmar Pörn.

Received December 21, 1987

dard semantics? And to what extent do we require any such changes to be independently motivated, say in virtue of the special subject matter at hand? Such questions can be raised for logical semantics in general, as well as for the case of a specific research tradition within logical semantics.

A second set of issues is closely related to the first. Is it reasonable to construe possible worlds semantics (and logical semantics generally) as an empirical science which provides general laws and principles to explain linguistic, mathematical, or logical 'facts'? And, if so, in what sense might its postulates and theories be said to be falsifiable or confirmable? In this context we can ask, for example, whether possible worlds semantics is a sufficiently flexible framework for giving a descriptive and not merely an idealized account of the propositional attitudes, say the epistemic and doxastic attitudes of knowing, believing, and the like. Can one, in particular, overcome the various problems of logical omniscience that beset the usual possible worlds treatment of epistemic logic?

In investigating these kinds of questions it is essential to develop a sufficiently abstract and general perspective on possible worlds semantics. Although Kripke's semantics for (normal) modal logics has acquired the status of a core or 'paradigm' model, many significantly different variants of possible worlds semantics have been proposed for other types of intensional logics. It is important, therefore, to provide as far as possible a *uniform* treatment of these different approaches and thereby determine what features are *characteristic* of possible worlds semantics as well as what constraints govern successful semantic modeling within this research tradition.

In this respect, notice that the usual criterion of semantic adequacy—namely the presence of a suitable completeness theorem—would be a purely vacuous constraint were it to turn out that the degrees of freedom permitted in defining model-types and the satisfaction relation were so great as to allow almost any logic to be completely semanticized. This point has been emphasized by van Benthem [5], who seems to have been the first to study the problem in detail by offering both a general framework for analyzing different variants of possible worlds semantics as well as by proposing global constraints on semantic modeling. One of van Benthem's most interesting observations is that there are modal logics which are not only incomplete with respect to the standard (Kripke) semantics but which remain incomplete when the standard truth-conditions and model-conditions are (within appropriate but wide-ranging limits) tampered with. In other words, in a certain sense there are logics incomplete with respect to possible worlds semantics *in general.* He concludes from this that there is indeed 'empirical content' or 'Popperian vitality' in the possible worlds program.

van Benthem's analysis is carried out in the framework of *correspondence theory* which, in the typical case, describes the formal connections that hold between an intensional propositional logic L_0 and the classical predicate language L_1 for which the L_0 Kripke-models are ordinary relational structures. In particular, since the theory establishes correspondences between the L_0 intensional operators (and L_0 axioms) and (first- or higher-order) sentences of L_1, it offers a suitable means to compare and contrast different versions of possible worlds semantics: variations in L_0 truth-conditions (involving possible worlds) give rise to differing L_1-sentences, variations in the model-conditions for L_0 correspond to different choices of L_1 (predicates). Since van Benthem is chiefly concerned with classical predicate languages (L_1), we may refer to his framework as that of classical or *standard* correspondence theory.

Taking normal modal logics as a test-case, van Benthem employs standard correspondence theory to defend three basic constraints (denoted by (C1)–(C3) in Section 2 below) on the setting up of an adequate possible worlds semantics. These define the relevant (generalized) notions of semantic (in)completeness. In addition, he investigates the sense in which such a semantics might be read as providing an *explanation* of intensional inference patterns. In short, he endeavors to characterize part of the heuristics and the explanatory aims and resources of possible worlds semantics as a scientific research program or research tradition. Though this exercise is directed at one semantical tradition in particular, it may be useful to regard it as part of a more general objective: that of clarifying the methodological role and function of logical semantics within the philosophy of logic and language, as well as its relation to general linguistics. At the same time, van Benthem's approach reveals important connections between these issues and problems and methods of the philosophy of science at large: Is 'semantic' explanation a *bona fide* species of scientific explanation? What kinds of empirical claims, if any, might be associated with a semantical research tradition?

In light of van Benthem's proposed semantic constraints, various methodological questions for logical semantics suggest themselves. Among them:

- Are the van Benthem constraints fully justified?
- To what extent are they appropriate outside the restricted domain of normal modal logic?
- In which directions can one usefully generalize the perspective of standard correspondence theory?
- Can correspondence theory be applied to all types of possible worlds semantics?
- How does the logical structure of the possible worlds research tradition resemble that of other scientific research traditions? What techniques and methods developed within the philosophy of science might fruitfully be applied to it?

In this paper we undertake a preliminary examination of these questions, focusing chiefly on van Benthem's analysis and on the correspondence theory. To make the discussion as widely accessible as possible, we summarize in Section 2 the salient points of [5]. Our coverage of the main themes is by no means intended to be exhaustive, and several of the issues raised here will be taken up in more detail in future publications. In particular, a sequel to the present paper [24] will deal with the semantics of nonnormal possible worlds as presented in [26] and [27], the analysis of which takes us beyond the resources of the correspondence theory.

2 Standard correspondence theory Let us turn first to the general treatment of possible worlds semantics provided by van Benthem in [5]; relevant also are [4] and [6]–[8]. The setting is that of *correspondence theory*, which, as we remarked, deals with the relation between formulas of a modal propositional language, L_0, and the 'corresponding' formulas in the (first- or higher-order) language L_1 of the possible worlds models for L_0. Thus, in the standard case, the primitives of L_0 will consist of propositional letters, $p, q, \ldots$, the usual classical connectives, $\neg, \&, \ldots$, and a modal operator, $\Box$ (necessity). And a (Kripke) model for L_0 is a structure $M = \langle W, R, V \rangle$, where W is a nonempty set (of possible worlds), R is a binary (accessibility) relation on W, and V is a valuation assigning at each world truth-values to the propositional letters. By considering, for each propositional letter p, a corresponding one-place predicate P, where 'Px' denotes 'p holds at world x', one can then express in the ordinary (classical) language of $M^* = \langle W, R, P, \ldots \rangle$ (now viewed as a first-order relational structure) conditions corresponding to modal statements in L_0.[1] Thus, corresponding to the Kripke clause for necessity:

(1) $V(\Box p, w) = 1$ iff $V(p, w') = 1$ for all $w' \in W$ such that wRw',

we have the first-order condition

$$\forall y (Rxy \to Py).$$

Formally, this correspondence is captured by a translation τ sending modal formulas A into formulas $\tau(A)$ in the language L_1 of M^*. In the standard case, as above, τ is recursively defined by

$$\tau(p) = Px;\ \tau(\neg A) = \neg \tau(A);\ \tau(A \ \& \ B) = \tau(A) \ \& \ \tau(B);$$
$$\tau(\Box A) = \forall y (Rxy \to \tau Ay);$$

whence it is clear that for any L_0-formula A, and model M,

(2) $M \models A$ (in the Kripkean sense) iff $M^* \models \overline{\tau(A)}$ classically,

where $\overline{\tau(A)}$ denotes the universal closure of $\tau(A)$. Though, in general, this kind of correspondence implies a reduction of modal logic to second-order logic, as is well-known there are many cases where a modal logical axiom or system is characterized by a first-order condition on frames, hence by a sentence in the first-order language of M^*. Thus, the translation of the T-axiom $\Box p \to p$ is $\forall y(Rxy \to Py) \to Px$, and the validity of this statement ($\forall P$) amounts to the well-known (first-order) condition that R is reflexive, i.e. $\forall x Rxx$.

Correspondence theory thus offers a suitably general perspective from which to consider the aim and scope of possible worlds semantics. van Benthem considers three features in particular.

First, there is the question of semantic *explanation*. If one thinks of the modal language L_0 as expressing certain valid inference patterns, in the simplest case of the form

$$A \vdash^+ B,$$

and invalidities, of the form

$$C \vdash^- D,$$

then one may plausibly regard the semantic explanation of such patterns to consist in providing an L_1-theory, T_1 say, from which the corresponding (non)inferences are deducible, i.e.,

$$T_1 \ \& \ \tau(A) \vdash^+ \tau(B);\ T_1 \ \& \ \tau(C) \vdash^- \tau(D).$$

In the standard case above, T_1 will be a first- (or higher-) order theory containing one binary relation R as its only 'theoretical' term.

Secondly, the perspective of correspondence theory brings out the interplay between truth-conditions and model-conditions. Different ways of defining the translation τ correspond to different truth-conditions for the modal operators (or the propositional connectives). The above standard translation amounts to a classical interpretation of the connectives and the usual Kripke interpretation (1) of necessity. But alternatives to this scheme are possible. As van Benthem observes, for an underlying propositional *quantum* logic, negation would receive a nonclassical reading, given by

$$\tau(\neg p) = \forall y(Py \rightarrow Rxy),$$

for R a suitable 'orthogonality' relation; and in cases like that of KB (minimal normal modal logic plus the Brouwer axiom $p \rightarrow \Box\Diamond p$), different combinations of truth/model-conditions may characterize a logic equally well. Since KB is determined by all symmetric Kripke-frames, it is also characterized by *all* frames under the nonstandard truth-condition

$$\tau(\Box A) = \forall y(Rxy \vee Ryx) \rightarrow \tau Ay.$$

This raises the question of the proper roles to be played by truth-conditions and model-conditions respectively.

A third, and related, question concerns the extent to which one may depart from the ordinary Kripke-semantics involving a single accessibility relation R (fixing the language L_1) and still retain a semantically adequate possible worlds framework. As we know, ternary relations have been proposed both for conditional logics of the Stalnaker–Lewis kind as well as for the usual modalities. What kinds of general restrictions here on T_1 or its set of theoretical terms are appropriate for semantic explanation?

In outline, van Benthem's approach to these issues can be summarized as follows. First, the presence of a translation τ suggests the possibility of applying Frank Ramsey's view of empirical theories within which one distinguishes between a primary (observational) language L_0 and a secondary language L_1 including theoretical terms (see [25]). There is assumed to be a dictionary available for translating L_0-terms into L_1, and a suitable L_1-theory T_1 is intended to explain a given set of L_0 'facts', T_0 say. In general, then, one requires

(4) $T_1 \restriction L_0 = T_0$.

In view of the earlier desiderata for semantic explanation, viz.

(5) $\vdash^+ A$ iff $\models \tau(A)$,

if we consider the case where T_0 is a system of propositional modal logic, condition (4) then corresponds to the presence of a completeness theorem for T_0 guaranteeing that (5) holds. Thus we obtain a natural reading of semantic explanation in terms of completeness.

The Ramsey view can be extended by further differentiating the model-theoretic component and by considering the problem of so-called *Ramsey-eliminability* (of theoretical terms).[2] Rather than starting with the full Kripke semantics for L_0, one might take modal algebras as the appropriate structures for the primary language, introducing Kripke models $M = \langle W, R, V\rangle$ or frames $F = \langle W, R\rangle$ (or their expansions M^*) only in the theoretical context L_1. In place of (2), this would yield (by a standard argument)

(6) $f(F) \models A$ (in the sense of algebraic semantics) iff $F \models A$ iff for all models M based on F, $M^* \models \tau(A)$,

for f a suitable functor sending Kripke frames to modal algebras. van Benthem also investigates an 'intermediate' L_0-semantics in the form of structures

$$M = \langle W, \mu, V\rangle$$

in which W is a set of possible worlds, μ is a selection function on subsets of W (essentially μ maps $[\![A]\!]$ to $[\![\Box A]\!]$ for each proposition A), and V is a valuation (formally, M is equivalent to a neighborhood or 'Scott-Montague' model). Taking for the primary L_0-theory the minimal modal logic K, T_0 becomes axiomatized by simple conditions on μ. The appropriate L_1-theory T_1 for this logic is then simply the classical predicate calculus, and van Benthem is able to show that the accessibility relation R is not Ramsey-eliminable in this case; i.e. whilst (5) holds, the class $(\mathrm{Mod}T_1) \upharpoonright L_0$ is not first-order definable. (R is, however, $L_{\infty\omega}$-eliminable in the sense of [23].)

Turning to the broader questions raised above concerning the scope of possible worlds semantics, let us for the time being focus our attention on the framework of normal and quasi-normal modal logics, i.e. logics extending K. Within this context van Benthem proposes the following three general constraints and obtains two formal results relevant to them:

(C1) The translation τ should be first-order; in particular, $\tau(\Box A)$ should be representable in the form $\alpha(x, \tau(A), \Sigma)$, where Σ is a set of relevant relations on possible worlds, and α is a first-order condition.

(C2) The truth-condition τ should account for the valid inferences of the minimal logic K; any further modal principles are to be semanticized by means of appropriate conditions C restricting the class of T_1-models.

(C3) Such conditions C must refer only to the predicates Σ involved in explaining the intensional operators.

The main motive behind restrictions (C1) and (C2) seems to be that the truth-conditions should be kept as 'simple' as possible, so that additional semantic complexities are always to be captured, where feasible, by appropriately adjusting the L_1 language and the relevant constraints on its models. (C3), on the other hand, is required in order to avoid gratuitous or 'circular' semantic explanation, since in virtue of

$$A \vdash_K B \text{ iff } \overline{\tau(A)} \models \overline{\tau(B)},$$

where $\vdash_K$ denotes derivability in K, any L_0 inference pattern $A \vdash_K B$ could be otherwise matched in L_1 by including $\overline{\tau(A)}$ among the conditions C.

Retaining these three assumptions, van Benthem is able to show that not all intensional logics possess a complete possible worlds semantics, even allowing for admissible changes in the truth-conditions τ and the model-conditions C. In short, not all intensional inference patterns in L_0 can be semantically 'explained' by choosing a suitable theoretical language L_1, 'dictionary' τ, and L_1-theory C.

As he points out, the well-known incomplete logics constructed by Fine [13] and Thomason [31] are incomplete also with respect to the neighborhood semantics (as shown by Gerson [14]). A simple argument suffices to demonstrate that this incompleteness phenomenon carries over to van Benthem's semantics: if either one of these logics *were* complete given some choice of truth-condition α, it would also have to be complete on frames whose neighborhood relation N is defined by α.

van Benthem's second result expresses a limitation on admissible truth-conditions, thereby answering a question posed earlier, now sharpened into the following: What syntactic form may α take given that T_1 is to capture at least the valid inferences of K? The answer is provided by a new preservation theorem for first-order logic. Once again, there is also a connection with the neighborhood semantics. One notes first that a formula of the kind

$$\alpha(P) = \forall y(Rxy \rightarrow Py),$$

i.e. the usual Kripke truth-clause (for $\Box p$), has the property of being conjunctive (preserved under arbitrary intersections of P) and monotone increasing in P. Further, this property also holds of any formula $\alpha(P)$ as above in which the predicate R is replaced by any P-free formula $\beta(y)$. Moreover, in the case of neighborhood models or their above equivalents $M = \langle W, \mu, V\rangle$, this property (of μ) is exactly the condition required for M to determine an equivalent Kripke-model, and thus for M to model K. van Benthem's preservation result then establishes the 'converse', namely that $\alpha(P)$ is conjunctive and monotone increasing in P *only if* it is equivalent to a formula of the form

$$\forall y(\beta(y) \rightarrow Py),$$

where β is P-free. It implies, therefore, that the admissible truth-conditions τ are restricted to those for which $\alpha(x,P,\Sigma)$ can be written in the form $\forall y(\beta(x,y) \rightarrow Py)$, where β contains only predicates from Σ.

3 Remarks on the methodology of semantic explanation This brief summary should suffice to indicate that van Benthem has provided a very elegant framework for handling the methodology of possible worlds semantics with the help of instruments drawn in part from the traditionally alien field of the philosophy of science. The latter in fact suggests several further questions and analogies not directly taken up by van Benthem. For example, in light of the above preservation result it would seem that possible worlds semantics exhibits an interesting feature characteristic of empirical theories, namely their *underdetermination* by all possible evidence. If one regards the set of all modal logics extending K as determining the range of potential intensional logical 'data', then the above result implies that many different, conceptually distinct 'T_1'-theories may adequately account for (appropriate parts of) this data.

If we turn to the interplay between truth-conditions and model-conditions, and the general constraints (C1)–(C3), several related issues spring to mind. For instance, in defending (C2) one might try to appeal to the methodological principle of preferring stronger to weaker theories. Of the mentioned alternative ways to semanticize KB, the customary truth-conditions are those which lead to the 'stronger' T_1-theory, given just the usual accessibility relation. However, the question of logical 'strength' here acquires an added dimension, making a straightforward comparison problematic. Viewed simply as first-order theories, the demand that R be symmetric is clearly stronger than the alternative approach imposing no constraints whatsoever on R. But, in moving to the 'nonstandard' theory, a subtle shift has occurred in what might be termed the "correspondence rules" (here represented by τ), so that if one considers the 'theory-plus-correspondence-rules' as a package, it turns out that the two approaches are logically equivalent.

A more convincing justification for (C2) can be derived from considerations not of logical strength but of logical *uniformity*. In the case of normal modal logics, the standard truth-conditions τ can be applied across the board to a wide range of logics of differing strength. This gives rise to the familiar *lattice* picture of normal logics, in which the minimal system K appears at the top of the lattice structure, and progressively stronger modal systems appear further down.[3] The attraction of the standard semantics here is that this picture remains the same whether we treat the modal systems purely syntactically, or whether we introduce the model-theoretic component to obtain an (isomorphic) lattice of first- (or higher-) order classical (T_1) theories. For each system, the interpretative or correspondence rules, linking theory to 'evidence', are the same.

At this point the methodology of semantics crosses paths with the semantics of (scientific) methodology. Currently, there is a consensus among philosophers of science that questions of rational theory appraisal and of scientific progress are best approached by looking not exclusively at individual laws and theories, but by taking account of the broader scientific "paradigms" (Kuhn) or "research traditions" (Laudan) to which the laws or theories concerned belong. Moreover, in physics, for example, there are certain theories, like classical mechanics and quantum electrodynamics, that already constitute the identifiable core of an entire paradigm; and it is often possible to regard these 'comprehensive' theories as being 'composed of' a homogenous and interrelated collection of individual, 'smaller' theories or 'theory-elements'.

This perspective has proved so fruitful in historical and methodological studies of science that it has been directly incorporated into the 'semantics' of empirical theories: in treating the logical structure of empirical science, such homogeneous collections of theories are often formally represented in the shape of "theory-nets", "theory-ensembles", and the like. In such representations there is a core or *base* theory T, containing all the fundamental laws, together with a succession of theories that in different ways extend T by adjoining additional laws (or even new concepts) in order to apply the base theory to a particular problem domain; the whole collection forming a tree-like or lattice-like 'network' of interconnected systems.[4]

The lattice structure of normal modal logics not only fits the formal pattern of such theory-ensembles, there are also many instances in which the intensional operators receive specific 'intuitive' interpretations in order to account for particular domains of application; consider, for instance, the various deontic, epistemic, and doxastic readings of '$\Box$'. Moreover, just as in classical mechanics, say, where there exist domains of application in which different concepts and laws of force jointly apply (e.g., systems subject to both gravitational and electromagnetic forces), so in intensional logic one also finds 'mixed' domains, e.g.

in which tenses and modalities are combined, as in temporal deontic logics of 'historical necessity'.

It seems to be a 'natural' process of scientific change that a research tradition seeks to develop, where possible, within the boundaries of some such 'net' or 'ensemble' structure (at least a semi-lattice), and to extend this structure to account for an ever wider range of empirical problems. The structure will, in fact, to a large extent reflect the heuristic and methodological constraints operating in the research tradition; establishing, for instance, the basic linguistic and ontological categories and the characteristic laws. Most significantly, it is the manner in which the different theories *fit together* (in the lattice structure) that seems to be decisive for characterizing the research tradition and for exhibiting its internal coherence and uniformity. Thus, if we consider possible worlds semantics as comprising nets or ensembles of interrelated theories, we obtain not only strong structural analogies to scientific research traditions of the customary sort, we also acquire additional semantic and methodological support in favor of the kind of *uniform* representation of intensional semantics that the standard picture provides. We have good grounds, therefore, to regard conditions (C1) and (C2) as well-grounded constraints.[5]

The lattice or theory-ensemble representation of theories presupposes, therefore, uniformity of the relevant interpretative or correspondence rules. One is also familiar in the philosophy of science with the idea that empirical anomalies or 'refutations' may usually be 'saved', if not by discovering a more adequate theory, then at least by suitably adjusting the interpretative basis to exclude the 'conflicting' evidence. But here several important methodological factors come into play. First, one tends to think of the interpretative rules as relatively 'stable'; in the first instance one looks for ways to repair the theory, or to reassess the evidence, rather than to reinterpret the entire basis on which evidence impinges on the theory. Secondly, changing the interpretative basis of some theory within the lattice is likely to remove that theory from the ensemble, thus creating a dishomogeneity symptomatic of a 'conceptual problem' (in Laudan's sense) for the research tradition. And, lastly, no change, either to theory or interpretative rules, must be allowed to be *ad hoc.* The first two principles reinforce the rationality of (C2); the third factor is relevant for constraint (C3): ruling out trivial or *ad hoc* semantic explanations.

4 Beyond standard correspondence theory So far we have been considering the paradigm case of normal modal logics. At this stage, therefore, it is natural to ask: To what extent can these proposals and results be adapted to other systems of intensional logic? A first point to make is that the perspective of correspondence theory is a quite general one, available for many types of logic and possible worlds semantics. The case of conditional logics is perhaps the easiest to handle in this respect; indeed, many of the standard treatments of conditionals fit into the present framework without further ado.[6] And a fairly well-developed correspondence theory has also been provided for intuitionistic logics, suggesting that constraints like (C1)–(C3), or at least analogous ones, might also be defended.[7] Moreover, the well-known Gödel translation from Heyting's propositional logic to modal *S4* establishes a natural link between the modal and intuitionistic frameworks.

However, the specific *results* referred to above have a decidedly 'classical' flavor to them. van Benthem's limitative theorem on admissible truth-conditions τ refers directly only to the 'α'-clause of the translation and is based on a preservation theorem for *classical* first-order logic. Likewise, his general incompleteness result depends on taking a classical reading of the propositional connectives, since the neighborhood semantics for which Fine's and Thomason's logics are incomplete assumes this. (Of course, this assumption is the natural one here, because the logics involved are extensions of K.) Moreover, it is *classical* (first- or higher-order) logic which provides the setting of the correspondence theory and the appropriate (model-theoretic or algebraic) instruments for its study. But one might speculate here on alternatives. For example, the idea that semantic explanation comes in the form of the usual completeness theorems was linked to the assumption (implicit above, but explicit in [5]) that the tasks of deducing *validities* and disproving *invalidities* are 'complementary'. This led to (5) as an adequacy criterion for explanations, with its focus on the 'positive' inference, $\vdash^+$. One can, however, take the view that the positive ($\vdash^+$) and the negative ($\vdash^-$) forms represent quite distinct, independent concepts of (non)inference, governed by separate rules.[8] This idea has in fact been explored in the literature, especially with propositional attitude contexts in mind; and semantic treatments are available (e.g. [32]) in which, for example, valuations V are regarded as *partial* functions on propositions, and one distinguishes between the situation $V \vDash A$ ("V supports A") and that of $V \Dashv A$ ("V rejects A"). In such cases, the natural setting for a suitable correspondence theory would appear to be not that of full classical logic, but rather that of *partial* (predicate) logic as developed, e.g., in [9].

Although partial logic thus offers an interesting weakening of the standard correspondence framework, in other contexts it may be appropriate to apply logics stronger than first-order. Let us consider for a moment systems of modal logics *weaker* than the 'minimal' normal logic K, for example the category of *classical* and *quasi-classical* logics. Though general in tone, van Benthem's constraints and observations are very specifically directed at the normal systems, and, as his incompleteness proof demonstates, his own semantics is reducible to the more general perspective of neighborhood models. Thus, although condition (C3) would presumably continue to be applicable in almost any context, (C1) and (C2) seem to make little sense outside the sphere of normal systems. In particular, the requirement that the truth-conditions τ be first-order does not appear to be satisfiable in the case of the 'minimal' classical modal logic, E. Yet the lattice of classical modal logics fits equally well the model of a 'theory-ensemble', based on E;[9] and for this ensemble the neighborhood semantics plays a role entirely analogous to that of the standard Kripke-semantics in the context of normal logics. This suggests immediately the problem of whether there exist suitable generalizations of the neighborhood semantics and appropriate methodological constraints, analogous to van Benthem's generalization of and constraints for the Kripke-semantics. The matter of incompleteness, for example, is already settled for the usual neighborhood semantics, but might still be open with respect to possible generalizations.

This problem can also be studied within the perspective of correspondence theory. In the case of classical modal logics, however, the appropriate correspondence setting is no longer that of ordinary predicate logic, but rather that of a stronger system: predicate logic with added (generalized) quantifiers, $L(Q)$. When the neighborhood or Scott-Montague semantics was first developed as a means to characterize some of the weaker modal systems (notably in [28]), one of the important features of the ordinary Kripke-semantics appeared to have been lost: the natural interpretation of the intensional operators as logical quantifiers. Thus, Hansson and Gärdenfors wrote in [15]:

> A disadvantage with the Scott-Montague semantics is that we no longer have the nice parallelism between $\Box$ and $\Diamond$ and the universal and the existential quantifier as in the Kripke case. (p. 157)

But the parallelism that Hansson and Gärdenfors refer to can be retrieved for the neighborhood semantics providing we think in terms of *generalized* quantifiers. For, just as the modal propositional operators in normal systems correspond to the usual universal and existential quantifiers of first-order logic, so in classical modal systems the modal operators $\Box$ and $\Diamond$ correspond to generalized quantifiers, say Q and Q' (where Q' is dual to Q). Among the quantifiers that come closest to filling this role are those developed in *topological* model theory and in *monotone* logic. For example, one could take as a suitable L_1-system the logic $L(Q)$, for Q an *interior* quantifier, as discussed by Sgro [29] and Ebbinghaus and Ziegler [12] (who use 'I' to denote the interior quantifier). This logic contains the formation rule that if $A(y)$ is a formula, so is $tQyA(y)$, for t a term. Thus, we can provide as before a translation τ from L_0-formulas into L_1-formulas, with the proviso that $\Box A$ is now mapped to $xQy\tau Ay$, where 'x' denotes the 'free' world of evaluation. The semantics for $L(Q)$ then ensures that

$$M \models_w A \text{ iff } M \models_{L(Q)} [x/w]\ Qy\tau Ay,$$

where M is a neighborhood model and '$\models_w$' denotes 'truth at world w'.[10]

In this manner, van Benthem's semantic contraints can be analogously formulated for the classical modal systems. Although the new translation τ above is no longer first-order, hence $L_{\omega\omega}$-elementary, it is nevertheless $L(Q)$-elementary (cf. (C1)). Moreover, by a suitable choice of Q (see note 10) we can arrange that (C2) is also satisfied in the sense that for classical modal systems the new truth-conditions τ account for the valid principles of the minimal logic E, while further modal axioms and rules are semanticized by conditions restricting the class of $L(Q)$ models. Lastly, (C3) can still be applied to rule out *ad hoc* semantic explanations, now under the special condition that no additional predicates Σ are required to explain the intensional operators.[11]

Taken in its widest sense, therefore, one might pursue the idea that the possible worlds 'paradigm' is constituted by several different, basic theory-ensembles, perhaps each with its own style of correspondence theory. The links between them would then also be open to investigation with the help of methods drawn from the philosophy of science. As a final point of interest, consider the dichotomy between theoretical and nontheoretical concepts, as it arises, for example, in the structuralist metascientific framework. As we saw, in van Benthem's treatment it was natural to construe the accessibility relation R, and similar predicates on worlds, as being theoretical in the context of the 'explanatory' theory T_1, whilst his world selection operator μ, introduced at the T_0 level, is intended to be 'nontheoretical' in the same context (of normal systems). Our earlier observation that the T_0-structures $\langle W, \mu, V \rangle$ are equivalent to neighborhood models (as is well-known) now takes on a new significance. Since the neighborhood models actually provide the *semantic* 'explanation' in the framework of *classical* systems, they correspond there already to T_1-structures; so, presumably, in the classical theory-ensemble, μ (or N) is to be regarded as a *theoretical* rather than *nontheoretical* term. This neatly illustrates a phenomenon familiar in the methodology of the empirical sciences, namely that a given concept may be 'observational' or nontheoretical in the context of one theory, yet theoretical in the context of another.

5 Concluding remarks Thus far we have seen that correspondence theory offers a flexible framework within which the principal branches of possible worlds semantics can be characterized and compared. It also reveals striking structural similarities between this and other kinds of scientific research traditions. Moreover, when this analogy with the empirical sciences at large is further developed, we obtain a clearer grasp of the heuristics of the possible worlds program, as brought out, for example, by the relevance of completeness for semantic explanation, and by van Benthem's general constraints on semantic modeling. The latter, indeed, turn out to be reinforced by considerations drawn from the formal methodology of science, and prove to be extendible beyond the confines of the standard correspondence theory.

There are of course limitations to this view and, in concluding, it seems only fair to mention two areas where misgivings are likely to arise. One concerns the strength of the semantic explanation discussed here, as compared say with the explanation of empirical laws and phenomena in the natural sciences. Viewed as explanatory 'theories', the various T_1-systems arising in the possible worlds context do not appear to yield a very deep level of explanation. If we say for example that the logic of obligation is accounted for by the theory of a serial, transitive, euclidean relation of deontic accessibility (i.e., deontic *S5*), it may well be argued that we have 'explained' very little about the nature of obligation. Up to a point this reaction is justified. But, in reply, it should also be born in mind that we are dealing with the *semantics* and not with the *metaphysics* of obligation, necessity and other modalities; and that our 'theory' offers at least a powerful systematization of its domain, even if it does not appear to reveal 'deep underlying causes'. (For that matter, neither does an inverse-square law of gravitational attraction.)[12]

Lastly, in response to a question raised earlier, we should also be aware that correspondence theory itself (standard or nonstandard) does not seem to provide a 'universal' view of possible worlds semantics in its entirety. For example, although there is no difficulty in extending this approach to include the 'nonnormal worlds' of [17] and [28], there is no obvious way to cater to the nonnormal or 'impossible' possible worlds of [16] and [26]–[27], since the very idea of a 'correspondence' seems to break down when we give up the requirement that truth be recursively defined in every world. It remains to be seen, therefore, how the methodology of possible worlds semantics is affected in such cases where the framework of correspondence theory no longer suffices. To this subject we shall return, however, at a later date (see [24]).

NOTES

1. Thus, M^* denotes the inessential variant of M, formed by replacing V by P ($= [\![p]\!] = \{w \in W\colon V(p, w) = 1\}$), Q, etc., for each proposition p, q, etc.
2. The question whether theoretical concepts are in principle dispensible or eliminable has been the subject of extensive analysis and debate in the philosophy of science. On the concept of 'Ramsey-eliminability', see [30], [3], [23].
3. For a pictorial representation of the usual systems in lattice form, see, e.g., [11], p. 132.
4. The term "theory-net" is used in the so-called *structuralist* approach to the philosophy of science, as developed by Sneed, Stegmüller, Balzer, Moulines and others; see, in particular, [1], [2]. A somewhat different, model-theoretic approach to the structure of empirical theories has been developed by Pearce and Rantala ([21], [22]; see also [20]), who use the term "theory-ensemble" in describing the fine structure of a scientific research tradition. For present purposes the differences between these two conceptions need not be discussed in detail. In each case, we are typically dealing with finite sets of theories partially ordered according to logical strength defined in terms of an appropriate 'extension' relation. Moreover, in the standard case, the 'net' or 'ensemble' of theories has a greatest element, and the structure as a whole has the shape of a lattice or semi-lattice.
5. Generally speaking, a scientific research tradition will comprise more than one net or ensemble of theories. This is so in the case of possible worlds semantics, where the lattice of normal modal logics represents merely one 'branch' (albeit a fundamental one) of the whole tradition, just as in, say, classical Newtonian physics. It would be an interesting exercise to try to reconstruct all the different branches of possible worlds semantics as theory nets or ensembles, and to identify the major links between these different structures.
6. We also readily obtain natural lattice structures for conditional logics, e.g. the "V-systems" of [18], p. 131.
7. See [6] and further references given there.
8. A possible analogy here with empirical science might be the asymmetry that is usually assumed between verification and falsification.
9. For a description of the lattice of classical modal logics, see e.g. [11], p. 237.
10. Actually, if we take Q to be a (one-place) interior quantifier in the sense of [12], then $L(Q)$ would correspond to the monotone modal logic M of [11], since they require of the neighborhood model $\langle W, N, V\rangle$ that for each $w \in W$, $N(w)$ is a monotone system over W. By dropping this last condition on N (or on Q), we would obtain an exact counterpart $L(Q)$ to the minimal classical modal system E. An appropriate logic of this kind, called $L(N)$, was first developed by Chang [10]. However, though Chang intends his semantics to generalize the usual semantics for modal logic, he gives an unusual reading to the new operator N: the expression '$NxAx$' is interpreted by Chang to mean "A is necessarily true for person x" instead of "$\Box A$ is true at world x". Moreover, while explicitly acknowledging the influence of Montague's work, Chang makes no reference to the neighborhood semantics of [28], nor does he appear to regard N as a generalized quantifier. By choosing Q to

be equivalent to Chang's quantifier N, the usual stronger (classical) modal systems, like the monotone and the regular logics, can readily be characterized in $L(Q)$ by additional axioms or model-conditions. For a survey and completeness results, see [19].

11. In effect, therefore, $L(Q)$ can be taken to be a *monadic* language, since the only predicates required (in the standard case) are those corresponding to the L_0 propositions; and, according to constraint (C3), they should not appear essentially in specifying the additional model-conditions.

12. The sense in which linguistics generally (and not only semantics) is an explanatory empirical science seems to us at the present time to be still very much open to clarification.

REFERENCES

[1] Balzer, W. and J. D. Sneed, "Generalized net structures of empirical theories, I," *Studia Logica*, vol. 36 (1977), pp. 195-212.

[2] Balzer, W., C. U. Moulines, and J. D. Sneed, *An Architectonic for Science*, D. Reidel, Dordrecht, 1987.

[3] van Benthem, J., "Ramsey eliminability," *Studia Logica*, vol. 37 (1978), pp. 321-336.

[4] van Benthem, J., "Logical semantics as an empirical science," *Studia Logica*, vol. 42 (1983), pp. 299-313.

[5] van Benthem, J., "Possible worlds semantics: A research program that cannot fail?," *Studia Logica*, vol. 43 (1984), pp. 379-393.

[6] van Benthem, J., "Correspondence theory," pp. 167-247 in *Handbook of Philosophical Logic*, Vol. II, ed. D. Gabbay and F. Guenthner, D. Reidel, Dordrecht, 1984.

[7] van Benthem, J., *Modal Logic and Classical Logic*, Bibliopolis, Naples, 1985.

[8] van Benthem, J., *Essays in Logical Semantics*, D. Reidel, Dordrecht, 1986.

[9] Blamey, S., "Partial logic," pp. 1-70 in *Handbook of Philosophical Logic*, Vol. III, ed. D. Gabbay and F. Guenthner, D. Reidel, Dordrecht, 1985.

[10] Chang, C. C., "Modal model theory," pp. 599-617 in *Proceedings of the Cambridge Summer School in Mathematical Logic, 1971*, Lecture Notes in Mathematics 337, ed. A. R. D. Mathias and H. Rogers, Springer-Verlag, Berlin, Heidelberg, New York, 1973.

[11] Chellas, B., *Modal Logic: An Introduction*, Cambridge University Press, Cambridge, 1980.

[12] Ebbinghaus, H.-D. and M. Ziegler, "Interpolation in Logiken monotoner Systeme," *Archiv für mathematische Logik*, vol. 22 (1982), pp. 1-17.

[13] Fine, K., "An incomplete logic containing S4," *Theoria*, vol. 40 (1974), pp. 23-29.

[14] Gerson, M., "The inadequacy of the neighbourhood semantics for modal logic," *The Journal of Symbolic Logic*, vol. 40 (1975), pp. 141-148.

[15] Hansson, B. and P. Gärdenfors, "A guide to intensional semantics," pp. 151-167 in *Modality, Morality and Other Problems of Sense and Nonsense*, CWK Gleerup Bokförlag, Lund, 1973.

[16] Hintikka, J., "Impossible possible worlds vindicated," *Journal of Philosophical Logic*, vol. 4 (1975), pp. 475–484.

[17] Kripke, S., "Semantical analysis of modal logic II, Non-normal propositional calculi," pp. 206–220 in *The Theory of Models*, ed. J. Addison, L. Henkin, and A. Tarski, North-Holland, Amsterdam, 1965.

[18] Lewis, D., *Counterfactuals*, Basil Blackwell, Oxford, 1973.

[19] Makowsky, J. A. and A. Marcja, "Completeness theorems for modal model theory with the Montague-Chang semantics I," *Zeitschrift für mathematische Logik und Grundlagen der Mathematik*, vol. 23 (1977), pp. 97–104.

[20] Pearce, D., *Roads to Commensurability*, D. Reidel, Dordrecht, 1987.

[21] Pearce, D. and V. Rantala, "New foundations for metascience," *Synthese*, vol. 56 (1983), pp. 1–26.

[22] Pearce, D. and V. Rantala, "Constructing general models of theory dynamics," *Studia Logica*, vol. 42 (1983), pp. 347–362.

[23] Pearce, D. and V. Rantala, "Ramsey-eliminability revisited," *Communication and Cognition*, vol. 18 (1985), pp. 157–176.

[24] Pearce, D. and H. Wansing, "On the methodology of possible worlds semantics, II: Nonnormal worlds and propositional attitudes," *Acta Philosophica Fennica*, forthcoming.

[25] Ramsey, F. P., "Theories" (1929), reprinted as pp. 101–125 in *Foundations: Essays in Philosophy, Logic, Mathematics and Economics*, ed. D. H. Mellor, Routledge, London, 1978.

[26] Rantala, V., "Impossible worlds semantics and logical omniscience," pp. 106–115 in *Intensional Logic: Theory and Applications* (*Acta Philosophica Fennica*, vol. 35), ed. I. Niiniluoto and E. Saarinen, Philosophical Society of Finland, Helsinki, 1982.

[27] Rantala, V., "Quantified modal logic: Non-normal worlds and propositional attitudes," *Studia Logica*, vol. 41 (1982), pp. 41–65.

[28] Segerberg, K., *An Essay in Classical Modal Logic*, Filosofiska Studier, University of Uppsala, 1973.

[29] Sgro, J., "The interior operator logic and product topologies," *Transactions of the American Mathematical Society*, vol. 258 (1980), pp. 99–112.

[30] Sneed, J. D., *The Logical Structure of Mathematical Physics*, D. Reidel, Dordrecht, 1971.

[31] Thomason, S., "An incompleteness theorem in modal logic," *Theoria*, vol. 40 (1974), pp. 30–34.

[32] Veltman, F., "Data semantics," pp. 541–565 in *Formal Methods in the Study of Language*, ed. J. Groenendijk, T. Janssen, and M. Stokhof, Mathematical Centre, Amsterdam, 1981.

Freie Universität Berlin
Fachbereich Philosophie u. Sozialwissenschaften I
Institut für Philosophie (WE 1)
D-1000 Berlin 33