May 1989

Uniform Inductive Improvement

Daniel N. Osherson
Massachusetts Institute of Technology

Michael Stob
Calvin College

Scott Weinstein
University of Pennsylvania, weinstein@cis.upenn.edu

University of Pennsylvania Department of Computer and Information Science Technical Report No. MS-CIS-89-29.

Uniform Inductive Improvement

Abstract

We examine uniform procedures for improving the scientific competence of inductive inference machines. Formally, such procedures are construed as recursive operators. Several senses of improvement are considered, including (a) enlarging the class of functions on which success is certain, and (b) transforming probable success into certain success.

Comments

University of Pennsylvania Department of Computer and Information Science Technical Report No. MS-CIS-89-29.

Uniform Inductive Improvement*

Daniel N. Osherson
M. I. T.

Michael Stob
Calvin College

Scott Weinstein
University of Pennsylvania

May 12, 1989

Abstract

We examine uniform procedures for improving the scientific competence of inductive inference machines. Formally, such procedures are construed as recursive operators. Several senses of improvement are considered, including (a) enlarging the class of functions on which sucess is certain, and (b) transforming probable success into certain success.

*Research support was provided by the Office of Naval Research under contract No. N00014-87-K-0401 to Osherson and Weinstein, and by NSF grants DMS-85-21712 and DMS-88-00030 to Stob. Correspondence to D. Osherson, E10-006, M.I.T., Cambridge, MA 02139; e-mail: dan@psyche.mit.edu

1 Introduction

In this article we consider uniform, recursive procedures for amplifying scientific competence. Such a procedure is called an "improvement operator" and may be conceived as a device that is coupled to scientists and behaves as follows. Data are presented to the improvement operator which delivers a possibly transformed version to its attached scientist. The scientist's responses to the data are examined and possibly modified before announced by the improvement operator. Thus, the improvement operator with its attached scientist amount to another scientist, whose competence may be compared to that of the original one. Such a device would be useful if it genuinely improved the competence of a wide class of scientists. Among other uses, improvement operators of this nature would facilitate division of labor in designing automated systems of scientific discovery, allowing the perfection of an approximate system at a later stage of design by application of a uniform algorithm. (Other schemes for the division of labor in designing automated scientists are considered in Pitt & Smith (1988) and Osherson, Stob & Weinstein (1989).) The present work offers some initial steps toward a theory of uniform improvement of scientific competence. Two senses of "improved competence" are investigated. The first sense concerns the range of problems that a scientist solves; the second bears on the probability of solving any given problem.

Our discussion is carried out in the context of the model of inductive inference initiated by Putnam (1975), Solomonoff (1964), Gold (1967), and Blum & Blum (1975). (For literature reviews, see Angluin & Smith (1983) and Osherson, Stob & Weinstein (1986).) In particular, we focus on the inference of total recursive functions. A wide variety of criteria of inductive success have been defined and investigated in this context (see Case & Smith (1983), Royer (1986)). We shall here rely on a success criterion that is more liberal than usual. However, the bulk of our results carry over to the standard, stricter criterion.

Formal scientists in the paradigm at issue are called "inductive inference machines." It will be seen that not every class of inductive inference machines lends itself to uniform improvement. There is a particularly improvable class, however, that will be of central concern in the sequel. These machines are called "rigorous" inasmuch as they always announce testable theories of the function they are investigating. The concept of rigor is a generalization of the concept of "popperian" introduced and investigated by Case & Ngo-Manguelle (1979). (Inductive inference machines conforming to other kinds of scientific strictures have been studied by Fulk (1988) and Weihagen (1976), among others.) Our discussion proceeds as follows. Section 2 introduces the paradigm of scientific inquiry within which our results are stated and proved. Section 3 is devoted to comparing the problem-solving abilities of rigorous and nonrigorous scientists. The class of algorithms for improving scientific competence is introduced in Section 4. Improving the competence of rigorous and nonrigorous scientists in the two senses indicated above is the subject of Sections 5 and 6. Section 7 presents additional results about uniform improvement.

2 A paradigm of scientific inquiry

2.1 Preliminaries

Functions are interpreted extensionally, as sets of ordered pairs. The set of natural numbers is denoted: N. We fix an acceptable ordering of some canonical means of computation over N (e.g., via Turing Machines), and we index the corresponding partial recursive functions φ_0, φ_1,... accordingly (for discussion of these concepts, see Machtey & Young, 1978). The index i may be conceived as a program for the function φ_i. The class of total functions in $\varphi_0, \varphi_1, \ldots$ is denoted: $\mathcal{T}$. $\mathcal{T}$ represents the universe of objects to be investigated by scientists. Let $n \in N$ and $f \in \mathcal{T}$ be given. We denote by "$f[n]$" the sequence: $(0, f(0)), (1, f(1)), \ldots, (n, f(n))$. The class $\{f[n] \mid n \in N \text{ and } f \in \mathcal{T}\}$ is denoted: *Seq*. *Seq* represents the data to which functions in $\mathcal{T}$ give rise.

Let $\sigma = (0, y_0), ..., (i, y_i), ..., (i+n, y_{i+n})$ be given. We denote by "$rng(\sigma)$" the set $\{(0, y_0), ..., (i, y_i), ..., (i+n, y_{i+n})\}$. The subsequence $(0, y_0), ..., (i, y_i)$ is denoted: $\sigma[i]$. For $\tau \in Seq$, "$lh(\tau)$" denotes the number of pairs in τ. Thus, $lh(\sigma) = i + n + 1$ (the length of a sequence always being 1 greater than its largest argument). We use "$\#(A)$" to denote the cardinality of the set A.

Let a recursive isomorphism between *Seq* and N be fixed. Implicit use of this isomorphism allows φ_0, φ_1,... to be construed as functions from *Seq* to N. In this guise, such functions represent the class of possible scientists. Scientists whose input-output behavior is not computer-simulable are thus excluded from the present model. (They are reconsidered in Section 7.2.) When construed as functions from *Seq* to N – i.e., as (formal) scientists – the class $\{\varphi_i \mid i \in N\}$ is denoted: $\mathcal{SCI}$.

2.2 Criterion of success

We now define one sense in which a scientist solves the induction problem represented by a total recursive function.

(1) DEFINITION: Let $f \in \mathcal{T}$, $\theta \in \mathcal{SCI}$ and $j \in N$ be given.

- (a) θ *converges on* f *to* j just in case:
  - i. for all $n \in N$, $\theta(f[n])$ is defined; and
  - ii. for all but finitely many $m \in N$, $\theta(f[m]) = j$.
- (b) θ is said to *weakly identify* f just in case there is $i \in N$ such that:
  - i. θ converges on f to i; and
  - ii. φ_i is an infinite subset of f.

Thus, to weakly identify f, θ's conjectures on successive, finite initial segments of f must eventually stabilize to some one index for an infinite subset of f. As noted above, weak identification is a less stringent concept than generally found in the inductive inference literature. Usually, clause (bii) above is replaced by: $\varphi_i = f$. Our weaker concept represents nontrivial scientific success without insisting that scientific knowledge be complete. For expositional ease, we henceforth abbreviate "weakly identify" to "identify."

(2) **Definition:** $\theta \in \mathcal{SCI}$ is said to *identify* $G \subseteq \mathcal{T}$ just in case θ identifies every $f \in G$. In this case, G is said to be *identifiable.*

Examples of identifiable and unidentifiable subsets of $\mathcal{T}$ are given in Section 2.4. We record the following proposition. Its proof is left for the reader.

(3) **Proposition:** Let $G \subseteq \mathcal{T}$ be identifiable, and let $F \subseteq \mathcal{T}$ be finite. Then $G \bigcup F$ is identifiable.

2.3 Totality

$\theta \in \mathcal{SCI}$ need not be total to identify $f \in \mathcal{T}$. (θ may be undefined on some $\sigma \in Seq$ not drawn from f.) On the other hand, the following proposition shows that we may pass effectively from the index of a scientist to the index of a total scientist that identifies at least as much. Proof of the proposition is straightforward.

(4) **Proposition:** There is a total recursive function g such that for all $i \in N$:

(a) $\varphi_{g(i)}$ is total; and

(b) for all $f \in \mathcal{T}$, if φ_i identifies f then $\varphi_{g(i)}$ identifies f.

2.4 Two subsets of $\mathcal{T}$

The following subsets of $\mathcal{T}$ figure prominently in the sequel.

(5) **Definition:** (based on Blum & Blum, 1975) Let $f \in \mathcal{T}$ be given.

(a) f is *zero-one stabilized* just in case for all $x \in N$, $f(x) \in \{0, 1\}$ and either for all but finitely many $n \in N$, $f(n) = 0$ or for all but finitely many $n \in N$, $f(n) = 1$. $ZOS = \{g \in \mathcal{T} \mid g \text{ is zero-one stabilized}\}$.

(b) f is *self-indexing* just in case for all $x \in N$, $f(x) \in \{0, 1\}$ and $f = \varphi_y$ where y is the least x such that $f(x) = 1$. $SI = \{g \in \mathcal{T} \mid g \text{ is self-indexing}\}$.

It is trivial to show that both ZOS and SI are identifiable. In particular, we now fix $\Gamma \in \mathcal{SCI}$ that identifies ZOS. Γ will figure in several proofs in later sections, as will the following lemma and proposition.

(6) **Lemma:** (Kleene's Recursion Theorem) Let total recursive function h be given. Then there is $n \in N$ such that $\varphi_n = \varphi_{h(n)}$.

Proof: See Rogers (1967, Theorem 11-I). ■

(7) **Proposition:** (based on Blum & Blum, 1975) $ZOS \bigcup SI$ is not identifiable.

PROOF: We adapt a proof given by Blum & Blum (1975), relying as well as on a technique introduced by Barzdins (1974). Suppose that $\theta \in \mathcal{SCI}$ identifies ZOS. We define a total recursive function h such that for all $y \in N$:

(8) $\varphi_{h(y)}(x) = 0$ for $x < y$;

(9) $\varphi_{h(y)}(y) = 1$;

(10) $\varphi_{h(y)}(z) \in \{0, 1\}$ for all $z \in N$; and

(11) θ does not identify $\varphi_{h(y)}$.

Application of Lemma (6) then yields $n \in N$ such that $\varphi_{h(n)} = \varphi_n \in SI$, and θ does not identify φ_n. Hence θ does not identify SI. Consequently, no $\theta \in \mathcal{SCI}$ identifies $ZOS \bigcup SI$.

It remains to specify the function h. Let $y \in N$ be given. Then h(y) is a uniformly constructed index for the function $\varphi_{h(y)}$ defined in stages as follows.

Stage 0: Let $\varphi_{h(y)}(x) = 0$ for all $x < y$. Let $\varphi_{h(y)}(y) = 1$.

Stage n+1: Suppose that $\varphi_{h(y)}[m]$ is defined. Let $s, s' \in ZOS$ be such that:

$s[m] = s'[m] = \varphi_{h(y)}[m]$;

$s(m + k) = 0$ for all $k > 0$; and

$s'(m + k) = 1$ for all $k > 0$.

Then, since θ identifies ZOS, there is a least $j \in N$ such that $\theta(s[m + j]) \neq \theta(s'[m + j])$, because no index is for a function that is an infinite subset of both s and s'. Consequently, one of $\theta(s[m + j]), \theta(s'[m + j])$ is distinct from $\theta(\varphi_{h(y)}[m])$, suppose $\theta(s[m + j])$. For all k with $m < k \leq m + j$, let $\varphi_{h(y)}(k) = s(k)$.

It is evident that $\theta_{h(y)}$ satisfies conditions (8) - (10). Regarding (11), the construction guarantees that θ changes its conjecture infinitely often on $\varphi_{h(y)}$. ■

3 Rigorous Scientists

This section introduces the class of "rigorous" scientists, which will be shown to be uniformly improvable in a strong sense. The concept of rigor was suggested to us by the analogous concept of "popperian" due to Case, J. & Ngo-Manguelle, S. (1979) (see also the concept of "accountability" in Osherson, Stob & Weinstein, 1986).

3.1 Characterization

(12) DEFINITION: Let $\theta \in \mathcal{SCI}$ be given.

(a) θ is *rigorous* just in case for all $\sigma \in Seq$ there is a pair (x, y) such that $(x, y) \in \varphi_{\theta(\sigma)} - rng(\sigma)$. Otherwise, θ is *nonrigorous*.

(b) The class $\{\theta \in \mathcal{SCI} \mid \theta$ is rigorous$\}$ is denoted: $\mathcal{RIG}$.

(c) $G \subseteq \mathcal{T}$ is *rigorously identifiable* just in case some $\theta \in \mathcal{RIG}$ identifies G.

Observe that $\theta \in \mathcal{RIG}$ implies θ total.

3.2 Comparison of $\mathcal{RIG}$ and $\mathcal{SCI}$

A simple construction suffices to verify that every identifiable $G \subseteq \mathcal{T}$ is identified by some nonrigorous $\theta \in \mathcal{SCI}$. On the other hand, since *SI* is identifiable, the following proposition shows that rigor can interfere with identification.

(13) PROPOSITION: *SI* is not rigorously identifiable.

PROOF: Let $\theta \in \mathcal{RIG}$ be given. Using θ we specify a total recursive function h such that for all $y \in N$:

(14) $\varphi_{h(y)}(x) = 0$ for all $x < y$;

(15) $\varphi_{h(y)}(y) = 1$;

(16) $\varphi_{h(y)}(z) \in \{0, 1\}$ for all $z \in N$; and

(17) θ does not identify $\varphi_{h(y)}$.

Application of Lemma (6) then yields $n \in N$ such that

(18) $\varphi_n \in SI$; and

(19) θ does not identify φ_n.

To specify h, let $y \in N$ be given. Then $h(y)$ is an effectively constructed index for the function $\varphi_{h(y)}$ that we now define in stages.

Stage 0: Let $\varphi_{h(y)}(x) = 0$ for all $x < y$. Let $\varphi_{h(y)}(y) = 1$.

Stage n+1: Suppose that $\varphi_{h(y)}[m] = \tau$ is defined. By choice of θ, let (a, b) be the first pair to emerge in some standard enumeration of $\varphi_{(\theta(\tau))} - rng(\tau)$.

case 1: $a \leq m$. (Then $\theta(\tau)$ is not an index for a subset of the function $\varphi_{h(y)}$ being defined.) Let $\varphi_{h(y)}(m+1) = 0$.

case 2: $a > m$. (Then $\theta(\tau)$ is a potential index for an infinite subset of $\varphi_{h(y)}$, so must be cancelled.) Let $\varphi_{h(y)}(r) = 0$ for all $m < r < a$. Let $\varphi_{h(y)}(a) = 1 - b$.

It may be seen that for infinitely many $m \in N$, $\theta(\varphi_{h(y)}[m])$ is not an index for a subset of $\varphi_{h(y)}$. Consequently, (17) is satisfied. It is evident that (14)-(16) are satisfied. ■

4 Improvement operators

Let $\mathcal{P}$ be the class of all functions from N to N. Members of $\mathcal{P}$ may be partial or total, recursive or nonrecursive. Officially, an *improvement operator* is any "recursive operator" in the sense of Rogers (1967, Section 9.8), that is, any effective, total mapping of $\mathcal{P}$ into $\mathcal{P}$. Rather than review the definition of recursive operator, we state the properties of improvement operators to be used in what follows.

(20) PROPERTY: Let J be an improvement operator. Then there is a total recursive function f such that for all $i \in N$, $J(\varphi_i) = \varphi_{f(i)}$.

(21) PROPERTY: Let M be a "Turing machine with oracle," or "oracle machine" in the sense of Rogers (1967, Section 9.2). Then there is an improvement operator J such that for all total $\theta \in \mathcal{SCI}$, $J(\theta)$ is the function computed by M^{θ}.

(22) PROPERTY: The class of improvement operators is closed under composition.

(23) PROPERTY: Let improvement operator J be given, and let $\theta, \Omega \in \mathcal{SCI}$ be such that $\theta \subseteq \Omega$. Then $J(\theta) \subseteq J(\Omega)$.

(24) PROPERTY: The class of improvement operators is countably infinite.

Properties (22) and (23) are taken directly from Rogers (1967). Property (24) follows immediately from the countability of the partial recursive functions. Properties (20) and (21) are easily deduced from the relevant definitions.

Note that improvement operators map functions into functions. They do not map programs into programs. Algorithms of this latter sort are of limited use since they can be applied only to scientists with known internal program. We return to this topic in Section 7.2.

5 Scope improvement

The present section addresses the first sense of improvement mentioned earlier, namely, expanding the subset of $\mathcal{T}$ that a scientist identifies.

5.1 Definition of scope improvement

(25) DEFINITION: Let $\theta \in \mathcal{SCI}$ be given. The *scope of* θ, denoted "$Sc[\theta]$", is the set $\{f \in \mathcal{T} \mid \theta \text{ identifies } f\}$.

(26) DEFINITION: Let $G \subseteq \mathcal{T}$, $\theta \in \mathcal{SCI}$, $\Sigma \subseteq \mathcal{SCI}$, and improvement operator J be given.

(a) J *G-improves* θ just in case $G \bigcup Sc[\theta] \subseteq Sc[J(\theta)]$.

(b) J *G-improves* Σ just in case J G-improves every $\theta \in \Sigma$.

(c) Σ is *G-improvable* just in case some improvement operator G-improves Σ.

The following proposition illustrates scope improvement; its simple proof is left for the reader (see Proposition (3)).

(27) PROPOSITION: Let finite $G \subseteq \mathcal{T}$ be given. Then, $\mathcal{SCI}$ is G-improvable.

5.2 Limits on scope improvement

It is obvious that if "finite" is suppressed in Proposition (27) then the resulting claim is false. For, in this case G may be taken to be an unidentifiable subset of $\mathcal{T}$, thereby excluding all hope of G-improvement. Similarly, $\theta \in \mathcal{SCI}$ cannot be G-improved for any $G \subseteq \mathcal{T}$ such that $Sc[\theta] \bigcup G$ is not identifiable. It may still be asked, however, whether $\Sigma \subset \mathcal{SCI}$ is G-improvable if for all $\theta \in \Sigma$, $G \bigcup Sc[\theta]$ is identifiable. The following proposition provides a (very) negative answer to this question. It exhibits a class of scientists each with finite scope that is not G-improvable, even though G itself is quite simple.

(28) PROPOSITION: $\{\theta \in \mathcal{SCI} \mid \#(Sc[\theta]) \leq 1\}$ is not ZOS-improvable.

PROOF: Let total recursive function d be such that for all $i \in N$,

$$\varphi_{d(i)} = \{(x, y) \mid \varphi_i(x) = y \ \& \ (\forall z \leq x)(\varphi_i(z) \text{ converges})\}.$$

Then, for all $i \in N$:

$$\varphi_{d(i)} = \begin{cases} \varphi_i & \text{if } \varphi_i \in \mathcal{T}; \\ \text{a finite function} & \text{otherwise.} \end{cases}$$

Let total recursive function h be such that for all $i \in N$, $\varphi_{h(i)}$ is the constant $d(i)$-function. Then, for all $i \in N$:

$$Sc[\varphi_{h(i)}] = \begin{cases} \{\varphi_i\} & \text{if } \varphi_i \in \mathcal{T}; \\ \emptyset & \text{otherwise.} \end{cases}$$

Hence, $Sc[\varphi_{h(i)}] \leq 1$, for all $i \in N$.

For a contradiction, suppose that improvement operator J ZOS-improves $S = \{\varphi_{h(i)} \mid i \in N\}$. By Property (20) let total recursive function g be such that for all $j \in N$, $J(\varphi_j) = \varphi_{g(j)}$. Let z be an index for the constant zero-function. Define $\Omega \in \mathcal{SCI}$ as follows. For all $\sigma \in Seq$:

$$\Omega(\sigma) = \begin{cases} \varphi_{g(h(i))}(\sigma) & \text{where } i \text{ is least such that } (i, 1) \in rng(\sigma), \text{ if such an } i \text{ exists;} \\ z & \text{otherwise.} \end{cases}$$

Then, it is easy to verify that Ω identifies $ZOS \bigcup SI$, contradicting Proposition (7). Hence, S is not ZOS-improvable. Hence, no superset of S is ZOS-improvable. ■

5.3 Scope improvement of $\mathcal{RIG}$

In contrast to the foregoing proposition, the next result indicates the extent to which rigorous scientists lend themselves to scope improvement:

(29) PROPOSITION: Let $G \subseteq \mathcal{T}$ be rigorously identifiable. Then $\mathcal{RIG}$ is G-improvable.

We rely on the following lemma.

(30) LEMMA: Let e be an index for $\emptyset$. Then there is an improvement operator J such that for all $\theta \in \mathcal{RIG}$:

(a) $Sc[\theta] = Sc[J(\theta)]$; and

(b) for all $f \in \mathcal{T}$, if $J(\theta)$ does not identify f then $J(\theta)(f[n]) = e$ for infinitely many $n \in N$.

Here and elsewhere, the following notation and terminology will be helpful. Given $i, j \in N$, the partial recursive function $\varphi_{i,j}$ is defined as follows. For all $k \in N$:

$$\varphi_{i,j}(k) = \begin{cases} \varphi_i(k) & \text{if program } i \text{ converges on input } k \text{ within } j \text{ steps;} \\ \text{undefined} & \text{otherwise.} \end{cases}$$

Let $X \subseteq N \times N$ and $Y \subseteq N \times N$ be given. We say that X and Y *conflict* just in case there are $n, m, m' \in N$ such that $(n, m) \in X$, $(n, m') \in Y$, and $m \neq m'$. Given $\sigma \in Seq$ of nonzero length, we denote by "$\sigma-$" the result of removing σ's last pair.

PROOF OF LEMMA (30): By Property (21) (and since rigorous scientists are total), it suffices to exhibit an oracle machine M such that for all $\theta \in \mathcal{RIG}$, conditions (a) and (b) of the lemma are satisfied if $J(\theta)$ is the function computed by M^θ. Let $\theta \in \mathcal{SCI}$ and $\sigma \in Seq$ be given. Then:

$$M^\theta(\sigma) = \begin{cases} e & \text{if } lh(\sigma) = 1; \\ \text{undefined} & \text{if either } \theta(\sigma) \text{ or } \theta(\sigma-) \text{ are undefined;} \\ \theta(\sigma) & \text{if } \theta(\sigma) = \theta(\sigma-) \text{ and } rng(\sigma) \text{ and } \varphi_{\theta(\sigma), lh(\sigma)} \text{ do not conflict;} \\ e & \text{otherwise.} \end{cases}$$

To see that M behaves as desired, let $\theta \in \mathcal{RIG}$ and $f \in \mathcal{T}$ be given. If θ identifies f then θ converges on f to some $i \in N$ such that θ_i is an infinite subset of f. In this case, $M^\theta(f[m]) = \theta(f[m])$ for all but finitely many $m \in N$, so M^θ identifies f. Suppose now that θ does not identify f. Since θ is rigorous, there is no $m \in N$ such that both φ_m is a finite subset of f, and θ converges on f to m. Consequently, either:

(31) θ converges to no index on f; or

(32) θ converges on f to an index for a function that is not a subset of f.

In both cases (31) and (32), $M^\theta(f[m]) = e$ for infinitely many $m \in N$. ∎

PROOF OF PROPOSITION (29): We rely on a construction employed by Minicozzi (see Blum & Blum, 1975). Let $\Omega \in \mathcal{RIG}$ identify G. By Property (21), it suffices to exhibit an oracle machine Z such that for all $\theta \in \mathcal{RIG}$, Z^θ identifies $G \bigcup Sc[\theta]$. Let J and e be as specified in Lemma (30). Define the function $P : \mathcal{SCI} \times Seq \to N$ as follows. For all $\theta \in \mathcal{SCI}$, $\sigma \in Seq$:

$$P(\theta, \sigma) = \begin{cases} 0 & \text{if } lh(\sigma) = 1; \\ \text{undefined} & \text{if any of } \theta(\sigma[0]) \ldots \theta(\sigma[lh(\sigma) - 1]) \\ & \text{are undefined;} \\ 0 & \text{if } \theta(\sigma[x]) \neq e \text{ for all } x < lh(\sigma); \\ \text{the greatest } x < lh(\sigma) & \text{otherwise.} \\ \quad \text{such that } \theta(\sigma[x]) = e & \end{cases}$$

It is clear that P is "recursive in θ", i.e., computable by a suitable oracle machine. Note that if $\theta \in \mathcal{RIG}$ (hence total), $P(\theta, \sigma)$ is defined for all $\sigma \in Seq$.

Z may now be specified as follows. For all $\theta \in \mathcal{SCI}$, $\sigma \in Seq$:

$$Z^{\theta}(\sigma) = \begin{cases} \text{undefined} & \text{if } P(J(\theta), \sigma) \text{ is undefined;} \\ J(\theta)(\sigma) & \text{if } P(J(\theta), \sigma) \leq P(J(\Omega), \sigma); \\ J(\Omega)(\sigma) & \text{otherwise.} \end{cases}$$

Informally, Z^{θ} adopts $J(\theta)$'s conjecture until $J(\theta)$ emits e, at which point Z^{θ} adopts $J(\Omega)$'s conjecture, etc. Z is a computable function in θ because J and P are.

To see that Z behaves as desired, let $f \in G \bigcup Sc[\theta]$ be given. If $f \in G - Sc[\theta]$, then $J(\Omega)$ converges to an index j for an infinite subset of f whereas $J(\theta)$ emits e infinitely often. Hence Z^{θ} converges on f to j, and thus identifies f. The cases $f \in Sc[\theta] - G$ and $f \in G \bigcap Sc[\theta]$ are similar. ■

Since *ZOS* is rigorously identifiable (as easily verified), Proposition (29) yields the following corollary, to be compared with Proposition (28).

(33) COROLLARY: $\mathcal{RIG}$ is *ZOS*-improvable.

5.4 Limits on scope improvement within $\mathcal{RIG}$

For $G \subseteq \mathcal{T}$ that is not rigorously identifiable, G-improvement within $\mathcal{RIG}$ is also limited in nontrivial ways. This is the content of the following proposition.

(34) PROPOSITION: $\{\theta \in \mathcal{RIG} \mid \#(Sc[\theta]) \leq 2\}$ is not *SI*-improvable.

Of course, since *SI* is identifiable, Proposition (3) implies that $SI \bigcup Sc[\theta]$ is identifiable for any $\theta \in \mathcal{SCI}$ with $\#(Sc[\theta]) \leq 2$. For proof of the proposition we rely on a strengthened form of Lemma (6). Let $\langle \cdot, \cdot \rangle$ be a recursive isomorphism between $N \times N$ and N.

(35) LEMMA: (Smullyan) Let total recursive functions g and h be given. Then there are $m, n \in N$ such that $\varphi_{g(\langle m,n \rangle)} = \varphi_m$ and $\varphi_{h(\langle m,n \rangle)} = \varphi_n$.

PROOF: See Rogers (1967, Theorem 11-X(a)). ■

PROOF OF PROPOSITION (34): For a contradiction, suppose that improvement operator J *SI*-improves $R = \{\theta \in \mathcal{RIG} \mid \#(Sc[\theta]) \leq 2\}$. Then by Property (20) there is a total recursive function f such that:

(36) for all $i \in N$, if $\varphi_i \in R$ then $SI \bigcup Sc[\varphi_i] \subseteq Sc[\varphi_{f(i)}]$.

We shall define total recursive functions g and h such that for all $y, z \in N$:

(37) $\varphi_{g(\langle y,z \rangle)}(x) = 0$ for all $x < y$;

(38) $\varphi_{g(\langle y,z \rangle)}(y) = 1$;

(39) $\varphi_{h(\langle y,z \rangle)} \in R$; and

(40) either:

(a) $\varphi_{g(\langle y,z\rangle)}(w) \in \{0,1\}$ for all $w \in N$, and $\varphi_{f(z)}$ does not identify $\varphi_{g(\langle y,z\rangle)}$; or

(b) $\varphi_{f(z)}$ does not identify $Sc[\varphi_{h(\langle y,z\rangle)}]$.

An application of Lemma (35) to g and h yield $m, n \in N$ such that $\varphi_{g(\langle m,n\rangle)} = \varphi_m$ and $\varphi_{h(\langle m,n\rangle)} = \varphi_n$. Hence by (37) – (40):

(41) $\varphi_m(x) = 0$ for all $x < m$;

(42) $\varphi_m(m) = 1$;

(43) $\varphi_n \in R$; and

(44) either:

(a) $\varphi_m(w) \in \{0,1\}$ for all $w \in N$, and $\varphi_{f(n)}$ does not identify φ_m; or

(b) $\varphi_{f(n)}$ does not identify $Sc[\varphi_n]$.

Since (41), (42), and (44)a imply that $\varphi_m \in SI$, (41) – (44) contradict (36).

Define $p : Seq \to N$ so that for all $\sigma \in Seq$, $x \in N$:

$$\varphi_{p(\sigma)}(x) = \begin{cases} 0 & \text{if } x \leq lh(\sigma) + 1; \\ \text{undefined} & \text{otherwise.} \end{cases}$$

Observe that a scientist who conjectures $\varphi_{p(\sigma)}$ on $\sigma \in Seq$ does not compromise his rigor thereby but that such a conjecture is incorrect about any $f \in \mathcal{T}$. In the construction to follow, p will be used to disrupt identification but ensure rigor.

Let $y, z \in N$ be given. We now specify, in stages, both $\varphi_{g(\langle y,z\rangle)}$ and $\varphi_{h(\langle y,z\rangle)}$. For notational ease, we abbreviate $g(\langle y,z\rangle)$ to g, $h(\langle y,z\rangle)$ to h, and $f(z)$ to f.

Stage 0: Let $\varphi_g(x) = 0$ for all $x < y$. Let $\varphi_g(y) = 1$. Let $\varphi_h(\sigma) = p(\sigma)$ for all $\sigma \in Seq$ with $lh(\sigma) \leq y + 1$.

Stage 2n+1: Let $m \in N$ be greatest such that both $\varphi_g[m]$ and $\varphi_h(\sigma)$ with $lh(\sigma) \leq m + 1$ are already defined. For all $v \in N$ such that $\varphi_{f,v}(\varphi_g[m])$ is divergent, let $\varphi_g(m + v + 1) = 0$, and let $\varphi_h(\sigma) = p(\sigma)$ for all $\sigma \in Seq$ with $lh(\sigma) = m + v + 2$. Let r be the least number, if such exists, for which $\varphi_{f,r}(\varphi_g[m])$ converges. Remark: If $\varphi_f(\varphi_g[m])$ diverges then the present stage does not terminate. In this case, $\varphi_g(w) \in \{0,1\}$ for all $w \in N$, $\varphi_h \in \mathcal{RIG}$, $Sc[\varphi_h] = \emptyset$, and φ_f does not identify φ_g. On the other hand, if $\varphi_f(\varphi_g[m])$ converges, then for the r above, $\varphi_g(x)$ is defined for all $x \leq m + r$, and $\varphi_h(\sigma)$ is defined for all σ with $lh(\sigma) \leq m + r + 1$. Assuming that the present stage terminates, carry these numbers m and r into the next stage.

Stage 2n+2: Let $s, s' \in \mathcal{T}$ be such that:

(a) $s(x) = s'(x) = \varphi_g(x)$ for all $x \leq m + r$;

(b) $s(m + r + k) = 0$ for all $k \geq 1$; and

(c) $s'(m+r+k) = 1$ for all $k \geq 1$.

Call $\langle u, v\rangle \in N$ "good" just in case $\varphi_{f,v}(s[m+r+u+1])$ and $\varphi_{f,v}(s'[m+r+u+1])$ are both convergent and they are not equal; call $\langle u, v\rangle$ "bad" otherwise. Search in increasing order for the least, good $\langle u, v\rangle \in N$. For each bad $\langle u, v\rangle$ encountered in this search, extend φ_h as follows. For all $\sigma \in Seq$ with $lh(\sigma) = m+r+u+1$:

$$\varphi_h(\sigma) = \begin{cases} \Gamma(\sigma) & \text{if } \sigma = s[m+r+u+1] \text{ or } \sigma = s'[m+r+u+1]; \\ p(\sigma) & \text{otherwise.} \end{cases}$$

Let $\langle t, d\rangle$ be the least, good number, if such exists. Then, either

(a) $\varphi_{f,d}(s[m+r+t+1]) \neq \varphi_f(\varphi_g[m+r+1])$, or

(b) $\varphi_{f,d}(s'[m+r+t+1]) \neq \varphi_f(\varphi_g[m+r+1])$.

Choose the smaller of $s[m+r+t+1]$, $s'[m+r+t+1]$ such that $\varphi_{f,d}$ of it is not equal to $\varphi_f(\varphi_g[m+r+1])$, say $s[m+r+t+1]$. Let $w = \max\{u \in N \mid (\exists v \in N)(\langle u, v\rangle \leq \langle t, d\rangle\}$. (Thus, $w \geq t$ and represents the longest initial segment of s examined in the search for a good number.) For all x with $m+r < x \leq m+r+w+1$, let $\varphi_g(x) = s(x)$. Let $\varphi_h(\sigma) = p(\sigma)$ for all $\sigma \in Seq$ of length $m+r+w+2$. (φ_h has already been defined for all shorter $\sigma \in Seq$.) Go to stage $2(n+1)+1$. Remark: If there is no good number, then the present stage does not terminate. In this case, $\varphi_h \in \mathcal{RIG}$, $Sc[\varphi_h] = \{s, s'\}$ and φ_f does not identify both s and s' (because no index is for a function that is an infinite subset of both s and s'), i.e., φ_f does not identify $Sc[\varphi_h]$. On the other hand, if there is a good $\langle t, d\rangle \in N$, then for the w defined above:

(a) $\varphi_g(x)$ is defined for all $x \leq m+r+w+1$;

(b) $\varphi_h(\sigma)$ is defined for all σ with $lh(\sigma) \leq m+r+w+2$;

(c) φ_f's convergence on φ_g is delayed; and

(d) φ_h's convergence on both s and s' is delayed.

It is clear from the associated remarks that if any stage in the foregoing construction fails to terminate, then (37) – (40) above are satisfied. On the other hand, if all stages terminate, then $\varphi_h \in \mathcal{RIG}$, $Sc[\varphi_h] = \emptyset$, $\varphi_g(w) \in \{0, 1\}$ for all $w \in N$, and φ_f converges to no index on φ_g. Consequently, (37) – (40) are satisfied in this case too. ■

6 Probability improvement

We now consider the second sense of improvement mentioned in the introduction, namely, transforming probable success into certain success.

6.1 Probabilistic identification

For purposes of the present section it is necessary to extend our model of scientific inquiry so as to incorporate less-than-certain identification. With minor departures, we follow Wiehagen, Freivald & Kinber (1984) and Pitt (1985).

(45) DEFINITION:

(a) The class of all ω-sequences over $\{0,1\}$ is denoted: *Coin*.

(b) Given $c \in Coin$ and $i \in N$, the ith member of c is denoted: c_i.

(c) Let $f \in \mathcal{T}$ and $c \in Coin$ be given. The function $\{(x, \langle y, c_x \rangle) \mid f(x) = y\}$ is denoted: fc.

Intuitively, fc contains information both about f and about the successive tosses of a fair coin.

(46) DEFINITION: Let $\theta \in \mathcal{SCI}$, $f \in \mathcal{T}$ and $c \in Coin$ be given. *θ identifies f in the presence of* c just in case there is $i \in N$ such that:

(a) θ converges on fc to i; and

(b) φ_i is an infinite subset of f.

To proceed, we let $\mathcal{M}$ be the natural probability measure on *Coin*. Specifically, $\mathcal{M}$ is taken to be the unique, complete probability measure such that for each finite sequence δ over $\{0,1\}$, $\mathcal{M}(\{c \in Coin \mid c \text{ begins with } \delta\}) = 2^{-lh(\delta)}$. The following lemma is straightforwardly proven.

(47) LEMMA: Let $\theta \in \mathcal{SCI}$ and $f \in \mathcal{T}$ be given. Then $\mathcal{M}(\{c \in Coin \mid \theta \text{ identifies } f \text{ in the presence of } c\})$ is defined.

(48) DEFINITION: Let $p \in [0,1]$, $\theta \in \mathcal{SCI}$, $f \in \mathcal{T}$, and $G \subseteq \mathcal{T}$ be given.

(a) *θ identifies f with probability* p just in case $\mathcal{M}(\{c \in Coin \mid \theta \text{ identifies } f \text{ in the presence of } c\}) \geq p$.

(b) *θ identifies G with probability* p just in case θ identifies every $f \in G$ with probability p.

(c) G is *identifiable with probability* p just in case some $\theta \in \mathcal{SCI}$ identifies G with probability p.

We continue to employ the term "identify" without probability qualification in the sense of Section 2. A simple construction suffices to prove the following proposition.

(49) PROPOSITION: If $G \subseteq \mathcal{T}$ is identifiable, then G is identifiable with probability 1.

A less trivial fact about probabilistic identification may be formulated as follows.

(50) PROPOSITION:

(a) $SI \bigcup ZOS$ is identifiable with probability $\frac{1}{2}$.

(b) For all $p > \frac{1}{2}$, $SI \bigcup ZOS$ is not identifiable with probability p.

The proof of clause (a) of the proposition is left for the reader. Clause (b) follows from Proposition (7) and the corollary to the following result, due to Wiehagen, Freivald & Kinber (1984).

(51) PROPOSITION: (Wiehagen, Freivald & Kinber, 1984) There is an improvement operator J such that for all total $\theta \in \mathcal{SCI}$ and $f \in \mathcal{T}$, if θ identifies f with probability $p > \frac{1}{2}$ then $J(\theta)$ identifies f.

PROOF: An adaptation of the proof of Proposition 10.6B of Osherson, Stob & Weinstein (1986). ■

(52) COROLLARY: Let $G \subseteq \mathcal{T}$ be given. If G is identifiable with probability $p > \frac{1}{2}$, then G is identifiable.

PROOF: By Proposition (51) and a simple adaptation of Proposition (4) to the probabilistic setting. ■

6.2 Rigor in the probabilistic context

The concept of "rigorous scientist" formulated in Definition (12) may be adapted to the present setting as follows.

(53) DEFINITION: Let $\theta \in \mathcal{SCI}$ be given.

(a) θ is *probabilistically rigorous* just in case for all $f \in \mathcal{T}$, $n \in N$, and $c \in Coin$, there is a pair (x, y) such that $(x, y) \in \varphi_{\theta((fc)[n])} - rng(f[n])$.

(b) The class $\{\theta \in \mathcal{SCI} \mid \theta$ is probabilistically rigorous$\}$ is denoted: *PRIG*.

Similarly to before, rigor can interfere with probabilistic identification.

(54) PROPOSITION:

(a) *SI* is identifiable with probability 1.

(b) No $\theta \in PRIG$ identifies *SI* with probability $p > 0$.

Clause (a) follows from Proposition (49). The proof of clause (b) is deferred to Section 6.5.

6.3 Definition of probability improvement

(55) DEFINITION: Let $p \in [0, 1]$ and $\theta \in \mathcal{SCI}$ be given. The *p-scope* of θ, denoted "$Sc_p[\theta]$", is the set $\{f \in \mathcal{T} \mid \theta$ identifies f with probability $p\}$.

(56) DEFINITION: Let $p \in [0, 1]$, $\theta \in \mathcal{SCI}$, $\Sigma \subseteq \mathcal{SCI}$, and improvement operator J be given.

(a) J *p-improves* θ just in case $Sc_p[\theta] \subseteq Sc[J(\theta)]$.

(b) J *p-improves* Σ just in case J p-improves every $\theta \in \Sigma$.

(c) Σ is *p-improvable* just in case some improvement operator p-improves Σ.

Observe that p-improvement gives rise to scientists that identify in the original sense of Section 2.2. To illustrate, Proposition (51) implies:

(57) PROPOSITION: Let $p > \frac{1}{2}$ be given. Then $\mathcal{SCI}$ is p-improvable.

6.4 Probability improvement in $\mathcal{SCI}$

Propositions (7) and (50) imply that $\mathcal{SCI}$ is not $\frac{1}{2}$-improvable. But even if we select just those scientists whose $\frac{1}{2}$-scopes *are* identifiable, the resulting set is not $\frac{1}{2}$-improvable. This is the content of the following proposition.

(58) PROPOSITION: $\{\theta \in \mathcal{SCI} \mid Sc_{\frac{1}{2}}[\theta]$ is identifiable$\}$ is not $\frac{1}{2}$-improvable.

The following notation will be helpful. The class of finite sequences over $\{0,1\}$ is denoted: *CoinSeq*. Given $\sigma = (0, y_0), ..., (n, y_n) \in Seq$ and $\alpha = x_0, ..., x_n$ in *CoinSeq*, the sequence $(0, \langle y_0, x_0 \rangle), ..., (n, \langle y_n, x_n \rangle) \in Seq$ is denoted: $\sigma\alpha$. (N.B. $\sigma\alpha$ is not the concatenation of α to the end of σ.) For $\alpha \in CoinSeq$ and $n < lh(\alpha)$, $\alpha[n]$ is the initial segment of length n in α.

PROOF: As in the proof of Proposition (28), let $d \in \mathcal{T}$ be such that for $i \in N$:

$$\varphi_{d(i)} = \begin{cases} \varphi_i & \text{if } \varphi_i \in \mathcal{T}; \\ \text{a finite function} & \text{otherwise.} \end{cases}$$

Let total recursive function h be such that for all $i \in N$, $\sigma \in Seq$, and $\alpha \in CoinSeq$ with $lh(\alpha) = lh(\sigma)$:

$$\varphi_{h(i)}(\sigma\alpha) = \begin{cases} d(i) & \text{if the first member of } \alpha \text{ is } 1; \\ \Gamma(\sigma) & \text{otherwise.} \end{cases}$$

It may be seen that for all $i \in N$, $Sc_{\frac{1}{2}}[\varphi_{h(i)}] = ZOS \bigcup \{\varphi_i\}$ if $\varphi_i \in \mathcal{T}$; $= ZOS$ otherwise. Consequently, for all $i \in N$, $Sc_{\frac{1}{2}}[\varphi_{h(i)}]$ is identifiable.

For a contradiction, suppose that improvement operator J $\frac{1}{2}$-improves $S = \{\varphi_{h(i)} \mid i \in N\}$. By Property (20) let total recursive function g be such that for all $j \in N$, $J(\varphi_j) = \varphi_{g(j)}$. Define $\Omega \in \mathcal{SCI}$ as follows. For all $\sigma \in Seq$:

$$\Omega(\sigma) = \begin{cases} \varphi_{g(h(i))}(\sigma) & \text{where } i \text{ is least such that } (i, 1) \in rng(\sigma), \text{ if such an } i \text{ exists;} \\ \varphi_{g(h(0))}(\sigma) & \text{otherwise.} \end{cases}$$

Then, it is easy to verify that Ω identifies $ZOS \bigcup SI$, contradicting Proposition (7). ■

Positive results about a related concept of probability improvement may be found in Pitt (1985).

6.5 Probability improvement in *PRIG*

In contrast to Proposition (58), rigorous scientists lend themselves to probability improvement in a strong sense.

(59) PROPOSITION: (based on Pitt, 1985) There is an improvement operator J such that for all $\theta \in PRIG$:

(a) $J(\theta) \in \mathcal{RIG}$;

(b) for all $f \in \mathcal{T}$, if there is $c \in Coin$ such that θ identifies f in the presence of c, then $J(\theta)$ identifies f.

That is, J creates from $\theta \in PRIG$ a rigorous scientist (in the nonprobabilistic sense) that identifies any $f \in \mathcal{T}$ that θ identifies in the presence of even one coin.

PROOF: By Property (21) (and since $PRIG \subseteq \{\theta \in \mathcal{SCI} \mid \theta \text{ is total}\}$), it suffices to exhibit an oracle machine M such that for all $\theta \in PRIG$, M^{θ} behaves as claimed for $J(\theta)$. In what follows we shall be concerned only with M's behavior on total $\theta \in \mathcal{SCI}$; $M^{\theta}(\sigma)$ may harmlessly diverge for nontotal θ. To compute $M^{\theta}(\sigma)$, let $\theta \in \mathcal{SCI}$ and $\sigma \in Seq$ be given. Let $i < lh(\sigma)$ be least (if such exists) such that there exists $\beta \in CoinSeq$ of length $lh(\sigma)$ such that:

(60) $rng(\sigma)$ and $\varphi_{\theta((\sigma\beta)[i]),lh(\sigma)}$ do not conflict.

(61) $i \leq j < lh(\sigma)$ implies $\theta((\sigma\beta)[j]) = \theta((\sigma\beta)[i])$.

The existence of such an i can be determined in finite time, since $\{\beta \in CoinSeq \mid lh(\beta) = lh(\sigma)\}$ is finite. If no such i exists, let $M^{\theta}(\sigma) = \theta(\sigma\alpha)$ where $\alpha \in CoinSeq$ has length $lh(\sigma)$ and consists entirely of 0's. Otherwise, let β be lexicographically least satsifying (60) and (61) and let $M^{\theta}(\sigma) = \theta((\sigma\beta)[i])$.

To verify (a) and (b) of the proposition, let $\theta \in PRIG$ be given. For (a), we must show that for all $f \in \mathcal{T}$ and $k \in N$, $\varphi_{M^{\theta}(f[k])} - rng(f[k]) \neq \emptyset$. There are two cases, corresponding to the two cases in the definition of $M^{\theta}(f[k])$. But in both cases, $M^{\theta}(f[k]) = \theta(f[k]\alpha)$ for some $\alpha \in CoinSeq$ of length $k+1$. Since $\theta \in PRIG$, $\varphi_{\theta(f[k]\alpha)} - rng(f[k]) \neq \emptyset$.

To verify (b), suppose that $f \in \mathcal{T}$ and $c \in Coin$ are such that θ identifies f in the presence of c. Let i be least such that there is $\beta \in CoinSeq$ of length $i+1$ such that for all $j \geq i$ there is $\alpha \in CoinSeq$ of length $j+1$ such that $\theta(f[j]\alpha) = \theta(f[i]\beta)$ and $\theta(f[i]\beta)$ is an index for a subset of f. Such an i exists since θ identifies f in the presence of c. Also, $\theta(f[i]\beta)$ is necessarily an index for an infinite subset of f since $\theta \in PRIG$. Fix the lexicographically least β with this property. Now it is apparent that M^{θ} converges on f to the index $\theta(f[i]\beta)$. ■

(62) COROLLARY: There is an improvement operator J such that for all $p > 0$, J p–improves *PRIG*.

PROOF OF PROPOSITION (54)b: Suppose that $\theta \in PRIG$ identifies $\mathcal{SCI}$ with probability $p > 0$. Then for the improvement operator J of the foregoing proposition, $J(\theta)$ rigorously identifies *SI* contradicting Proposition (13). ■

7 Related topics

The present section is devoted to a variety of topics that bear on uniform improvement of scientific competence.

7.1 Strategy improvement

Can nonrigorous scientists be mechanically transformed into rigorous ones? We consider this question in the larger context of "strategy improvement."

(63) DEFINITION:

(a) A *strategy* is any subset of $\mathcal{SCI}$.

(b) A strategy $X \subseteq \mathcal{SCI}$ is *restrictive* just in case there is $\theta \in \mathcal{SCI}$ such that for all $\Omega \in X$, not $Sc[\theta] \subseteq Sc[\Omega]$.

Thus, $\mathcal{RIG}$ is a strategy. By Proposition (13), $\mathcal{RIG}$ is restrictive. Note that we have returned to the nonprobabilistic conception of identification, introduced in Section 2.

(64) DEFINITION: Let $\theta \in \mathcal{SCI}$, $\Sigma \subseteq \mathcal{SCI}$, strategy X, and improvement operator J be given.

(a) *J improves θ to X* just in case $J(\theta) \in X$ and $Sc[\theta] \subseteq Sc[J(\theta)]$.

(b) *J improves Σ to X* just in case J improves every $\theta \in \Sigma$ *to* X.

(c) Σ is *improvable to X* just in case some improvement operator improves Σ to X.

It follows from our definitions that if strategy X is restrictive, then $\mathcal{SCI}$ is not improvable to X. Now consider $S = \{\theta \in \mathcal{SCI} \mid \#(Sc[\theta]) \leq 1\}$. For every $\theta \in S$ there is $\Omega \in \mathcal{RIG}$ such that $Sc[\theta] \subseteq Sc[\Omega]$. Nonetheless:

(65) PROPOSITION: $\{\theta \in \mathcal{SCI} \mid \#(Sc[\theta]) \leq 1\}$ is not improvable to $\mathcal{RIG}$.

PROOF: For a contradiction, suppose that improvement operator J improves $S = \{\theta \in \mathcal{SCI} \mid \#(Sc[\theta] \leq 1\}$ to $\mathcal{RIG}$. Then, $T = \{J(\theta) \mid \theta \in S\} \subseteq \mathcal{RIG}$. So, Corollary (33) implies that some improvement operator I *ZOS*-improves T. Thus, $I \circ J$ *ZOS*-improves S, contradicting Proposition (28). ■

Variants of the following strategies have been studied by other investigators.

(66) DEFINITION: Let $\theta \in \mathcal{SCI}$ be given.

(a) (Case & Ngo-Manguelle, 1979): θ is *popperian* just in case for all $\sigma \in Seq$, $\varphi_{\theta(\sigma)} \in \mathcal{T}$. $\{\theta \in \mathcal{SCI} \mid \theta$ popperian$\}$ is denoted: *POP*.

(b) (Weihagen, 1976): θ is *consistent* just in case for all $\sigma \in Seq$, $rng(\sigma) \subseteq \varphi_{\theta(\sigma)}$. $\{\theta \in \mathcal{SCI} \mid \theta$ consistent$\}$ is denoted: *CON*.

(c) (Minicozzi, cited in Blum & Blum, 1975): θ is *reliable* just in case θ is total and for all $f \in \mathcal{T}$, if θ converges on f to $i \in N$, then φ_i is an infinite subset of f. $\{\theta \in \mathcal{SCI} \mid \theta \text{ reliable}\}$ is denoted: *REL*.

Observe that $POP \subset \mathcal{RIG}$. Indeed, $\mathcal{RIG}$ may be conceived as a generalization of *POP* intended to capture the essence of the Popperian injunction of testability (see Popper, 1972).

It may be shown that each of *POP*, *CON*, and *REL* are restrictive. This is obvious in the case of *POP*; for *CON* and *REL*, see Weihagen (1976) and Blum & Blum (1975), respectively. On the other hand, for any finite $G \subseteq \mathcal{T}$ it is easy to specify $\theta \in POP \bigcap CON \bigcap REL$ such that θ identifies G. Parallel to Proposition (65) we have:

(67) Proposition: $\{\theta \in \mathcal{SCI} \mid \#(Sc[\theta]) \leq 1\}$ is improvable to neither *POP*, *CON*, nor *REL*.

The proof of (67) uses the same technique as the proof of Proposition (29), and is omitted.

Each of the strategies considered so far is restrictive, so we might ask whether there are nonrestrictive strategies for which improvement is similarly limited. The following proposition provides an affirmative answer to this question.

(68) Proposition: There is $X \subseteq \mathcal{SCI}$ such that:

(a) X is not restrictive; but

(b) $\mathcal{SCI}$ is not improvable to X.

Proof: Given $\theta \in \mathcal{SCI}$ and $m \in N$, let $\theta^m \in \mathcal{SCI}$ be defined as follows. For all $\sigma \in Seq$:

$$\theta^m(\sigma) = \begin{cases} m & \text{if } lh(\sigma) = 1; \\ \theta(\sigma) & \text{otherwise.} \end{cases}$$

Call $X \subseteq \mathcal{SCI}$ a "choice strategy" just in case for every $\theta \in \mathcal{SCI}$ there is $m \in N$ such that $\theta^m \in X$. Observe that every choice strategy is nonrestrictive.

By Property (24), let $J_0, J_1, \ldots$ be an enumeration of all improvement operators J such that $J(\mathcal{SCI})$ is a choice strategy. Let $\theta_0, \theta_1, \ldots$ be an enumeration without repetitions of $\mathcal{SCI}$. For all $i \in N$, let $s(i)$ be the least $m \in N$ such that $\theta_i^m \in J_i(\mathcal{SCI})$. Such an m must exist because $J_i(\mathcal{SCI})$ is a choice strategy. Let choice strategy $X = \{\theta_i^{s(i)+1} \mid i \in N\}$. Then for all $i \in N$, $\theta_i^{s(i)} \in J_i(\mathcal{SCI}) - X$, hence $J_i(\mathcal{SCI}) \not\subseteq X$. Consequently, $\mathcal{SCI}$ is not improvable to X. ■

7.2 Improvement via programs

Dropping the computability requirement on scientists yields the class $\mathcal{SCI}^*$ of all functions from *Seq* to N. Since improvement operators are just recursive operators in the sense of Rogers (1967, Section 9.8), they are defined on all of $\mathcal{SCI}^*$. As a consequence, some of our results carry over to the broader context of $\mathcal{SCI}^*$, notably Proposition (59). This feature of the present approach is related to the fact noted in Section 4 that improvement operators apply to computable scientists without having access to their programs.

In the context of artificial systems of scientific inquiry such programs are accessible, and we are led to consider a different form of algorithmic improvement. We illustrate with scope-improvement.

(69) DEFINITION: Let $G \subseteq \mathcal{T}$, $i \in N$, $S \subseteq N$, and partial recursive function δ be given.

(a) δ *G–improves* i just in case $\delta(i)$ is defined, and $G \bigcup Sc[\varphi_i] \subseteq Sc[\varphi_{\delta(i)}]$.

(b) δ *G–improves* S just in case δ G–improves every $i \in S$.

In the definition, i should be conceived as a program for the scientist φ_i.

Does access to programs facilitate scope-improvement? The following proposition provides an affirmative answer to this question.

(70) PROPOSITION: There is $S \subseteq N$ such that:

(a) some partial recursive function δ *ZOS*-improves S; but

(b) no improvement operator *ZOS*-improves $\{\varphi_i \mid i \in S\}$.

PROOF: Let p be a total recursive, one-one "padding" function. In particular, for all $x, y \in N$, $\varphi_{p(x,y)} = \varphi_x$ and $p(x, y+1) > p(x,y)$. Let S be the set of all numbers of form

$$\begin{array}{ll} p(x, i+1), & \text{where } Sc[\varphi_x] = \{\varphi_i\}; \textit{or} \\ p(x, 0), & \text{where } Sc[\varphi_x] = \emptyset. \end{array}$$

Then $\{\varphi_j \mid j \in S\} = \{\theta \in \mathcal{SCI} \mid \#(Sc[\theta]) \leq 1\}$. Consequently, by Proposition (28), S satisfies clause (b).

Regarding clause (a), observe first that $p(x, i+1) \in S$ implies $\varphi_i \in \mathcal{T}$. Now let partial recursive function δ be such that for all $z \in N$, $\sigma \in Seq$:

$$\varphi_{\delta(z)}(\sigma) = \begin{cases} \Gamma(\sigma) & \text{if } z \text{ is of form } p(x,0); \\ i & \text{if } z \text{ is of form } p(x, i+1) \text{ and } (\exists v)(rng(\sigma) \subseteq \varphi_{i,v}); \\ \Gamma(\sigma) & \text{if } z \text{ is of form } p(x, i+1) \text{ and} \\ & (\exists v, a, b)((a,b) \in \varphi_{i,v} \;\&\; a < lh(\sigma) \;\&\; (a,b) \notin rng(\sigma)); \\ \text{undefined} & \text{otherwise.} \end{cases}$$

It is easy to verify that δ satisfies clause (a). ■

Access to programs also facilitates strategy improvement. Let $\{\theta \in \mathcal{SCI} \mid \theta \text{ total}\} = TOT$. Then Proposition (4) may be read this way: There is total recursive δ that improves N to TOT, that is, for all $i \in N$, $\varphi_{\delta(i)} \in TOT$ and $Sc[\varphi_i] \subseteq Sc[\varphi_{\delta(i)}]$. In contrast, we have:

(71) PROPOSITION: $\mathcal{SCI}$ is not improvable to TOT.

PROOF: For a contradiction, suppose that improvement operator J improves $\mathcal{SCI}$ to TOT. Let $\Omega \in \mathcal{SCI}$ identify *SI*. Then $J(\Gamma)$ identifies *ZOS* and $J(\Omega)$ identifies *SI*. So by Proposition (7):

(72) $J(\Gamma) \neq J(\Omega)$.

Since $\emptyset \subseteq \Gamma$ and $\emptyset \subseteq \Omega$, Property (23) implies:

(73) $J(\emptyset) \subseteq J(\Gamma)$ and $J(\emptyset) \subseteq J(\Omega)$.

However, $J(\emptyset)$ is total, so (73) implies:

(74) $J(\emptyset) = J(\Gamma)$ and $J(\emptyset) = J(\Omega)$,

which contradicts (72). ■

We note that several of our results, e.g., Proposition (34) carry over to the case of improvement via programs.

7.3 Improvement via hypotheses

A natural class of improvement operators may be defined as follows.

(75) DEFINITION: Let $G \subseteq \mathcal{T}$, $\theta \in \mathcal{SCI}$, $\Sigma \subseteq \mathcal{SCI}$, and partial recursive $\Omega : N \times Seq \rightarrow N$ be given.

- (a) Ω *$G-$improves θ via hypotheses* just in case $G \bigcup Sc[\theta] \subseteq Sc[\lambda\sigma.\Omega(\theta(\sigma), \sigma)]$.
- (b) Ω *$G-$improves Σ via hypotheses* just in case Ω $G-$improves via hypotheses every $\theta \in \Sigma$.
- (c) Σ is *$G-$improvable via hypotheses* just in case some partial recursive function $G-$improves Σ via hypotheses.

It is easy to see that Ω in the above definition determines an improvement operator in the sense of Section 4. On the other hand, a trivial construction reveals that there are improvement operators J such that for all partial recursive $\Omega : N \times Seq \rightarrow N, J \neq \lambda\theta\sigma.\Omega(\theta(\sigma), \sigma)$.

Improvement via hypotheses provides a rough model of scientific cooperation. For, the function Ω may be conceived as a scientist who collaborates with colleague θ, benefiting from the theories announced by θ in response to accumulating data. Cooperation between Ω and θ is successful if $\lambda\sigma.\Omega(\theta(\sigma), \sigma)$ identifies the functions proper to the competencies of both Ω and θ (G and $Sc[\theta]$, respectively).

Does improvement via hypotheses constitute a normal form for scope-improvement? The following proposition provides a negative answer to this question.

(76) PROPOSITION: There is $G \subseteq \mathcal{T}$ and $\Sigma \subseteq \mathcal{SCI}$ such that:

- (a) Σ is $G-$improvable; but
- (b) Σ is not $G-$improvable via hypotheses.

PROOF: We rely on the following notation. Let $f \in \mathcal{T}$ and $G \subseteq \mathcal{T}$ be given. Then $f^0 =_{\text{def}} \{(x, \langle 0, y\rangle) \mid (x, y) \in f\}$ and $G^0 =_{\text{def}} \{f^0 \mid f \in G\}$. $Seq^0 =_{\text{def}} \{\sigma \in Seq \mid$ every member of $rng(\sigma)$ has the form $(x, \langle 0, y\rangle)\}$. Let total recursive d be as defined in the proof of Proposition (28). Fix total recursive function h such that for all $i \in N$, $\sigma \in Seq$:

$$\varphi_{h(i)}(\sigma) = \begin{cases} d(i) & \text{if } \sigma \in Seq^0; \\ \text{undefined} & \text{otherwise.} \end{cases}$$

Let $\Lambda = \{\varphi_{h(i)} \mid i \in N\}$. An easy adaptation of the proof of Proposition (28) yields:

(77) Λ is not ZOS^0-improvable.

Now we "complete" the $\varphi_{h(i)}$ in a way that depends on whether $\varphi_{d(i)} \in \mathcal{T}$ or $\varphi_{d(i)}$ is finite. Let e be an index for $\emptyset$ that is not in the range of d, and let total recursive h' be such that for all $i \in N$:

$$\varphi_{h'(i)} = \begin{cases} \varphi_{h(i)} \bigcup \{(\sigma, d(i)) \mid \sigma \in Seq - Seq^0\} & \text{if } \varphi_{d(i)} \in \mathcal{T} \\ \varphi_{h(i)} \bigcup \{(\sigma, e) \mid \sigma \in Seq - Seq^0\} & \text{if } \varphi_{d(i)} \text{ is finite.} \end{cases}$$

To witness Proposition (76), let $G = ZOS^0$ and let $\Sigma = \{\varphi_{h'(i)} \mid i \in N\}$.

To verify clause (b) of the proposition, suppose for a contradiction that Ω ZOS^0-improves Σ by hypotheses. We will show that Ω ZOS^0-improves Λ by hypotheses, hence that Λ is ZOS^0-improvable, contradicting (77). Since Ω ZOS^0-improves Σ by hypotheses, for all $\varphi_{h'(i)} \in \Sigma$ and all $f \in ZOS^0 \bigcup Sc[\varphi_{h'(i)}]$, $\lambda\sigma.\Omega(\varphi_{h'(i)}(\sigma), \sigma)$ identifies f. Since $Sc[\varphi_{h(i)}] \subseteq Sc[\varphi_{h'(i)}]$, for all $\varphi_{h'(i)} \in \Sigma$ and all $f \in ZOS^0 \bigcup Sc[\varphi_{h(i)}]$, $\lambda\sigma.\Omega(\varphi_{h'(i)}(\sigma), \sigma)$ identifies f. Now $\varphi_{h(i)}$ is the restriction of $\varphi_{h'(i)}$ to Seq^0, so $\lambda\sigma.\Omega(\varphi_{h(i)}(\sigma), \sigma)$ is the restriction to Seq^0 of $\lambda\sigma.\Omega(\varphi_{h'(i)}(\sigma), \sigma)$. Moreover, for all $f \in ZOS^0 \bigcup Sc[\varphi_{h(i)}]$, $f[i] \in Seq^0$ for all $i \in N$. Hence $\lambda\sigma.\Omega(\varphi_{h(i)}(\sigma), \sigma)$ identifies f.

Regarding clause (a), let $\Gamma^0 \in \mathcal{SCI}$ be the obvious function for identifying ZOS^0. Let oracle machine M work as follows when equipped with total $\theta \in \mathcal{SCI}$. Given $\sigma \in Seq$,

$$M^\theta(\sigma) = \begin{cases} \theta(\sigma) & \text{if } \sigma \notin Seq^0; \\ \Gamma^0(\sigma) & \text{if } \theta((0, \langle 1, 1\rangle)) = e; \\ d(i) & \text{if } \theta((0, \langle 1, 1\rangle)) = d(i) \text{ and } (\exists v)(rng(\sigma) \subseteq \varphi_{i,v}); \\ \Gamma^0(\sigma) & \text{if } \theta((0, \langle 1, 1\rangle)) = d(i) \text{ and} \\ & (\exists v, a, b)((a, b) \in \varphi_{i,v} \& a < lh(\sigma) \ \& \ (a, b) \notin rng(\sigma)) \\ \text{undefined} & \text{otherwise.} \end{cases}$$

It may be verified that for all $\theta \in \Sigma$, M^θ identifies $ZOS^0 \bigcup Sc[\theta]$. Hence, by Property (21), Σ is ZOS^0–improvable. ■

8 Concluding remarks

Senses of "improvement" alternative to those studied here come readily to mind. For example, algorithms may be sought that render scientists more resistant to noisy data or more efficient in data use. Alternatively, an improvement operator might strengthen the criterion of identification to which a scientist conforms, for example, by requiring convergence on $f \in \mathcal{T}$ to a program for (all of) f. Each sense of "improvement" may be examined in a fashion parallel to the preceding developments.

9 References

Angluin, D. & Smith, C. 1983. A survey of inductive inference: Theory and methods, *Computing Surveys.*

Barzdins, J. 1974. Two theorems on the limiting synthesis of functions, *Latvian State University Technical Report.*

Blum, L. & Blum, M. 1975. Toward a mathematical theory of inductive inference. *Information & Control* 28:125-155.

Case, J. & Ngo-Manguelle, S. 1979. Refinements of inductive inference by Popperian machines. Technical Report, Department of Computer Science, SUNY Buffalo.

Case, J. & Smith, C. 1983. Comparison of identification criteria for machine inductive inference, *Theoretical Computer Science*, 25:193-220.

Fulk, M. 1988. Saving the phenomena: Requirements that inductive inference machines not contradict known data, *Information & Computation* 79:193-209.

Gold, E. M. 1967. Language identification in the limit, *Information & Control* 10:447-474.

Machtey, M. & Young, P. 1978. *An Introduction to the General Theory of Algorithms.* New York: North-Holland.

Osherson, D., Stob, M. & Weinstein, S. 1986. *Systems that Learn.* Cambridge: MIT Press.

Osherson, D., Stob, M. & Weinstein, S., 1987. Modular learning, in Eric L. Schwartz (ed.) *Computational Neuroscience*, MIT Press.

Pitt, L. & Smith, C. 1988 Probability and plurality for aggregations of learning machines *Information & Computation* 77:77-92.

Pitt, L. 1985. *Probabilistic Inductive Inference.* Dissertation, Yale University.

Popper, K. 1972. *Objective Knowledge.* Oxford: Oxford University Press.

Putnam, H. 1975. Probability and confirmation, in *Mathematics, Matter and Method*, Cambridge: Cambridge University Press.

Rogers, H. 1967. *Theory of Recursive Functions and Effective Computability.* New York: McGraw-Hill.

Solomonoff, R. J. 1964 A formal theory of inductive inference, *Information & Control*, 7:1-22, 224-254.

Wiehagen, R. 1976. Limeserkennung Rekursiver Funktionen Durch Spezielle Stratgien *Elektronische Informationsverarbeitung und Kybernetik* 12:93-99.

Wiehagen, R., Freivald, R. & Kinber, E. 1984. On the power of probabilistic strategies in inductive inference. *Theoretical Computer Science*, 28:111-133.