

Aggregating Inductive Expertise*

DANIEL N. OSHERSON

MIT, Cambridge, Massachusetts 02139

MICHAEL STOB

Calvin College, Grand Rapids, Michigan 49506

AND

SCOTT WEINSTEIN

MIT, Cambridge, Massachusetts 02139

The aggregation problem is to design an inferential agent that makes intelligent use of the theories offered by a team of inductive inference machines working in a common environment. The present paper formulates several versions of the aggregation problem and investigates them from a recursion theoretic point of view. © 1986 Academic Press, Inc.

1. INTRODUCTION

Imagine that you have been appointed the director of a laboratory comprising several research teams. Each team consists of three scientists who examine data emanating from an unknown physical source. Different teams work on different problems of this nature. Each scientist elaborates a theory of the source underlying the data he receives, and he communicates the theory to you without consulting other team members; thus, you receive three theories from each team. This process of theory elaboration and communication is repeated indefinitely as more and more data become available.

Call a scientist "successful" in this setting just in case his successive conjectures eventually stabilize to a correct theory of his data source. Call a given team "successful" just in case a majority (i.e., at least two) of its members are successful. The successful members of a successful team need

* The research reported herein was supported by Grant SDF 590 B from the System Development Foundation, and by National Science Foundation Grant MCS 82-00032. We thank an anonymous referee for helpful comments.

not begin stabilization at the same moment, nor need they stabilize on identical axiomatizations of the same theory.

Each time a given team T presents you with its three, independently elaborated theories, you must formulate your own theory of the data source in question. For this purpose you may make whatever use you please of the theories communicated to you as well as of the data available to T . You are said to “aggregate” T just in case the following conditional is true: if T is successful then, likewise, your successive conjecture stabilize to an accurate theory of T 's data source. This latter stabilization need not be synchronous with that of T . Your job is to aggregate all the teams in your laboratory. The present paper investigates the prospects for success in this aggregation task.

The aggregation task is nontrivial for the following reason. Successful members of a team may stabilize to different formulations of the same theory. Since the problem of recognizing equivalent theories is itself nontrivial, it may not always be possible to identify the majority in a successful team.

We now recast the aggregation problem in the context of machine inductive inference. By an “inductive inference machine” is meant any computational agent that examines progressively larger data sets drawn from an unknown environment and emits, in response, a succession of hypotheses about the nature of that environment. The theory of inductive inference attempts to characterize the conditions under which the successive hypotheses of an inductive inference machine stabilize to an accurate theory of its environment. Such stabilization is called “identification (in the limit)” of the environment in question. Alternative formulations of the intuitive concepts “environment,” “stabilization,” “accurate theory,” etc. give rise to distinct inductive paradigms with distinct formal properties. Fundamental paradigms are introduced and examined in Gold (1967), Blum and Blum (1975), and Case and Smith (1983). Surveys of the field are available in Angluin and Smith (1983) and in Osherson, Stob, and Weinstein (1986).

Now consider a class E of potential environments. One typically attempts to construct a single inductive inference machine that identifies every member of E . However, in some circumstances it may be more feasible to construct a finite set S of machines with the property that for every environment $e \in E$ a majority of machines in S identifies e (where in general different machines participate in majorities proper to different members of E). In this case we say that S “majority identifies” E .

If in some environment majority identification is more feasible than identification, this suggests a need for an algorithm to aggregate the successive conjectures of an arbitrarily selected team of inductive inference machines. Such an algorithm would be required to identify any environment that is

majority-identified by any team of inductive inference machines attached to it. The existence of an aggregation algorithm of this kind would amount to the factorization of inductive inference problems into two pieces, one specific the other general. The specific piece is the construction of a team that majority-identifies the class of environments in question. The general piece is the universal aggregation algorithm applicable to any specific team. Naturally, aggregation algorithms somewhat less than universal are still of potential interest from this factorization point of view.

The aggregation problem discussed here is only indirectly related to “team-identification,” studied in Smith (1982). Aggregation per se is not required for successful team identification, nor do majoritarian considerations arise.

In the sections that follow we formulate and investigate the aggregation problem in recursion theoretic terms. Our discussion begins with formal preliminaries.

2. PRELIMINARIES

Notation and basic terminology are drawn, insofar as possible, from Osherson, Stob, and Weinstein (1986).

We fix an acceptable indexing $\varphi_0, \varphi_1, \dots$ of the partial recursive functions and associate indices accordingly (for details, see Machtey and Young, 1978). The corresponding indexing of the r.e. sets—viz., the domain of the partial recursive functions—is given as W_0, W_1, \dots . We let e be an index for \emptyset . The class $\{W_i | i \in N\}$ of all r.e. sets is denoted: RE. Members of RE are usually referred to as “languages,” and designated by: L, L' , etc. The set $\{\varphi_i | i \in N\}$ of all partial recursive functions is denoted: \mathcal{F}^{rec} . In the theory developed below, inductive inference machines are represented by indices of partial recursive functions; the hypotheses emitted by these machines take the form of indices for languages.

The set $0, 1, 2, \dots$ of natural numbers is denoted: N . We let $\# \notin N$ be a blank symbol. Let $L \in \text{RE}$ be given. A *text for L* is an ω -sequence in $L \cup \{\#\}$ and on L , that is, an infinite listing of all members of L , repetitions and blanks allowed, with no members of \bar{L} in the list. The class of all texts for L is denoted \mathcal{T}_L . Thus, if L is nonempty, \mathcal{T}_L is non-denumerable. Given $\mathcal{L} \subseteq \text{RE}$, the class $\bigcup_{L \in \mathcal{L}} \mathcal{T}_L$ is denoted: $\mathcal{T}_{\mathcal{L}}$. The set of numbers appearing in a text t is denoted: $\text{rng}(t)$. Thus, for all $t \in \mathcal{T}_{\text{RE}}$, $\# \notin \text{rng}(t) \in \text{RE}$. Texts represent the environments in which inductive inference machines work.

We let “ $\langle \cdot, \cdot \rangle$ ” code pairs as single integers. $L \in \text{RE}$ is said to *represent* the set $\{(x, y) | \langle x, y \rangle \in L\}$. The class $\{L \in \text{RE} | L \text{ represents a total function}\}$ is denoted: RE_{svt} (“svt” stands for “single-valued, total”). Thus, RE_{svt} represents the set of graphs of total recursive functions. Whereas it is

usual to conceive of inductive inference machines as operating directly on such graphs, it shall here be assumed that graphs are first coded as sets of (single) natural numbers. This will allow uniform treatment of language-identification and function-identification.

Let $t \in \mathcal{T}_{\text{RE}}$ and $n \in \mathbb{N}$ be given. The n th member of t is denoted: t_n . The finite initial sequence of length n in t is denoted \bar{t}_n . The set $\{\bar{t}_m \mid t \in \mathcal{T}_{\text{RE}} \text{ and } m \in \mathbb{N}\}$ of all finite initial sequences in any text is denoted: SEQ. Members of SEQ are often designated by “ σ ,” “ τ ,” etc., and may be thought of as finite, “evidential states.” The length of σ is denoted: $\text{lh}(\sigma)$. For $\sigma \in \text{SEQ}$, and $m < \text{lh}(\sigma)$, “ $\text{rng}(\sigma)$,” “ σ_m ” and “ $\bar{\sigma}_m$ ” are interpreted just as for text. (Notice that $\sigma_{\text{lh}(\sigma)}$ does not exist whereas $\bar{\sigma}_{\text{lh}(\sigma)} = \sigma$.)

We assume the existence of a fixed, recursive isomorphism between SEQ and \mathbb{N} . Tacit application of this isomorphism allows partial recursive functions to apply directly to sequences, yielding single natural numbers as outputs. A “course-of-values” notation will also be useful. Given $\psi \in \mathcal{F}^{\text{rec}}$ and $\sigma \in \text{SEQ}$, we define

$$\tilde{\psi}(\sigma) = (\psi(\bar{\sigma}_0), \dots, \psi(\bar{\sigma}_{\text{lh}(\sigma)})),$$

if each of $\psi(\bar{\sigma}_0), \dots, \psi(\bar{\sigma}_{\text{lh}(\sigma)})$ is defined:

$$\tilde{\psi}(\sigma) = \text{undefined} \quad \text{otherwise.}$$

Let $t \in \mathcal{T}_{\text{RE}}$, $j \in \mathbb{N}$, and $\psi \in \mathcal{F}^{\text{rec}}$ be given. ψ is said to be *defined on* t just in case for all $m \in \mathbb{N}$, $\psi(\bar{t}_m) \downarrow$. ψ is said to *converge on* t to j just in case (a) ψ is defined on t , and (b) for all but finitely many $m \in \mathbb{N}$, $\psi(\bar{t}_m) = j$. We now define identification. Intuitively, ψ identifies t just in case ψ converges to an index that “accurately represents” $\text{rng}(t)$. In order to formalize the latter concept, let any subset of $\text{RE} \times \mathbb{N}$ (the Cartesian product of RE and \mathbb{N}) be called an *accuracy criterion*. An accuracy criterion is to be conceived as a pairing of r.e. sets L and indices i such that W_i counts as being “close” to L . The two accuracy criteria of primary importance for present purposes are called “INT” and “FINT”, defined as follows:

$$\text{INT} = \{(L, i) \mid W_i = L\}.$$

$$\text{FINT} = \{(L, i) \mid (W_i - L) \cup (L - W_i) \text{ is finite}\}.$$

Thus, the INT criterion demands perfect accuracy since it pairs languages with their indices. FINT allows a finite margin of error. (For the significance of the “INT” and “FINT” terminology, see Osherson, Stob, and Weinstein, 1986, Chap. 6.)

Now let $t \in \mathcal{T}_{\text{RE}}$, accuracy criterion \mathcal{C} , and $\psi \in \mathcal{F}^{\text{rec}}$ be given. ψ *\mathcal{C} -identifies* t just in case there is $i \in \mathbb{N}$ such that (1) ψ converges on t to i , and (b) $(\text{rng}(t), i) \in \mathcal{C}$. ψ *\mathcal{C} -identifies* $\mathcal{E} \subseteq \mathcal{T}_{\text{RE}}$ just in case ψ *\mathcal{C} -identifies* every $t \in \mathcal{E}$.

3. BASIC AGGREGATION PARADIGMS

The aggregation problem discussed in Section 1 may be formalized in alternative ways depending on the kind of information made available to the aggregator. The two definitions that follow underlie all the paradigms to be discussed.

DEFINITION 3. A. Let $t \in \mathcal{T}_{\text{RE}}$, $m \in N$, $(i_1, \dots, i_m) \in N^m$, and accuracy criterion \mathcal{C} be given. (i_1, \dots, i_m) \mathcal{C} -identifies t just in case $\varphi_{i_1}, \dots, \varphi_{i_m}$ are all defined on t , and a majority (i.e., more than $m/2$) of $\varphi_{i_1}, \dots, \varphi_{i_m}$ \mathcal{C} -identify t . (i_1, \dots, i_m) \mathcal{C} -identifies $\mathcal{E} \subseteq \mathcal{T}_{\text{RE}}$ just in case (i_1, \dots, i_m) \mathcal{C} -identifies every $t \in \mathcal{E}$.

In Definition 3.A. it is intended that two occurrences of the same index in a team be counted twice when reckoning majorities. The requirement that $\varphi_{i_1}, \dots, \varphi_{i_m}$ all be defined on t simplifies our exposition but is not essential to later results.

DEFINITION 3.B. Let $m \in N$, $(i_1, \dots, i_m) \in N^m$, $\mathcal{E} \subseteq \mathcal{T}_{\text{RE}}$, $\psi \in \mathcal{F}^{\text{rec}}$ and accuracy criterion \mathcal{C} be given. ψ \mathcal{C} -aggregates (i_1, \dots, i_m) on \mathcal{E} just in case for all $t \in \mathcal{E}$, if (i_1, \dots, i_m) \mathcal{C} -identifies t then ψ \mathcal{C} -identifies t .

Alternative paradigms of aggregation arise from alternative analyses of the role of ψ in the foregoing definition. The subset \mathcal{E} of Definition 3.B will in practice be either \mathcal{T}_{RE} itself or $\mathcal{T}_{\text{RE}_{\text{svt}}}$. \mathcal{C} will be either INT or FINT.

The first form of aggregation to be considered provides the aggregator with maximum information: he has access to the programs of the team he must aggregate as well as to the data giving rise to their conjectures.

DEFINITION 3.C. Let $m \in N$, $\mathcal{E} \subseteq \mathcal{T}_{\text{RE}}$, and accuracy criterion \mathcal{C} be given. Then: $[\text{program}, \mathcal{C}, \mathcal{E}]^m = \{B \subseteq N^m \mid \text{there is computable } \Theta: N^m \times \text{SEQ} \rightarrow N \text{ such that for all } (i_1, \dots, i_m) \in B, \lambda\sigma.\Theta(i_1, \dots, i_m, \sigma) \mathcal{C}\text{-aggregates } (i_1, \dots, i_m) \text{ on } \mathcal{E}\}$.

We illustrate with $[\text{program}, \text{FINT}, \mathcal{T}_{\text{RE}_{\text{svt}}}]^3$. A set $B \subseteq N^3$ is a member of this collection just in case there is a computable function Θ , taking as arguments a triple of indices and a finite sequence, with the following property: for every triple (i, j, k) drawn from B and every text t representing the graph of a total recursive function, if (i, j, k) FINT-identifies t then $\lambda\sigma.\Theta(i, j, k, \sigma)$ also FINT-identifies t .

Note that the collections $[\text{program}, \mathcal{C}, \mathcal{E}]^m$ are closed downward under inclusion. The same will be true for all our aggregation paradigms.

The remaining forms of aggregation provide the aggregator with less information than program aggregation. In the following paradigm

aggregators have access to team members in the form of black boxes, with no information about their internal programs; the aggregator may also examine incoming text.

DEFINITION 3.D. Let $m \in N$, $\mathcal{E} \subseteq \mathcal{T}_{\text{RE}}$, and accuracy criterion \mathcal{C} be given. Then: $[\text{blackbox}, \mathcal{C}, \mathcal{E}]^m = \{B \subseteq N^m \mid \text{there is computable } \Theta: \mathcal{F}^{\text{rec}^m} \times \text{SEQ} \rightarrow N \text{ such that for all } (i_1, \dots, i_m) \in B, \lambda\sigma.\Theta(\varphi_{i_1}, \dots, \varphi_{i_m}, \sigma) \mathcal{C}\text{-aggregates } (i_1, \dots, i_m) \text{ on } \mathcal{E}\}$.

The function Θ invoked in the foregoing definition is a computable functional inasmuch as it can be computed by an ‘‘oracle machine’’ (Rogers, 1967, Chap. 9) with the graphs of $\varphi_{i_1}, \dots, \varphi_{i_m}$ as oracles.

The following paradigm limits aggregators to incoming text plus the hypotheses emitted by its associated team.

DEFINITION 3.E. Let $m \in N$, $\mathcal{E} \subseteq \mathcal{T}_{\text{RE}}$, and accuracy criterion \mathcal{C} be given. Then: $[\text{Hypothesis}, \mathcal{C}, \mathcal{E}]^m = \{B \subseteq N^m \mid \text{there is computable } \Theta: \text{SEQ}^m \times \text{SEQ} \rightarrow N \text{ such that for all } (i_1, \dots, i_m) \in B, \lambda\sigma.\Theta(\tilde{\varphi}_{i_1}(\sigma), \dots, \tilde{\varphi}_{i_m}(\sigma), \sigma) \mathcal{C}\text{-aggregates } (i_1, \dots, i_m) \text{ on } \mathcal{E}\}$.

Note that Definition 3.E allows aggregators access to the record of past hypotheses of team members; attention need not be limited to most recent conjectures.

Finally, we consider a ‘‘pure’’ form of aggregation in which incoming text cannot be directly examined.

DEFINITION 3.F. Let $m \in N$, $\mathcal{E} \subseteq \mathcal{T}_{\text{RE}}$, and accuracy criterion \mathcal{C} be given. Then: $[\text{pure}, \mathcal{C}, \mathcal{E}]^m = \{B \subseteq N^m \mid \text{there is computable } \Theta: \text{SEQ}^m \rightarrow N \text{ such that for all } (i_1, \dots, i_m) \in B, \lambda\sigma.\Theta(\tilde{\varphi}_{i_1}(\sigma), \dots, \tilde{\varphi}_{i_m}(\sigma)) \mathcal{C}\text{-aggregates } (i_1, \dots, i_m) \text{ on } \mathcal{E}\}$.

The foregoing definitions immediately yield such inclusions as: $[\text{pure}, \text{INT}, \mathcal{T}_{\text{RE}}]^m \subseteq [\text{hypothesis}, \text{INT}, \mathcal{T}_{\text{RE}}]^m \subseteq [\text{blackbox}, \text{INT}, \mathcal{T}_{\text{RE}}]^m \subseteq [\text{program}, \text{INT}, \mathcal{T}_{\text{RE}}]^m$, for all $m \in N$. We next examine some less obvious properties of the collections defined in this section.

4. BASIC RESULTS ABOUT AGGREGATION

Even in the best of circumstances, aggregation on \mathcal{T}_{RE} cannot be performed over all teams of inductive inference machines. This is the content of our first proposition.

PROPOSITION 4.A. For all odd $m \geq 3$:

- (i) $N^m \notin [\text{program}, \text{INT}, \mathcal{T}_{\text{RE}}]^m$.
- (ii) $N^m \notin [\text{program}, \text{FINT}, \mathcal{T}_{\text{RE}}]^m$.

The proof of Proposition 4.A relies on the following definition and lemma.

DEFINITION 4.A. Let $L \in \text{RE}$ be given:

- (i) The cardinality of $\{x \in L \mid \text{for some } y \in N, x = \langle 0, y \rangle\}$ is denoted: $\text{par}(L)$.
- (ii) L is *parity self-describing* just in case (a) there are $n, m \in N$ such that $\langle 1, n \rangle, \langle 2, m \rangle \in L$, and (b) either $\text{par}(L)$ is odd and $W_i = L$ or $\text{par}(L)$ is even and $W_j = L$ or $\text{par}(L)$ is infinite and $W_i = W_j = L$, where i is the least n such that $\langle 1, n \rangle \in L$ and j is the least m such that $\langle 2, m \rangle \in L$.
- (iii) $\{L \in \text{RE} \mid L \text{ is parity self-describing}\}$ is denoted: RE_{psd} .

LEMMA 4.A.

- (i) No member of \mathcal{F}^{rec} INT-identifies $\mathcal{T}_{\text{RE}_{\text{psd}}}$.
- (ii) No member of \mathcal{F}^{rec} FINT-identifies $\mathcal{T}_{\text{RE}_{\text{psd}}}$.

Proof. Part (i) follows from part (ii). Part (ii) is demonstrated in the proof of Proposition 6.4.1.A of Osherson, Stob, and Weinstein (1986). ■

DEFINITION 4.B. For $n \in N$, $\sigma \in \text{SEQ}$, $\mathbf{n}(\sigma) = \mu c[\langle n, c \rangle \in \text{rng}(\sigma)]$ if such exists and $= e$ otherwise.

Proof of Proposition 4.A. For notational ease we prove the proposition for the case of $m = 3$. Adaptation to the general case is straightforward. A similar policy applies to later proofs.

(i) We exhibit a single triple (i, j, k) such that for all computable $\theta: N^3 \times \text{SEQ} \rightarrow N$, $\lambda \sigma. \theta(i, j, k, \sigma)$ fails to INT-aggregate (i, j, k) on \mathcal{T}_{RE} . We may take (i, j, k) to be any indices such that for all $\sigma \in \text{SEQ}$:

$$\begin{aligned} \varphi_i(\sigma) &= \mathbf{1}(\sigma). \\ \varphi_j(\sigma) &= \mathbf{2}(\sigma). \\ \varphi_k(\sigma) &= \varphi_i(\sigma) \quad \text{if } \text{par}(\text{rng}(\sigma)) \text{ is odd;} \\ &= \varphi_j(\sigma) \quad \text{if } \text{par}(\text{rng}(\sigma)) \text{ is even.} \end{aligned}$$

It is easy to see that (i, j, k) INT-identifies $\mathcal{T}_{\text{RE}_{\text{psd}}}$. On the other hand, were there $\psi \in \mathcal{F}^{\text{rec}}$ that INT-aggregated (i, j, k) on \mathcal{T}_{RE} , then ψ would INT-identify $\mathcal{T}_{\text{RE}_{\text{psd}}}$ contracting Lemma 4.A(i).

(ii) Let $(i, j, k) \in N^3$ be as defined in part (i). Then (i, j, k) FINT-identifies $\mathcal{T}_{\text{RE}_{\text{psd}}}$. Consequently, Lemma 4.A(ii) implies that no $\psi \in \mathcal{F}^{\text{rec}}$ FINT-aggregates (i, j, k) on \mathcal{T}_{RE} . ■

The foregoing proof rests upon the existence of majority-identifiable collections of languages that are not identifiable by single machines. The next proposition shows that such collections are not essential to non-aggregability.

DEFINITION 4.C. Let $m \in N$ and accuracy criterion \mathcal{C} be given. The set $\{(i_1, \dots, i_m) \in N^m \mid \text{some } \psi \in \mathcal{F}^{\text{rec}} \text{ } \mathcal{C}\text{-aggregates } (i_1, \dots, i_m) \text{ on } \mathcal{T}_{\text{RE}}\}$ is denoted: $A_{\mathcal{C}}^m$.

Thus, a team in $A_{\mathcal{C}}^m$ \mathcal{C} -identifies no more than what can be \mathcal{C} -identified by a single inductive inference machine.

PROPOSITION 4.B. For all odd $m \geq 3$,

- (i) $A_{\text{INT}}^m \notin [\text{program, INT, } \mathcal{T}_{\text{RE}}]^m$.
- (ii) $A_{\text{FINT}}^m \notin [\text{program, FINT, } \mathcal{T}_{\text{RE}}]^m$.

Proof. (i) We show that $A_{\text{INT}}^3 \notin [\text{program, INT, } \mathcal{T}_{\text{RE}}]^3$. Given $a, b \in N$ we call a triple (i, j, k) “simple_{ab}” just in case for every $\sigma \in \text{SEQ}$:

$$\begin{aligned} \varphi_i(\sigma) &= a; \\ \varphi_j(\sigma) &= b; \\ \varphi_k(\sigma) &= a \quad \text{if } \text{par}(\text{rng}(\sigma)) \text{ is odd,} \\ &= b \quad \text{if } \text{par}(\text{rng}(\sigma)) \text{ is even.} \end{aligned}$$

Let f, g, h be recursive functions such that the triple $(f(\langle a, b \rangle), g(\langle a, b \rangle), h(\langle a, b \rangle))$ is simple_{ab} for every a, b . Note that for any $a, b \in N$, the triple $(f(\langle a, b \rangle), g(\langle a, b \rangle), h(\langle a, b \rangle))$ INT-identifies at most two languages and is therefore in A_{INT}^3 . Suppose that $L \in \text{RE}_{\text{psd}}$. Then if a, b are the least integers such that $\langle 1, a \rangle \in L$ and $\langle 2, b \rangle \in L$, respectively, then the corresponding triple $(f(\langle a, b \rangle), g(\langle a, b \rangle), h(\langle a, b \rangle))$ INT-identifies L .

Now suppose that $A_{\text{INT}}^3 \in [\text{program, INT, } \mathcal{T}_{\text{RE}}]^3$ and that Θ witnesses this. We show how to construct $\psi \in \mathcal{F}^{\text{rec}}$ which INT-identifies RE_{psd} , contradicting Lemma 4.A(i). Recall Definition 4.B, and define $\psi(\sigma) = \Theta(f(\langle \mathbf{1}(\sigma), \mathbf{2}(\sigma) \rangle), g(\langle \mathbf{1}(\sigma), \mathbf{2}(\sigma) \rangle), h(\langle \mathbf{1}(\sigma), \mathbf{2}(\sigma) \rangle), \sigma)$. Now if t is a text for $L \in \text{RE}_{\text{psd}}$, there is $n \in N$ such that \bar{t}_n contains the pairs $\langle 1, a \rangle$ and $\langle 2, b \rangle$ for the least such pairs in L . Then for all $m \geq n$, $\mathbf{1}(\bar{t}_m) = a$ and $\mathbf{2}(\bar{t}_m) = b$ so that $\psi(\bar{t}_m) = \Theta(f(\langle a, b \rangle), g(\langle a, b \rangle), h(\langle a, b \rangle), \bar{t}_m)$. Since Θ aggregates $(f(\langle a, b \rangle), g(\langle a, b \rangle), h(\langle a, b \rangle))$ and $(f(\langle a, b \rangle), g(\langle a, b \rangle), h(\langle a, b \rangle))$ INT-identifies L , we have that ψ INT-identifies t .

(ii) As in the preceding proposition, (ii) follows by adapting the argument of (i). ■

Aggregation is greatly facilitated by restricting attention to languages that represent total functions. In this case the aggregator may even ignore incoming data.

PROPOSITION 4.C. *For all odd $m \geq 3$,*

- (i) $N^m \in [\text{pure}, \text{INT}, \mathcal{F}_{\text{RE}_{\text{svt}}}]^m$.
- (ii) $N^m \in [\text{pure}, \text{FINT}, \mathcal{F}_{\text{RE}_{\text{svt}}}]^m$.

Proof. (i) We will show that $N^3 \in [\text{pure}, \text{INT}, \mathcal{F}_{\text{RE}_{\text{svt}}}]^3$ by constructing $\Theta: \text{SEQ}^3 \rightarrow N$ which witnesses this. We must define Θ on triples of sequences of the form τ^0, τ^1, τ^2 for which $\text{lh}(\tau^0) = \text{lh}(\tau^1) = \text{lh}(\tau^2)$. Fix such a triple and let $n = \text{lh}(\tau^0) - 1$. (Thus, $\tau^i = (\tau^i_0, \dots, \tau^i_n)$ for $i = 0, 1, 2$.) Given a set F of indices and an integer s , we say that F is “consistent at” s (as a set of indices for elements of RE_{svt}) if whenever $i, j \in F$, $\langle x, y \rangle \in \mathcal{W}_{i,s}$, and $\langle x, z \rangle \in \mathcal{W}_{j,s}$, then $y = z$. To define $\Theta(\tau^0, \tau^1, \tau^2)$, we define a sequence G_0, \dots, G_n of two element subsets of $\{0, 1, 2\}$ as follows: $G_0 = \{0, 1\}$, and given $G_{s-1} = \{i, j\}$, $G_s = G_{s-1}$ if $\tau^i_s = \tau^i_{s-1}$, $\tau^j_s = \tau^j_{s-1}$, and $\{\tau^i_s, \tau^j_s\}$ is consistent at s . Otherwise, G_s is the next set in the sequence $\{0, 1\}, \{1, 2\}, \{0, 2\}$. Now define Θ so that $\mathcal{W}_{\Theta(\tau^0, \tau^1, \tau^2)} = \bigcup \{\mathcal{W}_{\tau^i} \mid i \in G_n\}$.

Suppose now that t is a text for $L \in \text{RE}_{\text{svt}}$ and suppose that $\{i, j, k\}$ is a team that INT-identifies t . Then, for σ long enough in t , at least two of $\tilde{\varphi}_i(\sigma)$, $\tilde{\varphi}_j(\sigma)$, $\tilde{\varphi}_k(\sigma)$ will be constant and consistent with each other. Thus, for σ large enough in t , the sets G in the definition of Θ will settle to a pair of indices which give constant and consistent conjectures. Since at least one of these conjectures must be an index for $\text{rng}(t)$ (since the team INT-identifies t), taking the union in the definition of Θ guarantees that Θ will converge to an index for $\text{rng}(t)$.

(ii) We modify the proof of (i) in the following way: In the definition of G_s we must not abandon the pair in G_{s-1} simply because the conjectures of τ^i and τ^j are inconsistent at s . Else, since the team members that FINT-identify the text can make finitely many mistakes, we may abandon the correct pair of conjectures infinitely often. The solution is to abandon a pair i, j only if the conjectures have changed or if they are seen to be inconsistent at s on more pairs than the number of times the set G has changed before s . Thus, given a text t for some $L \in \text{RE}_{\text{svt}}$, the sets G will converge to the indices for two sets which are consistent with each other for all but finitely many pairs and such that one of the two sets differs finitely from t . Hence, taking the union again guarantees that Θ will converge, this time to an index for a set which is finitely different from t . ■

The next proposition reveals the importance to aggregation of being able to examine the programs of the inductive inference machines being aggregated—or, second best, of being able to interrogate them as black boxes.

PROPOSITION 4.D. *For all odd $m \geq 3$,*

(i) [hypothesis, INT, \mathcal{T}_{RE}] $^m \subset$ [blackbox, INT, \mathcal{T}_{RE}] $^m \subset$ [program, INT, \mathcal{T}_{RE}] m .

(ii) [hypothesis, FINT, \mathcal{T}_{RE}] $^m \subset$ [blackbox, FINT, \mathcal{T}_{RE}] $^m \subset$ [program, FINT, \mathcal{T}_{RE}] m .

Proof. (i) The inclusions are obvious: we need only show that they are strict. We first establish:

(a) [blackbox, INT, \mathcal{T}_{RE}] $^m \subset$ [program, INT, \mathcal{T}_{RE}] m . As usual, we show this for $m = 3$. Recall the definition of “simple $_{ab}$ ” from the proof of Proposition 4.B. Note that if (i, j, k) is simple $_{ab}$ and INT-identifies $t \in \mathcal{T}_{\text{RE}}$, then either $W_a = W_b$ and t is a text for W_a , or $W_a \neq W_b$ and either t is for W_a and $\text{par}(\text{rng}(t))$ is odd or t is for W_b and $\text{par}(\text{rng}(t))$ is even. Let $S = \{(i, j, k) \mid \text{there are } a, b \in N \text{ such that } (i, j, k) \text{ is simple}_{ab} \text{ and such that } W_a = W_b \text{ and } i \geq j \text{ or } W_a \neq W_b \text{ and } j > i\}$. We claim that S is our desired set of triples witnessing (a). To see first that $S \in [\text{program, INT, } \mathcal{T}_{\text{RE}}]^3$, define $\Theta: N^3 \times \text{SEQ} \rightarrow N$ by

$$\begin{aligned} \Theta(i, j, k, \sigma) &= \varphi_i(\sigma) && \text{if } i \geq j, \\ &= \varphi_k(\sigma) && \text{if } j > i. \end{aligned}$$

It is clear that Θ aggregates any triple in S .

To see that $S \notin [\text{blackbox, INT, } \mathcal{T}_{\text{RE}}]^3$, suppose to the contrary that $\Theta: \mathcal{F}^{\text{rec}^3} \times \text{SEQ} \rightarrow N$ aggregates S . We will show how to define $\psi: \text{SEQ} \rightarrow N$ which INT-identifies RE_{psd} , contradicting Lemma 4.A(i). Now given σ , let (i, j, k) be the triple $(f(\langle \mathbf{1}(\sigma), \mathbf{2}(\sigma) \rangle), g(\langle \mathbf{1}(\sigma), \mathbf{2}(\sigma) \rangle), h(\langle \mathbf{1}(\sigma), \mathbf{2}(\sigma) \rangle))$ where f, g , and h are the recursive functions defined in the proof of Proposition 4.B. Now define $\psi(\sigma) = \Theta(\varphi_i, \varphi_j, \varphi_k, \sigma)$. We claim that ψ identifies RE_{psd} . To see this, given $t \in \mathcal{T}_{\text{RE}}$ for $L \in \text{RE}_{\text{psd}}$, let n be large enough so that $\mathbf{1}(\bar{t}_n) = \mu a[\langle 1, a \rangle \in L]$ and $\mathbf{2}(\bar{t}_n) = \mu b[\langle 2, b \rangle \in L]$. Let a and b be the pair $\mathbf{1}(\bar{t}_n), \mathbf{2}(\bar{t}_n)$. Then for all $m \geq n$, the triple (i, j, k) computed in the definition of $\psi(\bar{t}_m)$ is simple $_{ab}$. Now it is easy to see that given this triple i, j, k there is a triple (i', j', k') which is in S and is such that $\varphi_i = \varphi_{i'}$, $\varphi_j = \varphi_{j'}$, and $\varphi_k = \varphi_{k'}$. Hence, for all $m \geq n$, $\Theta(\varphi_i, \varphi_j, \varphi_k, \bar{t}_m) = \Theta(\varphi_{i'}, \varphi_{j'}, \varphi_{k'}, \bar{t}_m)$ and since Θ aggregates (i', j', k') on t , ψ INT-identifies t .

(b) [hypothesis, INT, \mathcal{T}_{RE}] $^m \subset$ [blackbox, INT, \mathcal{T}_{RE}] m . The proof here is similar to that of (a). In (a) the set S was defined so that the indices

of the functions in S carried information, namely, the information of whether $W_a = W_b$. In the present situation we can define S so that the behavior of φ_i on some one special sequence carries the same information.

(ii) The FINT versions of (a) and (b) may be proven by an argument parallel to that given above. ■

We have not been able to determine wheter $[\text{pure, INT, } \mathcal{F}_{\text{RE}}]^m \subset [\text{hypothesis, INT, } \mathcal{F}_{\text{RE}}]^m$ or whether $[\text{pure, FINT, } \mathcal{F}_{\text{RE}}]^m \subset [\text{hypothesis, FINT, } \mathcal{F}_{\text{RE}}]^m$.

Propositions 2.A and 2.B prompt the search for natural categories of inductive inference machines that lend themselves to aggregation. The following definition and proposition illustrate this idea.

DEFINITION 4.D (Angluin, 1980). Let $\psi \in \mathcal{F}^{\text{rec}}$ be given:

- (i) ψ is consistent just in case for all $\sigma \in \text{SEQ}$, $\text{rng}(\sigma) \subseteq W_{\psi(\sigma)}$.
- (ii) $\{i \in N \mid \varphi_i \text{ is consistent}\}$ is denoted: CON.

Thus, inductive inference machines in CON always make conjectures that generate the data seen to date.

PROPOSITION 4.E. For all odd $m \geq 3$, $\text{CON}^m \in [\text{pure, INT, } \mathcal{F}_{\text{RE}}]^m$.

Proof. Let $m=3$. We construct computable $\Theta: \text{SEQ}^3 \rightarrow N$ that witnesses the proposition. Given sequences τ^0, τ^1, τ^2 , of length $n+1$, we choose 2-element subsets G_0, G_1, \dots, G_n of $\{0, 1, 2\}$ similar to the method of the proof of Proposition 4.C. Given $G_{s-1} = \{i, j\}$, $G_s = G_{s-1}$ unless $\tau_s^i \neq \tau_{s-1}^i$ or $\tau_s^j \neq \tau_{s-1}^j$. In this case G_s is the next 2-element subset of $\{0, 1, 2\}$. Given G_s we define Θ by

$$W_{\Theta(\tau^0, \tau^1, \tau^2)} = \cap \{W_{\tau_s^i} \mid i \in G_s\}.$$

Given a text t and $(i, j, k) \in \text{CON}$ that INT-identifies t , we have that at least two of $\varphi_i, \varphi_j, \varphi_k$ converge on t . By the definition of G_s , the sets G_s converge on t to (essentially) some two element subset $\{p, q\}$ of $\{i, j, k\}$ such that each of φ_p, φ_q converges on t . But, in general, if $r \in \text{CON}$ converges on t to $c \in N$, then $\text{rng}(t) \subseteq W_c$. Thus, since at least one of the two functions φ_p, φ_q INT-identifies t , the intersection taken in the definition of Θ guarantees that Θ INT-identifies t . ■

5. DIRECT AND RECENT AGGREGATION

We now consider the effect of constraining the aggregation process in two different ways. We begin by considering aggregators that must select

conjectures from those offered by their associated team by inductive inference machines.

DEFINITION 5.A. Let $m \in N$, $(i_1, \dots, i_m) \in N^m$, $\mathcal{E} \subseteq \mathcal{T}_{\text{RE}}$, $\psi \in \mathcal{F}^{\text{rec}}$ and accuracy criterion \mathcal{C} be given. ψ \mathcal{C} -aggregates (i_1, \dots, i_m) on \mathcal{E} directly just in case ψ \mathcal{C} -aggregates (i_1, \dots, i_m) on \mathcal{E} and for all $t \in \mathcal{E}$, $n \in N$, $\psi(\bar{t}_n) \in \{\varphi_{i_1}(\bar{t}_n), \dots, \varphi_{i_m}(\bar{t}_n)\}$.

DEFINITION 5.B. Let $m \in N$, $\mathcal{E} \subseteq \mathcal{T}_{\text{RE}}$, and accuracy criterion \mathcal{C} be given. Then: $[\text{program-direct}, \mathcal{C}, \mathcal{E}]^m = \{B \subseteq N^m \mid \text{there is computable } \Theta: N^m \times \text{SEQ} \rightarrow N \text{ such that for all } (i_1, \dots, i_m) \in B, \lambda\sigma.\Theta(i_1, \dots, i_m, \sigma) \mathcal{C}\text{-aggregates } (i_1, \dots, i_m) \text{ on } \mathcal{E} \text{ directly}\}$.

The next proposition shows program-direct aggregation to be more difficult than pure aggregation in the context of RE_{svt} . Compare Proposition 4.C(i).

PROPOSITION 5.A. For all odd $m \geq 3$,

- (i) $N^m \notin [\text{program-direct}, \text{INT}, \mathcal{T}_{\text{RE}_{\text{svt}}}]^m$.
- (ii) $N^m \notin [\text{program-direct}, \text{FINT}, \mathcal{T}_{\text{RE}_{\text{svt}}}]^m$.

Proof. (i) We exhibit (i, j, k) such that for all computable $\Theta: N^3 \times \text{SEQ} \rightarrow N$, Θ does not INT-aggregate (i, j, k) directly on $\mathcal{T}_{\text{RE}_{\text{svt}}}$. For the purpose of this proof, let $p: N^2 \rightarrow N$ be a recursive “padding” function with the following properties. For all $x, y \in N$.

$$W_{p(x, y)} = W_x,$$

and

$$p(x, y) > \max\{x, y\}.$$

Also, we define a recursive function $f: N^2 \times \{D \subseteq N \mid D \text{ finite}\} \rightarrow N$ as follows: For all $x, y \in N$, finite $D \subseteq N$,

$$W_{f(x, y, D)} = \{\langle 0, x \rangle, \langle 1, y \rangle\} \cup \{\langle n+2, 1 \rangle \mid n \in D\} \cup \{\langle n+2, 0 \rangle \mid n \notin D\}.$$

Finally, we recall the notation “ $\mathbf{n}(\sigma)$ ” introduced in the proof of Proposition 4.B.

Now we define the desired triple (i, j, k) as follows: For all $\sigma \in \text{SEQ}$,

$$\begin{aligned} \varphi_j(\sigma) &= \mathbf{0}(\sigma), \\ \varphi_k(\sigma) &= p(\mathbf{1}(\sigma), \varphi_j(\sigma)), \\ \varphi_i(\sigma) &= p(f(\mathbf{0}(\sigma), \mathbf{1}(\sigma), D), \varphi_k(\sigma)), \end{aligned}$$

where $D = \{n | \langle n+2, 1 \rangle \in \text{rng}(\sigma)\}$. Note that due to the use of padding function, $\varphi_j(\sigma) < \varphi_k(\sigma) < \varphi_i(\sigma)$ so that no pair of these is identical.

Now let $\Theta: N^3 \times \text{SEQ} \rightarrow N$ be given. We exhibit $t \in \mathcal{T}_{\text{RE}_{\text{svt}}}$ such that (i, j, k) identifies t but if for all $n \in N$, $\Theta(i, j, k, \bar{t}_n) \in \{\varphi_i(\bar{t}_n), \varphi_j(\bar{t}_n), \varphi_k(\bar{t}_n)\}$, then $\lambda\sigma.\Theta(i, j, k, \sigma)$ does not identify t . We shall define t by the recursion theorem. Thus we will actually define recursive functions g and h , and for each pair $\langle x, y \rangle$ a text t^{xy} . We will define $g(\langle x, y \rangle)$ while defining t^{xy} . The desired t will be one of the texts t^{xy} . For notational ease, denote $\Theta(i, j, k, \sigma)$ by $\Theta(\sigma)$.

Stage 0. $t_0^{xy} = \langle 0, x \rangle$; $t_1^{xy} = \langle 1, y \rangle$. Enumerate $\langle 0, x \rangle$, $\langle 1, y \rangle$ into $W_{g(\langle x, y \rangle)}$ and $W_{h(\langle x, y \rangle)}$.

Stage $n > 0$. There are four cases:

Case a. $\Theta(\bar{t}_{n+1}^{xy}) = \varphi_i(\bar{t}_{n+1}^{xy})$. In this case, set $t_{n+1}^{xy} = \langle n+1, 1 \rangle$ and enumerate $\text{rng}(\bar{t}_{n+2}^{xy})$ into both $W_{g(\langle x, y \rangle)}$ and $W_{h(\langle x, y \rangle)}$.

Case b. $\Theta(\bar{t}_{n+1}^{xy}) = \varphi_j(\bar{t}_{n+1}^{xy})$. In this case, set $t_{n+1}^{xy} = \langle n+1, 0 \rangle$ and enumerate $\text{rng}(\bar{t}_{n+2}^{xy})$ into $W_{h(\langle x, y \rangle)}$.

Case c. $\Theta(\bar{t}_{n+1}^{xy}) = \varphi_k(\bar{t}_{n+1}^{xy})$. In this case, set $t_{n+1}^{xy} = \langle n+1, 0 \rangle$ and enumerate $\text{rng}(\bar{t}_{n+2}^{xy})$ into $W_{g(\langle x, y \rangle)}$.

Case d. None of Cases a–c hold. Let $t_{n+k}^{xy} = \langle n+k, 0 \rangle$ for all $k > 0$ and enumerate $\text{rng}(t^{xy})$ into both $W_{g(\langle x, y \rangle)}$ and $W_{h(\langle x, y \rangle)}$. Do not proceed to stage $n+1$.

The above definition by cases is welldefined because of the remark about the padding function following the definition of (i, j, k) . Notice that for each x and y , t^{xy} is a text in $\mathcal{T}_{\text{RE}_{\text{svt}}}$. Let p and q be indices such that $W_p = W_{g(\langle p, q \rangle)}$ and $W_q = W_{h(\langle p, q \rangle)}$. We claim that t^{pq} is our desired text. To see this we first show that (i, j, k) INT-identifies t^{pq} . Notice that on t^{pq} , φ_j converges to p , and φ_k converges to $p(q, p)$ which is an index for W_q . There are now four cases to consider according to how t^{pq} is defined:

Case 1. Case a occurs infinitely often in the definition of t^{pq} . Then $W_p = W_q = \text{rng}(t^{pq})$ so that φ_k and φ_j both INT-identify t^{pq} .

Case 2. Case a occurs finitely often but Case b occurs infinitely often. In this case $\text{rng}(t^{pq}) = W_q$ so that φ_i and φ_k both INT-identify t^{pq} .

Case 3. Cases a and b occur finitely often but Case c occurs infinitely often. In this case $\text{rng}(t^{pq}) = W_p$ and φ_i and φ_j both INT-identify t^{pq} .

Case 4. None of Cases a–c occur infinitely often. Then Case d occurs once and $W_p = W_q = \text{rng}(t^{pq})$ and all of φ_i , φ_j , and φ_k INT-identify t^{pq} .

Finally we need to show that $\lambda\sigma.\Theta(\sigma)$ does not INT-identify t^{pq} directly. First, if $\Theta(\bar{t}_n^{pq}) = \varphi_i(\bar{t}_n^{pq})$ for cofinitely many n , then by Case a, $\lambda\sigma.\Theta(\sigma)$ (as well as φ_i) does not identify t^{pq} . On the other hand, if $\Theta(\bar{t}_n^{pq}) = \varphi_j(\bar{t}_n^{pq})$ or

$\varphi_k(\bar{t}_n^{pq})$ for cofinitely many n , then $\lambda\sigma.\Theta(\sigma)$ does not identify t^{pq} . Finally, if $\Theta(\bar{t}_n^{pq})$ is not any of $\varphi_i(\bar{t}_n^{pq})$, $\varphi_j(\bar{t}_n^{pq})$, or $\varphi_k(\bar{t}_n^{pq})$ cofinitely often, then either Θ does not converge or does not INT-identify t^{pq} directly.

(ii) Straightforward modifications to (i) suffice to prove (ii). The text t^{xy} needs to be modified as follows. Replace all pairs $\langle n, 0 \rangle$ and $\langle n, 1 \rangle$ by $\langle n, \langle x, y \rangle \rangle$ and $\langle n, \langle x, y \rangle + 1 \rangle$, respectively. Modify the definition of φ_i , φ_j , and φ_k accordingly. Then note that Θ either does not converge or converges to φ_j or φ_k , and in this case converges to an index for a finite function infinitely different from $\text{rng}(t^{pq})$. ■

From Propositions 4.C and 5.A we have

COROLLARY 5.A. *For all odd $m \geq 3$,*

- (i) $[\text{program-direct}, \text{INT}, \mathcal{T}_{\text{RE}}]^m \subset [\text{program}, \text{INT}, \mathcal{T}_{\text{RE}}]^m$.
- (ii) $[\text{program-direct}, \text{FINT}, \mathcal{T}_{\text{RE}}]^m \subset [\text{program}, \text{FINT}, \mathcal{T}_{\text{RE}}]^m$.

We turn next to aggregators that may examine only the most recent conjectures emitted by their associated teams.

DEFINITION 5.C. Let $m \in N$, $\mathcal{E} \subseteq \mathcal{T}_{\text{RE}}$, and accuracy criterion \mathcal{C} be given. Then,

- (i) $[\text{hypothesis-recent}, \mathcal{C}, \mathcal{E}]^m = \{B \subseteq N^m \mid \text{there is computable } \Theta: N^m \times \text{SEQ} \rightarrow N \text{ such that for all } (i_1, \dots, i_m) \in B, \lambda\sigma.\Theta(\varphi_{i_1}(\sigma), \dots, \varphi_{i_m}(\sigma), \sigma) \text{ } \mathcal{C}\text{-aggregates } (i_1, \dots, i_m) \text{ on } \mathcal{E}\}$.
- (ii) $[\text{pure-recent}, \mathcal{C}, \mathcal{E}]^m = \{B \subseteq N^m \mid \text{there is computable } \Theta: N^m \rightarrow N \text{ such that for all } (i_1, \dots, i_m) \in B, \lambda\sigma.\Theta(\varphi_{i_1}(\sigma), \dots, \varphi_{i_m}(\sigma)) \text{ } \mathcal{C}\text{-aggregates } (i_1, \dots, i_m) \text{ on } \mathcal{E}\}$.

The impact of the recency restriction is revealed in the following propositions.

PROPOSITION 5.B. *For all odd $m \geq 3$: $N^m \notin [\text{pure-recent}, \text{INT}, \mathcal{T}_{\text{RE}_{\text{svt}}}]^m$.*

Compare Proposition 4.C(i).

Proof. Fix $a \in N$ such that $W_a = \{\langle n, 0 \rangle \mid n \in N\}$; thus, W_a represents the constant 0-function. Fix total recursive function h such that for all $x \in N$, $W_{h(x)} \in \text{RE}_{\text{svt}}$ and $W_{h(x)} = W_a$ if and only if $W_x \subseteq W_a$. Also, let f be a recursive function such that $\varphi_{f(x)}(\sigma) = x$ for all $\sigma \in \text{SEQ}$. Finally, let g be a recursive function such that $\varphi_{g(x)}(\sigma) = a$ if $\text{lh}(\sigma)$ is even and $\varphi_{g(x)}(\sigma) = h(x)$ if $\text{lh}(\sigma)$ is odd.

Suppose now for a contradiction that $\Theta: N^3 \rightarrow N$ witnesses $N^3 \in [\text{pure-recent}, \text{INT}, \mathcal{T}_{\text{RE}_{\text{svt}}}]^3$.

CLAIM 1. For all $y \in N$, if $W_y \subseteq W_a$ then $\Theta(a, h(y), a) = \Theta(a, h(y), h(y))$.

Proof of Claim 1. Let t be any text for W_a . Under the hypothesis, for all $y \in N$, $(f(a), f(h(y)), g(y))$ is a triple that INT-identifies t . However, on t the conjectures of $(f(a), f(h(y)), g(y))$ alternate between $(a, h(y), a)$ and $(a, h(y), h(y))$. Hence, by the recency of Θ and the fact that Θ converges on t , $\Theta(a, h(y), a) = \Theta(a, h(y), h(y))$ for all $y \in N$.

CLAIM 2. For all $y \in N$, if $W_y \not\subseteq W_a$, then $\Theta(a, h(y), a) \neq \Theta(a, h(y), h(y))$.

Proof of Claim 2. Let t be any text for W_a . For all $y \in N$, the triple $(f(a), f(h(y)), f(a))$ INT-identifies t since it constantly conjectures $(a, h(y), a)$. Consequently, $\Theta(a, h(y), a)$ is an index for $\text{rng}(t)$, since Θ (by hypothesis) INT-aggregates $(f(a), f(h(y)), f(a))$. Similarly, $\Theta(a, h(y), h(y))$ is an index for $W_{h(y)}$, since $(f(a), f(h(y)), f(h(y)))$ INT-identifies any text for $W_{h(y)}$. Thus, since $W_{h(y)} \neq W_a$, $\Theta(a, h(y), a) \neq \Theta(a, h(y), h(y))$.

Claims 1 and 2 together exhibit $\{x \mid W_x \subseteq W_a\}$ as recursive, contracting Rice's theorem. ■

PROPOSITION 5.C. For all odd $m \geq 3$, $[\text{hypothesis-recent, INT, } \mathcal{T}_{\text{RE}}]^m \subset [\text{hypothesis, INT, } \mathcal{T}_{\text{RE}}]^m$.

Proof. The proof of this proposition is similar to that of Proposition 4.D. For $a, b, i, j, k \in N$, (i, j, k) is called “simple’ $_{ab}$ ” just in case (i, j, k) is simple $_{ab}$ except that for all $\sigma \in \text{SEQ}$, if $\text{lh}(\sigma) = 1$ then $\varphi_i(\sigma) = 0$ if and only if $W_a = W_b$. The set S witnessing that the inclusion of the proposition is proper is $S = \{(i, j, k) \mid \text{for some } a, b \in N, (i, j, k) \text{ is simple’}_{ab}\}$. It is clear that $S \in [\text{hypothesis, INT, } \mathcal{T}_{\text{RE}}]^3$ since Θ need only look at $\varphi_i(\bar{t}_1)$ to determine if $W_a = W_b$ and then choose conjectures from $\varphi_i(\sigma)$ or $\varphi_k(\sigma)$ accordingly. However, $S \notin [\text{hypothesis-recent, INT, } \mathcal{T}_{\text{RE}}]^3$ since otherwise we could construct ψ as in the proof of Proposition 4.D which INT-identifies RE_{psd} . To see this, let f, g, h be the recursive functions of this latter proof, namely, so that for all $a, b \in N$, the triple $(f(a, b), g(a, b), h(a, b))$ is simple $_{ab}$. Now given $\sigma \in \text{SEQ}$, let $i, j, k \in N$ be the triple $(f(\mathbf{1}(\sigma), \mathbf{2}(\sigma)), g(\mathbf{1}(\sigma), \mathbf{2}(\sigma)), h(\mathbf{1}(\sigma), \mathbf{2}(\sigma)))$. For $\sigma \in \text{SEQ}$, define $\psi(\sigma) = \Theta(\varphi_i(\sigma), \varphi_j(\sigma), \varphi_k(\sigma), \sigma)$, where Θ witnesses $S \in [\text{hypothesis-recent, INT, } \mathcal{T}_{\text{RE}}]^3$. To see that ψ INT-identifies RE_{psd} , note that there is a triple $(i', j', k') \in S$ such that $\varphi_{i'}(\sigma) = \varphi_i(\sigma)$, $\varphi_{j'}(\sigma) = \varphi_j(\sigma)$, $\varphi_{k'}(\sigma) = \varphi_k(\sigma)$ for all $\sigma \in \text{SEQ}$ with $\text{lh}(\sigma) \neq 1$. Consequently, Θ INT-aggregates (i, j, k) and hence ψ INT-identifies RE_{psd} . ■

The proofs of Propositions 5.A and 5.B yield

COROLLARY 5.B. For all odd $m \geq 3$, $[\text{program-direct, INT, } \mathcal{T}_{\text{RE}_{\text{svt}}}]^m$ and $[\text{pure-recent, INT, } \mathcal{T}_{\text{RE}_{\text{svt}}}]^m$ are incomparable.

It remains to be determined whether $N^m \in [\text{hypothesis, INT, } \mathcal{T}_{\text{RE}_{\text{svt}}}]^m$.

6. CLOSE AND EFFICIENT AGGREGATION

Let us now consider the effect of imposing accuracy and speed requirements on the aggregation process. To begin, we may require FINT-aggregators to converge to hypotheses at least as accurate as those of their associated team. Some additional notation is needed to make this idea precise.

DEFINITION 6.A. Total function $D: \mathcal{F}^{\text{rec}} \times \mathcal{T}_{\text{RE}} \rightarrow N \cup \{\omega\}$ is defined as follows. For $\psi \in \mathcal{F}^{\text{rec}}$, $t \in \mathcal{T}_{\text{RE}}$,

$$\begin{aligned} D(\psi, t) &= \omega && \text{if } \psi \text{ does not converge on } t; \\ &= \text{the cardinality of } (W_x - \text{rng}(t)) \cup (\text{rng}(t) - W_x) \\ &&& \text{otherwise, where } \psi \text{ converges on } t \text{ to } x. \end{aligned}$$

DEFINITION 6.B. Let $m \in N$, $(i_1, \dots, i_m) \in N^m$, $\mathcal{E} \subseteq \mathcal{T}_{\text{RE}}$, and $\psi \in \mathcal{F}^{\text{rec}}$ be given. ψ CLOSE-FINT-aggregates (i_1, \dots, i_m) on \mathcal{E} just in case for all $t \in \mathcal{E}$, if (i_1, \dots, i_m) FINT-identifies t then $D(\psi, t) \leq \min[D(\varphi_{i_1}, t), \dots, D(\varphi_{i_m}, t)]$.

Thus, CLOSE-FINT-aggregation requires the aggregator to converge to an index that is at least accurate as those selected by a successful, associated team.

DEFINITION 6.C. Let $m \in N$, and $\mathcal{E} \subseteq \mathcal{T}_{\text{RE}}$, be given. Then $[\text{program, CLOSE-FINT, } \mathcal{E}]^m = \{B \subseteq N^m \mid \text{there is computable } \Theta: N^m \times \text{SEQ} \rightarrow N \text{ such that for all } (i_1, \dots, i_m) \in B, \lambda \sigma. \Theta(i_1, \dots, i_m, \sigma) \text{ CLOSE-FINT-aggregates } (i_1, \dots, i_m) \text{ on } \mathcal{E}\}$.

The effect of the foregoing accuracy requirement is revealed in the next result; it should be compared to Proposition 4.C(ii).

PROPOSITION 6.A. For all odd $m \geq 3$: $N^m \notin [\text{program, CLOSE-FINT, } \mathcal{T}_{\text{RE}_{\text{svt}}}]^m$.

Proof. We exhibit $(i, j, k) \in N^3$ such that no $\psi \in \mathcal{F}^{\text{rec}}$ CLOSE-FINT-aggregates (i, j, k) on $\mathcal{T}_{\text{RE}_{\text{svt}}}$. Generalization to odd $m > 3$ is straightforward.

Let $\theta \in \mathcal{F}^{\text{rec}}$ be given. θ is called *self-naming* just in case $\theta = \varphi_{\theta(0)}$. θ is

called *almost self-naming* just in case $\theta(0) \downarrow$ and $\theta(x) = \varphi_{\theta(0)}(x)$ for all but at most one $x \in N$. The collection $\{L \in \text{RE}_{\text{svt}} \mid L \text{ represents a self-naming function}\}$ is denoted: RE_{sn} . The collection $\{L \in \text{RE}_{\text{svt}} \mid L \text{ represents an almost self-naming function}\}$ is denoted: RE_{asn} . Note that $\text{RE}_{\text{asn}} \subseteq \text{RE}_{\text{svt}}$. We rely on the following result, due to Case and Smith (1983):

$$\text{No } \psi \in \mathcal{F}^{\text{rec}} \text{ INT-identifies } \mathcal{T}_{\text{RE}_{\text{asn}}}. \quad (1)$$

For a proof of (1) see Osherson, Stob and Weinstein (1986, Proposition 6.2.3.A). The indices i, j, k may now be specified. Recall that e is an index for \emptyset . For all $\sigma \in \text{SEQ}$, let $\varphi_i(\sigma) = \varphi_j(\sigma) = e$ if there is not exactly one $x \in N$ such that $\langle 0, x \rangle \in \text{rng}(\sigma)$; $= x$ otherwise, where $\langle 0, x \rangle \in \text{rng}(\sigma)$. It is easy to see that both φ_i and φ_j INT-identify $\mathcal{T}_{\text{RE}_{\text{sn}}}$ and FINT-identify $\mathcal{T}_{\text{RE}_{\text{asn}}}$.

Choose k such that φ_k INT-identifies $\mathcal{T}_{\text{RE}_{\text{asn}}} - \mathcal{T}_{\text{RE}_{\text{sn}}}$. Informally, φ_k finds $\langle 0, p \rangle \in \text{rng}(t)$, assumes that φ_p and $\text{rng}(t)$ differ at exactly one argument, and then searches t for the needed “patch” to φ_p . (Using this procedure, φ_k does not converge on any $t \in \mathcal{T}_{\text{RE}_{\text{sn}}}$ since in this case its hypothesized patch changes infinitely often.)

Evidently, (i, j, k) FINT-identifies $\mathcal{T}_{\text{RE}_{\text{asn}}}$. Moreover, for all $t \in \mathcal{T}_{\text{RE}_{\text{asn}}}$, $\min[D(\varphi_i, t), D(\varphi_j, t), D(\varphi_k, t)] = 0$. However, the latter equation implies that if $\psi \in \mathcal{F}^{\text{rec}}$ CLOSE-FINT aggregates (i, j, k) on $\mathcal{T}_{\text{RE}_{\text{asn}}}$ then ψ INT-identifies $\mathcal{T}_{\text{RE}_{\text{asn}}}$, contradicting (1). ■

COROLLARY 6.A. For all odd $m \geq 3$, $[\text{program, CLOSE-FINT, } \mathcal{T}_{\text{RE}}]^m \subset [\text{program, FINT, } \mathcal{T}_{\text{RE}}]^m$.

The efficiency of aggregation is the second requirement examined in this section; attention is limited to the INT criterion of accuracy. We begin by defining the “identification point” (or “IP”) of an inductive inference machine on a text.

DEFINITION 6.D. (i) Total function $\text{IP}: \mathcal{F}^{\text{rec}} \times \mathcal{T}_{\text{RE}} \rightarrow N \cup \{\omega\}$ is defined as follows: For all $\psi \in \mathcal{F}^{\text{rec}}$, $t \in \mathcal{T}_{\text{RE}}$,

$$\begin{aligned} \text{IP}(\psi, t) &= \omega && \text{if } \psi \text{ does not INT-identify } t, \\ &= \mu n[\psi(\bar{i}_n) = \psi(\bar{i}_{n+k}) \text{ for all } k \in N] && \text{otherwise.} \end{aligned}$$

(ii) Given $m \in N$, IP is extended to $N^m \times \mathcal{T}_{\text{RE}}$ as follows. For all $(i_1, \dots, i_m) \in N^m$, $t \in \mathcal{T}_{\text{RE}}$,

$$\begin{aligned} \text{IP}((i_1, \dots, i_m), t) &= \omega && \text{if } (i_1, \dots, i_m) \text{ does not INT-identify } t, \\ &= \mu k[\text{for a majority } j_1, \dots, j_p \text{ of } (i_1, \dots, i_m), \\ &\quad \text{IP}(\varphi_{j_1}, t) \leq k, \dots, \text{IP}(\varphi_{j_p}, t) \leq k] && \text{otherwise.} \end{aligned}$$

Thus, $\text{IP}((i_1, \dots, i_m), t)$ is the earliest point of t (if such exists) at which a majority of $\varphi_{i_1}, \dots, \varphi_{i_m}$ begin to converge to indices for $\text{rng}(t)$. An efficient aggregator should reach its identification point on a given text no later than its associated team. This admonition may now be formalized as follows.

DEFINITION 6.E. Let $m \in \mathbb{N}$, $(i_1, \dots, i_m) \in \mathbb{N}^m$, $\mathcal{E} \subseteq \mathcal{T}_{\text{RE}}$, and $\psi \in \mathcal{F}^{\text{rec}}$ be given. ψ EFF-INT-aggregates (i_1, \dots, i_m) on \mathcal{E} just in case for all $t \in \mathcal{E}$, $\text{IP}(\psi, t) \leq \text{IP}((i_1, \dots, i_m), t)$.

It is easy to verify that if ψ EFF-INT-aggregates (i_1, \dots, i_m) on \mathcal{E} then ψ INT-aggregates (i_1, \dots, i_m) on \mathcal{E} .

DEFINITION 6.F. Let $m \in \mathbb{N}$, and $\mathcal{E} \subseteq \mathcal{T}_{\text{RE}}$, be given. Then $[\text{program}, \text{EFF-INT}, \mathcal{E}]^m = \{B \subseteq \mathbb{N}^m \mid \text{there is computable } \Theta: \mathbb{N}^m \times \text{SEQ} \rightarrow \mathbb{N} \text{ such that for all } (i_1, \dots, i_m) \in B, \lambda\sigma.\Theta(i_1, \dots, i_m, \sigma) \text{ EFF-INT-aggregates } (i_1, \dots, i_m) \text{ on } \mathcal{E}\}$.

The following result should be compared to Proposition 4.C(i). It highlights the impact of efficiency requirements on aggregation.

PROPOSITION 6.B. For all odd $m \geq 3$: $\mathbb{N}^m \notin [\text{program}, \text{EFF-INT}, \mathcal{T}_{\text{RE}_{\text{svt}}}]^m$.

Proof. Suppose for a contradiction that $\Theta: \mathbb{N}^3 \times \text{SEQ} \rightarrow \mathbb{N}$ witnesses $\mathbb{N}^3 \in [\text{program}, \text{EFF-INT}, \mathcal{T}_{\text{RE}_{\text{svt}}}]^3$. Let $W_a = \{\langle n, 0 \rangle \mid n \in \mathbb{N}\}$. We derive the same contradiction as in the proof of Proposition 5.B, namely, that $\{x \mid W_x \subseteq W_a\}$ is resursive. Let total recursive h be such that for all $x \in \mathbb{N}$, (a) $W_{h(x)} \in \text{RE}_{\text{svt}}$, (b) $W_{h(x)} = W_a$ if and only if $W_x \subseteq W_a$, and (c) $\{\langle 0, 0 \rangle, \langle 1, 0 \rangle\} \subseteq W_{h(x)}$. Let recursive f be defined by $\varphi_{f(x)}(\sigma) = x$ for all $x \in \mathbb{N}$ and $\sigma \in \text{SEQ}$. Let $\sigma^0 = \langle \langle 0, 0 \rangle \rangle$ and $\sigma^1 = \langle \langle 0, 0 \rangle, \langle 1, 0 \rangle \rangle$. Finally, let total recursive g be defined as follows. For all $x \in \mathbb{N}$, $\sigma \in \text{SEQ}$.

$$\begin{aligned} \varphi_{g(x)}(\sigma) &= a & \text{rng}(\sigma) \subseteq W_a, \text{ but } \sigma \neq \sigma^1, \\ &= h(x) & \text{otherwise.} \end{aligned}$$

CLAIM 1. For all $x \in \mathbb{N}$, if $W_x \subseteq W_a$ then $\Theta(f(a), f(h(x)), g(x), \sigma^0) = \Theta(f(a), f(h(x)), g(x), \sigma^1)$.

Proof of Claim 1. Let $x \in \mathbb{N}$ be such that $W_x \subseteq W_a$. Then $W_{h(x)} = W_a$. Consequently, for any text for W_a , $\varphi_{f(a)}$ and $\varphi_{f(h(x))}$ INT-identify t and $\text{IP}(\varphi_{f(a)}, t) = \text{IP}(\varphi_{f(h(x))}, t) = 0$; consequently, $\text{IP}(\lambda\sigma.\Theta(f(a), f(h(x)), g(x), \sigma), t) = 0$. Since σ^1 can be extended to a text for W_a , the claim follows.

CLAIM 2. For all $x \in N$, if $W_x \not\subseteq W_a$ then $\Theta(f(a), f(h(x)), g(x), \sigma^0) \neq \Theta(f(a), f(h(x)), g(x), \sigma^1)$.

Proof of Claim 2. Let $x \in N$ be such that $W_x \not\subseteq W_a$, and let t be a text for W_a which begins with σ^0 but not with σ^1 . Then, $\varphi_{f(a)}$ and $\varphi_{g(x)}$ INT-identify t and $\text{IP}(\varphi_{f(a)}, t) = \text{IP}(\varphi_{g(x)}, t) = 0$. Consequently, $\Theta(f(a), f(h(x)), g(x), \sigma^0)$ is an index for W_a . On the other hand, since $W_x \not\subseteq W_a$, there is a text t' for $W_{h(x)}$ such that t' begins with $(\langle 0, 0 \rangle, \langle 1, 0 \rangle, \langle y, z \rangle)$, where $z \neq 0$. For such a text t' , $\text{IP}(\varphi_{f(h(x))}, t') = 0$ and $\text{IP}(\varphi_{g(x)}, t') = 1$. Consequently, $\Theta(f(a), f(h(x)), g(x), \sigma^1)$ is an index for $W_{h(x)} \neq W_a$.

As in the proof of Proposition 5.B, Claims 1 and 2 provide the desired contradiction. ■

COROLLARY 6.B. For all odd $m \geq 3$, $[\text{program}, \text{EFF-INT}, \mathcal{F}_{\text{RE}}]^m \subset [\text{program}, \text{INT}, \mathcal{F}_{\text{RE}}]^m$.

7. PRESERVABILITY

Inductive inference machines sometimes have special properties that ought to be preserved by the systems that aggregate them. To illustrate, let \mathcal{P} be the consistent subset of \mathcal{F}^{rec} , in the sense of Definition 4.D. Then, aggregator A may be said to “preserve” \mathcal{P} just in case A implement a consistent function whenever each machine in its associated team does so. In this section we consider, for various properties \mathcal{P} , whether arbitrary aggregators can be replaced by systems that aggregate as much as the original system, and also preserve \mathcal{P} . The following definitions make this question precise.

DEFINITION 7.A. Let $m \in N$, $(i_1, \dots, i_m) \in N^m$, $\mathcal{E} \subseteq \mathcal{F}_{\text{RE}}$, $\psi \in \mathcal{F}^{\text{rec}}$, $\mathcal{P} \subseteq \mathcal{F}^{\text{rec}}$, and accuracy criterion \mathcal{C} be given. ψ \mathcal{C} -aggregates (i_1, \dots, i_m) on \mathcal{E} preserving \mathcal{P} just in case (a) ψ \mathcal{C} -aggregates (i_1, \dots, i_m) on \mathcal{E} , and (b) if $\{\varphi_{i_1}, \dots, \varphi_{i_m}\} \subseteq \mathcal{P}$ then $\psi \in \mathcal{P}$.

DEFINITION 7.B. Let $m \in N$, $\mathcal{E} \subseteq \mathcal{F}_{\text{RE}}$, $\mathcal{P} \subseteq \mathcal{F}^{\text{rec}}$, and accuracy criterion \mathcal{C} be given:

(i) $[\mathcal{P}\text{-program}, \mathcal{C}, \mathcal{E}]^m = \{B \subseteq N^m \mid \text{there is computable } \Theta: N^m \times \text{SEQ} \rightarrow N \text{ such that for all } (i_1, \dots, i_m) \in B, \lambda\sigma.\Theta(i_1, \dots, i_m, \sigma) \text{ } \mathcal{C}\text{-aggregates } (i_1, \dots, i_m) \text{ on } \mathcal{E} \text{ preserving } \mathcal{P}\}$.

(ii) $[\mathcal{P}\text{-hypothesis}, \mathcal{C}, \mathcal{E}]^m = \{B \subseteq N^m \mid \text{there is computable } \Theta: \text{SEQ}^m \times \text{SEQ} \rightarrow N \text{ such that for all } (i_1, \dots, i_m) \in B, \lambda\sigma.\Theta(\tilde{\varphi}_{i_1}(\sigma), \dots, \tilde{\varphi}_{i_m}(\sigma), \sigma) \text{ } \mathcal{C}\text{-aggregates } (i_1, \dots, i_m) \text{ on } \mathcal{E} \text{ preserving } \mathcal{P}\}$.

(iii) $[\mathcal{P}\text{-pure}, \mathcal{C}, \mathcal{E}]^m = \{B \subseteq N^m \mid \text{there is computable } \Theta: \text{SEQ}^m \rightarrow N \text{ such that for all } (i_1, \dots, i_m) \in B, \lambda\sigma.\Theta(\tilde{\varphi}_{i_1}(\sigma), \dots, \tilde{\varphi}_{i_m}(\sigma)) \text{ } \mathcal{C}\text{-aggregates } (i_1, \dots, i_m) \text{ on } \mathcal{E} \text{ preserving } \mathcal{P}\}$.

We now investigate the preservability of a natural subset of \mathcal{F}^{rec} .

DEFINITION 7.C. (Angluin, 1980). $\psi \in \mathcal{F}^{\text{rec}}$ is *conservative* just in case for all $\sigma \in \text{SEQ}$, if $\text{lh}(\sigma) > 1$ and $\text{rng}(\sigma) \subseteq W_{\psi(\bar{\sigma}_{\text{lh}(\sigma)-1})}$, then $\psi(\sigma) = \psi(\bar{\sigma}_{\text{lh}(\sigma)-1})$.

Thus, conservative inductive inference machines do not change conjectures until contradicted by new data.

PROPOSITION 7.A. Let $\mathcal{P} = \{\psi \in \mathcal{F}^{\text{rec}} \mid \psi \text{ is conservative}\}$, and let $m \geq 3$ be given. Then $[\mathcal{P}\text{-hypothesis, INT, } \mathcal{F}_{\text{RE}}]^m \subset [\text{hypothesis, INT, } \mathcal{F}_{\text{RE}}]^m$.

Proof. Let total recursive f and g be defined for all $\sigma \in \text{SEQ}$ and $x, y \in N$ by

$$\begin{aligned} \varphi_{f(x)}(\sigma) &= x; \\ \varphi_{g(x,y)}(\sigma) &= e \quad \text{if } \text{rng}(\sigma) = \emptyset \text{ or } \text{lh}(\sigma) \leq y, \\ &= x \quad \text{otherwise.} \end{aligned}$$

Recall that e is a fixed index for the empty set and let n be a fixed index for N . Define $S = \{(f(n), f(x), g(x, y)) \mid x, y \in N\} \cup \{(f(n), f(x), g(n, y)) \mid x, y \in N\}$. It is easy to see that $S \in [\text{hypothesis, INT, } \mathcal{F}_{\text{RE}}]^3$ and that for all $(i, j, k) \in S$, $\{\varphi_i, \varphi_j, \varphi_k\} \subseteq \mathcal{P}$. Now suppose that $\Theta: \text{SEQ}^3 \times \text{SEQ} \rightarrow N$ witnesses that $S \in [\text{hypothesis, INT, } \mathcal{F}_{\text{RE}}]^3$. For notational convenience, for $x, y, z \in N$, $\sigma \in \text{SEQ}$, define $\Theta_{xyz} = \lambda\sigma.\Theta(\tilde{\varphi}_{f(n)}(\sigma), \tilde{\varphi}_{f(x)}(\sigma), \tilde{\varphi}_{g(z,y)}(\sigma))$. Also, the concatenation of finite sequences is denoted in the obvious way by juxtaposition.

CLAIM 1. For all $x \in N$, if $W_x = N$, then

$$\begin{aligned} &\text{there is } \sigma \in \text{SEQ} \text{ such that } \text{rng}(\sigma) \subseteq W_x \text{ and for all } \tau, \gamma \in \text{SEQ}, \\ &\Theta_{x, \text{lh}(\sigma)+1, x}(\sigma\tau) = \Theta_{x, \text{lh}(\sigma)+1, n}(\sigma\gamma). \end{aligned} \quad (2)$$

Proof of Claim 1. Let $x \in N$ be such that $W_x = W_n$. Then, the triple $(f(n), f(x), g(z, y))$ INT-identifies N for any $z, y \in N$. In particular, fixing y , $(f(n), f(x), g(e, y))$ INT-identifies N . Thus, since Θ INT-aggregates this triple, there is $\sigma \in \text{SEQ}$ such that $\Theta_{xye}(\sigma)$ is an index for N . Let $y' = \text{lh}(\sigma) + 1$. Then $\Theta_{xy'x}(\sigma) = \Theta_{xy'n}(\sigma)$ and each is an index for N since $\tilde{\varphi}_{g(e,y)}(\sigma) = \tilde{\varphi}_{g(n,y')}(\sigma) = \tilde{\varphi}_{g(x,y')}(\sigma)$ for all $x \in N$. The claim then follows from the fact that Θ is supposed to preserve \mathcal{P} .

CLAIM 2. For all $x \in N$, if $W_x \neq N$ then (2) does not hold.

Proof of Claim 2. Otherwise, for x such that $W_x \neq N$ let σ witness that (2) holds, and let $y = \text{lh}(\sigma) + 1$. Now Θ_{xyx} INT-identifies W_x and Θ_{xyn} INT-identifies N . However (2) implies that Θ_{xyx} and Θ_{xyn} converge to the same index on any text that begins with σ . This contradicts $W_x \neq N$.

The claims together establish the desired contradiction to the existence of Θ . Namely, (2) is a Σ_2 -definition of the set $\{x \mid W_x = N\}$. No such definition exists (see Rogers, Theorem 13–VIII). ■

The next proposition gives an example of a property that can be preserved without loss of aggregating power.

DEFINITION 7.D. (Minicozzi, cited in Blum and Blum (1975). $\psi \in \mathcal{F}^{\text{rec}}$ is *reliable* just in case for all $t \in \mathcal{T}_{\text{RE}}$, if ψ converges on t then ψ identifies t .

Thus, reliable inductive inference machines never converge to an incorrect index.

PROPOSITION 7.B. Let $\mathcal{P} = \{\psi \in \mathcal{F}^{\text{rec}} \mid \psi \text{ is reliable}\}$, and let $m \geq 3$ be given. Then

- (i) $N^m \in [\mathcal{P}\text{-pure, INT, } \mathcal{T}_{\text{REsvt}}]^m$.
- (ii) $[\mathcal{P}\text{-pure, INT, } \mathcal{T}_{\text{REsvt}}]^m = [\text{pure, INT, } \mathcal{T}_{\text{REsvt}}]^m$.

Proof. (i) This requires only a slight modification of the proof of Proposition 4.C(i). Namely, for each s such that $G_s \neq G_{s-1}$, Θ conjectures e before shifting to its conjecture. Then Θ will never converge incorrectly on a text if two members of the team do not converge. Thus, Θ is reliable if the input team is.

- (ii) This follows from (i) and Proposition 4.C(i). ■

As the final topic in this section we consider inductive inference machines with limited memory of the input text. For $\sigma \in \text{SEQ}$ and $n \in \mathbb{N}$, the result of removing all but the last n members of σ is denoted: $\sigma - n$. If $n \geq \text{lh}(\sigma)$, then $\sigma - n = \sigma$. (Thus, if $\sigma = 3, 3, 8, 1, 9$ then $\sigma - 2 = 1, 9$.)

DEFINITION 7.E. (Wexler and Culicover, 1980). $\psi \in \mathcal{F}^{\text{rec}}$ is *memory-limited* just in case there is $n \in \mathbb{N}$ such that for all $\sigma, \tau \in \text{SEQ}$, if $\sigma - n = \tau - n$ and $\psi(\bar{\sigma}_{\text{lh}(\sigma)-1}) = \psi(\bar{\tau}_{\text{lh}(\tau)-1})$, then $\psi(\sigma) = \psi(\tau)$.

In other words, ψ is memory-limited just in case for some n and all σ , $\psi(\sigma)$ depends on no more than ψ 's previous conjecture and the last n members of σ .

PROPOSITION 7.C. Let $\mathcal{P} = \{\psi \in \mathcal{F}^{\text{rec}} \mid \psi \text{ is memory-limited}\}$, and let $m \geq 3$ be given. Then: $[\mathcal{P}\text{-program, INT, } \mathcal{T}_{\text{RE}}]^m \subset [\text{program, INT, } \mathcal{T}_{\text{RE}}]^m$.

Proof. Let A be any fixed nonrecursive, r.e. set and let $L = \{\langle 0, x \rangle \mid x \in A\}$. For each $n \in N$, let $L_n = L \cup \{\langle 1, n \rangle\}$ and let $L'_n = L \cup \{\langle 0, n \rangle, \langle 1, n \rangle\}$. Let $\mathcal{L} = \{L\} \cup \{L_n, L'_n \mid n \in N\}$. In Osherson, Stob, and Weinstein (1986, Proposition 4.4.1C) it is shown that:

no memory-limited member of \mathcal{F}^{rec} INT-identifies \mathcal{L} . (3)

In contrast, we now exhibit $(i, j, k) \in N^3$ that INT-identifies \mathcal{L} and is such that $\varphi_i, \varphi_j, \varphi_k$ are each memory-limited.

Let p be a fixed index for L . Let recursive functions f and g be defined so that for all $n \in N$, $W_{f(n)} = L_n$, and $W_{g(n)} = L'_n$. For $\sigma \in \text{SEQ}$, define

$$\begin{aligned} \varphi_i(\sigma) &= p && \text{if for all } n \in N, \langle 1, n \rangle \notin \text{rng}(\sigma), \\ &= f(n) && \text{if } \langle 1, n \rangle \text{ is the first pair } \langle 1, x \rangle \text{ occurring in } \sigma; \\ \varphi_j(\sigma) &= p && \text{if for all } n \in N, \langle 1, n \rangle \notin \text{rng}(\sigma), \\ &= g(n) && \text{if } \langle 1, n \rangle \text{ is the first pair } \langle 1, x \rangle \text{ occurring in } \sigma. \end{aligned}$$

The definition of φ_k uses the fixed isomorphism between SEQ and N allowing σ to be viewed as a natural number. For all $\sigma \in \text{SEQ}$.

$$\begin{aligned} \varphi_k(\sigma) &= \sigma && \text{if for all } n \in N, \langle 1, n \rangle \notin \text{rng}(\sigma), \\ &= f(n) && \text{if } \langle 1, n \rangle \text{ is the first pair } \langle 1, x \rangle \text{ occurring in } \sigma \text{ and} \\ &&& \text{for that } n, \langle 0, n \rangle \notin \text{rng}(\sigma), \\ &= g(n) && \text{if } \langle 1, n \rangle \text{ is the first pair } \langle 1, x \rangle \text{ occurring in } \sigma \text{ and} \\ &&& \text{for that } n, \langle 0, n \rangle \in \text{rng}(\sigma). \end{aligned}$$

It is easy to see that φ_i and φ_j are memory-limited. The operation of φ_k can be described as follows. φ_k uses its conjectures to store incoming data until a number of the form $\langle 1, n \rangle$ occurs in the text. Then φ_k conjectures $f(n)$ or $g(n)$ according to whether $\langle 0, n \rangle \in \text{rng}(\sigma)$ or not. This can be determined with the aid of φ_k 's preceding conjecture. If φ_k conjectures $f(n)$, it continues to conjecture $f(n)$ unless the pair $\langle 0, n \rangle$ occurs in the text. In this case, φ_k conjectures $g(n)$ forever. It is clear that (i, j, k) INT-identifies \mathcal{L} . For, φ_i and φ_j both INT-identify L , φ_i and φ_k both INT-identify L_n for every n , and φ_i and φ_k both INT-identify L'_n for every n . It is also easy to see that $(i, j, k) \in [\text{program}, \text{INT}, \mathcal{F}_{\text{RE}}]^3$.

Suppose however that $\{(i, j, k)\} \in [\mathcal{P}\text{-program}, \text{INT}, \mathcal{F}_{\text{RE}}]^3$ and that Θ witnesses this. Then $\lambda\sigma.\Theta(i, j, k, \sigma)$ is memory-limited and INT-identifies \mathcal{L} . However, this contradicts (3). ■

8. $[n, m]$ -AGGREGATION

Up to this point we have been examining “majority aggregation” in an obvious sense. The present discusses the more general case of “ n out of m ” aggregation. Attention is restricted to program-aggregation and the INT criterion of accuracy.

DEFINITION 8.A. Let $t \in \mathcal{T}_{RE}$, $n, m \in N$, and $(i_1, \dots, i_m) \in N^m$ be given. (i_1, \dots, i_m) $[n, m]$ -identifies t just in case at least n of $\varphi_{i_1}, \dots, \varphi_{i_m}$ INT-identify t . (i_1, \dots, i_m) $[n, m]$ -identifies $\mathcal{E} \subseteq \mathcal{T}_{RE}$ just in case (i_1, \dots, i_m) $[n, m]$ -identifies every $t \in \mathcal{E}$.

Thus, $[n, m]$ -identification is a collective form of INT-identification. As with Definition 3.A it is intended that two occurrences of the same index be counted twice.

DEFINITION 8.B. Let $n, m \in N$, $\Theta: N^m \times SEQ \rightarrow N$, and $\mathcal{E} \subseteq \mathcal{T}_{RE}$ be given:

(i) Θ $[n, m]$ -aggregates over \mathcal{E} just in case for all $t \in \mathcal{E}$ and all $(i_1, \dots, i_m) \in N^m$, if (i_1, \dots, i_m) $[n, m]$ -identifies t then $\lambda\sigma.\Theta(i_1, \dots, i_m)$ INT-identifies t .

(ii) $[n, m] = \{\mathcal{E} \subseteq \mathcal{T}_{RE} \mid \text{some computable } \Theta: N^m \times SEQ \rightarrow N \text{ } [n, m]\text{-aggregates over } \mathcal{E}\}$.

The following propositions illustrate this definition.

PROPOSITION 8.A. (i) $\mathcal{T}_{RE} \notin [2, 3]$.

(ii) $\mathcal{T}_{RE_{svt}} \in [2, 3]$.

Proof. Part (i) follows directly from the proof of Proposition 4.A(i). The function Θ described in the proof of Proposition 4.C(i) is easily adapted to witness part (ii). ■

PROPOSITION 8.B. $\mathcal{T}_{RE_{svt}} \notin [1, 2]$.

Proof. We use the following lemma due to Blum and Blum (1975); a proof may be found in Osherson, Stob and Weinstein (1986, Proposition 4.6.1C).

LEMMA. *There are collections $A \subseteq RE_{svt}$ and $B \subseteq RE_{svt}$ such that A and B are each INT-identifiable by recursive function but $A \cup B$ is not INT-identifiable by any recursive function.*

Proof. Let A and B be as in the lemma, and let φ_i and φ_j INT-identify A and B , respectively. Thus (i, j) $[1, 2]$ -identifies $\mathcal{T}_{A \cup B}$. Suppose however

that Θ witnesses that $\mathcal{T}_{\text{RE}_{\text{svt}}} \in [1, 2]$. Then $\lambda\sigma.\Theta(i, j, \sigma)$ identifies $A \cup B$, contradicting the lemma. \blacksquare

A more general result follows.

PROPOSITION 8.C. *Let $n, m \in N$ be given:*

- (i) *if $n/m > 2/3$ then $\mathcal{T}_{\text{RE}} \in [n, m]$.*
- (ii) *if $1/2 < n/m \leq 2/3$ then $[n, m] = [2, 3]$.*
- (iii) *if $n/m \leq 1/2$ then $[n, m] \subset [2, 3]$.*

Proof. (i) For every $i \in N$ and $\sigma \in \text{SEQ}$, if φ_i converges on every initial segment of σ , define $\text{CONV}(i, \sigma) =$ the least $t \in N$ such that for all $s \in N$, $t \leq s \leq \text{lh}(\sigma)$ implies $\varphi_i(\bar{\sigma}_s) = \varphi_i(\sigma)$. Also, define a recursive function $\text{MAJ}:\text{SEQ} \rightarrow N$ by $W_{\text{MAJ}(i_1, \dots, i_n)} = \{x \mid x \text{ is an element of a majority of the sets } W_{i_1}, \dots, W_{i_n}\}$, for all $(i_1, \dots, i_n) \in \text{SEQ}$. (It is easy to see that MAJ is recursive.) Now we define Θ , which witnesses that $\mathcal{T}_{\text{RE}} \in [n, m]$. Given $e_1, \dots, e_m \in N$ and $\sigma \in \text{SEQ}$, let $i_1, \dots, i_n \in \{e_1, \dots, e_m\}$ be such that the multiset $\{\text{CONV}(i_j, \sigma) \mid 1 \leq j \leq n\}$ contains the n least elements of the multiset $\{\text{CONV}(e_j, \sigma) \mid 1 \leq j \leq m\}$. Define $\Theta(e_1, \dots, e_m, \sigma) = \text{MAJ}(\varphi_{i_1}(\sigma), \dots, \varphi_{i_n}(\sigma))$.

Suppose now that e_1, \dots, e_m $[n, m]$ -identifies t . Then, in particular, at least n of the functions φ_{e_j} INT-identify t and hence converge. Let i_1, \dots, i_n be the n -many members of $\{e_1, \dots, e_m\}$ that index functions which converge earliest on t (not necessarily the functions that INT-identify t). Then $\lambda\sigma.\Theta(e_1, \dots, e_m, \sigma)$ converges on t to $\text{MAJ}(j_1, \dots, j_k)$, where φ_{i_k} converges to j_k for $k = 1, 2, \dots, n$. Now at most $m - n$ of the functions φ_{e_j} fail to INT-identify t so that at least $n - (m - n) = 2n - m$ of the functions $\varphi_{i_1}, \dots, \varphi_{i_n}$ INT-identify t . But $2n - m > m - n$ (since $2/3 < n/m$) so that a majority of the functions $\varphi_{i_1}, \dots, \varphi_{i_n}$ INT-identify t . Thus, $\text{MAJ}(j_1, \dots, j_n)$ is an index for $\text{rng}(t)$. This shows that Θ aggregates e_1, \dots, e_m on t .

(ii) We first show that if $n/m \leq 2/3$, $[n, m] \subseteq [2, 3]$. Suppose then that n and m are given such that $n/m \leq 2/3$ and suppose that computable $\Theta: N^m \times \text{SEQ} \rightarrow N$ $[n, m]$ -aggregates over $\mathcal{E} \subseteq \mathcal{T}_{\text{RE}}$. We define computable $\Omega: N^3 \times \text{SEQ} \rightarrow N$ which $[2, 3]$ -aggregates over \mathcal{E} . For $i, j, k \in N$, $\sigma \in \text{SEQ}$, define $\Omega(i, j, k, \sigma) = \Theta(i, j, k, i, j, k, \dots, l, \sigma)$, where l is i, j , or k depending on the remainder of m divided by 3. Now if 2 or 3 of i, j, k INT-identify $t \in \mathcal{E}$, then at least n of the arguments to Θ in the definition of Ω INT-identify t . Thus, since Θ $[n, m]$ -aggregates over \mathcal{E} , $\lambda\sigma.\Theta(i, j, k, i, j, k, \dots, l, \sigma)$ INT-identifies t . Hence Ω $[2, 3]$ -aggregates over \mathcal{E} .

Next we show that if $n/m > 1/2$, $[2, 3] \subseteq [n, m]$. The proof is by induction on $m \geq 3$. The cases of m even and m odd are treated separately.

Suppose then that m even is given and that the claim is true for $m' < m$. Now since m is even, $1/2 < n/m$ implies that $1/2 < (n-1)/(m-1)$. The

inductive hypothesis therefore guarantees that $[2, 3] \subseteq [n-1, m-1]$. But it is easy to see that $[n-1, m-1] \subseteq [n, m]$. For suppose that Θ $[n-1, m-1]$ -aggregates over \mathcal{E} ; then $\Omega(i_1, \dots, i_m, \sigma) = \Theta(i_1, \dots, i_{m-1}, \sigma)$ $[n, m]$ -aggregates over \mathcal{E} . Thus $[2, 3] \subseteq [n-1, m-1] \subseteq [n, m]$.

The inductive step for odd m is an intricate and unenlightening collection of number theoretic facts. We illustrate with the case $m = 5$ (and therefore $n \geq 3$). Consider the following three sets of triples chosen from the integers 1, 2, 3, 4, 5.

<i>A</i>	<i>B</i>	<i>C</i>
1, 2, 3	1, 3, 4	2, 3, 4
1, 2, 4	1, 3, 5	2, 3, 5
1, 2, 5	1, 4, 5	2, 4, 5

The key fact is this. Given any triple (i, j, k) from (1, 2, 3, 4, 5), at least two of the three triples in two of the three columns *A*, *B*, *C* contain at least two of (i, j, k) . This can be verified by examining the ten possibilities for (i, j, k) . Now suppose that computable $\Theta: N^3 \times \text{SEQ} \rightarrow N$ $[2, 3]$ -aggregates over \mathcal{E} . We define Ω which $[3, 5]$ -aggregates (and hence $[n, 5]$ -aggregates) over \mathcal{E} . Define computable $\mathbf{A}: N^5 \times \text{SEQ} \rightarrow N$ by

$$\mathbf{A}(e_1, \dots, e_5, \sigma) = \Theta(\Theta(e_1, e_2, e_3, \sigma), \Theta(e_1, e_2, e_4, \sigma), \Theta(e_1, e_2, e_5, \sigma), \sigma)$$

for all $e_1, \dots, e_5 \in N, \sigma \in \text{SEQ}$.

Define \mathbf{B} and \mathbf{C} similarly from the entries of columns *B* and *C* above. Finally, define

$$\Omega(e_1, \dots, e_5, \sigma) = \Theta(\mathbf{A}(e_1, \dots, e_5, \sigma), \mathbf{B}(e_1, \dots, e_5, \sigma), \mathbf{C}(e_1, \dots, e_5, \sigma), \sigma)$$

for all, $e_1, \dots, e_5 \in N, \sigma \in \text{SEQ}$.

The property mentioned above guarantees that Ω aggregates over \mathcal{E} .

For odd $m > 3$ the induction breaks down into a number of cases based on number-theoretic properties of m . For example, if $m = 4k + 1$ for some k (and hence $n \geq 2k + 1$), we can modify the proof for $m = 5$ as follows. Instead of 3 columns and 3 triples we use $2k + 1$ columns of $4k - 1$ ($4k - 1$)-tuples. The inductive hypotheses used are that $[2, 3] \subseteq [k + 1, k + 2]$ and $[2, 3] \subseteq [2k, 4k - 1]$. Computable $\Omega: N^{4k+1} \times \text{SEQ} \rightarrow N$ which witnesses $[2, 3] \subseteq [2k + 1, 4k + 1]$ is then defined as various compositions of functions which witness two inclusions in a manner entirely analogous to the case $m = 5$ above. The cases $m = 4k + 3$ are more difficult; there are a large number of special cases that must be verified individually before a general combinatorial principle can be applied.

(iii) If $n/m \leq 1/2$, the first part of the proof of (ii) guarantees that $[n, m] \subseteq [2, 3]$. That the inclusion is strict is a simple modification of the proof of Proposition 8.B. ■

9. CONCLUDING REMARKS

A vast number of aggregation paradigms may be defined by recombination of the concepts introduced above. Other paradigms result from more detailed consideration of the environments within which inductive inference takes place. For example, instead of texts for a language L , inductive inference machines may be presented with texts for the characteristic function of L , or with texts that provide imperfect information about L (see Osherson, Stob, and Weinstein, 1986, Chap. 5). Yet other paradigms arise from varying the criteria of accuracy that define successful identification. Royer (1985), for example, has proposed a concept of accuracy based on the "density" of one set in another. It can be shown that density-aggregation cannot be achieved on $\mathcal{T}_{RE,svt}$ for certain triples $(i, j, k) \in N^3$.

In addition to varying the accuracy criterion associated with successful identification, the stability of conjectures may also be taken into account. An illustration is provided by "BC-identification" in the sense of Case and Smith (1983) (called "EXT-identification" in Osherson, Stob, and Weinstein, 1986, Chap. 6). $\psi \in \mathcal{F}^{rec}$ BC-identifies $t \in \mathcal{T}_{RE}$ just in case $W_{\psi(\bar{i}_n)} = \text{rng}(t)$ for all but finitely many $n \in N$. It is easy to show that for all odd m , $N^m \in [\text{pure-recent, BC, } \mathcal{T}_{RE}]^m$.

A taxonomy of identification criteria is available in Osherson, Stob, and Weinstein (1986, Chap. 6).

RECEIVED: December 10, 1985; ACCEPTED: March 26, 1986

REFERENCES

- ANGLUIN, D. ((1980), Inductive inference of formal languages from positive data, *Inform. Control* **45**, 117–135.
- ANGLUIN, D., AND SMITH, C. (1983), Inductive inference: Theory and methods, *Comput. Surveys* **15**, No. 3, 237–269.
- BLUM, L., AND BLUM, M. (1975), Toward a mathematical theory of inductive inference, *Inform. Control* **28**, 125–155.
- CASE, J., AND SMITH, C. (1983), Comparison of identification criteria for machine inductive inference, *Theoret. Comput. Sci.* **25**, 193–220.
- GOLD, E. M. (1967), Language identification in the limit, *Inform. Control* **10**, 447–474.
- MACHTEY, M., AND YOUNG, P. (1981), "An Introduction to the General Theory of Algorithms," Prentice-Hall, Englewood Cliffs, N. J.

- OSHERSON, D., STOB, M., AND WEINSTEIN, S. (1986), "Systems that Learn: An Introduction to Learning Theory for Cognitive and Computer Scientists," MIT Press, Cambridge, Mass.
- ROGERS, H. (1967), "Theory of Recursive Functions and Effective Computability," McGraw-Hill, New York.
- ROYER, J. (1985), "On Machine Inductive Inference of Approximations," Technical Report 85-005, University of Chicago, Department of Computer Science.
- WEXLER, K., AND CULICOVER, P. (1980), "Formal Principles of Language Acquisition," MIT Press, Cambridge, Mass.