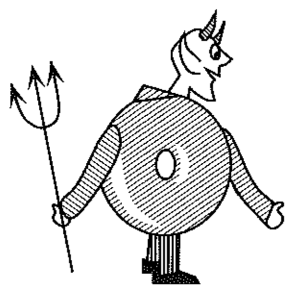
4

Topology and Ideal Hypothesis Assessment



1. Introduction

In this chapter, we move from the consideration of particular methods and problems to the characterization of problem solvability over entire paradigms. So while the issues treated in the preceding chapter were all located at level 3 of the taxonomy of questions presented at the end of chapter 2, the questions treated in this chapter belong to level 4. A characterization condition is a necessary and sufficient condition for the existence of a reliable method, given entirely in terms of the structures of $\mathcal{K}$, $\mathcal{C}$, and H. In other words, a characterization theorem isolates exactly the kind of background knowledge necessary and sufficient for scientific reliability, given the interpretation of the hypotheses and the sense of success demanded. To revive Kant's expression, such results may be thought of as *transcendental deductions* for reliable inductive inference, since they show what sort of knowledge is necessary if reliable inductive inference is to be possible.

To state a characterization condition, one must describe the structure of a set of possible data streams without reference to scientists or to reliability. We therefore require some framework in which to state such descriptions. It turns out that topology provides just the right concepts.

2. Basic Topological Concepts

A *topological space* is an ordered pair $\mathfrak{T} = (\mathcal{T}, T)$, such that $\mathcal{T}$ is an arbitrary set, and T is some collection of subsets of $\mathcal{T}$ with the following, simple

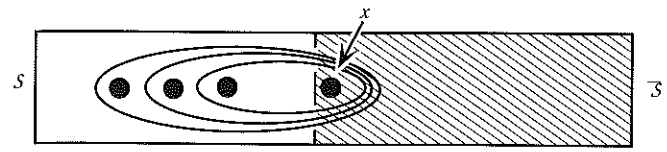


Figure 4.1

properties:

T is closed under arbitrary unions,

T is closed under finite intersections, and

$\mathcal{T}, \varnothing \in T$.

The elements of T are known as the *open sets* of $\mathfrak{T}$. If $x \in \mathcal{T}$, then $\mathcal{S} \subseteq \mathcal{T}$ is a *neighborhood* of x just in case $x \in \mathcal{S}$ and $\mathcal{S} \in T$. Claims about all neighborhoods of some point may be viewed, intuitively, as claims about what happens arbitrarily close to this point. All other topological concepts are defined in terms of the open sets of $\mathfrak{T}$, and hence are all relative to $\mathfrak{T}$, but explicit reference to $\mathfrak{T}$ will be dropped for brevity. A *closed set* is just the complement of an open set. A *clopen set* is both closed and open.

To simplify notation, when one speaks of subsets of $\mathfrak{T}$, one means subsets of $\mathcal{T}$. Of special importance to methodology is the notion of *limit points* of subsets of $\mathfrak{T}$. A limit point of a set $\mathcal{S}$ is an object $x \in \mathcal{T}$ that is so close to $\mathcal{S}$ that no open set (neighborhood) around x can fail to catch a point of $\mathcal{S}$ (*Fig. 4.1*).

x is a limit point of $\mathcal{S} \Leftrightarrow$ for each open O, if $x \in O$ then $O \cap \mathcal{S} \neq \varnothing$.

The *closure* of $\mathcal{S}$ is the result of adding all the limit points of $\mathcal{S}$ to $\mathcal{S}$. The closure of $\mathcal{S}$ will be denoted $cl(\mathcal{S})$. The following facts relating closed sets to the closure operation are among the most basic in topology. Since they will also turn out to be fundamental to the study of logical reliability, it is worth reviewing them here.

Proposition 4.1

(a) *$\mathcal{S}$ is closed $\Leftrightarrow \mathcal{S} = cl(\mathcal{S})$.*

(b) *$cl(\mathcal{S})$ is the least closed superset of $\mathcal{S}$.*

Proof: (a) ($\Rightarrow$) Suppose $\mathcal{S}$ is closed. Then $\bar{\mathcal{S}}$ is open. Suppose for reductio that some limit point x of $\mathcal{S}$ is missing from $\mathcal{S}$. Then $\bar{\mathcal{S}}$ is an open set containing x that is disjoint from $\mathcal{S}$, so x is not a limit point of $\mathcal{S}$ (*Fig. 4.2*). Contradiction. Thus $x \in \mathcal{S}$.

($\Leftarrow$) Suppose $\mathcal{S}$ contains all its limit points. Then for each $x \notin \mathcal{S}$, x is not

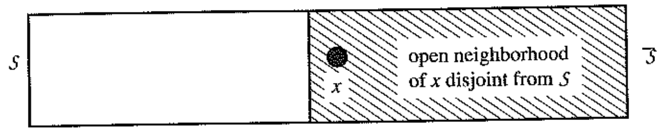


Figure 4.2

a limit point of $\mathcal{S}$. Thus for each $x \notin \mathcal{S}$, there is some open set $\mathcal{R}_x$ disjoint from $\mathcal{S}$ such that $x \in \mathcal{R}_x$. $\bar{\mathcal{S}}$ is identical to the union of all these sets $\mathcal{R}_x$ (*Fig. 4.3*). But since each $\mathcal{R}_x$ is open, so is the union. Thus, $\bar{\mathcal{S}}$ is open and hence $\mathcal{S}$ is closed.

(b) From (a) it is clear that $cl(\mathcal{S})$ is closed. Suppose for reductio that there exists some closed $\mathcal{R}$ such that $\mathcal{S} \subseteq \mathcal{R} \subset cl(\mathcal{S})$. Then some $x \in cl(\mathcal{S})$ is missing from $\mathcal{R}$ and hence is also missing from $\mathcal{S}$. Since x is a limit point of $\mathcal{S}$, there is no open O containing x that fails to contain some element of $\mathcal{S}$. But since $\mathcal{S} \subseteq \mathcal{R}$, there is no open O containing x that fails to contain some element of $\mathcal{R}$. Thus x is a limit point of $\mathcal{R}$ as well, so $\mathcal{R}$ is missing one of its own limit points. So by (a), $\mathcal{R}$ is not closed. Contradiction. ■

An *interior point* of $\mathcal{S}$ is a point that is a member of an open subset of $\mathcal{S}$ (*Fig. 4.4*). The *interior* of $\mathcal{S}$, denoted $int(\mathcal{S})$, is the set of all interior points of $\mathcal{S}$. The interior of $\mathcal{S}$ is dual to the closure of $\mathcal{S}$ is the sense that while the latter is the least closed superset of $\mathcal{S}$, the former is the greatest open subset of $\mathcal{S}$.

Proposition 4.2

$int(\mathcal{S})$ is the greatest open set contained in $\mathcal{S}$.

Proof: Each $x \in int(\mathcal{S})$ is contained in an open subset of $\mathcal{S}$. The union of these open subsets is open and is identical to $int(\mathcal{S})$. Suppose there is an open $\mathcal{R}$ such that $int(\mathcal{S}) \subset \mathcal{R} \subseteq \mathcal{S}$. Then there is some $x \in \mathcal{R} - int(\mathcal{S})$. But since $\mathcal{R}$ is open and $x \in \mathcal{S}$, it follows that $x \in int(\mathcal{S})$. Contradiction. ■

A *boundary point* of $\mathcal{S}$ is a point that is a limit point both of $\mathcal{S}$ and of $\bar{\mathcal{S}}$. The *boundary* of $\mathcal{S}$, denoted $bdry(\mathcal{S})$, is defined as the set of all boundary points of $\mathcal{S}$, which is evidently the same as $cl(\mathcal{S}) \cap cl(\bar{\mathcal{S}})$.

The boundary of $\mathcal{S}$ is the "residue" lying between the best closed

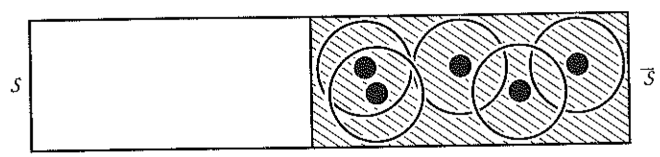


Figure 4.3

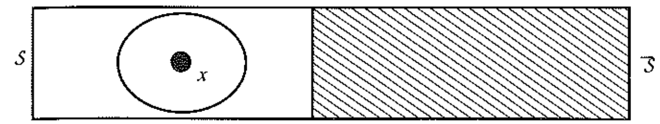


Figure 4.4

approximation of $\mathcal{S}$ from without and the best open approximation of $\mathcal{S}$ from within (*Fig. 4.5*).

Proposition 4.3

$bdry(\mathcal{S}) = cl(\mathcal{S}) - int(\mathcal{S})$.

Proof: ($\subseteq$) By definition, $bdry(\mathcal{S}) \subseteq cl(\mathcal{S})$. Suppose for reductio that some $x \in bdry(\mathcal{S})$ is also in $int(\mathcal{S})$. Then there is some open subset of $\mathcal{S}$ that includes x, so x is not a limit point of $\bar{\mathcal{S}}$, and hence is not in $bdry(\mathcal{S})$. Contradiction.

($\supseteq$) Suppose that $x \in cl(\mathcal{S}) - int(\mathcal{S})$. Then x is a limit point of $\mathcal{S}$. Suppose for reductio that x is not a limit point of $\bar{\mathcal{S}}$. Then for some open $Q \subseteq \mathcal{S}$, $x \in Q$. So $x \in int(\mathcal{S})$. Contradiction. Hence, x is a limit point of $\mathcal{S}$, so $x \in bdry(\mathcal{S})$. ■

It is therefore evident that a set is clopen just in case its boundary is empty; i.e., just in case it can be perfectly approximated both by a closed and by an open set.

$\mathcal{S}$ is *dense* in $\mathcal{R}$ just in case $\mathcal{R} \subseteq cl(\mathcal{S})$.[1] Thus, $\mathcal{S}$ is arbitrarily close to everything in $\mathcal{R}$ and may be thought of as spread throughout $\mathcal{R}$.

A *basis* for $\mathfrak{T}$ is a collection $\mathcal{B}$ of open sets of $\mathfrak{T}$ such that every open set of $\mathfrak{T}$ is a union of elements of $\mathcal{B}$. It will turn out that the topological concept of basis is closely related to the methodological concept of a finite data sequence.

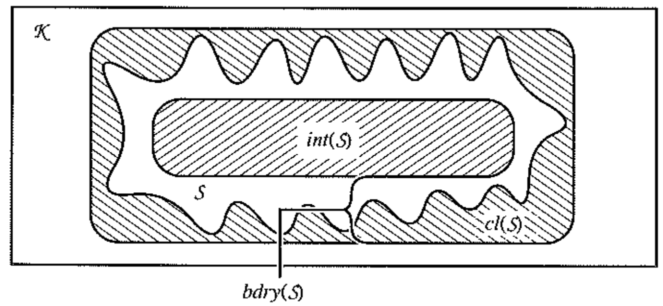


Figure 4.5

[1] This definition is nonstandard. The usual definition requires $\mathcal{R} = cl(\mathcal{S})$. But this clearly excludes cases in which $\mathcal{R}$ is not closed. The intuition of being "spread throughout" $\mathcal{R}$ does not require that $\mathcal{R}$ be closed, however. It suggests only that everything in $\mathcal{R}$ be arbitrarily close to something in $\mathcal{S}$, which accords with the condition $\mathcal{R} \subseteq cl(\mathcal{S})$.

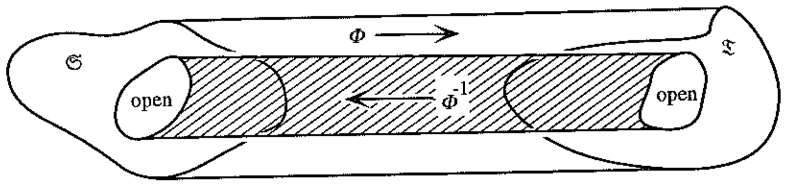


Figure 4.6

Let $\mathfrak{S}$ and $\mathfrak{T}$ be two topological spaces, and let Φ be a function from $\mathfrak{S}$ to $\mathfrak{T}$. Then Φ is *continuous* if and only if for each open subset S of $\mathfrak{T}$, $\Phi^{-1}(S)$ is open in $\mathfrak{S}$ (*Fig. 4.6*).

Topology was originally conceived as a kind of generalized geometry, in which the geometrical equivalence relation of *congruence* (one figure coincides with the other when laid on top of it) is replaced with the topological equivalence relation of *homeomorphism* (continuous, 1–1 deformability of one figure into another). A triangle and a square are homeomorphic in the Euclidean plane because each can be stretched or beaten into the other without cutting, tearing, or gluing. That is, regions that are arbitrarily close in the square remain arbitrarily close in the triangle (*Fig. 4.7*). On the other hand, a square cannot be stretched and beaten into a figure eight without cutting or gluing. Gluing brings into arbitrary proximity regions that were once apart, violating continuity (*Fig. 4.8*).

The apparatus of open sets provides a very general setting for the study of continuity. It is a remarkable fact that this abstract treatment of continuity also captures the local perspective of ideal inductive methods that distinguishes them from perfectly clairvoyant deities who can see the entire future all at once. Just as the conjecture sequence of a scientific method is determined by its conjectures on finite approximations of the infinite data stream, the value of a continuous function at a point in a topological space is determined by its values over open sets around the point. In the next section we examine a standard topological space in which points correspond to infinite data streams, sets of points become empirical hypotheses, elements of a countable basis correspond to finite data sequences, and continuity reflects the bounded perspective of the scientist.

3. The Baire Space

The *Baire space* $\mathfrak{N}$ is a topological space of special relevance to analysis, logic, and the study of inductive inference.[2] The Baire space is usually defined in

Figure 4.7

[2] The learning-theoretic relevance of continuity in the Baire space is also discussed in Osherson et al. (1986).

Figure 4.8

terms of a basis for its open sets. Let e be a finite sequence of natural numbers. Think of e as a finite data sequence, where data are encoded as natural numbers. Now consider the set $[e]$ of all infinite data streams in $\mathcal{N}$ that extend e. Thus, $[e] = \{\varepsilon: \varepsilon|lh(e) = e\}$. I refer to $[e]$ as the *fan* with *handle* e (*Fig. 4.9*).

We may think of the fan with handle e as representing the empirical uncertainty of a scientist who has just seen e, but whose background assumptions are vacuous (so that $\mathcal{K} = \mathcal{N}$). For all he knows, the actual data sequence may be any infinite extension of e. Let e, e' be finite data sequences. We say $e \subseteq e'$ just in case e' *extends* e or is identical to e. For any two fans, either one includes the other, or the two are disjoint. That is, $[e'] \subseteq [e] \Leftrightarrow e \subseteq e'$.

If one fan includes the other, then their intersection and union are also fans. Otherwise, the intersection is empty and the union may or may not be a fan (*Fig. 4.10*). The whole space $\mathcal{N}$ is the fan whose handle is the empty sequence $\mathbf{0}$. The empty set is not a fan. The intersection of a countable collection of fans is nonempty just in case there is some infinite data stream ε that extends the handles of all fans in the intersection (*Fig. 4.11*). If there is no bound on the lengths of the handles involved, then the intersection is $\varnothing$ or exactly $\{\varepsilon\}$. Otherwise the intersection is either empty or identical to the fan in the intersection with the longest handle.

The Baire space is just the space $\mathfrak{N} = (\mathcal{N}, B)$, where B is the collection of all sets formed by taking an arbitrary union of fans. Recall that ω^* denotes the set of all finite sequences of natural numbers. Then B is defined as follows:

$$\mathcal{P} \in B \Leftrightarrow \textit{there is a } G \subseteq \omega^* \textit{ such that } \mathcal{P} = \bigcup\{[e]: e \in G\}.$$

To see that $\mathfrak{N}$ is indeed a topological space, observe that $\mathcal{N} = [\mathbf{0}]$ where $\mathbf{0}$ is the empty data sequence, so the whole space is open since each fan is. $\varnothing$ is also open, since it is the trivial union $\bigcup\varnothing$ of the empty set of fans. Unions of open sets are trivially open since open sets are arbitrary unions of fans. Finally, consider a finite intersection of open sets. We have just seen that a finite

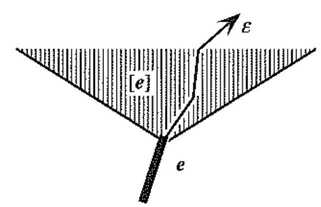


Figure 4.9

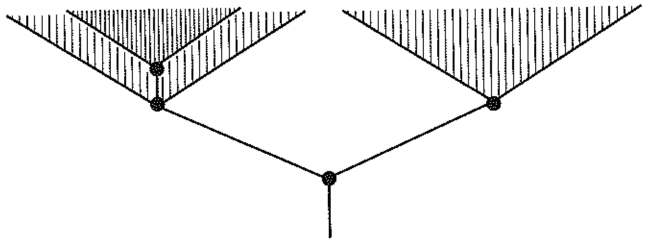


Figure 4.10

intersection of fans is either empty (and hence open) or a fan. Thus a finite intersection of unions of fans is either a union of fans or the empty set, both of which are open.

Since the fans constitute a basis for $\mathfrak{N}$, the open sets of $\mathfrak{N}$ are arbitrary unions of fans. Since there are only countably many distinct fans, all such unions are countable. An open set $\mathcal{S}$ may be thought of as a hypothesis verifiable with certainty through observation or, rather, as the set of all data streams for which such a hypothesis is correct. As soon as some handle of some fan in $\mathcal{S}$ is seen, the scientist knows that $\mathcal{S}$ is correct, since every extension of the handle is a data stream in $\mathcal{S}$. An example of such a hypothesis is "you will eventually see a 1." Once a 1 is seen, it doesn't matter what further data occurs. That is, once a 1 is seen by the scientist, he has entered a fan included in the open set.

The closed sets of $\mathfrak{N}$ are just the complements of open sets. Hence, a closed set corresponds to a hypothesis that can be refuted with certainty by finite observation if it is false. That is, the complement of a closed set is verified when one of its handles is seen in the data, and hence the closed set is refuted. Closed set correspond therefore to universal hypotheses such as "all ravens are black." Once a white raven is observed, the hypothesis is known with certainty to be false, no matter what further evidence is observed. A *clopen* set of $\mathfrak{N}$ is a set that is decidable with certainty.

Limit points in the Baire space also have an important methodological interpretation.

Proposition 4.4

ε is a limit point of $\mathcal{S}$ in $\mathfrak{N}$
$\Leftrightarrow$ for each n there is an $\varepsilon' \in \mathcal{S}$ such that $\varepsilon|n = \varepsilon'|n$.

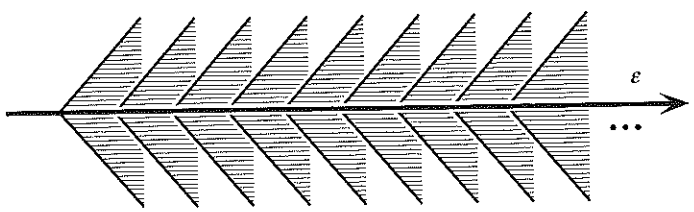


Figure 4.11

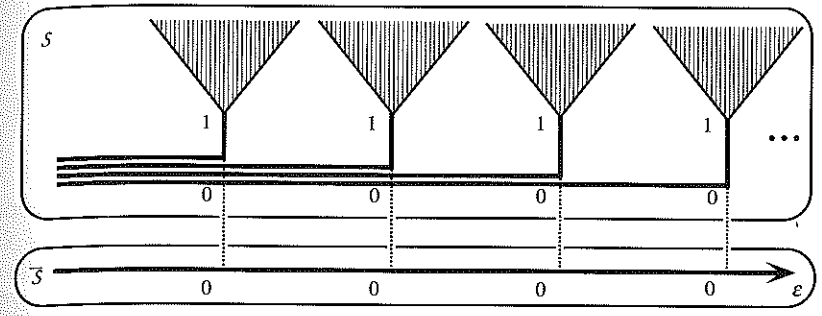


Figure 4.12

Proof: ($\Rightarrow$) Suppose ε is a limit point of $\mathcal{S}$ in $\mathcal{N}$. Then by definition, for each open O, if $\varepsilon \in O$ then $O \cap \mathcal{S} \neq \varnothing$. Then for each n, $[\varepsilon|n] \cap \mathcal{S} \neq \varnothing$, since $[\varepsilon|n]$ is an open set containing ε. So there is an $\varepsilon' \in \mathcal{S}$ such that $\varepsilon' \in [\varepsilon|n]$. Hence, $\varepsilon|n = \varepsilon'|n$. So for each n there is an $\varepsilon' \in \mathcal{S}$ such that $\varepsilon|n = \varepsilon'|n$.

($\Leftarrow$) Suppose that for each n there is an $\varepsilon' \in \mathcal{S}$ such that $\varepsilon|n = \varepsilon'|n$. Let $O \in B$. Suppose $\varepsilon \in O$. Then since O is a union of fans, there is an n such that $[\varepsilon|n] \subseteq O$. By assumption, there is $\varepsilon' \in \mathcal{S}$ such that $\varepsilon'|n = \varepsilon|n$. Hence, $\varepsilon' \in O$. So $O \cap \mathcal{S} \neq \varnothing$. Thus ε is a limit point of $\mathcal{S}$ in $\mathfrak{N}$. ■

A limit point of $\mathcal{S}$ in $\mathfrak{N}$ may be thought of as a data stream ε with data streams in $\mathcal{S}$ veering off infinitely often. *Figure 4.12* depicts the open set $\mathcal{S} = \bigcup\{[e]$: e is a finite string of 0s with a 1 at the end$\}$. $\mathcal{S}$ corresponds to the hypothesis that some non-0 will eventually be observed. $\bar{\mathcal{S}}$ is a closed set that corresponds to the universal hypothesis that 0 will always be observed. Since each initial segment of the everywhere 0 data stream is an initial segment of a data stream in $\mathcal{S}$, we have by proposition 4.4 that the everywhere 0 sequence is a limit point of $\mathcal{S}$. This is a case in which a limit point of $\mathcal{S}$ is missing from $\mathcal{S}$. It is also an instance of the problem of induction, for a demon can present data along ε until the scientist becomes sure that he is in $\bar{\mathcal{S}}$, after which the demon is free to veer back into $\mathcal{S}$. In fact, whenever a limit point ε of $\mathcal{S}$ is not in $\mathcal{S}$, the same argument can be given, in light of proposition 4.4. Therefore:

The problem of induction arises when either the hypothesis or its complement is not closed.

In other words, the problem of induction arises exactly in the *boundaries* of hypotheses.

In Figure 4.12, I did not overlap ε and the shared initial segments of data streams in $\mathcal{S}$ as in *Figure 4.13*. This is because the overlapped versions of such diagrams are dangerously misleading, since they depict *trees*, and tree structure does not uniquely determine the topological structure of the set so depicted. Let $G \subseteq \omega^*$ be a set of finite data sequences. G is a *tree* just in case each initial

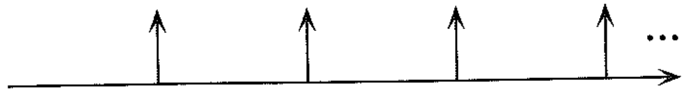


Figure 4.13

segment of a member of G is a member of G. The *tree generated by* $\mathcal{S}$ (denoted $Tree(\mathcal{S})$) is defined as the set of all finite, initial segments of elements of $\mathcal{S}$. $\mathcal{S}$ is not always uniquely determined by $Tree(\mathcal{S})$. For example, let $\mathcal{S}$ be the set of all data streams in which 1 occurs somewhere. Then $Tree(\mathcal{S}) = Tree(\mathcal{N})$. In general, $Tree(\mathcal{S}) = Tree(\mathcal{R})$ if and only if $\mathcal{S}$ is dense in $\mathcal{R}$. This fact is important to keep in mind when relying on tree diagrams. Since such diagrams are finite, they cannot indicate whether limit points are missing or present, but missing limit points are exactly what local underdetermination and the problem of induction are about.

To fix intuitions, it is useful to consider some simple examples of open and closed sets in the Baire space. $\varnothing$ and $\mathcal{N}$ are both clopen. Each singleton $\{\varepsilon\}$ is closed, since it is the complement of an open set (namely, the union of the fans whose handles are not extended by ε). Each finite union of closed sets is closed, so each finite subset of $\mathfrak{N}$ is closed and each cofinite subset of $\mathfrak{N}$ is open. Since a set cannot be both finite and cofinite, we also have that finite sets are properly closed (i.e., closed but not open) and cofinite sets are properly open (i.e., open but not closed). Closed sets can also be the closures of infinite open sets, as when we add the missing limit point ε to the set $\mathcal{S}$ in the preceding example.

Say that an open set is *n-uniform* just in case it is a union of fans whose handles are all bounded in length by n, and say that an open set is *uniform* just in case it is n-uniform for some n. Each uniform open set $\mathcal{S}$ is clopen, since its complement consists of the union of all fans not included in $\mathcal{S}$ whose handles are of length n (*Fig. 4.14*).

Not all clopen sets are uniform. For example, consider the union of all fans with handles of form $n*\boldsymbol{e}$, where $\boldsymbol{e}$ is a string of 0s of length n. This set is a union of fans whose handles come in every length, but all the handles are distinct at position 1. Evidently, this set is not uniform, since the handles of the fans in the union come in all lengths. That it is open can be seen from the fact that

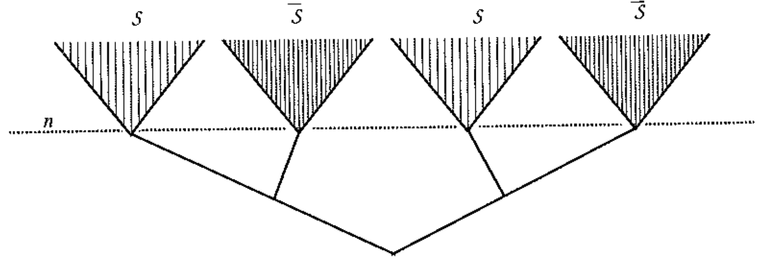


Figure 4.14

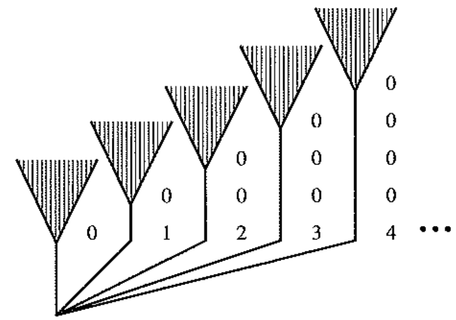


Figure 4.15

the set is a union of fans. That it is also closed can be seen from the fact that it is missing no limit point (*Fig. 4.15*). On the other hand, each clopen set is uniform if the space is finitely branching (cf. exercise 4.14).

4. Restricted Topological Spaces

Let $\mathfrak{T} = (\mathcal{T}, T)$ be a topological space, and let $\mathcal{K} \subseteq \mathcal{T}$. The *restriction* of $\mathfrak{T}$ to $\mathcal{K}$, denoted $\mathfrak{T}|\mathcal{K}$, is the space that results if we toss out of $\mathfrak{T}$ everything that is not in $\mathcal{K}$ while retaining essentially the same structure. The way to carry this out is to replace $\mathcal{T}$ with $\mathcal{T} \cap \mathcal{K}$ and to replace T with the set $T|\mathcal{K}$ of all intersections $O \cap \mathcal{K}$ such that $O \in T$. Accordingly, let $\mathfrak{T}|\mathcal{K} = (\mathcal{T} \cap \mathcal{K}, T|\mathcal{K})$. The elements of $T|\mathcal{K}$ are said to be *$\mathcal{K}$-open*. Sets of the form $\mathcal{K} - O$, where $O \in T|\mathcal{K}$, are said to be *$\mathcal{K}$-closed*. Sets that are both $\mathcal{K}$-open and $\mathcal{K}$-closed are *$\mathcal{K}$-clopen*. It is not too hard to show that $\mathfrak{T}|\mathcal{K}$ is itself a topological space, so all the results established previously apply to restricted spaces.

In the Baire space, restricting to some set $\mathcal{K}$ of data streams corresponds to adopting background assumptions about how the data will appear in the limit. The topological significance of background knowledge is to make sets simpler given the restriction than they are without it, and that is why local underdetermination depends crucially on what the scientist's background assumptions are.

5. A Characterization of Bounded Sample Decidability

It is now time to characterize the various notions of reliable success in topological terms. To do so, it is useful to let C_h denote the set of all data streams for which h is correct (according to C). We may think of C_h as the *empirical content* of h since C_h includes all the infinite data streams that might arise, given that h is correct (*Fig. 4.16*).

Bounded sample decidability is easy to characterize in terms of fans. Recall that $\mathcal{S} \subseteq \mathcal{N}$ is *n-uniform* just in case $\mathcal{S}$ is a union of fans whose handles are

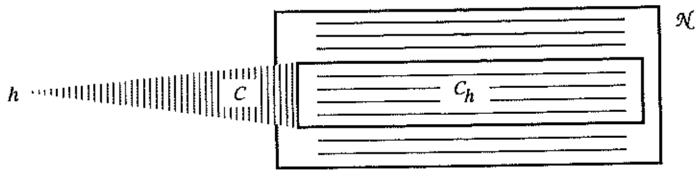


Figure 4.16

all bounded in length by n. Like all the other concepts, this one relativizes to background knowledge $\mathcal{K}$. Then it is simple to see that:

Proposition 4.5

H is decidable$_C$ by time n given $\mathcal{K}$ $\Leftrightarrow$ for each $h \in H$, $C_h \cap \mathcal{K}$ is $\mathcal{K}$-n-uniform.

Proof: ($\Leftarrow$) Suppose that $C_h \cap \mathcal{K}$ is $\mathcal{K}$-n-uniform. Thus, there is a $G_h \subseteq \omega^*$ such that each $e \in G_h$ is no longer than n and $C_h \cap \mathcal{K}$ is the union of all $[e] \cap \mathcal{K}$ such that $e \in G_h$. Define method α as follows (*Fig. 4.17*):

$$\alpha(h, e) = \begin{cases} 1 & \textit{if e extends some } e' \in G_h \\ 0 & \textit{otherwise.} \end{cases}$$

Let $\varepsilon \in \mathcal{K}$. Suppose $\varepsilon \in C_h$. Then for some $k \leq n$, $\varepsilon|k \in G_h$. Thus by time $k \leq n$, α stabilizes to 1 and hence conjectures 1 at time n. Suppose $\varepsilon \notin C_h$. Then $\varepsilon|n \notin G_h$, so $\alpha(h, \varepsilon|n) = 0$.

($\Rightarrow$) Suppose that for some $h \in H$, $\mathcal{K} \cap C_h$ is not $\mathcal{K}$-n-uniform. Then there are $\varepsilon, \varepsilon' \in \mathcal{K}$ such that h is correct for ε but not for ε' and $\varepsilon/n = \varepsilon'/n$. This is just what the demon needs to fool α (*Fig. 4.18*). For the demon merely checks the scientist's conjecture $\alpha(h, \varepsilon|n)$. If the conjecture is 1, the demon continues feeding with ε'. If it is anything else, the demon continues feeding ε. If α doesn't produce 1 or 0 at n, then α fails by default, and if α produces either

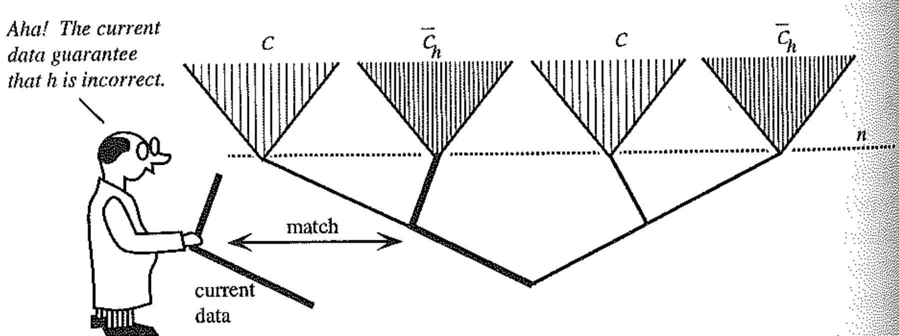


Figure 4.17

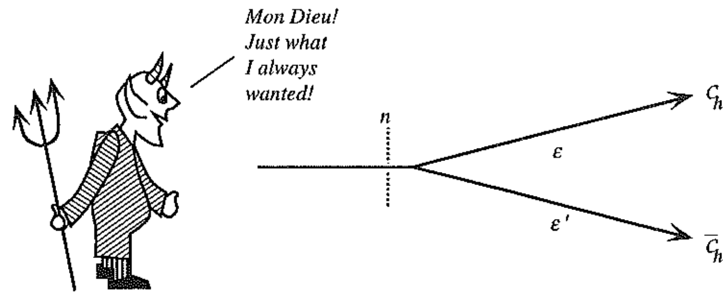


Figure 4.18

1 or 0 at n, then α is wrong. Since α is arbitrary, H is not decidable$_C$ by n given $\mathcal{K}$. ■

6. Characterizations of Certain Assessement

We have already seen that open and closed sets are like existential and universal hypotheses, respectively. It is therefore to be expected that open sets are verifiable with certainty and closed sets are refutable with certainty. The following proposition also shows that the converse is true: if a hypothesis is verifiable with certainty, it determines an open set and if a hypothesis is refutable with certainty, it determines a closed set.

Proposition 4.6

$$H \text{ is } \begin{bmatrix} \textit{verifiable}_C \\ \textit{refutable}_C \\ \textit{decidable}_C \end{bmatrix} \textit{ with certainty given } \mathcal{K}$$

$$\Leftrightarrow \textit{for each } h \in H,\ C_h \cap \mathcal{K} \textit{ is } \begin{bmatrix} \mathcal{K}\textit{-open} \\ \mathcal{K}\textit{-closed} \\ \mathcal{K}\textit{-clopen} \end{bmatrix}.$$

Proof: It suffices to show the verifiability case, since the refutation case follows by duality and the decision case follows immediately from the first two in light of proposition 3.3(b). ($\Leftarrow$) Suppose that $C_h \cap \mathcal{K}$ is $\mathcal{K}$-open. Then for some $G_h \subseteq \omega^*$, C_h is the union of all $[e] \cap \mathcal{K}$ such that $e \in G_h$. Let α conjecture 0 until some element of G_h is extended by e, at which time α conjectures the certainty mark '!' and continues producing 1s thereafter (*Fig. 4.19*).

$$\alpha(h, e) = \begin{cases} \text{'!'} & \textit{if } e \in G_h \\ 1 & \textit{if } e \textit{ properly extends some } e' \in G_h \\ 0 & \textit{otherwise.} \end{cases}$$

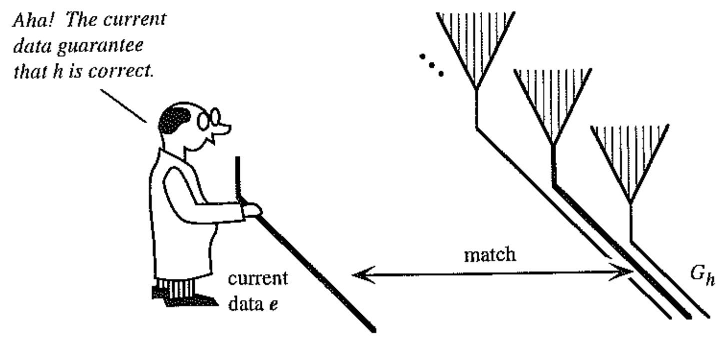


Figure 4.19

Let $\varepsilon \in \mathcal{K}$. Suppose that $\varepsilon \in C_h$. Then for some n, $\varepsilon|n \in G_h$. Thus, α produces 1 with certainty. Suppose $\varepsilon \notin C_h$. Then for no n is $\varepsilon|n$ in G_h. So α never produces 1 with certainty.

($\Rightarrow$) Suppose that for some $h \in H$, $C_h \cap \mathcal{K}$ is not $\mathcal{K}$-open. Thus $\overline{C_h} \cap \mathcal{K}$ is not $\mathcal{K}$-closed. So by proposition 4.1(a), $\overline{C_h}$ has a limit point ε in $\mathcal{K}$ that is not in $\overline{C_h} \cap \mathcal{K}$. But that is just the sort of situation that permits the demon to fool an arbitrary scientist regarding h given $\mathcal{K}$, as we saw in the preceding chapter (*Fig. 4.20*). The demon feeds ε until α produces its mark of certainty '!' followed by a 1. Otherwise he continues feeding ε forever. If α ever produces the mark of certainty followed by 1, then the demon presents a data stream for which h is incorrect$_C$, so α fails. Otherwise, the demon presents ε, for which h is correct, and α fails to produce 1 with certainty. Since α is arbitrary, no possible method can verify h (and hence H) with certainty given $\mathcal{K}$. ■

The demonic construction in the above proof is just Sextus' ancient argument for inductive skepticism, which we considered in the preceding two chapters. What we have just seen is that the argument arises for *every* hypothesis h such that C_h is not $\mathcal{K}$-open. It is interesting that this argument is an instance of the basic topological fact that each nonclosed set is missing one of its limit points (i.e., proposition 4.1).

I promised at the very end of the preceding chapter to apply topological concepts to the question whether some α that does not verify$_C$ H with certainty given $\mathcal{N}$ could nonetheless be an optimally reliable certain verifier for C, H.

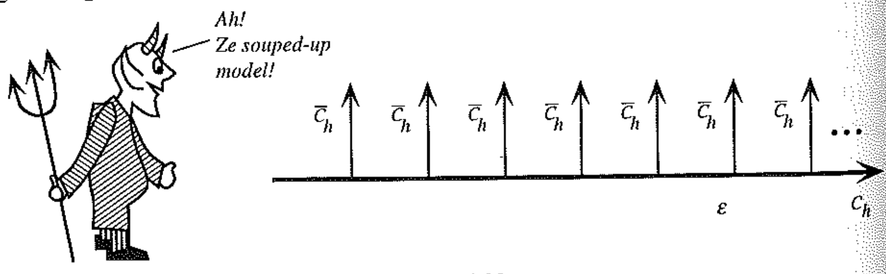


Figure 4.20

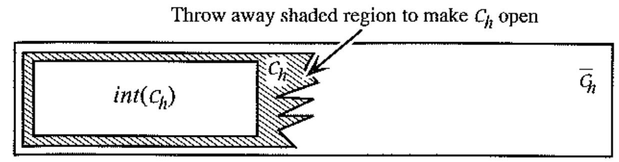


Figure 4.21

Now that we know that $\mathcal{K}$-openness characterizes verifiability with certainty given $\mathcal{K}$, the general resolution of the question is at hand.

Consider a single hypothesis in isolation. We will make use of the fact that the interior of a set is the *best* open approximation of the set from within. Then if we throw out of $\mathcal{N}$ every data stream in C_h that is not in the interior of C_h, we will have converted C_h into an open set in a *minimally destructive manner*, thereby arriving at minimal assumptions under which h is verifiable$_C$ with certainty (*Fig. 4.21*). In keeping with this strategy, define:

$$\mathcal{M}ax(C, h) = \mathcal{N} - (C_h - int(C_h)).$$

$C_h \cap \mathcal{M}ax(C, h)$ is $\mathcal{M}ax(C, h)$-open, since $C_h \cap \mathcal{M}ax(C, h) = int(C_h)$, so by proposition 4.6, there is some method α that verifies$_C$ h with certainty given $\mathcal{M}ax(C, h)$. By the following proposition, this method is optimal in terms of reliability.[3]

Proposition 4.7

h is verifiable$_C$ with certainty given $\mathcal{M}ax(C, h)$ but not given any proper superset of $\mathcal{M}ax(C, h)$.

Proof: Suppose $\mathcal{M}ax(C, h) \subset \mathcal{K}$. Then we may choose $\varepsilon \in (C_h \cap \mathcal{K}) - int(C_h)$. Since $\varepsilon \notin int(C_h)$, and $int(C_h)$ is the union of the set of all fans included in C_h, we have that ε is not an element of any fan included in C_h. Thus, for every n, $[\varepsilon|n]$ contains some element $\tau[n]$ of $\mathcal{N} - C_h$. But then $[\varepsilon|n]$ also contains some element of $\mathcal{M}ax(C, h) - C_h$ since $\overline{C_h} \subseteq \mathcal{M}ax(C, h)$ (*Fig. 4.22*).

Thus, ε is not included in any $\mathcal{K}$-fan included in $C_h \cap \mathcal{K}$, so $C_h \cap \mathcal{K}$ is not $\mathcal{K}$-open. So by proposition 4.6, h is not verifiable$_C$ with certainty given $\mathcal{K}$. ■

$\mathcal{M}ax(C, h)$ is also *uniquely* optimal, so we have a complete characterization of the optimal assumptions required for verification with certainty of an arbitrary hypothesis.

Proposition 4.8

If $\mathcal{K} \neq \mathcal{M}ax(C, h)$ and h is verifiable$_C$ with certainty given $\mathcal{K}$, then h is verifiable$_C$ with certainty given some proper superset of $\mathcal{K}$.

[3] For the refutation and decision with certainty cases, cf. exercise 4.5.

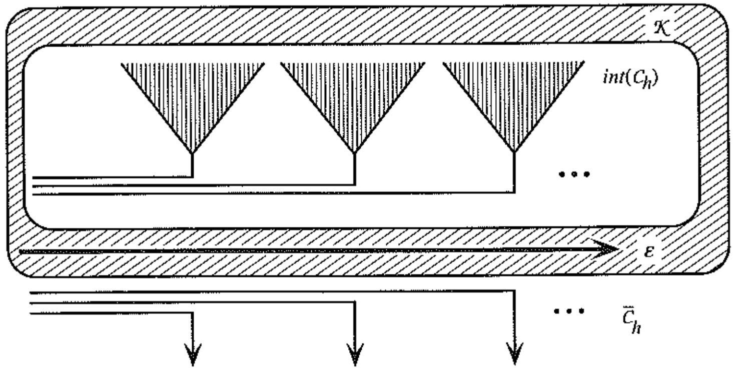


Figure 4.22

Proof: Suppose $\mathcal{K} \neq \mathcal{M}ax(\mathcal{C}, h)$. Then $\mathcal{K} \subset \mathcal{M}ax(\mathcal{C}, h)$ or there is some $\varepsilon \in \mathcal{K} - \mathcal{M}ax(\mathcal{C}, h)$. In the former case, we are done: h is verifiable$_\mathcal{C}$ with certainty given $\mathcal{M}ax(\mathcal{C}, h)$, by proposition 4.7. So let $\varepsilon \in \mathcal{K} - \mathcal{M}ax(\mathcal{C}, h)$. Then $\varepsilon \in \mathcal{C}_h$ and, as was shown in the preceding proof, $[\varepsilon|n]$ contains some element $\tau[n]$ of $\mathcal{N} - \mathcal{C}_h$, for each n. Since h is verifiable$_\mathcal{C}$ with certainty given $\mathcal{K}$, we have by proposition 4.6 that ε is contained in some $\mathcal{K}$-fan $[\varepsilon|n'] \subseteq \mathcal{C}_h \cap \mathcal{K}$. Let m be the least such n'. So for all $m' \geq m$, $\tau[m'] \notin \mathcal{K}$. Let $\mathcal{K}' = \mathcal{K} \cup \{\tau[m+1]\}$ (*Fig. 4.23*).

Let α verify$_\mathcal{C}$ h with certainty given $\mathcal{K}$, by hypothesis. Now define:

$$\alpha'(h, e) = \begin{cases} 0 & \text{if } e \subset \tau[m+1] \\ \alpha(h, e) & \text{otherwise.} \end{cases}$$

This verifies$_\mathcal{C}$ h with certainty given $\mathcal{K}'$. ■

7. Characterizations of Limiting Assessment

So far, we have isolated topological characterizations of verifiability, refutability, and decidability with certainty in terms of open, closed, and clopen sets, respectively. Open, closed, and clopen sets may be thought of as especially

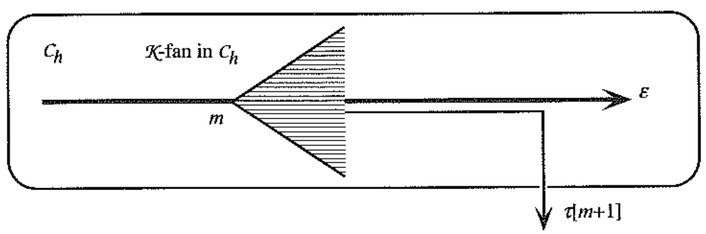


Figure 4.23

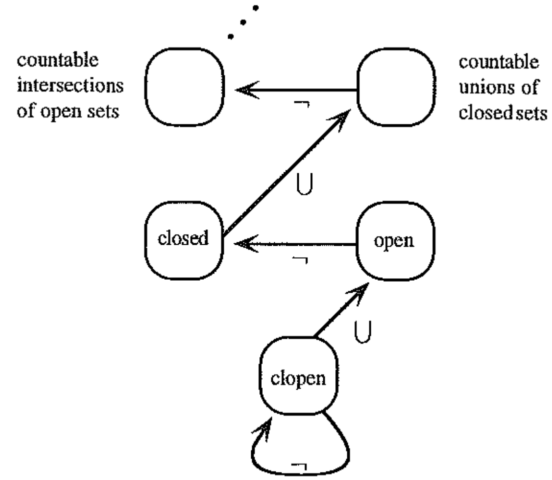


Figure 4.24

simple sets in a topological sense. Since we already know that limiting success is easier to guarantee than success with certainty, it is clear that we must consider weaker notions of topological simplicity if we are to characterize the other notions of reliability introduced in the preceding chapter. Happily, there is a standard scale of such notions, known as the *finite Borel hierarchy*. The Borel hierarchy is a system of mathematical cubbyholes that forms the lowest and most elementary part of the classificatory structure of *descriptive set theory*. The aim of descriptive set theory is to provide a kind of shipshape mathematics, in which there is a place for everything, and everything is put in its place. Each cubbyhole reflects a kind of intrinsic, mathematical complexity of the objects within it. Understanding is obtained by seeing the relative complexities of objects in terms of the cells in the hierarchy. The striking fact is that methodological success can be characterized *exactly* in terms of cubbyholes that are already familiar through other applications in logic, probability theory, and analysis.

To form the finite Borel hierarchy, we start out with the clopen sets. To build more complex sets, we iterate the operations of complementation and countable union (*Fig. 4.24*). The more countable unions and complementations it takes to generate a set, the higher the complexity of the set. That's all there is to it.[4] It is convenient to name the classes according to the following notational scheme due to Addison (1955). The first *Borel class* is just the collection of all clopen sets of the Baire space:

$$\mathcal{S} \in \Sigma_0^B \Leftrightarrow \mathcal{S} \text{ is clopen.}$$

[4] The full Borel hierarchy does not stop here, for we can form countable unions of sets drawn from each Π_n^B class, which are Borel sets that may reside at no finite level Π_n^B. We will have no use for such sets until chapter 13.

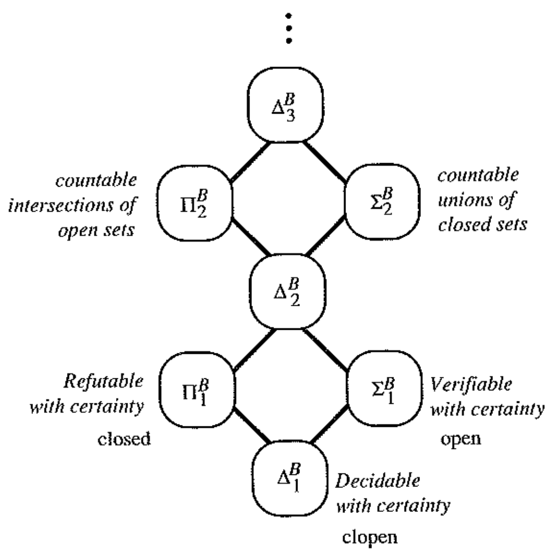


Figure 4.25

The other finite Borel classes are defined by the following induction:

$$\mathcal{S} \in \Sigma_{n+1}^B \Leftrightarrow \mathcal{S} \text{ is a countable union of complements of } \Sigma_n^B \text{ sets.}^5$$

The *dual Borel classes* are defined as follows:

$$\mathcal{S} \in \Pi_n^B \Leftrightarrow \bar{\mathcal{S}} \in \Sigma_n^B.$$

Finally, there are the *ambiguous Borel classes*:

$$\mathcal{S} \in \Delta_n^B \Leftrightarrow \mathcal{S} \in \Sigma_n^B \cap \Pi_n^B.$$

Thus $\Sigma_0^B = \Pi_0^B = \Delta_0^B =$ the class of clopen sets, the open sets are the Σ_1^B sets (since each open set is a union of fans and hence is a *countable* union of clopen sets), the closed sets are the Π_1^B sets (since they are complements of open sets), and the clopen sets are the Δ_1^B sets (both closed and open). Σ_2^B sets are countable unions of closed sets and Π_2^B sets are countable intersections of open sets. Since every singleton $\{\varepsilon\}$ is closed in $\mathcal{N}$, and hence is the complement of an open set, every countable subset of $\mathcal{N}$ is Σ_2^B, and every complement of a countable set is Π_2^B. It will be shown later that every level in this classificatory structure contains new subsets of $\mathcal{N}$ (*Fig. 4.25*).

[5] I.e., $\mathcal{S} \in \Sigma_{n+1}^B \Leftrightarrow$ there is a countable $\Gamma \subseteq \Sigma_n^B$ such that $\mathcal{S} = \bigcup \{\bar{\mathcal{R}}: \mathcal{R} \in \Gamma\}$.

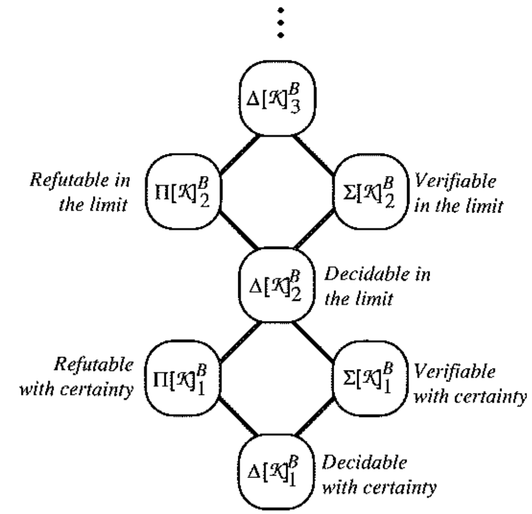


Figure 4.26

Proposition 4.9

For each n, $\Delta_n^B \subset \Sigma_n^B$.

Proof Deferred until chapter 7. ■

The finite Borel classes can be defined for restrictions of $\mathcal{N}$ to $\mathcal{K}$ just by starting out with $\mathcal{K}$-clopen sets instead of clopen sets and carrying through the induction from there. For $\mathcal{S} \subseteq \mathcal{K}$ define:

$$\mathcal{S} \in \Sigma[\mathcal{K}]_0^B \Leftrightarrow \mathcal{S} \text{ is } \mathcal{K}\text{-clopen.}$$

$$\mathcal{S} \in \Sigma[\mathcal{K}]_{n+1}^B \Leftrightarrow \mathcal{S} \text{ is a countable union of complements of } \Sigma[\mathcal{K}]_n^B \text{ sets.}$$

The dual and ambiguous classes are defined in terms of these as before. When we restrict the hierarchy to $\mathcal{K}$, it may collapse, depending on the character of $\mathcal{K}$. For a trivial example, if $\mathcal{K}$ is a singleton $\{\varepsilon\}$, then each $\mathcal{S} \subseteq \mathcal{N}$ is $\Delta[\mathcal{K}]_0^B$, and if $\mathcal{K}$ is countable, then each $\mathcal{S} \subseteq \mathcal{N}$ is $\Sigma[\mathcal{K}]_2^B$. Since the levels in the hierarchy will be seen to correspond to levels of local underdetermination by evidence, it is clear that to determine the sorts of assumptions that give rise to such collapse is a matter of the highest methodological importance.

The characterization of limiting reliability may now be stated (*Fig. 4.26*).[6]

[6] The following approach is a relativized, noncomputational version of the techniques introduced in Gold (1965) and Putnam (1965). A similar perspective is developed in Kugel (1977).

Proposition 4.10 (Gold 1965, Putnam 1965)

$$H \text{ is } \begin{bmatrix} \textit{verifiable}_C \\ \textit{refutable}_C \\ \textit{decidable}_C \end{bmatrix} \textit{ in the limit given } \mathcal{K}$$

$$\Leftrightarrow \textit{for each } h \in H,\ C_h \cap \mathcal{K} \in \begin{bmatrix} \Sigma[\mathcal{K}]^B_2 \\ \Pi[\mathcal{K}]^B_2 \\ \Delta[\mathcal{K}]^B_2 \end{bmatrix}.$$

Proof: It suffices to show the verifiability case. The refutability case follows by duality, and the ambiguous case follows from proposition 3.4(b). ($\Leftarrow$) Suppose that for each $h \in H$, $C_h \cap \mathcal{K} \in \Sigma[\mathcal{K}]^B_2$. So $C_h \cap \mathcal{K}$ is the union of some countable sequence $\mathcal{S}[h]_0, \mathcal{S}[h]_1, \ldots, \mathcal{S}[h]_n, \ldots$ of $\mathcal{K}$-closed sets. Thus, for each i, there is some $G[h]_i \subseteq \omega^*$ such that $\overline{\mathcal{S}[h]_i}$ is the union of all $\mathcal{K}$-fans $[e] \cap \mathcal{K}$ such that $e \in G[h]_i$ (*Fig. 4.27*).

When provided with hypothesis h and evidence e, α proceeds as follows. A pointer is initialized to 0. Then α scans e in stages $e|0, e|1, \ldots, e|lh(e)$. The pointer moves forward one step at a given stage if the data refutes the closed set $\mathcal{S}[h]_i$ that the pointer points to upon entering that stage, and the pointer stays put otherwise. More precisely:

$$pointer(h, \mathbf{0}) = 0$$

$$pointer(h, e|n+1) = \begin{cases} pointer(h, e|n) + 1 \\ \quad \textit{if } e \textit{ extends some member of } G[h]_{pointer(h, e|n)}. \\ pointer(h, e|n) \textit{ otherwise.} \end{cases}$$

The method α conjectures 0 when the pointer moves on reading the last datum in e, and conjectures 1 otherwise (*Fig. 4.28*):

$$\alpha(h, \mathbf{0}) = 0$$

$$\alpha(h, e^*x) = \begin{cases} 1 & \textit{if } pointer(h, e^*x) = pointer(h, e) \\ 0 & \textit{otherwise.} \end{cases}$$

Now we verify that α works. Let $h \in H$, $\varepsilon \in \mathcal{K}$. Suppose $\varepsilon \in C_h$. Then there is some $\mathcal{K}$-closed set in α's enumeration containing ε. Let $\mathcal{S}[h]_i$ be the first such. Since $\varepsilon \notin \overline{\mathcal{S}[h]_i}$, there is no n such that $\varepsilon|n \in G[h]_i$, so the pointer can never move past position i. Thus, there is some time after which the pointer stops moving and at this time α stabilizes correctly to 1 on ε. Suppose $\varepsilon \notin C_h$. Then ε is not in any $\mathcal{S}[h]_i$. So for each i, $\varepsilon \in \overline{\mathcal{S}[h]_i}$. Thus, for each i, there is an $e' \in G[h]_i$ such that $\varepsilon|lh(e') = e'$. It follows that the pointer never stops moving, so α does not stabilize to 1.

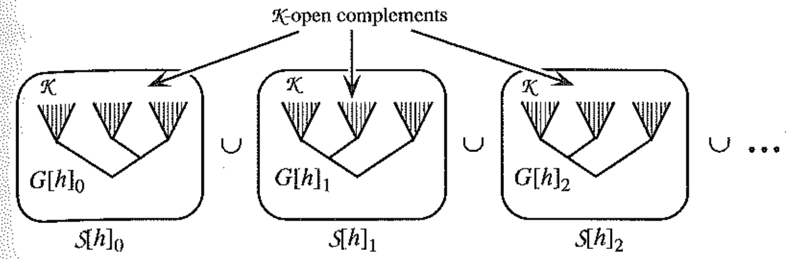


Figure 4.27

($\Rightarrow$) In previous characterizations, we have constructed demons to show that an arbitrary scientist can be fooled if the characterization condition is false. This time it is easier to do something different. We will use the fact that the scientist succeeds to decompose $C_h \cap \mathcal{K}$ into a countable union of $\mathcal{K}$-closed sets, showing directly that $C_h \cap \mathcal{K} \in \Sigma[\mathcal{K}]^B_2$. This technique was originally employed by Gold (1965) and Putnam (1965) in a computational setting, but it also works in the ideal case.

Suppose that some method α verifies H in the limit given $\mathcal{K}$. Then we have for each $h \in H$:

$$\forall \varepsilon \in \mathcal{K}[C(\varepsilon, h) \Leftrightarrow \exists n\ \forall m \geq n,\ \alpha(h, \varepsilon|m) = 1].$$

We can think of this fact as a definition of C_h over the restricted domain $\mathcal{K}$:

$$\forall \varepsilon \in \mathcal{K},\ \varepsilon \in C_h \cap \mathcal{K} \Leftrightarrow \exists n\ \forall m \geq n\ \alpha(h, \varepsilon|m) = 1.$$

The set $\mathcal{A}_h(m)$ of all $\varepsilon \in \mathcal{K}$ such that $\alpha(h, \varepsilon|m) = 1$ is $\mathcal{K}$-clopen, since we only

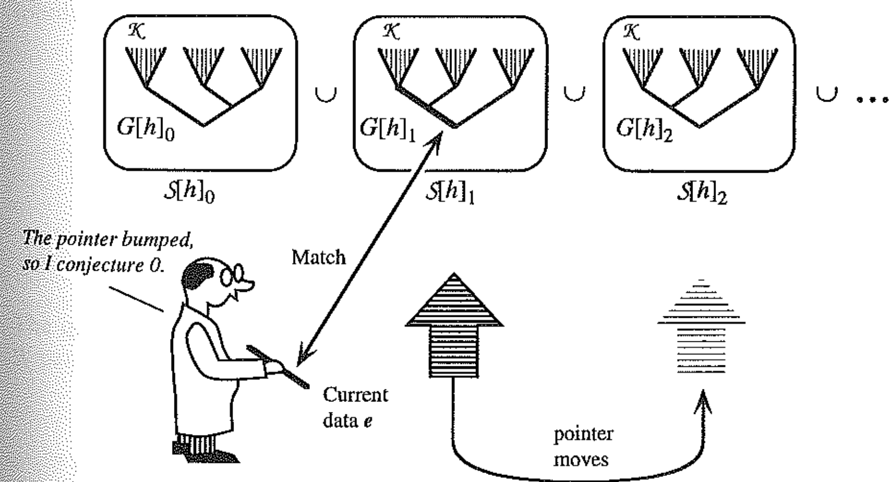


Figure 4.28

have to look at the data up to time m to determine what α will do. Also, for each $\varepsilon \in \mathcal{K}$ we have:

$$(\forall m \geq n\ \alpha(h, \varepsilon|m) = 1) \Leftrightarrow \varepsilon \in (\mathcal{A}_h(n) \cap \mathcal{A}_h(n+1) \cap \mathcal{A}_h(n+2), \ldots).$$

Let $\mathcal{B}_h(n)$ denote this countable intersection. Then we have for each $\varepsilon \in \mathcal{K}$:

$$\varepsilon \in C_h \cap \mathcal{K} \Leftrightarrow \exists n\, \forall m \geq n\ \alpha(h, \varepsilon|m) = 1$$
$$\Leftrightarrow \varepsilon \in (\mathcal{B}_h(0) \cup \mathcal{B}_h(1) \cup \mathcal{B}_h(2) \cup \cdots).$$

Thus $C_h \cap \mathcal{K} \in \Sigma[\mathcal{K}]^B_2$. ■

The method α constructed in this proof may be called the *bumping pointer method*. What we have just seen is that it is a *complete architecture* for verification in the limit, in the sense that for any inductive problem $(C, \mathcal{K}, H)$ verifiable in the limit by an ideal agent, the bumping pointer method employing the appropriate decomposition of each C_h into a countable union of $\mathcal{K}$-closed sets verifies $(C, \mathcal{K}, H)$ in the limit.

There is a very direct picture of how topological structure guarantees the bumping pointer method's success. Suppose that $C_h \cap \mathcal{K} \in \Sigma[\mathcal{K}]^B_2$. Then $C_h \cap \mathcal{K}$ is a countable union of $\mathcal{K}$-closed sets. Hence, the complement of $C_h \cap \mathcal{K}$ is a countable intersection of $\mathcal{K}$-open sets. Each $\mathcal{K}$-open set is a union of $\mathcal{K}$-fans. Recall that the intersection of two fans is empty unless one fan is a subset of the other, in which case the intersection is the fan with the longer handle. So a countable intersection of open sets can yield infinitely many paths determined by countable intersections of fans with shared handles, as in the case of ε in *Fig. 4.29*.

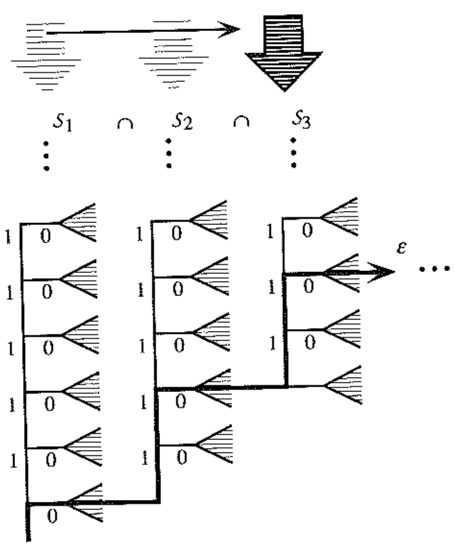


Figure 4.29

When $\varepsilon \notin C_h$, the data stream will "break through" each open set and the pointer will be moved infinitely often, so that the method emits infinitely many 0s. When $\varepsilon \in C_h$, there is some open set that ε never breaks through, so the pointer eventually remains fixed.

Karl Popper[7] has recommended holding onto a theory that has "passed muster" until it is refuted. Proposition 4.10 shows that this is not a very good idea if reliability is at stake. Recall the hypothesis h_{fin} that particle p is only finitely divisible. Popper's method would always stabilize to 1 on this hypothesis and hence would fail to verify it in the limit, since it is consistent with all possible data streams (given the setup described in chapter 3). But the bumping pointer architecture just described is guaranteed to verify$_C$ h_{fin} in the limit. In fact, it will beat Popper's proposal whenever $C_h \cap \mathcal{K} \in \Sigma[\mathcal{K}]^B_2 - \Pi[\mathcal{K}]^B_1$. When $C_h \cap \mathcal{K} \in \Delta[\mathcal{K}]^B_2$, then there are two bumping pointer methods α_1, α_2, such that α_1 verifies$_C$ h in the limit given $\mathcal{K}$ and α_2 refutes$_C$ h in the limit given $\mathcal{K}$. When these two methods are assembled together according to the strategy of proposition 3.4(b), we have a method that decides$_C$ h in the limit given $\mathcal{K}$. This method beats Popper's proposal whenever $C_h \cap \mathcal{K} \in \Delta[\mathcal{K}]^B_2 - \Pi[\mathcal{K}]^B_1$. Moreover, if the hypothesis is refutable with certainty, as Popper's method requires, the bumping pointer method can be set up so as to be guaranteed to succeed in this sense. The only modification required is to replace the infinite enumeration of $\mathcal{K}$-closed sets with a single $\mathcal{K}$-closed set and to have the method return '!', 0 when this set is refuted. Thus, the bumping pointer architecture can be modified so as to obey Popper's requirement whenever doing so doesn't get in the way of reliability, and to relax it just when it does.

Another application of this result concerns the question of optimal reliability.[8] We have already seen by proposition 3.15 that if h is not verifiable$_C$ in the limit given $\mathcal{N}$ but is verifiable in the limit given $\mathcal{K} \subset \mathcal{N}$, then there is an $\varepsilon \notin \mathcal{K}$ such that h is verifiable$_C$ in the limit given $\mathcal{K} \cup \{\varepsilon\}$. In purely topological terms, this amounts to:

Corollary 4.11

If $\mathcal{P} \notin \Sigma^B_2$ and $\mathcal{P} \in \Sigma[\mathcal{K}]^B_2$ then for some nonempty, finite $\mathcal{S} \subseteq \mathcal{N} - \mathcal{P}$, $\mathcal{P} \in \Sigma[\mathcal{K} \cup \mathcal{S}]^B_2$. ■

The question then arises whether the result can be strengthened to show that $\mathcal{K}$ can be weakened by an *infinite* set of possible data streams. In fact, this is

[7] Popper (1968).

[8] This issue was visited at the end of the preceding chapter as well as in proposition 4.7.

true, as the following proposition shows.[9] In some cases, it can even be shown that weakening by an uncountably infinite set of data streams is possible.[10]

Proposition 4.12 (with C. Juhl)

If $\mathcal{P} \notin \Sigma_2^B$ and $\mathcal{P} \in \Sigma[\mathcal{K}]_2^B$ then for some countably infinite $\mathcal{S} \subseteq \mathcal{N} - \mathcal{K}$, $\mathcal{P} \in \Sigma[\mathcal{K} \cup \mathcal{S}]_2^B$.

Proof: Given that $\mathcal{P} \notin \Sigma[\mathcal{K}]_2^B$ then $\mathcal{N} - \mathcal{K}$ is infinite, else by proposition 3.15, $\mathcal{P} \in \Sigma_2^B$. Then either (I) $(\mathcal{N} - \mathcal{K}) - \mathcal{P}$ is infinite or (II) $(\mathcal{N} - \mathcal{K}) \cap \mathcal{P}$ is infinite. Let α verify $\mathcal{P}$ in the limit given $\mathcal{K}$, as guaranteed by hypothesis and by proposition 4.10. Case (I). Either (A) there is an infinite, closed $\mathcal{S} \subseteq (\mathcal{N} - \mathcal{K}) - \mathcal{P}$ or (B) there is not. (A) Consider a method that verifies $\mathcal{P}$ in the limit given $\mathcal{K} \cup \mathcal{S}$ as follows. Let α' conjecture 0 until $\mathcal{S}$ is refuted, and agree with α thereafter. If $\varepsilon \in \mathcal{S}$ then α' correctly stabilizes to 0. Else, $\mathcal{S}$ is eventually refuted with certainty, after which α' agrees with α, which is guaranteed to work given $\mathcal{K}$. (B) Suppose there is no infinite, closed $\mathcal{S} \subseteq (\mathcal{N} - \mathcal{K}) - \mathcal{P}$. Then in particular, $(\mathcal{N} - \mathcal{K}) - \mathcal{P}$ is not closed, and hence has a limit point in $\mathcal{K} \cup \mathcal{P}$. Let Γ be a sequence of distinct elements of $(\mathcal{N} - \mathcal{K}) - \mathcal{P}$ that converges to some $\varepsilon \in \mathcal{K} \cup \mathcal{P}$. Let $\mathcal{S}$ be the range of Γ. Since Γ converges to ε, ε is the only limit point of $\mathcal{S}$ missing from $\mathcal{S}$ and hence (*) $\mathcal{S} \cup \{\varepsilon\}$ is closed. Let $b = 1$ if $\varepsilon \in \mathcal{P}$ and let $b = 0$ otherwise. Then construct the method α' that conjectures b until ε is refuted, that conjectures 0 until $\mathcal{S} \cup \{\varepsilon\}$ is refuted, and that agrees with α thereafter. On data stream ε, α' correctly stabilizes to b. If $\varepsilon' \in \mathcal{S}$, then α' correctly stabilizes to 0. If $\varepsilon' \notin \mathcal{S}$ and $\varepsilon' \neq \varepsilon$, then eventually the set $\mathcal{S} \cup \{\varepsilon\}$ is refuted by (*), and α' correctly agrees with α, which is correct given $\mathcal{K}$. Case (II) is similar. ■

8. Efficient Data Use

A standard objection to the limiting analysis of inductive methods is that success in the limit is consistent with any crazy behavior in the short run.[11] As far as

[9] Proposition 4.12 belongs to what we might describe as *inverted complexity theory*. In standard complexity theory, one starts with a fixed, "tidy" space (e.g., $\mathcal{N}$), and with a given object (e.g., $\mathcal{P} \subseteq \mathcal{N}$). The question is then to determine what the complexity (e.g., Borel complexity) of the object is in the fixed space. In propositions 4.12, 4.7, and 4.11, the issue is rather to specify an object $\mathcal{P}$ and a fixed complexity (e.g., Σ_2^B), and then to describe the set of all possible background spaces $\mathcal{K} \subseteq \mathcal{N}$ in which the object has the given complexity. This is what distinguishes epistemological applications of topology from standard results in descriptive set theory.

[10] We might ask, further, whether it is possible to expand $\mathcal{K}$ by an uncountable $\mathcal{S} \subseteq \mathcal{N} - \mathcal{K}$ when $\mathcal{P} \notin \Sigma_2^B$ and $\mathcal{P} \in \Sigma[\mathcal{K}]_2^B$. It is possible to do so if $\mathcal{N} - \mathcal{K} - \mathcal{P}$ is what is known as a *Suslin set*, for each such set has an uncountable, closed subset (Moschavakis 1980: 79, theorem 2C.2). Let $\mathcal{S}$ be this uncountable closed subset. Let α verify $\mathcal{P}$ in the limit given $\mathcal{K}$. Let α' conjecture 0 until $\mathcal{S}$ is refuted, and agree with α thereafter. Similarly, when $\mathcal{N} - \mathcal{K} \cap \mathcal{P}$ is a Suslin set, let α' conjecture 1 until $\mathcal{S}$ is refuted and agree with α thereafter.

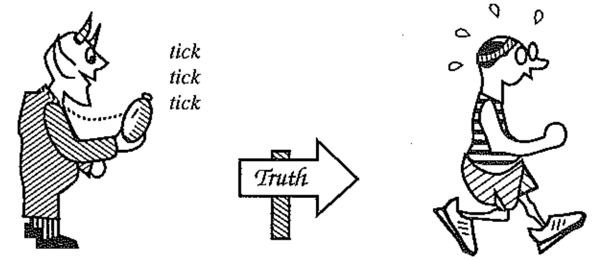


Figure 4.30

convergence is concerned, this is true. But that doesn't mean that any method reliable in the limit is as good as any other, for insanity in the short run may needlessly delay the onset of convergence (*Fig. 4.30*). Recall that

$$\textbf{\textit{modulus}}_\alpha(h, \varepsilon) = \text{the least } n \text{ such that for all } m \geq n, \alpha(h, \varepsilon|n) = \alpha(h, \varepsilon|m).$$

To simplify what follows, we will let the modulus function assume value ω when no such n exists. Define:

$$\beta \leq_{\mathcal{K}}^{H} \alpha \Leftrightarrow \text{for all } h \in H,\ \varepsilon \in \mathcal{K},\ modulus_\beta(h, \varepsilon) \leq modulus_\alpha(h, \varepsilon).$$

$$\beta <_{\mathcal{K}}^{H} \alpha \Leftrightarrow \beta \leq_{\mathcal{K}}^{H} \alpha \text{ but not } \alpha \leq_{\mathcal{K}}^{H} \beta.$$

In the former case we say that β is *as fast* as α on H given $\mathcal{K}$. In the latter case we say that β is *strictly faster*. β is strictly faster than α on H given $\mathcal{K}$ just in case β stabilizes as soon as α on each hypothesis in H and data stream in $\mathcal{K}$, and properly faster on some hypothesis in H and data stream in $\mathcal{K}$. This is a special case of what decision theorists call *weak dominance*.[12]

It would be nice to have a reliable method that is as fast as every other reliable method on H given $\mathcal{K}$. Let $\mathcal{M}$ be a class of methods. Define:

α is strongly data-minimal in $\mathcal{M}$ for H, $\mathcal{K}$
$\Leftrightarrow \alpha \in \mathcal{M}$ and for each $\beta \in \mathcal{M}$, $\alpha \leq_{\mathcal{K}}^{H} \beta$.

In other words, for each method β in $\mathcal{M}$, a strongly data-minimal method α gets to the truth as quickly as β on each data stream. We will be interested in cases in which $\mathcal{M}$ is the set of all solutions to the inductive problem in question. Strongly data-minimal limiting verifiers exist only for very trivial inductive problems. For example, if $C_h = \varnothing$ for each $h \in H$, then the constant method $\alpha(h, e) = 0$ is data-minimal over all limiting-verifiers of H. However, as soon as we have a nontrivial hypothesis (i.e., $C_h \neq \mathcal{K}, \varnothing\}$ and $\mathcal{K}$ has at least two data streams, we can forget about strongly data-minimal solutions. For suppose

[12] One act *weakly dominates* another if its outcome is as good in each possible world state and is better in some possible world state. Here, the possible world states are possible data streams in $\mathcal{K}$ and the outcome of using a method is better insofar as one gets to the truth sooner.

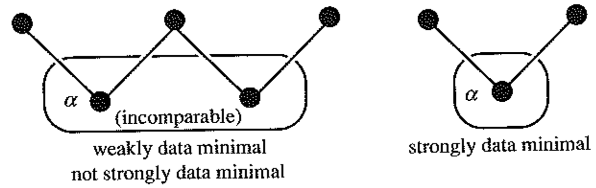


Figure 4.31

$\alpha(j, \mathbf{0}) = 1$. Then since $C_h \neq \mathcal{K}$, there is some $\varepsilon \in \mathcal{K} - C_h$, so α cannot be as fast on ε as the method that returns 0 until the current data e deviates from ε. A similar argument holds if $\alpha(h, \mathbf{0})$ is any other value.

The situation becomes more interesting when we lower our sights to methods that are not weakly dominated by any other method in $\mathcal{M}$ (*Fig. 4.31*):

$$\alpha \textit{ is weakly data-minimal}^{13} \textit{ in } \mathcal{M} \textit{ for } H \Leftrightarrow \alpha \in \mathcal{M} \textit{ and for each } \beta \in \mathcal{M},\ \beta \not<^{H}_{\mathcal{K}} \alpha.$$

In other words, there is no method in $\mathcal{M}$ that gets to the truth on each data stream as fast as α and that beats α on some data stream. Weak data minimality is weaker than strong data minimality because it countenances situations in which α gets to the truth sooner than β on some data stream and β also gets to the truth sooner than α on some other data stream. In decision theory, an *admissible* strategy is one that is not weakly dominated by any other strategy. Weak data minimality is a special case of admissibility if weak dominance is understood with respect to convergence time over all possible data streams, and the competing strategies are all in $\mathcal{M}$.

Weakly data-minimal solutions to problems of verifiability in the limit are much easier to find. Recall the hypothesis h_{fin}, which says that only finitely many 1s will occur and define the simple method:

$$\alpha_0(h, e) = \begin{cases} 1 & \textit{if } e \textit{ ends with } 0 \\ 0 & \textit{otherwise.} \end{cases}$$

Proposition 4.13

α_0 is weakly data-minimal among limiting verifiers$_C$ of h_{fin} given 2^ω.

Proof: It is evident that α_0 verifies$_C$ h_{fin} in the limit. Suppose β also verifies$_C$ h_{fin}. Suppose β's modulus is strictly smaller than that of α_0 on some $\varepsilon \in C_h$. Then ε stabilizes to 0 at some n, so the modulus of α_0 is n, and $\varepsilon_{n-1} = 1$. The modulus of β is strictly less than n, say n'. Let $\tau|n-1 = \varepsilon|n-1$ and let τ have only 1s thereafter, so that τ stabilizes to 1 with modulus $n-1$. Then the

[13] Cf. Gold (1967) and Osherson et al. (1986).

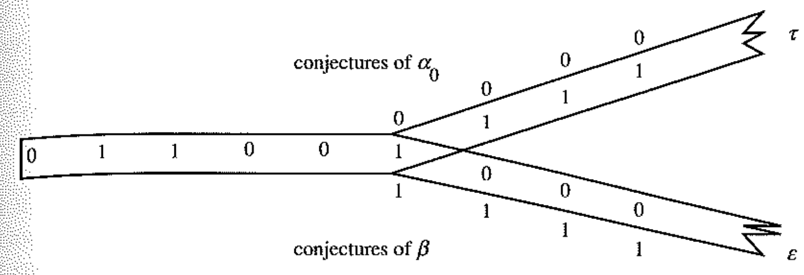


Figure 4.32

modulus of α_0 on τ is $n-1$, but the modulus of β on τ is at best n, so β is not strictly faster than α_0 (*Fig. 4.32*). So $\beta \not<^{H}_{\mathcal{K}} \alpha_0$.

Suppose then that β's modulus n is strictly smaller than that of α_0 on some $\varepsilon \in 2^\omega - C_h$. Then β stabilizes to 0 on ε. If α_0 stabilizes to 0 on ε, the argument is as before. If α_0 does not stabilize to 0 on ε, then there is some $k > n$ such that α conjectures 1 on $\varepsilon|k$ after β has stabilized to 0. Thus, $\varepsilon_k = 0$, by the definition of α_0. Extend $\varepsilon|k$ with all 0s thereafter to form τ. The modulus of α_0 is k on τ, but the modulus of β cannot be lower than $k+1$ since β conjectures 0 on $\tau|k$. So $\beta \not<^{H}_{\mathcal{K}} \beta$. ■

This is just one very simple example. Can it be extended to other cases? The surprising answer is that every problem solvable in the limit has a weakly data-minimal solution. I begin with the limiting decidability case, which is simpler, and then proceed to the more interesting case of limiting verifiability.

Proposition 4.14[14]

If H is decidable$_C$ in the limit given $\mathcal{K}$ then some weakly data-minimal method decides$_C$ H in the limit given $\mathcal{K}$.

Proof: By hypothesis, for each $h \in H$, $C_h \cap \mathcal{K} \in \Delta[\mathcal{K}]^B_2$. We define method α as follows. Let h and e be given. There are $\mathcal{K}$-closed sets $\mathcal{S}_0, \mathcal{R}_0, \mathcal{S}_1, \mathcal{R}_1, \ldots$ such that:

$$C_h \cap \mathcal{K} = \mathcal{S}_0 \cup \mathcal{S}_1 \cup \cdots \cup \mathcal{S}_n \cup \cdots$$
$$\overline{C_h} \cap \mathcal{K} = \mathcal{R}_0 \cup \mathcal{R}_1 \cup \cdots \cup \mathcal{R}_n \cup \cdots.$$

Enumerate the set of all pairs of form $(\mathcal{S}_i, 1)$, $(\mathcal{R}_j, 0)$ as $\pi_0, \pi_1, \ldots$. Let α conjecture the Boolean value associated with the first pair whose $\mathcal{K}$-closed set is not refuted by e (*Fig. 4.33*).

It is easily verified along the lines of proposition 4.10 that α decides$_C$ H in the limit given $\mathcal{K}$. Now it is shown that α is data-minimal. For suppose there is a β such that β decides$_C$ H in the limit given $\mathcal{K}$ and for some $h \in H$, $\varepsilon \in \mathcal{K}$, β converges sooner than α does. First, suppose that $\varepsilon \in C_h$. Let $modulus_\beta(h, \varepsilon) = n$.

[14] Similar to Gold (1967) and Osherson et al. (1986).

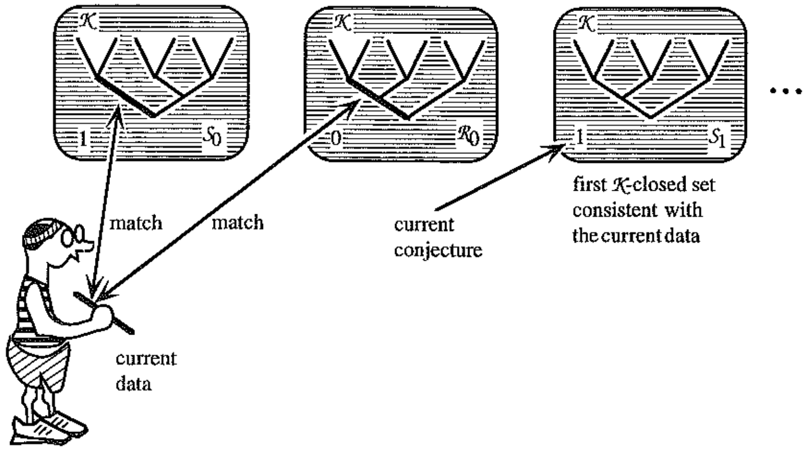


Figure 4.33

For some $m \geq n$, $\alpha(h, \varepsilon|m) = 0$. Hence, α's pointer points to some pair $\pi_k = (\mathcal{R}, 0)$ at stage m. So $\mathcal{R}$ is not refuted by $\varepsilon|m$. Hence, there is an $\varepsilon' \in \mathcal{K} - C_h$ such that ε' extends $\varepsilon|m$. α's pointer is never moved after stage m on ε', so $modulus_\alpha(h, \varepsilon') \leq m$. But $\beta(h, \varepsilon'|m) = 1$, so $modulus_\beta(h, \varepsilon') > m$. A similar argument works when $\varepsilon \notin C_h$. Since β is an arbitrary limiting decider$_C$ of H given $\mathcal{K}$, α is data-minimal among such solutions. ■

The preceding argument works because whenever α has not yet stabilized to 1 on some $\varepsilon \in \mathcal{K} \cap C_h$, it can be excused because there is some $\tau \in \mathcal{K} - C_h$ that it might have already stabilized to 0 on. It is not obvious that the same sort of argument can be given in the case of limiting verification. For suppose that h is verifiable but not decidable in the limit and that α verifies h in the limit. Then α fails to converge to 0 on infinitely many data streams that make h false. But then there is a method β that also converges to 0 on a data stream ε making h false on which α fails to converge to 0 (i.e., define β to conjecture 0 until the actual data veers off from ε and to pass control to α thereafter). On ε, the modulus of α is infinitely greater than that of β. It therefore seems that it should be possible to speed up any limiting verifier of h. But in fact, we can define α so that any β that converges to 0 when α does not must converge to 1 later than α on some data stream making h true.

Proposition 4.15

(a) *If H is verifiable$_C$ in the limit given $\mathcal{K}$ then some method is weakly data-minimal over limiting verifiers$_C$ of H in the limit given $\mathcal{K}$.*

(b) *The limiting refutation case is similar.*

Proof: (a) Suppose that for each $h \in H$, $C_h \cap \mathcal{K} \in \Sigma[\mathcal{K}]^B_2$. For each $h \in H$, choose $\mathcal{K}$-closed sets $\mathcal{S}[h]_i$ such that

$$C_h \cap \mathcal{K} = \mathcal{S}[h]_0 \cup \mathcal{S}[h]_1 \cup \cdots \cup \mathcal{S}[h]_n \cup \cdots$$

and if $C_h \cap \mathcal{K} \neq \varnothing$, then no $\mathcal{S}[h]_i = \varnothing$. The function *pointer*$(h, e)$ is set up just as in the proof of proposition 4.10 above. The method α is given as follows. For each h, e, let $\tau_{h,e}$ be an arbitrary, fixed choice of a data stream extending e in $\mathcal{K} - C_h$, if such a data stream exists, and let $\tau_{h,e}$ be undefined otherwise.

$\alpha(h, e)$:

(0) *if $C_h \cap \mathcal{K} = \varnothing$, then conjecture* 0

(1) *else if $e = \mathbf{0}$, then conjecture* 1

(2) *else if there is an $e' \subseteq e$ such that pointer$(h, e') \neq$ pointer(h, e) and $\tau_{h,e'}$ extends e, then conjecture* 0

(3) *else conjecture* 1.

The idea is that $\tau_{h,e}$ is "entertained" by α if the pointer bumps on e, and it continues to be entertained until data deviating from it is seen (*Fig. 4.34*). While some $\tau_{h,e}$ is entertained, α conjectures 0. Otherwise, α conjectures 1. This inductive architecture will be referred to as the *opportunistic architecture*, since it takes advantage of opportunities to stabilize to 0 in order to block the argument that some other method is properly faster.

First it must be shown that α verifies$_C$ H in the limit given $\mathcal{K}$. Let $h \in H$. Let $\varepsilon \in C_h \cap \mathcal{K}$. Then for some i, $\varepsilon \in \mathcal{S}[h]_i$. Thus the pointer cannot bump past position i, so the pointer bumps only finitely many times on ε, say on $\varepsilon|n_1$, $\varepsilon|n_2, \ldots, \varepsilon|n_k$. Since $\varepsilon \in C_h \cap \mathcal{K}$, $\varepsilon \neq \tau_{h,\varepsilon|n_1}, \ldots, \tau_{h,\varepsilon|n_k}$. So for some time m, $\varepsilon|m$ is not extended by $\tau_{h,\varepsilon|n_k}$, and hence (2) is never again satisfied after stage m. Moreover, (0) never applies, since for each n, $\varepsilon \in [\varepsilon|n] \cap C_h \cap \mathcal{K}$. So (3) applies forever after m, and α stabilizes correctly to 1 (*Fig. 4.35*).

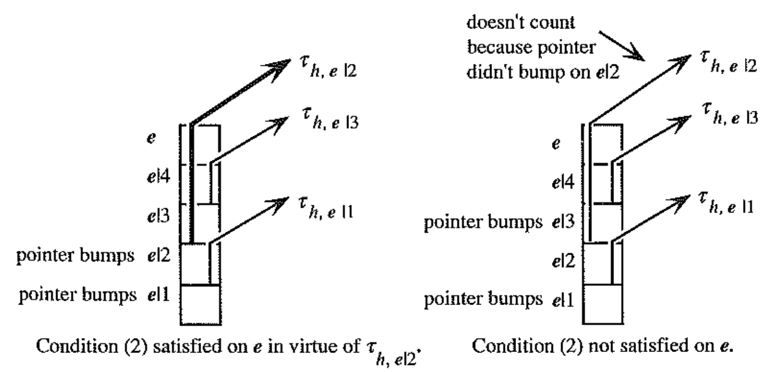


Figure 4.34

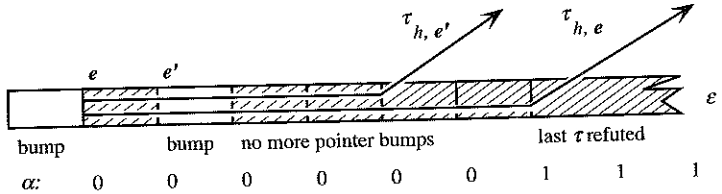


Figure 4.35

Now suppose $\varepsilon \in \mathcal{K} - C_h$. For each i, $\varepsilon \notin \mathcal{S}[h]_i$, so the pointer is bumped infinitely often. Moreover, ε itself witnesses the existence of $\tau_{h, \varepsilon|n}$, for each n, so condition (2) is satisfied infinitely often, and infinitely many 0s are conjectured by α on ε (*Fig. 4.36*).

It remains to show that α is weakly data-minimal. Suppose that for some $\varepsilon \in \mathcal{K}$ and for some limiting verifier β of h given $\mathcal{K}$, $n' = modulus_\beta(h, \varepsilon) < modulus_\alpha(h, \varepsilon) = n$. Suppose $\varepsilon \in C_h \cap \mathcal{K}$. Then both β and α stabilize to 1 on ε. So $\alpha(h, \varepsilon|n-1) = 0$. (0) is not satisfied by $\varepsilon|n-1$, since $\varepsilon \in C_h \cap \mathcal{K}$. So $\varepsilon|n-1$ satisfies (2). Then by condition (2), there is some $\tau_{h,\varepsilon|n-1} \in \mathcal{K} - C_h$ that extends $\varepsilon|n-1$, and by condition (2), $modulus_\alpha(h, \tau_{h,\varepsilon|n-1}) = n-1$. But since $\beta(h, \varepsilon|n-1) = 1$, $modulus_\beta(h, \tau_{h,\varepsilon|n-1}) > n-1$. So β is not strictly faster than α (*Fig. 4.37*).

Now suppose $\varepsilon \in \mathcal{K} - C_h$. Since $n' = modulus_\beta(h, \varepsilon) < modulus_\alpha(h, \varepsilon)$, β stabilizes to 0 and α either stabilizes to 0 later or never stabilizes. Either way, there exists $n \geq n'$ such that $\alpha(h, \varepsilon|n) = 1$. Suppose this happens due to clause (1). In that case, $n' = 0$. Since clause (0) was bypassed, $C_h \cap \mathcal{K} \neq \varnothing$, so $\mathcal{S}[h]_0 \neq \varnothing$, by our specification of the enumeration $\mathcal{S}[h]_i$. Hence, there exists $\tau \in \mathcal{S}[h]_0$ such that the pointer remains forever stuck at 0 on τ, so α stabilizes to 1 on τ with modulus 0. Since β starts out with 0, its modulus on τ can be no less than 1, so β is not strictly faster than α. Suppose, finally, that $\alpha(h, \varepsilon|n) = 1$ in virtue of clause (3). As in the proof of correctness for α, ε itself witnesses that for each k, $\tau_{h,\varepsilon|k}$ exists. Then (2) fails for $\varepsilon|n$ because $pointer(\varepsilon|n-1) = pointer(\varepsilon|n)$. Let k be the pointer position on $\varepsilon|n$. Then $\varepsilon|n$ is extended by some $\varepsilon' \in \mathcal{S}[h]_k \subseteq C_h$. $Modulus_\alpha(h, \varepsilon') = n$, since the pointer remains stuck at k on ε' and hence α stabilizes to 1 on ε at least by n. But since $\beta(h, \varepsilon|n) = 0$, $modulus_\beta(h, \varepsilon') > n$. So again β is not strictly faster than α. ■

Each opportunistic method is *consistent* in the sense that it stabilizes to 0 immediately when the hypothesis under test is refuted (i.e., when $C_h \cap \mathcal{K} \cap [e] = \varnothing$), and it stabilizes to 1 immediately when the hypothesis under test is

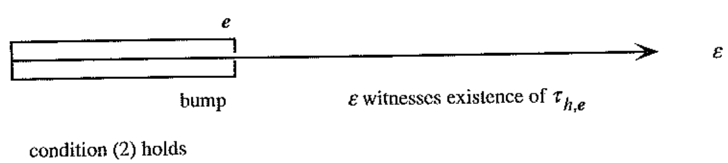


Figure 4.36

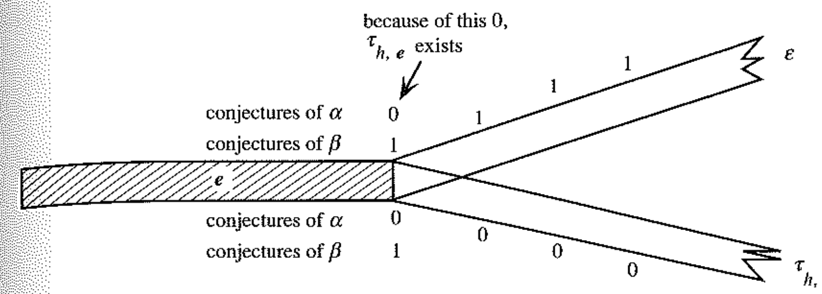


Figure 4.37

entailed by the data (i.e., when $\mathcal{K} \cap [e] \subseteq C_h \cap [e]$). In fact, consistency is necessary for weak data minimality in limiting verification problems (cf. exercise 4.7).

A remarkable feature of the above result is that no limiting decision procedure is strictly faster than an *arbitrary* specification of the opportunistic architecture for the problem in question, even though some other method may stabilize to 0 on more data streams. It is also interesting that such methods can be weakly data-minimal employing a pointer that bumps only one step each time a closed set in the enumeration is refuted. It would seem that a method employing a pointer that advances immediately to the first unrefuted closed set ought to be properly faster, but this is not the case.

9. A Characterization of *n*-Mind-Change Decidability

It is clear that if there is an n such that h is decidable with n mind changes given $\mathcal{K}$ then h is decidable in the limit given $\mathcal{K}$. The converse is false (cf. proposition 4.17 (a)). Since there are problems solvable with fixed numbers of mind changes that are neither verifiable nor refutable with certainty (section 3.5), the complexity of decidability with bounded mind changes is located above refutation and verification with certainty and below decidability in the limit (*Fig. 4.38*).

So we cannot characterize bounded mind change decidability in terms of Borel complexity, as there are no Borel classes left to characterize it. But as it turns out, there is another scale of topological complexity, the *finite difference hierarchy*,[15] that lies precisely in the required position. I will label this hierarchy with the superscript D. Whereas the finite Borel hierarchy builds complexity

[15] Cf. Kuratowski (1966). I am indebted to J. Tappenden for this reference, of which I was unaware when I worked out the results in the following two sections of this chapter. I was led to the question by Putnam (1965), which falls somewhat short of the full characterization, though providing such a characterization was not the point of the paper.

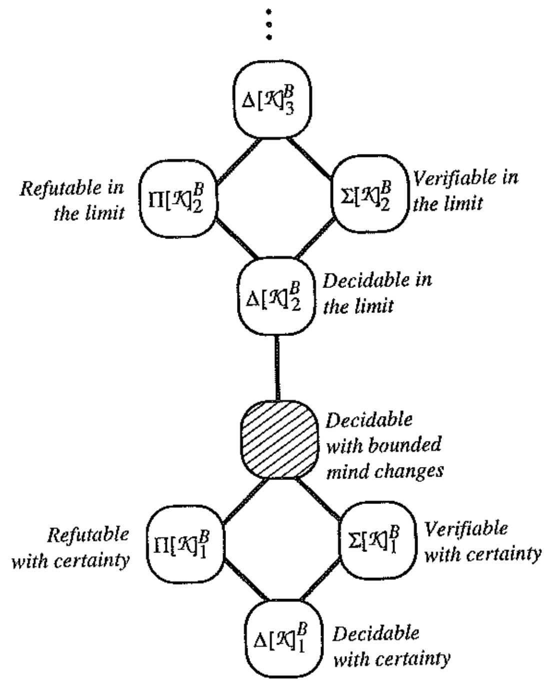


Figure 4.38

by complementation and countable union, the finite difference hierarchy builds complexity by finite intersection and union of alternating sequences of open and closed sets, so that $\mathcal{O}$ is simpler than $\mathcal{O} \cap \mathcal{P}$, which in turn is simpler than $(\mathcal{O} \cup \mathcal{P}) \cap \mathcal{O}'$, and so forth, where $\mathcal{O}, \mathcal{O}'$ are open and $\mathcal{P}$ is closed. Let $\mathcal{S} \subseteq \mathcal{K}$. Then we define:

$$\mathcal{S} \in \Sigma[\mathcal{K}]^D_0 \Leftrightarrow \mathcal{S} \text{ is } \mathcal{K}\text{-clopen.}$$

$$\mathcal{S} \in \Sigma[\mathcal{K}]^D_{n+1} \Leftrightarrow \text{for some } \mathcal{R} \in \Sigma[\mathcal{K}]^D_n, \text{ for some } \mathcal{K}\text{-open } \mathcal{O},\ \mathcal{S} = \bar{\mathcal{R}} \cap \mathcal{O}.$$

$$D_{\mathcal{K}} = \bigcup_{i \in \omega} \Sigma[\mathcal{K}]^D_i$$

The dual and ambiguous classes are defined in terms of the Σ classes in the usual way. It is now possible to state the characterization.

Proposition 4.16

For each $n \geq 0$, for each r such that $0 < r < 1$,

(a–c) *H is decidable with n mind changes starting with* $\begin{bmatrix} 0 \\ 1 \\ r \end{bmatrix}$ *given* $\mathcal{K}$

$$\Leftrightarrow \text{for each } h \in H,\ C_h \cap \mathcal{K} \in \begin{bmatrix} \Sigma[\mathcal{K}]^D_n \\ \Pi[\mathcal{K}]^D_n \\ \Delta[\mathcal{K}]^D_n \end{bmatrix}.$$

(d) *H is decidable with n mind changes given* $\mathcal{K}$
$\Leftrightarrow \forall h \in H,\ C_h \cap \mathcal{K} \in \Sigma[\mathcal{K}]^D_n \cup \Pi[\mathcal{K}]^D_n$.

Proof: (a) (⇒) Suppose that α decides$_C$ H given $\mathcal{K}$ in n mind changes starting with 0. Let

$$\alpha'(h, e) = \begin{cases} 1 & \text{if } \alpha(h, e) = 1 \\ 0 & \text{otherwise.} \end{cases}$$

α′ decides$_C$ H given $\mathcal{K}$ in n mind changes starting with 0, and α′ conjectures only 0s and 1s. Let $h \in H$. Define $\varepsilon \in \mathcal{O}(\varepsilon, n) \Leftrightarrow \alpha'$ changes its conjecture for h at least n times on ε, and define $\mathcal{P}(\varepsilon, n) \Leftrightarrow \alpha'$ changes its conjecture for h at most n times on ε. $\mathcal{O}(\varepsilon, n)$ is $\mathcal{K}$-open and $\mathcal{P}(\varepsilon, n)$ is $\mathcal{K}$-closed. First, consider the case when n is even. Then since α′ starts with conjecture 0 and never uses more than n mind changes over $\mathcal{K}$ on h, we have:

$$\begin{aligned}
\varepsilon \in C_h \cap \mathcal{K} &\Leftrightarrow \alpha' \text{ changes its mind some odd number of times} \\
&\quad \leq n-1 \text{ about } h \text{ on } \varepsilon \\
&\Leftrightarrow (\mathcal{O}(\varepsilon, 1) \,\&\, \mathcal{P}(\varepsilon, 1)) \vee (\mathcal{O}(\varepsilon, 3) \,\&\, \mathcal{P}(\varepsilon, 3)) \vee \cdots \\
&\quad \vee (\mathcal{O}(\varepsilon, n-1) \,\&\, \mathcal{P}(\varepsilon, n-1)) \\
&\Leftrightarrow [\mathcal{O}(\varepsilon, 1) \,\&\, \mathcal{P}(\varepsilon, 1)] \vee [\mathcal{O}(\varepsilon, 1) \,\&\, \mathcal{O}(\varepsilon, 3) \,\&\, \mathcal{P}(\varepsilon, 3)] \\
&\quad \vee [\mathcal{O}(\varepsilon, 1) \,\&\, \mathcal{O}(\varepsilon, 3) \,\&\, \mathcal{O}(\varepsilon, 5) \,\&\, \mathcal{P}(\varepsilon, 5)] \vee \cdots \\
&\quad \vee [\mathcal{O}(\varepsilon, 1) \,\&\, \mathcal{O}(\varepsilon, 3) \,\&\, \mathcal{O}(\varepsilon, 5) \,\&\, \cdots \\
&\quad \,\&\, \mathcal{O}(\varepsilon, n-3) \,\&\, \mathcal{O}(\varepsilon, n-1) \,\&\, \mathcal{P}(\varepsilon, n-1)] \\
&\Leftrightarrow \mathcal{O}(\varepsilon, 1) \,\&\, [\mathcal{P}(\varepsilon, 1) \vee [\mathcal{O}(\varepsilon, 3) \,\&\, [\mathcal{P}(\varepsilon, 3) \vee \cdots \\
&\quad \vee [\mathcal{O}(\varepsilon, n-1) \,\&\, \mathcal{P}(\varepsilon, n-1)]]]] \text{ (by factoring).}
\end{aligned}$$

So $C_h \cap \mathcal{K} \in \Sigma[\mathcal{K}]^D_n$. Now for the case in which n is odd. Since α starts out with conjecture 0 and never uses more than n mind changes over $\mathcal{K}$ on h, we have:

$$\begin{aligned}\varepsilon \in C_h \cap \mathcal{K} &\Leftrightarrow \alpha' \textit{ does not change its mind some even number of times} \\ &\quad \leq n \textit{ about } h \textit{ on } \varepsilon \\ &\Leftrightarrow \neg \mathcal{P}(\varepsilon, 0) \;\&\; \neg(O(\varepsilon, 2) \;\&\; \mathcal{P}(\varepsilon, 2)) \;\&\; \cdots \\ &\quad \& \neg(O(\varepsilon, n-1) \;\&\; \mathcal{P}(\varepsilon, n-1)) \\ &\Leftrightarrow O(\varepsilon, 1) \;\&\; [\mathcal{P}(\varepsilon, 1) \vee O(\varepsilon, 3)] \;\&\; [\mathcal{P}(\varepsilon, 3) \vee O(\varepsilon, 5)] \\ &\quad \& [\mathcal{P}(\varepsilon, 5) \vee O(\varepsilon, 7)] \;\&\; \cdots \;\&\; [\mathcal{P}(\varepsilon, n-3) \vee O(\varepsilon, n-1)] \\ &\Leftrightarrow O(\varepsilon, 1) \;\&\; [\mathcal{P}(\varepsilon, 1) \vee O(\varepsilon, 3)] \;\&\; [\mathcal{P}(\varepsilon, 1) \vee \mathcal{P}(\varepsilon, 3) \\ &\quad \vee O(\varepsilon, 5)] \;\&\; [\mathcal{P}(\varepsilon, 1) \vee \mathcal{P}(\varepsilon, 3) \vee \mathcal{P}(\varepsilon, 5) \vee O(\varepsilon, 7)] \\ &\quad \& \cdots \& [\mathcal{P}(\varepsilon, 1) \vee \mathcal{P}(\varepsilon, 3) \vee \cdots \vee \mathcal{P}(\varepsilon, n-5) \vee \cdots \\ &\quad \vee \mathcal{P}(\varepsilon, n-3) \vee O(\varepsilon, n-1)] \\ &\Leftrightarrow O(\varepsilon, 1) \;\&\; [\mathcal{P}(\varepsilon, 1) \vee [O(\varepsilon, 3) \;\&\; [\mathcal{P}(\varepsilon, 3) \vee \cdots \\ &\quad \& [\mathcal{P}(\varepsilon, n-3) \vee O(\varepsilon, n-1)]]]] \textit{ (by factoring).}\end{aligned}$$

So $C_\eta \cap \mathcal{K} \in \Sigma[\mathcal{K}]^D_n$.

($\Leftarrow$) Suppose that for each $h \in H$, $C_h \cap \mathcal{K} \in \Sigma[\mathcal{K}]^D_n$. Then $C_h \cap \mathcal{K}$ may be expressed in the forms

$$O_1 \cap [\mathcal{P}_2 \cup [O_3 \cap [\mathcal{P}_4 \cup \cdots [\mathcal{P}_{n-1} \cup O_n] \textit{ if } n \textit{ is odd},$$
$$O_1 \cap [\mathcal{P}_2 \cup [O_3 \cap [O_4 \cup \cdots [\mathcal{P}_{n-1} \cap O_n] \textit{ if } n \textit{ is even},$$

where each O_i is open and each $\mathcal{P}_i$ is closed. In either case, define α to conjecture 0 for h until O_1 is verified by the data, after which α says 1 until $\mathcal{P}_2$ is refuted by the data, after which α says 0 until O_3 is verified by the data, after which α says 1 until… α will succeed with at most n mind changes starting with 0.

(b) follows from (a) by duality, and (c) follows from (a) and (b) by proposition 3.12(b). (d) follows from (a), (b), and proposition 3.12(c). ■

The following proposition is an illustration of proposition 4.16 (*Fig. 4.39*).

Proposition 4.17

Let $\mathcal{K}$ be the set of all data streams that stabilize to some value. Then

(a) $D_{\mathcal{K}} \subset \Delta[\mathcal{K}]^B_2$.

(b) *For each n*, $\Sigma[\mathcal{K}]^D_n \subset \Sigma[\mathcal{K}]^D_{n+1}$.

(c) $D_{\mathcal{K}}$ *is the finitary Boolean closure of* $\Sigma[\mathcal{K}]^D_1$.

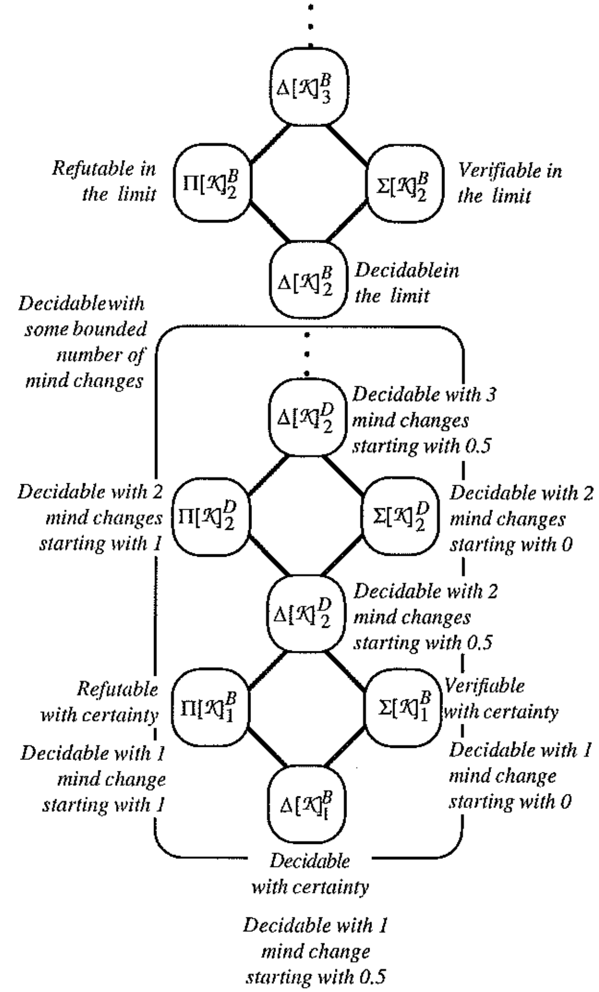


Figure 4.39

Proof: (b) To show that the inclusion is proper, define $\#0(\varepsilon)$ = the number of 0s occurring in ε. Let $H = \omega$ and define C as follows:

$$C(\varepsilon, h) = \begin{cases} \left(\bigvee_{k=1}^{h/2} \#0(\varepsilon) = 2k\right) \textit{ if } h \textit{ is even} \\ \left(\left(\bigvee_{k=1}^{(h-1)/2} \#0(\varepsilon) = 2k\right) \textit{ or } \#0(\varepsilon) \geq h+1\right) \textit{ if } h \textit{ is odd.} \end{cases}$$

h is readily seen to be decidable in h mind changes starting with 0. A simple demonic argument shows that h cannot be decided$_C$ with fewer mind changes starting with 0. Now apply proposition 4.16.

(a) Let C be defined as in the proof of (b). Add h^* to H, and define $C(\varepsilon, h^*) \Leftrightarrow C(\varepsilon, \varepsilon_1)$. Let α and n be given. The demon feeds $n + 1$ as the first datum and then forces α to change its mind concerning h^* at least $n + 1$ times, as in (b). By proposition 4.16, $C_{h^*} \cap \mathcal{K}$ is not in any class $\Sigma[\mathcal{K}]^D_n$, so $C_{h^*} \cap \mathcal{K} \notin D_{\mathcal{K}}$. But it is easy to decide$_C$ h^* in the limit. By proposition 4.10, $C_{h^*} \cap \mathcal{K} \in \Delta[\mathcal{K}]^B_2$.

(c) Each finite Boolean combination of open sets may be rewritten in disjunctive normal form. But by the closure of open [closed] sets under finite intersection, each disjunct can be rewritten in one of the following forms: $(O \cap \mathcal{P})$, O, or $\mathcal{P}$, where O is open and $\mathcal{P}$ is closed. Each disjunct of form $(O \cap \mathcal{P})$ is settled in two mind changes (say zero until O is verified; then say one until $\mathcal{P}$ is refuted). Each O and each $\mathcal{P}$ can be handled with one mind change, so the hole disjunction can be handled in some finite number of mind changes. Apply proposition 4.16. ■

10. A Demon-Oriented Characterization of *n*-Mind-Change Decidability

The proof of proposition 4.16 makes it more transparent how the scientist can succeed than how the demon can. What eactly does a demon need in order to be sure of fooling an arbitrary scientist n times about h given background assumptions $\mathcal{K}$? Let's consider the case of one mind change. We have seen in the proof of proposition 4.6 that the demon requires only the following pattern of "tracks" somewhere in $\mathcal{K}$ in order to fool at least twice a scientist who starts out conjecturing 0 (*Fig. 4.40*). To fool a scientist who starts out with conjecture 0 at least three times, the demon requires a slightly more complex system of tracks in $\mathcal{K}$ (*Fig. 4.41*).

The demon starts by steaming down the track depicted by the bold arrow. Eventually, the scientist says 1 because all tracks headed to the right make h true. At that point, the demon turns left onto a track for $\overline{C}_h$. He rolls down this track until the scientist says 0, which again must happen, else the scientist produces falsehood forever on this track. At this point, the demon again turns right, forcing the second mind change on pain, once again, of producing infinitely many false conjectures on a data stream in $\mathcal{K}$.

The demon's needs are becoming clear. What he requires is something like an infinite *feather*. The demon always starts out on the *shaft* of the feather, and then shunts back and forth along *barbs*, barbs of barbs, and so forth (*Fig. 4.42*). Each time the demon moves from a barb to a barb of a barb, the truth value of h changes. The *dimension* of the feather (the number of times one can turn

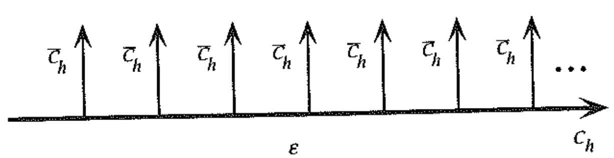


Figure 4.40

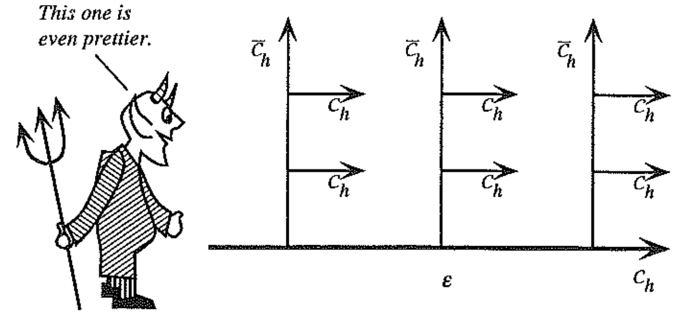


Figure 4.41

direction, traveling out from the shaft) determines how many times the demon can force the scientist to change his mind.

We may define n-feathers inductively, as follows. Let $\mathcal{K}, \mathcal{P} \subseteq \mathcal{N}$.

> $\mathcal{K}$ *is a* 1-*feather for* $\mathcal{P}$ *with shaft* $\varepsilon \Leftrightarrow \varepsilon \in \mathcal{P} \cap \mathcal{K}$.
>
> $\mathcal{K}$ *is an* $n + 1$-*feather for* $\mathcal{P}$ *with shaft* $\varepsilon \Leftrightarrow \varepsilon \in \mathcal{P} \cap \mathcal{K}$ *and*
>
> $\forall m\, \exists \varepsilon' \in \mathcal{K}$ *such that*
> $\varepsilon|m = \varepsilon'|m$ *and*
> $\mathcal{K}$ *is an n-feather with shaft* ε' *for* $\overline{\mathcal{P}}$.
>
> $\mathcal{K}$ *is an n-feather for* $\mathcal{P} \Leftrightarrow \exists\varepsilon$ *such that* $\mathcal{K}$ *is an n-feather for* $\mathcal{P}$ *with shaft* ε.

An $n + 1$ feather for $\mathcal{P}$ is a data stream for $\mathcal{P}$ that has n-feathers for $\overline{\mathcal{P}}$ branching off infinitely often (*Fig. 4.43*). An $n + 1$-feather for $\mathcal{P}$ has its shaft in $\mathcal{P}$, whereas an n-feather for $\overline{\mathcal{P}}$ has its shaft in $\overline{\mathcal{P}}$. We may now define the *feather dimension* of $\mathcal{K}$ for $\mathcal{P}$.

$$Dim_{\mathcal{P}}(\mathcal{K}) = n \Leftrightarrow \mathcal{K} \text{ is an } n\text{-feather for } \mathcal{P} \text{ and } \mathcal{K} \text{ is not an } n+1\text{-feather for } \mathcal{P}.$$

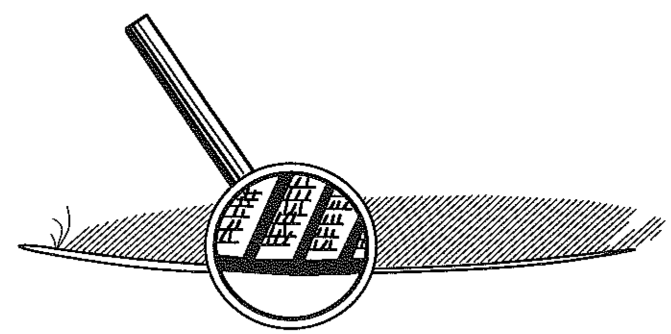


Figure 4.42

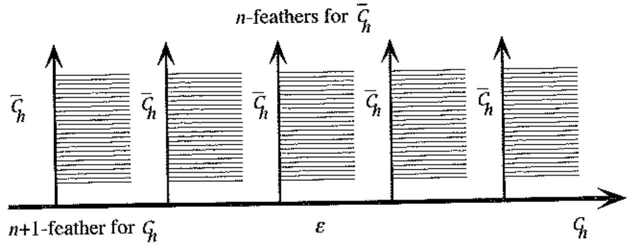


Figure 4.43

The obvious demonic argument yields that if $\mathcal{K}$ is an n-feather for C_h, then no scientist can decide h with n mind changes given $\mathcal{K}$, starting with conjecture 0. It is a bit less obvious that whenever $\mathcal{K}$ is not an n-feather for C_h, h is decidable with at most n mind changes given $\mathcal{K}$, starting with 0.

Proposition 4.18

(a) *H is decidable$_C$ given $\mathcal{K}$ in n mind changes starting with* 0 $\Leftrightarrow \forall h \in H$, *$\mathcal{K}$ is not an $n+1$-feather for C_h.*

(b) *H is decidable$_C$ given $\mathcal{K}$ in n mind changes starting with* 1 $\Leftrightarrow \forall h \in H$, *$\mathcal{K}$ is not an $n+1$-feather for $\overline{C_h}$.*

(c) *H is decidable$_C$ given $\mathcal{K}$ in n mind changes starting with* 0.5 $\Leftrightarrow \forall h \in H$, *$\mathcal{K}$ is not an $n+1$-feather for C_h or for $\overline{C_h}$.*

(d) *H is decidable given $\mathcal{K}$ in n mind changes* $\Leftrightarrow \forall h \in H$, *$\mathcal{K}$ is not an $n+1$-feather for C_h or $\mathcal{K}$ is not an $n+1$-feather for $\overline{C_h}$.*

It follows immediately from proposition 4.16 that:

Corollary 4.19

$C_h \cap \mathcal{K} \in \Sigma[\mathcal{K}]^D_n \Leftrightarrow$ *$\mathcal{K}$ is not an $n+1$-feather for C_h.*

$C_h \cap \mathcal{K} \in \Pi[\mathcal{K}]^D_n \Leftrightarrow$ *$\mathcal{K}$ is not an $n+1$-feather for $\overline{C_h}$.*

$C_h \cap \mathcal{K} \in \Delta[\mathcal{K}]^D_n \Leftrightarrow$ *$\mathcal{K}$ is not an $n+1$-feather for C_h and $\mathcal{K}$ is not an $n+1$-feather for $\overline{C_h}$.*

$C_h \cap \mathcal{K} \in \Sigma[\mathcal{K}]^D_n \cup \Pi[\mathcal{K}]^D_n \Leftrightarrow$ *$\mathcal{K}$ is not an $n+1$-feather for C_h or $\mathcal{K}$ is not an $n+1$-feather for $\overline{C_h}$.*

Proof: (a) & (b) ($\Rightarrow$) Prove the contrapositive by the obvious demonic argument.

($\Leftarrow$) Argument by induction on n. Base case for (b): Suppose that $\mathcal{K}$ is not a 1-feather for C_h. Then $C_h \cap \mathcal{K} = \varnothing$. Let $\varepsilon \in \mathcal{K}$. Then $\varepsilon \notin C_h$. So the trivial method always conjectures 0 no matter what succeeds in 0 mind changes. The base case for (a) is similar.

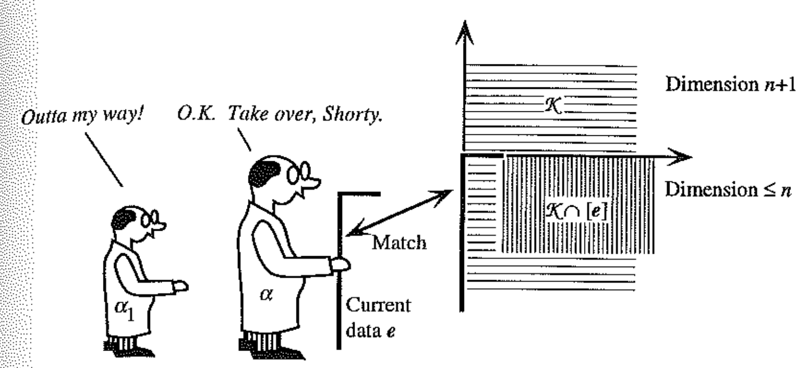


Figure 4.44

Now suppose (a) and (b) for each $m \leq n$. Suppose that for each $h \in H$, $\mathcal{K}$ is not an $n+2$-feather for C_h. Define the method (*Fig. 4.44*):

$$\alpha(h, e) = \begin{cases} 0 & \text{if } e = \mathbf{0} \\ 0 & \text{if } Dim_{C_h}(\mathcal{K} \cap [e]) > n \text{ and } Dim_{\overline{C_h}}(\mathcal{K} \cap [e]) > n \\ \alpha_1(h, e) & \text{if } Dim_{\overline{C_h}}(\mathcal{K} \cap [e]) \leq n \\ \alpha_0(h, e) & \text{otherwise (i.e., if } Dim_{C_h}(\mathcal{K} \cap [e]) \leq n). \end{cases}$$

where α_1 decides h with n mind changes starting with 1 given $\mathcal{K} \cap [e]$ and α_0 decides h with n mind changes starting with 0 given $\mathcal{K} \cap [e]$, as guaranteed by the induction hypothesis when $Dim_{C_h}(\mathcal{K} \cap [e]) \leq n$ or $Dim_{\overline{C_h}}(\mathcal{K} \cap [e]) \leq n$, respectively.

There are just two cases to consider.

(i) $\forall k$ *$Dim_{C_h}(\mathcal{K} \cap [\varepsilon|k]) > n$ and $Dim_{\overline{C_h}}(\mathcal{K} \cap [\varepsilon|k]) > n$.*

(ii) $\exists k$ *$Dim_{C_h}(\mathcal{K} \cap [\varepsilon|k] \leq n$ or $Dim_{\overline{C_h}}(\mathcal{K} \cap [\varepsilon|k]) \leq n$.*

Lemma 4.20

If (i) *then $\varepsilon \notin C_h$ and α stabilizes correctly to 0 on h, ε with no mind changes.*

Proof: For suppose that $\varepsilon \in C_h$. Then since for each k, $Dim_{\overline{C_h}}(\mathcal{K} \cap [\varepsilon|k]) > n$, we have that for each k, there is an ε' such that $\mathcal{K} \cap [\varepsilon|k]$ is an $n+1$ feather for $\overline{C_h}$ with shaft ε' and $\varepsilon'|k = \varepsilon|k$. Hence, $\mathcal{K}$ is an $n+2$-feather for C_h with shaft ε, contrary to the hypothesis. Now observe that for each k, the second clause of α is satisfied on $\varepsilon|k$, so for each k, $\alpha(h, \varepsilon|k) = 0$. ■

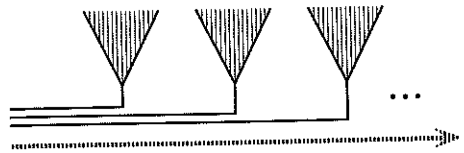


Figure 4.45

Lemma 4.21

If (ii) *then* α *succeeds in* $n + 1$ *mind changes.*

Proof: Let m be the least k such that $Dim_{C_h}(\mathcal{K} \cap [\varepsilon|k]) \leq n$ or $Dim_{\overline{C_h}}(\mathcal{K} \cap [\varepsilon|k]) \leq n$. Suppose that $Dim_{\overline{C_h}}(\mathcal{K} \cap [\varepsilon|k]) \leq n$. Then $\alpha(h, \varepsilon|m') = 0$, for all $m' < m$ and $\alpha(h, \varepsilon|m'') = \alpha_1(\varepsilon|m'')$ for all $m'' \geq m$. Since α_1 decides C_h over $\mathcal{K} \cap [\varepsilon|m]$ in n mind changes starting with 1 and α outputs only 0 prior to invoking α_1, α succeeds in $n + 1$ mind changes starting with 0. In case $Dim_{C_h}(\mathcal{K} \cap [\varepsilon|k]) \leq n$, we have a similar situation, except that α_0 starts with 0, so α succeeds in n mind changes. ■

The induction for (b) is similar, except that the method employed starts out with 1s until the submethods α_0 and α_1 are invoked. (c) and (d) may be obtained just as in proposition 4.16(c) and (d). ■

By corollary 4.19, the feather perspective and the *D*-hierarchy perspective coincide exactly. It is revealing to see how the correspondence works by constructing feathers out of intersections and unions of open and closed sets. The *D*-hierarchy starts at level Σ_1^D with open hypotheses. The open set depicted in (*Fig. 4.45*) clearly yields a 1-feather which affords the demon the chance to fool a scientist who starts with 0 once, and a scientist who starts with 1 twice.

To generate a properly Π_1^D set, we take the complement of the open set to arrive at a closed set, which allows the demon to fool twice a scientist who starts with conjecture 0, and to fool once a scientist who starts with conjecture 1 (*Fig. 4.46*).

To obtain a Σ_2^D set, we add an open set to out closed set. This yields a 2-feather, which affords the demon an opportunity to fool twice a scientist who starts with conjecture 1, and to fool three times a scientist who starts with conjecture 0 (*Fig. 4.47*).

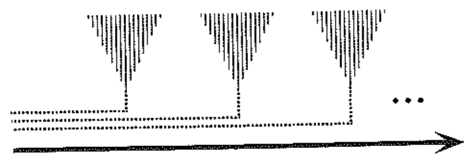


Figure 4.46

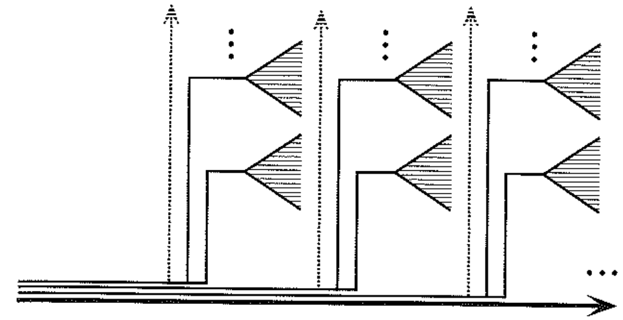


Figure 4.47

Complementing this set yields Π_2^D set and a dual 2-feather (*Fig. 4.48*). Now we are again free to add a dimension by augmenting this set with an open set. By successive complementations and open set additions, we can build a feather of arbitrary finite dimension.

11. Characterization of Gradual Assessment

The following characterization introduces a complete architecture that generalizes the bumping pointer architecture. Noteworthy features of this result are that Σ and Π switch sides regarding verification and refutation and decidability drops a level below $\Delta[\mathcal{K}]_3^B$.

Proposition 4.22

$$H \text{ is } \begin{bmatrix} verifiable_C \\ refutable_C \\ decidable_C \end{bmatrix} \text{ gradually given } \mathcal{K}$$

$$\Leftrightarrow \text{for each } h \in H,\ C_h \cap \mathcal{K} \in \begin{bmatrix} \Pi[\mathcal{K}]_3^B \\ \Sigma[\mathcal{K}]_3^B \\ \Delta[\mathcal{K}]_2^B \end{bmatrix}.$$

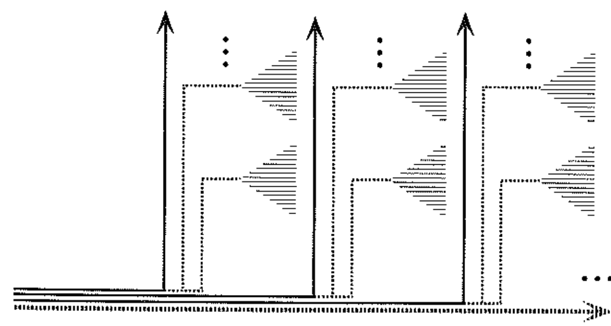


Figure 4.48

Proof: As usual, it suffices to prove the verification case. The refutation case follows by duality, and the decidability case follows from proposition 3.13(b).

($\Leftarrow$) Suppose that for each $h \in H$, $C_h \cap \mathcal{K} \in \Pi[\mathcal{K}]_3^B$. Then $C_h \cap \mathcal{K}$ is the intersection of some countable sequence $\mathcal{S}[h]_0, \mathcal{S}[h]_1, \ldots, \mathcal{S}[h]_n, \ldots$ of $\Sigma[\mathcal{K}]_2^B$ sets. Let $\mathcal{G}$ and H' be such that for each i there is an $h_i \in H'$ such that $\mathcal{G}_{h_i} = \mathcal{S}[h]_i$. Since each $\mathcal{G}_{h_i}$ is $\Sigma[\mathcal{K}]_2^B$, we have by proposition 4.10 that there is a method α_h that verifies$_{\mathcal{G}}$ H' in the limit given $\mathcal{K}$. Define method α as follows. On hypothesis $h \in H$ and finite data segment e of length k, α considers the sequence of conjectures $\alpha_h(h_0, e), \alpha_h(h_1, e) \cdots \alpha_h(h_k, e)$ until some conjecture less than 1 is produced. Let $n \leq k$ be the number of consecutive 1s occurring in this sequence. α conjectures $1 - 2^{-n}$ (*Fig. 4.49*). Let $h \in H$, $\varepsilon \in \mathcal{K}$. Suppose $\varepsilon \in C_h$. Then for each $h_i \in H'$, $\varepsilon \in \mathcal{G}_{h_i}$. Then since α_h verifies$_{\mathcal{G}}$ H' in the limit given $\mathcal{K}$, we have that for each h_i, α_h stabilizes to 1 on ε for h_i. Thus, for each i, there is a time after which α_h has stabilized to 1 for $h_0, \ldots, h_i$. Thereafter, α's conjecture never drops below $1 - 2^{-i}$. Since i is arbitrary, α correctly approaches 1 on ε for h. Suppose $\varepsilon \notin C_h$. Then for some $h_i \in H'$, $\varepsilon \notin \mathcal{G}_{h_i}$. Thus, α_h does not stabilize to 1 for h_i on ε, so infinitely often, the conjecture of α for h *on* ε is no greater than $1 - 2^{-i}$. Thus α correctly fails to approach 1 for h *on* ε.

($\Rightarrow$) Suppose that α verifies$_C$ H gradually given $\mathcal{K}$. Then for each $h \in H$, we have:

$$\forall \varepsilon \in \mathcal{K}, \quad C(\varepsilon, h) \Leftrightarrow \forall \text{ rational } r > 0\, \exists n\, \forall m \geq n\; 1 - \alpha(h, \varepsilon|m) < r.$$

The set $\mathcal{A}_h(m, r)$ of all $\varepsilon \in \mathcal{K}$ such that $1 - \alpha(h, \varepsilon|m) < r$ is $\mathcal{K}$-clopen. For all $\varepsilon \in \mathcal{K}$,

$$\forall m \geq n,\ 1 - \alpha(h, \varepsilon|m) < r$$
$$\Leftrightarrow \varepsilon \in (\mathcal{A}_h(n, r) \cap \mathcal{A}_h(n+1, r) \cap \mathcal{A}_h(n+2, r) \cap \cdots)$$

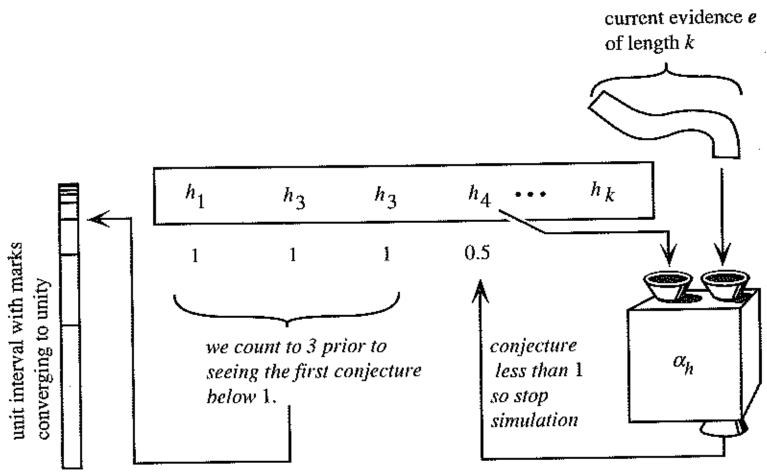


Figure 4.49

Let $\mathcal{B}_h(n, r)$ denote this countable intersection. $\mathcal{B}_h(n, r) \in \Pi[\mathcal{K}]_1^B$. Then for each $\varepsilon \in \mathcal{K}$ we have:

$$\exists n\, \forall m \geq n,\ 1 - \alpha(h, \varepsilon|m) < r$$
$$\Leftrightarrow \varepsilon \in (\mathcal{B}_h(0, r) \cup \mathcal{B}_h(1, r) \cup \mathcal{B}_h(2, r) \cup \cdots)$$

Let $\mathcal{D}_h(r)$ denote this union. $\mathcal{D}_h(r) \in \Sigma[\mathcal{K}]_2^B$. Let $r_0, r_1, \ldots, r_n, \ldots$ be an enumeration of the rationals greater than 0. Finally, for each $\varepsilon \in \mathcal{K}$ we have:

$$C_h = (\mathcal{D}_h(r_0) \cap \mathcal{D}_h(r_1) \cap \mathcal{D}_h(r_2) \cap \cdots).$$

Thus, $C_h \in \Pi[\mathcal{K}]_3^B$. ■

The complete assessment architecture introduced in the proof of this theorem may be referred to as the *consecutive ones* architecture, since it counts the consecutive 1s output by a simulated limiting verifier method to produce its conjecture.

12. The Levels of Underdetermination

The results of this chapter are summarized in *Figure 4.50*. Some hypotheses concerning the divisibility of matter are located in their proper places. The dots at the top of the diagram indicate that topological complexity continues upward forever (and then for infinitely many more eternities), even though our intuitions about what would count as limiting success for science pretty well peter out by level 3.

The characterizations bear on some traditional issues of empiricism. Recall that the empiricist draws a line in the sand so that hypotheses that fall on one side are gibberish and hypotheses that fall on the other are scientific. Empiricists have always had trouble deciding where the line should be, however. In his influential article on this subject, C. Hempel proposed the principle that empirical significance or meaningfulness should be closed under finitary Boolean operations.[16] He rejected $\Pi[\mathcal{K}]_1^B$ and $\Sigma[\mathcal{K}]_1^B$ as characterizations of empirical significance because neither class is closed under negation. B. Van Fraassen's (1980) version of instrumentalism countenances belief when there is no global underdetermination, but prohibits belief when there is, without regard to the level of local underdetermination. If a dichotomy has to be drawn, then perhaps it would be better to draw it at $\Delta[\mathcal{K}]_2^B$. This class contains both of the classes entertained by Hempel and is also closed under finitary Boolean operations, as Hempel required. And unlike Van Fraassen's line between local and global underdetermination, $\Delta[\mathcal{K}]_2^B$ characterizes both gradual decidability and decidability in the limit—a natural place for line drawing if lines must be drawn.[17]

[16] Hempel (1965): 102.
[17] Kugel (1977) emphasizes the importance of $\Delta[\mathcal{K}]_2^B$ for inductive inference.

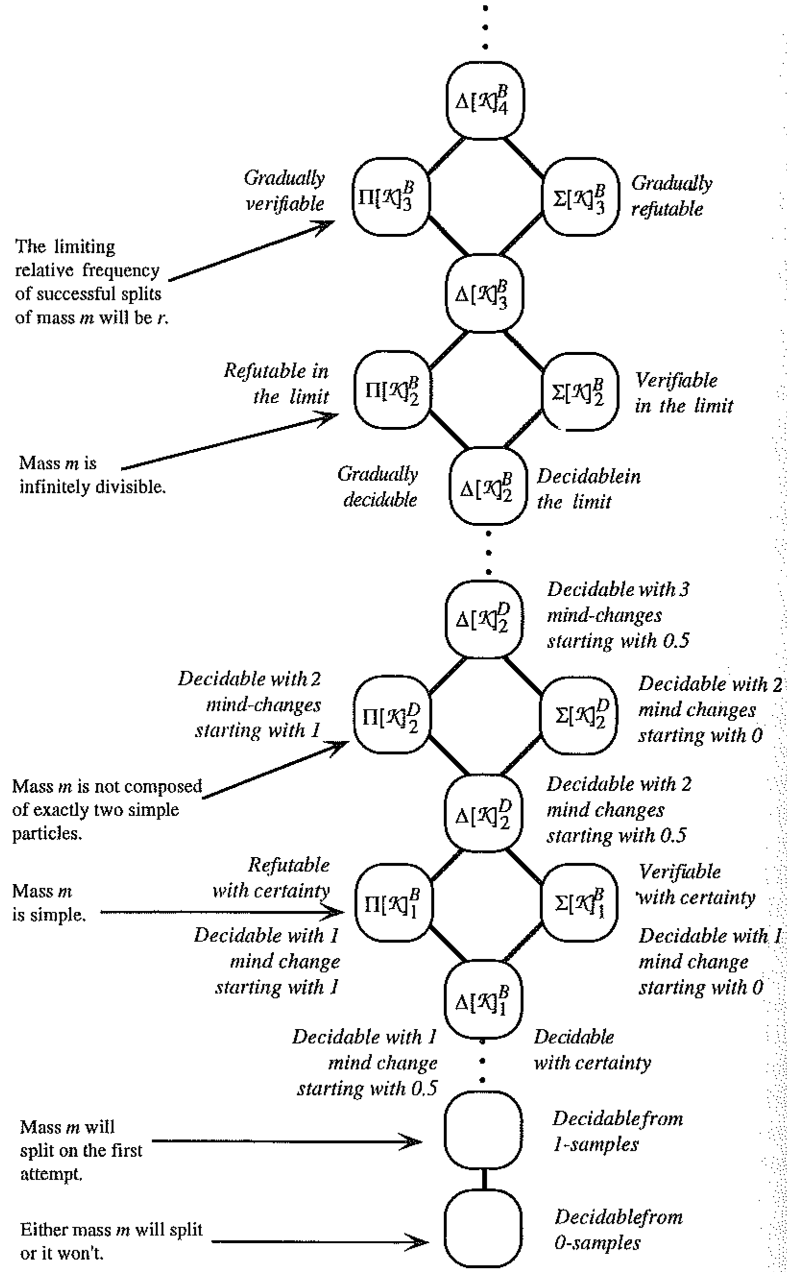


Figure 4.50

Someone could well demand that if we are going to drop the criterion of demarcation from global determination to $\Delta[\mathcal{K}]^B_2$, then we ought to drop it still further, for $\Delta[\mathcal{K}]^B_2$ characterizes success for *ideal* rather than for computational agents. If we do not entertain the fiction that we can scan the entire data

stream in an instant, then why should we entertain the fiction that we can decide all formal questions in an instant? This point is well taken, and will be explored in detail in chapter 7.

Exercises

4.1. Prove that the restriction of a topological space is a topological space.

4.2. Verify the closure laws summarized in *Figure 4.51*, for each $n > 0$.

4.3. Let $\mathcal{O}$ be open the let $\mathcal{P}$ be closed. Verify the following closure laws:

(a) *For each* $n > 0$:
- *if* $\mathcal{S} \in \Pi^D_n$ *then* $\bar{\mathcal{S}} \in \Sigma^D_n$. $\mathcal{S} \cup \mathcal{P} \in \Pi^D_n$; $\mathcal{S} \cap \mathcal{O} \in \Sigma^D_{n+1}$:
- *if* $\mathcal{S} \in \Sigma^D_n$ *then* $\bar{\mathcal{S}} \in \Pi^D_n$; $\mathcal{S} \cap \mathcal{O} \in \Sigma^D_n$; $\mathcal{S} \cup \mathcal{P} \in \Pi^D_{n+1}$.

(b) *For each odd* $n > 0$:
- *if* $\mathcal{S} \in \Pi^D_n$ *then* $\mathcal{S} \cap \mathcal{P} \in \Pi^D_n$; $\mathcal{S} \cup \mathcal{O} \in \Pi^D_{n+1}$:
- *if* $\mathcal{S} \in \Sigma^D_n$ *then* $\mathcal{S} \cup \mathcal{O} \in \Sigma^D_n$; $\mathcal{S} \cap \mathcal{P} \in \Sigma^D_{n+1}$.

(c) *For each even* $n > 0$,
- *If* $\mathcal{S} \in \Pi^D_n$ *then* $\mathcal{S} \cup \mathcal{O} \in \Pi^D_n$; $\mathcal{S} \cap \mathcal{P} \in \Pi^D_{n+1}$:
- *If* $\mathcal{S} \in \Sigma^D_n$ *then* $\mathcal{S} \cap \mathcal{P} \in \Sigma^D_n$; $\mathcal{S} \cup \mathcal{O} \in \Sigma^D_{n+1}$.

4.4. Show that the hypothesis $LRF(o)$ that a real-valued limiting relative frequency for outcome o exists is gradually verifiable. (Hint: a sequence ζ of rationals is a *Cauchy sequence* just in case for each rational $r > 0$ there is an n such that for all $m, m' \geq n$, $|\zeta_m - \zeta_{m'}| < r$. Recall from analysis that a sequence of rationals approaches a real if and only if it is Cauchy. Now design a method that can verify the hypothesis that the sequence of relative frequencies of o in ε is Cauchy.)

4.5. Prove an analogue of proposition 4.7 for refutation and decision with certainty.

	countable union	countable intersection	complement	finite intersection	finite union
Σ^B_n	yes	no	no	yes	yes
Δ^B_n	no	no	yes	yes	yes
Π^B_n	no	yes	no	yes	yes

Figure 4.51

4.6. A collection Θ of subsets of $\mathcal{N}$ is an *ideal* $\Leftrightarrow$

(i) $\varnothing \in \Theta$,

(ii) *if* $\mathcal{S} \in \Theta$ *and* $\mathcal{P} \subseteq \mathcal{S}$ *then* $\mathcal{P} \in \Theta$, *and*

(iii) *if* $\mathcal{S} \in \Theta$ *and* $\mathcal{P} \in \Theta$ *then* $\mathcal{S} \cup \mathcal{P} \in \Theta$.

Show that for some h, C, the collection $\Gamma = \{\mathcal{K}: h$ is verifiable$_C$ in the limit given $\mathcal{K}\}$ is not an ideal. Which of the conditions is Γ guaranteed to satisfy?

4.7. Find a data-minimal architecture for decision with n mind changes.

4.8. Show that each opportunistic limiting verifier is *consistent* in the sense that it stabilizes to 0 as soon as h is refuted by the data relative to $\mathcal{K}$, and it stabilizes to 1 as soon as h is entailed by the data relative to $\mathcal{K}$. (Hint: if $\mathcal{K} \cap [e] \subseteq C_h \cap [e]$, then $\tau_{h,e}$ does not exist, and if $\mathcal{K} \cap C_h \cap [\varepsilon|n] = \varnothing$, then $\tau_{h,\varepsilon|m}$ does not exist if $m \geq n$.) Show that consistency is necessary for weak data minimality among limiting verifiers for a given inductive problem.

4.9. Show that if $\mathcal{P} \in \Sigma[\mathcal{K}]_2^B$ and $\mathcal{S} \cap \mathcal{P} = \varnothing$ and $\mathcal{S} \in \Delta[\mathcal{K} \cup \mathcal{S}]_2^B$, then $\mathcal{P} \in \Sigma[\mathcal{K} \cup \mathcal{S}]_2^B$. (Hint: use a limiting decider for $\mathcal{S}$ given $\mathcal{K} \cup \mathcal{S}$ to determine whether you are in $\mathcal{S}$, and either output 0 or switch to a limiting verifier of $\mathcal{P}$ given $\mathcal{K}$ according to the current conjecture of the decider.) Show by example that the result cannot be improved to the case in which $\mathcal{S} \in \Sigma[\mathcal{K} \cup \mathcal{S}]_2^B - \Delta[\mathcal{K} \cup \mathcal{S}]_2^B$. (Another hint: let $\mathcal{P}$ = the set of all sequences in which infinitely many 1s occur, let $\mathcal{K} = \mathcal{P}$ and let $\mathcal{S} = \mathcal{N} - \mathcal{K}$).

4.10. Derive the upper bounds in exercise 3.1 by using the characterization theorems of this chapter. (Hint: construct C_h as a Borel set of the appropriate complexity.)

*4.11. I have assumed that the stages of inquiry are discrete. But we can imagine an inductive method implemented as an electronic device whose input at time t is the intensity of current in an input wire at t and whose conjecture at t is the position of a needle on a meter (scaled from 0 to 1) at t (actually a bit after t to allow the signal to travel through the circuits of the machine). Say that the method stabilizes just in case its value is fixed after some finite time interval. Similar definitions can be given in the bounded mind change and gradual cases. Extend the characterization results of this chapter to continuous methods in the paradigm just described, taking a fan to be the set of all unbounded extensions of a finite segment. Be careful when defining the Borel hierarchy, since open sets will not be countable unions of fans in this setting. What happens to the characterizations if the method is not assumed to be continuous?

*4.12. In the paradigm described in the preceding question, show that the set of all solutions to an ordinary, first-order differential equation is always closed in the corresponding topology. Discuss the methodological significance of this fact.

*4.13 (The epistemology of real analysis). Topologize the real line $\mathfrak{R}$ by taking open intervals with rational endpoints as basic open sets. Think of each real number as a complete "possible world." Then a *data stream* for $r \in \mathfrak{R}$ is a downward nested collection $\{I_k: k \in \omega\}$ of closed intervals such that $\bigcap\{I_k: k \in \omega\} = \{r\}$. Let $\mathcal{H}, \mathcal{K} \subseteq \mathfrak{R}$. Think of $\mathcal{H}$ as a hypothesis and of $\mathcal{K}$ as background knowledge. An *inductive method* α maps finite sequences of closed intervals into $[0, 1] \cup$ '!'. Define:

α *verifies* $\mathcal{H}$ *in the limit given* $\mathcal{K} \Leftrightarrow \forall r \in \mathcal{K}$, $\forall$ *data stream* $\{I_k: k \in \omega\}$ *for* r,
$r \in \mathcal{H} \Leftrightarrow \exists n\, \forall m \geq n,\ \alpha(I_m) = 1$,

and similarly for the other notions of reliability.

(i) Characterize the various notions of reliability in this paradigm.
(ii) Show that only $\varnothing$ and $[0, 1]$ are decidable with certainty given $[0, 1]$.
(iii) Show that each $\mathcal{H}$ is decidable with certainty given $\mathfrak{R} - \text{bdry}(\mathcal{H})$.
(iv) Contrast result (ii) with the fact that nontrivial subsets of 2^ω are decidable with certainty in the usual, Baire-space paradigm. Explain the difference.

*4.14. $\mathcal{K} \subseteq \mathcal{N}$ *is compact* $\Leftrightarrow$

for each collection Γ *of open sets such that* $\bigcup\Gamma = \mathcal{K}$, *there is a finite* $\Gamma' \subseteq \Gamma$ *such that* $\bigcup\Gamma' = \mathcal{K}$.

(i) Show that if $\mathcal{K}$ is compact, then $\forall \mathcal{H} \subseteq \mathcal{K}$, if $\mathcal{H}$ is $\mathcal{K}$-clopen then $\exists n$ such that $\mathcal{H}$ is n-uniform in $\mathcal{K}$. Rephrase this fact in methodological terms.
(ii) Using (a), show that $\mathcal{N}$ is not compact.
(iii) Show that $2^\omega \subseteq \mathcal{N}$ is compact.
(iv) Show that no properly open subset of 2^ω is compact.

*4.15 (Subjunctive learning theory). Recall the discussion of subjunctive reliability in chapter 2. We can investigate this proposal using techniques developed in chapter 4. Think of $\mathcal{K}$ as the set of possible worlds. A notion of "closeness" of possible worlds is a space $\mathfrak{K} = (\mathcal{K}, K)$, where K is some collection of subsets of $\mathcal{K}$. Let $\mathcal{A}, \mathcal{B} \subseteq \mathcal{N}$. Then define the subjunctive conditional $(\mathcal{A} \mapsto \mathcal{B})_{\mathfrak{K}} \subseteq \mathcal{K}$ as follows:

$\forall \varepsilon \in \mathcal{K}$, $\varepsilon \in (\mathcal{A} \mapsto \mathcal{B})_{\mathfrak{K}} \Leftrightarrow$ *either*
(i) $\forall \mathcal{S} \in K, \mathcal{S} \cap \mathcal{A} = \varnothing$ *or*
(ii) $\exists \mathcal{S} \in K$ *such that* $\mathcal{S} \cap \mathcal{A} \neq \varnothing$ *and* $\mathcal{S} \cap \mathcal{A} \subseteq \mathcal{B}$.

We may think of $(\mathcal{A} \mapsto \mathcal{B})_{\mathfrak{K}}$ as saying "if $\mathcal{A}$ were the case then $\mathcal{B}$ would be the case."[18] Now define:

α *verifies*$_C$ h *in the limit given* $\mathfrak{K}$ *on* $\varepsilon \Leftrightarrow$

(a) $(C_h \mapsto \alpha(h, \_)$ *stabilizes to* $1)_{\mathfrak{K}}$ *and*
(b) $(\overline{C_h} \mapsto \alpha(h, \_)$ *does not stabilize to* $1)_{\mathfrak{K}}$.

$\mathcal{LV}_C[h]_{\mathfrak{K}} = \{\varepsilon: \exists\alpha$ *such that* α *verifies*$_C$ h *in the limit given* $\mathfrak{K}$ *on* $\varepsilon\}$.

Similar definitions can be given for the other senses of reliability. Let $\mathcal{DC}$, $\mathcal{VC}$, and $\mathcal{RC}$ correspond to decidability, verifiability, and refutability with certainty, respectively.

[18] The definition is due to Lewis (1973). Lewis adds extra constraints on $\mathfrak{K}$.

Recall that $\mathfrak{N}$ is the Baire space $(\mathcal{N}, B)$. Let $\mathfrak{N} + \mathfrak{S}$ be the result of adding all singletons to $\mathfrak{N}$. Let $\mathfrak{D} = (\mathcal{N}, 2^{\mathcal{N}})$ be the discrete topology on $\mathcal{N}$. Let $\mathfrak{T} = (\mathcal{N}, \{\varnothing, \mathcal{N}\})$ be the trivial topology on $\mathcal{N}$. Let $\mathfrak{W} = (\mathcal{N}, \Delta)$, where Δ is the set of all closed balls under the distance function $\rho(\varepsilon, \tau) = \sup_n |\varepsilon_n - \tau_n|$, where the distance is ∞ if the sup does not exist.

(i) Show that $\forall \varepsilon$, h is verifiable$_C$ in the limit given $\mathfrak{T}$ on $\varepsilon \Leftrightarrow h$ is verifiable$_C$ in the limit given $\mathcal{N}$.

(ii) Let $C_h = \{\zeta\}$. Show that h is not verifiable$_C$ with certainty given $\mathfrak{N}$, $\mathfrak{T}$, or $\mathfrak{W}$, on ζ, whereas h is decidable$_C$ with certainty given $\mathfrak{D}$ on ζ. Show that a structure $\mathfrak{K}$ satisfies "h is decidable$_C$ with certainty given $\mathfrak{K}$ on ζ" just in case $\exists n$ such that $\zeta \in (\overline{\{\zeta\}} \mapsto \{\zeta\}$ is refuted by stage $n)_{\mathfrak{K}}$. Do you believe that if a universal hypothesis is actually true then there is some fixed time such that if the hypothesis were false it would be refuted by that time? What does this say about the prospects for Nozick's program (cf. section 2.7) as a response to the problem of local underdetermination? Show that for every $\mathfrak{K}$, $\zeta \in (\overline{\{\zeta\}} \mapsto \exists n$ such that $\{\zeta\}$ is refuted by stage $n)_{\mathfrak{K}}$. Why does the placement of the existential quantifier matter so much? (The former placement is said to be *de re*, or *of things*, while the latter placement is said to be *de dicto*, or *of words*.)

(iii) Show that for each ε, h_{fin} is not refutable in the limit given $\mathfrak{N}$, $\mathfrak{T}$, or $\mathfrak{W}$ on ε.

(iv) Show that:

(a) $\mathcal{DC}_C[h]_{\mathfrak{N}+\mathfrak{S}} = \mathcal{DC}_C[h]_{\mathfrak{N}} = C_h - bdry(C_h)$.
(b) $\mathcal{VC}_C[h]_{\mathfrak{N}+\mathfrak{S}} - \mathcal{DC}_C[h]_{\mathfrak{N}+\mathfrak{S}} = bdry(C_h) - C_h$.
(c) $\mathcal{RC}_C[h]_{\mathfrak{N}+\mathfrak{S}} - \mathcal{DC}_C[h]_{\mathfrak{N}+\mathfrak{S}} = bdry(C_h) - \overline{C_h}$.
(d) $int(\overline{C_h}) = \varnothing \Rightarrow \mathcal{DC}_C[h]_{\mathfrak{N}+\mathfrak{S}} = C_h$ (*i.e.*, "*h is decidable with certainty*" is the *same proposition* as h).

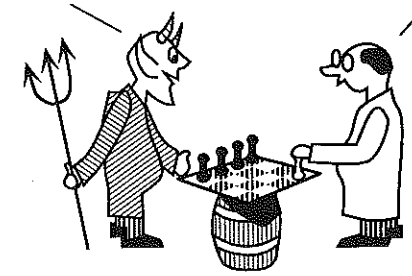
(v) What happens to (b) and (c) in $\mathfrak{N}$?

(vi) Relate results (iv a–c) to the discussion in this chapter in which it was claimed that the problem of induction arises in the boundaries of hypotheses.

(vii) What is the significance of (d) for naturalistic epistemology, the proposal that we should use empirical inquiry to find out if we are or can be reliable?

5

Reducibility and the Game of Science



1. Introduction

Many of the negative arguments considered so far have involved demons who attempt to feed misleading evidence to the inductive method α, as though science were an ongoing game between the demon and the scientist, where winning is determined by what happens in the limit. This game-theoretic construal of inquiry is familiar in skeptical arguments from ancient times, and it is the purpose of this chapter to explore it more explicitly than we have done so far. One reason it is interesting to do so is that the theory of infinite games is central to contemporary work in the foundations of mathematics, so that the existence of winning strategies for inductive demons hinges on questions unanswered by set theory.[1] In set theory, infinite games are known to be intimately related to the notion of continuous reducibility among problems and inductive methods turn out to determine continuous reductions. It is therefore natural to consider games and continuous reducibility all at once. The discussion of continuous reducibility in this chapter will help to underscore the analogy between reliable inductive inference on the one hand and ordinary computation on the other, which will be developed in the next chapter.

[1] Material presented in this chapter assumes some elementary issues in set theory, but may be skipped without loss of comprehension in the chapters that follow.