九州大学学術情報リポジトリ
Kyushu University Institutional Repository

Inferability of Recursive Real-Valued Functions

Hirowatari, Eiju
Department of Informatics Kyushu University

Arikawa, Setsuo
Department of Informatics Kyushu University

https://hdl.handle.net/2324/3006

出版情報：DOI Technical Report. 137, 1997-04-01. Department of Informatics, Kyushu University
バージョン：
権利関係：

DOI-TR-137

DOI Technical Report

Inferability of Recursive Real-Valued Functions

by

E. Hirowatari and S. Arikawa

April 1, 1997

Department of Informatics
Kyushu University
Fukuoka 812-81, Japan

Email: {eiju,arikawa}@i.kyushu-u.ac.jp Phone: +81-92-642-2697

Inferability of Recursive Real-Valued Functions

Eiju Hirowatari
Department of Informatics
Kyushu University
Fukuoka 812-81
Japan
eiju@i.kyushu-u.ac.jp

Setsuo Arikawa
Department of Informatics
Kyushu University
Fukuoka 812-81
Japan
arikawa@i.kyushu-u.ac.jp

April 1, 1997

Abstract

This paper presents a method of inductive inference of real-valued functions from given pairs of observed data of $(x, h(x))$, where h is a target function to be inferred. Each of such observed data inevitably involves some ranges of errors, and hence it is usually represented by a pair of rational numbers which show the approximate value and the error bound, respectively. On the other hand, a real number called a recursive real number can be represented by a pair of two sequences of rational numbers, which converges to the real number and converges to zero, respectively. These sequences show an approximate value of the real number and an error bound at each point. Such a real number can also be represented by a sequence of closed intervals with rational end points which converges to a singleton interval with the real number as both end points.

In this paper, we propose a notion of recursive real-valued functions that can enjoy the merits of the both representations of the recursive real numbers. Then we present an algorithm which approximately infers real-valued functions from numerical data with some error bounds, and show that there exists a rich set of real-valued functions which is approximately inferable in the limit from such numerical data. We also discuss the precision of the guesses from the machine when sufficient data have not yet given.

1 Introduction

This paper discusses the inductive inference of real-valued functions from examples. The most important problem we should solve is how to represent real numbers and real-valued functions in computers. A real number is represented by a sequence of rational numbers that converges to the number. Hence it is reasonable to regard a real-valued function as a function that maps a sequence of rational numbers to another sequence of rational numbers. The notion of computable real-valued functions was introduced by Grzegorczyk [4], and then several other formulations different from but equivalent to his have been reported [5, 8, 10, 12].

Roughly speaking, a real number is said to be computable if there exists a Turing machine which generates a sequence of rational numbers that converges to the real number, while a real-valued function h is computable if there exists a Turing machine that computes, for any real number x and any output precision 2^{-n}, a rational number s_n such that $|s_n - h(x)| \leq 2^{-n}$. Hence the machine to compute the function needs to know all the information on the number x. The above definition of computable real-valued functions is convenient for discussing the computational complexity of such functions. However it is not always amenable to learning or inductive inference from examples.

Examples of a real-valued function is a set of numerical data which are obtained through some scientific observations or experiments. Such numerical data inevitably involve some ranges of errors. Hence each of such numerical data can be represented as a pair of a rational number approximating an exact value and an error bound. Unfortunately we can not reduce the error bound as we like. The error bounds depend on the devices, environments and the like in which the observations and experiments are made. If we could realize a machine that computes a real-valued function, it would be impossible for the machine to know all the information on a real number as a value of the variable of the function.

Now each of the numerical data can also be represented as a pair of an upper and lower bounds to the exact value. Thus it is regarded as an interval number [9, 1], that is, a closed interval that contains the exact value. Then we can define a real-valued function as an interval function from interval numbers to interval numbers. An interval function h receives a closed interval I as an input, and produces as an output a closed interval J such that $h(I) \subseteq J$ by using only the information on the interval I.

In this paper, we propose a notion of recursive real-valued functions that can enjoy the merits of both computable real-valued functions and interval functions. Our functions are rich enough to express all the elementary functions with computable real coefficients. We present a learning model which learns real-valued functions from numerical data in terms of the recursive real-valued functions thus defined.

Note here that Apsītis, Freivalds and Smith [2] showed two approaches to modeling the learning real-valued functions from examples by using computable analytic functions and arbitrary computable functions of recursive real numbers, respectively. They proved that the set of continuous functions defined over an interval is learnable iff the interval is closed on both ends, and that the same is true for monotonic functions. Furthermore Haussler [6] considered the problem of learning a function from X into Y, as a generalization of the PAC learning model. In the model the learner receives randomly drawn example $(x, h_0(y)) \in X \times Y$ for some unknown target function h_0, and tried to find a decision rule $h : X \to A$, in order to minimize the expectation of a loss $l(y, a)$, where X, Y and A are arbitrary sets of reals, and l is a real-valued function. Our learning model is different from their models [2, 6].

Now our model can be considered as an inductive inference which is a process

of hypothesizing recursive real-valued functions which explain the sets of numerical data. As a successful identification criterion we follow Gold's identification in the limit [3]. An inference machine requests input data from time to time, and identifies an algorithm that computes the target function in the limit. As we deal with real-valued functions as target functions, we need to consider the precision of the guesses from the inference machine. For this purpose, we introduce the notion of approximate inductive inference as a successful identification criterion. The approximate inductive inference is the same as the ordinary inductive inference except that each guess from the machine satisfies a certain learning precision that indicates a rate of convergence to the hypothesis.

We show that a set of recursive real-valued functions whose domains are the same rational interval is approximately inferable as long as the set is recursively enumerable. For example, the set of all elementary functions whose domains are the same closed interval and coefficients are recursive real numbers is shown to be approximately inferable.

This paper is organized as follows: In Section 2, we give the definitions of recursive real numbers and related concepts necessary for our discussion. In Section 3, we introduce the notion of recursive real-valued functions and show some properties of them. In Section 4, we introduce the notion of approximate inductive inference as a successful identification criterion. In Section 5, we show which sets of real-valued functions are approximately inferable in the limit. In Section 6, we discuss the precision of the guesses when sufficient data have not yet given, and then refer to some related works.

2 Recursive Real Numbers

The notion of recursive real numbers has been defined from various aspects [7, 11, 13, 14, 15]. In this paper we adopt the following definition: A recursive real number is a pair of two sequences of rational numbers which converge to the number and zero, respectively.

Sets N, Q and R are the sets of all natural numbers, rational numbers and real numbers, respectively. By Q^+ we denote the set of all positive rational numbers.

Definition 1 *Let f be a function from N to Q, and g be a function from N to Q^+. A pair $\langle f, g\rangle$ is an approximate expression of a real number x, if f and g satisfy the following conditions:*

1. $\lim_{n\to\infty} g(n) = 0$.
2. $|f(n) - x| \leq g(n)$ *for any* n.

The number x is a recursive real, if there is an approximate expression $\langle f, g\rangle$ of x such that f and g are recursive.

$f(n)$ and $g(n)$ show an approximate value of the real number and an error bound at each point, respectively.

It is easy to show that every rational number is a recursive real number.

Example 1 *Napier's number e is a recursive real number. In fact, we can construct the following recursive functions f and g from N to Q:*

$$f(n) = \sum_{i=0}^{n-1} \frac{1}{i!}, \qquad g(n) = \frac{3}{n!}.$$

For any $n > 0$,

$$e - f(n) = \frac{1}{n!} + \frac{1}{(n+1)!} + \cdots < \frac{(1 + \frac{1}{2} + \frac{1}{4} + \cdots)}{n!} = \frac{2}{n!}.$$

For $n = 0$, $0 < e < 3$. Thus we have $f(n) < e < f(n) + g(n)$. Furthermore it is clear that $\lim_{n\to\infty} g(n) = 0$. Hence, the pair $\langle f, g\rangle$ is an approximate expression of the number e.

3 Recursive Real-Valued Functions

Hereafter we deal with the real-valued functions h such that $h(x)$ is recursive real for any recursive real number x in the domain of h. In this section, we define a rationalized domain of the function h, denoted by Dom_h, and a rationalized function on the set Dom_h, denoted by $\mathcal{F}_h$. Then the real-valued function h can be taken as a computable function on a set of recursive real numbers by using Dom_h and $\mathcal{F}_h$.

By a rational interval we mean an interval whose end points are rational. We sometimes call it just an interval if no confusion occurs.

For any subset S of R, we define a set $Approx(S)$ as the set of all approximate expressions $\langle f, g\rangle$ such that $[f(m) - g(m), f(m) + g(m)] \subseteq S$ for some n and any $m \geq n$.

Definition 2 *Let S be a subset of R, and h be a function from S to R. A rationalized domain of h, denoted by Dom_h, is a recursive subset of Q^2 which satisfies the following conditions:*

1. *For any $\langle f, g\rangle \in Approx(S)$, there is a number $n \in N$ such that $\langle f(m), g(m)\rangle \in Dom_h$ for any $m \geq n$.*
2. *$[p - \alpha, p + \alpha] \subseteq S$ for any $\langle p, \alpha\rangle \in Dom_h$.*

A rationalized function on Dom_h, denoted by $\mathcal{F}_h$, is a computable function from Dom_h to Q^2 which satisfies the following condition:

For any approximate expression $\langle f, g\rangle$ of a number x, there exists an approximate expression $\langle f_0, g_0\rangle$ of the number $h(x)$ such that $\langle f(n), g(n)\rangle \in Dom_h$ implies $\mathcal{F}_h(\langle f(n), g(n)\rangle) = \langle f_0(n), g_0(n)\rangle$.

If the above f and g are recursive, then so are f_0 and g_0. Thus the above function h from S to R satisfies the condition that $h(x)$ is recursive real for any recursive real number $x \in S$.

Definition 3 *Let h be a real-valued function. Then h is said to be a recursive real-valued function, if there exist a rationalized domain Dom_h of h, and a rationalized function $\mathcal{F}_h$ on Dom_h.*

From these definitions we can design the following algorithm $\mathcal{A}_h$ which computes the recursive real-valued function: For any recursive real-valued function h, the algorithm $\mathcal{A}_h$ takes a pair $\langle p, \alpha\rangle$ of rational numbers as an input, and produces $\mathcal{F}_h(\langle p, \alpha\rangle)$ and stops if $\langle p, \alpha\rangle \in Dom_h$, else it stops without any output. Thus our recursive real-valued functions are computable.

Example 2 *Let r be a recursive real number. Then a constant function $c_r(x) = r$ is a recursive real-valued function.*

Since r is a recursive real, there is an approximate expression $\langle f, g\rangle$ of r such that f and g are recursive. We define that $Dom_{c_r} = Q \times Q^+$, and a computable function $\mathcal{F}_{c_r}$ from Dom_{c_r} to Q^2 by $\mathcal{F}_{c_r}(\langle p, \alpha\rangle) = \langle f(n), g(n)\rangle$, where n is the least natural number such that $\frac{1}{\alpha} \leq n$. Then Dom_{c_r} is a rationalized domain of c_r, and $\mathcal{F}_{c_r}$ is a rationalized function on Dom_{c_r}. In fact, we can construct the following algorithm that computes c_r:

```
Algorithm: A_{c_r}
begin
  let ⟨p, α⟩ be an input;
  if ⟨p, α⟩ ∈ Q × Q^+ then begin
    n := 1;
    while nα < 1 do n := n + 1;
    output ⟨f(n), g(n)⟩
  end
end.
```

We call the following functions *basic functions*: x, $-x$, $\frac{1}{x}$, e^x, $\log x$, $\sin x$, $\arctan x$, $x^{\frac{1}{2}}$, $\arcsin x$ and c_r. As usual, $\frac{1}{x}$ for $x = 0$, $\log x$ for $x \leq 0$, $x^{\frac{1}{2}}$ for $x < 0$, and $\arcsin x$ for $|x| > 1$ are left undefined. These basic functions which generate the elementary functions are shown to be recursive real-valued functions.

Although, in general, functions with empty domain are not considered, in this paper, we admit such functions and add them to the list of the basic functions.

Theorem 1 *Let h be a real-valued function, and let Dom_h and $\mathcal{F}_h$ be a rationalized domain of h and a rationalized function on Dom_h, respectively. Then $h([p-\alpha, p+\alpha]) \subseteq [q-\beta, q+\beta]$, where $\langle p, \alpha\rangle \in Dom_h$ and $\langle q, \beta\rangle = \mathcal{F}_h(\langle p, \alpha\rangle)$.*

Proof. For any $x \in [p-\alpha, p+\alpha]$, there is an approximate expression $\langle f, g\rangle$ of x such that $f(0) = p$ and $g(0) = \alpha$, and hence an approximate expression $\langle f_0, g_0\rangle$ of the number $h(x)$ that $\langle f(n), g(n)\rangle \in Dom_h$ implies $\mathcal{F}_h(\langle f(n), g(n)\rangle) = \langle f_0(n), g_0(n)\rangle$. Since $f_0(0) = q$ and $g_0(0) = \beta$, $h(x) \in [q-\beta, q+\beta]$. Hence $h([p-\alpha, p+\alpha]) \subseteq [q-\beta, q+\beta]$. □

Theorem 2 *Let I be a rational interval on R, and h be a recursive real-valued function from I to R. Then h is continuous on I.*

Proof. Let Dom_h and $\mathcal{F}_h$ be a rationalized domain of h and a rationalized function on Dom_h. For any $x \in I$, there is an approximate expression $\langle f, g\rangle$ of x satisfying the conditions: (1) For any n, $[f(n)-g(n), f(n)+g(n)] \subseteq I$, and (2) if x is an interior of I, then $x \in (f(n)-g(n), f(n)-g(n))$ for any n, else $x = f(n)-g(n)$ or $x = f(n)+g(n)$ for any n. Furthermore there is an approximate expression $\langle f_0, g_0\rangle$ of the number $h(x)$ such that $\langle f(n), g(n)\rangle \in Dom_h$ implies $\mathcal{F}_h(\langle f(n), g(n)\rangle) = \langle f_0(n), g_0(n)\rangle$. For any $\varepsilon > 0$, there is a natural number l such that $2g_0(l) < \varepsilon$, and hence there is a number $\delta > 0$ such that $(x-\delta, x+\delta) \cap I \subseteq [f(l)-g(l), f(l)+g(l)]$. Therefore $|x-y| \leq \delta$ implies $|h(x)-h(y)| \leq \varepsilon$. Hence h is continuous on I. □

Theorem 1 and 2 assert that the recursive real-valued functions satisfy the conditions required in the interval analysis [1, 9].

We can extend our discussions on recursive real-valued functions of one variable to those of several variables. Note here that $x+y$ and xy are recursive real numbers if x and y are recursive real numbers.

Now we consider the compositions of recursive real-valued functions. Let h_1 and h_2 be recursive real-valued functions. Then we define the composite functions h_1+h_2, $h_1 \times h_2$ and $h_1 \circ h_2$ as follows:

$$\begin{aligned}(h_1+h_2)(x) &= h_1(x)+h_2(x),\\ (h_1 \times h_2)(x) &= h_1(x) \times h_2(x),\\ (h_1 \circ h_2)(x) &= h_1(h_2(x)).\end{aligned}$$

Theorem 3 *If h_1 and h_2 are recursive real-valued functions on some unions of open intervals, then so are the functions h_1+h_2, $h_1 \times h_2$ and $h_1 \circ h_2$.*

Let H be the set of all recursive real-valued functions on some unions of open intervals. Then, by Theorem 3, the set H is closed under the three compositions above.

Let h be a recursive real-valued function, and $\mathcal{A}_h$ be the algorithm that computes h. By $\mathcal{A}_h(\langle p,\alpha\rangle)$, we denote the output of $\mathcal{A}_h$ for an input $\langle p,\alpha\rangle$. We construct sets D^n_δ in the following way:

$$\begin{array}{rcl} A^0 & = & \{p-\alpha, p, p+\alpha\}, \\ A^n & = & A^{n-1} \cup \{a \in [p-\alpha, p+\alpha] \mid a = b + \frac{\alpha}{2^n} \text{ or } a = b - \frac{\alpha}{2^n} \text{ for } b \in A^{n-1}\}, \\ B^n & = & \{\langle a,\gamma\rangle \in Q^2 \mid a \in A^n, \gamma = \frac{\alpha}{2^n}, \text{ and } \mathcal{A}_h \text{ has an output for } \langle a,\gamma\rangle\}, \\ D^n_\delta & = & \{\langle a,\gamma\rangle \in \bigcup_{i=0}^n B^i \mid \mathcal{A}_h(\langle a,\gamma\rangle) = \langle b,\beta\rangle \text{ and } \beta < \frac{\delta}{2}\}, \end{array}$$

where $\langle p,\alpha\rangle \in Q \times Q^+$ and $\delta \in Q^+$. We call the set D^n_δ *the n-th division set of the interval* $[p-\alpha, p+\alpha]$ *w.r.t.* $\mathcal{A}_h$ *and* δ.

Proposition 1 *Let S be a subset of R, h be a recursive real-valued function from S to R, and $\langle p,\alpha\rangle \in Q \times Q^+$ such that $[p-\alpha, p+\alpha] \subseteq S$. Then there is a number $l \in N$ such that*

$$[p-\alpha, p+\alpha] \subseteq \bigcup_{\langle a,\gamma\rangle \in D^l_\delta} [a-\gamma, a+\gamma].$$

Proof. Let $A^* = \bigcup_{n\in N} A^n$ and $D^*_\delta = \bigcup_{n\in N} D^n_\delta$. We show that, for any $x \in (p-\alpha, p+\alpha)$, there is an $\langle a,\gamma\rangle \in D^*_\delta$ such that $x \in (a-\gamma, a+\gamma)$. (1) In case $x \in A^*$. Because of $x \in Q$, there is an approximate expression $\langle f,g\rangle$ of x such that $f(n) = x$ and $g(n) = \dfrac{q}{2^n}$. There is a natural number s such that $\langle f(n), g(n)\rangle \in D^*_\delta$ for any $n \geq s$. (2) In case $x \notin A^*$. We consider the following functions:

$$\begin{array}{rcl} f_0(n) & = & \begin{cases} p & \text{if } n = 0 \\ f_0(n-1) + g_0(n) & \text{if } n > 0 \text{ and } x > f_0(n-1) \\ f_0(n-1) - g_0(n) & \text{otherwise ,} \end{cases} \\ g_0(n) & = & \dfrac{\alpha}{2^n}. \end{array}$$

It is obvious that $\lim_{n\to\infty} g_0(n) = 0$ and $g_0(n) > 0$ for any n. We show that $|f_0(n) - x| \leq g_0(n)$, for any n, by a mathematical induction on n. In case $n = 0$. Since $x \in (p-\alpha, p+\alpha)$, $|f_0(0)-x| = |p-x| < \alpha = g_0(0)$. In case $n > 0$. We assume the claim for $(n-1)$. By the inductive hypothesis, $f_0(n-1)-g_0(n-1) \leq x \leq f_0(n-1)+g_0(n-1)$. Note that $2g_0(n) = g_0(n-1)$. If $x > f_0(n-1)$, then $f_0(n) = f_0(n-1) + g(n)$. Thus,

$$\begin{array}{rcl} f_0(n) - g_0(n) & = & f_0(n-1) + g_0(n) - g_0(n) \\ & = & f_0(n-1) \ < x \ \leq f_0(n-1) + g_0(n-1) \\ & = & f_0(n-1) + 2g_0(n) \\ & = & f_0(n) + g_0(n). \end{array}$$

If $x \leq f_0(n-1)$, then $f_0(n) = f_0(n-1) - g(n)$. Thus,

$$\begin{array}{rcl} f_0(n) - g_0(n) & = & f_0(n-1) - g_0(n) - g_0(n) \\ & = & f_0(n-1) - 2g_0(n) \\ & = & f_0(n-1) - g_0(n-1) \ \leq x \ \leq f_0(n-1) \\ & = & f_0(n) + g_0(n). \end{array}$$

Consequently $\langle f_0, g_0\rangle$ is an approximate expression of x.

By the definition of f_0, $f_0(n) \in A^n$ for any n. Therefore there is a natural number t such that $\langle f_0(n), g_0(n)\rangle \in D^*_\delta$ for any $n \geq t$. Since $x \notin A^*$, $x \in (f_0(n) - g_0(n), f_0(n) + g_0(n))$ for any n. Hence, for any $x \in (p-\alpha, p+\alpha)$, there is an $\langle a, \gamma\rangle \in D^*_\delta$ such that $x \in (a-\gamma, a+\gamma)$.

For any $q > 0$, it holds that $[p-\alpha+q, p+\alpha-q] \subseteq \bigcup_{\langle a,\gamma\rangle \in D^*_\delta}(a-\gamma, a+\gamma)$. By the theorem of Hine-Borel, there are finite elements $\langle a_1, \gamma_1\rangle, \cdots, \langle a_j, \gamma_j\rangle \in D^*_\delta$ such that $[p-\alpha+q, p+\alpha-q] \subseteq \bigcup_{1\leq i\leq j}(a_i-\gamma_i, a_i+\gamma_i)$. By taking q small enough, there are $\langle b_1, \beta_1\rangle, \langle b_2, \beta_2\rangle \in D^*_\delta$ such that $[p-\alpha, p-\alpha+q] \subseteq [b_1-\beta_1, b_1+\beta_1]$ and $[p+\alpha-q, p+\alpha] \subseteq [b_2-\beta_2, b_2+\beta_2]$. Thus we have

$$[p-\alpha, p+\alpha] \subseteq \bigcup_{1\leq i\leq j}(a_i-\gamma_i, a_i+\gamma_i) \cup [b_1-\beta_1, b_1+\beta_1] \cup [b_2-\beta_2, b_2+\beta_2].$$

Hence there exists a natural number l such that

$$[p-\alpha, p+\alpha] \subseteq \bigcup_{\langle a,\gamma\rangle \in D^l_\delta}[a-\gamma, a+\gamma].$$

□

Proposition 2 *Let S be a subset of R, h be a recursive real-valued function from S to R, and let $\mathcal{A}_h$ be an algorithm that computes h. Furthermore, let $\langle p, \alpha\rangle \in Q \times Q^+$ and $\delta \in Q^+$. Then there is an algorithm $\mathcal{B}_h$ which works as follows: If $[p-\alpha, p+\alpha] \subseteq S$, then $\mathcal{B}_h$ outputs a set U and stops, else it never stops, where U is D^n_δ, that is, the n-th division set of $[p-\alpha, p+\alpha]$ w.r.t. $\mathcal{A}_h$ and δ such that $[p-\alpha, p+\alpha] \subseteq \bigcup_{\langle a,\gamma\rangle \in U}[a-\gamma, a+\gamma]$ for some n.*

Proof. We consider the following algorithm $\mathcal{B}_h(p, \alpha, \delta, U)$ that requests $\langle p, \alpha\rangle \in Q \times Q^+$ and $\delta \in Q^+$:

Algorithm: $\mathcal{B}_h(p, \alpha, \delta, U)$
begin
  $T := 0$; $n := 0$;
  while $T = 0$ **do begin**
    $division_h(p, \alpha, n, \delta, U)$;
    $covering(p, \alpha, U, x)$;
    if $x = 1$ **then** $T := 1$ **else** $n := n + 1$
  end;
  output U
end.

The algorithm $division_h(p, \alpha, n, \delta, U)$ requests $\langle p, \alpha\rangle \in Q \times Q^+$, $n \in N$ and $\delta \in Q^+$ as inputs, and outputs a set U, i.e., the n-th division set D^n_δ of $[p-\alpha, p+\alpha]$ w.r.t. $\mathcal{A}_h$

and δ. The algorithm $covering(p, \alpha, U, x)$ requests $\langle p, \alpha \rangle \in Q \times Q^+$ and $U \subseteq Q \times Q^+$ as inputs, and works as follows: If $[p - \alpha, p + \alpha] \subseteq \bigcup_{\langle a,\gamma \rangle \in U} [a - \gamma, a + \gamma]$, then the algorithm outputs $x = 1$ and stops, else it outputs $x = 0$ and stops. In fact, we can construct the algorithms as follows:

Algorithm: $division_h(p, \alpha, n, \delta, U)$
begin
  $A := \{p\}$; $U := \emptyset$;
  for $k := 0$ **to** n **do begin**
    for each $a \in A$ **do begin**
      $v := a - \frac{\alpha}{2^k}$; $w := a + \frac{\alpha}{2^k}$;
      if $p - \alpha \leq v$ **then** $A := A \cup \{v\}$;
      if $w \leq p + \alpha$ **then** $A := A \cup \{w\}$
    end;
    for each $a \in A$ **do if** $\mathcal{A}_h$ has
      an output for $\langle a, \frac{\alpha}{2^k} \rangle$ **do begin**
        $\langle b, \beta \rangle := \mathcal{A}_h(\langle a, \frac{\alpha}{2^k} \rangle)$;
        if $\beta < \frac{\delta}{2}$ **then** $U := U \cup \{(a, \frac{\alpha}{2^k})\}$
    end
  end;
  output U
end.

Algorithm: $covering(p, \alpha, U, x)$
begin
  $l := p - \alpha$;
  while $U \neq \emptyset$ **do begin**
    $V := \emptyset$;
    for each $\langle a, \gamma \rangle \in U$ **do**
      if $a - \gamma \leq l$ **then begin**
        $V := V \cup \{\langle a, \gamma \rangle\}$;
        $U := U \setminus \{\langle a, \gamma \rangle\}$
      end;
    if $V = \emptyset$ **then** $U := \emptyset$
    else for each $\langle a, \gamma \rangle \in V$ **do**
      if $l < a + \gamma$ **then** $l := a + \gamma$
  end;
  if $p + \alpha \leq l$ **then** $x := 1$ **else** $x := 0$;
  output x
end.

Now we show that the algorithm $\mathcal{B}_h(p, \alpha, \delta, U)$ stops iff $[p - \alpha, p + \alpha] \subseteq S$.

($\Rightarrow$) The proof is given by the contraposition. Assume that $[p-\alpha, p+\alpha] \nsubseteq S$. Then there is a real number $y \in [p - \alpha, p + \alpha] \setminus S$. The algorithm $\mathcal{A}_h$ has no output for any $\langle a, \gamma \rangle$ such that $y \in [a - \gamma, a + \gamma]$. Thus, for any n and $\delta > 0$, $division_h(p, \alpha, n, \delta, U)$ outputs U such that $[p-\alpha, p+\alpha] \nsubseteq \bigcup_{\langle a,\gamma \rangle \in U} [a-\gamma, a+\gamma]$. Therefore $covering(p, \alpha, U, x)$ outputs $x = 0$ and stops. Consequently $\mathcal{B}_h(p, \alpha, \delta, U)$ never stops, if $[p - \alpha, p + \alpha] \nsubseteq S$.

($\Leftarrow$) Assume that $[p - \alpha, p + \alpha] \subseteq S$. By Proposition 1, there is a natural number l such that $[p - \alpha, p + \alpha] \subseteq \bigcup_{\langle a,\gamma \rangle \in D^l_\delta} [a - \gamma, a + \gamma]$. Thus $division_h(p, \alpha, l, \delta, U)$ outputs U, and then $covering(p, \alpha, U, x)$ outputs $x = 1$, where U is D^l_δ. Consequently $\mathcal{B}_h(p, \alpha, \delta, U)$ outputs U, that is D^l_δ such that $[p - \alpha, p + \alpha] \subseteq \bigcup_{\langle a,\gamma \rangle \in D^l_\delta} [a - \gamma, a + \gamma]$, and stops, if $[p - \alpha, p + \alpha] \subseteq S$. $\square$

4 Inductive Inference Machines

In our scientific activities we can not observe the exact value of a real number x but can observe some approximate values of x. Such approximate values can be captured by a pair $\langle p, \alpha \rangle$ of rational numbers such that p is an approximate value of the number x and α is its error bound, i.e., $x \in [p - \alpha, p + \alpha]$. We call such a pair $\langle p, \alpha \rangle$ a datum of x.

Definition 4 *Let S be a subset of R and h be a function from S to R. A datum of a function h is a pair $\langle\langle p,\alpha\rangle,\langle q,\beta\rangle\rangle$ of data $\langle p,\alpha\rangle$ and $\langle q,\beta\rangle$ of the numbers $x \in S$ and $h(x)$, respectively.*

A *presentation* of the function h is an infinite sequence $w_1, w_2, \cdots$ of data of h in which, for any number x in the domain of h and any positive number ζ, there is a $w_i = \langle\langle p,\alpha\rangle,\langle q,\beta\rangle\rangle$ such that $x \in [p-\alpha, p+\alpha] \subseteq (x-\zeta, x+\zeta)$ and $h(x) \in [q-\beta, q+\beta] \subseteq (h(x)-\zeta, h(x)+\zeta)$. By σ we denote such a presentation, and by $\sigma[n]$ we denote the σ's initial segment of length n.

Definition 5 *An inductive inference machine (IIM, for short) is a procedure that requests inputs from time to time and produces algorithms that compute recursive real-valued functions from time to time. A guess is an output produced by the machine.*

For an IIM $\mathcal{M}$ and a finite sequence $\sigma[n] = w_1, w_2, \cdots, w_n$, by $\mathcal{M}(\sigma[n])$ we denote the last guess of the IIM $\mathcal{M}$ requested data $w_1, w_2, \cdots, w_n$ as inputs. In this paper, we assume that $\mathcal{M}(\sigma[n])$ is defined for any n.

An IIM $\mathcal{M}$ is said to *converge* to an algorithm $\mathcal{A}$ for a presentation σ, if there exists a number $n \in N$ such that $\mathcal{M}(\sigma[m])$ is equal to the algorithm $\mathcal{A}$ for any $m \geq n$.

Definition 6 *Let S_0 and S be subsets of R, h_0 be a function from S_0 to R, and h be a function from S to R. The h_0 is a restriction of h (in S_0), or equivalently h is an extension of h_0 (to S), iff $S_0 \subseteq S$ and $h_0(x) = h(x)$ for any $x \in S_0$.*

By $\mathcal{T}_0 \preceq \mathcal{T}$ we mean that $\mathcal{T}$ is a set of extensions of functions in $\mathcal{T}_0$, that is, for any $h_0 \in \mathcal{T}_0$, there is an extension of h_0 in $\mathcal{T}$.

Definition 7 *Let $\mathcal{T}$ be a set of functions. An IIM $\mathcal{M}$ is said to approximately infer $\mathcal{T}$ in the limit from data, if the following conditions are satisfied for any $h \in \mathcal{T}$ and any presentation σ of h: (1) the IIM $\mathcal{M}$ converges to an algorithm that computes an extension of h, and (2) for any $\langle\langle p,\alpha\rangle,\langle q,\beta\rangle\rangle \in \sigma[n]$, there is an $x \in [p-\alpha, p+\alpha] \cap S$ such that $[h_n(x)-\beta, h_n(x)+\beta] \cap [q-\beta, q+\beta] \neq \emptyset$, where h_n is a recursive real-valued function the algorithm $\mathcal{M}(\sigma[n])$ computes.*

A set $\mathcal{T}$ is said to be approximately inferable in the limit from data, if there is an IIM $\mathcal{M}$ that approximately infers $\mathcal{T}$ in the limit from data.

Let h be a target function, $\sigma = w_1, w_2, \cdots$ be a presentation of h, and $\sigma[n]$ be the data so far received by an IIM $\mathcal{M}$ as inputs. Then, at the point n, the machine can not tell whether the presentation σ is of h itself or of its extension. In other words, for any presentation of a function, the IIM $\mathcal{M}$ can not distinguish the function from its extensions. Hence, from now on, we do not distinguish functions from their extensions.

On the other hand, a set $\mathcal{T}$ of recursive real-valued functions is said to be recursively enumerable if there is a recursive function Ψ such that the set $\mathcal{T}$ is equal to the set of all functions computed by algorithms $\Psi(1), \Psi(2), \cdots$. Therefore we may say that a set $\mathcal{T}_0$ of recursive real-valued functions is recursively enumerable, if there is a recursively enumerable set $\mathcal{H}$ of recursive real-valued functions which is a set of extensions of functions in $\mathcal{T}_0$, that is, $\mathcal{T}_0 \preceq \mathcal{H}$ holds.

5 Approximate Inductive Inference

Now we discuss the inferability of sets of functions, and show which sets are approximately inferable in the limit from data.

Theorem 4 *Let $\mathcal{T}$ be a recursively enumerable set of recursive real-valued functions on the same rational closed interval I. Then $\mathcal{T}$ is approximately inferable in the limit from data.*

Proof. Since $\mathcal{T}$ is the recursively enumerable, there is a recursively enumerable set $\mathcal{H}$ of recursive real-valued functions such that $\mathcal{T} \preceq \mathcal{H}$. Hence there exists a recursive function Ψ such that $\mathcal{H}$ is equal to the set of all functions computed by algorithms $\Psi(0), \Psi(1), \cdots$.

Let h_0 be a target function in $\mathcal{T}$, and $\sigma = w_1, w_2, \cdots$ be a presentation of h_0. We consider the following IIM $\mathcal{M}$ that requests data $w_1, w_2, \cdots, w_n, \cdots$ as inputs from time to time:

```
IIM: M
begin
  D := ∅;   n := 1;
  repeat
    D := D ∪ {w_n};
    m := 1;   T := 0;
    while T = 0 do begin
      check_{Ψ(m)}(I, D, x);
      if x = 0 then m := m + 1
        else T := 1;
    end;
    output A_{Ψ(m)};
    n := n + 1
  forever
end.
```

```
Algorithm: check_h(I, D, x)
begin
  x := 1;
  let ⟨i, ξ⟩ ∈ Q^2 such that [i − ξ, i + ξ] = I;
  for each ⟨⟨p, α⟩, ⟨q, β⟩⟩ ∈ D do begin
    y := 0;   B_h(i, ξ, β, U);
    for each ⟨a, γ⟩ ∈ U such that
    [a − γ, a + γ] ∩ [p − α, p + α] ≠ ∅ do begin
      ⟨b, λ⟩ := A_h(⟨a, γ⟩);
      if [b − λ, b + λ] ∩ [q − β, q + β] ≠ ∅
        then y := 1
    end;
    x := xy
  end;
  output x
end.
```

The algorithm $check_h(I, D, x)$ requests a rational interval I and a set D of data as inputs, and works as follows: It outputs $x = 0$ if D is not a set of data of h,

and outputs $x = 1$ if there exists a real number $x \in I \cap [p-\alpha, p+\alpha]$ such that $[h(x)-\beta, h(x)+\beta] \cap [q-\beta, q+\beta] \neq \emptyset$ for any $\langle\langle p,\alpha\rangle, \langle q,\beta\rangle\rangle \in D$.

First, we consider the case that the IIM $\mathcal{M}$ outputs the guess $\mathcal{M}(\sigma[n])$. Let l_n be a natural number such that $\mathcal{M}(\sigma[n]) = \Psi(l_n)$. The **while** loop in the IIM $\mathcal{M}$ is repeated at most l_n times. Note that the set D, an input to the algorithm $check_{\Psi(l_n)}(I, D, x)$, is equal to the set $\sigma[n]$. Thus the algorithm $check_{\Psi(l_n)}(I, D, x)$ in the **while** loop outputs $x = 1$. Consequently, for any data $\langle\langle p,\alpha\rangle, \langle q,\beta\rangle\rangle \in \sigma[n]$, there exists a real number $x \in I \cap [p-\alpha, p+\alpha]$ such that $[h_n(x)-\beta, h_n(x)+\beta] \cap [q-\beta, q+\beta] \neq \emptyset$, where h_n is a recursive real-valued function the algorithm $\mathcal{M}(\sigma[n])$ computes.

Now we assume that the IIM $\mathcal{M}$ never converges to any algorithm that computes an extension of h. Since $h_0 \in \mathcal{T}$, $\mathcal{T} \preceq \mathcal{H}$ and $\mathcal{H} = \{\Psi(i) \mid i \in N\}$, there exists a natural number l such that the target function h_0 is a restriction of the function computed by the algorithm $\Psi(l)$. Thus the algorithm $check_{\Psi(l)}(I, D, x)$ outputs $x = 1$, if D is a set of data of the target function h_0. Consequently we have $\mathcal{M}(\sigma[n]) = \Psi(l)$ for any $n \geq l$, which gives a contradiction.

Hence the set $\mathcal{T}$ is approximately inferable in the limit from data. □

For any set $\mathcal{T}$ of functions, we define $\mathcal{T}_I$ as the set of all functions in $\mathcal{T}$ such that the domains include the same rational interval I.

Theorem 5 *Let $\mathcal{T}$ be a recursively enumerable set of recursive real-valued functions, and let I be a rational closed interval. If the set $\mathcal{T}_I$ is not empty, then the set $\mathcal{T}_I$ is approximately inferable in the limit from data.*

Proof. Assume that $\mathcal{T}_I$ is not empty. Note that $\mathcal{T}_I$ is a set of recursive real-valued functions. Since $\mathcal{T}$ is recursively enumerable, there is a recursively enumerable set $\mathcal{H}$ of recursive real-valued functions such that $\mathcal{T} \preceq \mathcal{H}$. Hence there is a recursive function Ψ such that the set $\mathcal{H}$ is equal to the set of all functions computed by algorithms $\Psi(0), \Psi(1), \cdots$.

In order to complete the proof, it suffices to show that there is a recursively enumerable set $\mathcal{H}_0$ of recursive real-valued functions such that $\mathcal{T}_I \preceq \mathcal{H}_0$. Consider the following algorithm:

Algorithm: $\mu_I(n, m, x)$
begin
  let $\langle p,\alpha\rangle \in Q^2$ such that $[p-\alpha, p+\alpha] = I$;
  $division_{\Psi(m)}(p, \alpha, n, 1, U)$;
  $covering(p, \alpha, U, x)$;
  output x
end.

The algorithm $\mu_I(n, m, x)$ requests natural numbers n and m as inputs, and computes the n-th division set D_n of the interval I w.r.t. $\mathcal{A}_{\Psi(m)}$ and the number 1, and then

outputs $x = 1$ if $I \subseteq \bigcup_{\langle a,\gamma\rangle \in D_n}(a-\gamma, a+\gamma)$. Let h_m be a function the algorithm $\Psi(m)$ computes. Given a natural number m, if the domain of h_m contains the interval I, then the algorithm $\mu_I(n,m,x)$ stops with $x = 1$ for some natural number n.

Hence we can construct a procedure which enumerates all the pairs of n and m, makes $\mu_I(n,m,x)$ run on each of such pairs, and then outputs the algorithm $\Psi(m)$ if $x = 1$.

Thus the set of all functions in $\mathcal{T}$ such that the domains include I is recursively enumerable. Let $\mathcal{H}_0$ be such a set. Then $\mathcal{T}_I \preceq \mathcal{H}_0$. $\square$

For a set of recursive real-valued functions $\mathcal{T}$, we denote by $\mathcal{T}^*$ the smallest set that satisfies the following conditions:

1 If $h \in \mathcal{T}$, then $h \in \mathcal{T}^*$.

2 If $h_1, h_2 \in \mathcal{T}^*$, then $h_1 + h_2, h_1 \times h_2, h_1 \circ h_2 \in \mathcal{T}^*$.

Proposition 3 *For sets of functions $\mathcal{T}$ and $\mathcal{U}$, if $\mathcal{T} \preceq \mathcal{U}$, then $\mathcal{T}^* \preceq \mathcal{U}^*$.*

Let $\mathcal{BF}$ is the set of all basic functions we have defined in Section 3. Then $\mathcal{BF}^*$ is the set of all elementary functions whose coefficients are recursive real numbers. Note that domains of the functions $x^{\frac{1}{2}}$ and $\arcsin x$ are not unions of open intervals. Now we consider the following functions:

$$\psi_1(x) = \begin{cases} x^{\frac{1}{2}} & \text{if } x \geq 0 \\ 0 & \text{otherwise,} \end{cases} \qquad \psi_2(x) = \begin{cases} -\frac{\pi}{2} & \text{if } x < -1 \\ \arcsin x & \text{if } -1 \leq x \leq 1 \\ \frac{\pi}{2} & \text{otherwise .} \end{cases}$$

The functions ψ_1 and ψ_2 are recursive real-valued functions. Let $\mathcal{EF}$ be the set of the following functions: x, $-x$, $\frac{1}{x}$, e^x, $\log x$, $\sin x$, $\arctan x$, ψ_1, ψ_2 and c_r, where $\frac{1}{x}$ for $x = 0$, and $\log x$ for $x \leq 0$ are undefined. By Theorem 3, $\mathcal{EF}^*$ is the set of all recursive real-valued functions. By Proposition 3, it holds that $\mathcal{BF}^* \preceq \mathcal{EF}^*$ and $\mathcal{EF}^*$ is recursively enumerable. Hence we have the following theorem:

Theorem 6 *Let $\mathcal{BF}^*$ be the set of all elementary functions whose coefficients are recursive real numbers, and let I be a rational closed interval. Then the set $\mathcal{BF}^*_I$ is approximately inferable in the limit from data.*

For any datum $\langle\langle p,\alpha\rangle, \langle q,\beta\rangle\rangle$ and any rational interval I, we can determine whether $[p-\alpha, p+\alpha] \subseteq I$ or not, even if I is not closed interval. Therefore, for any presentation σ, we can take a presentation σ_I such that σ_I is a subsequence of σ and $\sigma_I[n] = \{\langle\langle p,\alpha\rangle, \langle q,\beta\rangle\rangle \in \sigma[n] \mid [p-\alpha, p+\alpha] \subseteq I\}$ for each n. If the presentation σ is of a function h from I to R, then so is the presentation σ_I. Hence we can deal with only such presentations σ_I. Now we state our final result.

Theorem 7 *Let $\mathcal{T}$ be a recursively enumerable set of recursive real-valued functions on the same rational interval I. Then the set $\mathcal{T}$ is approximately inferable in the limit from data.*

Proof. We can construct an IIM $\mathcal{M}_0$ which approximately infers $\mathcal{T}$:

```
IIM: M_0
begin
  D := ∅;   n := 1;
  repeat
    D := D ∪ {w_n};
    m := 0;   T := 0;
    while T = 0 do begin
      check2_Ψ(m)(D, x);
      if x = 0 then m := m + 1
        else T := 1
    end;
    output A_Ψ(m);
    n := n + 1
  forever
end.
```

```
Algorithm: check2_h(D, x)
begin
  x := 1;
  for each ⟨⟨p, α⟩, ⟨q, β⟩⟩ ∈ D do begin
    y := 0;   B_h(p, α, β, U);
    for each ⟨a, γ⟩ ∈ U do begin
      ⟨b, λ⟩ := A_h(⟨a, γ⟩);
      if [b − λ, b + λ] ∩ [q − β, q + β] ≠ ∅
        then y := 1
    end;
    x := xy
  end;
  output x
end.
```

The set D, an input to the algorithm $check_h(D,x)$, is $\sigma_I[n]$ for some n. Therefore, for any $\langle\langle p,\alpha\rangle,\langle q,\beta\rangle\rangle \in D$, $[p-\alpha,p+\alpha] \subseteq I$. Thus the algorithm $check2_h(D,x)$ requests a set D of data as an input, and works as follows: It outputs $x=0$ if D is not a set of data of h, and outputs $x=1$ if there is a real number $x \in [p-\alpha,p+\alpha]$ such that $[h(x)-\beta,h(x)+\beta] \cap [q-\beta,q+\beta] \neq \emptyset$ for any $\langle\langle p,\alpha\rangle,\langle q,\beta\rangle\rangle \in D$. By a similar discussion to that in Theorem 4, we can show that $\mathcal{M}_0$ approximate infers $\mathcal{T}$ in the limit from data. □

6 Concluding Remarks

We conclude this paper by discussing the precision of guesses. Data of a function h we have considered are pairs of data of numbers x and $h(x)$. Let us deal with only rational numbers as the value of x, and ignore the error bound of data of x. Then the data of the function h can be denoted by $\langle\langle x,0\rangle,\langle q,\beta\rangle\rangle$. In this paper, we have allowed the IIM to receive sufficient data, which is not natural when we try to apply the IIM to practical problems. There should be some precision $\delta > 0$ which satisfies $2\beta \geq \delta$ for any data $\langle\langle x,0\rangle,\langle q,\beta\rangle\rangle$ of the function h. Therefore, we need to discuss the inferability from such data.

Let D_δ be the set of all data $\langle\langle x,0\rangle,\langle q,\beta\rangle\rangle$ of a function h from S to R such that $x \in S \cap Q$, $h(x) \in [q-\beta,q+\beta]$ and $2\beta \geq \delta$. Here we modify the notion of presentations. We define a presentation of h as an infinite sequence $w_1, w_2, \cdots$ of data

of h such that $\{w_1, w_2, \cdots\} = D_\delta$.

Let $\mathcal{T}$ be a set of recursive real-valued functions on the same rational interval I. Even if the set $\mathcal{T}$ is approximately inferable in the limit from data, the set $\mathcal{T}$ is not approximately inferable in the limit from D_δ. However, an IIM $\mathcal{M}$, which approximately infers the set $\mathcal{T}$ in the limit from data, converges to an algorithm $\mathcal{A}$ in the limit from D_δ. Let h_0 and h be a target function in $\mathcal{T}$ and a function the algorithm $\mathcal{A}$ computes, respectively. Then we have $|h(x) - h_0(x)| \leq \frac{3}{2}\delta$ for any $x \in I$. Thus the set $\mathcal{T}$ can be learned in the limit from D_δ within the precision $\frac{3}{2}\delta$.

Apsītis, Freivalds and Smith[2] dealt with arbitrary computable functions of recursive real numbers to model the learning by example of real functions. They introduced the maximum metric that determines the distance of functions, and then showed that the set of all continuous functions on the interval $[0, 1]$ is EX-learnable within precision ϵ for any $\epsilon > 0$. They also showed that the set of all computable non-increasing functions defined on the interval $(0, 1]$ cannot be learned in the limit within integral precision ϵ for any $\epsilon > 0$, even if they allow the integral precision (i.e., the Lebesgue metric) instead of the maximum metric. The set we considered here does not contain all the continuous functions. However, our set contains all the elementary functions with recursive real coefficients, and hence it is a rich set of functions. Furthermore, we have shown that any set of recursive real-valued functions on the same rational interval is approximately inferable in the limit from data, as far as the set is recursively enumerable, where the interval needs not be closed.

References

[1] G. Alefeld and J. Herzberger. *Introduction to Interval Computations.* Academic Press, London, England, 1983.

[2] K. Apsītis, R. Freivalds, and C. H. Smith. On the inductive inference of real valued functions. In *Proceedings of the Eighth Annual Conference on Computational Learning Theory*, pp. 170–177, 1995.

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

[4] A. Grzegorczyk. Computable functionals. *Fundamenta Mathematicae*, Vol. 42, pp. 168–202, 1955.

[5] A. Grzegorczyk. On the definitions of computable real continuous functions. *Fundamenta Mathematicae*, Vol. 44, pp. 61–71, 1957.

[6] D. Haussler. Decision theoretic generalizations of the PAC model for neural net and other learning applications. *Information and Computation*, Vol. 100, pp. 78–150, 1992.

[7] K. Ko. On the definitions of some complexity classes of real numbers. *Mathematical Systems Theory*, Vol. 16, pp. 95–109, 1983.

[8] K. Ko. *Complexity Theory of Real Functions.* Birkhäuser, 1991.

[9] R.E. Moore. *Interval Analysis.* Prentice-Hall, 1966.

[10] A. Mostowski. On computable sequences. *Fundamenta Mathematicae*, Vol. 44, pp. 37–51, 1957.

[11] R. Péter. *Recursive functions* (*third revised edition*). Academic Press, New York and London, 1967.

[12] M. B. Pour-El and J.I. Richards. *Computability in Analysis and Physics.* Springer-Verlag, 1988.

[13] H.G. Rice. Recursive real numbers. In *Proceedings of the American Mathematical Society* Vol. 5, pp. 784–791, 1954.

[14] R.M. Robinson. Review of "R. Péter : 'Rekursive funktionen', Académiai Kiadó (Akademisher Verlag), Budapest, 1951. *Journal of Symbolic Logic*, Vol. 16, pp. 280–282, 1951.

[15] A.M. Turing. On computable numbers, with an application to the Entscheidungsproblem. In *Proceedings of the London Mathematical Society* Vol. 42, pp. 230–265, 1936.