JOHAN VAN BENTHEM

THE LOGICAL STUDY OF SCIENCE*

ABSTRACT. The relation between logic and philosophy of science, often taken for granted, is in fact problematic. Although current fashionable criticisms of the usefulness of logic are usually mistaken, there are indeed difficulties which should be taken seriously – having to do, amongst other things, with different "scientific mentalities" in the two disciplines (section 1). Nevertheless, logic is, or should be, a vital part of the theory of science. To make this clear, the bulk of this paper is devoted to the key notion of a "scientific theory" in a logical perspective. First, various formal explications of this notion are reviewed (section 2), then their further logical theory is discussed (section 3). In the absence of grand inspiring programs like those of Klein in mathematics or Hilbert in metamathematics, this preparatory ground-work is the best one can do here. The paper ends on a philosophical note, discussing applicability and merits of the formal approach to the study of science (section 4).

CONTENTS

Synthese **51** (1982) 431–472. 0039-7857/82/0513-0431 $04.20

1. INTRODUCTION

1.1 *Logic and Philosophy of Science*

That there exists an intimate connection between logic and the philosophy of science would seem to be obvious, given the fact that the two subjects were still one until quite recently. Bolzano and Mill exemplify this in their work; but, even in the twentieth century, main currents like Hempel's 'hypothetico-deductive' view of science, or Popper's 'falsificationism' presuppose an obvious, if often implicit, link with logic. This is true to an even greater extent of Carnap's program of 'logical reconstruction' of science, still echoing in contemporary research like that of Sneed 1971. So, what is the problem?

These contacts are, on the whole, rather superficial – going no deeper than elementary logic. The Carnap-Suppes-Sneed tradition is a favourable exception; but, there as well, advanced applications of logic remain isolated examples: occasionally, one encounters Padoa's Method (1901), Beth's Theorem (1953) or Craig's Theorem (1953). Highlights of modern logic, like Cohen's forcing technique or non-standard model theory, have found no applications at all.[1] Moreover, the technical work which is being done often seems to lack contact with actual science.

Are these just maturation problems in an otherwise promising marriage, or is there something fundamentally wrong – calling for a rapid divorce? The latter diagnosis is given by more and more people, following Kuhn and, especially, Feyerabend. In the philosophy of science, they say, the 'logical point of view' is inappropriate, or, at best, inadequate (to the extent of being useless). The scientific reality one should be concerned with is either too 'dynamic', or too 'complex' to be captured by formal tools.[2]

This type of criticism will not be discussed here. Either it amounts to such authors stating *general* personal preferences for different approaches, say history or sociology of science – indeed quite honourable subjects – or, when *specific* complaints are adduced, these will invariably be found to illustrate not so much the inadequacy of logic as that of the author's logical maturity.[3] The problem to be treated here is rather that logicians and philosophers of good will have not yet been able to get a successful enterprise going comparable to foundational research in mathematics.

In the final analysis, there may be deep (and, no doubt, dark) reasons for this failure – prohibiting a non-trivial logic of science. And in fact, some logicians prefer to accept defeat without a struggle, under the cover of manoeuvres like the following. 'Logic is the philosophy of *mathematical* or *deductive* science, the philosophy of *natural* science is for the philosophers.' Or again: 'Philosophy of Science is, by definition, the pragmatic component of (applied) logic'. Both ways, the logician can remain at home. As I see it, however, it is far too early for such conclusions. Logic never had a good try at the theory of science – and this paper is devoted to clearing the ground for such attempts. Like some bearded German once said, the important thing is not to re-interpret the problematic situation, but to change it.

1.2 *Logicians and Philosophers of Science*

It would be an interesting historical project to describe the adventures of logicians in mathematics as compared with those of formal philosophers of science in their fields of study. Subterranean grumblings apart, mathematics has been quite hospitable to logic – even absorbing whole logical sub-disciplines like set theory, model theory or recursion theory. No similar development has taken place, however, in, e.g., physics or biology. Far from experiencing *Schadenfreude*, logicians should worry about this – making common cause with those philosophers of science engaged in foundational or methodological research.[4] Why has this not happened long ago already? Again, historical causes may be advanced but there are also some serious methodological obstacles. These should be mentioned first, so as to look each other straight in the face.

To begin with, there is a difference in 'mentality' between logicians and many formal philosophers of science. Briefly, logicians want *theorems* where these philosophers often seem content with *definitions*. This observation reveals more than just the usual academic animosity between closely related disciplines. To see this, it suffices to compare an (admirable) book like Reichenbach 1956 with Suppes 1973. Reichenbach discusses various formal issues concerning Time without ever formulating problems admitting of deductive solutions in the form of elegant theorems. In the Suppes volume, in contrast, one finds conceptual analyses leading to (and probably guided by) beautiful formal results – say representation theorems coupling Robb's causal analysis of Space-Time with the Special Theory of Relativity.

This difference in mentality may well reflect a difference in goals: say, formal 'explications' in Carnap's sense as an aim in itself versus formal definitions as a means for obtaining desired theorems. Put another way, Frege and Hilbert did not formulate their formal conception of theories in order to publish papers called 'the logical structure of arithmetic', but in order to carry out their *programs* (derivation of all arithmetical statements as laws of pure logic; proofs of consistency). But, does not the philosophy of science know such guiding programs as well, say that of the 'Unity of Science'? In a sense, yes, but there is a subtle difference. The above mentioned logical programs made claims which were *falsifiable*[5]; and indeed they were falsified – witness, e.g., Gödel's Incompleteness Theorems. It is mainly this characteristic which made them so fruitful. (Compare the analogous case of 'squaring the circle'.) In contrast, it is hard to see how programs like the Unity of Science, or even its implementations like 'fitting each scientific theory into the Sneed formalism'[6] could be refuted at all. Who is going to unfold the really inspiring banner?

To a certain extent, the preceding paragraphs amounted to a polite invitation to philosophers of science: invest more heavily in technical logical theory. But, on the other hand, there looms the equally deplorable obstacle of the self-imposed isolation of logic. Certainly, nowadays many first-rate logicians are opening up to problems outside the familiar circle of mathematics, notably those of the semantics of natural language. But the boundaries of logic should be set 'wider still and wider', or so I will now try to argue.

1.3 *Logica Magna*

Contemporary logic is a flourishing discipline, both as 'mathematical logic' (witness the handbook Barwise 1977) and as 'philosophical logic'. Therefore, organisational schemes like the one to be presented below might well be thought superfluous – being rather the symptom of a subject in its infancy. (The most grandiose conceptions of logic were drawn up in times of logical stagnation.) The only excuse for the following is its modest aim, namely to make people realize (or remember) what *more* logic is, or could be.

Logic I take to be the study of reasoning, wherever and however it occurs. Thus, in principle, an ideal logician is interested both in that *activity* and its *products*, both in its *normative* and its *descriptive* aspects, both in *inductive* and *deductive* argument.[7] That, in all this, she is looking for stable patterns ('forms', if one wishes) to study is inevitable, but innocent: the assumption of regularity underlies any science. These patterns assume various shapes: the 'logical form' of a sentence of inference, the 'logical structure' of a book or theory, 'logical rules' in discourse or debate.

Given any specific field of reasoning, the ideal logician chooses her weapons. Which level of complexity will be attacked: sentences, inferences, texts, books, theories?[8] Furthermore, which perspective is most suitable; syntactic, semantic or pragmatic? Finally, which tools are to be used in the given perspective: which formal language, which type of theory of inference and of which strength? Thus she decides, e.g., to study certain ethical texts using a tensed deontic predicate logic with a Kripkean "world course" semantics – or a theory of quantum mechanics using a propositional language receiving a pragmatic interpretation in terms of verification games. Even so, many aspects of the chosen field may remain untouched by such analyses, of course. Fortunately, then, there are various neighbouring disciplines with related interests to be consulted.

On this view there is no occasion for border clashes, but rather for mutual trade with not just *philosophy*, *mathematics* and *linguistics*, but also, e.g., with *psychology* and *law*. These are not idle recommendations, but important tasks. An enlightened logician like Beth, for instance, realized the danger of intellectual sterility in a standard gambit like separating the *genesis* of knowledge in advance from its *justification* (thus removing psychology to beyond the logical horizon, by definition) – witness Beth and Piaget 1966. Another type of project which should become culturally respectable among logicians is the systematic comparison of *mathematical* and *juridical* modes of reasoning (cf. Toulmin 1958). But, not even mathematical logic itself covers its chosen field in its entirety. A book like Lakatos 1976 makes it clear how eminently logical subjects – on the present view of logic, that is – have fallen out of fashion with the orthodox mathematico-logical community.

Finally, it may be noted that such cross-connections would also provide the dishes which, after all, a spice like 'logical awareness' is supposed to flavour. Only the most insensitive palates tolerate spices in pure form ....

2. FORMAL NOTIONS OF 'THEORY'

The key concept in the study of science would seem to be that of a 'theory'.[9] In accordance with the remarks made in section 1.2, we will approach its formal study in this section keeping in mind both definitions and pleasant results. First, a short historical sequence of definitions is given (2.1), which may already be richer than most logicians are aware of. Nevertheless, it should be emphasized that this is not a representative historical account, but a didactical tale for a rather special purpose. Then, there is a review showing how versatile and flexible an arsenal modern logic provides for the systematic development of such definitions (2.2). Further logical theorizing, in order to produce subsequent results, is the subject of the following section (3). Hopefully, the intellectual interest of the enterprise advocated here will become clear as we proceed.

2.1 *A Short History from Hilbert to Sneed*

To many people, 'the logical view' of a theory is that of a *formal system*, whose components are a formal language, a set of axioms and an apparatus of deduction deriving theorems from these. It is one of the amazing achievements of the early modern logicians that they managed to do so much with this extremely unrealistic notion. Nowadays, it has become fashionable to blame the formalists for the 'poverty' of this concept, when compared with actual practice (that they would never have denied this is conveniently forgotten). But, given their aims, one should rather congratulate them for their happy choice of this austere but fruitful notion. Like so often in the development of science, it paid to be simple-minded.

Nevertheless, different aims may call for richer concepts. E.g., in many cases one needs the additional observation that developing a theory consists in a judicious interplay of proof *and definition*. Thus, definability becomes a logical concern of equal importance with derivability. The complications to be considered here are of a different kind, however – as will appear from the following sequence, whose main theme is how to account for the additional complexity of *empirical* theories in natural science. Still, our story begins inside mathematics proper:

2.1.1 *David Hilbert*

As is well-known, Hilbert's Program of consistency proofs presupposed the above-mentioned view of mathematical theories, which had developed in the course of millennia of geometrical studies. But, moreover, it was based upon a global view of mathematics as consisting of a 'finitistic' core surrounded by more abstract hull theories like analysis or set theory. The core consists of simple concrete manipulations with numbers; say, encoded in some fragment of arithmetic. These threatening to become extended beyond human comprehension, 'higher' theories (possibly adding infinite objects) are invented, amongst others, to speed up proofs and indeed the very process of arithmetical discovery. (Cf. Smorynski 1977.)

Thus, one could formalize a typical mathematical theory as a *two-stage* affair: a 'concrete' part T_1 (with language L_1) translated into some 'abstract superstructure' T_2 (with language L_2), or maybe contained in some mixed theory $T_{1,2}$ (with language $L_1 + L_2$). These two set-ups are obviously related: for convenience, the latter will be discussed henceforth.

For Hilbert, the consistency of T_1 was beyond doubt: but that of $T_{1,2}$ was not[10] – whence the attempt to prove it by means within the range of T_1. Another side of the matter was pointed out by Kreisel: Hilbert assumed that such abstract extensions did not create new concrete insights (they only make it easier to discover proofs for them). Formally, this means that $T_{1,2}$ is a *conservative extension* of T_1:

if $T_{1,2} \vdash \varphi$ *then* $T_1 \vdash \varphi$, *for all* L_1-*sentences* φ.

All this was refuted by Gödel. (Take for T_1: Peano Arithmetic, and for $T_{1,2}$, say, Zermelo-Fraenkel Set Theory – φ being the relevant Liar Sentence.) Nevertheless, the above notion of 'theory' remains interesting, especially in view of later developments.

Before turning to these, one more aspect of Hilbert's views is to be noted. It would be rash to assume that the underlying *logic* of T_1 will be predicate logic. For one thing, one might want to have a more constructive logic at this level[11] – but this is not the issue here. It is rather that the *complexity* of L_1-statements is important. Clearly, *individual* statements, like $7+5=13$ are 'concrete' – but what about quantification? Hilbert seems to have allowed *universal* statements, like $\forall xy\, x+y<x\cdot y$; but not, e.g., quantifier combinations like $\forall x \exists y\, \varphi(x, y)$ – unless these could be reformulated constructively. That such combinations are rather 'abstract' is brought out by the Skolem-equivalence introducing theoretical functions:

$$\forall x \exists y\, \varphi(x, y) \leftrightarrow \exists f \forall x\, \varphi(x, fx).$$

Only when such an f may be given constructively – say $\forall x \exists y x<y$ as $\forall x x<Sx$ – is this permitted. Thus, one uses only a fragment of predicate logic in formulating T_1.[12,13]

2.2.2 *Frank Ramsey*

Hilbert's view of theories was inspired by *mathematics*. One could say, perhaps, that the 'finistic' part introduces *empirical* considerations (actual manipulation of concrete symbols), but only in a very weak sense. A little later, in 1929, Ramsey wrote a short paper called 'Theories', which lay dormant for a long time. (Cf. Ramsey 1978). In it, he describes the role of theories, also and notably empirical theories, as follows. First, there is the so-called 'primary language' containing the prima facie descriptions of our various experiences. These descriptions may be observed regularities (*laws*) or reports of individual observations (*consequences*). Then, in order to systematize this complex field, one introduces a so-called 'secondary language', with a higher level of abstraction. A *dictionary* will translate 'primary' concepts into 'secondary' ones, and the latter will be organized by means of some simple set of *axioms*, from which the original laws and consequences are derivable.

Evidently, there are similarities with the Hilbert view. Notice, for instance, how 'laws' correspond to the above *universal* statements. For empirical theories, this restriction is even more reasonable than in the original mathematical context. E.g., one may observe particles at certain positions, or state universal extrapolations from these, like 'in this trajectory each position is occupied at most thrice'. But, a localization principle like 'each particle is *somewhere* at each instant' amounts to the theoretical claim that a *position function* exists. Another analogy lies in the possible uses of various conditions on the relevant Skolem functions. In the Hilbert case, one would like these to be 'constructive'. In the empirical case, one would like them to be *continuous* (and hence approximable by means of constructive or measurable functions).[14]

In order to bring out the postulated character of the secondary language, Ramsey stressed that a theory could be regarded as an *existentially quantified second-order sentence*, viz. $\exists Y(A(Y)$ & $D(Y, X))$ – where Y is the secondary vocabulary, A the 'axioms', and D the 'dictionary' coupling Y to the primary vocabulary X. Thus, not too much ontological reality should be ascribed to theoretical entities: a methodological turn which is quite typical for Ramsey's philosophy in general. Notice that the whole point is based upon the following simple logical observation:

$$\textit{if } A(Y), D(Y, X) \vdash S(X), \textit{ then } \exists Y(A(Y) \ \& \ D(Y, X)) \vdash S(X);$$

where S is any primary statement. There is a connection here with Hilbert's Program (cf. Smorynski 1977): if you have 'primary' reasons for believing in $\exists Y(A(Y)$ & $D(Y, X))$ (*consistency*) then you have primary reasons for believing in the truth of any primary consequence $S(X)$ of the whole secondary theory (*conservation*).

This syntactic point of view is quite true to the spirit of Ramsey's work, which appeared in the pre-semantical stage of modern logic. Its full impact only became clear in a semantical setting, however, as we shall see later on.

Intermezzo. In authoritative text books on the philosophy of science, like Nagel 1961, one finds the so-called 'statement view' of scientific theories. There, such a theory is taken to possess an *observational* as well as a *theoretical* vocabulary. The first is interpreted by actual inspection, measurement, and so on – the second receives a 'partial' interpretation derived from the first through so-called 'correspondence principles'. This division (taken for granted in Ramsey's approach) turns out to be difficult to implement precisely. E.g., original strong proposals, like *explicit definability* ('reduction') of theoretical vocabulary to observational terms, turned out to be untenable. Nagel is not unaware of relevant logical work. He discusses Craig's Theorem, which may be taken to mean that, if one can find a recursive axiomatization for some mixed 'observational/theoretical' theory, then one can also find such an axiomatization for its observational part alone: a weaker form of reduction.[15] Although there are certain 'semantic' elements in this picture, its main emphasis is still syntactical.

It is interesting to find that Nagel does refer to Ramsey's views, be it in passing, in his discussion of *instrumentalist* views of theories versus *realist* ones. On his account, the Ramsey sentence was designed to turn an otherwise undetermined theoretical statement *form* (meant instrumentalistically as a 'convenient short-hand' for some complex set of observation reports) into a determinate *statement*. (Our account turned things the other way around: an otherwise too strong statement about specific theoretical entities was weakened to a merely existential statement. Nevertheless, Nagel is right in classifying Ramsey on the instrumentalist side – and it is ironical that the revived 'semantical' Ramsey has become popular with latter day realists.)

In order to pick up the main thread of our story, we go back to 1947, when E. W. Beth proposed a 'semantic conception of theories' (cf. Beth 1947), still alive in the work of authors like F. Suppe or B. C. van Fraassen. Briefly, Beth proposes to view theories as describing certain *structures* (physical systems, 'histories', or what you like), on the pattern of Tarski semantics. Thus, e.g., *laws* of the theory may be viewed as restrictions on the trajectories open to the system in phase space. These were prophetic ideas – and that before the first flowering of logical model theory.

A more influential author propagating similar views is P. Suppes, whose 'set-theoretic predicate' approach to theories was formulated and applied in the fifties already. (Cf. Suppes 1957, Suppes 1960.)[16] E.g., Newtonian particle mechanics would be defined as a class of set-theoretic structures (time-dependent systems) satisfying certain Newtonian conditions formulated in set-theoretic terminology. These may be taken as the (syntactic) axioms of mechanics – set theory then serving as the underlying apparatus of deduction; but this is not usually made explicit – let alone a full first-order predicate-logical axiomatization. We will have more to say about such a 'language-free' approach (as it is sometimes called) below. For the moment, just note the historical precedent of Klein's 'Erlanger Program' in which geometry was approached purely 'structurally' through the study of (invariants of) certain groups of transformations on geometric spaces.

Suppes did use model-theoretic tools occasionally – e.g., Padoa's method in his discussion of Mach (cf. Suppes 1957). Nevertheless, the set-theoretic predicate approach does not fit in smoothly with the perspective of logical model theory – being biased on the structural side, so to speak. Therefore, our next author is more in line with the present exposition.

2.1.3 *Marian Przełęcki*

There is a strong Polish tradition in formal semantics of scientific theories, and the little book Przełęcki 1969 provides an elegant example of this. (Its inspiration is derived partly from earlier work by Ajdukiewicz. Cf. Giedymin 1978.) The author presents a model-theoretic account of empirical theories in the following steps. To begin with, there is an *observational language* L_0 whose predicates are interpreted in a certain concrete domain U_0 by means of *ostension*. (More precisely, by means of paradigmatic cases and extrapolation from these.) This domain U_0 is gradually enlarged by us to an empirical domain U of 'physical objects'. (Thus, electrons and galaxies join dishes and babies.) L_0-structures will be all those structures with domain U which contain U_0 (with its fixed interpretation) as a *substructure*. Next, a theoretical vocabulary L_t is added, calling for suitable enrichment of the L_0-structures. The relevant model-theoretic concept is that of expansion, yielding $L_0 + L_t$-structures consisting of some L_0-structure to which interpretations have been added (without changing the domain U) for the L_t-terms. Yet not all such expansions will do: 'meaning postulates' are necessary in terms of the L_t-vocabulary. Finally, the syntactic part of the theory is a set T of first-order $L_0 + L_t$-sentences defining some class of $L_0 + L_t$-structures.

This framework gives rise to various logical questions, e.g. concerning different kinds of *definability* of L_t-vocabulary in terms of L_0-vocabulary. Indeed, the book contains a wealth of notions and results which cannot be reviewed here. (Recall that this is not meant to be a comprehensive historical survey.) Probably the most striking one is Przełęcki's attempt to separate a theory T into an 'analytical' part A, containing meaning postulates without empirical import, and a 'synthetical' part S doing the real work of making empirical claims (by excluding certain L_0-models). The People's Republic of Poland has invested many years of research into this elusive equation $T =$

$A+S$, without finding satisfactory answers. Nevertheless, in this connection, some interesting questions have been raised. E.g., A should clearly be 'semantically non-creative'; in the sense that it should not exclude L_0-structures: each L_0-structure must be expandable to an L_0+L_t-structure which is a model for A. A true model-theorist, Przełęcki wants to find the syntactic side of this coin. Is it that A is 'syntactically non-creative', in the sense that the only L_0-sentences derivable from it are the universally valid ones?[17] This type of question is the precursor of a more general one (treated by Przełęcki in related publications) to which we now turn.

The L_0+L_t-theory T has a class of models MOD(T). Accordingly it may be said to describe the class MOD(T) $\upharpoonright L_0$ of those L_0-structures which can be expanded to L_0+L_t-structures in which T holds. Notice the way in which Ramsey's idea is implemented: this class of L_0-structures receives a 'second-order existential' description in this way. Now, let us return to the original syntactic version of Ramsey's ideas, in the following form. We have an L_0-theory T_0 which is contained in the L_0+L_t-theory T; i.e.,

(1) *if* $T_0 \vdash \varphi$ *then* $T \vdash \varphi$, *for each* L_0*-sentence* φ. (*Extension*)

Then, what about the connection between MOD(T_0) (those L_0-structures which are models of T_0) and the class MOD(T) $\upharpoonright L_0$ described by T? Clearly, at least, it should hold that T does not exclude models of T_0:

(2) MOD(T_0) *is contained in* MOD(T) $\upharpoonright L_0$. (*Ramsey Extension*)

But, the above syntactic requirement does not guarantee this: maybe T has L_0-consequences outside of T_0. Now, there was also the old Hilbert requirement of conservative extension, reading

(3) *if* $T \vdash \varphi$ *then* $T_0 \vdash \varphi$, *for each* L_0*-sentence* φ. (*Conservative Extension*)

And, indeed, (2) implies (3), as is easily seen. (Notice that (2) does nót imply (1).) Do the two conditions match precisely? Unfortunately, the answer is negative.

Counter-example: Let T_0 be the complete first-order theory stating that Time has a beginning 0, and proceeds from there via a 1-1 successor operation S ('to-morrow') never making loops. (Infinitely many axioms are needed to secure this.) The latter phenomenon is explained in the (finitely axiomatizable) theory T having a transitive relation B ('before') such that $\forall xy(Sx = y \rightarrow Bxy)$, as well as a property E ('early') such that E0, $\forall x(Ex \rightarrow ESx)$, $\forall x \exists y(Bxy \ \& \ \neg EY)$, $\forall x(\neg Ex \rightarrow \forall y(Bxy \rightarrow \neg Ey))$. (Time goes from 'early' to 'late', never to return.) Now, the natural numbers $\langle N, 0, S\rangle$ form a model for T_0 which cannot be expanded to a model for T.[18] (More 'empirical' counter-examples would take too much space to present in full here.)

Thus, this typical model-theoretic attempt to establish a duality has failed. Still, there is a result which one can prove, viz. the equivalence of Conservative Extension with the following weaker version of Ramsey Extension:

(4) *each L_0-structure in* MOD(T_0) *has an L_0-elementary extension to some $L_0 + L_t$-structure in* MOD(T).

Thus, the condition that the original empirical situations (L_0-structures) be 'enrichable' with suitable theoretical entities to $L_0 + L_t$-structures which are models for T, is now related to their being thus 'extendable' – addition of individuals becoming allowed as well. Given the fact that this is a not unknown procedure in physics (postulating new particles, or whole planets) model-theoretic curiosity has led to a not unreasonable amendment on Ramsey's views.

These considerations will have given an impression of Przełęcki's model-theoretic spirit. Now, finally, we turn to maybe the most influential author on formal philosophy of science in recent years:

2.1.4 *Joseph Sneed*

In his book Sneed 1971, this author gave an analysis of classical particle mechanics using a formal machinery which transcends the above in various interesting respects. Since this work has become well-known (largely through allied work; cf. Stegmüller 1979), it will suffice to mention only those ideas which are relevant to the present discussion.

In the above picture of a theory, with ingredients T_0, MOD(T_0), T, MOD(T), it is implicitly assumed that a division into non-theoretical ('observational', if you wish) vocabulary (L_0) and theoretical vocabulary (L_t) has been effected satisfactorily. In practice, however, this is a difficult problem, as we have seen – and Sneed gives an ingenious solution (the exact nature of which need not concern us here).

Another fruitful practical perspective of his work is the emphasis on what it means to *apply* a theory to a given empirical situation. On the Ramsey view, adopted by Sneed, this becomes: to show how theoretical ('super-natural') concepts may be introduced (forces, wishes, and so on) turning the situation into a model for the whole theory T.

Will such an addition always be possible? More precisely, will any model of $T \upharpoonright L_0 = \{\varphi \in L_0 | T \vdash \varphi\}$ (clearly, the most 'fitting' choice for the empirical part T_0) be expandable to a model for T? The answer is negative: evidently, T is a conservative extension of this T_0, and we are asking if it is also a Ramsey extension. Thus, the problem remains like in the preceding sub-section. In Sneed's terminology, then, the theoretical vocabulary need not always be 'Ramsey eliminable' from the theory. If one wishes this, it will have to be postulated.[19]

Next, we turn from questions about 'at least' to questions about 'at most'. Will just any theoretical expansion do? (Recall that Przełęcki introduced 'meaning postulates' at this point.) This seems implausible. Thus, suitable restrictions are to be formulated. Sneed considers the requirement that, given any model for T_0, there be a *unique* expansion of it to a model for T. Model Theory gives one the cash value of this proposal for a large class of languages. By Beth's Definability Theorem, this 'implicit definability' amounts to *explicit definability* of L_t in terms of L_0 (on the basis of T). Thus, old-fashioned reductionism would re-emerge!

This line of thought should not be pursued, as Sneed realized, because it is too 'local'. The kind of restriction one wants is a more 'global' one, concerning cross-connections between expansions of different empirical situations. What we have in reality is a *class* of empirical situations, which are expanded *simultaneously* in such a way that certain 'constraints' are obeyed. Two typical examples of such constraints are the following

(C1) particles in different empirical situations should receive the same mass values through different expansions. ('Mass' is a theoretical function in Sneed's analysis.)

Thus, the earth gets the same mass in the Solar System as in the Milky Way.

(C2) the physical join of two particles receives the sum of their separate mass values
(whether present – at some earlier or later time – in the situation considered, or not).

Thus, the picture of an empirical theory now becomes, say, $\langle T_0$, MOD(T_0), T, MOD(T), $C\rangle$ – where C is the set of constraints. *Applying* such a theory to a class E of empirical situations means finding a simultaneous expansion of E satisfying C such that the resulting class is contained in MOD(T). (Obviously, at least, E will have to be a 'suitable' candidate, in the sense that it is contained in MOD(T_0).)

In later publications (cf. Balzer and Sneed 1977/8), this picture is stripped of its linguistic content, leaving a purely set-theoretic 'language free' formulation in the Suppes tradition. E.g., MOD(T) is replaced by an arbitrary class X of L_t-structures while C becomes a class of subclasses of the $L_0 + L_t$-universe satisfying certain conditions. (Closure under set-theoretic inclusion is one of these.) This move will be discussed in section 2.2.

Clearly, it is the notion of 'constraint' which complicates the above picture from a model-theoretic point of view. Thus, it is worthwhile to speculate a little about possible alternatives. In fact, there seem to be several promising escape routes, open to further exploration. The first is to have *one single domain* for the language L_0 (as already proposed in Przełęcki 1969), considering the 'empirical situations' as its subdomains. Ramsey expansion would then apply to the whole 'universal' L_0-structure: immediately guaranteeing constraints like (C1) above.[20] Possible objections might be that this becomes *too* global, and also that it is not obvious how *all* constraints may be captured without further complications. As for the 'globality', notice that one need not think of a literally 'universal' domain: it would suffice to consider a class of L_0-structures directed by model-theoretic inclusion – introducing their *direct union* for convenience only. This observation opens up a second route which is conceptually attractive as well.

Up to now, we have followed ordinary model theory in assigning a class MOD(T_0) of 'isolated' L_0-structures to a theory T_0. But, already in mathematical practice, one is normally confronted with 'structured' classes of such models, with relations and operations connecting them. Thus, it would be natural, e.g., to replace MOD(T_0) by the *category* CAT(T_0) of T_0-models *plus* their connecting natural morphisms. The particular choice of morphisms will then be dictated by the necessities of the theory considered. E.g., one might think of *isomorphic embeddings* (as substructures). On this view, constraints come in as follows. One wants to expand a class of L_0-structures *together with* its categorial structure: i.e., the expanded class should be an L_t-subcategory – the original morphisms remaining morphisms.[21] Put fashionably, the *forgetful functor* should yield CAT(T_0) when applied to CAT(T). Some reflection shows that this takes care of the constraint (C1) as well.[22] As it stands, the requirement will even be too strong in many cases. E.g., it seems reasonable that *isomorphic* empirical situations (being equivalent up to a co-ordinate transformation, say) receive isomorphic expansions; but not each single L_0-isorphism need remain an $L_0 + L_t$-isomorphism! (E.g., not each purely *topological* automorphism of Space need remain a *metrical* isomorphism, when Space is metricized in accordance with the topology.) This second possibility may well have considerable model-theoretic interest in itself, whether it accommodates all constraints or not. (An example will be found in the appendix, section 5.)

This concludes our historical survey. A very rich and promising formal notion of 'theory' has emerged – whose logical study is still in its infancy. For example, the fate of the usual *properties* of and *relations* between such theories (normally defined in terms of *formal systems*) is still to be explored. Some relevant suggestions and discussions will be found in section 3. Finally, it should be remarked that this survey is by no means representative of contemporary logical views of theories, being inspired by a model-theoretic standpoint. Thus, interesting different tracks have remained outside our scope (cf., e.g., Wessel 1977). That there are new frontiers for logicians will have become abundantly clear, however; and that was our goal.

2.2 *A Systematic Logical Perspective*

A phrase like 'the logical structure' of a theory is misleading in that it suggests one single (or one preferred) logical approach to the study of science. Certainly, there is one single logical research mentality, but an aspect of that is precisely the *plurality* of available logical approaches. Thus, depending on one's specific aims, there are many choices available (cf. section 1.3). For the study of scientific theories, in particular, the following perspectives should be kept in mind.

2.2.1 *Syntax*

It was remarked in 2.1 already how the syntactic notion of a *formal system* turned out to be a happy choice for many enquiries. It was also recalled that this notion arose from a long development of Euclidean geometry – in which the axiomatic organisation of deductive knowledge proved its value as a means for efficient organisation and a stimulus for rapid development of deductive knowledge. Even so, one has a variety of syntactic approaches. E.g., in the actual computer formalization of mathematics in the Eindhoven AUTOMATH project (cf. Van Benthem Jutting 1979), one needed a syntactic *text logic* consisting of a version of the typed lambda-calculus in a natural deduction presentation. Apart from providing a catalogue of truths, such a system also provides a model for actual mathematical prose. (Indeed, modern computer languages are a natural and interesting extension at the text level of the syntactic approach in logic.) Similar projects for the *natural* sciences would no doubt inject invaluable practical experience into the now often purely a priori debate concerning the possibilities of their formalization.

2.2.2 *Structures*

Before the 19th century, no distinction was made in geometry between syntactic derivation and spatial intuition. After the advent of non-Euclidean geometry, this separation led to the concept of geometrical *structures* ('spaces') modelling certain purely syntactic sets of geometrical sentences. The resulting interplay is exemplarily illustrated in, e.g., Tarski 1959. But, purely 'structural' views of theories arose very early too: witness Klein's 'Erlanger Program' mentioned in section 2.1, or Poincaré's group-theoretic approach to geometry and mechanics.

In the natural sciences, lip service was paid for a long time to axiomatic ideals; but, in practice, structural views lie even more at hand there. E.g., Newtonian mechanics might easily be identified with a certain class of real 'mechanical systems'. Which brings one to the question just which 'structures' are most suitable for use in the logic of *natural* science.

For a start, let us quickly dispose of a 'philosophical' pseudo-objection. 'Mathematics is about abstract structures, natural science about non-formal Reality'. This is true; only it is too true. What natural scientists deal with are *models* or *representations* of reality, and these go into the logical study of science. So, which models? True, the ordinary model-theoretic notion of 'structure' seems rich enough in principle to accommodate any kind of system (cf. Suppes 1960, Montague 1974) – but it would also be natural to try and work with *systems* in the sense of general system theory, because of their explicit *time-dependent* presentation. (Cf. Arbib and Padulo 1974.) E.g., system-theoretical *representation theory* contains many 'logical' lessons.[23]

In this connection, it may be of interest to observe that a classical paper in the logical study of quantum mechanics formulates its main result on 'hidden variables' (a topic with a syntactical ring) in structural Hilbert space terms, rather than as a result about classical and quantum-mechanical axiomatic theories. (Cf. Kochen and Specker 1967.)

Does not all this mean that the ladder of linguistic form can be thrown away, now that one has arrived at the structural 'reality' behind scientific theories? This would seem to be the 'set-theoretic view' discussed earlier on. And indeed for many problems this is a perfectly sensible thing to do – witness much work by Sneed and Suppes. E.g., the obvious problem which presentations of systems describe the same 'real' system (in different co-ordinates, say) is a structural problem about suitable isomorphism classes. Or, when Ashby talks about relations of analogy between finite machines (cf. Ashby 1976), the insight that mutually homomorphic machines are isomorphic owes nothing to underlying syntactic theories. Other interesting questions may require the full structural-linguistic perspective, however: as will be illustrated below.

If this be so, then why the friction one sometimes observes between the 'set-theoretic' and the 'model-theoretic' approach? (Cf. Przełęcki 1974.) Inevitably, in such cases, there are priority disputes poisoning the atmosphere: whose perspective is 'better'?[24] E.g., on the set-theoretic side, it is claimed that one loses *nothing* – except maybe troubles (with non-standard models for mathematical theories, which one did not want anyway). This seems short-sighted, however. Why not leave the door open, e.g., to 'non-standard mechanics', when mathematicians are just discovering the delights of non-standard analysis?

Throwing away the linguistic ladder forever means cutting oneself off from the original *motivation* of many structural concepts.[25] A similar warning was given by Dieudonné to mathematicians wanting to apply logical gadgets like ultraproducts without being bored with stories about their logico-linguistic origin. This is possible, of course, but one deprives oneself from understanding their full importance, as well as the heuristic fuel for further discoveries in the same vein.

2.2.3 *Semantics*

Since logical model theory is such a well-known field (cf. Chang and Keisler 1973), it will suffice here to draw attention to just a few relevant points. To begin with, here is a short list of questions concerning theories which are meaningful only in a semantical perspective.

(1) Does mechanical 'determinism' as a structural notion – that is, given a past history, only one continuation is possible into the future (in the relevant set of structures) – imply 'determinism' as a linguistic notion: i.e., the relevant state variables at time $t+1$ are explicitly definable in terms of those at time t – on the basis of the mechanical axioms? (Cf. Montague 1976: 'Deterministic Theories'.)

(2) In geometry, "a point relation is said to be objective if it is invariant with respect to every automorphism. In this sense the basic relations are objective, and so is any relation logically defined in terms of these. [...] Whether every objective relation may be so defined raises a question of logical completeness [...]' (Weyl 1963, p. 73): a model-theoretic question, to be sure. (Cf. the appendix (section 5) for some relevant results.) Finally, the direction may also be reversed, as in

(3) Do linguistic 'relativity principles' like Einstein's – viz. that 'physical equations' be invariant under the Lorentz transformation – admit of a structural characterization?

One could ask scores of similar questions, say about the relation between 'linear' *functions* and 'linear' polynomial *forms*, all of them illustrating the same peculiarity of model theory: to look for a systematic *duality* between syntactic and structural points of view.

Another point which deserves constant attention is the following.

The great success of plain Tarski semantics should not make one forget that, in principle, *any* link between language and structures is of model-theoretic interest. Thus, e.g., non-classical semantics of the forcing variety, or the 'vague' semantics of Fine 1975 are equally respectable. In particular, then, the choice of a certain class of structures does not commit one yet to a fixed underlying apparatus of deduction: classical, intuitionist, or otherwise.[26] This observation, by the way, provides one with an excellent opportunity to make more exact sense of the famous 'Revisability Thesis' in Quine 1951.[27]

2.2.4 *Pragmatics*

Between them, syntax and structures (and hence: Model Theory) would seem to exhaust all possible logical perspectives upon theories. But, moreover, many pragmatic themes concerning our actual *handling* of the above notions can be studied by logical means. Thus, theories as scientific *activities* rather than *products* of such activities are not irrevocably outside the scope of logic. Just a few examples will have to provide the backing for this claim here.

In the purely syntactic perspective, there is already the matter of the *heuristic* of proofs. (Cf. Lakatos 1976, mentioned in section 1.3, or Hintikka and Remes 1974.) Moreover, the logical study of the actual defense or attack of sentences in dialogues has been initiated in works like Lorenzen and Lorenz 1978.

In the semantic perspective too, there is room for pragmatic studies. E.g., model theory presupposes that successful interpretation has taken place already. How? Here is where 'game-theoretical semantics' in the style of Hintikka may become useful (cf. Saarinen 1979). There is also a connection here with actual *measurement* (cf. note 15), as well as with Giles' approach to physical theories (cf. Giles 1979). Indeed, the volume Hooker 1979 signals an interesting 'logico-pragmatic' turn in the study of physical theories. Nevertheless, it will be clear that these references are only the first landmarks in a hopefully fruitful new area of logic.

Finally, the pragmatic perspective provides an appropriate occasion for stating a concluding remark about the notions considered here. Exclusive concentration upon theories as intellectual products may reveal a lot about the 'edifice of science', but it may also generate pseudo-problems. E.g., is the rationality of science in jeopardy when two of its beams do not fit exactly – say, Newtonian mechanics and Einsteinian mechanics? Will 'continuous growth' be endangered? Of course not: the continuity of science is guaranteed by the (re-) construction worker's know-how of common intellectual rules, strategies, and the like. Take away specific languages, axioms, or even whole theories: this cat will still retain its rational grin.

3. FORMAL QUESTIONS CONCERNING THEORIES

The preceding section illustrates how logic may be viewed as the study of non-trivial scientific theories.[28] More specifically a whole *range* of logical considerations and tools was reviewed. A theory should always be studied with specific questions in mind and these call for a suitable *choice* of logical perspective. This brings us back to one of the observations made in section 1.2 about logical research being guided by the lure of theorems on the horizon. Here a related question will be treated. What becomes of the logician's pet theorems when her subject is used in the present context? To take just one curious weed which is rather dear to her: does *Lindström's Theorem* have any significance in this more general area? In a sense, yes, insofar as this theorem (and abstract model theory in general) is concerned with the general properties of logical languages plus semantics. It has been argued that scientific theories may be based upon quite different logics; and abstract model theory may help provide reasons for choices, by telling one which pleasant meta-theorems will be gained or lost in each case. (Cf. Pearce 1980A for an application of abstract model theory to the Sneed framework.)

More systematically, some favourite topics and results of logical research will be reviewed now (be it very sketchily) with respect to their possible relevance to the study of scientific theories in general. This will serve as a first measure of the required angle of logical re-orientation.

To be sure, there exist some important logical topics already in the metascience of empirical theories, like causal explanation or analyticity. The point is that we want more.

To begin with, classical research into the foundations of mathematics has produced mainly results in the following area.

3.1 *Properties of Theories*

In the syntactic tradition, there is a multitude of notions with a logical pedigree. Let us just follow the traditional text book list: special *forms of axioms*, *derivability*, *definability*, then the familiar triad *consistency*, *completeness*, *independence*, and, finally, *decidability* (or complexity of the theorem set). Many of the well-known standard results of logic are in this area: Craig's Theorem, Gödel's Theorems, Church's Theorem. Some of the original problems inspiring these would seem to become less urgent in the case of empirical theories. E.g., consistency will be a minor concern – the much-praised 'physical intuition' being there to keep us straight.[29] Similarly, one might want to 'neutralize' Gödel's results, so to speak, searching rather for 'partial completeness' for the non-mathematical parts of our theories.

Questions in the structural tradition are often mathematical ones in their own right, having become respectable in empirical science even before the adolescence of modern logic. One example of a pertinent result which might be called 'logical' in spirit is the theorem in Zeeman 1964 to the effect the that the only mappings between physical reference frames preserving causal connectibility are those of the Lorentz group.

Some relevant model-theoretic topics were mentioned already in section 2.2.3. A more systematic list follows the semantic duals of the above topics.[30] Forms of axioms (more generally: complexity of definition), are connected with the *preservation theorems* so dear to logicians. This topic of the relation between structural closure conditions and explicit definability seems of wide interest. In this connection, key results like Keisler's characterization of elementary classes, conjuring up first-order definitions out of structural data in terms of closure under isomorphisms and ultraproducts, should have great inspirational value. Next, *completeness theorems* linking syntactic derivability with structural consequence retain an obvious interest. (A nice example is the finding in Giles 1979 that the logic of a certain 'physical game theory' is precisely Łukasiewicz's system $Ł_\infty$.) Padoa's Method and Beth's Theorem tying up explicit definability with structural *determination* have already been applied repeatedly in the theory of empirical science (cf. Suppes 1960, Montague 1974, Rantala 1977). Of the above-mentioned triad, consistency retreats into the back-ground – the possession of *models* for theories being presupposed model-theoretically. Moreover, independence has fallen out of favour with the logicians themselves already, being of purely academic interest. Completeness, however, is connected with the important structural phenomenon of *categoricity*, which may have wider importance.

What becomes clear from this quick look is not that all these logical results remain interesting *as they stand* – only that this type of result remains of potential interest *if* it can be formulated and proven in the more complex setting of empirical theories as developed in section 2.1. Here is one example, arbitrarily chosen. Take Tennenbaum's beautiful theorem stating that the standard natural numbers are the only countable model of Peano Arithmetic with *recursive* operations of addition and multiplication. Here is a more 'empirical' one in the same spirit. The only countable models for linearly ordered Time which are *homogeneous* (in some natural sense) are the rationals, the integers and their lexicographic product (in that order). (Cf. van Benthem 1980.) What about a result in this spirit for mechanics?

Interest in the edifice of science, or the succession of scientific theories ('scientific progress', if one wishes) will soon lead into the following less settled area.

3.2 *Relations between Theories*

Some syntactic relations between theories have been discussed in section 2.1 already, notably *extension* and *conservative extension*. Important relevant results are, e.g., the theorem of Craig and Vaught 1958 stating that each recursively axiomatized theory (in a language with identity) is finitely axiomatizable using additional predicates; or Gödel's non-conservation results. Such questions may well be transferred to the empirical sciences. E.g., is classical mechanics a conservative extension of classical analysis? (If not, this would be an interesting example of 'feed-back' from application to applied theory.)

These notions of extension may be regarded as special cases of the relations of *interpretability* and *embeddability* (respectively) whose formal definitions are as follows. T_1 (in L_1) is *interpretable in* T_2 (in L_2) if there exists some L_2-formula U with one free variable (the "universe") as well as some effective translation τ from the non-logical constants in L_1 to fitting (possibly complex) L_2-expressions, such that $T_2 \vdash (\tau(\alpha))^U$ for each axiom α of T_1. (Here $\tau(\alpha)$ is the result of replacing each non-logical constant in α by its τ-counterpart; while the superscript "U" indicates subsequent relativisation of all quantifiers in $\tau(\alpha)$ to U.) For technical reasons, one also requires that $T_2 \vdash \exists x U(x)$. It then follows that $(\tau(\varphi))^U$ becomes derivable in T_2 for each theorem φ of T_1. In case this implication can also be reversed, i.e., only T_1-theorems are provable (through τ, U) in T_2, T_1 has been *embedded* in T_2. These arose already in the early history of Euclidean versus non-Euclidean geometries. Notice, for example, that Klein's 'circle interior' model establishes an interpretation from hyperbolic geometry into a conservative extension of Euclidean geometry (obtained by adding an individual constant for a circle). As an example of a relevant logical result, here is a kind of 'compactness theorem for inner models' (cf. van Benthem 1980a): "If T_1 (in L_1), T_2 (in L_2) are theories such that each *finite* part of T_1 is interpretable in T_2 (constants in $L_1 \cap L_2$ being translated identically), then the whole theory T_1 is interpretable in some conservative extension of T_2".[31] In general, there is a scarcity of results concerning these notions. Lately, however, logicians have been studying questions like the above for special cases. E.g., for extensions of Peano Arithmetic, one has the stronger 'Orey Compactness Theorem' (cf. Lindström 1979).

Speaking structurally, empirical science is replete with 'reduction phenomena' calling for logical analysis: thermodynamics versus statistical mechanics, classical mechanics versus relativistic or quantum mechanics, etc. Very often no uniform satisfactory logical notions have been forged yet to get a good grip on these. (Cf. Pearce 1981, however.)

To illustrate this, here is a small example concerning Time. Suppose one has always worked with *dense* time – on the pattern of the rationals $\langle \mathbb{Q}, < \rangle$, say. Now it becomes clear that time is *discrete* at some deeper level, on the pattern on the integers $\langle \mathbb{Z}, < \rangle$. The resulting structure could be the lexicographic product $\mathbb{Q} \times \mathbb{Z}$ consisting of a dense unbounded linear sequence of copies of the integers. Notice that the theory of this structure is that of *discrete* time! Should one say that the original $\mathbb{Q}$ is reducible to the new 'richer' $\mathbb{Q} \times \mathbb{Z}$ – and, if so, in which sense? There are various relations to choose from. E.g., $\mathbb{Q}$ is clearly *isomorphically embeddable* into $\mathbb{Q} \times \mathbb{Z}$. Conversely, the obvious contraction mapping is a $\leqslant$-homomorphism from $\mathbb{Q} \times \mathbb{Z}$ onto $\mathbb{Q}$, satisfying, in its backward direction, the 'p-morphism' clause of modal logic (cf. Segerberg 1971). (Notice that the contraction is *not* a $<$-homomorphism.) The correct view of the situation would seem to be the following, however: $\mathbb{Q}$ is reducible to a certain 'level' of $\mathbb{Q} \times \mathbb{Z}$, in the sense of being *isomorphic to a quotient of* the latter structure under some suitable equivalence relation. (Notice that the above equivalence was *not* a congruence in the model-theoretic sense.) This notion of reduction seems of potential wide application.[32]

Third comes the model-theoretic perspective upon the previous two sets of notions. Duality relations between syntactic 'conservative extension' and structural 'Ramsey extension' were already treated in section 2.1. As for interpretability, an interesting duality result will be found in Pearce 1980B, which also contains many useful references to the technical literature. (Even so, powerful criteria are still lacking to *disprove* interpretability or embeddability in given cases.) One partial result is the following easy consequence of the Compactness Theorem: 'A finitely axiomatized theory T_1 is interpretable in T_2 *up to disjunction*[33] if and only if each model for T_2 contains a substructure with L_2-definable domain and predicates which is a model for T_1.' Thus, structural 'local' definability is equivalent to 'global' definability up to disjunction. A simple syntactic trick enables one to get rid of the latter phrase: any 'disjunctive' interpretation may be replaced by a 'single' one – and so we have a second duality result for interpretability. Unfortunately, its 'structural' clause is of a $\forall\exists$-form which does not make for easy counter-examples (unlike in the case of Gödel's Completeness Theorem or Beth's Definability Theorem). Cf. the appendix (section 5) for more related results.

Even in this unsatisfactory state, there is an interesting phenomenon to be observed here, namely a 'reversal of direction'. This is a common event in semantics: the experienced reader will have pondered already about cases like Sneed's account of how to 'apply' a theory, where *content* and *applicability* threaten to become inversely proportional. In the present case, the relevant observation is that syntactic 'reducibility' from T_1 to T_2 amounts to structural 'reducibility' from models of T_2 to models of T_1.[34] Thus, e.g., structural relations like the nonembeddability of orthomodular Hilbert space lattices into Boolean algebras (cf. Kochen and Specker 1967) do *not* correspond smoothly to non-interpretability of the quantum-mechanical axiomatic theory into that of classical mechanics! Similarly, the structural formulation in Suppes 1960 of the thesis that biology may be reduced to physics' in the following terms: 'for any model of a biological theory it [is] possible to construct an isomorphic model within physical theory' is very hard to fathom in syntactic terms.

Given these uncertainties at such an elementary level, we will refrain from formulating the far more involved (and numerous) notions of 'reducibility' arising from the richer concept of 'theory' developed in section 2. Nevertheless, picking out the fruitful ones from among the many Sneedian proposals (cf. Balzer and Sneed 1977/8) will be one of the logicians' first tasks.

To conclude, let us return once more to the above temporal illustration. The structural reduction relation between $\mathbb{Q}$ and $\mathbb{Q} \times \mathbb{Z}$ was that of division by the equivalence relation of 'being only finitely many discrete steps apart'. What about a corresponding syntactic formulation? At first sight, this seems hopeless, the two syntactical theories being *inconsistent*: density versus discreteness. So, a scientific revolution? That would be a spectacular, but evidently rather shallow conclusion. A better description is the following. From the theory of unbounded dense linear Time (T_1, language L_1) one passes on to a *two-sorted* theory T_2 retaining T_1 at one level and having the theory of unbounded discrete linear Time at the other – together with some obvious 'bridge principles' between the two sorts. Further model-theoretic questions might concern the matter just when the 'upper sort' is explicitly definable in terms of the 'lower sort'. (In the present example, it is clearly too aristocratic for that.)

4. PHILOSOPHICAL AFTERMATH

This paper has been extremely general, not only in its background considerations (section 1), but also in its technical development (sections 2, 3; as well as the appendix to follow). Nevertheless, production of more logical generalities is not advocated here – to the contrary. (The present author has taken a vow that this paper has been his last sin against his own precepts.) Recall the observation made in section 1.2: logical research should be guided by specific results one wants to obtain. Now, it has often been noted that both the internal dynamic of a theoretical subject and the demands from its applications may set such goals. The latter source of inspiration deserves the main emphasis at the present juncture: witness also Suppes' editorial exhortation in Suppes 1973 to carry the banners from mathematics to the logical study of Space, Time and Kinematics. A route of march is not the same as a Cause, certainly; but, it is all we have to propose by way of a general Program. The remainder of this section will be devoted to some feelings about the scope of such an undertaking.

4.1 *What Is 'Application'*?

In section 1.1 it was argued, as against proponents of the a priori inadequacy of logic, that the only way to find out how far logic will take us in the study of science is to go and try. But, in order to avoid unnecessary disillusions, some prior thought about what is *meant* by 'application' of logic seems appropriate. Quite generally, its frequent use in scientific talk – and its usefulness at the academic money tap – should not hide the fact that 'application' is a very diffuse term, in urgent need of clarification. For the latter purpose, indeed, a separate paper (or book) would be required: here only a few relevant points will be made.

'Applying logic' in the study of a certain scientific theory means using logical tools: a neat, but un-informative statement. For, these 'tools' may be anything from methods and theorems to mere notions or notations. Sometimes, the only 'tool' would even be that esoteric (though real) quality called 'logical sophistication' – everything else being supplied by hard work. This situation is not different in principle from that of any formal discipline. Maybe some engineers are able to 'apply' mathematics to problems as a kind of magic wand – but most applications require the creation of new mathematics on the spot.

The preceding paragraph should not be misunderstood to imply that some vague 'logical point of view' is all that matters. For instance, if one merely used logical *notions* in the study of science, then what would be the difference with, say, the information-theoretic sauce ('sender', 'receiver', 'channel') which has poisoned so many infant sciences in the cradle by substituting labels for analyses? (Cf. the danger of a renascent Aristoteleanism referred to in note 6.) It was precisely the contention of sections 1.1 and 1.2 that application of logic in the philosophy of science should not stop here: notions should be tied up with 'regulative' theorems (already existing, or created especially for the purpose). And – an esoteric logical sophistication able to produce these requires a very down-to-earth training in technical logic ....

Thus, the question arises what it means to apply a logical *theorem*.

Very often, these will serve as no more than 'methodical traffic lights'. E.g., recursion theory tells us that certain grammars are too weak to generate given languages, whereas others are worth trying. (Cf. the famous ascending hierarchy in Chomsky 1957.) Similarly, Beth's Theorem warned us that a certain formulation of constraints in section 3.1 paved the way to Reductionism. Notice the familiar Popperian point exemplified here: failure is the only definitive form of success.[35] Positive results may be less informative, because the reason for their success remains unclear. (Cf. note 27 about the success of 'classical' calculi of deduction.)

It may be of interest to observe how critics of modern logic fail to acknowledge this point. E.g., the influential polemic Perelman 1976 establishes the fundamental inadequacy of formal logic in the study of juridical reasoning by means of examples like the following. In the French legal tradition after the Code Napoléon judges were required to give their verdicts on the presuppositions that the Law contained neither 'conflicts' nor 'gaps', while a decision could be reached in a simple methodical objective way. (By the way, the honourable motive behind this strictness was to minimize juridical arbitrariness; thus protecting the citizens. Nevertheless, this ideal turned out untenable in practice: there remained an irreducible component of *interpretation* on the part of judges.) Now, as for the methodological background of this ideal, Perelman remarks (correctly) that it amounts to requiring the Law to be a consistent, complete and decidable theory: a rare species even in mathematics, and so . . . one more example of the inadequacy of logic. In the light of the preceding paragraph, however, an *application* of logic has been given, demonstrating the inherent limitations of certain conceptions of Law; thus creating room for more sophisticated methodical conceptions, or supplementary considerations of *content* (rather than mere *form*).

Another reason why positive results may be of doubtful applicability is their (necessarily) general nature. E.g., having a general recursion theory requires some distance from practice – which results in paradigms of 'constructivity' like primitive recursive functions easily becoming physically non-computable. Many logical existence theorems are proven by means of quite un-realistic 'constructions', as was noted already in connection with Craig's Theorem (cf. note 15). Thus, logical *methods* (as embodied in such proofs) may be largely inspirational.

In the present author's experience, however, an optimistic search for manageable instances will usually be a successful strategy in concrete cases. Reality may be harsh, but she is not mean.

4.2 *In Praise of Formalism*

Nobody devoid of 'logical sensitivity' will be converted by apologies for formal methods. Thus, one *should* never make them to unsympathetic audiences. (Unless one has to do so in order to make a living, of course.) As one *need* not make them to sympathetic audiences, a rather silent conclusion seems to follow . . . .

Still, to whomever it may concern, a few virtues of 'the formal approach' will be extolled here, which are praised too seldom in the prevailing cultural climate (whether by opponents or defenders). I do not take 'formal' or 'exact' philosophy to be a road to instant rationality and rock-bottom insights. It is, if anything, a means for re-fitting our raft (precariously afloat on Neurath's ocean), gauging the depth of our intellectual ignorance underneath. Accordingly, the choice for this type of work (instead of all-embracing philosophies) has always seemed to me a matter of intellectual *honesty* rather than *poverty*.

Honesty is a virtue, of course, but not a particularly exciting one. Let me, therefore, be more explicit about the merits of being formal and precise – or, better, about making things a little *more* formal and precise. For one thing, it forces us to become clearer about cherished intuitions – teaching us the invaluable art of being wrong. (Saying that the world is an 'Organic Whole' means risking, and learning, nothing: saying that it is a Finite State Machine is a heroic and instructive mistake.) Moreover, as we unravel our concepts, their real wealth is unveiled, and hitherto unexplored possibilities are opened up. Hence formal precision is a stimulus to creative phantasy – as has been stressed in Piaget 1973. The fashionable opposition of 'creative freedom' and 'logical armour' does not do justice to logic (nor, one fears, to creativity).

Here is also where a task of logic emerges which I take to be of vital importance to the philosophy of science. It should provide a 'conceptual laboratory' where ideas are tested out under ideal circumstances. Very importantly too, there should be a 'conceptual sanctuary' (or 'mental asylum'?), where old discarded scientific ideas are kept alive[36] – like strains in a pollen bank for grain, which may be needed again at some future time. Hopefully, in this way, logic could help a great deal in closing the self-imposed gap between the worlds of common sense and science from which our culture is suffering.

5. TECHNICAL APPENDIX

The main text of this paper is rather sketchy, detailed technical arguments having been suppressed. In order to establish at least some logical credentials, here are some samples of more technical work around our general theme.

I. *Invariance From a Model-Theoretic Perspective*

A typical model-theoretic line of research might be the following. One starts studying a certain domain, say the discrete two-dimensional lattice $\mathbb{Z} \times \mathbb{Z}$; endowed with a certain structure, say the ternary relation 'further (from) than'. First comes a structural question. Which bijections of the lattice are *automorphisms* with respect to this structure? (The answer would be, in this case, precisely the group of translations, rotations and reflections.) These automorphisms give rise to *invariants*: n-ary relations on the domain which are mapped onto themselves by all automorphisms in the group. (Cf. previous remarks about Klein's Program.) These invariants form a very interesting class, being closed under operations like *complement*, *intersection* or *projection*. Could, then, all invariants be characterized linguistically using some suitable language? Clearly, at least, the original relation is an invariant, together with all relations which are predicate-logically definable from it. (Cf. the earlier quotation from Weyl 1949.) Does the converse hold too? This is a difficult question, calling for 'internal' definability results (relativized to some mother structure), rather than 'external' ones like Keisler's characterization of elementary classes. Here come a few pertinent results. (Cf. also Rantala 1977a.)

PROPOSITION 1. In *finite* structures, invariance for automorphisms implies first-order definability.

Proof: Let $\mathcal{D}$ be some finite structure, and $A \subseteq D^n$ some invariant. A direct combinatorial argument will produce a first-order definition. Here we just observe, however, that this also follows from proposition 3 to be proven below, *if* we can only show that $(\mathscr{D}, A)$ is $(L+A)$-saturated. Thus, let $\vec{x}$ be any sequence of, say, s variables, and $\Sigma = \Sigma(\vec{x})$ some set of $(L+A)$-formulas involving finitely many parameters in D which is finitely satisfiable in $\mathscr{D}$. Saturation requires that Σ be simultaneously satisfiable. Suppose it were not. Then, for each s-tuple $\vec{d}$ in D^s, there exists some formula $\sigma_{\vec{d}}$ in Σ such that $\sigma_{\vec{d}}$ is false at $\vec{d}$. But, there being only finitely many such s-tuples, it would follow that some *finite* subset of Σ is non-satisfiable in $\mathscr{D}$: contradicting the original assumption on Σ. Q.E.D.

In *infinite* structures, proposition 1 may fail, however. *Counter-example*: In the structure $\mathbb{N} \oplus \mathbb{Z}$ consisting of the natural numbers followed by a copy of the integers, in the usual ordering, $\mathbb{N}$ is invariant for automorphisms, without being first-order definable in terms of the ordering relation.

Often, model-theoretic results arise only when single structures are replaced by their *theories*. Thus, let T be some first-order theory in a language $L+A$, where A is some n-ary relation symbol. Indeed, a duality between *invariance* and *definability* is now forthcoming:

PROPOSITION 2. $A^{\mathscr{D}}$ is an L-automorphic invariant in each model for T if and only if A is explicitly definable in T up to disjunction.

Proof: From right to left, clearly, if

$$T \vdash \forall \vec{x}(A\vec{x} \leftrightarrow \delta_1((\vec{x})) \vee \ldots \vee \forall \vec{x}(A\vec{x} \leftrightarrow \delta_m(\vec{x}))$$

where $\delta_1, \ldots, \delta_m$ are L-formulas – then, in each model for T, A will be first-order defined by some δ_i. Hence, it will be invariant in the above sense.

Secondly, for the converse, *Svenonius' Theorem* may be applied (Chang and Keisler 1973; theorem 5.3.3). Let $\mathscr{D}$ be an L-structure with $(L+A)$-isomorphic expansions $(\mathscr{D}, A)$, $(\mathscr{D}, A')$ to models for T. By the invariance of A, then $A = A'$ ($(L+A)$-automorphisms being L-automorphisms). By the above-mentioned theorem this implies explicit definability up to disjunction. Q.E.D.

In order to apply such changed results to 'internal' problems like the original geometrical one, it would have to be shown that, e.g., any given invariant A in $\mathbb{Z} \times \mathbb{Z}$ gives rise to a theory $T = Th((\mathbb{Z} \times \mathbb{Z}, A))$ in each of whose models the interpretation of A is an invariant. We do not solve this problem here.

Finally, there does exist a kind of structure inside which invariance implies explicit definability without further ado – an observation which will be used below:

PROPOSITION 3. If A is an L-automorphic invariant in an $(L+A)$-saturated structure $\mathcal{D}$, then it is first-order L-definable in $\mathcal{D}$.

Proof: Let this A be an n-ary relation. Consider any n-ary sequence $\vec{a}$ of objects in A. Now, the L-type of $\vec{a}$ in $\mathcal{D}$ cannot be satisfied outside of A. Otherwise, L-homogeneity (a consequence of the given saturation property) would provide an L-automorphism of $\mathcal{D}$ *not* mapping A onto itself. Put differently, the $(L+A)$-set consisting of the L-type of $\vec{a}$ together with the formula $\neg A\vec{x}$ is not satisfiable in $\mathcal{D}$. Hence some *finite* subset of it will not be satisfiable – $\mathcal{D}$ being $(L+A)$-saturated. Say

$$\{\tau_1(\vec{x}), \ldots, \tau_s(\vec{x}), \neg A\vec{x}\}$$

is not satisfiable in $\mathcal{D}$; where $\tau_1, \ldots, \tau_s$ are true of $\vec{a}$. Equivalently,

(1) $\forall\vec{x}(\tau_{\vec{a}}(\vec{x}) \rightarrow A\vec{x})$ is true in $\mathcal{D}$;

where $\tau_{\vec{a}} =_{\text{def}} \tau_1 \,\&\, \ldots \,\&\, \tau_s$.

Moreover, since each $\vec{a}$ in A will satisfy some such L-formula $\tau_{\vec{a}}$, the $(L+A)$-set

$$\{\neg\tau_{\vec{a}} \mid \vec{a} \in A\} \cup \{A\vec{x}\}$$

is not satisfiable in $\mathcal{D}$ either. Hence, again by saturation, there exist finitely many sequences $\vec{a}_1, \ldots, \vec{a}_t$ in A such that

(2) $\forall\vec{x}(A\vec{x} \rightarrow \tau(\vec{x}))$ is true in $\mathcal{D}$;

where $\tau =_{\text{def}} \tau_{\vec{a}_1} \vee \ldots \vee \tau_{\vec{a}_t}$.

In combination with (1), it then follows that

$$\forall\vec{x}(\tau(\vec{x}) \leftrightarrow A\vec{x}) \text{ is true in } \mathcal{D}:$$

A has been L-defined. Q.E.D.

II. *The Categorial Mode of Thinking*

The categorial perspective mentioned in connection with Sneed's constraints may be used to provide a characterization of first-order definability inside a structure. To see this, let $\mathfrak{D}$ be some L-structure, A some n-ary invariant in it. Now, $\mathfrak{D}$ belongs to the category CAT (L) of all L-structures, with L-isomorphic embeddings (not necessarily surjective) for its morphisms, and the ultraproduct construction as a characteristic operation. (Unfortunately, I do not know of any categorial *definition* of ultraproducts in terms of their 'morphological' behaviour.)

PROPOSITION 4. A is first-order L-definable in $\mathfrak{D}$ if and only if CAT(L) is simultaneously expandable to a sub-category of CAT ($L + A$) containing ($\mathfrak{D}$, A) such that all morphisms and operations of CAT(L) remain the same in CAT($L + A$).

Proof: If A is first-order L-definable, say $A = \varphi^{\mathfrak{D}}$, then the obvious expansion of each L-structure $\mathfrak{D}'$ to $(\mathfrak{D}', \varphi^{\mathfrak{D}'})$ will do. For, L-isomorphisms remain $(L + A)$-isomorphisms by their fundamental preservation property, and Łoś' Theorem guarantees the same for ultraproducts.

Conversely, suppose that CAT(L) has been expanded as indicated. Let U be any countably incomplete α-good ultrafilter over some index set I; where $|I| = \max(\aleph_0, |D|)$ and $\alpha = |I|^+$. Consider the expanded ultrapower $\Pi_U \mathfrak{D}$ in CAT($L + A$) – which equals $\Pi_U(\mathfrak{D}, A)$, by assumption. Notice that this ultraproduct is $(L + A)$-saturated; by Chang and Keisler 1973, theorem 6.1.8. Moreover, $A^{\Pi_U \mathfrak{D}}$ is invariant in the ultraproduct: L-automorphisms of this structure being automatically $(L + A)$-automorphisms, by assumption. The promised conclusion then follows at once from proposition 3. Q.E.D.

Hopefully, this line of thought will yield more results. (This speculation in the first draft of this paper was verified in Pearce 1980B.)

III. *Varieties of Reduction*

The notion of 'reduction' has been shown to admit of various logical explications in section 3.2. First, a *syntactic* result of that section will be proven here.

Let T_1 (T_2) be a first-order theory in language L_1 (L_2). Translations τ from L_1 to L_2 will assign, possibly complex, L_2-predicates to L_1-primitives – those in $L_1 \cap L_2$ being mapped identically. (Thus, one implements the idea that a partial correspondence is given in advance.) 'Interpretation' of T_1 in T_2 means that some translation $\tau(T_1)$ may be proven in T_2, possibly relativized to a unary L_2-predicate denoting some sub-domain. Since T_1 may be infinitely axiomatized, a 'compactness theorem' would be helpful here. But, in fact, as was stated in note 31, no such result holds in general. (Cf. Lindström 1979, however.) What we do have is the following:

PROPOSITION 5. If each finite subset of T_1 is interpretable in T_2, then the whole theory T_1 is interpretable in some *conservative extension* of T_2.

Proof: Take a new unary predicate constant U not occurring in either L_1 or L_2. Now, consider $T_2 \cup T_1^U$ (where all quantifiers occurring in T_1 are relativized to U). T_1 is trivially interpretable in this new theory. Thus, it only matters to show that the latter is a conservative extension of T_2. To see this, let φ be any L_2-sentence such that $T_2 \cup T_1^U \vdash \varphi$. It follows that, for finitely many $\alpha_1, \ldots, \alpha_k \in T_1$, $T_2 \cup \{\alpha_1^U, \ldots, \alpha_k^U\} \vdash \varphi$. Equivalently, taking α to be the conjunction of these formulas, $T_2 \vdash \alpha^U \to \varphi$. Then, since neither U nor the $(L_1 - L_2)$-vocabulary of α occur in T_2, it follows for the universal second-order closure $\forall(\alpha^U \to \varphi)$ – taken with respect to U and $L_1 - L_2$ – that $T_2 \vdash \forall(\alpha^U \to \varphi)$.

Next, by the assumption on T_1 and T_2, there exists some translation τ of L_1 into L_2, as well as some unary L_2-predicate β such that $T_2 \vdash (\tau(\alpha))^\beta$. From the previous paragraph, it then follows that $T_2 \vdash (\tau(\alpha)^\beta \to \varphi$, and hence $T_2 \vdash \varphi$. Q.E.D.

In section 3.2 the following *model-theoretic* interpretation result was stated, for finitely axiomatized theories T_1:

PROPOSITION 6. T_1 is interpretable in T_2 up to disjunction if and only if each model for T_2 contains an L_2-definable sub-domain together with L_2-definable predicates which is a model for T_1.

Proof: From left to right, this is obvious. Conversely, if T_1 is not interpretable as indicated, then T_2 united with the set of all negations of L_2-definable translations of T_1 (relativized to all possible L_2-definable domains) is a finitely satisfiable set of formulas. By compactness, then, it will be simultaneously satisfiable. In other words, T_2 possesses a model without any definable submodel for T_1. Q.E.D.

Next, the 'syntactic trick' mentioned in section 3.2 works as exemplified by the following special case. Suppose that

$$T_2 \vdash (\tau_1(T_1)^{U_1} \wedge \exists x U_1) \vee (T_2(T_1)^{U_2} \wedge \exists x U_2).$$

Call the first disjunct α, and set

$$U(x) =_{\mathrm{def}} (U_1(x) \wedge \alpha) \vee (U_2(x) \wedge \neg\alpha).$$

Notice that $T_2 \vdash \exists x U(x)$. Then, for arbitrary L_1-predicates P, set

$$\tau(P)(\vec{x}) =_{\mathrm{def}} (\tau_1(P)(\vec{x}) \wedge \alpha) \vee (\tau_2(P)(\vec{x}) \wedge \neg\alpha).$$

A simple argument about models of T_2 establishes that $T_2 \vdash \tau(T_1)^U$.

Further refinements, using 'cross-structural' categorial conditions as in II above, yield a duality result for interpretability 'tout court'. Instead of embarking upon this course, however, we prefer to review the whole topic from a different angle, namely that of *combination* of theories.

One has got two first-order theories T_1 (in L_1) and T_2 (in L_2), possibly sharing some vocabulary $L_1 \cap L_2$. Let us assume that both are intended to describe the *same* kind of objects. The weakest form of combination would seem to be:

(1) the deductive union of T_1 and T_2 is consistent in $L_1 + L_2$.

Robinson's Joint Consistency Theorem tells us that this occurs just when $T_1 \upharpoonright L_1 \cap L_2$ and $T_2 \upharpoonright L_1 \cap L_2$ contain no mutually contradictory theorems. Notice that $T_1 \cup T_2$ need not be a conservative extension of either T_1 or T_2: both theories may have learnt in the process. Some reflection upon Robinson's proof shows that, e.g., $T_1 \cup T_2$ will be a conservative extension of T_2 if and only if $T_1 \upharpoonright L_1 \cap L_2$ is contained in $T_2 \upharpoonright L_1 \cap L_2$.

(For, the direction from left to right is obvious. Conversely, for any L_2-sentence φ, if $T_2 \nvdash \varphi$, then $T_2 \cup \{\neg\varphi\}$ has got a model $\mathfrak{D}$. By the assumption, then, the $(L_1 \cap L_2)$-theory of $\mathfrak{D}$ united with T_1 will be finitely satisfiable – and hence it has a model $\mathfrak{D}'$. Starting from $\mathfrak{D}$ and $\mathfrak{D}'$, one builds alternating elementary chains, whose unions may be folded together into an $(L_1 + L_2)$-structure verifying $T_1 \cup T_2$ while falsifying φ. Cf. Chang and Keisler 1973.)

Next, stronger connections between T_1 and T_2 may arise upon addition, such as:

(2) $T_1 \cup T_2$ is contained in some *definitional extension* of T_2.

I.e., given certain L_2-definitions of $(L_1 - L_2)$-vocabulary, T_1 becomes derivable from T_2. This is the earlier definition of 'interpretation' – leaving out the complication of possible relativization to sub-domains, that is. In this case, automatically $T_1 \cup T_2$ is a conservative extension of T_2 (cf. the proof of proposition 5 above), though not necessarily of T_1. The latter need not even happen on the stronger connection that 'definitional reducibility' occurs:

(3) $T_1 \cup T_2$ coincides with some definitional extension of T_2.

Contrary to our experience with the mis-match between Conservative Extension and Ramsey Extension, the latter connection between theories admits of an elegant structural characterization:

PROPOSITION 7. T_1 is definitionally reducible to T_2 if and only if each model for T_2 admits of exactly one expansion to a model for T_1.

Proof: First, suppose that T_1 is definitionally reducible to T_2 – say, $T_1 \cup T_2$ is axiomatizable by T_2 plus L_2-definitions δ as above. Now, consider any model $\mathfrak{D}$ for T_2. By interpreting the $(L_1 - L_2)$-vocabulary through δ, $\mathfrak{D}$ is expanded to an $(L_1 + L_2)$-structure $\mathfrak{D}^+$ which satisfies (T_2 plus δ and hence) $T_1 \cup T_2$. A fortiori, then, $\mathfrak{D}^+$ is a model for T_1: and the 'at least' has been taken care of. Next, as for 'at most'. If $\mathfrak{D}^+$, $\mathfrak{D}^{+\prime}$ are any two expansions of $\mathfrak{D}$ to models for T_1, then ($T_1 \cup T_2$, and hence) δ will hold in both. Therefore, the respective interpretations of the $(L_1 - L_2)$-vocabulary must coincide: i.e., $\mathfrak{D}^+ = \mathfrak{D}^{+\prime}$.

Conversely, suppose that the above structural condition holds. Its 'at most' side means that the implicit definability clause of Beth's Theorem is satisfied – and hence explicit L_2-definitions δ of the $(L_1 - L_2)$-vocabulary are derivable in $T_1 \cup T_2$. It follows that $T_1 \cup T_2$ may be (re-)axiomatized as the union of its L_2-part $(T_1 \cup T_2) \upharpoonright L_2$ with δ. For, clearly, both of these are derivable from $T_1 \cup T_2$. Moreover, vice versa, if $T_1 \cup T_2 \vdash \varphi(L_1, L_2)$, then $T_1 \cup T_2 \vdash \varphi(\delta(L_1), L_2)$ – where '$\delta(L_1)$' refers to suitable L_2-replacements through δ-, whence the L_2-formula $\varphi(\delta(L_1), L_2)$ belongs to $(T_1 \cup T_2) \upharpoonright L_2$. In conjunction with δ, then, the latter theory reproduces the original formula $\varphi(L_1, L_2)$.

Thus, in order to prove definitional extension of T_1 to T_2, it now suffices to show that $(T_1 \cup T_2) \upharpoonright L_2$ coincides with T_2. Put differently, it remains to be shown that $T_1 \cup T_2$ is a conservative extension of T_2. But this follows from the 'at least' side of the above structural condition. For, if φ is any L_2-formula such that $T_2 \nvdash \varphi$, then $T_2 \cup \{\neg\varphi\}$

will have a model: which can be expanded to some model for $T_1 \cup T_2 \cup \{\neg\varphi\}$ – whence $T_1 \cup T_2 \nvdash \varphi$. Q.E.D.

These were pleasing notions – but, rather un-realistic. For, in actual practice, the theories to be added will usually refer to *different* kinds of objects. Consider, e.g., Elementary Geometry (cf. Tarski 1959) with primitives (ternary) 'betweenness' and (quaternary) 'equidistance' for T_1, and the algebraic theory of the reals ($\mathbb{R} = \langle R, 0, 1, +, \bullet\rangle$) for T_2. As was noted already in section 3.2, a *two-sorted* combination is more appropriate – possible *sortal connections* becoming the main issue. E.g., in this case, equidistance induces an equivalence relation between pairs of points allowing for a bridge notion of 'length' satisfying certain bridge principles for 'addition' (etc.). The relation between such more complex syntactic notions of reduction and the usual representation theorems will not be investigated here.

Rijksuniversiteit Groningen

NOTES

* I would like to thank David Pearce and Veikko Rantala for their helpful comments.

[1] An exception to this rule is V. Rantala's paper 'Correspondence and Non-Standard Models: A Case Study', in *Acta Philosophica Fennica* **30** (1979), 366–378. As for 'ordinary' applications, there is, of course, a variety of publications by Robinson himself and co-workers.
[2] Regrettably, defeatist (or elitist) logicians have used similar excuses to justify their exclusive concentration upon *pure mathematics*.
[3] What is usually criticized is some static elementary text book version of logic. E.g., in the otherwise very interesting work *Psychology of Reasoning* (Batsford, London, 1972), P.C. Wason and P.N. Johnson-Laird give the following example to show that 'actual reasoning transcends any formal modelling.' The two sentences 'if prices increase, the firm goes bankrupt' and 'prices increase only if the firm goes bankrupt' are stated to be 'logically equivalent', having the same logical *form* '$I \rightarrow B$'. But, in practice, people feel a difference, due to additional temporal or causal *content*, and so . . . . The authors seem unaware of the fact that observations like these are precisely at the basis of modern logical semantics. E.g., in tense logic, the two logical forms would be '$I \rightarrow$ *future* B' and '$I \rightarrow$ *past* B', respectively.
[4] And that without expecting red carpets. There seems to be a growing tendency among such philosophers of science to regard logic as just one auxiliary discipline among many: set theory, topology, game theory, system theory, etc.
[5] This is Popper's main insight – applied to the philosophy of science itself.

[6] In this connection, one sometimes fears that a New Aristoteleanism is in the air, confusing *labeling* phenomena with *understanding* them.

[7] It almost goes without saying that her interest will also cover the most diverse *purposes* of reasoning: rationalization, justification, extraction of information, refutation, explanation. In particular, the latter two form an obvious link with Hempelian or Popperian philosophy of science.

[8] Notice how, on this view, there is a continuous spectrum from the logical study of (natural) language to that of scientific theories.

[9] Yet more complex entities have been proposed, like 'theory nets' (Sneed) or 'research programs' (Lakatos). But, even then, the level of theories remains a convenient starting point.

[10] Speaking historically, this view may have been inspired by Kant's doctrine that only knowledge which is related to (possible) experience is free from contradiction, whereas 'pure reason' runs the constant risk of antinomies.

[11] Or should one argue to the opposite effect: non-constructive logic is harmless for constructively given domains: one has to be more careful once *non*-constructive domains are considered (the intuitionist position)? As so often, metaphors point either way.

[12] There might be a problem here. If T_1 is axiomatized by means of universal statements, and one is only allowed to use universal statements in proofs, could not it be that certain universal theorems become unprovable: because their predicate-logical proof requires intermediate steps of higher quantifier complexity? Fortunately, the answer is negative: as may be seen, for instance, by using *semantic tableaus.*

[13] An interesting border-line case is, e.g., Goldbach's Conjecture to the effect that each even natural number greater than two is the sum of two primes. We do not know any explicit functions f, g yielding only primes such that each even number n equals $f(n)+g(n)$. But, there are (trivially) *recursive* functions f, g such that, if n can be written as a sum of two primes at all, then $f(n)+g(n)$ will be such a sum. Does this make Goldbach's Conjecture an acceptable 'concrete' statement? Or should one require *primitive* recursive Skolem functions?

[14] A relevant example (A.S. Troelstra) is the following. In the real numbers, it is constructively provable that the equation $y^3-3y+x=0$ has solutions for each value of the parameter x. But, there is no continuous function f: $\mathrm{IR} \to \mathrm{IR}$ such that

$$\forall x(fx)^3 - 3(fx) + x = 0.$$

(A 'cusp catastrophe' occurs.)

[15] By the way, here is an instructive counter-example to the claim that philosophers can get by with a knowledge of just logical *results.* For, as it happens, Craig's method of *proof* produces a 'tricky' set of axioms for the observational sub-theory, without any practical use. Thus, although it is good to know that a 'reduction' exists in principle, a lot of work remains to be done to produce convincing and interesting *examples.*

[16] This contribution by no means exhausts Suppes' important work in the philosophy of science. To mention just one other example, his papers on *measurement theory* have a clear interest for the semantics of 'observational' terms.

[17] The answer is negative, maybe surprisingly. Cf. Przełęcki 1969.

[18] Notice that, in this case, $T \vdash \varphi$ if and only if $T_0 \vdash \varphi$, for each L_0-sentence φ (T_0 being

complete). Hence, the example also refutes the conjecture that the *conjunction* of (1) and (3) would be a syntactic counterpart to the structural condition that MOD(T_0) *equals* MOD(T) $\upharpoonright L_0$.

[19] Originally, it was hoped that, at least, it would be a *decidable* matter, given a theory T, if its theoretical vocabulary is Ramsey eliminable. This conjecture was refuted, however, in van Benthem 1978.

[20] As for C2, it would have to be decomposed into a *law*

$$\forall xy\, m(x \cup y) = m(x) + m(y)$$

(with '∪' denoting physical join) plus the identity constraint C1.

[21] For an example of this type of requirement from a quite different area, cf. H. Ehrig, H.-J. Kreowski and P. Padawitz: 'Stepwise Specification and Implementation of Abstract Data Types', in *Proceedings ICALP* 5 (1978), Springer Lecture Notes in Computer Science 62, 205-226. Here the morphisms are algebraic *homomorphisms*.

[22] Let p occur in both D_1 and D_2. Now, D_1 and D_2 will have a common empirical *superstructure* D containing p. Since this relation must be preserved at the theoretical level, p will receive the same theoretical function values in the expansions of D, D_1 and D_2.

[23] One would also like to see the Kripke structures of modern *tense logic* applied in the semantics of scientific theories. (The 'empirical structures' of recent work by R. Wójcicki are of this kind.)

[24] There are parallel discussions concerning the status of natural *laws*: linguistic expressions or structural 'sequencing relations'? Cf. Suppe 1976.

[25] E.g., the 'subset condition' on Sneed's set-theoretic constraints is intelligible (and even correct) *only if* the constraint has ('had'?) a *universal* explicit formulation. 'Universal-existential' constraints like '*any* two particles occur together in *at least one* empirical situation' do not satisfy the condition.

Another case in point is the example in Sneed 1971 of a *balance* in equilibrium with finitely many objects on it. Some structuralists will take the predicate-logical undefinability of finiteness (and hence of this simple type of empirical situation) to be the ultimate verdict upon first-order formalisation. But, what such an analysis would do for us is force one to think about the cash value of 'finite' in this particular case. E.g., a simple first-order formalisation of the above example would involve a *recursion* on the number of objects, corresponding, in a natural fashion, to progressive addition of weights on the balance.

Finally, in Pearce 1980B it is argued that the incorporation of linguistic aspects into the Sneedian frame-work would make it 'still more expressive and comprehensive'.

[26] It need not even commit one to classical logic on the Tarski semantics. Cf. Vaught 1960, where the class of formulas classically true in all *countable* structures with *decidable* predicates is shown to be non-recursively axiomatizable.

[27] E.g., the preferred role of classical logic may be due largely to its excessive strength blasting a way from premises to conclusions – making people ignore the possibility that the same theory might have been organized using weaker logical means, allowing for subtler distinctions to be made. One traditional example is intuitionist Heyting Arithmetic, which enables one to derive the Law of Excluded Middle for atomic statements as a (*mathematical*) *theorem*. A recent exciting example is the proof in M. Dunn:

'Quantum Mathematics' (Department of Philosophy, Bloomington, 1980) that Peano Arithmetic axiomatized with *quantum logic* produces all additional classical logical laws as mathematical theorems.

[28] One would like to see at least one text book of model theory organized around this theme.

[29] Although . . . would you board a rocket to Mars, immediately after the publication of an inconsistency in the theory of differential equations?

[30] The organisational principle of Chang and Keisler 1973, viz. that of various methods of model construction, seems less suitable from the present perspective (cf. note 28). Still, awareness of such 'proliferation principles' gives a feeling for the ontological abundance of the structural universe.

[31] The latter addition is necessary: without it, counter-examples may be given. E.g., take for T_2 the first-order theory of $\langle N, O, S\rangle$ and for T_1: T_2 plus $\{c \neq s^i 0, i = 0, 1, 2, \ldots\}$. (The required calculation is not as trivial as it seems at first sight.)

[32] To get better acquainted with these various notions, notice that $\mathbb{Z}$ is isomorphically embeddable in $\mathbb{Q}$, as well as being isomorphic to one of its quotients (contract all 'pins' $(n, n+1]$ for integer n). Both converses fail. As for the model-theoretic notion of reducibility to be introduced below, $\mathbb{Z}$ is not isomorphic to any *definable* substructure of $\mathbb{Q}$- and the converse fails as well.

[33] I.e., there exist L_2-formulas $U_1, \ldots, U_k$ as well as translations $\tau_1, \ldots, \tau_k$ of the non-logical constants in L_1 into L_2 such that $T_2 \vdash (\tau_1(T_1)^{U_1} \wedge \exists x U_1) \vee \ldots \vee (\tau_k(T_1)^{U_k} \wedge \exists x U_k)$.

[34] One might try to restore symmetry by formulating dualities concerning *embeddability* rather than interpretability. E.g., 'for each L_1-sentence φ, $T_1 \vdash \varphi$ *if and* only if $T_2 \vdash (\tau(\varphi))^U$, is equivalent to: [some *symmetric* condition on MOD(T_1), MOD(T_2), τ, U]'.

[35] As the famous Dutch Calvinist prime minister Colijn once remarked to his political opponents: 'In negation lies our strength'.

[36] A prime example is the Leibnizian 'relational' view of Space and Time; cf. Suppes 1973.

REFERENCES

Arbib, M. and Padulo, L.: 1974, *System Theory*, Saunders, Philadelphia.

Ashby, W. R.: 1976, *Introduction to Cybernetics*, Methuen, London.

Balzer, W. and Sneed, J. D.: 1977/8, Generalized Net Structures of Empirical Theories', *Studia Logica* **36**, 195–211, and **37**, 167–177.

Barwise, J. (ed.): 1977, *Handbook of Mathematical Logic*, North-Holland, Amsterdam.

Benthem, J. F. A. K. van: 1978, 'Ramsey Eliminability', *Studia Logica* **37**, 321–336.

Benthem, J. F. A. K. van: 1980, The Logic of Time', Filosofisch Instituut, Rijksuniversiteit, Groningen.

Benthem, J. F. A. K. van: 1980a, Modeltheorie voor Wetenschapsfilosofen', Filosofisch Instituut, Rijksuniversiteit, Groningen.

Benthem Jutting, L. S. van: 1979, 'Checking Landau's 'Grundlagen' in the AUTOMATH System', Mathematical Centre Tracts **83**, Amsterdam.

Beth, E. W.: 1949, 'Towards an Up-to-date Philosophy of the Natural Sciences', *Methodos* **1**, 178–185.

Beth, E. W. and Piaget, J.: 1966, *Mathematical Epistemology and Psychology*, D. Reidel, Dordrecht.
Chang, C. C. and Keisler, H. J.: 1973, *Model Theory*, North-Holland, Amsterdam.
Craig, W. and Vaught, R. L.: 1958, 'Finite Axiomatizability using Additional Predicates', *Journal of Symbolic Logic* 23, 289–308.
Chomsky, N.: 1957, *Syntactic Structures*, Mouton, Den Haag.
Fine, K.: 1975, 'Vagueness, Truth and Logic', *Synthese* **30**, 265–300.
Giedymin, J. (ed): 1978, K. Ajdukiewicz – *The Scientific World-Perspective and Other Essays*', Reidel, Dordrecht.
Giles, R.: 1979, 'Formal Languages and Physics', in Hooker (ed.) 1979, pp. 19–87.
Henkin, L., Suppes, P. and Tarski, A. (eds.): 1959, '*The Axiomatic Method: with Special Reference to Geometry and Physics*, North-Holland, Amsterdam.
Hilbert, D.: 1899, *Grundlagen der Geometrie*, Teubner, Leipzig.
Hintikka, J. and Remes, U.: 1974, *The Method of Analysis, Its Philosophical Significance and Geometrical Origins*, D. Reidel, Dordrecht.
Hooker, C. A. (ed.): 1979, *Physical Theory as Logico-Operational Structure*, D. Reidel, Dordrecht.
Kochen, S. and Specker, E.: 1967, 'The Problem of Hidden Variables in Quantum Mechanics', *Journal of Mathematics and Mechanics* 17, 59–67.
Lakatos, I. M.: 1976, *Proofs and Refutations*, Cambridge University Press, Cambridge.
Lindström, P.: 1979, 'Some Results on Interpretability', Department of Philosophy, University of Göteborg.
Lorenzen, P. and Lorenz, K.: 1978, *Dialogische Logik*, Wissenschaftliche Buchgesellschaft, Darmstadt.
Montague, R.: 1974, *Formal Philosophy*, Yale University Press, New Haven.
Nagel, E.: 1961, *The Structure of Science*, Harcourt, Brace & World, New York.
Pearce, D.: 1980, 'Logical Properties of the Structuralist Concept of Reduction' *Erkenntnis* (forthcoming) (referred to as 1980A).
Pearce, D.: 1980, 'A Mathematical Characterization of Interpretation, Reduction and Definitional Equivalence', to appear in *Studia Logica* (referred to as 1980B).
Pearce, D.: 1981, 'Some Relations between Empirical Systems', *Epistemologia* **4**, 363–380.
Perelman, C.: 1976, *Logique Juridique*, Jurisprudence Générale Dalloz, Paris.
Piaget, J.: 1973, *Principles of Genetic Epistemology*, Routledge and Kegan Paul, London.
Przełęcki, M.: 1969, *The Logic of Empirical Theories*, Routledge and Kegan Paul, London.
Przełęcki, M.: 1974, 'A Set-theoretic versus a Model-theoretic Approach to the Logical Structure of Physical Theories', *Studia Logica* 33, 91–105.
Przełęcki, M., Szaniowski, K. and Wójcicki, R. (eds.): 1976, *Formal Methods in the Methodology of Empirical Sciences*, D. Reidel, Dordrecht.
Quine, W. V. O.: 1951, 'Two Dogmas of Empiricism', in Quine 1963, pp 20–46.
Quine, W. V. O.: 1963, *From a Logical Point of View*, Harper and Row, New York.
Ramsey, F. P.: 1978, *Foundations: Essays in Philosophy, Logic, Mathematics and Economics*, Routledge and Kegan Paul, London.
Rantala, V.: 1977, 'Definability in Empirical Sciences', *Acta Philosophica Fennica* **29**, 213–222.

Rantala, V.: 1977, *Aspects of Definability*, North-Holland, Amsterdam (referred to as 1977a).
Rantala, V.: 1980, 'On the Logical Basis of the Structuralist Philosophy of Science', *Erkenntnis* **15**, 33–53.
Reichenbach, H.: 1956, *The Philosophy of Space and Time*, Dover, New York.
Saarinen, E. (ed.): 1979, *Game-theoretical Semantics*, D. Reidel, Dordrecht.
Segerberg, K.: 1971, *An Essay in Classical Modal Logic*, Filosofiska Studier **13**, Uppsala.
Smorynski, C.: 1977, 'The Incompleteness Theorems', in Barwise 1977, pp. 821–865.
Smullyan, R. M.: 1977, *The Tao Is Silent*, Harper and Row, New York.
Sneed, J. D.: 1971, *The Logical Structure of Mathematical Physics*, D. Reidel, Dordrecht (2nd ed. rev. 1979).
Stegmüller, W.: 1979, *The Structuralist View of Theories*, Springer, Berlin.
Suppe, F.: 1976, 'Theoretical Laws', in Przełęcki, Szaniowski, and Wójcicki (eds.) 1976, pp 247–267.
Suppes, P.: 1957, *Introduction to Logic*, Van Nostrand, New York.
Suppes, P.: 1960, 'A Comparison of the Meaning and Uses of Models in Mathematics and the Empirical Sciences', *Synthese* **12**, 287–301.
Suppes, P. (ed.): 1973, *Space, Time and Geometry*, D. Reidel, Dordrecht.
Tarski, A.: 1959, 'What is Elementary Geometry?', in Henkin, Suppes and Tarski (eds.) 1959, pp. 16–29.
Toulmin, S. E.: 1958, *The Uses of Argument*, Cambridge University Press, Cambridge.
Vaught, R. L.: 1960, 'Sentences True in All Constructive Models', *Journal of Symbolic Logic* **25**, 39–53.
Wessel, H. (ed.): 1977, *Logik und empirische Wissenschaften*, Akademie Verlag, Berlin (DDR).
Weyl, H.: 1963, *Philosophy of Mathematics and Natural Science*, Atheneum, New York.
Zeeman, E. C.: 1964, 'Causality Implies the Lorentz Group', *Journal of Mathematical Physics* **5**, 490–493.