Generalized Logical Consequence: Making Room for Induction in the Logic of Science

Samir Chopra (schopra@cse.unsw.edu.au)
Knowledge Systems Group
School of Computer Science and Engineering
University of New South Wales
Sydney, NSW 2052, Australia

Eric Martin (emartin@cse.unsw.edu.au)
School of Computer Science and Engineering
University of New South Wales
Sydney, NSW 2052, Australia

Abstract. We present a framework that provides a logic for science by generalizing the notion of logical (Tarskian) consequence. This framework will introduce hierarchies of logical consequences, the first level of each of which is identified with deduction. We argue for identification of the second level of the hierarchies with inductive inference. The notion of induction presented here has some resonance with Popper's notion of scientific discovery by refutation. Our framework rests on the assumption of a restricted class of structures in contrast to the permissibility of classical first-order logic. We make a distinction between deductive and inductive inference via the notions of compactness and weak compactness. Connections with the arithmetical hierarchy and formal learning theory are explored. For the latter, we argue against the identification of inductive inference with the notion of learnable in the limit. Several results highlighting desirable properties of these hierarchies of generalized logical consequence are also presented.

Keywords: Induction, deduction, logical consequence, learning theory

1. Introduction

Here is an *idealized* notion of science. Hypotheses are proposed, experiments designed and conducted, and observations taken. These either confirm or refute the hypotheses. Upon confirmation, the scientist turns to the consequences of the hypotheses and investigates them through further hypothesis formation and experimentation. Upon refutation of proposed hypotheses, the scientist devises alternate hypotheses and tests those. On the basis of successful experiments, scientists *deduce* or *induce* laws and other consequences of evidential statements. These serve as the foundation of theoretical superstructures as they guide further scientific discovery. And so the cycle continues, heading inexorably towards truth and theories that are always closer to it and explain more than the previous ones. In this picture of scientific practice, scientists work with confirmed hypotheses—often elevated to the status of laws—tentative hypotheses, and a vast vocabulary that refers to observables and theoretical terms. For the scientist, a 'scientific statement' will either record an observation, represent a hypothesis, or be a member of a possible theory.

This paper presents a framework for *generalized logical consequence* which formalizes the *logical* aspects of the picture presented above i.e., that part of science that consists of deducing and inducing scientific statements. We introduce *hierarchies* of logical consequences, one level of which will be deductive and another that we will take as corresponding to inductive inference; there will be others, but these two are the most significant. Even if the latter identification is disagreed with, the general notion of consequence presented will be of independent philosophical interest in divorcing logical consequence from deductive consequence. The levels of these hierarchies correspond to the difficulty of inference of consequences contained in them: the higher a sentence in the hierarchy, the more difficult it is to infer it. The picture of logical consequence here is Tarskian i.e., a sentence follows from a set of premises if all models of the premises are models of the conclusion. We do not take a position on the debate (Etchemendy, 1999) on whether such a notion of consequence is the correct one. The crucial formal move in our framework is the adoption of a class of possible worlds or structures that are in general more restricted than those considered admissible by first-order logic. This is accompanied by the selection of a restricted class of possible theories—as opposed to first-order logic's unrestricted version of a theory being any set of sentences. We then relativize the notion of consequence to our chosen class of structures and use the set of possible theories to define levels of the hierarchies, showing how inferences made at one level can serve as premises for subsequent levels. The demarcation of the class of possible worlds and theories correspond, in our opinion, to the decision made by a scientific investigator to mark out a domain of investigation by a notion of the physically possible. The layers of inference correspond to the iterated nature of the scientific enterprise.

The outline of the paper is as follows. We first provide some background to the problem of induction and its history. We then move on to a discussion of why this history does not provide us with a 'logic' of induction as we understand it. In Section 2, after specifying the desiderata for a logic of science, we discuss the reasons for the inadequacy of first-order logic for our purpose. We propose a solution via a restriction of the class of admissible structures and provide some basic definitions. In Section 3 we then show how the hierarchies of logical consequences can be constructed and investigate their properties. In Section 4 we explore connections with the arithmetical hierarchy and learning theory and conclude with some pointers to future work.

1.1. THE PROBLEM OF INDUCTION

Our task of providing a logic of science brings us face to face with the problem of induction, since in general, the confirmation of a scientific hypothesis is an inductive conclusion from evidence statements. The problem of induction is the problem of justifying the formation of universal statements from particular ones—with incomplete data at hand. The universal statements in science are the hypotheses and laws that make up theories; the particular statements are observations or reports on experiments. The rich and complex history of the problem of induction and inductive inference is taken up with attempts to either justify induction (grue puzzles, the correct form of inductive rules of inference (Goodman, 1954)) or to provide an adequate analysis of its use in scientific practice (e.g., the Carnap-Popper debate on the correct logic of scientific discovery). The difficulty of this problem leads to many answers, ranging from positive attempts like Carnap's inductive logic (Carnap, 1950) to pragmatic approaches such as the *vindication* approach (Reichenbach, 1949; Salmon, 1968)—that the use of an inductive rule of inference is justified by its success in achieving some desired means.

As might be expected, there is no uniform understanding of the logic of induction or its role in science. One contrast in conceptions of induction is between induction by elimination and induction by enumeration (von Wright, 1957). In the former, a statement follows as inductive conclusion from a body of evidence when all or 'enough' alternative conclusions have been eliminated. In the latter, a statement follows as inductive conclusion from a body of evidence when 'enough' cases in its favour have been enumerated ("all emeralds observed till now are green; therefore all emeralds including unobserved ones are green"; here 'enough' is interpreted as 'all observations till the current point in time'). Another contrast arises in understanding induction via probabilistic considerations: an inductive conclusion can be accepted on the basis of evidence that confirms it to an acceptably high degree, where this notion of 'acceptably high-degree' is linked with notions of rational belief. Or, conclusions can be accepted as more or less probable according to evidence—essentially, the degree of confidence in the putative conclusion is interpreted as some sort of probability measure. Numerous arguments for this approach have been provided, the most prominent being a form of Bayesian conditionalization (Howson and Urbach, 1989; Earman, 1992). In the probabilistic approach to inductive inference there is considerable disagreement over the nature of the probabilistic relation between evidence and conclusion—for a modern treatment of this issue, see (Pitt, 1989).

The historical Carnap-Popper debate is plausibly viewed as one about the role of induction in science. Carnap held that the proper 'logic of science' was inductive, and more particularly, the logic which provides a logical foundation for probability—inductive logic was viewed as the theory of the concept of probability or confirmation. (See (Kyburg, 1970) for similar views.) Since probability was taken to be objective, its explicatum would refer only to the sentences (evidence and hypothesis) and their logical properties in a given formal system. Carnap then, attempted to analyze probability or the degree of confirmation, as a logical relation between the evidence and the hypothesis and regarded it as the central concept in inductive logic: the so-called c-functions $c(h, e) = r$ are read as the degree of confirmation of hypothesis h by evidence e to degree r. In Carnap's inductive logic, where the degree of confirmation of every universal law is always zero, 'inductive' is understood as 'non-deductive' and its logic as quantitative.

Popper (Popper, 1959), in contrast, held that there was no logic of induction (i.e., Hume's problem was insuperable) and that science progressed by cycles of conjecture formation and attempts at refutation—the basis of the notion of corroboration. The notion of refutability, that the ever increasing corroboration of a highly falsifiable hypothesis follows from the unsuccessful attempts at falsifying it, aimed to establish the potential deductive refutation of a hypothesis by a single failed experiment as the hallmark of scientific discovery. A scientific hypothesis was never actually confirmed—a universal generalization cannot be confirmed with certainty but it can be refuted. A single black swan, as legend goes, refutes the hypothesis that all swans are white. However, such an analysis was also unlikely to explain scientific progress without resort to induction since the truth of the so-called reliable observations depended on prior theorizing. It is clear then, that some ampliative inference occurs in the formation of scientific laws and corroboration of hypotheses in the cycles described by Popper. Indeed, the problem of answering why among unfalsified hypotheses, one is to select the most corroborated one brings us back to the methodological problem of justification.

1.2. A 'LOGIC' OF INDUCTION?

It should be clear that the approaches to the logic of induction described above are, strictly speaking, not logics. Instead, the term 'logic' is used to refer to a language for talking about scientific practice—what are scientists doing when they carry out induction and can this process be rationally reconstructed? The 'logic' appears at the meta-level. For instance, Carnap's inductive logic generalizes truth values to a real-valued degree of confirmation, which indicates the extent to which the inductive conclusion is confirmed by the premises. Proofs in such a logic consist of calculating the degree of confirmation pertaining to pairs of premises and conclusions. This is not a terribly new concern: part of the debate in the early history of induction was whether it was even appropriate to use the word 'inference' to describe induction, in light of the criticism that we make here:

> He [Salmon] is a little afraid logicians will say: but no inference is any good unless what is inferred follows *logically* [italics added] from the data.[1]

In this study, we concern ourselves with the problem of providing a purely formal, qualitative system akin to classical logic, in which deductive consequence can be separated from non-deductive consequence. We argue that portions of this non-deductive part can be identified as inductive. (Note the implicit understanding in the quote above that logical consequence just is deductive consequence.) We will not argue that the logic of induction is to be cast precisely in terms of 'logics' as logicians understand them. Rather, as just indicated, our approach will be to devise a system in which a sound argument can be made for identifying a notion of consequence within it as approximating our notion of inductive inference. The similarity of this brand of inductive inference with previous conceptions of induction will be a point of interest. Of equal interest will be the hierarchies of logical consequences that emerge, corresponding to the iterative nature of the scientific enterprise.

Our road to this identification takes us through formal learning theory where an alternative picture of inductive inference is provided. Learning theory originated in a numerical setting (Gold, 1967); see (Jain et al., 1999) for an overview of the extensive work carried out on variations of the numerical paradigm. An offshoot of this work was an extension to logical settings by, among others, (Glymour, 1985; Kelly, 1996; Martin and Osherson, 1998). Our study bears links with the latter. In these frameworks, the learning theorist identifies successful inductive inference as a *limit case of learning*. Given a sentence φ whose truth or untruth in the underlying world W is unknown prior to any investigation, and given an infinite stream e of data representing the evidence gathered from the environment (passively by observations or actively by conducting experiments), a learner will react to each finite initial segment of e by either issuing φ or not. Assume as an idealization that the evidence gathered characterizes W uniquely. Success is defined in terms of stabilization of the sequence of hypotheses to φ if and only if φ is correct in W. What makes this task non-trivial is that the learner has to succeed for any W taken from a class $\mathcal{W}$ representing the class of all possible worlds. This guarantees the existence of a convergence point for worlds in which φ is true, but the learner is usually unaware of whether it has reached this point. Note that in worlds in which φ is not true, we can impose the additional restriction that the learner must converge to $\neg\varphi$. Such a criterion for learning is merely one amongst several alternatives.

What does this picture of induction afford us? Note that φ is the inductive conclusion to be drawn from *any* set of evidence sentences for worlds in which φ is true. This paradigm, then, identifies or analyzes inductive infer-

[1] (Hacking, 1968)

ence by defining it as learning in the limit. Such a conception of induction however, is dissatisfying. Convergence in the limit should be implied by the definition of inductive inference—itself independently grounded in logical, purely model-theoretic notions. Defining induction as formal learning theory does, renders it a purely operational notion.

2. Generalized logical consequence and a unified logic of science

2.1. The desiderata for a logic of science

What happens if we cast scientific discovery and investigation in a logical framework? Both deduction and induction—corresponding to the activity of the scientist—should appear as kinds of inference in such a framework. The notion of an experiment leading to repeatable results which approach the truth should be taken seriously and quantitative notions such as probabilities disregarded. The logical system that emerges should provide a unified framework for both deduction and induction. Classical first-order logic, while clearly inadequate for this task, should hold its ground as a particular case. It equates logical consequence with deductive consequence; a situation that arises because of the compactness theorem and because of first-order logic's well known failure to restrict the class of admissible first-order interpretations. An adequate notion of generalized logical consequence will need to address both these concerns and it is to this issue that we now turn.

Consider a situation in which every object in the world is denoted by a closed term of the formal language used to describe and model the world. In physical contexts, this assumption is reasonable if every object in the world has a definite location in some coordinate system, and if a location can be determined through measurements[2]. Even in the mathematical universe, this assumption is often justified. For instance, a number theorist is mainly interested in one structure, namely the set $\mathbb{N}$ of natural numbers with its usual operations, and first-order logic provides formalizations in which every natural number is interpreted by a closed term. One such is based on a vocabulary that enables us to represent the numerals i.e., for each $n \in \mathbb{N}$ a closed term $\overline{n}$, and an interpretation such that every natural number n interprets $\overline{n}$. We then select a set $T_{\mathbb{N}}$ of formulas (axioms) true of the natural numbers, together with the definition of the models of $T_{\mathbb{N}}$. The structure $\mathbb{N}$ is just one of these models; the other structures are *non-standard* models.

Non-standard models, or *unintended interpretations*, are not recognized by some number theorists as part of their object of investigation (other mathematicians do study non-standard models of arithmetic and hence are inter-

[2] Though some interpretations of quantum mechanics tell us that this is not possible, some, such as David Bohm's theory, do make this possible.

ested in the broader class of interpretations (Kaye, 1991)). So from the classical number theorist's point of view, the completeness theorem for first-order logic is not of great importance. What is the point of having a completeness theorem for a class of structures—the class of models of $T_{\mathbb{N}}$—which is not regarded as an object of investigation? This remark is clearly not applicable to all first-order formalizations. A group theorist studies precisely the class of structures defined as the class of models of the (first-order) axioms for groups. In that case, the completeness theorem is of value. And contrary to the completeness theorem, the incompleteness theorem is of direct relevance to the classical number theorist, because it says something about $\mathbb{N}$ itself and not about a class of structures which contains $\mathbb{N}$.

In first-order logic it is impossible to restrict the class of possible worlds to those structures all of whose individuals interpret a closed term, even with an infinite set of formulas. It is impossible to express that the domain of the structure(s) under investigation consists exactly of the numbers $0, 1, 2 \ldots$ (up to isomorphism of course) and to rule out the existence of non-standard integers. First-order logic forces us to accept the unintended interpretations. When we have a definite interpretation or class of interpretations in mind, for instance an interpretation in which every individual interprets a closed term, should we also accept the enlargement of the associated class of possible worlds to a more inclusive class of structures? If we do, we should be rewarded for this sacrifice.

To discover what the reward could be, consider how a mathematician who accepts a first-order formalization of his field of interest which makes it a subfield of the class of all possible interpretations, is rewarded, or supposed to be. The usual answer is: by the completeness theorem, by the downward Löwenheim-Skolem theorem, and by the compactness theorem (the third is a consequence of the first one). But the value of the completeness theorem lies in the acceptance of the extended class of structures as the new object of investigation, so this answer actually begs the question. The downward Löwenheim-Skolem theorem is implicit in any approach which equates intended interpretations with some class of countable structures: if every possible world is such that all individuals interpret a closed term, then every possible world is of cardinality the maximum of $\aleph_0$ and the cardinality of the vocabulary (hence $\aleph_0$ in every natural choice of the latter).

So, only the compactness theorem can be rewarding. Its content is that if a formula φ is a logical consequence of theory T, then φ is a logical consequence of a *finite subset* of T. This property guarantees that φ can be discovered *with certainty* from T. Indeed, since at any given time, only some finite amount of information (some finite part of T) is available, there would be no way to discover φ with certainty from T if φ were not a logical consequence of some finite subset of T. The compactness theorem says that the truth of the conclusion can always be discovered with certainty: *logical* *consequence* and *deductive consequence* are *equivalent notions in the first-order framework*. First-order logic is all about deduction. This is a property that mathematicians can praise, since their activity consists exclusively of deducing statements from other statements. But scientists whose activity involves both some form of deduction and some form of induction will not be so enthusiastic, and note that first-order logic certainly has something to offer to the former, but less to the latter. First-order logic might serve as a foundation for mathematics, but a 'logic' for science, even at the purely formal, idealized level is still not with us.

There are numerous so-called 'inductive logics', of course (see (Hintikka and Suppes, 1966; Jeffrey, 1980) for discussions). These are logics in which we are not concerned with just deduction, not necessarily logics only for induction. But they differ considerably from first-order logic in their treatment of both consequence and proof as Carnap's inductive logic shows. We set ourselves the following goal:

> A first-order logic of deduction and induction which generalizes the notion $T \models \varphi$ (φ is a logical consequence of theory T), so that classical first-order logic becomes a particular case of the former.

Classical first-order logic would correspond to the case where the class of intended interpretations is the class of all structures. How would the issue of compactness be tackled in such a logic?

2.2. Languages and Possible worlds

Consider a vocabulary (a set of function and predicate symbols) S. We refer to S-structures as *structures*, and to closed S-terms, as *closed terms*. (See (Shoenfield, 1967; Ebbinghaus et al., 1994) for clarification of basic notions from model theory.) The natural class of languages on the basis of which the concepts of our framework can be expressed turns out to be the *countable fragments of* $\mathcal{L}^S_{\omega_1\omega}$. Remember that $\mathcal{L}^S_{\omega_1\omega}$ is the extension of $\mathcal{L}^S_{\omega\omega}$, the set of first-order formulas, that accepts countable disjunctions and conjunctions (see for instance (Ebbinghaus et al., 1994)). A countable fragment of $\mathcal{L}^S_{\omega_1\omega}$ is a countable subset of $\mathcal{L}^S_{\omega_1\omega}$ closed under subformulas (see (Barwise, 1975) for a complete account). So $\mathcal{L}^S_{\omega\omega}$ is the smallest countable fragment of $\mathcal{L}^S_{\omega_1\omega}$. As an illustration, suppose that S consists of a unary predicate P, a constant $\overline{0}$, and a unary function symbol s (successor). The concept 'P holds for all even numbers' can be expressed by a finite sentence namely, $P(\overline{0}) \wedge \forall x(P(x) \rightarrow P(s(s(x))))$. On the other hand, if S does not contain s, but rather a constant $\overline{n}$ for every natural number n, then the same concept cannot be expressed by a finite formula, but it can be expressed by the infinite formula $\bigwedge_{n\in 2\mathbb{N}} P(\overline{n})$. Such a formula can be included in a countable fragment of $\mathcal{L}^S_{\omega_1\omega}$. We can actually just consider $\mathcal{L}^S_{\omega\omega} \cup \{\bigwedge_{n\in 2\mathbb{N}} P(\overline{n})\}$, which *is* a countable fragment of $\mathcal{L}^S_{\omega_1\omega}$, since it is obviously countable and closed under subformulas. From now on, we denote by $\mathcal{L}$ an arbitrary fragment of $\mathcal{L}^S_{\omega_1\omega}$. We refer to the closed members of $\mathcal{L}$ as *sentences*.

The framework we develop rests on the assumption that we are free to consider an *arbitrary class* of structures, instead of being committed to the class of *all* structures, as is the case for elementary developments of classical logic. So, we denote by $\mathcal{W}$ a non-empty class of structures and given $T \subseteq \mathcal{L}$, we denote by $\mathrm{Mod}_{\mathcal{W}}(T)$ the class of all members of $\mathcal{W}$ that are models of T. The class of structures under consideration, $\mathcal{W}$, is called *the class of possible worlds* to emphasize its intended meaning. In doing so, we parametrize the logic by a class of structures and classical logic emerges as a particular case—when $\mathcal{L} = \mathcal{L}^S_{\omega\omega}$ and the class of structures is taken to be the class of all set-theoretic structures. In this framework *Henkin* structures will play an important role (but other choices are always possible). These are defined as those structures all of whose individuals interpret a closed term. When we consider Henkin structures, we tacitly assume that S contains at least one constant. Such structures are used in the Henkin proof of the completeness theorem for first-order logic which says that if a theory has a model, then it has a Henkin model in a language extended with constants bearing witness to existential statements. Henkin structures, when the underlying language is devoid of the equality symbol, are referred to as *Herbrand* structures; these are instrumental in developing a theory of automated theorem proving in artificial intelligence.

Our suggestion that we stick to so-called intended interpretations is not without philosophical or scientific precedent (see for example discussions such as (Putnam, 1983; Klenk, 1976; Shapiro, 1991; Parson, 1990) centered around issues arising from the Löwenheim-Skolem theorem). We do not take an explicit stance on what these intended interpretations could be, but will often work with a class especially useful for formal developments of some parts of our framework. As noted above, a desirable class of structures is used in the fundamental development of first-order logic and artificial intelligence. Similarly, a scientist commencing investigation makes the tacit assumption of a class of physically possible situations which will form the subject of further investigation. In these situations, certain 'non-standard' scenarios such as the physical laws of the world being radically different are ruled out. These non-standard scenarios might not be expressible in first-order logic and hence their exclusion is effected by the choice of structures.

With $\mathcal{W}$ as the class of intended interpretations, the notion of logical consequence which corresponds to '$T \models \varphi$' is defined as follows.

Definition 1. *Let $T \subseteq \mathcal{L}$ and $\varphi \in \mathcal{L}$ be given. We say that φ is a* logical consequence of T in $\mathcal{W}$*, and we write $T \models_{\mathcal{W}} \varphi$, iff every member of* $\mathrm{Mod}_{\mathcal{W}}(T)$ *is a model of φ. Given a set T of sentences, we denote by* $\mathrm{Cn}_{\mathcal{W}}(T)$ *the set of all sentences φ such that $T \models_{\mathcal{W}} \varphi$.*

As an example, assume that $\mathcal{W}$ consists exclusively of Henkin structures, that S contains a unary predicate symbol Phys, and consider the theory $T_P = \{\mathrm{Phys}(t) \mid t \text{ is a closed term}\}$, which says that our language is only capable of denoting physical objects. By this choice of $\mathcal{W}$ the converse is also true, that every physical object is denoted by a term in our language. We do not accept non-standard structures because we know, or assume, that such a 'non-standard' object does not exist. From a physical point of view this is a reasonable assumption to make in some circumstances. Note that the restriction that every physical object is denoted by a closed term is not first-order definable and also justifies the choice of Henkin structures as the class of intended interpretations. With our indicated choice of $\mathcal{W}$, $\forall x \mathrm{Phys}(x)$ is a logical consequence of T_P in $\mathcal{W}$, but not a compact one since $\forall x \mathrm{Phys}(x)$ is not a logical consequence of any finite subset of T_P in $\mathcal{W}$. This shows that in general, the notion of logical consequence defined above is *not* compact.

2.3. Compactness, and Weak Compactness

We interpret the compactness theorem of first-order logic as saying that logical consequence and deductive consequence are equivalent, and claim that the compactness property is the hallmark of deduction. Starting from a generalized notion of logical consequence which is no longer compact, we can then *define* the deductive consequences of T in $\mathcal{W}$ as those logical consequences of T in $\mathcal{W}$ that enjoy the compactness property i.e., the logical consequences, in $\mathcal{W}$, of some finite subset of T. Consider again the theory $T_P = \{\mathrm{Phys}(t) \mid t \text{ is a closed term}\}$. According to our definition of 'deduction,' $\forall x \mathrm{Phys}(x)$ is not a deductive consequence of T_P in $\mathcal{W}$. This is neither surprising nor disappointing. We cannot discover *with certainty* that $\forall x \mathrm{Phys}(x)$ is a logical consequence of T_P in $\mathcal{W}$; we cannot *deduce* that $\forall x \mathrm{Phys}(x)$ is true. But we can expect that $\forall x \mathrm{Phys}(x)$ is an *inductive consequence* of T_P in $\mathcal{W}$, in some reasonable sense of 'inductive.' As we will see later, some inductive conclusions can be drawn in the case when the class of structures is such that all objects in their domain can be named.

Being an inductive consequence of T in $\mathcal{W}$ should certainly not just mean being a logical consequence of T in $\mathcal{W}$, whether a deductive consequence or a non-deductive one. There are logical consequences of T in $\mathcal{W}$ which are too difficult to infer to be tagged with the qualifier 'inductive.' Inductive consequences should lie somewhere between logical (compact or non-compact) consequences in $\mathcal{W}$ and deductive (compact) consequences in $\mathcal{W}$. It is impossible to discover *with certainty* that formula φ is an inductive consequence of theory T in $\mathcal{W}$ unless φ is actually a deductive consequence of T in $\mathcal{W}$ i.e., unless the compactness property is satisfied. But in order to claim that φ is an inductive consequence of T in $\mathcal{W}$, it should be possible to at least *reasonably believe* in φ. The intuition behind 'we can reasonably believe in φ' is formally defined by a property of *weak compactness*. So the equation 'induction = weak compactness' applied to the notion 'φ is a logical consequence of T in $\mathcal{W}$' summarizes a fundamental intuition—to be clarified when our formal machinery is in place.

In case the choice of $\mathcal{W}$ is such that the notion 'φ is a logical consequence of T in $\mathcal{W}$' is compact, then the three notions 'φ is a deductive consequence of T in $\mathcal{W}$', 'φ is an inductive consequence of T in $\mathcal{W}$', and 'φ is a logical consequence of T in $\mathcal{W}$' collapse. This is the case in particular for classical first-order logic, where $\mathcal{L} = \mathcal{L}^S_{\omega\omega}$ and $\mathcal{W}$ is the class of all structures. When the notion 'φ is a logical consequence of T in $\mathcal{W}$' is not compact, then the notion 'φ is a deductive consequence of T in $\mathcal{W}$' becomes strictly weaker than the notion 'φ is a logical consequence of T in $\mathcal{W}$,' and there is room for a specific notion of induction in-between.

2.4. Paradigms

We have relativized the notion of logical consequence to $\mathcal{W}$, one of the key parameters of our framework. Similarly, we relativize the notion of consistency to $\mathcal{W}$:

Definition 2. *Let $T \subseteq \mathcal{L}$ be given. If $\mathrm{Mod}_{\mathcal{W}}(T) \neq \emptyset$ then we say that T is* $\mathcal{W}$-consistent*; otherwise we say that T is* $\mathcal{W}$-inconsistent*.*

$\mathcal{W}$-consistency is stronger than consistency *tout court*. A simple example which exemplifies the difference, with S the kind of vocabulary considered in the previous section, is easily provided. Consider the example of the previous section, where $\mathcal{W}$ is the class of all Henkin structures. Then $\{\mathrm{Phys}(t) \mid t \text{ is a closed term}\} \cup \{\exists x \neg \mathrm{Phys}(x)\}$ is consistent, but $\mathcal{W}$-inconsistent. There is an important case where both notions are equivalent:

Proposition 1. *Suppose that S contains infinitely many constants and $\mathcal{W}$ is the class of all Henkin structures. Then no subset of $\mathcal{L}$ in which infinitely many constants in S do not occur is both consistent and $\mathcal{W}$-inconsistent.*

A theory in classical logic is any set of sentences. This is quite liberal. Just as we need to be able to consider a restricted class of structures—the class of possible worlds—, we also need to be able to consider a restricted class of theories, that we call *possible theories*. One intuition is to conceive of a possible theory as potential evidence which can be gathered from a possible world. A possible theory could also contain some kind of background knowledge: it is not just an arbitrary set of formulas, but a starting point for logical investigation, for a search for the truth, and such a starting point is not free from epistemological constraints. In full generality, we should be free to, *a priori*, accept any set of sets of sentences as the starting point of the logical investigation, and classical logic, which accepts all sets of sentences, is a particular case. So the class of possible theories denoted by $\mathcal{T}$, a set of sets of sentences, is another parameter of our framework. Some results will require $\mathcal{T}$ to fulfill certain conditions. We will describe in the remainder of this section some especially useful conditions that can be imposed on $\mathcal{T}$.

We now have the basic ingredients of our framework to put together. We denote by $\mathcal{P}$ the triple $(\mathcal{L}, \mathcal{W}, \mathcal{T})$, and term it the *paradigm* under investigation, with $\mathcal{L}$, $\mathcal{W}$ and $\mathcal{T}$ based on a given vocabulary S. The choice of this terminology is deliberate—the selection of a language, a class of possible worlds and a set of possible theories reflects the methodological presuppositions of the scientist and the choice of a particular investigative framework. The most general paradigm sets $\mathcal{L}$ equal to $\mathcal{L}^S_{\omega\omega}$, $\mathcal{W}$ equal to the class of all structures, and $\mathcal{T}$ equal to the set of all sets of sentences i.e., $\mathcal{P}$ can be identified with classical, first-order logic. It is a limit case—and we will soon see what that means formally—as a consequence of the identity between 'logical consequence' and 'deductive consequence' in classical logic.

Two classes of paradigms play a prominent role. One of them is a subclass of the other, obtained by imposing some conditions on $\mathcal{T}$. Since $\mathcal{T}$ can be intuitively viewed as all potential evidence obtainable (via passive observation or experiments) from an underlying possible world, we can speak as well of the selection of a set $\mathcal{O}$ of sentences, called the set of *possible observations*:

Definition 3. *Given a set $\mathcal{O}$ of sentences, we say that $\mathcal{P}$ is* observational (for $\mathcal{O}$) *iff $\mathcal{T} = \{D_{\mathcal{O}}(\mathfrak{M}) \mid \mathfrak{M} \in \mathcal{W}\}$, where $D_{\mathcal{O}}(\mathfrak{M})$ denotes the* $\mathcal{O}$-diagram *of $\mathfrak{M}$ i.e., the set of all members of $\mathcal{O}$ true in $\mathfrak{M}$.*

Observational paradigms correspond to the simplified picture of investigating the properties of a world on the basis of the set of all possible observations true in it. The only background knowledge is implicit, represented by $\mathcal{W}$. The paradigms defined in formal learning theory are similar to our observational paradigms, with the notion of *language* (non-empty subset of $\mathbb{N}$) and *text for* a language L (enumeration of the members of L) corresponding to the possible theory defined from a possible world corresponding to L. In section 4.2 we examine this connection in more detail. But in formal learning theory, no text is for more than one language, whereas for some choices of $\mathcal{W}$ and $\mathcal{O}$, two distinct members of $\mathcal{W}$ could have the same $\mathcal{O}$-diagram. This leads to the following notion:

Definition 4. *Given a set $\mathcal{O}$ of sentences, we say that $\mathcal{P}$ is* perfectly observational (for $\mathcal{O}$) *iff $\mathcal{P}$ is observational for $\mathcal{O}$ and for all $T \in \mathcal{T}$, $\mathrm{Cn}_{\mathcal{W}}(T)$ is a complete theory.*

For all sets of sentences X, φ is a logical consequence of X in $\mathcal{W}$ if and only if for all $T \subseteq \mathcal{L}, X \subseteq T \Rightarrow T \models_{\mathcal{W}} \varphi$. The following result shows that in observational paradigms, the set of possible theories enjoys a certain kind of density property relative to the notion of logical consequence in $\mathcal{W}$.

That is, the relation specified above holds for a set of possible observations X, even when T is drawn from the restricted set of theories $\mathcal{T}$ instead of being an arbitrary set of sentences.

Proposition 2. *Suppose that $\mathcal{P}$ is observational for $\mathcal{O}$. Let $X \subseteq \mathcal{O}$ and $\varphi \in \mathcal{L}$ be given. The following clauses are equivalent.*

- *$X \models_{\mathcal{W}} \varphi$.*
- *For all $T \in \mathcal{T}$, if $X \subseteq T$ then $T \models_{\mathcal{W}} \varphi$.*

3. Hierarchies of Logical Consequences

3.1. Definitions

Having defined a generalized notion of logical consequence, namely the notion of *logical consequence in* $\mathcal{W}$, we now formalize the fundamental concept of our framework: *the difficulty of inferring a given sentence φ, from a possible theory T*, presuming that all models of T in $\mathcal{W}$ are models of φ. More precisely, given $T \in \mathcal{T}$, we define a *hierarchy* over T of members of $\mathrm{Cn}_{\mathcal{W}}(T)$, that we denote by $\mathcal{H}(T)$. These hierarchies have Σ_n and Π_n levels, for any $n \in \mathbb{N}$, starting from T at level 0. For every $n > 0$, the Π_n level is potentially above the Σ_n level and below the levels Σ_m for all $m > n$, but some levels can collapse. Intuitively, we can think of the hierarchy as providing a measure of the complexity of the inference of sentences: the higher the first level at which a sentence φ occurs at in $\mathcal{H}(T)$, the 'harder' the task of discovering that φ is a generalized logical consequence of T.

To build the hierarchy over T, we apply the following strategy. Given $n > 0$, suppose that the hierarchies over all possible theories have been built up to level Π_{n-1}. Let $T \in \mathcal{T}$ and $\varphi \in \mathrm{Mod}_{\mathcal{W}}(T)$ be given.

- Let finite sets of sentences E and H be such that E is included in T and all members of H occur in $\mathcal{H}(T)$ below level Σ_n. Suppose that for every possible theory T', the following holds. If $E \subseteq T'$ and $H \subseteq \mathrm{Cn}_{\mathcal{W}}(T')$, then $T' \models_{\mathcal{W}} \varphi$. In that case we can be sure, on the basis of E and H, that φ is a logical consequence of T in $\mathcal{W}$. Thus φ is placed on level Σ_n of $\mathcal{H}(T)$.
- Let finite sets of sentences E and H be such that E is included in T and all members of H occur in $\mathcal{H}(T)$ at level Σ_n or below. Suppose that for every possible theory T', the following holds. If $E \subseteq T'$, $H \subseteq \mathrm{Cn}_{\mathcal{W}}(T')$, and $T' \not\models_{\mathcal{W}} \varphi$, then $\neg\varphi$ occurs in $\mathcal{H}(T')$ at level Σ_n or below. We can then, on the basis of E and H, believe with confidence that φ is a logical consequence of T in $\mathcal{W}$, knowing that we would be refuted if we were wrong. Thus φ is placed on level Π_n of $\mathcal{H}(T)$.

The description of the Σ_n level is based on a compactness property. Finite information—*evidence* (subset of the theory) and *hypotheses* (which at the meta-level, can be assessed as correct; this fact is usually not known to the scientist)—enable the conclusion that φ is a logical consequence of T in $\mathcal{W}$. The description of the Π_n level is based on a property of *weak compactness*. Finite information—of the same kind as before—enables the conclusion that φ could not belong to the set of sentences that are not logical consequences of T in $\mathcal{W}$, without this fact being among those that could be discovered before. The compactness property enables us to conclude with certainty that φ is true; the property of weak compactness enables us to believe confidently in the truth of φ. The certainty and confident belief in question are not absolute, but relative to that part of the hierarchy below the level which is currently built. Only the sentences on level Σ_1 of $\mathcal{H}(T)$ can be discovered to be true with absolute certainty, and only the sentences on level Π_1 of $\mathcal{H}(T)$ can be believed to be true with absolute confidence, as far as T is reliable; in the case of observational paradigms, as far as T contains all possible observations true in the underlying world. The following definitions make clear how the properties of compactness and weak compactness are manifested in the Σ_n and Π_n levels.

Definition 5. *Given $n > 0$, $T \in \mathcal{T}$ and finite $E, H \subseteq \mathcal{L}$, we say that sentence φ is a* compact consequence of T (in $\mathcal{P}$) (on the basis of (E, H)) at level n *just in case $E \subseteq T$, H is a set of weakly compact consequences of T at level $n-1$, and for all $T' \in \mathcal{T}$:*

if $E \subseteq T'$ and $T' \models_{\mathcal{W}} H$ then $T' \models_{\mathcal{W}} \varphi$.

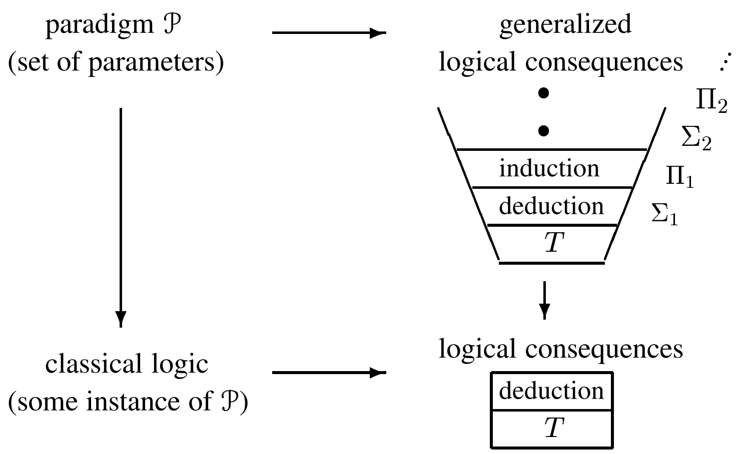
Obviously, H should be set to $\emptyset$ in the definition above when $n = 1$.

Definition 6. *Given $n > 0$, $T \in \mathcal{T}$ and finite $E, H \subseteq \mathcal{L}$, we say that sentence φ is a* weakly compact consequence of T (in $\mathcal{P}$) (on the basis of (E, H)) at level n *just in case $E \subseteq T$, H is a set of compact consequences of T at level n, $T \models_{\mathcal{W}} \varphi$, and for all $T' \in \mathcal{T}$:*

if $E \subseteq T'$, $T' \models_{\mathcal{W}} H$ and $T' \not\models_{\mathcal{W}} \varphi$, then $\neg\varphi$ is a compact consequence of T' at level n.

The definitions above let us establish the following notation. Let $T \in \mathcal{T}$ be given. Set $\Sigma_0^{\mathcal{P}}(T) = \Pi_0^{\mathcal{P}}(T) = T$. Given $n > 0$, $\Sigma_n^{\mathcal{P}}(T)$ denotes the set of all sentences that are compact consequences of T at level n, and $\Pi_n^{\mathcal{P}}(T)$ denotes the set of all sentences that are weakly compact consequences of T at level n.

The following pictorially represents the framework that emerges (note that we have identified the Σ_1 level with deduction and the Π_1 level with induction; in the sections to follow, we provide further justification for this):



In the picture above, classical first-order logic only determines hierarchies of one level. The basis of refutation of the sentences in the higher levels of the hierarchies become progressively weaker as we move up the hierarchies. If the application of the weak compactness property at a Π_n level entails the acceptance of φ while an application of the compactness property at the corresponding Σ_n levels entails the rejection of φ, then it is the compactness property that prevails. Of course, such a rejection is provisional as well, since it itself is based on prior applications of the weak compactness property at Π_m $(m < n)$ levels.

3.2. Properties of the hierarchy

We now show that the hierarchies of inferences so defined have desirable properties. Given $n > 0$ and $T \in \mathcal{T}$, members of $\Sigma_n^{\mathcal{P}}(T)$ and $\Pi_n^{\mathcal{P}}(T)$ are not at the same level in contrast to the arithmetic-like hierarchies. Despite this, there will be some strong structural similarities.

Property 1. *Let $T \in \mathcal{T}$ be given.*

- *For all $m, n \in \mathbb{N}$ with $m \leq n$, $\Sigma_m^{\mathcal{P}}(T) \subseteq \Pi_n^{\mathcal{P}}(T)$.*
- *For all $m, n \in \mathbb{N}$ with $m < n$, $\Pi_m^{\mathcal{P}}(T) \subseteq \Sigma_n^{\mathcal{P}}(T)$.*

Clearly, the notion of generalized logical consequence in $\mathcal{W}$ is *monotone*. Therefore, if possible theory T is included in possible theory T', then the formulas that occur at level Σ_n—respect. Π_n—in the hierarchy over T also occur, at the same level (and maybe below), in the hierarchy over T':

Property 2. *Let $T, T' \in \mathcal{T}$ with $T \subseteq T'$ be given. Then for all $n \in \mathbb{N}$, $\Sigma_n^{\mathcal{P}}(T) \subseteq \Sigma_n^{\mathcal{P}}(T')$ and $\Pi_n^{\mathcal{P}}(T) \subseteq \Pi_n^{\mathcal{P}}(T')$.*

As explained above, the Σ_n levels of the hierarchies are defined using a compactness property. How does this compare to the usual compactness property of first-order logic? Since the latter refers to the underlying theory T alone, the question is relevant only for the Σ_1 level of the hierarchies. These are defined on the basis of T alone as well, contrary to the Σ_n ($n > 1$) levels, which involve not just T but all consequences generated at previous levels. The result below shows that the compactness property for Σ_1 levels is defined in terms of not just T, but also in terms of all possible theories.

Property 3. *Let $T \in \mathcal{T}$ and $\varphi \in \mathcal{L}$ be given. Then $\varphi \in \Sigma_1^{\mathcal{P}}(T)$ iff there exists a finite subset E of T such that for all $T' \in \mathcal{T}$, $T' \models_{\mathcal{W}} \varphi$ whenever $E \subseteq T'$.*

Property 3 is of epistemological significance. A sentence is a deductive consequence of a theory T if and only if the truth of φ can be discovered from T with certainty i.e., the truth of φ can be discovered from a finite subset of T. So the Σ_1 level of the hierarchy over a possible theory T consists precisely of those consequences of T in $\mathcal{W}$ which can be tagged with the qualifier 'deductive'. In short, Σ_1 *is the level of deduction*. When $\mathcal{L}$ is first-order, it comprises all sentences that are (classical) logical consequences of the underlying theory:

Proposition 3. *Suppose that $\mathcal{L} = \mathcal{L}^S_{\omega\omega}$. For all $T \in \mathcal{T}$ and $\varphi \in \mathcal{L}$, if $T \models \varphi$ then $\varphi \in \Sigma_1^{\mathcal{P}}(T)$.*

When $\mathcal{L}$ is first-order, the Σ_1 level of the hierarchies can include more than (classical) logical consequences because the compactness property used to define level Σ_1 of the hierarchies is stronger than the compactness property of first-order logic. Getting back to arbitrary choices of $\mathcal{L}$ and considering all the non-null natural numbers (not just $n = 1$), the following proposition provides another justification for describing the formulas in the Σ_n levels as being compact consequences of the preceding levels:

Proposition 4. *For all $n > 0$, $T \in \mathcal{T}$ and finite $H \subseteq \Sigma_n^{\mathcal{P}}(T)$, $\mathrm{Cn}_{\mathcal{W}}(H) \subseteq \Sigma_n^{\mathcal{P}}(T)$.*

Another important property of the Σ_1 level of the hierarchies that holds under some assumptions is the equivalent of the *deduction theorem*. For the latter to make sense, a special assumption is required.

Proposition 5. *Let $T \in \mathcal{T}$ and finite $D \subseteq \mathcal{L}$ be such that $T \cup D \in \mathcal{T}$. For all $\varphi \in \mathcal{L}$, the following holds.*

1. *If $\bigwedge D \rightarrow \varphi$ belongs to $\Sigma_1^{\mathcal{P}}(T)$ then φ belongs to $\Sigma_1^{\mathcal{P}}(T \cup D)$.*
2. *Suppose that $\mathcal{T}$ contains every $\mathcal{W}$-consistent set of the form $T' \cup D$ for $T' \in \mathcal{T}$. Then $\bigwedge D \rightarrow \varphi$ belongs to $\Sigma_1^{\mathcal{P}}(T)$ iff φ belongs to $\Sigma_1^{\mathcal{P}}(T \cup D)$.*

The following property establishes an important feature of the Π_1 level of the hierarchies: that inferences at this level are based on finite (incomplete) evidence and are directly refutable by evidence statements. Hence, inferences at this level are defeasible.

Property 4. *Let $T \in \mathfrak{T}$ and $\varphi \in \mathrm{Cn}_{\mathcal{W}}(T)$ be given. Then $\varphi \in \Pi_1^{\mathcal{P}}(T)$ iff there exists a finite subset E of T such that for all $T' \in \mathfrak{T}$:*

$$\text{if } E \subseteq T' \text{ and } T' \not\models_{\mathcal{W}} \varphi \text{ then } \neg\varphi \in \Sigma_1^{\mathcal{P}}(T').$$

It is our contention that the formulas inferred at the Π_1 level are inductive conclusions of the evidence statements. Why is this plausible? Firstly, note that statements of the form 'all Xs are Ys' that are deductively refutable on the basis of further evidence will belong to this level. More formally, as we point out later in this paper, the Π_1 levels of the hierarchies are closely related to the Π_1 level of the arithmetical hierarchy in containing all universal formulas that are consequences of the background theory. These assertions are true under the assumption that the evidence statements include all atomic sentences and their negations, and that the class of possible worlds is made up of Henkin structures (for the latter, a sufficient condition for the refutation of a universal sentence is that we have enough names for all objects in the domain of the model). The scientist that concludes that 'all emeralds are green' after carrying out a set of observations is carrying out a Π_1 inference: his conclusion is refutable in the light of future evidence and is a generalization about unobserved entities on the basis of a finite set of evidence statements about observed entities. The property of weak compactness enables the scientist to believe confidently in the truth of φ, relative to that part of the hierarchies below the level which is currently built. The sentences on level Π_1 of $\mathcal{H}(T)$ can be believed to be true with absolute confidence, as far as T is reliable; inductive conclusions are reasonable in the light of background knowledge and the current 'best' theory available to the scientist. That such an inference is defeasible is well-known to the scientist, but that is the status of all scientific knowledge. In the case of observational paradigms, the Π_1 inferences are reasonable insofar as T contains all possible observations true in the underlying world.

Contrary to the Σ_n levels, the Π_n levels of the hierarchies do not enjoy the property of being closed under any natural form of compact extended logical consequence. Generally they are not even closed under logical consequence (in the classical sense). But we still have the following:

Proposition 6. *For all $n > 0$ and $T \in \mathfrak{T}$, $\Pi_n^{\mathcal{P}}(T)$ is closed under disjunction and conjunction.*

We can use Propositions 4 and 6 to obtain the following simplification of the Σ_n (for $n > 1$) and Π_n (for $n > 0$) levels of the hierarchies; formulas will be

Σ_n and Π_n consequences on the basis of finite evidence and a *single* formula among the previously derived consequences.

Proposition 7. *Let $T \in \mathcal{T}$ and $\varphi \in \mathrm{Cn}_{\mathcal{W}}(T)$ be given.*

- *For all $n > 0$, φ belongs to $\Pi_n^{\mathcal{P}}(T)$ iff there exists a finite $E \subseteq T$ and $\psi \in \Sigma_n^{\mathcal{P}}(T)$ such that for all $T' \in \mathcal{T}$:*

$$\text{if } E \subseteq T',\ T' \models_{\mathcal{W}} \psi, \text{ and } T' \not\models_{\mathcal{W}} \varphi \text{ then } \neg\varphi \in \Sigma_n^{\mathcal{P}}(T').$$

- *For all $n > 1$, φ belongs to $\Sigma_n^{\mathcal{P}}(T)$ iff there exists a finite $E \subseteq T$ and $\psi \in \Pi_{n-1}^{\mathcal{P}}(T)$ such that for all $T' \in \mathcal{T}$:*

$$\text{if } E \subseteq T' \text{ and } T' \models_{\mathcal{W}} \psi \text{ then } T' \models_{\mathcal{W}} \varphi.$$

3.3. Properties of Observational Paradigms

A simpler characterization of some levels of the hierarchies can be obtained by taking $\mathcal{P}$ to be observational or perfectly observational. We start with the former and move on to the latter. First note the following formal simplification of the Σ_1 level, which follows directly from Proposition 2 and Property 3, and which involves the notion of logical consequence in $\mathcal{W}$.

Proposition 8. *Suppose that $\mathcal{P}$ is observational. Let $T \in \mathcal{T}$ and $\varphi \in \mathcal{L}$ be given. Then $\varphi \in \Sigma_1^{\mathcal{P}}(T)$ iff there exists finite $E \subseteq T$ such that $E \models_{\mathcal{W}} \varphi$.*

Proposition 8 characterizes the Σ_1 level of the hierarchies in terms of a compactness property that is still closer to the compactness property of first-order logic than the form of compactness given in Property 3. But when are the logical consequences of a possible theory T precisely the Σ_1 consequences of T in $\mathcal{P}$? When $\mathcal{L}$ is first-order, observational paradigms suggest natural examples of sufficient conditions. Here is one:

Proposition 9. *Suppose that $\mathcal{L} = \mathcal{L}_{\omega\omega}^S$, $\mathcal{P}$ is observational, and no finite subset of $\mathcal{L}$ is both consistent and $\mathcal{W}$-inconsistent. Then for all $T \in \mathcal{T}$ and sentences φ, $T \models \varphi$ iff $\varphi \in \Sigma_1^{\mathcal{P}}(T)$.*

Proposition 9 can be applied to two more concrete classes of paradigms. First we can require $\mathcal{W}$ to be just what it is in classical logic, namely the class of all structures:

Proposition 10. *Suppose that $\mathcal{L} = \mathcal{L}_{\omega\omega}^S$, $\mathcal{W}$ is the class of all structures, and $\mathcal{P}$ is observational. For all $T \subseteq \mathcal{L}$ and sentences φ, $T \models \varphi$ iff $\varphi \in \Sigma_1^{\mathcal{P}}(T)$.*

Then using Proposition 1, Proposition 9 can be applied to the class of all Henkin structures with an additional constraint placed on the vocabulary S:

Proposition 11. *Suppose that S contains infinitely many constants, $\mathcal{L} = \mathcal{L}^S_{\omega\omega}$, $\mathcal{W}$ is the class of all Henkin structures, and $\mathcal{P}$ is observational. For all $T \in \mathcal{T}$ and sentences φ, $T \models \varphi$ iff $\varphi \in \Sigma^{\mathcal{P}}_1(T)$.*

The propositions above exhibit cases where we do not go beyond classical first-order consequences at the Σ_1 level, providing another intuitive justification for terming this the level of deduction.

Now we consider perfectly observational paradigms. Proposition 8 provides the simplest characterization of the Σ_1 level. But there is an interesting characterization of the Π_1 level for perfectly observational paradigms that does not hold for observational paradigms in general. Indeed, in that case, the formulas occurring on the Π_1 level but not on the Σ_1 level, can be defined directly in terms of the class of possible worlds and the set of possible observations (in terms of which observational and perfectly obse rvational paradigms are defined):

Proposition 12. *Suppose that $\mathcal{P}$ is perfectly observational for $\mathcal{O}$. Let $T \in \mathcal{T}$ and $\varphi \in \mathrm{Cn}_{\mathcal{W}}(T)$ be given. Then φ belongs to $\Pi^{\mathcal{P}}_1(T)$ iff $\varphi \in \Sigma^{\mathcal{P}}_1(T)$ or there exists a finite $E \subseteq T$ and a non-empty set X of non-empty finite subsets of $\mathcal{O}$ such that:*

$$E \models_{\mathcal{W}} \varphi \leftrightarrow \bigwedge\{\neg \bigwedge D \mid D \in X\}.$$

3.4. Refutability and Π_n Levels of the Hierarchy

We now address the question of the 'difficulty' of inferring a sentence φ in general, from an arbitrary possible theory. A natural approach is to identify the difficulty of inferring φ, in general (in a given paradigm), with the least level which is not strictly below any level at which φ can be inferred from a a possible theory T. This suggests the following:

Definition 7. *Let $n \in \mathbb{N}$ and φ be given. We say that φ is Σ_n* in $\mathcal{P}$ *iff for all $T \in \mathcal{T}$, if $T \models_{\mathcal{W}} \varphi$ then $\varphi \in \Sigma^{\mathcal{P}}_n(T)$. We say that φ is Π_n* in $\mathcal{P}$ *iff for all $T \in \mathcal{T}$, if $T \models_{\mathcal{W}} \varphi$ then $\varphi \in \Pi^{\mathcal{P}}_n(T)$.*

Another important notion is related to a sentence being Σ_n in $\mathcal{P}$. Recall that the levels Σ_n of the hierarchies are defined from a compactness property, and the levels Π_n from a property of weak compactness. Both properties involve finite sets of formulas derived previously in the hierarchies. For the Σ_n levels with $n > 1$, the finite sets in question have to be non-empty in order to generate new formulas i.e., formulas not previously derived below Σ_n. But this does not have to be true of the Π_n levels. Indeed, having new formulas appear at level Σ_n raises the possibility that some new sentences appear at level Π_n, such that if they are not logical consequences of the underlying theory in $\mathcal{W}$, then their negations appear on the Σ_n level of the hierarchy built from that theory. For these formulas, there is no need to 'filter' the theories with a non-empty finite set of previously occurring formulas. The property of weak compactness can be non-trivial in that new consequences can be derived, at every Π_n level, even when $E, H = \emptyset$. We introduce a definition for this notion:

Definition 8. *Let $n \in \mathbb{N}$ and sentence φ be given. We say that φ is n-*refutable *just in case for all $T \in \mathcal{T}$, if $T \not\models_{\mathcal{W}} \varphi$ then $\neg\varphi \in \Sigma_n^{\mathcal{P}}(T)$.*

Property 5. *For all $n > 0$, every n-refutable sentence is Π_n in $\mathcal{P}$.*

The Π_n level of the hierarchy over a possible theory T includes n-refutable sentences: formulas which are in $\Pi_n^{\mathcal{P}}(T)$ on the basis of $(\emptyset, \emptyset)$ (see Definition 6). In particular, the Π_1 level includes the 1-refutable sentences which are the counterparts of what we term, in the Σ_1 level, *generalized tautologies*: formulas which follow, in the general sense, from any possible theory—thus conveying no information. Here is an immediate example of 1-refutable sentences, interesting only in the case when the set of possible observations is not closed under negation:

Property 6. *If $\mathcal{P}$ is perfectly observational for $\mathcal{O}$ then for all $\varphi \in \mathcal{O}$, $\neg\varphi$ is 1-refutable.*

Clearly, the refutation of a formula is 'more difficult' as we move higher in the hierarchy. There are straightforward relations between the notions of n-refutability and the notion of a formula being Σ_n in $\mathcal{P}$:

Proposition 13. *For all $n > 0$, a sentence φ is n-refutable iff:*

1. *$\neg\varphi$ is Σ_n in $\mathcal{P}$, and*
2. *for all $T \in \mathcal{T}$, either $T \models_{\mathcal{W}} \varphi$ or $T \models_{\mathcal{W}} \neg\varphi$.*

By the second point above, n-refutability implies a form of completeness of any possible theory T w.r.t. φ.

Now, to the question: "given possible theory T, how difficult is it to infer a given logical consequence of T in $\mathcal{W}$, say φ, from T?" the hierarchy over T provides the following answer: "the smallest level (either Σ_n or Π_n) at which φ occurs". If φ does not occur in the hierarchy over T, then inferring φ from T is more difficult than the hierarchy can measure. The following definition formalizes this assessment of the difficulty of an inference in the hierarchy:

Definition 9. *The* index of $\mathcal{P}$ *is $n \in \mathbb{N}$ if n is least such that every sentence is Σ_n in $\mathcal{P}$; if there is no $n \in \mathbb{N}$ such that every sentence is Σ_n in $\mathcal{P}$ the index is ω.*

The following provides necessary and sufficient conditions for hierarchies to possess finite indices:

Proposition 14. *Let $n > 0$ be such that $\Pi_n^{\mathcal{P}}(T) = \Sigma_n^{\mathcal{P}}(T)$ for all $T \in \mathcal{T}$. Then the index of $\mathcal{P}$ is at most equal to n.*

The index is finite then, if and only if induction comes to an end at some level of all hierarchies in the paradigm (for all possible theories). It is certainly an important issue, to be addressed later, whether the hierarchy over T contains all sentences which are logical consequences of T in $\mathcal{W}$.

4. Connections with the arithmetical hierarchy and learning theory

In this section we explore connections between our hierarchies of generalized logical consequence and the arithmetical hierarchy; we also explore connections with an influential framework, that of formal learning theory. The arithmetical hierarchy classifies first-order sentences with respect to their syntactic complexity; our framework constructs hierarchies with respect to the difficulty of inferring sentences. Similarities in these hierarchies provide an interesting connection between the syntactic complexity of sentences and their inference. Such a connection is necessarily incomplete, of course, since the arithmetical hierarchy assumes the language of arithmetic along with its standard interpretation, whereas our framework leaves these unspecified. A comparison with learning theory is useful in enabling a determination of the similarity (and important differences) between the concept of 'learnable in the limit' and our notion of inductive inference. Since we have already defined observational paradigms in our framework as the counterpart of paradigms in learning theory, it would be useful to see how much of the framework's properties carry over in our setting.

4.1. THE ARITHMETICAL HIERARCHY

We now derive relationships between our hierarchies of generalized consequence and the arithmetical hierarchy. This will justify the use of 'Σ_n' and 'Π_n' for denoting the levels of the hierarchies of generalized logical consequences. More specifically, we investigate the question of the *completeness* of the hierarchies i.e., whether any logical consequence in $\mathcal{W}$ of some possible theory T belongs to the hierarchy over T, at some level. The following two propositions get us part of the way:

Proposition 15. *Suppose that $\mathcal{P}$ is perfectly observational. Let $n > 0$, $T \in \mathcal{T}$, and non-empty set X of sentences be given.*

1. *If $X \subseteq \Sigma_n^{\mathcal{P}}(T)$, then $\bigvee X$ belongs to $\Sigma_n^{\mathcal{P}}(T)$.*
2. *If all members of X are n-refutable in $\mathcal{P}$, then $\bigwedge X$ is n-refutable in $\mathcal{P}$.*

When $\mathcal{W}$ consists exclusively of Henkin structures, $\exists x \varphi(x)$ is logically equivalent in $\mathcal{W}$ to $\bigvee\{\varphi[t/x] \mid \text{closed term } t\}$ and $\forall x \varphi(x)$ is logically equivalent in $\mathcal{W}$ to $\bigwedge\{\varphi[t/x] \mid \text{closed term } t\}$. Hence the previous applies:

Proposition 16. *Suppose that $\mathcal{W}$ is a class of Henkin structures and $\mathcal{P}$ is perfectly observational. Let $n > 0$, $T \in \mathcal{T}$, and formula φ be given. Denote by X the set of all closed instances of φ.*

1. *If $X \subseteq \Sigma_n^{\mathcal{P}}(T)$, then the existential closure of φ belongs to $\Sigma_n^{\mathcal{P}}(T)$.*
2. *If all members of X are n-refutable in $\mathcal{P}$, then the universal closure of φ is n-refutable in $\mathcal{P}$.*

Applying the propositions above, and under certain assumptions about the paradigm $\mathcal{P}$ under investigation—in particular, assuming that $\mathcal{P}$ is perfectly observational—we obtain a result that shows that the location of a sentence in the hierarchy of generalised logical consequences may be determined by inspection of the syntactic complexity of the sentence. This is precisely akin to the location of sentences in the arithmetical hierarchy, which is the content of the following proposition. It uses the notion of being Σ_n or Π_n, applied to members of $\mathcal{L}$ some of which can be infinite formulas, since we are dealing with countable fragments of $\mathcal{L}^S_{\omega_1\omega}$. Basically, countable disjunctions play the same role as existential quantifiers, and countable conjunctions the same role as universal quantifiers. The inductive proof of the proposition is based on a formal definition of that notion.

Proposition 17. *Suppose that $\mathcal{W}$ is a class of Henkin structures and that $\mathcal{P}$ is observational for the set of atomic sentences and their negations. For all $n > 0$, every Σ_n sentence is Σ_n in $\mathcal{P}$, and every Π_n sentence is n-refutable, hence Π_n, in $\mathcal{P}$.*

There are important distinctions as well between the syntactic complexity of a sentence and the level at which it can occur in a hierarchy of generalized logical consequences. Assume that our vocabulary is that of arithmetic with the usual successor, zero and addition and multiplication function symbols. Assume that $\mathcal{L}$ is first-order and $\mathcal{W}$ is the class of all Henkin structures. One possible theory of the paradigm is the set of all atomic sentences and their negations that are true of the natural numbers, say T_N. Robinson arithmetic is a finite set of Π_1 sentences, say D. Hence by Proposition 17, all members of D occur at level Π_1 of the hierarchy built over T_N. This set D characterizes natural numbers up to isomorphism and every sentence true in $\mathbb{N}$ will be a compact consequence of D in $\mathcal{W}$. Hence, every sentence true in $\mathbb{N}$ will occur at level Σ_2 of the hierarchy of logical consequences over T_N. So the arithmetical hierarchy, consisting of the formulas true of the natural numbers, collapses to level Σ_2 of the hierarchy of logical consequences over T_N. Note that it has only collapsed to this particular hierarchy, and not to others, built over other possible theories T.

Thanks to Proposition 17 we can derive the following result which indicates a form of completeness of the hierarchy of logical consequences over any possible theory i.e., that the consequences in $\mathcal{W}$ of a background theory T occur at some level of the hierarchy. That is, all consequences, compact or otherwise, will be found in the hierarchy.

Proposition 18. *Suppose that $\mathcal{W}$ is a class of Henkin structures. Suppose that $\mathcal{P}$ is observational for the set of atomic sentences and their negations. Then for all $T \in \mathcal{T}$, $\bigcup_{n\in\mathbb{N}} \Sigma_n^{\mathcal{P}}(T) = \mathrm{Cn}_{\mathcal{W}}(T)$.*

4.2. Learning Theory

In this section we highlight some connections that may be drawn between our framework and concepts in formal learning theory. Our study bears close links with learning theory in a logical setting. For the sake of simplicity of a short introduction to learning theory, we briefly review some basic concepts in the numerical setting. A *language* is a non-empty set of natural numbers. We can identify such such a language L with a Henkin structure $\mathfrak{M}_L$ satisfying the set of formulas $T_L = \{P(\overline{n}) \mid n \in L\} \cup \{\neg P(\overline{n}) \mid n \notin L\}$ (assuming that the vocabulary contains a unary predicate P and has terms representing the numerals). Corresponding to sets of languages, $\mathbb{L}$, considered by learning theory, we have natural choices available for the class of intended structures i.e., $\mathcal{W} = \{\mathfrak{M}_L \mid L \in \mathbb{L}\}$ and for the set of possible theories i.e., $\mathcal{T} = \{T_L \mid L \in \mathbb{L}\}$. Learning theory aims to provide *descriptions* of languages supplied as sequences of inputs, that is, given an enumeration of a language L, converge in the limit to a description of L. If $\mathbb{L}$ consists of say, r.e. languages then the description of the language could be the index of a Turing machine that computes the characteristic function for L. Corresponding to this task, in our framework would be the following: given an enumeration of a theory, converge in the limit to a description of the theory or the underlying Henkin model. For instance, assume that L is the set of even numbers. The corresponding theory is $T_L = \{P(\overline{0}), P(\overline{2}) \ldots\}$. Assume that the numerals are represented using the constant $\overline{0}$ and the unary function symbol s (successor). Then, a representation of T_L might be $\varphi_L = P(\overline{0}) \wedge \neg P(s(\overline{0})) \wedge \forall x(P(x) \rightarrow P(s(s(x)))) \wedge \forall x(\neg P(x) \rightarrow \neg P(s(s(x))))$. Infinite formulas can also be used for this purpose when no finite formula is suitable.

We now examine the connections between levels of inference and such a convergence. In particular, we examine the similarities between the concept of learning in the limit and Σ_2 inference, and the concept of finite learning and Σ_1 inference. We will show that 'inferred at Σ_2 level of the hierarchy of logical consequences' implies learning in the limit. The converse is also true in the particular case of observational paradigms, provided that the language is made rich enough by possible adding (at most countably many) infinite formulas. We first clarify some basic notation. The range of a (finite or infinite) sequence e is denoted $\mathrm{rng}(e)$. Given a sequence e indexed by $\mathbb{N}$, the initial segment of e of length k is denoted by $e[k]$. For simplicity, and in order not to have to add an additional symbol to $\mathcal{L}$ that would be interpreted as 'pause,' or 'no datum observed,' we assume in this part that no member of $\mathcal{T}$ is empty.

Definition 10. *Given $T \in \mathcal{T}$, an* environment for T *is a sequence indexed by $\mathbb{N}$ whose range is T. A* learner *is a partial function from the set of finite sequences of sentences into $2^{\mathcal{L}}$.*

Our notion of environment corresponds to the *texts* of formal learning theory in the case when $\mathcal{O}$ is the set of atomic sentences. When $\mathcal{O}$ is the set of atomic sentences and their negations, environments correspond to the *informants* of formal learning theory (Jain et al., 1999). Since $\mathcal{O}$ is arbitrary, we prefer the inclusive notion of an environment (as used in (Martin and Osherson, 1998)).

Having defined the notion of a 'learner' we can now define different notions of successful learning. The following establishes the first kind:

Definition 11. *Let learner f and $\varphi \in \mathcal{L}$ be given. We say that f* identifies φ in the limit in $\mathcal{P}$ *iff for all $T \in \mathcal{T}$ and environments e for T, the following clauses are equivalent:*

1. $T \models_{\mathcal{W}} \varphi$.
2. $\{k \in \mathbb{N} \mid \varphi \notin f(e[k])\}$ *is finite.*

In the case when $T \not\models_{\mathcal{W}} \varphi$, note that the learner should output a response other than φ infinitely many times. The learner is still able to output φ infinitely many times. A stricter criterion would require that when $T \not\models_{\mathcal{W}} \varphi$, the learner outputs φ only finitely many times. If we are interested in discovering the truth of φ but not its untruth, this is perhaps less natural than the criterion formalized in Definition 11. Generalized to sets of sentences rather than single sentences, the stricter criterion is embodied in the following definition.

Definition 12. *Let learner f and $X \subseteq \mathcal{L}$ be given. We say that f* is X-reliable in $\mathcal{P}$ *iff for all $T \in \mathcal{T}$ and environments e for T, $f(e[k]) \cap X$ is equal to $\{\varphi \in X \mid T \models_{\mathcal{W}} \varphi\}$ for all but finitely many $k \in \mathbb{N}$.*

Definition 12 defines a learner that is capable of correctly identifying all and only those formulas from a given set X that are consequences of the background theory. That is, it enables us to reach a point in time where we can successfully identify those consequences of the background theory that are members of X. In contrast, with the less powerful learners of Definition 10, we might have to be content with the piece-wise identification of members of X that are consequences of the background theory. Let us record the obvious fact that the second criterion is stronger than the first one:

Property 7. *Let learner f and $X \subseteq \mathcal{L}$ be given. If f is X-reliable in $\mathcal{P}$, then f identifies in the limit every member of X in $\mathcal{P}$.*

Given a set X of sentences, and for all $\varphi \in X$, a learner f_φ that learns φ in the limit, we can aggregate all such f_φ to obtain a learner whose competence encompasses all of them.

Property 8. *Let $X \subseteq \mathcal{L}$ be given. The following clauses are equivalent.*

1. *Some learner identifies all members of X in the limit in $\mathcal{P}$.*
2. *For all $\varphi \in X$, there exists a learner which identifies φ in the limit in $\mathcal{P}$.*

The following result provides another relationship between the criteria of success for learners in paradigms based on learning theory. Given a choice of $\mathcal{T}$ and $\mathcal{W}$ as indicated in our introduction to this section, we now show that with a natural choice of X as a set of descriptions of $\mathcal{T}$ (ergo, any member of X is implied in $\mathcal{W}$ by a unique member of $\mathcal{T}$), the notion of learning in the limit and X-reliability are equivalent. The property above asserted that they are equivalent in the finite case—the result at hand shows that they are equivalent given a particular choice of X which may or may not be finite.

Proposition 19. *Let $X \subseteq \mathcal{L}$ be given. Suppose that for all $T \in \mathcal{T}$, there exists a unique $\varphi \in X$ such that $T \models_{\mathcal{W}} \varphi$. The following clauses are equivalent.*

1. *Some learner is X-reliable in $\mathcal{P}$.*
2. *Some learner identifies all members of X in the limit in $\mathcal{P}$.*

We now establish a notion from learning theory that corresponds to the non-defeasibility of deductive inference.

Definition 13. *Let learner f and $\varphi \in \mathcal{L}$ be given. We say that f is f* identifies φ in $\mathcal{P}$ with no mind changes *just in case f identifies φ in $\mathcal{P}$ in the limit and for all $T \in \mathcal{T}$, environments e for T, and $k \in \mathbb{N}$, if $f(e[k])$ is defined then $f(e[k]) = f(e[k+1])$.*

A learner as defined above, is such that once it has issued a response to an initial segment of an input stream, it continues to issue the same response in the face of any new data in subsequent segments of the same input stream. The following proposition asserts that sentences at the Σ_1 level are precisely those sentences that can be learned without mind changes. This connection confirms again our decision to treat the Σ_1 level as that of deductive inference, since these formulas are non-defeasible in the face of new data: their inference is certain given the evidence. This correspondence between learning theory and our framework is independent of the choice of paradigm.

Proposition 20. *Let sentence φ be given. The following are equivalent.*

1. *φ is Σ_1 in $\mathcal{P}$.*

2. *Some learner identifies φ in the limit in $\mathcal{P}$ with no mind changes.*

The following result establishes a compactness property for characterizing those sets of formulas that can be reliably learnt without mind changes with no conditions laid on the underlying paradigm.

Proposition 21. *Given a set X of sentences, the following are equivalent.*

1. *Some learner is X-reliable in $\mathcal{P}$ without mind changes.*

2. *For all $T \in \mathcal{T}$, there exists finite $D \subseteq T$ such that for all $T' \in \mathcal{T}$:*

$$D \subseteq T' \implies \mathrm{Cn}_{\mathcal{W}}(T) \cap X = \mathrm{Cn}_{\mathcal{W}}(T') \cap X.$$

The following proposition asserts that formulas occurring at the Σ_2 level are learnable in the limit as defined above. Since the Σ_2 level includes the Π_1 and Σ_1 levels as well, this means we have subsumed the concept of learning in the limit to several levels of our hierarchy.

Proposition 22. *Let a sentence φ be Σ_2 in $\mathcal{P}$. Then some learner identifies φ in the limit in $\mathcal{P}$.*

If $\mathcal{O}$ is a set of sentences that is closed under negation, it can be shown that the converse of Proposition 22 is also true for paradigms that are observational for $\mathcal{O}$, provided that $\mathcal{L}$ is made expressive enough *i.e.*, by extending the language of first-order logic to the appropriate fragment of $\mathcal{L}^S_{\omega_1\omega}$. The argument is a variation on the characterization of learnability as appears in (Martin and Osherson, 1998, Prop. 3.(35), p. 75).

Proposition 23. *Suppose that $\mathcal{P}$ is observational for some set of sentences closed under negation. There exists a countable fragment $\mathcal{L}'$ of $\mathcal{L}^S_{\omega_1\omega}$ that extends $\mathcal{L}$ such that the following holds. Set $\mathcal{P}' = (\mathcal{L}', \mathcal{W}, \mathcal{T})$. For all sentences φ, if some learner identifies φ in the limit in $\mathcal{P}$, then φ is Σ_2 in $\mathcal{P}'$.*

While we have established a correspondence between the Σ_1 and Σ_2 levels and concepts in learning theory, no such correspondence has been made for the Π_1 level. This is not accidental and has import for our enterprise. Our classification of the latter as the level of induction means that inferences at this level do not correspond to any fundamental category of learning theory, and certainly not learnable in the limit. The concept of learnable in the limit is too weak for our purposes since it does not make the crucial distinction between those inferences that are purely inductive and those that consist of induction followed by deduction. Consider as an example, the physicist that induces a law from evidence statements and then using this law deduces several of its consequences for further scientific investigation. This kind of inference, we believe, corresponds to a Π_1 inference followed by Σ_2 inferences. The conditions under which the scientist discards the newly deduced laws corresponds to the discovery of new evidence that refutes the law originally induced. The concept of learnability in the limit is too crude to make a distinction between these two kinds of laws. In essence, the learning theorist identifies all learning in the limit with inductive inference. We draw a finer distinction, and only identify a part of it with induction. The kind of inductive inference we propose retains a certain measure of defeasibility at all times. The picture of the hierarchies then, is as follows: the first level, Σ_1, corresponds to deductive or non-defeasible inferences. The second, Π_1, that of inductive inference includes those formulas which are are present by virtue of direct defeasibility i.e., the evidence statements lead to the refutation of the putative conclusion itself. The third level, Σ_2, corresponds to those formulas that are deduced from the Π_1 level and are indirectly defeasible i.e., the evidence statements lead to the refutation of the formulas on which those inferences are based.

5. Conclusion

The framework presented in the current paper is designed to present the foundation for a logic of both deduction and induction. We have sacrificed full generality in the formalism presented here for clarity of exposition. Full generality would utilize an expanded notion of generalized consequence. This requires the consideration of only those models that satisfy some minimality criteria such as minimizing the extension of the set of all formulas of a given kind true in the underlying world. The full model also generalizes the hierarchies to measure the difficulty of measuring whether a given sentence is, or *is not*, a generalized logical consequence of a theory. In this manner, deeper connections with the concepts of formal learning theory as well as more general results can be obtained. In our framework, we have not deviated from a model-theoretic Tarskian view of logical consequence. The debate on whether such a notion is the correct one is certainly worth engaging in, and the issues that it raises are rich and interesting. For the current study, however, such an engagement would draw us too far afield. It would clearly be of much theoretical and philosophical interest to investigate the possibility of a similar framework being constructed for a non-Tarskian notion of logical consequence.

The picture of scientific discovery and investigation that is portrayed is of the scientist demarcating a class of possible worlds as the object of inquiry, and selecting a set of sentences as the possible observations made in those worlds. We can think of these actions as establishing a paradigm for further investigation—a demarcation of the physically and methodologically possible from the merely logical possible. With these paradigms in hand, the scientist commences experiment design and testing. Simple deduction is but the first level of the scientist's logical activity, to be followed by successive acts of inductive and deductive inferences based on the conclusions reached in the hierarchy. With the full machinery in place, being learnable in the limit and occurring on level Σ_2 or below of the hierarchies of generalized logical consequences, are equivalent notions for classes of paradigms that are extremely general, much more than the observational paradigms classically considered by formal learning theory. Our framework raises the important issue of whether generalized logical consequences that occur on higher levels can still be the object of some form of scientific inquiry.

Acknowledgements

The authors would like to thank Daniel Osherson for his useful comments on our article. Eric Martin is supported by the Australian Research Council Grant A49803051. Samir Chopra is supported by Australian Research Council Grant A49805142.

Appendix

Proofs of all major results are presented in this section.

Proof of Proposition 1: Let T be a consistent subset of $\mathcal{L}$ in which infinitely many constants in S do not occur. Denote by C the set of constants in S that do not occur in T. Set $S' = S \setminus C$. By the Löwenheim-Skolem theorem for $\mathcal{L}_{\omega_1\omega}$ (see (Ebbinghaus et al., 1994, p. 139)) there exists a denumerable S'-structure $\mathfrak{M}$ that is a model of T. Since C is infinite, there exists a total mapping f from C onto the domain of $\mathfrak{M}$. Denote by $\mathfrak{N}$ the structure that is the expansion of $\mathfrak{M}$ obtained when c is interpreted by $f(c)$, for all $c \in C$. Clearly, $\mathfrak{N}$ is a Henkin structure and a model of T. $\square$

Proof of Proposition 2: Trivially, the first clause implies the second one. Conversely, suppose that for all $T \in \mathcal{T}$ with $X \subseteq T$, $T \models_{\mathcal{W}} \varphi$. Let $\mathfrak{M} \in \mathcal{W}$ be a model of X. It suffices to show that $\mathfrak{M} \models \varphi$. Since $X \subseteq \mathcal{O}$, X is a subset of $D_{\mathcal{O}}(\mathfrak{M})$. Since $\mathcal{P}$ is observational, $D_{\mathcal{O}}(\mathfrak{M})$ is a member of $\mathcal{T}$, and as $\mathfrak{M}$ is a model of $D_{\mathcal{O}}(\mathfrak{M})$ in $\mathcal{W}$, $\mathfrak{M} \models \varphi$, as wanted. $\square$

Proof of Proposition 4: Let $n > 0$, $T \in \mathcal{T}$ and finite $H \subseteq \Sigma_n^{\mathcal{P}}(T)$ be given. Let sentence φ be such that $H \models_{\mathcal{W}} \varphi$. By Definition 5, for all $\psi \in H$ we can choose a finite subset E_ψ of T and a finite subset H_ψ of $\bigcup_{m<n} \Pi^{\mathcal{P}}_m(T)$ such that ψ belongs to $\Sigma^{\mathcal{P}}_n(T)$ on the basis of (E_ψ, H_ψ). Define E as $\bigcup_{\psi \in H} E_\psi$ and H' as $\bigcup_{\psi \in H} H_\psi$. Since $E \subseteq T$ and $H' \subseteq \bigcup_{m<n} \Pi^{\mathcal{P}}_m(T)$, it follows easily from Definition 5 again that φ belongs to $\Sigma^{\mathcal{P}}_n(T)$ on the basis of (E, H'). $\square$

Proof of Proposition 5: Let $\varphi \in \mathcal{L}$ be given. Clause *1.* follows immediately from Property 2 and Proposition 4. We prove clause *2.* Suppose that $\mathcal{T}$ contains every $\mathcal{W}$-consistent set of the form $T' \cup D$ for $T' \in \mathcal{T}$. Assume that $\varphi \in \Sigma^{\mathcal{P}}_1(T \cup D)$. By Definition 5, choose a finite subset E of T such that φ belongs to $\Sigma^{\mathcal{P}}_1(T \cup D)$ on the basis of $E \cup D$. Let $T' \in \mathcal{T}$ be such that $E \subseteq T'$. Assume that $T' \cup D$ is $\mathcal{W}$-consistent. Then $T' \cup D \in \mathcal{T}$ by hypothesis. With the facts that $E \cup D \subseteq T' \cup D$ and φ belongs to $\Sigma^{\mathcal{P}}_1(T \cup D)$ on the basis of $E \cup D$, we infer that $T' \cup D \models_{\mathcal{W}} \varphi$. So we have shown that for all $\mathfrak{M} \in \mathrm{Mod}_{\mathcal{W}}(T')$, if $\mathfrak{M} \models D$ then $\mathfrak{M} \models \varphi$, and we conclude with Definition 5 again that $\bigwedge D \rightarrow \varphi$ belongs to $\Sigma^{\mathcal{P}}_1(T)$ on the basis of E. This completes the proof of one half of clause *2.*, and the other half follows from clause *1.* $\square$

Proof of Proposition 6: Let $n > 0$, $T \in \mathcal{T}$, and $\varphi_1, \varphi_2 \in \Pi^{\mathcal{P}}_n(T)$ be given. By Definition 6, let finite $E_1, E_2 \subseteq T$ and finite $H_1, H_2 \subseteq \Pi^{\mathcal{P}}_n(T)$ be such that φ_1 belongs to $\Pi^{\mathcal{P}}_n(T)$ on the basis of (E_1, H_1), and φ_2 belongs to $\Pi^{\mathcal{P}}_n(T)$ on the basis of (E_2, H_2). Set $E = E_1 \cup E_2$ and $H = H_1 \cup H_2$. Let $T' \in \mathcal{T}$ be such that $E \subseteq T'$, $T' \models_{\mathcal{W}} H$, and $T' \not\models_{\mathcal{W}} \varphi_1 \vee \varphi_2$. Then $T' \not\models_{\mathcal{W}} \varphi_1$ and $T' \not\models_{\mathcal{W}} \varphi_2$. Hence $\{\neg\varphi_1, \neg\varphi_2\} \subseteq \Sigma^{\mathcal{P}}_n(T')$, which implies that $\neg(\varphi_1 \vee \varphi_2) \in \Sigma^{\mathcal{P}}_n(T')$. We conclude with Definition 6 that $\varphi_1 \vee \varphi_2$ belongs to $\Pi^{\mathcal{P}}_n(T)$ on the basis of (E, H). Let $T' \in \mathcal{T}$ be such that $E \subseteq T'$, $T' \models_{\mathcal{W}} H$, and $T' \not\models_{\mathcal{W}} \varphi_1 \wedge \varphi_2$. Then either $T' \not\models_{\mathcal{W}} \varphi_1$ or $T' \not\models_{\mathcal{W}} \varphi_2$. Hence $\neg\varphi_1 \in \Sigma^{\mathcal{P}}_n(T')$ or $\neg\varphi_2 \in \Sigma^{\mathcal{P}}_n(T')$, which implies that $\neg(\varphi_1 \wedge \varphi_2) \in \Sigma^{\mathcal{P}}_n(T')$. We conclude with Definition 6 that $\varphi_1 \wedge \varphi_2$ belongs to $\Pi^{\mathcal{P}}_n(T)$ on the basis of (E, H). $\square$

Proof of Proposition 9: Let set $\mathcal{O}$ of sentences be such that $\mathcal{P}$ is observational for $\mathcal{O}$. Let $T \in \mathcal{T}$ and sentence φ be given. If $T \models \varphi$ then $\varphi \in \Sigma^{\mathcal{P}}_1(T)$ by Proposition 3. Conversely, suppose that $T \not\models \varphi$. Let finite $E \subseteq T$ be given. Then $E \cup \{\neg\varphi\}$ is consistent, hence $\mathcal{W}$-consistent by hypothesis. If $\mathfrak{M} \in \mathcal{W}$ is a model of $E \cup \{\neg\varphi\}$ then $\mathcal{T}$ contains $D_{\mathcal{O}}(\mathfrak{M})$. Since $E \subseteq D_{\mathcal{O}}(\mathfrak{M})$ and $D_{\mathcal{O}}(\mathfrak{M}) \not\models_{\mathcal{W}} \varphi$, we infer that φ does not belong to $\Sigma^{\mathcal{P}}_1(D_{\mathcal{O}}(\mathfrak{M}))$ on the basis of $(E, \emptyset)$. So by Property 3, $\varphi \notin \Sigma^{\mathcal{P}}_1(T)$, which completes the proof of the proposition. $\square$

Proof of Proposition 12: Suppose that $\varphi \in \Pi^{\mathcal{P}}_1(T) \setminus \Sigma^{\mathcal{P}}_1(T)$. By Definitions 6 and 5, we can choose finite $E \subseteq T$ and finite $H \subseteq \Sigma^{\mathcal{P}}_1(T)$ such that φ belongs to $\Pi^{\mathcal{P}}_1(T)$ on the basis of (E, H), and for all $\psi \in H$, ψ belongs to $\Sigma^{\mathcal{P}}_1(T)$ on the basis of $(E, \emptyset)$. By Proposition 8, for all $T' \in \mathcal{T}$ such that $E \subseteq T'$, $T' \models_{\mathcal{W}} H$ and $T' \not\models_{\mathcal{W}} \varphi$, we can choose finite $E_{T'} \subseteq T'$ such that $E_{T'} \models_{\mathcal{W}} \neg\varphi$. Let X be the set of all sets of form $E_{T'}$, for all $T' \in \mathcal{T}$ with $E \subseteq T'$, $T' \models_{\mathcal{W}} H$ and $T' \not\models_{\mathcal{W}} \varphi$. Since $\varphi \notin \Sigma_1^{\mathcal{P}}(T)$, X is a non-empty set of non-empty subsets of $\mathcal{O}$, and it is easy to verify that $E \models_{\mathcal{W}} \varphi \leftrightarrow \bigwedge\{\neg \bigwedge D \mid D \in X\}$.

Conversely, let finite subset E of T and non-empty set X of non-empty finite subsets of $\mathcal{O}$ be such that $E \models_{\mathcal{W}} \varphi \leftrightarrow \bigwedge\{\neg \bigwedge D \mid D \in X\}$. Note that for all $T' \in \mathcal{T}$ and $D \in X$, if $E \cup D \subseteq T'$ then $\neg\varphi$ belongs to $\Sigma_1^{\mathcal{P}}(T')$ on the basis of $(D \cup E, \emptyset)$. To complete the proof of the proposition it suffices to show that φ belongs to $\Pi_1^{\mathcal{P}}(T)$ on the basis of $(E, \emptyset)$. Let $T' \in \mathcal{T}$ be such that $E \subseteq T'$ and $T' \not\models_{\mathcal{W}} \varphi$. Since $\mathcal{P}$ is perfectly observational, $T' \models_{\mathcal{W}} \neg\varphi$. Since $\mathcal{P}$ is perfectly observational again, there exists $D \in X$ such that $T' \models_{\mathcal{W}} D$. So $E \cup D \subseteq T'$, which implies that $\neg\varphi \in \Sigma_1^{\mathcal{P}}(T')$, and we are done. $\square$

Proof of Proposition 13: Let $n > 0$ and sentence φ be given. Suppose that φ is n-refutable. For all $T \in \mathcal{T}$, if $T \models_{\mathcal{W}} \neg\varphi$ then $T \not\models_{\mathcal{W}} \varphi$, hence $\neg\varphi \in \Sigma_n^{\mathcal{P}}(T)$. So $\neg\varphi$ is Σ_n for $\mathcal{P}$. Let $T \in \mathcal{T}$ be given, and suppose that $\varphi \notin \mathrm{Cn}_{\mathcal{W}}(T)$. Then $\neg\varphi \in \Sigma_n^{\mathcal{P}}(T)$, so $\neg\varphi \in \mathrm{Cn}_{\mathcal{W}}(T)$. Conversely, suppose that clauses *1.* and *2.* are satisfied. If $T \not\models_{\mathcal{W}} \varphi$ then $\neg\varphi \in \mathrm{Cn}_{\mathcal{W}}(T)$ (by *1.*), hence $\neg\varphi \in \Sigma_n^{\mathcal{P}}(T)$ (by *2.*). So φ is n-refutable. $\square$

Proof of Proposition 14: Assume for a contradiction that $\Sigma_m^{\mathcal{P}}(T) \neq \Sigma_n^{\mathcal{P}}(T)$ for some $T \in \mathcal{T}$ and $m > n$. There are two cases to consider.

1st Case: there exists $m \geq n$ and $T \in \mathcal{T}$ such that $\Pi_m^{\mathcal{P}}(T) = \Sigma_n^{\mathcal{P}}(T)$ and $\Sigma_{m+1}^{\mathcal{P}}(T) \neq \Pi_m^{\mathcal{P}}(T)$. Choose $\varphi \in \mathcal{L}$, finite $E \subseteq T$, and a finite subset H of $\Pi_m^{\mathcal{P}}(T)$ such that φ belongs to $\Sigma_{m+1}^{\mathcal{P}}(T)$ on the basis of (E, H) and $\varphi \notin \Pi_m^{\mathcal{P}}(T)$. Since $H \subseteq \Sigma_n^{\mathcal{P}}(T)$, it follows easily that φ belongs to $\Pi_n^{\mathcal{P}}(T)$ on the basis of (E, H). Contradiction.

2nd Case: there exists least $m > n$ such that $\Sigma_m^{\mathcal{P}}(T) = \Sigma_n^{\mathcal{P}}(T)$ for all $T \in \mathcal{T}$, and $\Pi_m^{\mathcal{P}}(T) \neq \Sigma_m^{\mathcal{P}}(T)$ for some $T \in \mathcal{T}$. Choose $T \in \mathcal{T}$, $\varphi \in \mathcal{L}$, finite $E \subseteq T$, and finite $H \subseteq \Sigma_m^{\mathcal{P}}(T)$ such that φ belongs to $\Pi_m^{\mathcal{P}}(T)$ on the basis of (E, H) and $\varphi \notin \Sigma_m^{\mathcal{P}}(T)$. Let $T' \in \mathcal{T}$ be such that $E \subseteq T'$, $T' \models_{\mathcal{W}} H$, and $T' \not\models_{\mathcal{W}} \varphi$. Then $\neg\varphi \in \Sigma_m^{\mathcal{P}}(T')$, and it follows from the choice of m that $\neg\varphi \in \Sigma_n^{\mathcal{P}}(T')$. Since $H \subseteq \Sigma_n^{\mathcal{P}}(T)$, we infer that φ belongs to $\Pi_n^{\mathcal{P}}(T)$ on the basis of (E, H). Contradiction. $\square$

Proof of Proposition 15: If $X \subseteq \Sigma_n^{\mathcal{P}}(T)$ then it follows immediately from Proposition 4 that $\bigvee X$ belongs to $\Sigma_n^{\mathcal{P}}(T)$. Suppose that all members of X are n-refutable in $\mathcal{P}$. Let $T \in \mathcal{T}$ be such that $T \not\models_{\mathcal{W}} X$. Choose $\varphi \in X$ with $T \not\models_{\mathcal{W}} \varphi$. Since $\mathcal{P}$ is perfectly observational, $T \models_{\mathcal{W}} \neg\varphi$. By hypothesis, φ is n-refutable, hence $\neg\varphi \in \Sigma_n^{\mathcal{P}}(T)$, which implies with Proposition 4 again that $\neg \bigwedge X \in \Sigma_n^{\mathcal{P}}(T)$. We conclude that $\bigwedge X$ is n-refutable in $\mathcal{P}$. $\square$

Proof of Proposition 17: Let $\mathcal{O}$ denote the set of atomic sentences and their negations. Proof is by induction. Let $n \in \mathbb{N}$ and $\mathfrak{M} \in \mathcal{W}$ be given. In case $n > 0$ suppose that for all $\psi \in \mathcal{L}$, if ψ is Π_n and $\mathfrak{M} \models \psi$ then ψ

is Π_n in $\mathcal{P}$. Assume that φ is Σ_{n+1}. Let formula ψ be such that ψ is in Π_n prenex form and φ is logically equivalent in $\mathcal{W}$ to the existential closure of ψ. Suppose that $\mathfrak{M} \models \varphi$. Since $\mathfrak{M}$ is Henkin, there exists a closed instance ψ' of ψ such that $\mathfrak{M} \models \psi'$. First assume that $n = 0$. Let E be the set of atomic formulas occurring in ψ'. Let D be the set of formulas χ such that $\mathfrak{M} \models \chi$, and either χ belongs to E or χ is the negation of a member of E. Since $D \models \varphi$ and $D \subseteq D_{\mathcal{O}}(\mathfrak{M})$ by the assumption on $\mathcal{P}$, it follows from Proposition 4 that $\varphi \in \Sigma_1^{\mathcal{P}}(D_{\mathcal{O}}(\mathfrak{M}))$. If $n > 0$, it follows from the inductive hypothesis that $\psi' \in \Pi_n^{\mathcal{P}}(D_{\mathcal{O}}(\mathfrak{M}))$, which implies again with Proposition 4 that $\varphi \in \Sigma_{n+1}^{\mathcal{P}}(D_{\mathcal{O}}(\mathfrak{M}))$. The proof is similar if φ is logically equivalent in $\mathcal{W}$ to $\bigvee X$ for some non-empty set X of sentences that are in Σ_n prenex form.

Now assume that φ is Π_{n+1}. Let formula ψ be such that ψ is in Σ_n prenex form and φ is logically equivalent in $\mathcal{W}$ to the universal closure of ψ. Suppose that $\mathfrak{M} \not\models \varphi$. Since $\mathfrak{M}$ is Henkin, there exists a closed instance ψ' of ψ such that $\mathfrak{M} \not\models \psi'$. Since $\neg\psi'$ is Σ_{n+1}, it follows from the former that $\neg\psi' \in \Sigma_{n+1}^{\mathcal{P}}(D_{\mathcal{O}}(\mathfrak{M}))$, which implies with Proposition 4 that $\neg\varphi \in \Sigma_{n+1}^{\mathcal{P}}(D_{\mathcal{O}}(\mathfrak{M}))$. Hence φ is $(n+1)$-refutable. The proof is similar if φ logically equivalent in $\mathcal{W}$ to $\bigwedge X$ for some non-empty set X of sentences that are in Π_n prenex form. $\square$

Proof of Proposition 19: The concatenation of finite sequences is denoted by $\star$ and the set of finite sequences of members of $\mathcal{L}$ by $\mathcal{L}^{<\mathbb{N}}$. The proof is trivial if $X = \emptyset$, so suppose otherwise. Fix an enumeration $(\varphi_i)_{i\in\mathbb{N}}$ of X in which every member of X occurs infinitely many times. Let learner f be such that f identifies all members of X in the limit in $\mathcal{P}$. We define $h : \mathcal{L}^{<\mathbb{N}} \to \mathbb{N}$ inductively, as follows. Set $h(()) = 0$. Let $\sigma \in \mathcal{L}^{<\mathbb{N}}$ and $x \in \mathcal{L}$ be given, and suppose that $h(\sigma)$ has been defined. If $\varphi_{h(\sigma)}$ belongs to $f(\sigma \star x)$ then $h(\sigma \star x) = h(\sigma)$; otherwise $h(\sigma \star x) = h(\sigma) + 1$. Let g denote the (unique) learner such that for all non-empty $\sigma \in \mathcal{L}^{<\mathbb{N}}$, $g(\sigma) = \{\varphi_{h(\sigma)}\}$. We show that g is X-reliable in $\mathcal{P}$. Let $T \in \mathcal{T}$ and environment e for T be given. Let φ be the (unique) member of X such that $T \models_{\mathcal{W}} \varphi$. Since f identifies φ in the limit in $\mathcal{P}$, we can choose $n \in \mathbb{N}$ such that $\varphi \in f(e[k])$ for all $k \geq n$. Set $h(e[n]) = q$. If $\varphi_q = \varphi$ then it is clear that $h(e[k]) = q$ for all $k \geq n$, hence $g(e[k]) = \{\varphi\}$ for all $k \geq n$, so suppose otherwise. Let $p > q$ be least with $\varphi_p = \varphi$. Since f identifies each member of $\{\varphi_q, \ldots, \varphi_{p-1}, \varphi_p\}$ in the limit in $\mathcal{P}$, there exits $n_{q+1}, \ldots, n_{p-1}, n_p \in \mathbb{N}$ with $n < n_{q+1} < \ldots < n_{p-1} < n_p$ such that:

- $h(e[n]) = \ldots = h(e[n_{q+1} - 1])$,
- $\ldots$
- $h(e[n_{p-1} - 1]) + 1 = h(e[n_{p-1}]) = \ldots = h(e[n_p - 1])$,
- $h(e[n_p - 1]) + 1 = h(e[n_p]) = p$.

It is clear that for all $k \geq n_p$, $h(e[k]) = p$, hence $g(e[k]) = \{\varphi\}$. So g is X-reliable in $\mathcal{P}$, as wanted. $\square$

Proof of Proposition 20: Suppose that φ is Σ_1 in $\mathcal{P}$. Let a learner f be defined as follows. Let $\sigma \in \mathcal{L}^{<\mathbb{N}}$ be given. If there exists $T \in \mathcal{T}$ and finite $D \subseteq T$ such that φ belongs to $\Sigma_1^{\mathcal{P}}(T)$ on the basis of D and $D \subseteq \mathrm{rng}(\sigma)$, then $f(\sigma) = \{\varphi\}$; otherwise $f(\sigma)$ is undefined. It follows immediately from Property 3 that f identifies φ in the limit in $\mathcal{P}$ with no mind changes.

Suppose that some learner learns φ in the limit in $\mathcal{P}$ with no mind changes. Let $T \in \mathcal{T}$ be such that $T \models_{\mathcal{W}} \varphi$, and let e be an environment for T. Let $k \in \mathbb{N}$ be such that $f(e[k])$ is defined, and let D be the set of formulas that occur in $e[k]$. It is easy to verify that for all $T' \in \mathcal{T}$, $D \subseteq T'$ implies $T' \models_{\mathcal{W}} \varphi$. Hence by Property 3 again, φ is Σ_1 in $\mathcal{P}$. $\square$

Proof of Proposition 21: Let learner f be X-reliable in $\mathcal{P}$ with no mind changes. Let $T \in \mathcal{T}$ be given. Choose $\sigma \in T^{<\mathbb{N}}$ such that $f(\sigma) \cap X = \mathrm{Cn}_{\mathcal{W}}(T) \cap X$. Suppose there exists $T' \in \mathcal{T}$ such that $\mathrm{rng}(\sigma) \subseteq T'$ and $\mathrm{Cn}_{\mathcal{W}}(T') \cap X \neq \mathrm{Cn}_{\mathcal{W}}(T) \cap X$. Choose an environment e for T' that begins with σ. Since f is X-reliable in $\mathcal{P}$, there exists k greater than the length of σ such that $f(e[k]) \cap X = \mathrm{Cn}_{\mathcal{W}}(T') \cap X$. Hence $f(\sigma) \cap X \neq f(e[k]) \cap X$, which contradicts the hypothesis that f is X-reliable in $\mathcal{P}$ with no mind changes.

Conversely, suppose that clause 2. holds. Let Z be the set of all finite subsets D of $\mathcal{L}$ for which there exits $Y \subseteq X$ such that for all $T \in \mathcal{T}$, $D \subseteq T$ implies $\mathrm{Cn}_{\mathcal{W}}(T) \cap X = Y$. Given $D \in Z$, denote by X_D the unique subset of X such that for all $T \in \mathcal{T}$, $D \subseteq T$ implies $\mathrm{Cn}_{\mathcal{W}}(T) \cap X = X_D$. Let scientist f be defined as follows. Let $\sigma \in \mathcal{L}^{<\mathbb{N}}$ be given. If there is $D \in Z$ such that $D \subseteq \mathrm{rng}(\sigma)$ then $f(\sigma) = X_D$; otherwise $f(\sigma)$ is undefined. It is easy to verify that f is well defined and is X-reliable in $\mathcal{P}$ with no mind changes. $\square$

Proof of Proposition 22: Fix an enumeration $((E_i, H_i))_{i>0}$ of all pairs of form (E, H), where E and H are finite sets of sentences and for all $T \in \mathcal{T}$, if $E \subseteq T$ and $H \subseteq \Pi_1^{\mathcal{P}}(T)$, then φ belongs to $\Sigma_2^{\mathcal{P}}(T)$ on the basis of (E, H). We define a partial function $g : \mathcal{L}^{<\mathbb{N}} \to \mathbb{N}$ as follows. Let $\sigma \in \mathcal{L}^{<\mathbb{N}}$ be given. Suppose there exists least $i > 0$ such that for some $T \in \mathcal{T}$:

1. $E_i \subseteq \mathrm{rng}(\sigma) \subseteq T$, and
2. all members of H_i belong to $\Pi_1^{\mathcal{P}}(T)$ on the basis of E_i.

Then $g(\sigma) = i$. If there is no such i then $g(\sigma) = 0$. Now we define learner f as follows. $f(()) = \emptyset$. Let $\sigma, \tau \in \mathcal{L}^{<\mathbb{N}}$ be such that τ extends σ and the length of τ is equal to 1 plus the length of σ. If $g(\tau) = g(\sigma) \neq 0$ then $f(\tau) = \{\varphi\}$; otherwise $f(\tau) = \emptyset$. We show that f identifies φ in the limit in $\mathcal{P}$.

Let $T \in \mathcal{T}$ and environment e for T be given. It is easy to verify that if $T \models_{\mathcal{W}} \varphi$ then $f(e[k]) = \{\varphi\}$ for cofinitely many $k \in \mathbb{N}$. Suppose that

$T \not\models_{\mathcal{W}} \varphi$. To complete the proof it suffices to show that for all $i > 0$, there exists $k \in \mathbb{N}$ such that for all $k' \geq k$ and $T' \in \mathfrak{T}$, either $E_i \not\subseteq \mathrm{rng}(e[k'])$, or $\mathrm{rng}(e[k']) \not\subseteq T'$, or some member of H_i does not belong to $\Pi_1^{\mathfrak{P}}(T')$ on the basis of E_i. So let $i > 0$ be given. Suppose that $E_i \subseteq T$ and there exists $T' \in \mathfrak{T}$ such that all members of H_i belong to $\Pi_1^{\mathfrak{P}}(T')$ on the basis of E_i (otherwise, we are done). Then $T \not\models_{\mathcal{W}} H_i$ (since otherwise, φ would belong to $\Sigma_2^{\mathfrak{P}}(T)$ on the basis of (E_i, H_i)). Hence we can choose $\psi \in H_i$ and finite $E \subseteq \mathfrak{L}$ such that $\neg\psi$ belongs to $\Sigma_1^{\mathfrak{P}}(T)$ on the basis of E. Let $k \in \mathbb{N}$ be such that $E \cup E_i \subseteq \mathrm{rng}(e[k])$. Let $k' \geq k$ and $T' \in \mathfrak{T}$ be such that $\mathrm{rng}(e[k']) \subseteq T'$. Then $\neg\psi$ belongs to $\Sigma_1^{\mathfrak{P}}(T')$ on the basis of E. So for all $k' \geq k$ and $T' \in \mathfrak{T}$, either $\mathrm{rng}(e[k']) \not\subseteq T'$ or $T' \not\models_{\mathcal{W}} H_i$. This implies trivially what we wanted to show. □

References

Barwise, J.: 1975, *Admissible Sets and Structures*. Berlin: Springer-Verlag.

Carnap, R.: 1950, *Logical Foundations of Probability (2nd Edition)*. Chicago, IL: University of Chicago Press.

Earman, J.: 1992, *Bayes or Bust?* Cambridge MA: MIT Press.

Ebbinghaus, H. D., J. Flum, and W. Thomas: 1994, *Mathematical Logic, Second Edition*. Berlin: Springer-Verlag.

Etchemendy, J.: 1999, *The Concept of Logical Consequence*. Stanford: CSLI Press.

Glymour, C.: 1985, 'Inductive Inference in the Limit'. *Erkenntnis* **22**, 23–31.

Gold, E. M.: 1967, 'Language identification in the limit'. *Information and Control* **10**, 447–474.

Goodman, N.: 1954, *Fact, Fiction and Forecast*. London: Athlone Press.

Hacking, I.: 1968, 'Reply to Salmon's 'The Justification of Inductive Rules of Inference''. In: I. Lakatos (ed.): *The Problem of Inductive Logic*. North-Holland.

Hintikka, J. and P. Suppes: 1966, *Aspects of Inductive Logic*. Amsterdam: North-Holland.

Howson, C. and P. Urbach: 1989, *Scientific Reasoning: The Bayesian Approach*. La Salle: Open Court.

Jain, S., D. Osherson, J. Royer, and A. Sharma: 1999, *Systems that Learn, 2nd Edition*. Cambridge MA: MIT Press.

Jeffrey, R.: 1980, *Studies in Inductive Logic and Probability*. Berkeley: University of California Press.

Kaye, R.: 1991, *Models of Peano Arithmetic*. Oxford University Press.

Kelly, K. T.: 1996, *The Logic of Reliable Inquiry*. New York, NY: Oxford University Press.

Klenk, V.: 1976, 'Intended Models and the Löwenheim-Skolem theorem'. *Journal of Philosophical Logic* **5**, 475–489.

Kyburg, H.: 1970, *Probability and Inductive Logic*. Macmillan Company.

Martin, E. and D. N. Osherson: 1998, *Elements of Scientific Inquiry*. MIT Press.

Parson, C.: 1990, 'The Uniqueness of the Natural Numbers'. *Iyyun, A Jerusalem Philosophical Quarterly* **39**, 13–44.

Pitt, L.: 1989, 'Probabilistic inductive inference'. *Journal of the ACM* **36**(2), 383–433.

Popper, K.: 1959, *The Logic of Scientific Discovery*. Hutchinson.

Putnam, H.: 1983, 'Models and Reality'. In: P. Benacerraf and H. Putnam (eds.): *Philosophy of Mathematics*. Cambridge University Press, pp. 421–44.

Reichenbach, H.: 1949, *The Theory of Probability*. University of California Press.
Salmon, W.: 1968, 'The Justification of Inductive Rules of Inference'. In: I. Lakatos (ed.): *The Problem of Inductive Logic*. North-Holland.
Shapiro, S.: 1991, *Foundations without Foundationlism*. Oxford: Clarendon Press.
Shoenfield, J.: 1967, *Mathematical Logic*. Addison-Wesley.
von Wright, G.: 1957, *The Logical Problem of Induction*. Oxford: Basil Blackwell.