Theory Choice, Theory Change, and Inductive Truth-Conduciveness

Konstantin Genin

April 24, 2015

Contents

To Kevin Kelly, whom I am proud to consider a mentor.

Acknowledgements

This work would not have been possible without the encouragement, support and inspiration of my advisor, Kevin Kelly. Many others have also been instrumental. I am grateful to Hanti Lin for setting such a high standard, and to Thomas Icard, for indulging me in long discussions. I would also like to thank our friends in Amsterdam, Alexandru Baltag, Sonja Smets and Nina Gierasimczuk for their assistance and hospitality. I am gifted with many talented and inspiring colleagues. I would also like to thank the Philosophy Department for supporting the travel that has made these connections possible.

Chapter 1

Introduction

This work is concerned with three things: the synchronic norms of theory *choice*, the diachronic norms of theory *change*, and the justification of these norms by their *reliability,* or *truth-conduciveness.*

Synchronic norms of theory choice restrict the theories one can choose in light of given, empirical information. They are the traditional purview of the philosophy of science. But how do they facilitate arrival at true theories? Are they better than other means toward that end and, if so, in what sense? For example, it is widely agreed that scientific theory choice proceeds in accordance with a bias toward simpler or more sharply testable theories. However, the truth might not be simple, in which case that bias would probably lead to error. So how can one argue that the characteristic scientific biases toward simplicity or testability are truth-conducive, without begging the question with material assumptions?

Diachronic norms of theory change restrict how one should change one's current beliefs in light of new information. The crucial difference from synchronic norms is dependency upon one's prior beliefs. Such norms are studied propositionally in belief revision theory and non-monotonic logic and quantitatively in Bayesian epistemology. The question of truth-conduciveness arises for such norms, just as it does for norms of theory choice. Furthermore, there is the additional question how diachronic norms relate to the more traditional, synchronic ones.

If one insists upon an overly strict standard of truth-conduciveness in inductive contexts, the crucial question of truth-conduciveness becomes intractable. Theoretical virtues are not guaranteed to *indicate* truth with a low chance of error, the way litmus paper indicates pH—inductive inferences in accordance with the rationality principles are subject to arbitrarily high chances of error, because the available information can probably be arbitrarily similar, regardless of which conclusion is true.

Maturity is a matter of ceasing to demand the impossible. In that spirit, it makes more sense to adjust the standards of truth-conduciveness to the intrinsic difficulty of the inference problem one faces—a view we call *feasibility-contextualism.* Feasibility contextualism presupposes a spectrum of alternative concepts of truth-conduciveness. Just such a spectrum is routinely studied in the subject known as formal learning theory, which studies concepts of truth-conduciveness ranging from the very strict standard of truth-indicativeness to the very weak standard of mere convergence to the truth in the limit. Neither extreme suffices for epistemic justification. The former is too strict to apply without question-begging assumptions and the latter, notoriously, mandates no short-run norms at all, since convergence in the limit is compatible with any inductive behavior whatever in the short run. In between these extremes, however, are more nuanced concepts of optimally *direct* convergence to the truth. We consider two such concepts in this paper: convergence with minimal reversals of opinion and convergence with minimal cycles of opinion. Since a strategy is conducive to a goal insofar as it leads as directly as possible to the goal, we view directness of approach to the truth as *constitutive* of truth-conduciveness and, hence, of epistemic justification, rather than as an auxiliary, "pragmatic" consideration.

In this work, we show how the rationality principles of belief revision can be thought of as truth-conducive norms of maximally direct convergence in the sense just described. Furthermore, we prove that preferring simple, falsifiable theories (Ockham's razor) is a necessary condition for achieving optimally truth-conducive performance. The results forge deep and, perhaps, surprising connections between synchronic rationality norms, diachronic rationality norms, and the truth-conduciveness of both.

1.1 Reliability and the Norms of Theory Choice

It is commonplace to observe that science seeks true theories about the world. But that banal observation raises a profound question about scientific method: how, and in what sense, do such hallmark scientific values as simplicity, precision, scope, and novelty help one find true theories? To demand an answer to that question is to demand an *epistemic justification* of scientific values (Hempel, 1983; Laudan, 1984; Douglas, 2000). One goal of this work is to provide such a justification for Ockham's razor, the pervasive scientific bias in favor of simple theories.

An epistemic justification of Ockham's razor is traditionally understood to be a demonstration that simpler theories are *more likely to be true*.[1] That narrow, synchronic concept of justification makes a hopeless conundrum out of Ockham's razor. It has led theorists into metaphysical speculations less plausible than the scientific conclusions they are meant to justify. Both Kepler and Dirac expressed the conviction that Nature loves mathematical elegance, and that physicists ought to adopt the same passion to more surely uncover her secrets (Morrison, 2007; Dirac, 1939). It has led others to give up on epistemic justification entirely: "no one has shown that any of these rules is more likely to pick out

[1]Baker (2013) is representative: "justifying an epistemic principle requires answering an epistemic question: why are parsimonious theories more likely to be true?" The demand is trivial if "likely" is understood subjectively, so we understand it objectively, in the sense of a guaranteed low chance of error.

true theories than false ones. It follows that none of these rules is epistemic in character" (Laudan, 2004).

There is no shortage of non-epistemic justifications. Predictive accuracy, not truth, is the target of frequentist justifications: a bias toward simple theories prevents over-fitting and improves prediction when extrapolating from small samples (Akaike, 1974; Forster and Sober, 1994; Vapnik, 1998). But while they may be more predictively accurate at small sample sizes, simple theories are not more likely to be true in any objective sense of likelihood. Akaike's method does not even converge to the true theory in the limit of infinite data. Some frequentists take that to be a design feature, rather than a flaw, and warn against methods that both impose penalties on complexity *and* converge to the truth (Leeb and Pötscher, 2008). Bayesians, on the other hand, explicate Ockham's razor as the result of conditioning over a wide class of plausible, prior probabilities that impose flattish distributions over theoretical parameters (Jeffreys, 1961; Bandyopadhayay et al., 1996; Wasserman, 2000; Myrvold, 2003). But that does not begin to explain how such prior probabilities lead one to true theories better than alternative biases would—unless one begs the question by appealing to the prior probabilities themselves. The question of *epistemic* justification, if not begged, is dodged.

The point can be sharpened with a bit of terminology. Say that a method is *truth-indicative* if at every stage of inquiry the theory it selects is probably true. But truth-indicative performance is impossible in inductive inference problems—insisting on achieving that impossible synchronic standard leads to inductive skepticism. More plausibly, one can entertain a range of weaker concepts of *diachronic truth-conduciveness*, and understand epistemic justification as achievement of the strongest performance possible for the problem one faces.[2] Weaker demands do not fall short of epistemic demands. They are, rather, the *appropriate* epistemic demands, in light of the intrinsic difficulty of the task at hand.

Over half a century ago, Carnap already sketched the idea in *On Inductive Logic* (1945). He recognized, that for inductive methods, synchronic truth-indicativeness is *too high* a standard: "the fact that the truth of the predictions reached by induction cannot be guaranteed does not preclude a justification in a weaker sense". On the other hand, Reichenbach's diachronic norm of limiting convergence is *too low* a standard:

> Reichenbach is right in the assertion that any procedure which does not [converge to the truth in the limit] is inferior to his rule of induction. However,

[2]In statistics, the distinction between truth-indicative and truth-conducive methods is closely tracked by the distinction between uniformly and point-wise consistent methods. Both types of consistency entail convergence to the truth in the limit, but for uniformly consistent methods, the probability and severity of error can be quantified and bounded at each sample size. For point-wise convergent methods no such guarantees can be given. Of course, we *prefer* uniformly convergent methods, but these do not always exist. No statistician claims that using a point-wise convergent method is not epistemically justified when there is no better alternative.

> his rule, which he calls "the" rule of induction, is far from being the only one possessing the characteristic. The same holds for an infinite number of other rules of induction. ... Therefore we need a more general and stronger method for examining and comparing any two given rules of induction ... (1945, p.)

The relevant notions of truth-conduciveness have to lie somewhere between those two extremes. If they are to be feasible, they must relax truth-indicativeness. If they are to mandate interesting short-run methodological principles like Ockham's razor, they must demand more than mere limiting convergence. We propose that such notions can be developed by adapting and refining existing concepts from formal learning theory.

1.2 Learning Theory and Truth-Conduciveness

Formal learning theory is a mathematical framework for studying inductive problems and the methods that solve them (see Putnam (1965), Gold (1967), Osherson et al. (1986), Kelly (1996)). As in computational complexity theory, inductive problems are classified by their intrinsic difficulty. Inductive methods are justified so long as they solve a problem as efficiently as problems of comparable complexity can be solved. From the perspective of formal learning theory, it makes no more sense to demand truth-indicative performance in inductive problems than it does to demand general polynomial time solutions to NP-hard problems. The demands of epistemic justification are kept proportionate to epistemic complexity.

The baseline notion of truth-conduciveness in formal learning theory is limiting convergence: methods eventually settle on the truth as information accumulates, without ever becoming certain that the future holds no surprises in store. As Carnap observed, limiting convergence is compatible with any arbitrary behavior in the short-run. To narrow the field of admissible methods, learning theorists have developed several refinements of limiting convergence that fall short of short-run guarantees. One possible refinement is to require methods to minimize the number of *mind changes* on the way to convergence (Putnam, 1965; Case and Smith, 1983; Sharma et al., 1997; Luo and Schulte, 2006; Kelly, 2004, 2007, 2011). Another refinement guards against "U-shaped learning," wherein learners conjecture a theory, reject it, and then return to it again (Carlucci et al., 2005; Carlucci and Case, 2013). We propose that these convergence criteria—lying midway between truth-indicativeness and mere limiting convergence—can answers Carnap's challenge. If scientists are in pursuit of truth, then virtuous inquiry ought to exhibit the virtues of pursuit. If the target of pursuit is evasive, then a certain amount of swerving may be expected. But false starts and U-turns ought to be avoided if possible—*virtuous pursuit is as direct as the problem situation allows.* We show how optimal truth-conduciveness mandates a preference for simple theories. That solves the traditional puzzle of justification for Ockham's Razor, a hallmark virtue of theory choice.

1.3 Belief Revision and Scientific Inquiry

Belief revision is an alternative, formal framework in which to analyze belief change driven by new information. Reliability and truth-conduciveness are not central to the belief revision framework. Gärdenfors seems indifferent—if not hostile—to concerns about truth:

> [T]he concepts of truth and falsity are *irrelevant* for the analysis of belief systems. These concepts deal with the relation between belief systems and the external world, which I claim is not essential for an analysis of epistemic dynamics. ... My negligence of truth may strike traditional epistemologists as heretical. However, one of my aims is to show that many epistemological problems can be attacked without using the notions of truth and falsity (Gärdenfors, 1988, p. 20).

Instead, belief revision theorists derive epistemic justification from conformity with a set of idealized postulates that govern rational belief change. The postulates are usually motivated by considerations of *preservation*, or *minimal change*: the injunction to (1) add only those new beliefs and (2) remove only those old beliefs, that are absolutely compelled by incorporation of new information.[3] It is not obvious that those postulates of rationality have anything to do with truth-conduciveness. In §5.2.1, we demonstrate that such a connection does, in fact, exist: we show that a weakened version of the rationality postulates is equivalent to a truth-conduciveness norm from formal learning theory, once the requirement of limiting convergence has been imposed.

It is also not obvious that there should be any connection between the rationality postulates that belief revision theorists take to govern theory change, and the synchronic theoretical virtues investigated by philosophers of science. Rott (2000) expresses a hope that such connections exist:

> In his joint book with J.S. Ullian, The Web of Belief (1978), Quine has added more virtues that good theories should have: modesty, generality, refutability, and precision. Again, belief revision as studied so far has little to offer to reflect the quest for these intuitive desiderata. Except for the issue of conservatism, Quine's list is one of theory choice rather than theory change in that it lists properties that a good posterior theory should have, independently of the properties of the prior theory. It is a strange coincidence that the philosophy of science has focussed on the monadic (nonrelational) features of theory

[3] Rott (2000) questions whether the rationality postulates of belief revision are correctly thought of as principles of minimal change. Alternatively, one can think of them as monotonicity principles. That also suggests a rapprochement with the learning-theoretic viewpoint: adherence to the rationality principles throughout the course of inquiry guarantees a certain degree of monotonicity—or "directness" – of convergence to the truth.

> choice, while philosophical logic has emphasized the dyadic (relational) features of theory change. I believe that it is time for researchers in both fields to overcome this separation and work together on a more comprehensive picture (2000, p. 15).

In §5.2.1 and §5.2.2, we demonstrate that theoretical refutability is a necessary condition for the rationality postulates of belief revision, in light of the requirement of limiting convergence. What's more, we show that theoretical refutability is equivalent to a version of theoretical simplicity.

1.4 The Plan of the Work

The subsequent three chapters of this work are concerned with developing a fairly idealized, but flexible model of inquiry. In Chapter 2, we show how topological structure arises from the tendency of empirical information to accumulate.[4] We develop an epistemological interpretation of standard topological concepts and outline how methodological difficulty arises from topological complexity. In Chapter 3, we show how the structure of empirical underdetermination is captured topologically. We define a general notion of empirical simplicity that arises from the topological structure of underdetermination. We also show how that notion of simplicity is related to Popper's degrees of falsifiability. Chapter 4 focuses on the empirical *questions* that frame scientific inquiry. In that chapter we characterize the class of empirical questions that can be successfully answered in the limit of inquiry. In Chapter 5, we develop notions of optimally direct convergence that refine the basic convergence notion of Chapter 4. We generalize the learning-theoretic notions of mind changes and U-shapes, to *reversals* and *cycles*, respectively. We show how cycle avoidance can be thought of as a weakening of some standard rationality principles from belief revision. We define two versions of Ockham's razor and show that they are individually necessary conditions for avoidance of cycles, and minimization of reversals, respectively. That performance can fall far short of truth-indicativeness, but it is, in our sense, optimally truth-conducive, so the argument steers past inductive skepticism. On the other hand, it mandates intuitive, methodological principles in the short run, so it also avoids methodological triviality. In Gandenberger's words, that justification "may not give us everything we want, but it does seem to give us everything we need" (2012). Finally, we characterize the classes of problems that can be solved without cycles and where a notion of minimization of reversals applies. In Chapter 6 we turn to a natural class of problems for which the simplicity order is a particularly apt epistemic representation. We demonstrate several examples of such problems. Chapter 7 concludes with some final, philosophical considerations.

[4]The general topological setting is consonant with the increasing awareness that topology is a fruitful setting for learning-theoretic analysis (Kelly, 1996; Martin et al., 2006; Schulte et al., 2007; Yamamoto and de Brecht, 2010; Case and Kötzing, 2013; Baltag et al., 2014).

Chapter 2

The Topology of Inquiry

Topology is usually thought of as a "rubber-sheet geometry," or the study of continuous deformations of a space. Topological learning theory begins with a re-interpretation of topology as the mathematics of verifiability by finite information. Once the fundamental translations are made, topology emerges as a natural setting for epistemology. Many of the same translations are made in Kelly (1996) and Vickers (1996). We adopt and expand upon them here.

2.1 Possible Worlds

Let W be a set of possible worlds. In a particular scientific context, possible worlds may be—among other things—probability distributions, points in a parameter space, or functions in a Hilbert space. A ***proposition*** is a set $P \subseteq W$. The set of all propositions is denoted $\mathcal{P}(W)$. The contradictory proposition is $\varnothing$ and the necessary proposition is W. We assume the usual correspondence between logical and set-theoretic operations: $P \wedge Q = P \cap Q$, $P \vee Q = P \cup Q$, $P^c = W \setminus P$ and P entails Q iff $P \subseteq Q$.

2.2 Information States

Some propositions correspond to what we think of as ***information states***. Propositional information is understood to be true. Examples include propositions concerning discrete experimental outcomes and inexact measurements of continuous quantities. Let $\mathcal{O} \subseteq \mathcal{P}(W)$ be the set of all possible information states one might be in.

Definition 1. *We denote the set of all information states in world w as:*

$$\mathcal{O}(w) = \{O \in \mathcal{O} : w \in O\}.$$

Definition 2. *The set of information states $\mathcal{O}$ is a **countable topological basis** iff the following postulates are satisfied.*

O1. $\bigcup \mathcal{O} = W$;

O2. If $A, B \in \mathcal{O}(w)$, *then* $A \cap B \in \mathcal{O}(w)$.[1]

O3. $|\mathcal{O}| \leq \omega$.

The first thesis is rather uncontroversial: vacuous information is information. If a world presented no information, it would be of no interest for empirical inquiry. The second thesis is material. It says that if A and B are both information states in world w, then so is their conjunction. Essentially: information accumulates. The second thesis is not a conceptual fact about the nature of information, e.g. quantum theory tells us that although we can be informed about the position and about the momentum of a particle, we cannot be informed of both at once. But information *does* accumulate in most ordinary, scientific contexts. In statistical estimation, more data is collected from the same distribution. In curve fitting, we observe more instances of a process governed by a general law. Most empirical inquiry does not get off the ground without some assumption about the tendency of information to accumulate. The third thesis is partly for mathematical tractability. But it is also well-motivated by Alan Turing's (1936) argument that infinite gradations of input information are indistinguishable. In what follows, we take for granted that the set of all information states form a countable topological basis.

2.3 Verifiable Propositions

Definition 3. *Proposition P is **verifiable** iff for every world $w \in P$ there is some information state $O \in \mathcal{O}(w)$ such that O entails P.*

Methodologically, a proposition is verifiable iff there is an empirical method that returns 'true' whenever the proposition is true. The natural method simply waits for information that entails the proposition. 'Bread will cease to nourish' is an archetypal verifiable proposition: in every world in which it is true we eventually observe that bread no longer works as it used to. 'The empirical law is a polynomial of degree 2 or higher' is another such proposition: if it is true, we eventually observe some non-linear effect. We show that the set of verifiable propositions is simply the closure of the set of information states under arbitrary union.

Proposition 1. *Let $\mathcal{V}$ be the set of verifiable propositions and let $\mathcal{O}^*$ be the closure of O under arbitrary union. Then $\mathcal{V} = \mathcal{O}^*$.*

[1] This can be weakened to [O2'.] If $A, B \in \mathcal{O}(w)$, then there is $C \in \mathcal{O}(w)$ such that $C \subseteq A \cap B$. Nothing essential would be changed, though the statements of the theorems and their proofs would be more cumbersome.

Proof. Suppose $E \in \mathcal{V}$. Then for every $w \in E$ there is $O_w \in \mathcal{O}(w)$ such that $O_w \subseteq E$. Therefore $E = \bigcup_{w \in E} O_w \in \mathcal{O}^*$. Now suppose $E \in O^*$. Then $E = \bigcup_{i \in I} O_i$ where I is some index set and $O_i \in \mathcal{O}$ for each i. Then if $w \in E$, there is some i such that $w \in O_i \subseteq E$ as required. □

The following are four theses about verifiability.

V1. The contradictory proposition $\varnothing$ is verifiable.

V2. The trivial proposition W is verifiable.

V3. The verifiable propositions are closed under finite conjunction.

V4. The verifiable propositions are closed under arbitrary disjunction.

The first follows trivially, since the contradictory proposition contains no world. The second follows from the fact that every world presents *some* information. The fourth is equally immediate: if every disjunct is entailed by some information state in each world, then so is the disjunction. Only the third proposition is substantive, because only the third relies essentially on the second thesis of information.

Proposition 2. *The set of verifiable propositions, $\mathcal{V}$, satisfies (V1-V4).*

Proof. It remains only to show (V3). Suppose A and B are verifiable. If $A \cap B = \varnothing$, then it is verifiable by $(V1)$. Suppose there is $w \in A \cap B$. Then there are information states O_A, O_B entailing A and B respectively, such that $w \in O_A \cap O_B$. So $w \in O_A \cap O_B \subseteq A \cap B$. By (O2) $O_A \cap O_B \in \mathcal{O}(w)$ as required. □

Perhaps the most striking thing about (V1-V4) is the asymmetry between (V3) and (V4). Although we can verify that the sun will rise every day this week if we wake up early enough, we cannot verify that the sun will rise every morning from now on. And although we can verify that any finite number of ravens are black by examining each of them, we cannot verify the universal generalization that all ravens are black. Universal laws are not verifiable, though their instances are. The asymmetry is just right for the study of verifiability.

What we have shown is that the verifiable propositions form a *topological space.*

Definition 4. *If $\mathcal{V} \subseteq \mathcal{P}(W)$, then $(W, \mathcal{V})$ is a* ***topological space*** *iff $\mathcal{V}$ satisfies the following four axioms.*

T1. $\varnothing \in \mathcal{V}$*;*

T2. $W \in \mathcal{V}$*;*

T3. If $A, B \in \mathcal{V}$, then $A \cap B \in \mathcal{V}$;

T4. If $\mathcal{S} \subseteq \mathcal{V}$, then $\bigcup \mathcal{S} \in \mathcal{V}$.

In topology, the elements of $\mathcal{V}$ are called ***open sets***. We say that $\mathcal{O}$ is a ***basis*** for $(W, \mathcal{V})$ iff $\mathcal{V}$ is the closure of $\mathcal{O}$ under arbitrary union. The elements of the basis are called ***basic open sets***. Every topological basis is a basis for the topology $(W, \mathcal{O}^*)$ where $\mathcal{O}^*$ is the closure of $\mathcal{O}$ under arbitrary union. Proposition 1 shows that the information states are a basis for the topology of verifiable propositions. We also need the following technical notion: if $A \in \mathcal{P}(W)$, then the ***subspace topology on*** A is given by $(A, \mathcal{V}\big|_A)$ where $\mathcal{V}\big|_A = \{E \cap A : E \in \mathcal{V}\}$.

2.4 Falsifiable Propositions

Definition 5. *Proposition P is **falsifiable** iff P^c is verifiable.*[2]

Equivalently, every world in P^c presents some information entailing P^c. Methodologically, a proposition is falsifiable so long as there is a method that returns 'false' whenever the proposition is false. The following four theses axiomatize falsifiability.

F1. The contradictory proposition $\varnothing$ is falsifiable.

F2. The trivial proposition W is falsifiable.

F3. The falsifiable propositions are closed under arbitrary conjunction.

F4. The falsifiable propositions are closed under finite disjunction.

In topology, the falsifiable propositions are called ***closed*** sets. The complement of every closed set is open. If $\mathcal{F} \subseteq \mathcal{P}(\mathcal{W})$ and $\mathcal{F}$ satisfies (F1-F4), then letting $\mathcal{V} = \{W \setminus F : F \in \mathcal{F}\}$ we have that $(W, \mathcal{V})$ is a topological space. Finally, we say that $Q \in \mathcal{P}(W)$ is ***closed in*** P iff $\varnothing \neq Q \cap P$ is closed in the subspace topology defined on P.

2.5 Decidable Propositions

Definition 6. *Proposition P is **decidable** iff it is both verifiable and falsifiable.*

Methodologically, a proposition is decidable iff there is an empirical method that returns 'true' whenever it is true and 'false' otherwise. In topology, decidable propositions are called ***clopen*** sets.

[2] We will use *refutable* and *falsifiable* interchangeably.

2.6 Conditionally Falsifiable Propositions

Definition 7. *Proposition P is **conditionally falsifiable** iff it is the intersection of a verifiable and a falsifiable proposition.*

Equivalently, $P = A \cap B^c$ for A, B both verifiable. In topology, such sets are said to be ***locally closed***. The methodological interpretation of conditional falsifiability is somewhat subtle. If P is true, then A will be verified, and B will not. If P is false, then either B will be verified or A will not. Every conditionally falsifiable proposition can be thought of as a default rule: conclude P so long as A has been verified but B has not. This is best illustrated with an example. Suppose you are interested in the polynomial form of an empirical law. The proposition 'the law is quadratic' is not refutable if the truth is linear—no amount of true linear data rules out the appearance of a subtle quadratic effect at larger sample sizes. However, if you *do* observe a quadratic effect, the proposition becomes refutable by the observation of any higher-order effect. So, if you observe a quadratic effect you can fallibly conclude that the law is quadratic. If you are wrong, your conjecture will be refuted by future data. The following proposition express that epistemic feature topologically.

Proposition 3. *P is locally closed in $(W, \mathcal{V})$ iff for every $w \in P$ there is $O \in \mathcal{O}(w)$ such that $P \cap O$ is closed (falsifiable) in the subspace topology on O.*

Proof. $\Rightarrow$: Suppose P is locally closed. Then $P = O \cap C$ for $O, C^c \in \mathcal{V}$. So for all $w \in P$, there is $O' \in \mathcal{O}(w)$ such that $O' \subseteq O$. Therefore $P \cap O' = O' \cap C$ is closed in O'. $\Leftarrow$: Suppose the right hand side holds. Then for each $w \in P$ there is $O_w \in \mathcal{O}(w)$ such that P is closed in O_w, and therefore $O_w \setminus P$ is open in O_w and in W as well. Then $P = \bigcup_{w \in P}(O_w) \setminus \bigcup_{w \in P}(O_w \setminus P)$ is locally closed as required. □

2.7 Constructible Propositions

Definition 8. *P is **constructible** iff it is a finite union of locally closed (conditionally verifiable) propositions.*

The constructible sets form a *Boolean algebra.*[3] In fact the constructible sets are precisely the algebra generated if we close the open (verifiable) and closed (falsifiable) propositions under finite conjunction and complementation. The methodological interpretation

[3]Hempel (1965) posited that the empirically significant propositions must be closed under negation and disjunction i.e. they must form an algebra. Finding that neither the falsifiable nor verifiable propositions were closed under negation, he despaired of any sharp criterion of empirical significance. The constructible propositions go part way in answering that objection. Methodologically, these are the propositions that are decidable with finitely many *mind-changes* (Kelly, 1996). However, they do not include propositions of higher complexity e.g. 'for every substance there exists some solvent'. Propositions of that kind reside in the second level of the Borel hierarchy. In a sense, Hempel was correct that cognitive significance comes in degrees—but it is in degrees of topological complexity.

of the constructible propositions will not play a large role in what follows. However, Kelly (1996) shows that these are exactly the propositions that are decidable with finitely many mind changes. We have the following useful theorem about constructible propositions.

Proposition 4. *P is constructible iff there is a least integer n such that P admits a decomposition into n disjoint, locally closed sets.*

Proof. See the main (and only) Theorem in Allouche (1996). □

2.8 The Generalized Borel Hierarchy

The constructible propositions do not exhaust the propositions of interest for empirical inquiry. Propositions of higher topological complexity are classified according to the Borel hierarchy.[4]

Definition 9. *The successive levels of the Borel hierarchy are defined recursively.*

$$\Sigma^0_1 = \mathcal{V}.$$
$$\Sigma^0_\alpha = \{Y : Y = \bigcup_{i\in\omega} U_i \setminus V_i\}.$$

where $U_i, V_i \in \Sigma^0_{\beta_i}$ for $\beta_i < \alpha$ and $\alpha < \omega_1$. Furthermore, we define $\Pi^0_\alpha = \{Y : Y^c \in \Sigma^0_\alpha\}$ and $\Delta^0_\alpha = \Sigma^0_\alpha \cap \Pi^0_\alpha$.

For our purposes, we will not be concerned with propositions residing above the second level of the hierarchy. So the most complex propositions we will deal with are countable unions of locally closed sets (Σ^0_2 sets) and their complements (Π^0_2 sets). The methodological interpretation of these sets will play a large role in what follows.

[4]We follow de Brecht and Yamamoto (2009) in our definition. This definition is in general different from the classical definition, but more appropriate for our purposes. The two definitions coincide for metrizable spaces (see Proposition 3.7 in Kechris (1995)).

Chapter 3

Simplicity and the Structure of Underdetermination

Underdetermination arises when finite information cannot distinguish between possibilities. The *problem of induction* is just another name for underdetermination by information. Hume teaches that we cannot be certain that bread will always nourish—everything we observed in such a world would be consistent with bread ceasing to nourish sometime in the future (Hume, 2011, Part II, §28). Finite information can establish that bread has nourished up until now, but underdetermines whether it will continue to nourish. Where there is underdetermination, inquiry must proceed by ampliative leaps beyond the data. In this section, we show how topological operators like *closure*, *boundary*, *interior* and *frontier* capture the various aspects of underdetermination in an elegant and concise fashion. We then show how a notion of simplicity arises from the structure of underdetermination in a topological space.

3.1 The Topological Closure

For every proposition $P \subseteq \mathcal{P}(W)$, we can define the set of information states consistent with P as follows:

$$\mathcal{O}(P) = \bigcup_{w \in P} \mathcal{O}(w).$$

The ***closure*** of P, written $\overline{P}$, is the set of worlds in which P is never falsified, or equivalently, the set of worlds that only present information consistent with P.

Definition 10. *The closure of P is $\overline{P} = \{w : \mathcal{O}(w) \subseteq \mathcal{O}(P)\}$.*

We rehearse some standard facts about the closure operator. Proposition Q entails $\overline{P}$ if and only if every information state consistent with Q is consistent with P.

Proposition 5. *$Q \subseteq \overline{P}$ if and only if $\mathcal{O}(Q) \subseteq \mathcal{O}(P)$.*

Proof. $Q \subseteq \overline{P}$ iff for every $w \in Q$, $\mathcal{O}(w) \subseteq \mathcal{O}(P)$ iff $\mathcal{O}(Q) = \bigcup_{w \in Q} \mathcal{O}(w) \subseteq \mathcal{O}(P)$. □

Proposition 6. *P is falsifiable if and only if $P = \overline{P}$.*

Proof. P is falsifiable iff every world in P^c presents information falsifying P iff $P^c = \overline{P}^c$ iff $P = \overline{P}$. □

Proposition 7. *If $P \subseteq Q$, then $\overline{P} \subseteq \overline{Q}$.*

Proof. Suppose $P \subseteq Q$. Then $\mathcal{O}(P) \subseteq \mathcal{O}(Q)$. So $\overline{P} = \{w : \mathcal{O}(w) \subseteq \mathcal{O}(P) \subseteq \mathcal{O}(Q)\}$. So if $w \in \overline{P}$, $w \in \overline{Q}$ as required. □

3.2 Interior, Boundary, and Frontier

There are finer epistemological distinctions to be made. The ***boundary*** of P, written ∂P, is the set of worlds where neither P nor P^c is ever refuted, i.e. $\overline{P} \cap \overline{P^c}$. In boundary worlds, the question *whether* P is underdetermined by information. The boundary of any proposition is closed. The ***interior*** of P, written P°, is the set of worlds where P is eventually verified i.e. $\overline{P} \cap \overline{P^c}^c$. The interior of any proposition is open. Clearly, the closure of P is the union of its interior and its boundary. The ***frontier*** of P, sometimes written $\overset{\vee}{P}$, is the set of worlds in which P is false, but never refuted i.e. $\overline{P} \cap P^c$. In the frontier, P is false, but the question *whether* P is underdetermined. Denote n iterations of the frontier operator as $\overset{n\vee}{P}$. We can now give a more detailed statement of Proposition 4:

Proposition 8. *P is constructible iff there is a least integer n such that $\overset{2n\vee}{P} = \varnothing$ and furthermore, P admits the following canonical decomposition into n disjoint locally closed sets:*

$$P = (\overline{P} \setminus \overline{\overset{\vee}{P}}) \cup \left(\overline{\overset{\vee\vee}{P}} \setminus \overline{\overset{\vee\vee\vee}{P}} \right) \cup \cdots \cup \left(\overline{\overset{(2n-2)\vee}{P}} \setminus \overline{\overset{(2n-1)\vee}{P}} \right).$$

Proof. See the main (and only) Theorem in Allouche (1996). □

3.3 Separation Axioms

Topological separation axioms constrain how underdetermined worlds (or maximally informative propositions) can be by information. We say that the space $(W, \mathcal{V})$ is T_0 iff for every distinct pair of worlds w and v, we have that $\mathcal{O}(w) \neq \mathcal{O}(v)$. This has the effect of ruling out *metaphysical* distinctions between worlds: if we have distinguished between w and v, it is because there is some informational difference between them. We can say more. The space $(W, \mathcal{V})$ is T_1 iff for every distinct pair of worlds w and v, $\mathcal{O}(w) \nsubseteq \mathcal{O}(v)$. Thus, for any distinct w and v, world w presents information refuting v and vice-versa.

The T_D axiom is a separation axiom introduced by Aull and Thron (1962) that is strictly stronger than T_0 and weaker than T_1. We say that the space $(W, \mathcal{V})$ is T_D iff every singleton $w \in W$ is locally closed.[1] We will discuss the epistemological significance of the T_D axiom in what follows.

3.4 The Specialization Order

If we live in a world where bread will always nourish, everything we ever observe will be consistent with bread ceasing to nourish sometime in the future. But the situation is asymmetrical: if we lived in a world where bread ceases to nourish eventually, we would find out sooner or later. There is a natural order that captures the structure of inductive underdetermination between worlds.

Definition 11. *The* ***specialization order*** *is given by setting $w \preceq v$ iff $\mathcal{O}(w) \subseteq \mathcal{O}(v)$.*

Then we have that $w \preceq v$ iff all information in w is consistent with v. If that is the case, we say that w has a *problem of induction* with respect to v. Furthermore, let $w \prec v$ if $w \preceq v$ and $v \npreceq w$. Then we say that w has a *strict* problem of induction with respect to v. The specialization order is a pre-order on W for every topological space $(W, \mathcal{V})$. So long as $\mathcal{V}$ is T_0, the specialization is also a partial order i.e. if $w \preceq v$ and $v \preceq w$, then $w = v$. If $\mathcal{V}$ is T_1, the specialization order is trivial and uninteresting: nothing is ordered with anything else.[2]

We can lift that order to arbitrary proposition in an obvious way.

Definition 12. *For $X, Y \in \mathcal{P}(W)$ set $X \preceq Y$ iff $\mathcal{O}(X) \subseteq \mathcal{O}(Y)$.*[3]

Then $X \preceq Y$ iff any information consistent with X is consistent with Y. For any topological space $(W, \mathcal{V})$, that defines a pre-order over $\mathcal{P}(W)$. It is not in general a partial order, even if the topology is T_0. Crucially, that order does not collapse in T_1 spaces. We identify $\preceq$ with the simplicity order over propositions.

Definition 13. *For all $X, Y \in \mathcal{P}(W)$, X is* ***simpler*** *than Y if and only if $X \preceq Y$.*

[1] I am grateful to Thomas Icard for turning my attention to these spaces.

[2] Many of the topological spaces studied in the empirical sciences have a T_1 structure, so their specialization order is uninteresting. Lifting the simplicity order to propositions in Definition 12 allows us to abstract away from the T_1 structure of such spaces.

[3] Or, equivalently, $X \preceq Y$ iff $X \subseteq \overline{Y}$.

So X is simpler than Y if and only if all information states compatible with X are compatible with Y.[4] This straightforward relation captures many of our simplicity intuitions. Exactly that relation holds between sets of polynomials of lower and higher degree; between nested statistical models; and between universal generalizations like "all ravens are black" and their negations. We give some examples for which the simplicity relation renders an intuitive verdict.

Example 1. *Let W be the set of all binary sequences. Think of a 0 in the n-th position as encoding the observation 'the n-th raven is black' and 1 in the n-th position as encoding the observation 'the n-th raven is not black'. Let the information states $\mathcal{O}$ be given by the cylinder sets determined by finite initial sequences. Then $(W, \mathcal{O}^*)$ is a topological space. Let B be the set containing the constant zero sequence. This proposition says that all ravens are black. Then $B \prec B^c$, since any finite number of black-raven observations are consistent with the next raven being non-black.*

Example 2. *In learning theory a **concept space** is a collection of subsets of $\mathbb{N}$. Concept spaces are used to model language learning problems (Luo and Schulte, 2006; Yamamoto and de Brecht, 2010). An element of a concept space is called a **concept**. Given a concept space W and $S \subseteq \mathbb{N}$, we define $\uparrow S = \{L \in W : S \subseteq L\}$. The information states are given by the set of all such up-sets: $\mathcal{O} = \{F : F = \uparrow S \text{ for some } S \subseteq \mathbb{N}\}$. Then $(W, \mathcal{O}^*)$ is a topological space. It is straightforward to show that richer concepts are more complex i.e. if $S, S' \in W$, then $S \preceq S'$ iff $S \subseteq S'$.*

Example 3. *Consider the paradigmatic problem of inferring polynomial degrees. Let W_{cts} be the set of all continuous functions with support on $[a, b]$. Let $\mathbf{a} = (a_1, \ldots, a_n)$ be a finite vector of real numbers such that $a_n \neq 0$. A polynomial function is a function expressible as $f_{\mathbf{a}} = \sum_{i \leq n} a_i x^i$. Let $W_{\mathsf{poly}} \subset W_{\mathsf{cts}}$ denote the set of all polynomial functions with support on the interval $[a, b]$. Then n is the* degree *of $f_{\mathbf{a}}$ and the* form *of $f_{\mathbf{a}}$ is the set S of non-zero positions in $\mathbf{a}$. Let D_n denote all $f \in W_{\mathsf{poly}}$ of degree n and let F_S denote the set of all $f \in W_{\mathsf{poly}}$ with form S. Define $F_{\leq S}$ to be the union of all $F_{S'}$ such that $S' \subseteq S$. Define $F_{<S}$ similarly, for strict subsets of S. Observations are finite sets of open coordinate rectangles in the real plane, which are just non-empty cross products $(x_1, x_2) \times (y_1, y_2)$. Given finite set R of such rectangles, the information E_R is the set of all $f \in W_{\mathsf{cts}}$ that have non-empty intersection with each element of R. Let $\mathcal{O}_{\mathsf{rec}}$ denote the set of all such information states. Then $\mathcal{O}_{\mathsf{rec}}$ is a topological basis, for given $E, E' \in \mathcal{O}_{\mathsf{rec}}$, the conjunction $E \cap E'$ is the set of all continuous functions that pass through $R \cup R'$, for corresponding, finite sets of rectangles R, R'. Thus, $\mathfrak{I}_{\mathsf{cts}} = (W_{\mathsf{cts}}, \mathcal{O}^*_{\mathsf{rec}})$ is a topological space. Similarly, $\mathfrak{I}_{\mathsf{poly}} = \mathfrak{I}_{\mathsf{cts}}\big|_{W_{\mathsf{poly}}}$ is a topological space. Note that the topological space is not imposed—it emerges from the sort of measurements one intends to make.*

[4]Note that that relationship depends only on the topological structure of information and is therefore preserved by any homeomorphism of the space. This makes it invariant e.g. under *grue*-like transformations.

Let $\mathcal{O}_{\mathsf{L}^2}$ be the information basis induced by the L^2 norm on W_{cts}. Then $\mathfrak{I}_{\mathsf{L}^2} = (W_{\mathsf{cts}}, \mathcal{O}_{\mathsf{L}^2})$ is a Hilbert space. As one would expect, the set of polynomials of degree $n+1$ is more complex than the polynomials of degree n.

Lemma 1. *If S is a finite subset of $\mathbb{N}$, $F_{<S}, F_{\leq S}$ are closed in $\mathfrak{I}_{\mathsf{L}^2}$.*

Proof. Every finite dimensional subspace of a Hilbert space is closed. Finite unions of closed sets are closed. □

Lemma 2. *If $D \subseteq W_{\mathsf{cts}}$ is closed in $\mathfrak{I}_{\mathsf{L}^2}$, it is closed in $\mathfrak{I}_{\mathsf{cts}}$.*

Proof. Suppose $D \subseteq W_{\mathsf{cts}}$ is not closed in $\mathfrak{I}_{\mathsf{cts}}$. Then there is $f \in \overline{D} \setminus D$ where the closure is in terms of $\mathfrak{I}_{\mathsf{cts}}$. Exploiting continuity and the compactness of $[a, b]$, we can construct a Cauchy sequence $\{f_i\} \subset D$ converging to f in the L^2 norm. Since every closed subset of a complete metric space is complete, it must be that D is not closed in $\mathfrak{I}_{\mathsf{L}^2}$. □

Proposition 9. *If S, S' are finite subsets of $\mathbb{N}$ and $f \in F_S$, then $F_S \preceq_{\mathfrak{I}_{\mathsf{cts}}} F_{S'}$ iff $S \subseteq S'$.*

Proof. $\Leftarrow$: It suffices to show that for all $f \in F_S$ and $\epsilon > 0$, there exists $g \in F_{S'}$ such that $\sup |f - g| < \epsilon$. Since the x^i are continuous, $|\sum_{i \in S' \setminus S} x^i|$ is continuous as well. Since $[a, b]$ is closed and bounded, $|\sum_{i \in S' \setminus S} x^i|$ attains a maximum M on $[a, b]$ by the extreme value theorem. So letting $a = \epsilon/M$, and $g = f + \sum_{i \in S' \setminus S} ax^i$, $\sup |f - g| < \epsilon$. $\Rightarrow$: Immediate from lemmas 1 and 2. □

Proposition 10. *For all $n \in \mathbb{N}$, $D_n \preceq D_{n+1}$.*

Proof. Follows immediately from Proposition 9. □

Our notion of simplicity has some pedigree in the philosophy of science. In *The Logic of Scientific Discovery*, Popper proposes to define simplicity in terms of the falsifiability relation:

> A statement x is said to 'falsifiable in a higher degree' or 'better testable' than a statement y ... if and only if the class of potential falsifiers of x includes the class of the potential falsifiers of y as a proper subclass' ... The epistemological questions which arise in connection with the concept of simplicity can all be answered if we equate this concept with *degree of falsifiability* (Popper, 1959).

Popper's notion and ours are equivalent. To see that, define the class of potential falsifiers of $X \subseteq W$ as the set of all information states inconsistent with X:

$$\mathcal{F}(X) = \mathcal{O} \setminus \mathcal{O}(X).$$

Then X is *more falsifiable* than Y, if $\mathcal{F}(Y) \subseteq \mathcal{F}(X)$. It is easy to show that $X \preceq Y$ if and only if $\mathcal{F}(Y) \subseteq \mathcal{F}(X)$. Much of what follows is an elaboration of the consequences of identifying the simplicity order with the falsifiability order.

3.5 Simplicity and Falsifiability

Definition 14. *Nonempty proposition P is* ***minimal in*** $\preceq$ *iff* $\varnothing \neq A \preceq P$ *implies that* $A = P$.

Since simplicity is defined in terms of falsifiability, it is not surprising that minimal elements in the simplicity order are falsifiable.

Proposition 11. *If a proposition P is minimal in the simplicity order, it is closed (falsifiable).*

Proof. Suppose P is not closed. Let $w \in \overline{P} \setminus P$. Then $\{w\} \preceq P$ and $\{w\} \neq P$ so P is not minimal. □

The converse, however, is not true. For an easy counterexample consider that the entire space W is closed, but every proposition is simpler than W.

Definition 15. *Define the* **downward closure** *of a proposition P as follows:*

$$P_{\preceq} = \bigcup\{Q \preceq P : Q \in \mathcal{P}(W)\}$$

Say that a proposition P is **downward closed** *iff* $P = P_{\preceq}$.

If a proposition is downward-closed, all simpler propositions are logically stronger. A downward closed proposition is minimal so far as empirical—though not logical—considerations are concerned. It is straightforward to show that the downward closure of a proposition in the simplicity order is its topological closure i.e. $P_{\preceq} = \overline{P}$. The following proposition follows immediately.

Proposition 12. *Proposition P is downward closed in the simplicity order iff P is closed (falsifiable).*

To summarize: if a proposition is *simplest*, then it is falsifiable, and if a proposition is falsifiable, then any simpler proposition is logically stronger.

Constructible and locally closed propositions interact nicely with the simplicity order. Locally closed propositions have a special feature: the set of all disjoint, simpler propositions is falsifiable.

Proposition 13. *P is locally closed iff $\overset{\vee}{P} = P_{\preceq} \setminus P$ is closed (falsifiable).*

Proof. $\Rightarrow$: If P is locally closed, $P = A \cap B^c$ for $A, B \in \mathcal{V}$. So

$$\overline{P} \cap P^c = \overline{A \cap B^c} \cap (A^c \cup B) = (\overline{A \cap B^c} \cap A^c) \cup (\overline{A \cap B^c} \cap B) = \overline{A \cap B^c} \cap A^c,$$

since $\overline{A \cap B^c} \cap B \subseteq \overline{A} \cap \overline{B^c} \cap B = \overline{A} \cap B^c \cap B = \varnothing$. And since A is open, $\overline{A \cap B^c} \cap A^c$ is closed. $\Leftarrow$: In general, $P = \overline{P} \setminus \overset{\vee}{P}$. If $\overset{\vee}{P}$ is closed, then P is locally closed. □

As an immediate Corollary, we have that, in every world in a T_D space, the set of all simpler worlds is falsifiable.

Corollary 1. *$(W, \mathcal{V})$ is a T_D space iff for all $w \in W$, $\overset{\vee}{\{w\}}$ is closed (falsifiable).*

Furthermore, the simplicity relation is antisymmetric for disjoint, constructible propositions.

Proposition 14. *$A, B \in \mathcal{P}(W)$ are non-empty, constructible, and disjoint, then $B \not\preceq A$ if $A \preceq B$.*

Proof. First note that for all A, $\overset{\vee\vee}{A} \subseteq A$, since $\overset{\vee\vee}{A} = \overline{\overline{A} \cap A^c} \cap (\overline{A} \cap A^c)^c = \overline{\overline{A} \cap A^c} \cap A \cup \overline{\overline{A} \cap A^c} \cap \left(\overline{A}\right)^c$. The first disjunct is a subset of A and the second disjunct $\overline{\overline{A} \cap A^c} \cap \left(\overline{A}\right)^c \subseteq \overline{A} \cap \overline{A^c} \cap \left(\overline{A}\right)^c = \varnothing$.

Let A, B be constructible. Let n be the least natural such that $\overset{2n\vee}{A} = \varnothing$. Such an n exists by Proposition 8. Now suppose $B \subseteq \overline{A}$ and $A \subseteq \overline{B}$. Since A, B are disjoint, we have that $B \subseteq \overset{\vee}{A}$ and $\overline{B} \subseteq \overline{\overset{\vee}{A}}$. Therefore $A \subseteq \overline{\overset{\vee}{A}}$. Since for all propositions A, $A \subseteq \left(\overset{\vee}{A}\right)^c$, we have that $A \subseteq \overset{\vee\vee}{A}$. So $A = \overset{\vee\vee}{A}$. But then $\overset{(2n-2)\vee}{A} = \overset{2n\vee}{A} = \varnothing$. So n is not the least natural such that $\overset{2n\vee}{A} = \varnothing$. Contradiction. $\square$

Chapter 4

Empirical Problems and their Solutions

4.1 The Context of Inquiry

It is a philosophical truism that science is a problem-solving activity aimed at answering a theoretical *question* (Van Fraassen, 1980; Laudan, 1984). Laudan (1984) gives a representative articulation of that view:

> If problems are the focal point of scientific thought, theories are its end result. Theories matter, they are *cognitively* important, insofar as—and only insofar as—they provide adequate solutions to problems. If problems constitute the questions of science, it is theories which constitute the answers (1984, p. 13).

To adequately flesh out the context of inquiry, we must include the question that inquiry is meant to answer. We adopt a partition semantics for questions—standard in theoretical linguistics—wherein a question sets up a choice between the mutually exclusive and jointly exhaustive propositions that constitute its answers. (Hamblin, 1973; Groenendijk and Stokhof, 1984).

Definition 16. *A **question** is a partition of W into a countable set of **answers**.*

The topological basis specifies what information is potentially *provided*. The question under discussion regiments the output information *demanded*. An empirical problem arises from the tension between the information demanded by the question and the information supplied by the basis.

Definition 17. *An **empirical problem** is a triple $\mathfrak{P} = (W, \mathcal{O}, \mathcal{Q})$, where W is a set of possibilities, $\mathcal{O} \subseteq \mathcal{P}(W)$ is a countable topological basis, and $\mathcal{Q}$ is a question partitioning W.*

In other words, the context specifies the possibilities the inquirer takes seriously (W), the information Nature is expected to provide as input ($\mathcal{O}$), and the information the inquirer is expected to output ($\mathcal{Q}$).[1]

4.2 Empirical Methods

We are not interested in all of the features of agents performing empirical inquiry. Neither do we care so much about what belief is, or for traditional distinctions between belief, acceptance, and so on. We represent agents only by their *method*, which we take to be their disposition to respond to information with *conjectures*.

Definition 18. *An **empirical method** is a total function* $\lambda : \mathcal{O} \to \mathcal{P}(W)$.

Methods can conjecture any proposition in response to information. This allows a method to make many more distinctions than are relevant to the question. It is helpful to define operators that pick out the answers entailed by any given proposition.

Definition 19. *For* $A \in \mathcal{P}(W)$*, define* $\widehat{\mathcal{Q}}(A)$ *to be the set of all answers compatible with* A *i.e.* $\{Q \in \mathcal{Q} : A \cap Q \neq \varnothing\}$ *and* $\mathcal{Q}(A) = \bigcup \widehat{\mathcal{Q}}(A)$.

When $w \in W$ we abuse notation somewhat, by letting $\widehat{\mathcal{Q}}(w) = \widehat{\mathcal{Q}}(\{w\})$. We also make the following idealization on the nature of the learning methods under consideration.

Definition 20 (Deductive cogency)**.** *We say that a method* λ *is* **deductively cogent** *iff* $\varnothing \neq \mathcal{Q}(\lambda(E)) \subseteq \mathcal{Q}(E)$ *for* $E \in \mathcal{O} \setminus \{\varnothing\}$.

A deductively cogent method never admits an answer as a possibility if it is inconsistent with the data. We assume all methods are deductively cogent from now on.

4.3 Success in the Limit

A method solves the problem if in every world, the method eventually converges on the true answer. That is, in every world there is some information state such that for all stronger information states consistent with that world, the method conjectures the true answer.

Definition 21. *Method* λ ***solves*** $\mathfrak{P}$ ***in the limit*** *iff for all* $w \in W$*, there is a **locking information*** $E \in \mathcal{O}(w)$ *such that if* $F \in \mathcal{O}(w)$*, then* $\varnothing \neq \lambda(E \cap F) \subseteq \mathcal{Q}(w)$.

It is convenient to refer to the set of all information states for a particular world on which a given method locks on to the correct answer.

[1] It has been widely observed that observations are often *theory laden*—what is considered information may itself be determined by theory. We do not represent that dependence.

Definition 22 (Locking information)**.** *For arbitrary empirical method* λ *and* $w \in W$*, define* $\mathsf{Lock}(\lambda, w)$ *as the set of all* $E \in \mathcal{O}(w)$ *such that if* $F \in \mathcal{O}(w)$*, then* $\lambda(E \cap F) \subseteq \mathcal{Q}(w)$*.*

The following proposition is immediate.

Proposition 15. λ *solves* $\mathfrak{P}$ *in the limit iff* $\mathsf{Lock}(\lambda, w) \neq \varnothing$ *for each* $w \in W$*.*

As in all learning-theoretic analyses, ours is an *externalist* notion of success: a method that solves the problem is guaranteed to converge to the true answer, no matter which world is the true one. However, there is no sense in which the method becomes subjectively certain that it has converged to the truth.[2] The set of all problems that are solvable in the limit can be given an elegant, topological characterization.[3] We first define a special class of questions.

Definition 23. *Question* $\mathcal{Q}$ *is* ***locally closed*** *iff each answer* $Q \in \mathcal{Q}$ *is locally closed. Problem* $\mathfrak{P} = (W, \mathcal{O}, \mathcal{Q})$ *is locally closed iff* $\mathcal{Q}$ *is.*

Definition 24. *Question* $\mathcal{Q}'$ *refines another question* $\mathcal{Q}$ *iff for each answer* $Q' \in \mathcal{Q}'$ *there is an answer* $Q \in \mathcal{Q}$ *such that* $Q' \subseteq Q$*.*

Proposition 16 (Characterization of Solvability)**.** *The following propositions are equivalent*

1. *There is a method that solves* $\mathfrak{P} = (W, \mathcal{O}, \mathcal{Q})$ *in the limit.*
2. *Each answer* $Q \in \mathcal{Q}$ *is a* Δ^0_2 *proposition.*
3. *The question* $\mathcal{Q}$ *is refined by a locally closed question* $\mathcal{Q}'$*.*

Proof. To see that (1) implies (2), suppose λ solves $\mathfrak{P}$. For each w, choose $O_w \in \mathsf{Lock}(\lambda, w)$. Then $\{O_w : w \in W\}$ is a (countable) cover of W by locking information. Now let $O'_w = \bigcup\{E \in \mathcal{O} : E \subset O_w \text{ and } \lambda(E) \nsubseteq \mathcal{Q}(w)\}$. Then $O_w \setminus O'_w$ is locally closed. We claim that $Q = \bigcup_{w \in \mathcal{Q}} O_w \setminus O'_w$ for each $Q \in \mathcal{Q}$. Let $w \in Q$, then $w \in O_w$ and since O_w is locking for w, $w \notin O'_w$. So $w \in O_w \setminus O'_w \subseteq \bigcup_{w \in \mathcal{Q}} O_w \setminus O'_w$. Now suppose $v \in \bigcup_{w \in Q} O_w \setminus O'_w$. Then for some $w \in Q$, $v \in O_w \setminus O'_w$. Suppose $v \in Q' \neq Q$, then there is $O_v \in \mathsf{Lock}(\lambda, v)$ and $\lambda(O_v \cap O_w) \subseteq Q' \neq Q$. But then $v \in O'_w$. Contradiction. We have

[2] One may well be interested in stronger notions of success, in which subjective certainty is possible. Finite mind-change decidability is one such notion. It is investigated in Kelly (1996) and Luo and Schulte (2006). The notion of success investigated in this work is weak enough to be achievable in a broad class of problems and strong enough to be of methodological interest.

[3] This result is also given in de Brecht and Yamamoto (2009, Theorem 5), where it is couched in the terms of computable analysis. Many of the moves made in this theorem were anticipated in the manuscript of Kelly and Lin (2011, Theorem 3) in a first-countable setting. It is also to be published in Baltag et al. (2014, Theorem 8). It is proven for the Baire space in Kelly (1996, Proposition 4.10). Except where noted, this proof was arrived at independently by the author.

shown that each $Q \in \mathcal{Q}$ is Σ^0_2. Since there are only countably many answers $Q^c \in \Sigma^0_2$ as well. So each Q is Δ^0_2.

To see that (2) implies (3), suppose Q is a Σ^0_2 proposition. Then $Q = \bigcup_{i \in \mathbb{N}} O_i \cap C_i$ where $O_i, C_i^c \in \mathcal{V}$. Since the constructible sets are closed under complementation, letting $Q_i = O_i \cap C_i \setminus (\bigcup_{j<i} O_j \cap C_j)$, we have that the Q_i are a disjoint family of constructible sets.[4] By Proposition 4, each Q_i can be decomposed into finitely many disjoint locally closed sets as required.

To see that (3) implies (1), we construct a method that solves an arbitrary locally closed question. In broad outline, the method enumerates all of the answers and conjectures the first one in the enumeration that is refutable in the subspace of the current data. The nature of locally closed questions ensures that no matter which answer is true, it eventually becomes (and remains) the first refutable answer in the enumeration. Let $\mathcal{Q}'$ be locally closed and refine $\mathcal{Q}$. Enumerate the elements of $\mathcal{Q}'$ as $\{Q'_i\}_{i \in \mathbb{N}}$. For $E \in \mathcal{O}$, define

$$\sigma(E) = \begin{cases} \text{the least } i \text{ such that } Q'_i \cap E \text{ is closed in } E & \text{if it exists,} \\ \omega & \text{otherwise.} \end{cases}$$

Now define the following method:

$$\lambda_{\mathsf{MinFals}}(E) = \begin{cases} Q'_{\sigma(E)} & \text{if } \sigma(E) \neq \omega, \\ E & \text{otherwise.} \end{cases}$$

Let $w \in Q'_i \in \mathcal{Q}'$. Let $C := \{j : j < i \text{ and there is } E \in \mathcal{O}(w) \text{ such that } Q'_j \cap E \text{ is closed in } E\}$. For each $j \in C$, let $E_j \in \mathcal{O}(w)$ be such that Q'_j is closed in E_j. Then for $j \in C$, Q'_j is closed (or falsified) in $\bigcap_{j \in C} E_j := E \in \mathcal{O}(w)$. Furthermore, for all $j \in C$, Q'_j is falsified in $\bigcap_{j \in C} E_j \cap \bigcap_{j \in C} (E \setminus Q'_j) := F \in \mathcal{O}(w)$. Furthermore, by Proposition 3 there is $O_w \in \mathcal{O}(w)$ such that $Q'_i \cap O_w$ is closed in O_w. Then $O_w \cap F$ is locking for $\lambda_{\mathsf{MinFals}}$ in w.

□

The following proposition is an immediate corollary.

Corollary 2. *If $|W| \leq \omega$ and $\mathcal{Q}_\perp$ is the discrete partition on W, then $\mathfrak{P} = (W, \mathcal{O}, \mathcal{Q}_\perp)$ is solvable in the limit iff $(W, \mathcal{O}^*)$ is a T_D space.*

According to the Duhem-Quine thesis, a scientific theory becomes refutable by experience only when sufficiently *articulated* with auxiliary hypotheses and initial conditions. Furthermore, a theory can always be saved in the face of an anomalous observation by

[4] I am grateful to Alexandru Baltag, for showing me that construction in private communication.

tweaks to these auxiliary hypotheses.[5] Proposition 16 shows that inquiry can proceed successfully even in light of the Duhem-Quine thesis. The hallmark of a solvable problem is that each competing theory can be decomposed into at most countably many conditionally falsifiable *articulations*. One can think of each of these articulations as a theory fortified with enough auxiliary hypotheses as to make it refutable by experiment. Though a scientific theory may itself be unfalsifiable, each of its articulations must be (conditionally) falsifiable. Each competing theory may be re-articulated infinitely many times, but the problem remains solvable—inquiry eventually arrives at the true articulation and, by consequence, at the true theory.

Locally closed questions interact nicely with the simplicity order.

Proposition 17. *If $\mathfrak{P} = (W, \mathcal{O}, \mathcal{Q})$ is locally closed, then the elements of $\mathcal{Q}$ are partially ordered by $\preceq$.*

Proof. Follows immediately from Proposition 14 and the fact that $\preceq$ is a pre-order over all propositions. □

We conclude this section with a technical result, demonstrating that—in a sense to be made precise—any restriction of a solvable problem is also solvable. Problem restrictions can be thought of as representing the state of inquiry after some new information has been acquired, or some new presupposition has been imposed.

Definition 25 (Problem Restriction)**.** *If $\mathfrak{P} = (W, \mathcal{O}, \mathcal{Q})$ and $C \in \mathcal{P}(W)$, then $\mathcal{O}\big|_C = \{E \cap C : E \in \mathcal{O}\}$, $\mathcal{Q}\big|_C = \{Q \cap C : Q \in \mathcal{Q}\} \setminus \{\varnothing\}$ and $\mathfrak{P}\big|_C = (C, \mathcal{O}\big|_C, \mathcal{Q}\big|_C)$.*

Lemma 3. *If $\mathfrak{P}$ is a solvable problem and $C \in \mathcal{P}(W)$, then $\mathfrak{P}\big|_C$ is a solvable problem.*

Proof. Follows immediately from the equivalence between (1) and (3) demonstrated in Proposition 16 and the fact that a set is locally closed in the subspace topology $\mathcal{O}\big|_C$ iff it is the intersection of C with a locally closed set of $\mathcal{O}$. □

4.4 Simplicity and Learning

The method $\lambda_{\mathsf{MinFals}}$ succeeds by conjecturing the least answer in the enumeration that is downward closed in the simplicity order (falsifiable) in the subspace of current information. The method $\lambda_{\mathsf{MinFals}}$ relies on *both* the simplicity order and on the additional bias provided by the (arbitrary) enumeration of the answers. It is natural to ask: when does the simplicity order itself provide enough guidance to solve the problem? Define

$$\mathsf{Fals}(E) = \bigcup \{Q \in \mathcal{Q} : Q \cap E \text{ is closed in } E\}.$$

[5] For a classic discussion see Lakatos (1970, pp. 101-2)

Now define the method λ_{Fals} as follows:

$$\lambda_{\mathsf{Fals}}(E) = \begin{cases} \mathsf{Fals}(E) & \text{if } \mathsf{Fals}(E) \neq \varnothing \\ E & \text{otherwise.} \end{cases}$$

Though plausible, that method is not universal. It will only solve a problem if every answer eventually becomes the uniquely falsifiable answer and remains uniquely falsifiable until it is refuted.

Definition 26. *Problem $\mathfrak{P} = (W, \mathcal{O}, \mathcal{Q})$ is **simplicity driven** iff for every $w \in W$, there is $E \in \mathcal{O}(w)$ such that for all $F \in \mathcal{O}(w)$, $\mathcal{Q}(w) = \mathcal{Q}(\mathsf{Fals}(E \cap F))$.*

Proposition 18. *If $\mathfrak{P} = (W, \mathcal{O}, \mathcal{Q})$ is* **simplicity driven**, *then $\mathcal{Q}$ is locally closed.*

Proof. This is an immediate consequence of Proposition 3. □

Proposition 19. *Problem $\mathfrak{P}$ is solved by λ_{Fals} iff it is simplicity driven.*

Proof. $\Rightarrow$: Suppose $\mathfrak{P}$ is not simplicity driven. Then there is $w \in W$ such that for all $E \in \mathcal{O}(w)$, there is $F \in \mathcal{O}(w)$ such that $\mathcal{Q}(w) \neq \mathcal{Q}(\mathsf{Fals}(E \cap F))$. Suppose $E \in \mathsf{Lock}(\lambda_{\mathsf{Fals}}, w)$. Then there is F such that $\mathcal{Q}(w) \neq \mathcal{Q}(\mathsf{Fals}(E \cap F))$. That implies that $\mathcal{Q}(w) \neq \mathcal{Q}(\lambda(E \cap F))$, since otherwise $\mathcal{Q}(\lambda(E \cap F)) = \mathcal{Q}(\mathsf{Fals}(E \cap F))$. So we have that $\lambda_{\mathsf{Fals}}(E \cap F) \neq \mathcal{Q}(w)$, and, hence, $E \notin \mathsf{Lock}(\lambda_{\mathsf{Fals}}, w)$. $\Leftarrow$: Straightforward, by reversing the previous argument. □

Not all problems are simplicity driven. For such problems, some bias in addition to simplicity must be applied in order to solve the problem. Here is a straightforward example of such a problem.

Example 4. *Let $\mathcal{Q} := \{\{n\} : n \in \mathbb{N}\}$, $\mathcal{O}$ be the co-finite topology on $\mathbb{N}$, and $\mathfrak{P}_{\mathsf{CofinNat}} = (\mathbb{N}, \mathcal{O}, \mathcal{Q})$.*

This problem is not simplicity driven, since every answer compatible with the data is simplest. However, it is solvable in a straightforward way: conjecture the least natural number consistent with the data.

Chapter 5

Refining Success in the Limit

5.1 Varieties of Non-Monotonicity

Scientific inquiry is not monotonic: an accepted theory can be retracted or repudiated as new evidence comes to light. But while some degree of non-monotonicity is mandated by the pressure to converge to the truth, other varieties of non-monotonicity can be minimized, and sometimes, avoided altogether. Optimally truth-conducive methods are as monotonic as the problem allows. In this section we develop a range of non-monotonicity concepts that generalize mind changes and U-shapes from learning theory. First let us define monotonicity.

Definition 27. *Method λ is* **monotonic** *for a problem $\mathfrak{P} = (W, \mathcal{O}, \mathcal{Q})$ iff $\mathcal{Q}(\lambda(E \cap F)) \subseteq \mathcal{Q}(\lambda(E))$ for all consistent $E, F \in \mathcal{O}$.*

Monotonicity is not achievable unless each answer is verifiable (i.e., no inductive inference is required to arrive at the true answer).

Proposition 20. *Problem $\mathfrak{P} = (W, \mathcal{O}, \mathcal{Q})$ is solved by a monotonic method iff each $Q \in \mathcal{Q}$ is open.*

Proof. $\Rightarrow$: Suppose λ solves $\mathfrak{P}$ and $Q \in \mathcal{Q}$ is not open. Let $w \in \mathcal{Q} \setminus \mathcal{Q}^\circ$. Then $w \in \partial Q$. Let $E \in \mathsf{Lock}(\lambda, w)$. Since $w \in \partial Q$, there is $w' \in E \cap Q^c$. Let $Q' := \mathcal{Q}(w')$ and $F \in \mathsf{Lock}(\lambda, w')$. Then $\mathcal{Q}(\lambda(E \cap F)) = Q' \nsubseteq Q = \lambda(E)$. So λ is not monotonic. $\Leftarrow$: This direction follows straightforwardly by considering the method defined by $\lambda(E) = E$ for $E \in \mathcal{O}$. $\square$

We define three different forms of non-monotonicity, in decreasing generality. *Retraction*, the most general form of non-monotonicity, involves any belief change that is not a logical strengthening.

Definition 28 (Retractions)**.** *A* **λ-retraction sequence** *is a sequence of conjectures* $(\lambda(E_i))_{i=1}^n$ *such that for* $1 \leq i \leq n$*,* $\varnothing \neq E_i \in \mathcal{O}$ *and for* $1 \leq i < n$*,* $E_i \supset E_{i+1}$ *and* $\mathcal{Q}(\lambda(E_i)) \not\supseteq \mathcal{Q}(\lambda(E_{i+1}))$*.*

A retraction that also repudiates what was previously believed is called a *reversal*. This is a "sharper" form of non-monotonicity, since it involves believing the negation of what was previously believed.[1]

Definition 29 (Reversals)**.** *A* **λ-reversal sequence** *is a retraction sequence* $(\lambda(E_i))_{i=1}^n$ *such that* $\mathcal{Q}(\lambda(E_i))$ *is disjoint from* $\mathcal{Q}(\lambda(E_{i+1}))$ *for* $1 \leq i < n$*.*

A retraction sequence that terminates in a conjecture that entails the initial conjecture is called a *cycle*.[2]

Definition 30 (Cycles)**.** *A* **λ-cycle sequence** *is a reversal sequence* $(\lambda(E_i))_{i=1}^n$ *such that* $\mathcal{Q}(\lambda(E_n)) \subseteq \mathcal{Q}(\lambda(E_1))$*.*

We need a way to compare the retraction sequences of different methods.

Definition 31 (Comparing Conjecture Sequences)**.** *If* σ *is a* λ*-retraction sequence of length* n *and* δ *is a* λ'*-retraction sequence of length* n*, then* $\sigma \leq \delta$ *iff* $\mathcal{Q}(\sigma_i) \subseteq \mathcal{Q}(\delta_i)$ *for* $1 \leq i \leq n$*.*

If $\sigma \leq \delta$, we say that σ retracts *at least as sharply* as δ. Note that the relation $\leq$ is a partial order over conjecture sequences. Finally, we define the technically useful notion of a *singleton* retraction sequence.

Definition 32 (Singleton Sequence)**.** *A* **singleton** *reversal sequence* δ *is reversal sequence in which for each* i*,* $\mathcal{Q}(\delta_i) \in \mathcal{Q}$*.*

5.2 Avoiding Cycles

We have shown in Proposition 20 that reversals are not avoidable so long as any answer is not verifiable. Cycles, however, are avoidable in a broad class of problems. In this section we will show that the avoidance of cycles is equivalent to a weakening of some standard rationality principles from belief revision. We also show that this postulate mandates a version of Ockham's razor. Finally, we will characterize the class of problems which are solvable without cycles. First, let us define what it means to be cycle free.

[1] In the formal learning theory literature, learners are defined so as to make maximally specific conjectures. In that setting, every retraction is a reversal. We allow methods to entertain disjunctive beliefs about the true theory, so the distinction does not collapse.

[2] In formal learning theory, learning performance that exhibits cycles is called "U-shaped," referring to the linguistic proficiency curves of children that learn a grammatical construction, unlearn it, and then learn it again (Carlucci et al., 2005; Carlucci and Case, 2013).

Definition 33 (Cycle Free)**.** *Method λ is* ***$\mathcal{Q}$-cycle free*** *for $\mathfrak{P} = (W, \mathcal{O}, \mathcal{Q})$ iff there exist no λ-cycle sequences for $\mathfrak{P}$.*[3]

5.2.1 As a Principle of Rational Revision

We first restate two rationality postulates—familiar from belief revision and nonmonotonic logic—in our own terms.[4]

Definition 34 ($\mathcal{Q}$-Conditionalization)**.** *Method λ satisfies* ***$\mathcal{Q}$-conditionalization*** *iff $\mathcal{Q}(\lambda(E)) \cap \mathcal{Q}(E \cap F) \subseteq \mathcal{Q}(\lambda(E \cap F))$.*

This postulate says that the new belief state is no stronger than the conjunction of the old belief state with the new data, so long as there are any answers compatible with both. In slogan form, it says *no induction, without refutation.* Conditionalization is meant to prevent "jumping to conclusions"—but that does not seem right for inductive inquiry. If the relation between two measured variables satisfies repeated tests of linearity $(e_1, ..., e_n)$ it should eventually be reasonable to conclude h_{lin}, that the relationship is linear, even though the observations do not logically entail the conclusion. That seems true even if no prior theory was refuted, as the inquiry may well have begun with only a very general hypothesis about the form of the relationship. The only way that can happen in the AGM framework is if the conditional belief $\bigwedge_{1 \leq i \leq n} e_i \rightarrow h_{\mathsf{lin}}$ was already in the original belief state of the researchers (Schurz, 2011). Thus, so long as no refutation occurs, there is no induction—scientific inquiry is simply a matter of working out the deductive consequences of pre-existing commitments. That does not strike us as descriptively adequate for scientific belief revision. More plausibly, it is a matter of leaping to the simplest hypothesis compatible with the data.

Definition 35 (Rational $\mathcal{Q}$-monotonicity)**.** *Method λ is* ***rationally $\mathcal{Q}$-monotone*** *iff $\mathcal{Q}(\lambda(E \cap F)) \subseteq \mathcal{Q}(\lambda(E)) \cap \mathcal{Q}(E \cap F)$ whenever $\mathcal{Q}(\lambda(E)) \cap \mathcal{Q}(E \cap F) \neq \varnothing$.*

This postulate says that so long as there are any answers compatible with both the new information and the previous belief state, the new belief state entails them. In slogan form, it says *no retraction, without refutation.* There are good reasons to think that rationality principle is also too strong for inductive inquiry. Many have observed that this principle precludes undercutting defeat[5] of inductive inferences (Gärdenfors, 1988;

[3]Cycle sequences are problem relative, but we often leave out mention of the problem when it is clear from context.

[4]A learner can be thought of as a skeptical consequence relation by setting $A \mid\!\sim B$ iff $\lambda(A) \subseteq B$. The analogy is not perfect since the learner is only defined on the topological basis, not on a Boolean algebra. Nevertheless, rational $\mathcal{Q}$-conditionalization and $\mathcal{Q}$-monotonicity are intended as an analogues to principles (K^*7) and (K^*8) of AGM revision respectively and the defeasible inference principles of the same name.

[5]An undercutting defeat does not license inference to the contradictory conclusion, but calls into doubt a *prima facie* reason for the conclusion (Koons, 2014).

Schurz, 2011; Koons, 2014). We give an example in which rational $\mathcal{Q}$-monotonicity is violated by a plausible application of Ockham's razor.

Example 5. *Suppose that a scientific discipline considers the following three mutually exclusive hypotheses possible, where the free parameters are allowed to range over all non-zero values.*

$$\begin{array}{ccccccccc} & H_1 & & & H_2 & & & H_3 & \\ Y & = & \mathsf{sin}(X) & Y & = & \mathsf{sin}(X) & Y & = & X^2 \\ Z & = & bX + a & Z & = & cX^2 + bX + a & Z & = & cX^2 + bX + a. \end{array}$$

Hypothesis H_1 is clearly simpler than H_2. Any finite, inexact but arbitrarily precise data set compatible with H_1 is also compatible with H_2 (tune c as close to 0 as required). On the other hand, H_1 is not clearly simpler than H_3—the quadratic function is nothing like the sine function—and a competent scientist can distinguish the two with a sufficiently large or accurate sample. Suppose laboratory A is gearing up its accelerator to settle whether H_3 is true, and will complete the test by next year. For some time, laboratory B has been diligently measuring X and Z with increasing precision, and the results confirm (without certainty) that the relationship is linear. It is at least plausibly rational to reason as follows. If Y is a sine function of X, then H_3 must be false. Since that question will be settled next year, there is no harm now in concluding that either H_1 or H_3 is true. One *might* leap immediately to H_1, but it suffices for our purposes that believing the disjunction of H_1 and H_3 is not irrational. Now, suppose that laboratory B discovers non-linearity of Z prior to the date of laboratory A's experiment. Now H_1 is refuted, but the disjunction $H_1 \vee H_3$ is not. Rational $\mathcal{Q}$-monotonicity mandates belief in H_3, prior to the fateful experiment. That seems rash, since H_2 appears to be equally simple with H_3, and within the year laboratory A will decide between them. It seems more rational to continue to suspend belief between H_2 and H_3, depending on the outcome observed in laboratory A.[6] We can describe the situation by saying that H_2 being more complex than H_1 was the only *reason* we had to exclude H_2 (by Ockham's razor), and refutation of H_1 defeats that reason without literally refuting the disjunctive belief H_1 or H_3.

Neither rational monotonicity nor conditionalization seem quite suited for inductive inquiry. One can plausibly[7] weaken these two postulates in the following way:

Definition 36 (Reversal $\mathcal{Q}$-monotonicity)**.** *Method λ is* ***reversal $\mathcal{Q}$-monotone*** *iff $\mathcal{Q}(\lambda(E\cap F))$ meets $\mathcal{Q}(\lambda(E)) \cap \mathcal{Q}(E \cap F)$ whenever $\mathcal{Q}(\lambda(E)) \cap \mathcal{Q}(E \cap F) \neq \varnothing$.*

[6] This is exactly what a simplicity-driven method like λ_{Fals} would do.

[7] It is at least *prima facie* plausible that if one receives data compatible with a prior belief state, one should not conclude the *negation* of what one previously believed. But there are reasons to doubt even that very weak principle of rational belief change. Suppose I believe that H_2 is true, but have only seen data compatible with H_1. Then the pressure to converge to the truth can make we conclude H_1, although I have seen no data incompatible with H_2. A more defensible principle might be that if one receives data compatible with a prior belief state, one should not conclude the *negation* of what one previously believed, unless that is absolutely mandated by convergence to the true theory.

In slogan form that postulate says *no reversal, without refutation.* So long as there are some answers compatible with both the new information and the previous belief state, the new belief state ought to entertain some previous answer as a possibility. This postulate insists that *something* be preserved from the old belief state, but does not say what, or how much. Once the requirement of convergence is imposed, that very weak rationality principle is equivalent to the total avoidance of cycles.

Proposition 21. *If a method λ solves $\mathfrak{P}$, then λ is $\mathcal{Q}$-cycle free iff λ is reversal $\mathcal{Q}$-monotone.*

Proof. $\Leftarrow$: Suppose λ performs a cycle. Then for some consistent $E, F \in \mathcal{O}$, $\mathcal{Q}(\lambda(E \cap F))$ is disjoint from $\mathcal{Q}(\lambda(E))$. Furthermore, for some $G \in \mathcal{O}$, such that G consistent with $E \cap F$, $\mathcal{Q}(\lambda(E \cap F \cap G)) \subseteq \mathcal{Q}(\lambda(E))$. By deductive cogency, $\varnothing \neq \mathcal{Q}(\lambda(E \cap F \cap G)) \subseteq \mathcal{Q}(E \cap F \cap G) \subseteq \mathcal{Q}(E \cap F)$. So $\mathcal{Q}(E \cap F) \cap \mathcal{Q}(\lambda(E)) \neq \varnothing$. So λ fails to be reversal $\mathcal{Q}$-monotone. $\Rightarrow$: Suppose λ is not reversal $\mathcal{Q}$-monotone. Then there are consistent $E, F \in \mathcal{O}$ such that $\mathcal{Q}(\lambda(E \cap F))$ is disjoint from $\mathcal{Q}(\lambda(E)) \cap \mathcal{Q}(E \cap F)$ although $\mathcal{Q}(\lambda(E)) \cap \mathcal{Q}(E \cap F) \neq \varnothing$. Let $w \in E \cap F$ such that $\mathcal{Q}(w) \subseteq \mathcal{Q}(\lambda(E)) \cap \mathcal{Q}(E \cap F)$ and $G \in \mathsf{Lock}(\lambda, w)$. Then $\mathcal{Q}(\lambda(E \cap F \cap G)) = \mathcal{Q}(w) \subseteq \mathcal{Q}(\lambda(E))$ and λ performs a cycle. $\square$

So long as you restrict to methods that solve the problem in the limit, endorsing any of the two rationality principles implicitly endorses the avoidance of cycles. This result sharpens our suggestion, at the outset, of a rapprochement between belief revision theory and learning theory—avoiding"U-shaped" learning is shown to be equivalent to a weakened rationality principle, given that the method converges to the truth.

5.2.2 A Necessary Ockham Principle

In this section we show that reversal monotonicity—and therefore any stronger rationality principle—entails a version of Ockham's razor, so long as attention is restricted to methods that solve the problem in the limit.

Definition 37 ($\mathcal{Q}$-Vertical Ockham)**.** *Method λ is* **$\mathcal{Q}$-vertical Ockham** *iff for all $E \in \mathcal{O}$, $\mathcal{Q}(\lambda(E))$ is downward-closed in E.*

This methodological condition stops short of requiring that the learner's conjecture be *minimal* in the simplicity order, mandating only that any simpler answer consistent with the data be admitted as a possibility.[8] The vertical razor is content-neutral: it does not say whether the conjecture has to be weak or strong, requiring only that it be downward-closed. By Proposition 12, the vertical razor is equivalent to the requirement that the

[8]In other words, if one pays an epistemic *compliment* to a complex possibility by including it in one's conjecture, one must pay the same compliment to any simpler answer consistent with the data.

learner's conjecture be *falsifiable* at every stage of inquiry. That is precisely Popper's requirement of falsifiability—an enjoiner for bold conjectures, vulnerable before the tribunal of experience.

Proposition 22. *If a method λ solves $\mathfrak{P}$ and is reversal $\mathcal{Q}$-monotone, it is $\mathcal{Q}$-vertical Ockham.*

Proof. Suppose that λ is not $\mathcal{Q}$-vertical Ockham. Then for some $E \in \mathcal{O}$ there is $w \in E \cap \overline{\mathcal{Q}(\lambda(E))} \setminus \mathcal{Q}(\lambda(E))$. Let $F \in \mathsf{Lock}(\lambda, w)$, then $\mathcal{Q}(\lambda(E \cap F)) = \mathcal{Q}(w)$ is disjoint from $\mathcal{Q}(\lambda(E))$, although $\mathcal{Q}(E \cap F)$ meets $\mathcal{Q}(\lambda(E))$, since $E \cap F$ meets $\mathcal{Q}(\lambda(E))$. $\square$

The moral of the previous result is that combining the principles of rational belief change with a mandate of limiting convergence leads to highly intuitive, methodological principles that are not mandated by either norm individually. Perhaps most surprising is that if you endorse both limiting convergence and some weak rationality principle, you must also endorse Popper's falsificationist program. It was not our aim at the outset to vindicate Popper's views. Rather, they emerged in a surprising way from our topological investigation of simplicity and optimal truth-conduciveness.

To see that the converse of Proposition 22 does not hold, consider the following example.

Example 6. *Let $W = \{a, b, c\}$, $\mathcal{O} = \{\{a, b, c\}, \{a\}, \{b, c\}, \{c\}, \{a, c\}\}$, and $\mathcal{Q} = \{\{a\}, \{b\}, \{c\}\}$. Let $\lambda(\{a, b, c\}) = \{a\}, \lambda(\{a\}) = \{a\}, \lambda(\{b, c\}) = \{b\}, \lambda(\{c\} = \{c\}$ and $\lambda(\{a, c\}) = \{c\}$. Then λ is vertical $\mathcal{Q}$-Ockham and solves $\mathfrak{P} = (W, \mathcal{O}, \mathcal{Q})$, but it is not reversal $\mathcal{Q}$-monotone, since $\lambda(\{a, b, c\}) \cap \lambda(\{a, c\}) = \varnothing$ although $\{a, b, c\}$ and $\{a, c\}$ are consistent.*

The Ockham Principle Stated Diachronically

We had originally defined the vertical razor as a synchronic condition: all conjectures must be falsifiable in the current information state. We can also state an equivalent, *diachronic* formulation, characterizing how vertical Ockham methods *respond* to new data.

Proposition 23. *Method λ is $\mathcal{Q}$-vertical Ockham iff for all consistent $E, F \in \mathcal{O}$, $\mathcal{Q}(\lambda(E \cap F)) \not\subseteq \mathcal{Q}(\mathcal{Q}(\check{\lambda}(E)))$.*

Proof. $\Rightarrow$: If $\mathcal{Q}(\lambda(E))$ is closed in E, then $\mathcal{Q}(\mathcal{Q}(\check{\lambda}(E))) = \mathcal{Q}(\check{\lambda}(E)) = \varnothing$. $\Leftarrow$: Suppose $\mathcal{Q}(\lambda(E))$ is not closed in E. Let $w \in E \cap \mathcal{Q}(\check{\lambda}(E))$. Let $F \in \mathsf{Lock}(\lambda, w)$. Then $\mathcal{Q}(\lambda(E \cap F)) \subseteq \mathcal{Q}(\mathcal{Q}(\check{\lambda}(E)))$. $\square$

Intuitively, what the method conjectures *now* is not consistent with any possibility simpler than (but not entailing) what it conjectured *before*. The vertical razor says that the complexity of conjectures *can only increase* in additional information—more information will never elicit a simpler conjecture.

5.2.3 Is Cycle-Avoidance Restrictive?

Not all problems can be solved without cycles. Consider the following example.

Example 7. *Let* $\mathcal{Q}_{\mathsf{prty}} := \{\{2n : n \in \mathbb{N}\}, \{2n+1 : n \in \mathbb{N}\}\}$*,* $\mathcal{O}_{\mathsf{upset}} := \{\{m : m \geq n\} : n \in \mathbb{N}\}$ *and* $\mathfrak{P}_{\mathsf{prty}} = (\mathbb{N}, \mathcal{O}_{\mathsf{upset}}, \mathcal{Q}_{\mathsf{prty}})$*.*

This problem asks whether the true natural number is even or odd. Every learner can be forced to perform arbitrarily many cycles. The difficulty is that the question is too coarse. It can be refined in an obvious way to yield a problem that is solvable without cycles:

Example 8. *The problem* $\mathfrak{P}_{\mathsf{prty}}$ *is refined by the natural problem* $\mathfrak{P}_{\mathsf{num}} = (\mathbb{N}, \mathcal{O}_{\mathsf{upset}}, \mathcal{Q}_{\mathsf{num}})$ *where* $\mathcal{Q}_{\mathsf{num}} := \{\{n\} : n \in \mathbb{N}\}$*.*

This problem asks which is the true natural number. It is solved without cycles in the obvious way: at each stage, conjecture the least natural compatible with the data. The example illustrates the general fact that every solvable problem can be refined to a problem that is solvable without cycles. The remainder of this section proves that result. A related result is demonstrated in Baltag et al. (2014), where the technical notion of *pseudo-stratification* is first defined.

Definition 38 (Psuedo-Stratified Problems)**.** *Problem* $\mathfrak{P} = (W, \mathcal{O}, \mathcal{Q})$ *is* **pseudo-stratified** *iff it is locally closed and there exists a total well-order* $<$ *on* $\mathcal{Q}$*, with order type less than or equal to* ω*, such that for all* $A, B \in \mathcal{Q}$*, we have that*

$$\textit{if } A < B, \textit{ then either } B \subseteq \overline{A} \textit{ or } B \subseteq \overline{A}^c.$$

Proposition 24. *If* $\mathfrak{P} = (W, \mathcal{O}, \mathcal{Q})$ *is a solvable problem, then there is* $\mathfrak{P}' = (W, \mathcal{O}, \mathcal{Q}')$ *such that* $\mathcal{Q}'$ *refines* $\mathcal{Q}$ *and* $\mathfrak{P}'$ *is pseudo-stratified.*

Proof. See Proposition 2 in Baltag et al. (2014). □

Proposition 25. *If* $\mathfrak{P} = (W, \mathcal{O}, \mathcal{Q})$ *is pseudo-stratified, then it is solvable by a rationally* $\mathcal{Q}$*-monotone learner.*

Proof. Define $\mathsf{OrdMin}(E)$ to be the set

$$\{B \in \widehat{\mathcal{Q}}(E) : B \text{ is closed in } E \text{ and, for each } A \in \widehat{\mathcal{Q}}(E), \text{ if } A < B, \text{ then } B \subseteq \overline{A}.\}$$

Now define:

$$\lambda_{\mathsf{OrdMin}}(E) = \begin{cases} \text{the } <\text{-least element of } \mathsf{OrdMin}(E) & \text{if } \mathsf{OrdMin}(E) \neq \varnothing \\ E & \text{otherwise.} \end{cases}$$

First, we show that $\lambda_{\mathsf{OrdMin}}$ is a learner. Let $w \in Q \in \mathcal{Q}$. Since Q is locally closed, it is a closed subset of some $E \in \mathcal{O}(w)$. Furthermore, there is $F \in \mathcal{O}(w)$ such that if $A < Q$ and $Q \subseteq \overline{A}^c$, then $A \notin \mathcal{Q}(F)$. Then $E \cap F \in \mathsf{Lock}(\lambda_{\mathsf{OrdMin}}, w)$, since for any $G \in \mathcal{O}(w)$, Q remains closed in $E \cap F \cap G$ and—since $\mathfrak{P}$ is psuedo-stratified—if $A < Q$, then $Q \subseteq \overline{A}$.

It remains to show that $\lambda_{\mathsf{OrdMin}}$ is rationally $\mathcal{Q}$-monotone. Note that for all $E \in \mathcal{O}$, $\lambda(E) = \mathcal{Q}(\lambda(E))$. Now, suppose for consistent $E, F \in \mathcal{O}$, $\lambda(E \cap F) \nsubseteq \lambda(E)$. By deductive cogency, it cannot be the case that $\lambda(E) = E$, so there exists $Q \in \mathcal{Q}$ such that $\lambda(E) = Q$. Suppose for contradiction that $Q \in \widehat{\mathcal{Q}}(E \cap F)$. Then it is still the case that $Q \in \mathsf{OrdMin}(E \cap F)$. So there exists $A \in \mathsf{OrdMin}(E \cap F)$ such that $A < Q$. But since $Q \subseteq \overline{A}$, by assumption, and Q, A are disjoint, A cannot be closed in $E \cap F$. Contradiction. So $\mathcal{Q}(\lambda(E)) \cap \mathcal{Q}(E \cap F) = \varnothing$ as required. □

Since every rationally $\mathcal{Q}$-monotone learner is also reversal $\mathcal{Q}$-monotone, we have the following Proposition as a consequence of the previous two.

Proposition 26. *If $\mathfrak{P} = (W, \mathcal{O}, \mathcal{Q})$ is a solvable problem then there is $\mathfrak{P}' = (W, \mathcal{O}, \mathcal{Q}')$ such that $\mathcal{Q}'$ refines $\mathcal{Q}$ and $\mathfrak{P}'$ is solved in the limit by a $\mathcal{Q}'$-cycle free method.*

By the previous Proposition and Proposition 22 we also have the following.

Proposition 27. *If $\mathfrak{P} = (W, \mathcal{O}, \mathcal{Q})$ is a solvable problem, then there is $\mathfrak{P}' = (W, \mathcal{O}, \mathcal{Q}')$ such that $\mathcal{Q}'$ refines $\mathcal{Q}$ and $\mathfrak{P}'$ is solved in the limit by a $\mathcal{Q}'$-vertical Ockham method.*

5.3 Minimizing Reversals

Proposition 20 shows that reversals cannot be avoided in genuinely inductive problems. But that does not mean that they cannot be minimized. We propose that a good method performs only those reversals that are mandated by the requirement of limiting convergence. In other words, performing a reversal is epistemically excusable if and only if every method that solves the problem can be forced to do the same. We make that notion precise in the following definition.

Definition 39 (Forcible Sequences)**.** *A λ-reversal sequence δ is* **forcible** *iff for every λ' that solves $\mathfrak{P}$, there is a λ'-reversal sequence σ, such that $\sigma \leq \delta$.*

We say that a method is *optimal* if all of its reversals are forcible.

Definition 40 (Optimality)**.** *Method λ is* ***reversal optimal*** *for $\mathfrak{P}$ iff λ solves $\mathfrak{P}$ and every λ-reversal sequence is forcible.*

Forcible sequences are illustrated by the following example. Suppose that you are trying to learn the polynomial degree of a function from inexact data with bounded error. Nature can present linear-looking data until you conjecture the “linear” hypothesis. Then, nature can present quadratic-looking data until you conjecture "quadratic." It is easy to see how to continue. In that problem, the sequence "linear", "quadratic", "cubic", ..., is *forcible* by Nature. The forcible sequences in any given problem can be given a topological characterization that extrapolates what is essential in the example.

Proposition 28. *A λ-reversal sequence $\delta = ((\lambda(E_i))_{i=1}^n$ is forcible iff*

$$\mathcal{Q}(\lambda(E_1)) \cap \overline{\ldots \cap \overline{\mathcal{Q}(\lambda(E_{n-2})) \cap \overline{\mathcal{Q}(\lambda(E_{n-1})) \cap \overline{\mathcal{Q}(\lambda(E_n))}}}} \neq \varnothing.$$

Proof.
$\Leftarrow$: Suppose λ' is a learner. Let

$$w_1 \in \mathcal{Q}(\lambda(E_1)) \cap \overline{\ldots \cap \overline{\mathcal{Q}(\lambda(E_{n-2})) \cap \overline{\mathcal{Q}(\lambda(E_{n-1})) \cap \overline{\mathcal{Q}(\lambda(E_n))}}}}$$
$$w_{i+1} \in \bigcap_{j=1}^{i} \mathsf{Lock}(\lambda', w_j) \cap \mathcal{Q}(\lambda(E_{i+1})) \cap \overline{\ldots \cap \overline{\mathcal{Q}(\lambda(E_{n-2})) \cap \overline{\mathcal{Q}(\lambda(E_{n-1})) \cap \overline{\mathcal{Q}(\lambda(E_n))}}}}$$

Then $\sigma = (\lambda'(\bigcap_{j=1}^{i} \mathsf{Lock}(\lambda', w_j)))_{i=1}^n$ is a reversal sequence and $\sigma \leq \delta$.

$\Rightarrow$: Suppose $\mathcal{Q}(\lambda(E_1)) \cap \overline{\ldots \cap \overline{\mathcal{Q}(\lambda(E_{n-2})) \cap \overline{\mathcal{Q}(\lambda(E_{n-1})) \cap \overline{\mathcal{Q}(\lambda(E_n))}}}} = \varnothing$. Let

$$C := \overline{\ldots \cap \overline{\mathcal{Q}(\lambda(E_{n-2})) \cap \overline{\mathcal{Q}(\lambda(E_{n-1})) \cap \overline{\mathcal{Q}(\lambda(E_n))}}}}.$$

The restricted problems $\mathfrak{P}\big|_C$, $\mathfrak{P}\big|_{C^c}$ are solvable by Lemma 3. Let λ' solve $\mathfrak{P}\big|_C$ and λ'' solve $\mathfrak{P}\big|_{C^c}$. Now define a new method for the original problem:

$$\lambda^*(E) = \begin{cases} \lambda'(E \cap C) & \text{if } E \cap C \neq \varnothing \\ \lambda''(E) & \text{if } E \cap C = \varnothing \end{cases}$$

Since C is refutable, it is possible to focus on the $\mathfrak{P}\big|_C$ subproblem until it is refuted. So λ^* solves $\mathfrak{P}$. Now, since by assumption $\mathcal{Q}(\lambda(E_1)) \cap C = \varnothing$, if $\mathcal{Q}(\lambda^*(E)) \subseteq \mathcal{Q}(\lambda(E_1))$, C has been refuted. So there can be no λ^*-reversal sequence σ such that $\sigma \leq \delta$. □

Note that the same proof strategy as in the right-to-left proof of Proposition 28 also demonstrates the following Proposition.

Proposition 29. *δ is a forcible λ-reversal sequence iff there is a forcible singleton λ-reversal sequence σ such that $\sigma \leq \delta$.*

Locally closed problems have a special feature: no singleton cycle sequences are forcible. We prove that result by means of the following two propositions.

Lemma 4. *If $Q_1, Q_2 \in \mathcal{Q}$ are distinct and locally closed, then $Q_1 \cap \overline{Q_2 \cap \overline{Q_1}} = \varnothing$.*

Proof. Note that $Q_2 \cap \overline{Q_1} \subseteq \overset{\vee}{Q_1}$ and $Q_1 \subseteq \left(\overset{\vee}{Q_1}\right)^{\mathrm{c}}$. So by Proposition 13, $\left(\overset{\vee}{Q_1}\right)^{\mathrm{c}}$ is an open set containing Q_1, but disjoint from $Q_2 \cap \overline{Q_1}$. Therefore $Q_1 \cap \overline{Q_2 \cap \overline{Q_1}} = \varnothing$. □

As a direct corollary of the previous Lemma, we have the following proposition.

Corollary 3. *If $\mathfrak{P}$ is a locally closed problem, then no singleton cycle sequence is forcible.*

5.3.1 A Necessary Ockham Principle

In this section we show that a version of Ockham's razor is necessary for reversal optimality. We state our second Ockham principle as follows.

Definition 41 (Horizontal Ockham)**.** *Method λ is* $\mathcal{Q}$**-horizontal Ockham** *iff for all $E \in \mathcal{O}$ and $Q \in \widehat{\mathcal{Q}}(E)$, there is $Q' \in \widehat{\mathcal{Q}}(\lambda(E))$ such that $Q' \cap \overline{Q} \neq \varnothing$.*

Recall that in Example 5, we had that $H_1 \preceq H_2$ and that H_3 was unordered with the other two. At the outset, our scientists conjectured the disjunction $H_1 \vee H_3$. Their conjecture entails $\neg H_2$. In our story, the scientists accepted the negation of H_2 *because* they conceded the simpler possibility H_1. Similarly, if a horizontal Ockham conjecture entails the negation of a particular answer, it is because it concedes a simpler possibility. A horizontal Ockham method always has a simplicity-based *reason* for *dis*believing an answer.

Unlike the vertical razor, the horizontal razor favors logically weak conjectures. If there are many answers compatible with the data, it must allow a possibility as simple as each of them. While the vertical razor is satisfied as long as the method conjectures *any* currently refutable answer, the horizontal razor is satisfied only if the method disjoins all of them.

The horizontal razor is a necessary condition for reversal optimality. If reversal optimality is an epistemic norm, the horizontal razor is *necessary* for inductive, epistemic justification.

Proposition 30. *If method λ is reversal optimal for $\mathfrak{P} = (W, \mathcal{O}, \mathcal{Q})$, then λ is $\mathcal{Q}$-horizontal Ockham.*

Proof. Suppose λ is not $\mathcal{Q}$-horizontal Ockham. Then for some $E \in \mathcal{O}$ there is $Q \in \widehat{\mathcal{Q}}(E)$ such that for all $Q' \in \widehat{\mathcal{Q}}(\lambda(E))$, $Q' \cap \overline{Q} = \varnothing$. Therefore $\mathcal{Q}(\lambda(E)) \cap \overline{Q} = \varnothing$. Let $w \in E \cap Q$ and $F \in \mathsf{Lock}(\lambda, w)$. Then $\mathcal{Q}(\lambda(E \cap F)) = Q$ is disjoint from $\mathcal{Q}(\lambda(E))$. So $(\lambda(E), \lambda(E \cap F))$ is a λ-reversal sequence, but $\mathcal{Q}(\lambda(E)) \cap \overline{\mathcal{Q}(\lambda(E \cap F))} = \varnothing$. So by Proposition 28, that sequence is not forcible, and λ is not reversal optimal. □

The previous result shows that reversal optimality comes at the cost of logical content. To avoid unnecessary reversals, inquiry must proceed cautiously. This caution is justified so long as reversal-optimality is the operative norm. To see that the converse of Proposition 30 does not hold, consider the following example.

Example 9. *Let $W = \{a, b, c, d\}$. Let $Q_1 = \{a\}$, $Q_2 = \{b, c\}$, $Q_3 = \{d\}$ and $\mathcal{Q} = \{Q_1, Q_2, Q_3\}$. Let $\mathcal{O}$ be the closure of $\{\{a, b, d\}, \{b\}, \{c, d\}, \{d\}\}$ under unions. Now let $\lambda(W) = \{a\}$, $\lambda(\{c, d\}) = \{c\}$, and $\lambda(\{d\}) = \{d\}$. Then $\lambda(W) = Q_1$ is horizontal Ockham since $Q_1 \cap \overline{Q_2} = Q_1 \neq \varnothing$ and $Q_1 \cap \overline{Q_3} = Q_1 \neq \varnothing$. But $(\lambda(W), \lambda(\{c, d\}, \lambda(\{d\}))$ is not forcible since $\mathcal{Q}(\lambda(\{c, d\}) \cap \overline{\mathcal{Q}}(\lambda(\{d\}) = Q_2 \cap \overline{Q_3} = \{c\}$ is disjoint from $\mathcal{Q}(\lambda(W)) = Q_1 = \{a\}$.*

The Ockham Principle Stated Diachronically

We can also give a diachronic formulation of the horizontal Ockham principle.

Proposition 31. *If method λ solves $\mathfrak{P}$, then λ is* **$\mathcal{Q}$-horizontal Ockham** *iff for all consistent $E, F \in \mathcal{O}$, $\mathcal{Q}(\lambda(E)) \cap \overline{\mathcal{Q}(\lambda(E \cap F))} \neq \varnothing$.*

Proof. $\Rightarrow$: Suppose that λ is $\mathcal{Q}$-horizontal Ockham. Let $Q \in \widehat{\mathcal{Q}}(\lambda(E \cap F))$. By deductive cogency, $Q \in \widehat{\mathcal{Q}}(E \cap F) \subseteq \widehat{\mathcal{Q}}(E)$. Then, since λ is horizontal Ockham, there is $Q' \in \widehat{\mathcal{Q}}(\lambda(E))$ such that $Q' \cap \overline{Q} \neq \varnothing$. So $\mathcal{Q}(\lambda(E)) \cap \overline{\mathcal{Q}(\lambda(E \cap F))} \neq \varnothing$ as required. $\Leftarrow$: Suppose that λ is not $\mathcal{Q}$-horizontal Ockham. Then there is $E \in \mathcal{O}$ and $Q \in \widehat{\mathcal{Q}}(E)$ such that $\mathcal{Q}(\lambda(E)) \cap \overline{Q} = \varnothing$. Letting $w \in Q$ and $F \in \mathsf{Lock}(\lambda, w)$ we have that $\mathcal{Q}(\lambda(E)) \cap \overline{\mathcal{Q}(\lambda(E \cap F))} = \varnothing$. □

In other words, whatever a horizontal Ockham principle conjectures *now*, it previously entertained possibility at least as simple.

5.3.2 Is Reversal Optimality Restrictive?

Not all problems can be solved by a reversal optimal learner. For a dramatic example, we reuse Example 4.

Example 4. *Let $\mathcal{Q} := \{\{n\} : n \in \mathbb{N}\}$, $\mathcal{O}$ be the co-finite topology on $\mathbb{N}$, and $\mathfrak{P}_{\mathsf{CofinNat}} = (\mathbb{N}, \mathcal{O}, \mathcal{Q})$.*

If λ is $\mathcal{Q}$-horizontal Ockham, then $\mathcal{Q}(\lambda(E)) = \mathcal{Q}(E)$ for all $E \in \mathcal{O}$. A $\mathcal{Q}$-horizontal Ockham learner can never perform the inductive leap necessary to converge in that problem. Intuitively, horizontal Ockham learners rely on the eventual emergence of a uniquely simplest answer. In this section we characterize the set of problems that can be solved by a reversal optimal method. It will turn out that a problem is solvable by a reversal optimal method if and only if it is ***simplicity driven***, as defined in Section 4.4. First we need to get a handle on exactly which singleton sequences (henceforth *paths*) are forcible.

Definition 42 (Forcible Paths)**.** *We define the **forcible paths in** $\mathfrak{P}$ by a recursion on the length of paths.*

$$\Pi_1 = \{Q : Q \in \mathcal{Q}\}$$
$$\Pi_n = \{Q \cap \overline{P} : Q \in \mathcal{Q} \text{ and } P \in \Pi_{n-1}\} \setminus \left\{ \bigcup_{i=1}^{n-1} \Pi_i \cup \{\varnothing\} \right\}.$$

So the set of all forcible paths $\Pi = \bigcup_{i=1}^{\infty} \Pi_i$. For $E \in \mathcal{O}$, define the restriction $\Pi\big|_E$ to be the set of all forcible paths in $\mathfrak{P}\big|_E$.

Maximal paths are downward-closed in the simplicity order.

Definition 43 (Maximal Paths)**.** *A forcible path $P \in \Pi\big|_E$ is **maximal** iff P is closed (falsifiable) in E. Define $\mathsf{Max}(E)$ to be the set of all maximal paths in $\Pi\big|_E$.*

The maximal paths are exactly those that cannot be extended by a longer path. The following makes that notion precise.

Definition 44 (Path Extensions)**.** *One forcible path P_2 **extends** another path P_1 iff $P_2 = Q_1 \cap \overline{Q_2 \cap \overline{\ldots \cap \overline{P_1}}}$.*

For example, the path $Q_1 \cap \overline{Q_2 \cap \overline{Q_3}}$ extends the path $Q_2 \cap \overline{Q_3}$ which in turn extends Q_3.

Proposition 32. *Path P has no non-trivial extension iff P is maximal.*

Proof. Suppose $P = Q_2 \cap \overline{Q_3 \cap \overline{\ldots \cap \overline{Q_n}}}$ is maximal. Then for any $Q_1 \in \mathcal{Q}$, the extension of P, $Q_1 \cap \overline{Q_2 \cap \overline{Q_3 \cap \overline{\ldots \cap \overline{Q_n}}}} = Q_1 \cap \overline{Q_2} \cap \overline{Q_3 \cap \overline{\ldots \cap \overline{Q_n}}}$, which—since $\mathcal{Q}$ is a partition—is nonempty iff $Q_1 = Q_2$. Suppose P is not maximal. Let $w \in \overline{Q_2 \cap \overline{Q_3 \cap \overline{\ldots \cap \overline{Q_n}}}} \setminus Q_2 \cap \overline{Q_3 \cap \overline{\ldots \cap \overline{Q_n}}}$. Then, since $w \in \overline{Q_3 \cap \overline{\ldots \cap \overline{Q_n}}}$, $w \notin Q_2$. Letting $Q_1 := \mathcal{Q}(w)$, we have that $Q_1 \cap \overline{Q_2 \cap \overline{Q_3 \cap \overline{\ldots \cap \overline{Q_n}}}}$ is a non-trivial extension of P. □

Definition 45 (Path Roots)**.** *The **root** of a path P is defined as $\mathcal{Q}(P)$. Let $\mathsf{RootMax}(E) = \{\mathcal{Q}(P) : P \in \mathsf{Max}(E)\}$.*

Definition 46 (Tree-like Problems)**.** *Problem $\mathfrak{P}$ is **tree-like** iff (1) every forcible path in $\mathfrak{P}$ has a maximal extension and (2) every maximal path is rooted in the same answer i.e. $\bigcap \mathsf{RootMax}(E) \neq \varnothing$.*

We now have the resources to characterize the class of problems solved by reversal optimal methods.

Definition 47. *Problem $\mathfrak{P}$ is* ***eventually tree-like*** *iff for every $w \in W$ there is $E \in \mathcal{O}(w)$ such that for all $F \in \mathcal{O}(w)$, $\mathfrak{P}\big|_{E \cap F}$ is tree-like and rooted in $\mathcal{Q}\big|_{E \cap F}(w)$.*

Eventually tree-like problems are locally closed.

Proposition 33. *If $\mathfrak{P} = (W, \mathcal{O}, \mathcal{Q})$ is eventually tree-like, it is locally closed.*

Proof. If $\mathfrak{P}$ is eventually tree-like, then for every $w \in W$ there is $E \in \mathcal{O}(w)$ such that $\mathfrak{P}\big|_E$ is tree-like and rooted in $\mathcal{Q}\big|_E(w)$. If $\mathfrak{P}\big|_E$ is rooted in $\mathcal{Q}\big|_E(w)$, then $\mathcal{Q}\big|_E(w) \cap \overline{\mathcal{Q}\big|_E(w)} = \mathcal{Q}\big|_E(w)$ is closed. So $\mathcal{Q}(w)$ is locally closed by Proposition 3. □

Proposition 34. *Problem $\mathfrak{P}$ is solved by some reversal optimal method iff $\mathfrak{P}$ is eventually tree-like.*

Proof. First we show that every eventually tree-like problem has a reversal-optimal solution. Let

$$\lambda_{\mathsf{root}}(E) = \begin{cases} \bigcap \mathsf{RootMax}(E) & \text{if } \mathfrak{P}\big|_E \text{ is tree-like,} \\ E & \text{otherwise.} \end{cases}$$

To see that λ_{root} is a learner, Let $w \in W$. Since $\mathfrak{P}$ is eventually tree-like, there is $E \in \mathcal{O}(w)$ such that for all $F \in \mathcal{O}(w)$, $\mathfrak{P}\big|_{E \cap F}$ is tree-like and rooted in $\mathcal{Q}(w)$. So $E \in \mathsf{Lock}(\lambda_{\mathsf{root}}, w)$ as required.

We now show that λ_{root} is reversal-optimal. By Lemma 29 and the transitivity of $\leq$, it suffices to show that every singleton λ_{root}-reversal sequence is forcible. We proceed by induction on the length of singleton reversal sequences. For the base case, let $\delta = (\lambda_{\mathsf{root}}(E_1), \lambda_{\mathsf{root}}(E_2))$ be a singleton reversal sequence. So $\mathcal{Q}(\lambda_{\mathsf{root}}(E_1)) = \mathcal{Q}(\bigcap \mathsf{RootMax}(E_1)) := Q_1$ and $\mathcal{Q}(\lambda_{\mathsf{root}}(E_2)) = \mathcal{Q}(\bigcap \mathsf{RootMax}(E_2)) := Q_2$. Since $\mathfrak{P}\big|_{E_1}$ is tree-like, $\overline{Q_2 \cap E_1}$ has a maximal extension rooted in $Q_1 \cap E_1$. So there is some path $P = Q_1 \cap \overline{E_1 \cap \ldots \cap \overline{Q_2 \cap E_1}}$. Therefore $Q_1 \cap \overline{Q_2} \neq \varnothing$, so δ is forcible. Now suppose $\delta = (\lambda_{\mathsf{root}}(E_i))_{i=1}^n$ is a singleton reversal sequence. By the inductive hypothesis $(\lambda_{\mathsf{root}}(E_i))_{i=2}^n$ is forcible. Since $E_1 \supset E_2$, $\mathcal{Q}(\lambda_{\mathsf{root}}(E_2)) \cap \overline{\mathcal{Q}(\lambda_{\mathsf{root}}(E_3)) \cap \ldots \cap \overline{\overline{\mathcal{Q}(\lambda(E_n)}}} \in \Pi\big|_{E_1}$. Since $\mathfrak{P}\big|_{E_1}$ is tree-like, that forcible path is extended by one rooted in $\mathcal{Q}(\lambda_{\mathsf{root}}(E_1))$. Therefore $(\lambda_{\mathsf{root}}(E_i))_{i=1}^n$ is forcible, as required.

Now we show that if $\mathfrak{P}$ is solved by a reversal optimal method, it is eventually tree-like. We will need the following lemma about locally closed problems.

Lemma 5. *If $\mathfrak{P}$ is locally closed and λ is a reversal optimal method for $\mathfrak{P}$, then $\mathcal{Q}(\lambda(F))$ is closed in F, for $F \in \mathsf{Lock}(\lambda, w)$.*

Proof. Suppose $F \in \mathsf{Lock}(\lambda, w)$, but $\mathcal{Q}(\lambda(F))$ is not closed in F. Let $v \in F \cap \overline{\mathcal{Q}(\lambda(F))} \setminus \mathcal{Q}(\lambda(F))$ and $G \in \mathsf{Lock}(\lambda, v)$. Finally, since $v \in \overline{\mathcal{Q}(\lambda(F))}$, there is $x \in F \cap G \cap \mathcal{Q}(\lambda(F))$. Let $H \in \mathsf{Lock}(\lambda, x)$. Then $(\lambda(F), \lambda(F \cap G), \lambda(F \cap G \cap H))$ is a singleton cycle sequence where $\mathcal{Q}(\lambda(F)) = \mathcal{Q}(\lambda(F \cap G \cap H))$ and $\mathcal{Q}(\lambda(F)), \mathcal{Q}(\lambda(F \cap G))$ are locally closed. Therefore, by Lemma 4, that cycle is not forcible, and λ is not reversal optimal. $\square$

To conclude, we show that if λ is a reversal optimal method for $\mathfrak{P}$, then for every $w \in W$ and $E_1 \in \mathsf{Lock}(\lambda, w)$, $\mathfrak{P}\big|_{E_1}$ is tree-like. Let $Q_1 = \mathcal{Q}(\lambda(E_1))$. By Lemma 5, Q_1 is closed in E_1. If $\mathfrak{P}\big|_{E_1}$ is not tree-like, then either (1) there is some forcible path $P \in \Pi\big|_{E_1}$ that has no maximal extension or (2) there is some maximal path not rooted in Q_1. Suppose (1) is the case and $P = Q_2 \cap \overline{Q_3 \cap \ldots \cap \overline{Q_n}}$. We have that $Q_1 \cap \overline{Q_2 \cap \overline{Q_3 \cap \ldots \cap \overline{Q_n}}} = \varnothing$, since otherwise that would be a maximal extension for P. If (2) is the case and P is maximal, then $\overline{Q_1 \cap \overline{Q_2 \cap \overline{Q_3 \cap \ldots \cap \overline{Q_n}}}} = Q_1 \cap Q_2 \cap \overline{Q_3 \cap \ldots \cap \overline{Q_n}} = \varnothing$, since $Q_1 \cap Q_2$ are disjoint. In either case there is a λ-reversal sequence $((\lambda(E_i))_{i=1}^{n}$ such that $\mathcal{Q}(\lambda(E_i)) = Q_i$ that is not forcible, since $Q_1 \cap \overline{Q_2 \cap \overline{Q_3 \cap \ldots \cap \overline{Q_n}}} = \varnothing$. To construct that sequence, follow the steps in the proof of Proposition 28. $\square$

We conclude this section by proving that the eventually tree-like problems are exactly the simplicity driven problems.

Proposition 35. *Problem $\mathfrak{P} = (W, \mathcal{O}, \mathcal{Q})$ is eventually tree-like iff $\mathfrak{P}$ is simplicity driven.*

Proof. $\Rightarrow$: If $\mathfrak{P}$ is eventually tree-like, then for every $w \in W$ there is $E \in \mathcal{O}(w)$ such that for all $F \in \mathcal{O}(w)$, $\mathfrak{P}\big|_{E\cap F}$ is tree-like and rooted in $\mathcal{Q}\big|_{E\cap F}(w)$. If $\mathfrak{P}\big|_{E\cap F}$ is rooted in $\mathcal{Q}\big|_{E\cap F}(w)$, then $\mathcal{Q}\big|_{E\cap F}(w) \cap \overline{\mathcal{Q}\big|_{E\cap F}(w)} = \mathcal{Q}\big|_{E\cap F}(w)$ is closed. Furthermore, no other answer is closed, since $\mathfrak{P}\big|_{E\cap F}$ is rooted in $\mathcal{Q}\big|_{E\cap F}(w)$. So $\mathcal{Q}(w) = \mathcal{Q}(\mathsf{Fals}(E \cap F))$ as required. $\Leftarrow$: Suppose $\mathfrak{P}$ is simplicity driven. Then for every $w \in W$, there is $E \in \mathcal{O}(w)$ such that for all $F \in \mathcal{O}(w)$, $\mathcal{Q}(w) = \mathcal{Q}(\mathsf{Fals}(E \cap F))$. Suppose there is a path $P = Q_1 \cap \overline{Q_2 \cap \overline{Q_3 \cap \ldots \cap \overline{Q_n}}}$ in $\mathfrak{P}\big|_{E\cap F}$ with no maximal extension. Then $\mathcal{Q}(w) \cap \overline{Q_1 \cap \overline{Q_2 \cap \overline{Q_3 \cap \ldots \cap \overline{Q_n}}}} = \varnothing$. Since $\mathcal{Q}_1$ is locally closed, it is closed in the open set $\left(\overset{\vee}{\mathcal{Q}_1}\right)^c$. So both $\mathcal{Q}(w)$ and $\mathcal{Q}_1$ are closed in $E \cap F \cap \left(\overset{\vee}{\mathcal{Q}_1}\right)^c$. Contradiction, since $\mathcal{Q}(w)$ can be the only closed answer, since $\mathfrak{P}$ is simplicity driven. Suppose some path $P = Q_1 \cap \overline{Q_2 \cap \overline{Q_3 \cap \ldots \cap \overline{Q_n}}}$ in $\mathfrak{P}\big|_{E\cap F}$ is not rooted in $\mathcal{Q}(w)$. Then $Q_1 \neq \mathcal{Q}(w)$ is closed in $E \cap F$. But $\mathcal{Q}(w)$ is the only closed answer in $E \cap F$. Contradiction.

$\square$

Chapter 6

Stratified Problems

6.1 Simplicity and Falsifiability, Revisited

In the preceding, we have relied on falsifiability as the operative notion of simplicity. The learning methods we have constructed wait for an answer to become falsifiable before it is entertained. Falsifiability is connected to the simplicity order by Proposition 12: a proposition is falsifiable if and only if it is downward-closed in the simplicity order. But one might hope for a tighter connection between falsifiability and the simplicity order. In this section we define a natural class of problems for which the simplicity order lines up with falsifiability in a natural way.

Definition 48. *A locally closed question $\mathcal{Q}$ is* ***stratified*** *iff for all $A, B \in \mathcal{Q}$, $A \cap \overline{B} \neq \varnothing$ implies that $A \subseteq \overline{B}$ and hence $A \preceq B$. Problem $\mathfrak{P} = (W, \mathcal{O}, \mathcal{Q})$ is stratified iff $\mathcal{Q}$ is.*[1]

The answers in a stratified problems enjoy a certain epistemic *homogeneity*: if for some $Q \in \mathcal{Q}$, $\{w\} \preceq Q$, then $\mathcal{Q}(w) \preceq Q$. In stratified problems, an answer is falsifiable if and only if it is minimal in the simplicity order among all other answers.

Definition 49. *Answer $Q \in \mathcal{Q}$ is* ***$\mathcal{Q}$-minimal*** *iff for all $Q' \in \mathcal{Q}$ distinct from Q, $Q' \not\preceq Q$.*

Proposition 36. *If $\mathfrak{P} = (W, \mathcal{O}, \mathcal{Q})$ is stratified, then $Q \in \mathcal{Q}$ is closed (falsifiable) iff Q is $\mathcal{Q}$-minimal.*

Proof. $\Rightarrow$: This direction is immediate, and is true for all problems. $\Leftarrow$: Suppose Q is not falsifiable. Then there is $w \in \overline{Q} \setminus Q$. But since $\mathcal{Q}$ is stratified, we have that $\mathcal{Q}(w) \preceq Q$. So $\mathcal{Q}$ is not $\mathcal{Q}$-minimal. $\square$

[1] The concept of a stratification of a topological space is familiar in algebraic geometry (Stratification.; Stacks Project, http://stacks.math.columbia.edu/tag/09XY), which studies the numerical stability of solutions to polynomial equations. We arrived at it independently, with the application to simplicity in mind.

To see that the preceding does not hold in general, consider the following example.

Example 6. *Let* $W = \{a, b, c\}$*,* $\mathcal{O} = \{\{a\}, \{b, c\}, \{c\}\}$*, and* $\mathcal{Q} = \{\{a, b\}, \{c\}\}$*.* $\mathfrak{P} = (W, \mathcal{O}, \mathcal{Q})$ *is locally closed but not stratified. Note that* $\{a, b\} \not\preceq \{c\}$ *although* $\{c\}$ *is not closed, since* $\{c\} \neq \overline{\{c\}} = \{b, c\}$*.*

6.2 Simplicity and Problem Restriction

In general, answers that were previously unordered in simplicity may become ordered when the problem is restricted to new information. In stratified problems, the simplicity order remains rigid under restriction.

Proposition 37. *If* $\mathfrak{P} = (W, \mathcal{O}, \mathcal{Q})$ *is stratified, then for* $w, v \in W$ *and* $E \in \mathcal{O}$*,* $\mathcal{Q}(w) \preceq \mathcal{Q}(v)$ *iff* $Q\big|_E(w) \preceq Q\big|_E(v)$*.*

Proof. $\Rightarrow$: This direction is immediate, and is true for all problems. $\Leftarrow$: If $\mathcal{Q}(w) \cap E \preceq \mathcal{Q}(v) \cap E$, then $\mathcal{Q}(w) \preceq \mathcal{Q}(v)$ by stratification, as required. □

As a result, we have the following Corollary.

Corollary 4. *If* $\mathfrak{P} = (W, \mathcal{O}, \mathcal{Q})$ *is stratified, then for* $E \in \mathcal{O}$*, the restricted problem* $\mathfrak{P}\big|_E$ *is stratified.*

6.3 Ockham's Razors

In stratified problems, the vertical and horizontal razors can be satisfied in an especially natural way. Any conjecture that disjoins a (finite) set of minimal answers is vertical Ockham. Furthermore, a conjecture is *co-initial* in the simplicity order over the answers iff it is horizontal Ockham.

Proposition 38. *If* $\mathfrak{P}$ *is stratified and for all* $E \in \mathcal{O}$*,* $\widehat{\mathcal{Q}}(\lambda(E))$ *is finite and each* $Q \in \widehat{\mathcal{Q}}(\lambda(E))$ *is* $\mathcal{Q}$*-minimal, then* λ *is* $\mathcal{Q}$*-vertical Ockham.*

Proof. Follows from Proposition 36 and the fact the finite unions of closed propositions are closed. □

Definition 50. *Proposition* P *is* ***$\mathcal{Q}$-co-initial*** *in the simplicity order on* $\mathfrak{P} = (W, \mathcal{O}, \mathcal{Q})$ *iff for all* $Q \in \mathcal{Q}$ *there is* $Q' \in \mathcal{Q}(P)$ *such that* $Q' \preceq Q$*.*

Proposition 39. *If* $\mathfrak{P} = (W, \mathcal{O}, \mathcal{Q})$ *is stratified, then* λ *is* $\mathcal{Q}$*-horizontal Ockham iff* $\lambda(E)$ *is* $\mathcal{Q}$*-co-initial in the simplicity order on* $\mathfrak{P}\big|_E$ *for all* $E \in \mathcal{O}$*.*

Proof. Immediate from the definitions. □

6.4 Simplicity and Reversals

The forcible paths in stratified problems are exactly the *chains* in the simplicity order over answers. To see that this is not true in general, consider Example 6. There we have that $\{a, b\} \cap \overline{\{c\}} \neq \varnothing$ but $\{a, b\} \npreceq \{c\}$.

Proposition 40. *If $\mathfrak{P} = (W, \mathcal{O}, \mathcal{Q})$ is stratified, then $P = Q_1 \cap \overline{Q_2 \cap \overline{Q_3 \cap \overline{\ldots \cap \overline{Q_n}}}} \neq \varnothing$ iff $Q_1 \preceq Q_2 \preceq \ldots \preceq Q_n$.*

Proof. $\Rightarrow$: By induction on the length of paths. For the base case, if $Q_1 \cap \overline{Q_2} \neq \varnothing$, then $Q_1 \preceq Q_2$ as required. By the inductive hypothesis if $Q_2 \cap \overline{Q_3 \cap \overline{\ldots \cap \overline{Q_n}}} \neq \varnothing$, then $Q_2 \preceq Q_3 \preceq \ldots \preceq Q_n$. If $Q_1 \cap \overline{Q_2 \cap \overline{Q_3 \cap \overline{\ldots \cap \overline{Q_n}}}} \neq \varnothing$, then $Q_1 \cap \overline{Q_2} \neq \varnothing$ so $Q_1 \preceq Q_2 \preceq \ldots \preceq Q_n$ as required. $\Leftarrow$: Immediate.

□

Each answer in a stratified problem can be represented canonically in terms of the simplicity order.

Definition 51. *For $\mathfrak{P} = (W, \mathcal{O}, \mathcal{Q})$ and $Q \in \mathcal{Q}$, define $Q_{\preceq} = \bigcup\{Q' : Q' \preceq Q\}$ and $Q_{\prec} = \bigcup\{Q' \in \mathcal{Q} : Q' \prec Q\}$. Define $Q_{\succeq}$ and $Q_{\succ}$ similarly.*

Proposition 41. *If $\mathfrak{P} = (W, \mathcal{O}, \mathcal{Q})$ is stratified, then each $Q \in \mathcal{Q}$ can be written canonically as $Q = Q_{\preceq} \setminus Q_{\prec}$.*

Proof. Suppose $\mathfrak{P} = (W, \mathcal{O}, \mathcal{Q})$ is stratified and $Q \in \mathcal{Q}$. Then $\overline{Q} = \bigcup\{w : \{w\} \subseteq \overline{Q}\} = \bigcup\{Q'(w) : \{w\} \subseteq \overline{Q}\} = Q_{\preceq}$, by stratification. Furthermore, by Proposition 14, $\overset{\vee}{Q} = Q_{\prec}$. Since every proposition is the difference of its closure and its frontier, $Q = \overline{Q} \setminus \overset{\vee}{Q} = Q_{\preceq} \setminus Q_{\prec}$. □

We can now characterize the simplicity driven stratified problems in a natural way.

Proposition 42. *If $\mathfrak{P} = (W, \mathcal{O}, \mathcal{Q})$ is stratified, then $\mathfrak{P}$ is simplicity driven iff for each $Q \in \mathcal{Q}$, $Q = Q_{\succeq} \setminus Q_{\succ}$ with $Q_{\succeq}, Q_{\succ}$ open.*

Proof. $\Rightarrow$: If $\mathfrak{P} = (W, \mathcal{O}, \mathcal{Q})$ is stratified and simplicity driven, then for each $Q \in \mathcal{Q}$ and $w \in Q$ there is $O_w \in \mathcal{O}(w)$ such that Q is uniquely falsifiable in O_w. By Proposition 36, Q is also uniquely minimal in O_w. Then we have that $Q_{\succeq} = \bigcup_{w \in Q_{\succeq}} O_w$ is open. Likewise, $Q_{\succ} = \bigcup_{w \in Q_{\succ}} O_w$ is open as well and $Q = Q_{\succeq} \setminus Q_{\succ}$ as required. $\Leftarrow$: Let $\mathfrak{P} = (W, \mathcal{O}, \mathcal{Q})$ be stratified, $w \in Q$ and $Q_{\preceq}$ open. Then there is $O \in \mathcal{O}(w)$ such that $O \subseteq Q_{\succeq}$. By Proposition 36 and Corollary 4, Q is uniquely falsifiable in O and in all $O \cap O'$ for $O' \in \mathcal{O}(w)$. □

So the stratified problems solved by reversal optimal methods are exactly those that have open up-sets in the simplicity order over answers.

6.5 Stratification and Aptness of Representation

The preceding results demonstrate that the simplicity order is a particularly apt epistemic description of a stratified problem. These problems have such natural properties that one hopes that every solvable problem is refined by one. We have not been able to prove or disprove the following conjecture.

Conjecture 1. *If $\mathfrak{P} = (W, \mathcal{O}, \mathcal{Q})$ is solvable in the limit, then there is a stratified question Q' refining Q.*

However, we can show that many natural problems *are* stratified.

Example 7. *If $|W| \leq \omega$ and $\mathfrak{P} = (W, \mathcal{O}, \mathcal{Q})$ is solvable in the limit, then the discrete partition $\mathcal{Q}_\perp$ is a stratified question refining $\mathcal{Q}$. By Corollary 2, $(W, \mathcal{O}^*)$ is a T_D space, so every singleton is locally closed. Since the discrete partition is countable, it is a stratified question.*

Example 8. *Let W_{cts} and O_{rec} be defined as in Example 3. Then question $\mathcal{Q}_{\mathsf{deg}} = \{D_n : n \in \mathbb{N}\}$ asks what the degree of the true law is and $\mathcal{Q}_{\mathsf{form}} = \{F_S : S$ is a finite subset of $\mathbb{N}\}$ asks what its form is.*

Proposition 43. *The questions $\mathcal{Q}_{\mathsf{deg}}, \mathcal{Q}_{\mathsf{form}}$ are stratified*

Proof. By Lemmas 1 and 2, since $F_{<S}, F_{\leq S}$ are closed in $\mathfrak{I}_{\mathsf{L}^2}$, they are closed in $\mathfrak{I}_{\mathsf{cts}}$ and therefore closed in the restriction $\mathfrak{I}_{\mathsf{poly}}$. Therefore $F_S = F_{\not< S} \setminus F_{\not\leq S}$ is locally closed in $\mathfrak{I}_{\mathsf{poly}}$. Since $D_n = F_{\not\leq \{1,\ldots,n-1\}} \cap F_{\leq \{1,\ldots,n\}}$ it is locally closed by Lemmas 1 and 2. Stratification of both questions follows from Proposition 9. □

Example 9. *Consider a situation in which one receives discrete inputs through time. Let I denote the countable set of all possible such inputs, and let W be some collection of infinite sequences of inputs. Let $w \upharpoonright t$ denote $(w_0, \ldots, w_{t-1})$. The information imparted by observation of finite sequence e of inputs is just $[e]$ = the set of all $w \in W$ that extend e. Let $\mathcal{O}_{\mathsf{seq}} = \{[w \upharpoonright n] : w \in W$ and $n \in \mathbb{N}\}$. Then $\mathcal{O}_{\mathsf{seq}}$ is a topological basis. One can formulate many inductive problems by varying W and the question asked. For example, suppose that $I = \mathbb{N} \cup \{*\}$, where asterisk is a non-numeric input. Let W_{fin} contain all infinite sequences of inputs that have finite range. For finite set S of natural numbers, let $w \in C_S$ iff the set of numbers occurring in w is exactly S, and let the range question $\mathcal{Q}_{\mathsf{rng}}$ denote the set of all C_S such that S is a finite subset of $\mathbb{N}$. Let $w \in C_n$ iff $w \in C_S$ and $|S| = n$ and let the counting question $\mathcal{Q}_{\mathsf{cnt}}$ denote the set of all C_n such that $n \in \mathbb{N}$. Then we have the intuitive simplicity relations:*

Proposition 44.

$$\begin{aligned} C_n \preceq C_{n'} \quad &\textit{iff} \quad n \leq n'; \\ C_S \preceq C_{S'} \quad &\textit{iff} \quad S \subseteq S'. \end{aligned}$$

The proof is intuitive—no matter which inputs one has seen so far, one could always see a new input later. Furthermore:

Proposition 45. *The questions $\mathcal{Q}_{\mathsf{cnt}}$ and $\mathcal{Q}_{\mathsf{rng}}$ are stratified.*

Proof. To see that each answer is locally closed, note that $C_S = A_S \setminus B_S$, where A_S is the disjunction of all $[e]$ such that the numbers occurring in e are exactly those in S and B_S is the disjunction of all $[e]$ such that some number missing from S occurs in e. Stratification follows from Proposition 44. □

Stratified problems are so aptly represented by their simplicity relations that Kelly et al. (2014b,a) propose that we consider an admissible ***simplicity concept*** for a problem to be a stratified problem that refines it.[2] That approach was not taken in this work for a number of reasons. Not least among them is the fact that in general, the existence of stratifications is an unsettled question. Secondly, it is known that there is not always a unique simplicity concept that refines a problem (Kelly et al., 2014a). This invites uncomfortable questions about arbitrary impositions of simplicity relations. The approach taken in this work steps around that problem. The simplicity relation is an entirely topological feature of the problem—there is no question of arbitrary impositions. Finally, although there are distinct representational advantages to stratified problems, we have been unable to demonstrate any *methodological* advantages to recommend them. Every success concept defined in this work is equally well achieved in locally closed problems as in stratified problems. Furthermore, every solvable problem is known to be refined by a locally closed problem. For those reasons, simplicity is not defined as an order over answers in a stratified question, but as an order over the algebra of all propositions.

[2]The condition is actually a bit more subtle than refinement, but refinement captures most of what is intended.

Chapter 7

Ockham's Razor: Pragmatic or Epistemic?

We have shown that our two versions of Ockham's razor are necessary for achieving optimally monotonic convergence to the truth. If you want to avoid going circles on the way to the truth, you must respect the vertical razor. If you want to make as few course reversals as possible on the way to the truth, you must respect the horizontal razor. These results are natural developments and generalizations of Kelly's (2004; 2007; 2011).[1] Like Kelly's previous arguments, the arguments in this work may leave some unsatisfied. For example, Fitzpatrick (2013) objects that such arguments do not show *enough*.

> A natural reaction to Kelly's proposal is that retraction efficiency appears to be a purely *pragmatic* consideration and not a genuinely *epistemic* one. How can concerns about possibly having to change one's mind a greater number of times provide good reason to believe, or even tentatively accept, the claims of the theory currently selected as the simplest one consistent with the data? Kelly is explicit that his efficiency argument solely concerns the worst-case convergence properties of methods, not features of simpler theories themselves. It certainly doesnt show that simpler theories are more likely to be true. How, then, can simplicity be a genuine criterion for theory choice on Kelly's account?

The objection assumes that the only way to provide an *epistemic* justification for an inductive methodology is to demonstrate that it probably selects true theories. Call that thesis *indicativism*. We respond that the indicativist badly misconstrues both scientific inquiry and epistemology. In many empirical domains, there simply cannot be truth-indicative methods—the inductive problems are too difficult. In the causal inference

[1] Perhaps the largest difference from that previous work is the identification of the simplicity order with Popper's falsifiability order. This leads to the reduction of the vertical razor to Popper's dictum to make only falsifiable conjectures. I claim that the connections to belief revision also represent something of an advance.

literature, it is well known that unless strong assumptions are imposed on the generating model, no truth-indicative methods exist for inferring the effects of manipulations from observational data (Zhang and Spirtes, 2002). Without additional assumptions, the best you can do is to converge to the truth in the limit without unnecessary vacillations.Therefore, insistence upon that strict standard results either in metaphysical speculations sufficient to make simplicity truth-conducive, or in skeptical denial that inductive justification is possible at all. But those standard, unappealing outcomes are not necessary. Feasibility contextualism demands that we scale our normative demands to be the strongest *feasible* demands in light of the intrinsic difficulty of the problem, and provides a non-circular account of epistemic justification of inductively inferred belief.

Equally fatal is that indicativism provides no justification, in terms of truth-conduciveness, of diachronic principles of theory change. The monotonicity postulates of defeasible inference cannot be epistemic, since they do not guarantee that the inferences are probably true. The rationality postulates of belief revision theory cannot be epistemic either, since they do not guarantee that belief states are probably true. This puts the indicativist out of step with much of formal and mainstream epistemology, where truth-conducive belief dynamics are considered crucial to the analysis of knowledge. The *locus classicus* for that position is Plato's *Meno*:

> True opinions too are a fine thing and altogether good in their effects so long as they stay with one, but they won't willingly stay long and instead run away from a person's soul, so they're not worth much until one ties them down by reasoning out the explanation. And when they've been tied down, then for one thing they become items of knowledge, and for another, permanent. And that's what makes knowledge more valuable than right opinion, and the way knowledge differs from right opinion is by being tied down (Plato and Sharples, 1985, 97e-98a).

Plato's point is that it is *stable* true belief, rather than true belief *simpliciter* that constitutes knowledge.[2] Stability is a dynamical concept, and there is no sense in which stable beliefs are *probably true*. But from the perspective we have developed in this work, stability is the strongest kind of truth-conduciveness that is feasible in inductive settings. And to be justified is exactly to have achieved the best kind of truth-conduciveness possible. Thus—for the feasibility-contextualist—stable belief is *exactly* justified belief, and our esteem for stability is no more mysterious than our esteem for justification. That unified perspective can never be available to the indicativist.

[2] In contemporary epistemology, that position is best defended by Lehrer (1990) and Rott (2001, 2005).

Bibliography

Hirotugu Akaike. A new look at the statistical model identification. *Automatic Control, IEEE Transactions on*, 19(6):716–723, 1974.

Jean-Paul Allouche. Note on the Constructible Sets of a Topological Space. *Annals of the New York Academy of Sciences*, 806(1):1–10, 1996.

CE Aull and WJ Thron. Separation axioms between T_0 and T_1. *Indag. Math*, 24:26–37, 1962.

Alan Baker. Simplicity. In Edward N. Zalta, editor, *The Stanford Encyclopedia of Philosophy*. Fall 2013 edition, 2013.

Alexandru Baltag, Nina Gierasimczuk, and Sonja Smets. Epistemic Topology (to appear). *ILLC Technical Report*, 2014.

Prasanta S Bandyopadhayay, Robert J Boik, and Prasun Basu. The curve fitting problem: A Bayesian approach. *Philosophy of Science*, pages S264–S272, 1996.

Lorenzo Carlucci and John Case. On the Necessity of U-Shaped Learning. *Topics in Cognitive Science*, 5(1):56–88, 2013.

Lorenzo Carlucci, John Case, Sanjay Jain, and Frank Stephan. Non U-shaped vacillatory and team learning. In *Algorithmic Learning Theory*, pages 241–255. Springer Berlin Heidelberg, 2005.

Rudolf Carnap. On Inductive Logic. *Philosophy of Science*, 12(2):72, 1945.

John Case and Timo Kötzing. Topological separations in inductive inference. In *Algorithmic Learning Theory*, pages 128–142. Springer, 2013.

John Case and Carl Smith. Comparison of identification criteria for machine inductive inference. *Theoretical Computer Science*, 25(2):193–220, 1983.

Matthew de Brecht and Akihiro Yamamoto. Interpreting Learners as Realizers for Σ^0_2-Measurable Functions. *(Manuscript)*, 2009.

Paul AM Dirac. The relation between mathematics and physics. In *Proc. Roy. Soc. Edinburgh*, volume 59, page 122, 1939.

Heather Douglas. Inductive risk and values in science. *Philosophy of Science*, pages 559–579, 2000.

Simon Fitzpatrick. Kelly on Ockhams Razor and Truth-Finding Efficiency. *Philosophy of Science*, 80(2):298–309, 2013.

Malcolm Forster and Elliott Sober. How to tell when simpler, more unified, or less ad hoc theories will provide more accurate predictions. *The British Journal for the Philosophy of Science*, 45(1):1–35, 1994.

Greg Gandenberger. Methodism and Ockham's Razor. `http://gsganden.blogspot.com/2012/07/methodism-and-ockhams-razor.html`, 2012.

Peter Gärdenfors. *Knowledge in Flux, Bradford Book.* MIT Press, Cambridge, Mass, 1988.

E Mark Gold. Language identification in the limit. *Information and control*, 10(5):447–474, 1967.

Jeroen Groenendijk and Martin Stokhof. On the semantics of questions and the pragmatics of answers. *Varieties of Formal Semantics*, 3:143–170, 1984.

Charles L Hamblin. Questions in montague english. *Foundations of Language*, pages 41–53, 1973.

Carl G Hempel. *Aspects of Scientific Explanation.* Free Press, 1965.

G. Hempel. Valuation and Objectivity in Science (1983). Reprinted in Fetzer, J.(ed.): The Philosophy of Carl G, 2001, 1983.

David Hume. *An Enquiry Concerning Human Understanding.* Broadview Press, 2011.

Harold Jeffreys. Theory of probability, 1961.

Alexander S Kechris. *Classical Descriptive Set Theory*, volume 156. Springer-Verlag New York, 1995.

Kevin Kelly. *The Logic of Reliable Inquiry.* Oxford University Press, 1996.

Kevin Kelly and Hanti Lin. A Simple Theory of Theoretical Simplicity. *Manuscript*, 2011.

Kevin Kelly, Konstantin Genin, and Hanti Lin. Simplicity, Truth, and Topology. *(Manuscript)*, 2014a.

Kevin Kelly, Konstantin Genin, and Hanti Lin. Realism, Rhetoric, and Reliability. *(Manuscript)*, 2014b.

Kevin T Kelly. Justification as truth-finding efficiency: how Ockham's Razor works. *Minds and Machines*, 14(4):485–505, 2004.

Kevin T Kelly. Ockham's razor, empirical complexity, and truth-finding efficiency. *Theoretical Computer Science*, 383(2):270–289, 2007.

Kevin T Kelly. Simplicity, Truth and Probability. In Prasanta S Bandyopadhyay and Malcolm Forster, editors, *Handbook of the Philosophy of Science Volume 7: Philosophy of Statistics*. North Holland, 2011.

Robert Koons. Defeasible Reasoning. In Edward N. Zalta, editor, *The Stanford Encyclopedia of Philosophy*. Spring 2014 edition, 2014.

Imre Lakatos. Falsification and the methodology of scientific research programmes. *Criticism and the Growth of Knowledge*, 4:91–196, 1970.

Larry Laudan. *Science and values*, volume 87. Cambridge Univ Press, 1984.

Larry Laudan. The epistemic, the cognitive, and the social. *P. Machamer and G. Wolters, Science, Values, and Objectivity*, pages 14–23, 2004.

Hannes Leeb and Benedikt M Pötscher. Sparse estimators and the oracle property, or the return of Hodges estimator. *Journal of Econometrics*, 142(1):201–211, 2008.

Keith Lehrer. *Theory of Knowledge*. Westview Press, 1990.

Wei Luo and Oliver Schulte. Mind change efficient learning. *Information and Computation*, 204(6):989–1011, 2006.

Eric Martin, Arun Sharma, and Frank Stephan. Unifying logic, topology and learning in parametric logic. *Theoretical Computer Science*, 350(1):103–124, 2006.

Margaret Morrison. *Unifying Scientific Theories: Physical concepts and mathematical structures*. Cambridge University Press, 2007.

Wayne C Myrvold. A Bayesian account of the virtue of unification. *Philosophy of Science*, 70(2):399–423, 2003.

Daniel N Osherson, Michael Stob, and Scott Weinstein. *Systems that learn: An introduction to learning theory for cognitive and computer scientists.* The MIT Press, 1986.

Plato and RW Sharples. *Plato: Meno.* Aris and Phillips, 1985.

Karl R. Popper. *The Logic of Scientific Discovery.* London: Hutchinson, 1959.

Hilary Putnam. Trial and Error Predicates and the Solution to a Problem of Mostowski. *Journal of Symbolic logic*, pages 49–57, 1965.

Hans Rott. Two dogmas of belief revision. *The Journal of Philosophy*, pages 503–522, 2000.

Hans Rott. *Change, Choice and Inference: A study of belief revision and nonmonotonic reasoning.* Oxford University Press, 2001.

Hans Rott. Stability, strength and sensitivity: converting belief into knowledge. In *Contextualisms in Epistemology*, pages 327–351. Springer, 2005.

Oliver Schulte, Wei Luo, and Russell Greiner. Mind Change Optimal Learning of Bayes Net Structure. In *20th Annual Conference on Learning Theory (COLT), San Diego, CA, June 12-15*, 2007.

Gerhard Schurz. Abductive belief revision in science. In *Belief Revision meets Philosophy of Science*, pages 77–104. Springer, 2011.

Arun Sharma, Frank Stephan, and Yuri Ventsov. Generalized notions of mind change complexity. In *Proceedings of the tenth annual conference on Computational learning theory*, pages 96–108. ACM, 1997.

Stacks Project. Partitions and Stratifications, http://stacks.math.columbia.edu/tag/09XY.

Stratification. *The Encyclopedia of Mathematics.* http://www.encyclopediaofmath.org/index.php?title=

Alan Mathison Turing. On computable numbers, with an application to the Entscheidungsproblem. *J. of Math*, 58:345–363, 1936.

Bas C Van Fraassen. *The Scientific Image.* Oxford University Press, 1980.

Vladimir Vapnik. *Statistical learning theory*, volume 2. Wiley New York, 1998.

Steven Vickers. *Topology Via Logic.* Cambridge University Press, 1996.

Larry Wasserman. Bayesian model selection and model averaging. *Journal of Mathematical Psychology*, 44(1):92–107, 2000.

Akihiro Yamamoto and Matthew de Brecht. Topological properties of concept spaces (full version). *Information and Computation*, 208(4):327–340, 2010.

Jiji Zhang and Peter Spirtes. Strong faithfulness and uniform consistency in causal inference. In *Proceedings of the nineteenth conference on uncertainty in artificial intelligence*, pages 632–639. Morgan Kaufmann Publishers Inc., 2002.