Scientific Discovery from the Point of View of Acceptance[1]

Eric Martin
University of New South Wales

Daniel Osherson
Princeton University

May 1, 2003

[1] This work was supported by Australian Research Council Grant #A49803051 to Martin and by NSF Grant #IIS-9978135 to Osherson. We thank Scott Weinstein for generously checking proofs and offering improvements to several arguments. Thanks also to Sebastian Rahtz for developing the LaTeX package `hyperref.sty`, used to produce the current document. Authors' addresses: emartin@cse.unsw.edu.au, osherson@princeton.edu.

In the four papers available on our web site (of which this is the first), we propose to develop an inductive logic. By "inductive logic" we mean a set of principles that distinguish between successful and unsuccessful strategies for scientific inquiry. Our logic will have a technical character, since it is built from the concepts and terminology of (elementary) model theory. The reader may therefore wish to know something about the kind of results on offer before investing time in definitions and notation. Providing such an informal overview is the purpose of the present essay. We begin with discussion of the central concept under investigation, namely, *theory acceptance*.

1. Acceptance

In this section we aim to motivate a distinction between two epistemological states, which will be termed "belief" and "acceptance." As a preliminary, it might be helpful to consider a parallel distinction. Consider the difference between desiring that your business rival suffer a car accident versus consenting or acquiescing to this mishap. Suppose that you are like most of us, and have not yet succeeded in ridding yourself of every vile thought. When this one floats into your mind you may say to yourself: "Aha! A wicked desire. Such things do occur to me from time to time despite efforts to improve my character. But not to worry. I don't intend to act on the idea." In other words, you can take note of your dark desire for an accident without experiencing any *new* guilt. For, you seem not to be guilty of anything more than the same old failure to rebuild your character so as to avoid nasty ideas. In contrast, you are genuinely guilty of a new moral lapse (and should be ashamed of yourself) if you somehow consent to the accident (e.g., by nodding to someone who claims they can arrange it), or if you acquiesce to it (e.g., by failing to signal a wobbly wheel on your rival's vehicle). Unlike desire, consent and acquiescence are willful acts that have moral consequences beyond any previous efforts to reform one's character. Willful acts involve choice, and each one is ethically valent (whether or not the choice ends up having its intended consequences). We hope that you see the matter our way, and feel the moral distinction being drawn.

Various authors have noticed an epistemological distinction that parallels the one between desire and consent (as in [4, Ch. II]). Like desire, *belief* seems to be largely involuntary at any given moment. Although you can influence which beliefs tend to enter your mind over time (e.g., by choosing to read certain newspapers rather than others), you can't alter your current beliefs by a mere act of will. (Try to believe that you are eight feet tall.) For this reason, beliefs play a different social role from what can be termed *acceptance*, which does involve a decision. For example, suppose you buy a lottery ticket. Once the purchase is made, you can't prevent yourself from doubting that the ticket will be a winner, that is, from believing it is a loser. But you haven't accepted the thesis that the ticket is a loser (if you did, you would throw the ticket away). Indeed, for every ticket sold you doubt that it is a winner, but it's not true that for every ticket sold you accept that it is a loser. To see this, note that your doubts about each individual ticket don't commit you to questioning the fairness of the draw. But if for each ticket you accepted the thesis that it is a loser, then just such questions ought to come to mind.[1]

Another way to perceive the distinction between belief and acceptance is to reflect on their consequences. Accepting a hypothesis entails willingness to adopt it as a basis for action. Thus, if you accepted the hypothesis that the ticket will lose, you would be inclined to destroy it (prior to the draw). Your mere belief that it will lose does not, by itself, so incline you. The character of the action connected to acceptance varies with the social context. Because of scant admissible evidence, for example, a judge may be led to officially accept a defendant's innocence but strongly believe the contrary. In this case, acceptance entails the intention to make certain pronouncements in court, sign certain documents, and so forth. In the case of scientific inquiry, accepting a theory would seem to involve commitment to some combination of (a) expounding the theory (however cautiously) in

[1] This kind of example is discussed more thoroughly in [23, §6.2.1].

public, (b) exploring its consequences in view of further test, (c) being biased to retain large fragments of the theory when it is contradicted by data, (d) trying to base technology on the theory, and (e) attempting to use the theory to set background parameters when developing models of new phenomena. Of course, belief is also an important factor in theory acceptance. (Often we refuse to accept a theory we strongly believe is false.) But there may well be other factors, such as the scientist's impression that one theory is more interesting than another, as well as her suspicion that postulating a certain theory will promote inquiry leading to a better one. Indeed, in some circumstances, all the theories we can think of strike us pitifully implausible. Yet we may go on to choose one of the theories as the most interesting to explore, and defend it as the best currently available.[2]

Acceptance seems to be categorical inasmuch as it rests on a definite choice among alternative theories. Thus, we accept a theory or not, even though our confidence that the theory is true might be a graded affair. There need be no dogmatism in acceptance, however. New data (or new reflection) can undermine our decision to accept a theory, and push us towards another.[3]

Perhaps this brief discussion is enough to persuade you that there is a meaningful distinction between belief and acceptance, and that the latter may be important to understanding empirical inquiry. We have not offered an *analysis* of acceptance, that is, a full account of the character of this epistemological state. There is considerable room for disagreement about such an analysis; rival accounts are available in [4, 23, 22, 19]. Let us not rehearse this valuable material, nor take sides on contentious points. We wish simply to assume that all parties to the discussion are referring to the same, intuitively felt distinction, and turn

[2]Thus, the astronomer Simon Lilly says of current cosmological theory: "We have these simpleminded ideas and won't be surprised if it's all different than we think" (reported in *Science*, 7 June 1996, vol. 272, p. 1436).

[3]In placing the accent on categorical acceptance, we don't intend to rule out the logical possibility of *graded acceptance* that is somehow different from graded credibility. Our theory is built, however, on the more familiar idea of "full" acceptance.

immediately to the task of building an inductive logic for theory acceptance. In particular, we adopt a picture in which acceptance is a recognizable mental act that applies to just one theory (if any) at a time. Within this perspective, scientific success consists not in gradually increasing one's confidence in the true theory, but rather in ultimately accepting it, and then holding on to it in the face of new data. Our stance on empirical inquiry is similar to well known views expressed by Popper [25]. But Popper might have been sceptical about the further project (to be pursued here) of constructing an inductive logic for inquiry.[4]

All this is very different from focussing on the *credibility* (belief) that scientists attach to alternative theories, and on the evolution of these credibilities under the impact of data. Interpreting credibility as probability leads to the Bayesian analysis of inquiry, which has greatly illuminated diverse aspects of scientific practice.[5] Bayesians sometimes find the concept of acceptance to be epistemologically inert since they think that distributing credibilities among theories is enough to represent scientists' attachment to them (see [18, Ch. 2]). Other scholars in the Bayesian tradition attempt to make sense of acceptance in terms of probabilities and utilities (e.g., [23, Ch. 6]).

Rather than embracing a Bayesian stance on acceptance, we prefer to investigate the logic of acceptance without making additional assumptions about background probabilities or utilities. One reason for this attitude is that Bayesianism faces unresolved challenges involving the impact of evidence on belief. The problem of "old evidence," for example, is still much debated [11, 9, 8, 7, 19, 15, 16].[6] A more fundamental motive for our work is that scientists sometimes can't or won't specify the probability and utility functions needed to get the Bayesian analysis off the ground (e.g., priors and likelihoods). Even when it

[4] For insightful discussion of the relation of Popperianism to the kind of project carried out here, see [20, 21].

[5] For illustrations, see [6, 15, 16, 24, 27]. A classic example of Bayesian reconstruction of scientific history is provided in [5].

[6] Here is the problem: once an evidence-statement is verified, its truth becomes nearly certain, so conditioning the probability of theoretical statements on this basis should have marginal impact.

Home Page

Title Page

Page 6 of 38

Go Back

Full Screen

Close

Quit

seems possible to reconstruct the needed functions after the fact (on behalf of the scientist), such cases still justify an analysis of inquiry in the simpler terms of theory proposal and modification. Such is the mission of the four essays available on our site.

So let us affirm that for purposes of the present study we accept the acceptance picture of science, and will attempt to work out some of its consequences. The consequences will take the form of principles that describe the prospects for scientific success in various circumstances by various kinds of scientists. Of course, both "problems" and "scientists" will be idealized models of the real thing, built from bits of set-theory and logic. Once we have formally defined a class of inductive problems and a kind of scientist, we ask whether scientists of the given kind can solve the problems in the class.

What we mean by "solving" a problem will become clearer in Section 3, below, where we present an informal overview of the theory. You might wish to proceed immediately to that section. Readers unfamiliar with the field, however, may benefit from discussion of a simpler paradigm of inductive inference that shares some fundamental concepts with the more general framework to be developed later. The simple paradigm is presented next. It is drawn from research that began last century with the seminal papers [29, 12, 26, 3].[7]

2. A numerical paradigm

2.1. The numbers game

Our simple inductive paradigm can best be explained by playing a game. It will involve subsets of *natural numbers*. Recall that a natural number is a nonnegative integer, that is, one of the numbers $0, 1, 2, \cdots$. We'll play with the following sets:

[7] For a review of the voluminous literature that ensued, see [17].

(1) all the natural numbers except for 0 (call this set S_0);
all the natural numbers except for 1 (call this set S_1);
all the natural numbers except for 2 (call this set S_2);
and so forth.

In general, for every natural number n, the set S_n (which contains every natural number except n) appears in our collection (1).

Here's how we'll play. The authors will huddle together and choose one subset from (1). Then we'll give you a clue about the set we chose, and you have to guess it. After the first clue, we'll give you another clue (still about the same set chosen at the outset), and you get to guess again. We'll go on like this to a third clue and third guess, etc. Got it?

O.K. We've chosen our subset from (1). Here is your first clue. *The set we chose includes the number 2.* Go ahead and guess. What subset did we choose?

Excellent guess! Now here is your second clue. *The set we chose at the outset includes the number* 0. Guess again. (You're allowed to make the same guess as before if you like.)

Here is your third clue. *The set we chose at the outset includes the number* 3. Guess.

To move the game along, let us provide the next ten clues all at once; imagine that you received them one by one. *The set we chose at the outset includes the numbers* $4, 5, 6, 7, 8, 9, 10, 11, 12, 13$. Now tell us: Has your guess begun to stabilize to S_1, that is, to the set containing all the natural numbers except for 1? Good. Here is your next clue: *The set we chose at the outset includes the number* 1. So guess again!

We neglected to mention that the game goes on forever, so let us break it off at this point and tell you more about the rules. A given clue consists of a declaration that a single number is in the chosen set; clues about what is outside the set are illegal. As the game unfolds, we are obliged to make a declaration about every number in our set. In other words, for every number x in our set, there has to be a clue that declares x to be in it. There is no constraint on when we get around to announcing the number x; it must simply appear among the clues sooner or later. Recall that we waited until the fourteenth clue before telling you that 1 belongs to our set.[8] We can repeat old clues if we choose. Regarding the rules governing your behavior, at most one guess may be made after each clue. You are allowed to refrain from guessing altogether, or even to stupidly guess a set outside the collection (1). But no funny business, like guessing "the set the authors happened to choose." It must be unambiguous to a third party which numbers are in your set, and which are out.

2.2. Winning the numbers game

You win the game just in case the following is true.

(2) Sooner or later you announce the correct set (that is, the one we chose) and never deviate from that guess thereafter.

For example, if we had chosen the set S_{20}, then you win the game just in case there is a clue after which you announce S_{20} and then announce S_{20} after every new clue, forever. Recall that a *cofinite* set of natural numbers is a subset that includes all but finitely many of them (e.g., the set missing just 4, 8, 10 and 25). So, the criterion (2) can be stated this way: you win the game just in case you conjecture the set we chose cofinitely many times. Let's say that you *make a mistake* on a given clue if you fail to make a conjecture, or if your conjecture is wrong (not our chosen set). Then another way to state the criterion is

[8] There is an elementary confusion to be avoided here. Stipulating that every member of the set must eventually be declared is not the same thing as stipulating that eventually we must have declared every member of the set. It is possible to abide by the former stipulation but not the latter (there are too many numbers).

as follows. You win the game just in case you make only finitely many mistakes. There is one more paraphrase of (2) worth pursuing. Let us say that you *converge* to a subset S of natural numbers just in case you announce the set S after one of the clues and forever afterwards. Then you win the game just in case you converge to the set we chose at the outset. Otherwise, you lose (there are no ties).

In fact, the set we chose was S_{14}, all natural numbers except 14. Would you have won the game if we kept playing? Actually, a more poignant question is the following. Would you have won the game no matter which set we selected from (1)? That is, would you have converged to any possible initial choice of set? We think that you *would* have won the game in all circumstances, most likely through some variant of the following reasoning.

(3) YOUR REASONING: "One number is missing from the chosen set, but I'm only shown which numbers are *in* the set. This means that after each clue there are always infinitely many numbers that I've not been shown, any one of which could be the magic missing number. I can nonetheless persistently guess that some specific number is absent, and be confident that if I'm wrong, I'll eventually be contradicted. So, I'll guess that 0 is absent until I see 0 among the clues. If I see 0, then I'll guess that 1 is absent until I see 1, among the clues, and so forth. In other words, if clues show me that the numbers $x_1, x_2, x_3 \cdots x_n$ are present in the chosen set, I'll guess that the set is S_y, where y is the first natural number that is not one of $x_1, x_2, x_3 \cdots x_n$. No matter what the missing number is, I'll eventually start conjecturing the correct set, namely, once I've seen all the numbers smaller than the missing one. And once I start conjecturing the correct set, I'll never stop because I'll never be shown the missing number!"

Anyone who follows strategy (3) will win the game in all cases, no matter what set we choose at the outset. Indeed, use of the strategy is successful not only regardless of the choice of initial set, but also regardless of how the clues are delivered. No matter how we string out the members of our chosen set (which was S_{14}, recall), we will never show you the number 14, and we will eventually show you all of $0, 1, 2 \cdots 13$. You will at that point seize upon the hypothesis S_{14}, from which the clues will never dislodge you thereafter. You win no matter what we do.

Winning in such an unqualified sense is a central concept in our theory, so it deserves a separate name. Let us say that you *solve* a game just in case you are using a strategy that would cause you to win the game no matter what choice is made of initial set, and no matter how the clues are delivered to you. This definition applies to any game, not just the one based on the collection (1). A game is determined by spelling out the subsets of natural numbers from which the initial choice can be made. For example, a new game is given by the collection of all finite subsets of natural numbers (except the empty set). Can you figure out a strategy that solves the finite-subset game?

Yes, of course. At each stage of the game, you need merely guess the (finite) collection of numbers you've seen so far. If the chosen set was, say, $3, 5, 8$ then your conjectures will be correct as soon as all three numbers show up in the data, and then stay correct forever after.

If you operate according to a strategy that solves the game you are playing then each of your conjectures might considered to be "justified." At least, you could defend any one of them by saying that it was selected according to a strategy that guarantees success. Perhaps such a justification yields some peace of mind. It does not, however, warrant confidence that any particular conjecture is true. In the game consisting of all finite subsets, for example, you are never in a position to affirm that no new number will show up in a future clue. This will be a fundamental feature of our inductive logic. Except for degenerate cases, it will not be possible to infer the accuracy of any given conjecture. (Otherwise, our logic would be more deductive than inductive.) Nor will there be any basis for attaching a degree-of-belief to conjectures. (Otherwise, we would be developing a probabilistic theory, instead of one based on acceptance.) Justification will only accrue to entire strategies.

2.3. Scientists

The strategy (3), which solves the game (1), determines a function. To see what function it is, think of all the situations you might be faced with in the course of a game. You might have seen the clues $5, 3, 109, 21$ in that order, or maybe in the order $3, 5, 21, 109$. Or you might have seen the clues $9, 1, 9, 1, 9, 9, 9$. Any such sequence of clues can be called a *data state*, or *evidential position*, or *information situation*. Let us just call them *clue-sequences*. The strategy (3) operates on any clue-sequence, and maps it to a set of natural numbers, namely, the set S_y, where y is the least natural number that does not appear in the clue-sequence. So (3) defines a particular function with *domain* the set of all clue-sequences, and *range* the collection of all subsets of natural numbers. We can say that this function represents strategy (3) *in extension*, or conversely that strategy (3) provides an *intension* for the function.

When we contemplate intensions, we are close to the *process* of reasoning whereas extensions represent no more than the *results* of reasoning. Our inductive logic will have much to say about intensions, but there is also something to learn from studying just extensions. So we will consider any function from the set of clue-sequences to the collection of subsets of natural numbers to be a *scientist*. Thus, for now, we won't distinguish two strategies that have the same function as extension. They will be considered to be the same scientist. And we'll say that scientists (considered as functions) solve or fail to solve a given game.

Of course, use of the term "scientist" tips our hand. We would like you to feel that our games resemble scientific inquiry. Thus, a subset of numbers might be codes for a more concrete reality. Each natural number, for example, might represent a distinctive kind of genome, and a set of numbers might code the class of biological possiblities. The goal of the game is then to figure out which class of genomes includes all those that could yield viable organisms on Earth.

In such a perspective, the starting collection of sets is like the range of possibilities from which Nature will choose reality. (The authors were playing Nature's role in the game we played at the outset.) The range of possibilities need not embrace every subset of natural numbers. Instead, Nature's choice will typically be constrained by what we've already discovered about the world, prior to the game. For this reason, we are confident that Nature will choose, say, from just the subsets in (1). The selected set is Nature's primeval choice of reality. The clues are data that she makes available to us. Winning the game is like discovering which potential reality is the "actual one." Solving a game is like implementing a successful scientific strategy. Scientists are embodiments of strategies.

Have we convinced you that our games are a little like science? If not, don't give up yet. The present section is just a preliminary to Section 3 and beyond, which present an overview of the more sophisticated theory that will occupy the sequel.

2.4. Unsolvability

The kind of games we've been discussing are interesting because they are not always solvable. For example, suppose we add the set $0, 1, 2, \cdots$ of all natural numbers to our first game. The game defined by (1) is thereby transformed into:

(4) all the natural numbers (call this set N);
all the natural numbers except for 0 (call this set S_0);
all the natural numbers except for 1 (call this set S_1);
all the natural numbers except for 2 (call this set S_2);
and so forth.

You can't rely on strategy (3) to win the expanded game because Nature might have chosen the set N (all natural numbers). In this case, your hypotheses will keep changing forever, since every time you conjecture that a given number is missing it shows up later on. Hence, strategy (3) does not solve the game (4) inasmuch as it fails to win for some choice of initial set (namely, N). In fact, a more general fact can be demonstrated.

(5) THEOREM: No scientist solves (4). In other words, game (4) is *unsolvable*.

It is important to appreciate the generality of this theorem. No matter how you try to reason about the game (4), your reasoning will determine some function from clue-sequences to subsets of natural numbers. The function need not be *total*; for certain clue-sequences it need not yield any conjecture at all (perhaps because your reasoning is incomplete in some cases).[9] Total or not, however, the function determined by your reasoning counts as a scientist, and *no scientist* solves (4), so you can't either.

Another example of the unsolvability phenomenon was discovered by E. Mark Gold [12] in a remarkable paper that launched much of the ensuing literature on acceptance based discovery.

(6) THEOREM: Let the game $\mathcal{G}$ consist of every finite set of natural numbers along with any infinite set of natural numbers. Then no scientist solves $\mathcal{G}$. Any such game is unsolvable.

Theorems (5) and (6) are proved in [17, §3.6]. They follow as corollaries to a general characterization of the solvable versus unsolvable games. The question "What is solvable?" will similarly loom large in the more general theory to be developed here. Within the acceptance based perspective, a central task of inductive logic is to reveal the kind of inductive problems that can be solved by specified kinds of scientists.

[9] If your reasoning yields more than one conjecture on a give clue-sequence, this is considered as equivalent to yielding no conjecture. According to the rules, you can't make two conjectures at the same time.

2.5. Special kinds of scientists

Of course, the most general kind of scientist is "any kind whatsoever." But we are also interested in the inductive power of narrower classes, those that meet some special condition not applicable to scientists in general. One such condition is *computability.* Roughly speaking, a scientist is computable if his input-output behavior can be simulated by a computer program. Putting the matter precisely would require several definitions. Suffice it to say here that the bulk of research on numerical paradigms of inductive inference addresses the scope and limits of computable scientists; see [17].

A simpler class of special scientists is called *memory limited.* A scientist with memory limited to 3 clues (for example) forms his guesses on the basis of (a) the last 3 clues received (or fewer clues if 3 haven't yet arrived), and (b) the last conjecture he made. Thus, if such a scientist makes the same conjectures at the end of clue-sequences (a) and (b), below, then he will make the same conjectures at the end of (c) and (d).

(a)	8	9	5	3	6	
(b)	1	12	4	3	6	
(c)	8	9	5	3	6	13
(d)	1	12	4	3	6	13

This is because the second-to-last conjectures made in response to (c) and (d) are the same [namely, the identical responses made at the end of (a) and (b)], and the last three clues in (c) and (d) are the same (namely, $3, 6, 13$).[10]

The question arises "Does memory limitation stand in the way of solving certain games?" That is, are there any games that can be solved, but not by any member of the restricted

[10] Memory limited scientists were introduced and studied in [30]. Variants on the same idea have also been investigated.

class of memory limited scientists? In this case, memory limitation is *restrictive*. It turns out that memory limitation is indeed restrictive. Our first game (1) witnesses this fact. It can be solved, e.g., by scientists who implement the reasoning described in (3). But no memory limited scientist solves it, regardless of the size of its memory window.[11]

Any class of scientists is known technically as a *strategy*. The computable scientists are a strategy, the memory limited scientists are a strategy, and the class of computable scientists that are also memory limited is yet another strategy. Dozens of strategies have been investigated in recent years. Similarly, there is a profusion of results about modifications of the rules for the basic game sketched above. In variant games, clues might involve noise, scientists might be allowed to work in teams, conjectures may need to be no more than approximately right, and so forth. We refer the reader to [17] for a review of findings on these topics and others. Our more general theory of inductive logic will involve some of the same issues, transposed to the richer context of predicate logic. We hope it has been helpful to encounter some fundamental concepts in the numerical setting. Now it is time to move to higher ground.

3. A first-order paradigm

The goal of the remainder of the essay is to communicate the character of the theory that will be presented in greater detail in the three remaining papers available on our web site. We proceed informally, eschewing generality in favor of clarity. Only a few results will be sketched, with proofs deferred to the other papers. Where possible we make contact with concepts introduced in the simpler numerical context described above in Section 2.

To proceed, the present section introduces our first-order paradigm of inductive in-

[11] In contrast, the game consisting of all finite sets is *not* a witness to the restrictiveness of memory limitation.

quiry.[12] In the next section we'll describe an interesting collection of scientists. They have the particularity of operating via *schemes for hypothesis selection and revision*, which will be explained. Finally, the success of the scientists for solving inductive problems will be assessed. All the concepts discussed below will be taken up again from scratch in the remaining papers, and treated more formally. To understand the "official" theory, you don't have to remember anything from the present essay except the general picture.

As before, inductive problems may be conceived as a guessing game between two players. We still call them "Nature" and "the scientist." In the numerical paradigm the two players agree upon a collection of sets of numbers [as in (1)]. In the present context, they agree on a predicate logic vocabulary and on a *background theory* written in the resulting language. The theory describes some aspects of reality but doesn't pin everything down. For example, it might provide principles governing the way genomes unfold into organisms but leave open a variety of questions about this process. The game then takes place in one of the *realizations* of the background theory. Technically, a realization is a set-theoretical structure. It interprets the logical language used to formulate the background theory.

To illustrate, suppose that our predicate logic embraces just one symbol (aside from the usual logical notation, including the identity symbol $=$). The symbol in question is R, denoting a binary relation. The background theory agreed to by the two players imposes constraints on how R is interpreted (e.g, whether or not R is transitive). For example, our theory might be the following.

(7) Background theory:

(a) R is transitive, irreflexive, asymmetric, and connected.[13]

[12] The qualifier "first-order" refers to the use of predicate logic in the formulation of the theory.

[13] Transitivity means that Rxy, Ryz implies Rxz. Irreflexivity means that Rxx never holds. Asymmetry means that Rxy excludes Ryx. Connectedness means that either Rxy or Ryx must hold if x and y are distinct.

(b) R has a greatest point or a least point.

(c) R does not have *both* a greatest point and a least point.

This theory describes the class of *total orderings* with either a greatest point or least point but not both. There are just three sentences in (7), but in general theories can have any number, even infinitely many.

There are many ways to interpret the binary relation symbol R, and not all of them will be realizations of the theory (7). An interpretation of R will be called a *structure* in what follows. Here are some structures.

(8) Structures:

(a) The positive integers ordered by *less than*.

(b) The negative integers ordered by *greater than*.

(c) The rationals in $(1, 2]$ ordered by *less than*.

(d) The positive and negative integers ordered by *less than*.

(e) The rationals in $[7, 9]$ ordered by *less than*.

The first three structures in (8) are realizations of the background theory (7) since in each case the domain is totally ordered and there is either a least or greatest point but not both.[14] The fourth and fifth structures do not realize the theory. The fourth structure lacks both a greatest and least point, the fifth has both. The domains of these structures are countable, which is the only case we'll consider here.[15]

[14] The *domain* of structure (8)a is the set of positive integers; in (8)e the domain is the set of rational numbers in $[7, 9]$, etc.

[15] A set is *countable* if its members can be associated with unique natural numbers.

There are many structures that realize the background theory (7). The differences between two such structures are often of no interest to us. For example, the set $2, 3, 4 \cdots$ ordered by *less than* is so similar to (8)a that we may wish to group the two together and consider them essentially the same. For this reason, Nature and the scientist agree to *partition* the structures realizing the background theory into non-intersecting classes. For example, the realizations of Theory (7) might be partitioned as follows.

(9) Partition of the realizations of the background theory:

(a) The total orders that have a least point (but no greatest point).

(b) The total orders that have a greatest point (but no least point).

This partition signifies that the two players care only about whether a given realization of the background theory has a least point or a greatest point.

To summarize, the game between Nature and the scientist is played with two pieces. There is a background theory written in the language of predicate logic. And there is a partition of the class of structures that realize the theory.

Nature prepares herself for the game as follows. First she chooses one structure from some cell of the partition. This structure is "reality." Pursuing our illustration, she may choose (8)a, the structure consisting of the positive integers ordered by *less than*. She then gives a temporary name to each element in the domain of her chosen structure. For convenience, we think of Nature as using the individual variables of our predicate language as a stock of temporary names. Suppose that the individual variables are the symbols $v_0, v_1, v_2 \cdots$. Nature might then use the variable v_{32} to name the positive integer 1, the variable v_{238} to name 2, both v_9 and v_{101} as two names for 3, and so forth. You see that Nature is allowed to give more than one name to the same object (the use of the identity sign in the data will signal whether two names refer to the same thing). All the variables must be assigned as names but *Nature does not reveal to us how temporary names are assigned.* Next, Nature selects some way to list all the *basic formulas* true in her chosen structure (and relative to the names she's assigned). Basic formulas, you recall, are atomic formulas or their negations.

An example should make everything clear. Suppose that Nature chooses the positive integers ordered by *less than* as her structure. And suppose that she names the number 8 with both v_8 and v_{23}, and the number 20 with v_2. Then she might begin to list the basic facts as follows.

$$(10)\quad Rv_8v_2 \qquad Rv_8v_3 \qquad v_8 \neq v_2 \qquad \neg Rv_3v_8 \qquad v_8 = v_{23} \qquad \ldots$$

It's as if Nature first tells us:

> "Here's a thing (let's call it v_8) and here's another thing (let's call it v_2), and notice that v_8 is below v_2."

By v_8 being "below" v_2, Nature means that the things named by v_8 and v_2 are related (in the order stated) by the relation named by R. Since we don't know which structure Nature chose at the outset, we can't see behind the symbols v_8, v_2, and R. All we're told is what can be expressed in our predicate logic language (which was fixed at the start of the game).

For the remaining formulas, Nature goes on to say:

> "Here's a thing we can call v_3, and it is above v_8. By the way, the things named v_8 and v_2 are different. The thing named v_3 is *not* below the thing named v_8. The things named v_8 and v_{23} are actually the same thing,"

and so forth. These are the data made available to the scientist. Nature need not apply temporary names in a predictable way, nor list the basic facts in convenient order. Note also that the vocabulary of the background theory will typically be richer, involving multiple predicates, function symbols, and names of its own.

For his part, the scientist examines more and more of Nature's list of data. After each formula, he designates a cell of the partition as his guess about the origin of the structure that Nature chose at the outset. In this sense, the scientist is a function from data into partition-cells. For example, faced with the data in (10), the scientist might respond as follows.

$S(Rv_8v_2)$	$=$	the class of total orders with a least point
$S(Rv_8v_2,\ Rv_8v_3)$	$=$	the class of total orders with a least point
$S(Rv_8v_2,\ Rv_8v_3,\ v_8 \neq v_2)$	$=$	the class of total orders with a greatest point

and so forth. The scientist is here portrayed as a function S that maps each initial segment of the data into one of the two cells in the partition (9). At each stage of the game, the chosen cell is the scientist's guess about the location of the structure Nature chose at the outset. A given guess is correct if the cell holds Nature's choice.

The scientist wins the game just in case all but finitely many of his conjectures are correct. So, in our illustrative game — in which Nature is supposed to have picked the positive integers ordered by *less than* — the scientist wins if his conjectures ultimately stabilize to the cell containing the orders with a least point. Success thus requires that only finitely many mistakes are made, where a "mistake" is conjecturing the wrong cell of the partition or conjecturing nothing at all.

The data available to the scientist from Nature's list of formulas might never imply the correct cell of the partition. In our illustration, at no stage of the game is it made clear whether the order has a least versus greatest point. To see this, consider the first five formulas in the list (10). Depending on what the variables v_8, v_2, v_{23} name, the formulas might be generated from the positive integers ordered by *less than* or from the negative integers ordered by *greater than*. In the first case, the correct cell is (9)a; in the second, it is (9)b. The same ambiguity arises at every stage of the game since you only get to see a finite amount of data at any time.

Nonetheless, the scientist has a reliable strategy for winning the game. He may reason as follows.

(11) THE SCIENTIST'S REASONING: "Nature and I agreed that the chosen structure will contain either a least element or a greatest element but not both. Nature also agreed to give a name to every element in the domain, and to show me all the true basic formulas. So, one of the variables in the data is the name of the *extremal* element of the domain (either its greatest or its least member). I'll start off by supposing that v_0 names the extremal element, in fact, that v_0 names the least member. Hence, I'll start off by conjecturing (9)a in our partition. If my conjecture is wrong, then v_0 is not least, hence, there is some element below it in the R-ordering. This other element also has a name, say y, so eventually I will see Ryv_0 in the data. This will prove to me that my initial conjecture is wrong. I'll then switch to supposing that v_0 names the greatest element, and I'll start conjecturing (9)b. If this new guess is wrong, then I'll eventually find a datum of the form Rv_0y (indeed, I'll check to see whether this datum has already appeared in the data currently available). If Rv_0y shows up, then I'll know that v_0 is not the extremal element of the structure Nature chose. In this case, I'll switch to supposing that v_1 is the extremal element, in fact, that v_1 is least. I'll then behave just as I did with respect to v_0. If v_1 is revealed to be neither greatest nor least (because I see data of the forms Ryv_1 and Rv_1y), then I'll move on to v_2. Proceeding in this way, I'll eventually reach the true extremal element, say v_{324}. If v_{324} is the least element, I'll never see a datum of the form Ryv_{324} so I'll stick with the cell (9)a, which will be right! And if v_{324} is the greatest element, I'll never see a datum of the form $Rv_{324}y$

so I'll stick with the cell (9)b, which will be right! Since I'll reach v_{324} after finitely many incorrect conjectures, I'm guaranteed to make cofinitely correct ones. So I win!

A scientist who relies on the strategy described in (11) can be confident of winning the game, but has no reason to be confident about any particular conjecture made in the course of play. After all, the next datum may well provoke the conjecture's withdrawal. Each conjecture is thus *accepted* by the scientist (provisionally, of course), even though he may have minimal *belief* in its veracity. This illustrates the sense in which our inductive logic concerns acceptance rather than belief.

Let us adapt our example to illustrate a simple game that *cannot* be won. It suffices to remove the last sentence from our background theory (7), the sentence stipulating that R does not have both a greatest and a least point. In other words, we play with the background theory shown here:

(12) New background theory:

(a) R is transitive, irreflexive, asymmetric, and connected.

(b) R has a greatest point or a least point.

The new theory imposes weaker constraints on R than before, so it has more realizations. For example, the set $[3, 5]$ ordered by *less than* satisfies all the clauses of (12) but not (7) since it has both a least point and a greatest point. The set $[3, 5]$ therefore fits into neither of the two cells of the original partition (9). Our new game relies on the following partition instead.

(13) Partition of the realizations of the new background theory:

(a) The total orders that have a least point (but no greatest point).

(b) The total orders that have a greatest point (but no least point).

(c) The total orders that have both a least point and a greatest point.

Now it can be shown that no scientist succeeds reliably in the game based on this expanded partition. In other words, no function from data to cells of the expanded partition is guaranteed to converge to the correct cell. Nature can always choose a realization of the background theory and give names to its elements so that the scientist is led to the wrong conjecture infinitely often. A given scientist might win the game for some choices; but no scientist wins in all cases. For the expanded partition the data are simply not rich enough to support reliable induction.

Let's give a name to reliable success.

(14) DEFINITION: A given scientist *solves* a given partition just in case for every structure in the partition, every way of temporarily naming elements of the structure, and every listing of the basic formulas true in the structure, the scientist's successive conjectures converge to the cell from which the structure was drawn.

Thus, the initial partition (9) is *solvable* whereas the expanded partition (13) is not.

Can the distinction between solvable and unsolvable partitions be characterized in a revealing way? This question is central to inductive logic. The next essay offers combinatorial conditions on solvability, but we won't pause here to summarize the results. Let us also resist the temptation to investigate variants of the inductive games described above (for example, in which success is made a more graded affair). Variants will occupy us later on. We turn instead to a special kind of scientist that plays a central role in our theory. The scientists in question rely on schemes of theory-selection and revision to choose a conjecture following each new datum.

4. Theory revision

4.1. The basics of theory revision

Scientists based on theory revision have two working parts. The first is a starting theory, which the scientist chooses prior to encountering data. Formally, the starting theory is an arbitrary set of predicate logic formulas. We typically denote it by X. The second part is a scheme for revising the starting theory under the impact of data. The scheme is called a *revision function*, typically denoted by the dotted plus $\dotplus$. A revision function $\dotplus$ maps a given theory X and given data d into another theory, $X \dotplus d$.

How the revision function works will be explained shortly. To fill in the general picture, imagine a scientist based on the starting theory X and revision function $\dotplus$ working his way along the data stream (10). His successive conjectures take the form of "revisions," illustrated as follows.

$$
\begin{array}{ll}
\textit{First revision:} & X \dotplus Rv_8v_2 \\
\textit{Second revision:} & X \dotplus (Rv_8v_2,\ Rv_8v_3) \\
\textit{Third revision:} & X \dotplus (Rv_8v_2,\ Rv_8v_3,\ v_8 \neq v_2) \\
 & \quad\text{ETC.}
\end{array}
$$

Each revision is a first-order theory in the sense of being a collection of first-order formulas. The class of such theories is the range of a revision function. Its domain consists of starting theories paired with data.

Now let us conceive of the pair X and $\dotplus$ as a scientist. At each stage of the game, the scientist views an increasing sequence d of basic formulas. In response, he consults his starting theory X and his revision function $\dotplus$. He computes $X \dotplus d$, which is a revision of the starting theory X. Then he asks: "Does the new theory imply a unique cell of the game's partition?" If so, this cell is the scientist's conjecture. In the absence of such a cell, no conjecture is made. Call scientists who operate in this way *revision based*.

We illustrate using our familiar game with partition (9) and data (10). At the third stage, our revision based scientist applies $(Rv_8v_2,\ Rv_8v_3,\ v_8 \neq v_2)$ to the starting theory X. The revision function $\dotplus$ determines how X needs to be revised (if at all) to reconcile it with the three formulas. If the revised theory $X \dotplus (Rv_8v_2,\ Rv_8v_3,\ v_8 \neq v_2)$ implies that R is a total order with a least point and no greatest point then cell (9)a is conjectured. If $X \dotplus (Rv_8v_2,\ Rv_8v_3,\ v_8 \neq v_2)$ implies that R is a total order with a greatest point and no least point then (9)b is conjectured. Otherwise, nothing is conjectured at all.

It remains to describe the revision functions that equip revision based scientists. A wide variety of revision functions have been introduced and analyzed in the literature (see [14] for an extremely useful overview). Our approach is a little different from others.

Our revision functions (like most) are driven by contradiction. If a starting theory is consistent with the current data, then the data are just thrown into the theory; no formulas are withdrawn. But in case of contradiction, theoretical retrenchment is called for before the data can be consistently added. The data themselves are never called into question since they really do come from Nature's chosen structure. It is therefore necessary to abandon some formulas from the background theory. A choice is typically required, however, since there may be more than one way to cut back a theory in view of making it compatible with the data. We can appreciate the issue with an example from sentential logic. Consider the theory and data shown here:

$$(15)\quad \begin{array}{ll} \text{Theory:} & \{\ s,\quad p \rightarrow q,\quad q \rightarrow r\ \} \\ \text{Data:} & (\ p,\quad \neg r\ ) \end{array}$$

The theory contradicts the data because datum p and the theory imply r, which is contradicted by datum $\neg r$. There are several ways to avoid the contradiction while incorporating the data into the theory. Here are some of them.

(16)

One possible revision:	$\{\ s,\quad p \rightarrow q,\quad p,\quad \neg r\ \}$
Another possible revision:	$\{\ s,\quad q \rightarrow r,\quad p,\quad \neg r\ \}$
A third possible revision:	$\{\ s,\quad p,\quad \neg r\ \}$
A fourth possible revision:	$\{\ p,\quad \neg r\ \}$

We can suppress the conditional $q \rightarrow r$, as in the first revision. Alternatively, we could suppress the conditional $p \rightarrow q$, as in the second revision. We could even suppress both conditionals, as in the third revision — or for that matter suppress the entire theory, ending up with just the data, as in the fourth revision.

You can see that the fourth revision has an odd character since the formula s plays no role in the contradiction between data and theory in (15). We say that s is *innocent* in this conflict, a concept that we define as follows.

(17) Definition: Let X be a starting theory, and let d be data. Suppose that f is a member of X. Then f is *innocent* (with respect to X and d) just in case there is no subset A of X such that:

A is inconsistent with d, and

$A - \{f\}$ (that is, A with f removed) is consistent with d.

The idea is that f is innocent of contradicting d if it figures in no non-redundant proof of the negation of d. (For what follows, it's enough to retain just a rough idea of innocence.)

Our conception of theory revision requires that innocent formulas be included in the successor theory. This is the only substantive constraint we impose on revision. The matter can be put as follows.

Home Page

Title Page

Page 27 of 38

Go Back

Full Screen

Close

Quit

(18) Definition: A *revision* of a starting theory X in the face of data d is any consistent set of the form $Y \cup d$, where Y is a subset of X that includes all of X's innocent members.

You see that a revision of a theory in the face of data is the result of adding the data to any subset of the theory that is consistent with the data and includes the innocent formulas. Note that if the data are consistent with the theory to start with, then every member of the theory is innocent, so there is just one possible revision, namely, the theory plus the data. At last we can define a revision function.

(19) Definition: Any mapping of theories X and data d into a revision of X in the face of d is called a *revision function*.

Intuitively, a revision function looks at a given theory and given data, chooses among the allowable retrenchments of the theory, dumps in the data, and returns the result.

4.2. Special forms of theory revision

Definition (19) imposes genuine constraints on revision functions. They must, for example, return a logically consistent theory. The definition is nonetheless liberal enough to admit some unwelcome members — whose choices of successor theory don't appear particularly rational.

For example, a revision function might hold onto *nothing but* the innocent members of a theory whereas more of the theory could be saved. Such a scorched earth policy is illustrated by the third revision in (16). The first two revisions in (16) seem more reasonable. To define away this kind of mischief, let us call a revision function "maxichoice" if it preserves a largest possible subset of the initial theory.[16] The concept can be formulated as follows.

(20) DEFINITION: A revision function $\dot{+}$ is *maxichoice* if and only if for every starting theory X and data-set d, there is no revision function $\oplus$ such that $X \oplus d$ is a proper superset of $X \dot{+} d$.

Maxichoice revision functions implement the conservative policy of inflicting minimal change on existing theories.

There is another revision anomaly that we would like to set aside. Suppose that a revision function behaves as summarized here:

(21)

Theory $T1$	:	$\{ q, \neg p, (q \wedge r) \to p \}$
Datum	:	r
Chosen revision $R1$	:	$\{ q, \neg p, r \}$.
Rejected revision $R2$	:	$\{ q, (q \wedge r) \to p, r \}$.

In the face of datum r, the revision function retrenches Theory $T1$ by throwing out the conditional $(q \wedge r) \to p$ before adding in r. This yields the revision $R1$. A different option would have been to throw out $\neg p$ from $T1$ before adding r, yielding the revision $R2$. We conclude that the revision function prefers to embrace the theory $R1$ compared to $R2$.

Now suppose that the same revision function behaves as summarized here:

(22)

Theory $T2$	:	$\{ r, \neg p, (q \wedge r) \to p \}$
Datum	:	q
Chosen revision $R2$	:	$\{ q, (q \wedge r) \to p , r \}$.
Rejected revision $R1$	:	$\{ q, \neg p, r \}$.

[16] The terminology derives from [1, 10].

That is, faced with datum q and Theory $T2$, the revision function retrenches $T2$ by throwing out $\neg p$ before adding q. This yields the revision $R2$. A different option would have been to throw out $(q \wedge r) \rightarrow p$ from $T2$ before adding q, yielding $R1$. So now we must conclude that there is a preference for theory $R2$ over $R1$.

Nothing prevents the same revision function from behaving as indicated in (21) and (22). In each case, the choice is a revision of the starting theory in the sense of Definition (18). Indeed, even a maxichoice revision function can behave in this way. But the two preferences are inconsistent since the revision function prefers the successor theory $R1$ to $R2$ in the first situation but $R2$ to $R1$ in the second situation. (Both options are available both times.)

To put a stop to such behavior, we can require that a revision function establish once and for all a total ordering of all possible theories, with earlier theories being preferred to later ones. The chosen revision of any given theory must then be the first member in the ordering that satisfies the definition of a revision. In this case, revealed preferences for successor theories will be consistent and transitive. Let us call revision functions that choose their revisions in this way "definite," and put the matter as follows.

(23) DEFINITION: A revision function $\dotplus$ is *definite* just in case there is a linear ordering of the entire class of potential theories in our predicate language such that:

> for any starting theory X and data d, $X \dotplus d$ is the first theory in the ordering that is a revision of X in the face of d.

The expression "revision of X in the face of d" is meant to be interpreted according to Definition (18).

Of course, for there to be definite revision functions, the class of all theories must be ordered so that there always *is* a most preferred one among the potential revisions. This condition can be met. Indeed, it is even possible to arrange matters so that the most preferred revision is never a subset of some other potential revision. That is, there are definite revision functions that are also maxichoice.

Maxichoice definite revision satisfies important criteria of rationality. It manifests transitive preferences among theories, and revises them so as to minimize change and maximize acceptability. In at least this sense, maxichoice definite revision represents a reasonable policy of theory selection.

4.3. Revision used for inquiry

In Section 4.1 we qualified a scientist as "revision based" if his conjectures were determined by the application of a revision function to a starting theory and current data. Now that we have introduced revision functions in a more systematic way, we can define more directly the way they can be used as scientists.

(24) Definition: Let revision function $\dotplus$, starting theory X and partition $\mathcal{P}$ be given. We say that the function

$$R(d) = X \dotplus d$$

solves $\mathcal{P}$ just in case the following scientist solves $\mathcal{P}$. In response to data d, the scientist conjectures the unique cell $\mathcal{P}$ that includes the realizations of $X \dotplus d$. (The scientist must be undefined on d whenever there is no such cell.)

Not all scientists can be conceived in these terms. For example, a scientist who switches conjectures upon receiving the datum $v_0 = v_0$ is not revision based. Other conditions on the behavior of revision based scientists will be discussed in the third essay. The point for now is that scientists based on maxichoice definite revision embody rational strategies of inquiry. Just how rational this makes them can, of course, be debated. But let us turn instead to a more interesting issue. What kinds of partitions can be solved by maxichoice definite revision?

5. The power of revision

The prospects for using revision to solve a given partition depend on the starting theory. Surprisingly, if the starting theory consists entirely of closed formulas (no free variables) then not much can be accomplished.[17] In particular, *no* revision function can be used in such circumstances to solve the partition (9), consisting of the total orders with just a least point, and the total orders with just a greatest point. That is:

(25) FACT: No function of the form $R(d) = X \dotplus d$ solves (9) if X is restricted to closed formulas.

In contrast, we know that other scientists succeed perfectly well on this partition.

A starting theory without free variables gets no traction because it is unable to formulate hypotheses about which particular object is greatest or least. With this defect corrected, it is easy to specify a starting theory that allows a revision based scientist to solve partition (9). In fact, the same is true of *every* solvable partition. The matter can be put as follows.

(26) THEOREM: Let $\mathcal{P}$ partition the structures that realize a given background theory T. If $\mathcal{P}$ is solvable (by any kind of scientist) then there is a consistent extension

[17] A formula like $\exists x(Py \rightarrow Hx)$ is open because the variable y is not governed by a quantifier. The formula $\forall y \exists x(Py \rightarrow Hx)$ is closed.

X of T such that for every maxichoice definite revision function $\dotplus$, the function $R(d) = X \dotplus d$ solves $\mathcal{P}$.

To explain the theorem, take a background theory T. It consists of closed formulas of predicate logic (no free variables). Let $\mathcal{P}$ be any partition of the structures that realize T. We can add to T formulas with free variables to get an expanded theory X. (The expansion must be carried out cleverly, of course, if X is to witness the theorem.) Any revision function that is maxichoice and definite may then be attached to the theory X. The resulting scientist will solve the partition if it is solvable at all. That is, if some scientist can solve the partition, then *any* of our maxichoice definite revision functions can do it, starting with the expanded background theory.

The result does not claim that every partition is solvable in the maxichoice definite way. This stronger claim is is false because there are partitions that can be solved by no scientist whatsoever, e.g., (13). Theorem (26) says rather that maxichoice definite revision yields an optimal method of inquiry (within our first-order paradigm). No scientist can solve anything that can't be solved this way.

As a corollary, any single maxichoice definite revision function suffices for all solvable partitions. There is no need to tailor the revision function to the problem at hand. All that needs tailoring is which particular formulas are added to the background theory T to construct a starting point for revision. The added formulas use free variables to announce hypotheses about witnesses to existential sentences (details in the third essay.) Once these hypotheses are present, maxichoice definite revision of the starting theory is guaranteed to converge to the correct cell of the partition — if such successful induction is possible at all.

There is also a sense in which maxichoice definite revision solves a given partition as fast as possible. Recall that Nature has wide latitude in constructing a stream of data after she chooses a cell of the partition and a structure from that cell. The "stream" is Nature's list of basic formulas that are true in the chosen structure relative to the way she chooses to name elements of the structure.

If a scientist is successful on a given data stream there will come an earliest moment after which all of his conjectures are correct; he will have converged to the correct cell of the partition at that moment. If one scientist never converges later than another on the data streams associated with a given partition, and sometimes converges earlier, we say that the first scientist *dominates* the other. Finally, we say that a successful scientist that is dominated by no other is called *efficient*. To summarize:

(27) DEFINITION: Let S_1 and S_2 be scientists that solve a given partition $\mathcal{P}$. S_1 *dominates* S_2 on $\mathcal{P}$ if the convergence points of S_1 on data-sequences drawn from $\mathcal{P}$ never come later than those for S_2, and sometimes come sooner. A scientist solves $\mathcal{P}$ *efficiently* if no scientist dominates it on $\mathcal{P}$.

We have the following fact.

(28) THEOREM: Let $\mathcal{P}$ partition the structures that realize a given background theory T. If $\mathcal{P}$ is solvable then there is a consistent extension X of T and a maxichoice definite revision function $\dotplus$ such that the function $\Psi(d) = X \dotplus d$ solves $\mathcal{P}$ efficiently.

Thus, every solvable partition can be solved efficiently using maxichoice definite revision. Different revision functions may be needed for different partitions — unlike Theorem (26) (without efficiency) in which every maxichoice definite function works on all solvable partitions. Nonetheless, we see that the strategy of maxichoice definite revision is sufficient for efficient inquiry. This kind of revision is *general purpose*.

There is one more accomplishment of maxichoice definite revision that is worth raising here. You may have noticed that revision based scientists always revise the same starting theory X, put in place at the beginning of inquiry. Progressively longer sequences of data are used for this purpose. Revision is exploited in more iterative fashion by using just Nature's latest datum to revise a continually updated theory. The two schemes can be contrasted by picturing their application to the data stream $d_1, d_2, d_3, d_4 \cdots$. Noniterative revision based scientists apply as shown in (29)a. Iterative revision based scientist work as shown in (29)b.

(29) (a) $X \dotplus d_1,\ X \dotplus (d_1, d_2),\ X \dotplus (d_1, d_2, d_3)\ \cdots$

(b) $X \dotplus d_1,\ (X \dotplus d_1) \dotplus d_2,\ ((X \dotplus d_1) \dotplus d_2) \dotplus d_3,\ \cdots$

Iterated use of maxichoice definite revision can also be used to solve partitions efficiently. Specifically:

(30) THEOREM: Let $\mathcal{P}$ partition the structures that realize a given background theory T. If $\mathcal{P}$ is solvable then there is a consistent extension X of T and a maxichoice definite revision function $\dotplus$ such that the iterative function $S(d_1 \cdots d_n) = (((X \dotplus d_1) \dotplus d_2) \cdots) \dotplus d_n$ solves $\mathcal{P}$ efficiently.

Iteration brings a hint of realism to our conception of scientists since past data are assimilated into the current theory; they need not be forever remembered in their original form. Scientists that rely on iteration resemble the memory limited scientists discussed in the numerical paradigm (Section 2.5).

More can be said about the inductive powers of revision based scientist, both iterated and otherwise. Indeed, more *will* be said in the three essays to follow. But to close the present essay let us return to more general issues.

6. Concluding remarks

It is commonplace to conceive of inductive logic as a generalization of deductive logic. Whereas the latter concerns inferences that can be drawn *with certainty*, the former bears on inferences that can be drawn *with confidence* (see, e.g., [28, 2]). Such a perspective encourages an approach to scientific methodology based on warranted belief. Hypotheses, it would be seem, should be justified by inductive logic, and all the latter can deliver is confidence. In contrast, the approach explored in the present essays abandons the idea of justifying individual hypotheses in terms of warranted belief (or its derivatives, like cognitive utility [22, 23]). Indeed, individual hypotheses are not justified at all in our approach; only scientists are, conceived as schemes for hypothesis selection. Faced with a specific inductive problem, a scientist is praiseworthy if he stabilizes reliably (and efficiently) on the correct hypothesis. Scant justification can be offered to any given hypothesis issued along the way. You saw this for yourself when you played the game (1) in the numerical paradigm of Section 2, and the game based on Partition (9) in the first-order paradigm of Section 3. To win the game, all you could do at a given stage was accept a specific hypothesis on the basis of a larger plan, then hope for the best.

Such is the viewpoint of the inductive logic to be elaborated in the three remaining essays. It attempts to characterize the kinds of "larger plans" that succeed on specified classes of inductive problems. Our logic will be no more than a formal object, of course, the palest reflection of genuine empirical inquiry. In its defense we recall the curious history of formal theories of inference, both deductive and probabilistic. Their clarification and analysis have had remarkable consequences in both philosophical and practical realms [13]. We do not expect the same glorious destiny for acceptance-based inductive logic. But is it too much to hope that our theory will throw light on aspects of scientific inquiry that are not fully illuminated by alternative accounts?

References

[1] C. E. Alchourrón, P. Gärdenfors, and D. Makinson. On the logic of theory change: Partial meet contraction and revision functions. *Journal of Symbolic Logic*, 50:510–530, 1985.

[2] D. Baird. *Inductive Logic: Probability and Statistics*. Prentice Hall, Englewood Cliffs, NJ, 1992.

[3] L. Blum and M. Blum. Toward a mathematical theory of inductive inference. *Information and Control*, 28:125–155, 1975.

[4] L. J. Cohen. *Belief & Acceptance*. Oxford University Press, Oxford, UK, 1992.

[5] J. Dorling. Bayesian personalism, the methodology of scientific research programmes, and Duhem's problem. *Studies in History and Philosophy of Science*, 10:177–87, 1979.

[6] J. Earman. *Bayes or Bust?* MIT Press, Cambridge MA, 1992.

[7] E. Eells. Problems of Old Evidence. *Pacific Philosophical Quarterly*, 66:283–302, 1985.

[8] B. Van Fraassen. The Problem of Old Evidence. In D. Autin, editor, *Philosophical Analysis*, pages 153 – 65. Kluwer Academic Publishers, Dordrecht, 1988.

[9] D. Garber. Old Evidence and Logical Omniscience in Bayesian Confirmation Theory. In J. Earman, editor, *Testing Scientific Theories*, pages 99 – 131. University of Minnesota Press, Minneapolis MN, 1983.

[10] P. Gärdenfors. *Knowledge in Flux: Modeling the Dynamics of Epistemic States*. MIT Press, Cambridge MA, 1988.

[11] C. Glymour. *Theory and Evidence*. Princeton University Press, Princeton, NJ, 1980.

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

[13] J. Halpern, X. Harper, N. Immerman, X. Kolaitis, M. Vardi, and X. Vianu. On the Unusual Effectiveness of Logic in Computer Science. *BSL*, 2001.

[14] S. O. Hansson. *A Textbook of Belief Dynamics: Theory Change and Database Updating*. Kluwer Academic Publishers, Dordrecht, 1999.

[15] P. Horwich. *Probability and Evidence*. Cambridge University Press, Cambridge, 1982.

[16] C. Howson and P. Urbach. *Scientific Reasoning: The Bayesian Approach (2nd Edition)*. Open Court, La Salle, Illinois, 1993.

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

[18] R. Jeffrey. *Probability and the Art of Judgment*. Cambridge University Press, Cambridge, England, 1992.

[19] M. Kaplan. *Decision Theory as Philospohy*. Cambridge University Press, Cambridge UK, 1998.

[20] K. T. Kelly. The Logic of Discovery. *Philosophy of Science*, 54:435–452, 1987.

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

[22] I. Levi. *Gambling with Truth*. MIT Press, Cambridge MA, 1967.

[23] P. Maher. *Betting on Theories*. Cambridge University Press, Cambridge, UK, 1993.

[24] W. C. Myrvold. Bayesianism and Diverse Evidence: A Reply to Andrew Wayne. *Philosophy of Science*, 63:661 – 665, 1996.

[25] K. Popper. *The Logic of Scientific Discovery*. Hutchinson, London, 1959.

[26] H. Putnam. Probability and confirmation. In *Mathematics, Matter, and Method*. Cambridge University Press, 1975.

[27] R. Rosenkrantz. *Inference, Method, and Decision*. Reidel, Boston MA, 1977.

[28] Brian Skyrms. *Choice & Chance: An Introduction to Inductive Logic (3rd Edition)*. Wadsworth, Belmont CA, 1986.

[29] R. J. Solomonoff. A formal theory of inductive inference. *Information and Control*, 7:1–22, 224–254, 1964.

[30] K. Wexler and P. Culicover. *Formal Principles of Language Acquisition*. M.I.T. Press, Cambridge MA, 1980.