数理解析研究所講究録
754巻 1991年 1-14

On Inferability of Functions by Derivatives from A Finite Number of Input-Output Samples

Teruyasu Nishizawa
Faculty of Economics, Niigata University

§1. Introduction and Notations

The key notion developed in this paper is the concept 'Finite Derivative Closure' formulated in the papers [1],[2] and [3], which is applicable to synthesis of recursive programs from given (complete or incomplete) expressions of functions, and is also applicable to the problem of identifying functions from a finite number of input-output samples.

The concept 'Finite Derivative Closure' is obtained by extending and generalizing the process in which a finite automaton recognizing a regular set is synthesized by taking all derivatives of a regular expression of the regular set.

In this paper, we rewrite the essense of the content of the papers [1], [2] and [3], and give more smart and inproved formulation, introducing the concept of 'inferability' of functions by derivatives and adding some new results.

In this section, we give some basic notations needed throughout this paper.

(1) Let ϕ be a partal functions of $A \to B$. Then the domain of ϕ, denoted by $D(\phi)$, is the set $\{x \in A ; (\exists y \in B)\ \phi(x)=y\}$. For arbitarary $X \subset A$, $y \in B$ and $Y \subset B$, we set

$\phi(X)=\{\phi(x) ; x \in X\}$,

$\phi^{-1}(y)=\{x \in A ; \phi(x)=y\}$ and

$\phi^{-1}(Y)=\{x \in A ; \phi(x) \in Y\}$.

'ϕ is one-to one (or 1-1)' means $(\forall x, y \in D(\phi))\ [\phi(x)=\phi(y) \to x=y]$.

'ϕ is onto' means $\phi(A)=B$.

(2) For partial functions ϕ and ψ of $A \to B$, $\phi \sqsubset \psi$ means that $D(\phi) \subset D(\psi)$ and $(\forall x \in D(\phi))\ \phi(x) = \psi(x)$. If $\phi \sqsubset \psi$ holds, we say that ϕ is a restriction of ψ, or ϕ is an extension of ψ.

(3) Composition of partial functions is denoted by an operator $\circ$. Thus, for example, if g is a partial function $A^n \to A$ and h is a partial function of $A \to B$ then $h \circ g$ denotes a partial function of $A^n \to B$ such that $(h \circ g)(x_1, \ldots, x_n) = h(g(x_1, \ldots, x_n))$.

(4) For a nonempty set W and a nonnegative integer m, $\Pi_m(W)$ denotes the family of all partial functions of $W \to W^m$. W^0 is constituted of one trivial element. So, we identify an element of $\Pi_0(W)$ with a subset W, that is, $\phi \in \Pi_0(W)$ is identified with $D(\phi)$.

For $\vec{f} \in \Pi_m(W)$, we define the dimension $d(\vec{f})$ of $\vec{f}$ as the m.

We put $\Pi(W) = \bigcup_{m=0}^{\infty} \Pi_m(W)$. For $\vec{f} \in \Pi(W)$, $\vec{f} = (f_1, \ldots, f_m)$ means that

(a) $d(\vec{f}) = m$

(b) for each $i = 1, \ldots, m$, $f_i \in \Pi_1(W)$ and $D(f_i) = D(\vec{f})$

(c) $(\forall x \in D(\vec{f}))\ \vec{f}(x) = (f_1(x), \ldots, f_m(x))$.

(5) Throughout this paper we use the upper-case W and A as nonempty sets and lower-case i, j, k, l, m and n as nonnegative integers.

§2. Finite Derivative Closure of Unary Partial Functions

First we give some definitions.

<u>Definition 1</u> (Differentialbility and Derivatives)

Let ω be a partial function of $W \to A$ and $\vec{f}$ be in $\Pi_m(W)$.

We say that ω is differentialble by $\vec{f}$ if and only if, for each x and y in $D(\vec{f}) \cap D(\omega)$, it holds that $\vec{f}(x) = \vec{f}(y)$ implies $\omega(x) = \omega(y)$.

When ω is differentiable by $\vec{f}$, a partial funcion $\partial_{\vec{f}}\, \omega$ of $W^m \to A$, called

the derivative of ω by $\vec{f}$, is defined by $D(\partial_{\vec{f}}\,\omega)=\vec{f}(D(\vec{f})\cap D(\omega))$ and $(\forall x\in D(\vec{f})\cap D(\omega))\ [\partial_{\vec{f}}\,\omega(\vec{f}(x)=\omega(x)]$.

By the above definition 1, the following proposition is obvious.

<u>Proposition 1</u>

Let $\vec{f}$ be in $\Pi(W)$ and ω be a partial funcion of $W\rightarrow A$.

(1) If $D(\omega)\cap D(f)$ is empty then ω is differentiable by $\vec{f}$ and $D(\partial_{\vec{f}}\,\omega)=\phi$.

(2) If ω is differentialble by $\vec{f}$ then, for an arbitrary $\vec{x}$ in $D(\partial_{\vec{f}}\,\omega)$, $\omega(\vec{f}^{-1}(\vec{x}))=\{\partial_{\vec{f}}\,\omega(\vec{x})\}$ holds.

(3) If $\vec{f}$ is 1-1 then an arbitraly ω is differentiable by $\vec{f}$.

(4) If $\vec{f}$ has a constant value over $D(f)\cap D(\omega)$ then ω is differentiable by $\vec{f}$ if and only if ω has also a constant value over the set $D(\vec{f})\cap D(\omega)$.

(5) If ω is a function and differentiable by an onto $\vec{f}$ then $\partial_{\vec{f}}\,\omega$ is also a function.

(6) If $d(\vec{f})=0$ then it is a special case of the above (4) and ω is differentiable by $\vec{f}$ if and only if ω has a constant value, say α . In its case, $\partial_{\vec{f}}\,\omega$ is the constant α .

<u>Definition 2</u> **(Case-splitting Transformation)**

(1) A finite subfamily F of $\Pi(W)$ is called a case-splitting transformation over W (abbreviated cst$/W$) if the $\{D(\vec{f})\,;\,\vec{f}\in F\}$ constitutes a finite division of W .

(2) Let W be a well-ordered set with a partial order $\prec$ and F be a cst$/W$. F is said to be descending if and only if, for each $\vec{f}=(f_1,\ldots,f_m)$ in F and for each $i=1,\ldots,m$, $(\forall x\in D(\vec{f}))\ f_i(x)\precneqq x$ holds.

(3) We say that a partial function ω of $W\rightarrow A$ is differentiable by a cst F over

W if ω is differentiable by each $\vec{f}$ in F .

Here we note that, if a partial function ω of $W \to A$ is differentiable by a cst F over W , then, for every $\vec{f}$ in F and each x in W , it holds that $x \in D(\vec{f})$ implies $\omega(x) = \partial_{\vec{f}}\, \omega\, (\vec{f}(x))$.

For the definition of a finite derivative closure, we give some notations.

Let A and B be nonempty sets. The family of all partial functions of $B^m \to A$ is denoted by $\Omega_m(B, A)$ and we set

(a) $\Omega(B, A) = \bigcup_{m=0}^{\infty} \Omega_m(B, A)$

(b) $\Omega_m(A) = \Omega_m(A, A)$

(c) $\Omega(A) = \Omega(A, A)$

<u>Definition 3</u> **(Finite Derivative Closure of Unary Partial Functions)**

Let H be a subset of $\Omega(A)$, E be a subset of $\Omega_1(W, A)$, $\omega_0, \omega_1, \ldots, \omega_n$ be elements of $\Omega_1(W, A)$ and $F_0, F_1, \ldots F_n$ be cst's over W . Then a system $\Gamma = (H\,;\, \omega_0, \omega_1, \ldots, \omega_n\,;\, F_0, F_1, \ldots, F_n) / E$ is called a finite derivative closure including ω_0 if and only if the following conditions hold:

(a) Every ω_j is differentiable by F_j .

(b) For each $\vec{f} \in F_j$, supposing $d(\vec{f}) = m$, there exist $\phi_1, \ldots, \phi_m \in \{\omega_0, \omega_1, \ldots, \omega_n\} \cup E$ and an $h \in H$ such that, for each $(x_1, \ldots, x_m) \in D(\partial_{\vec{f}}\, \omega_j)$, $\partial_{\vec{f}}\, \omega_j(x_1, \ldots, x_m) = h(\phi_1(x_1), \ldots, \phi_m(x_m))$ holds. (The equslity is denoted by $\partial_{\vec{f}}\, \omega_j \sqsubset h \circ (\phi_1, \ldots, \phi_m)$.)

When the all F_j 's are descending, Γ is said to be halting.

Considering H and E are families of known partial functions (that is, we know how to compute them), the above system Γ is a recursive program computing $\omega_0, \omega_1, \ldots, \omega_n$, and if Γ is halting then the program is halting. In this case we have functions $\tilde{\omega}_0, \tilde{\omega}_1, \ldots, \tilde{\omega}_n$ computed by Γ such that $\omega_j \sqsubset \tilde{\omega}_j$ for $j = 0, \ldots, n$. For a given ω_0 which is a finite number of input-output samples or

is a partial function specified by some incomplete or complete expression, if we find a halting finite derivative closure including the ω_0 , it means that we success to identify the function $\widetilde{\omega}_0$ which is an extension of ω_0 or to synthesize a program computing ω_0 . So, we give the following definition.

<u>Definition 4</u> (Inferability by Derivatives)

Let H be a subfamily of $\Omega(A)$, E be a subfamily of $\Omega_1(W,A)$ and $\mathfrak{F}$ be a finite family of descenfing cst's over W . (We suppose that we know how to compute partial functions in H , E and each $F \in \mathfrak{F}$.)

A (partial) function ω of $W \rightarrow A$ is said to be inferable by derivatives with the environment $(H, E, \mathfrak{F})$ if there exists a halting finite derivative closure $\Gamma = (H\,;\omega_0, \omega_1, \ldots, \omega_n\,; F_0, F_1, \ldots, F_n)\,/E$ where $\omega_0 = \omega$ and all F_j 's are in $\mathfrak{F}$.

The above definition states nothing about the infering process of ω from a finite number of input-output samples. But when the above Γ exists and a suitable restriction ω_0' of ω with a finite domain is given, we can construct a halting finite derivative closure $\Gamma' = (H\,;\omega_0', \omega_1', \ldots, \omega_n'\,; F_0, F_1, \ldots, F_n)\,/E$ where each ω_j' is a restriction of ω_j with a finite domain, and the functions $\widetilde{\omega}_0', \widetilde{\omega}_1', \ldots, \widetilde{\omega}_n'$ computed by Γ' as extensions of $\omega_0', \omega_1', \ldots, \omega_n'$ respectively coincide with (or are extensions of) the $\omega_0, \omega_1, \ldots, \omega_n$ respectively.

For the reason why we can construct such a Γ' , it is enough to point out that the minimum derivative closure including ω' is a finite family, where the minimum closure means the family Φ defined by the following (a), (b), (c).

(a) $\omega' \in \Phi$.

(b) If a ϕ is in Φ and ϕ is differentiable by an $F \in \mathfrak{F}$ then, for each $\vec{f} \in F$, $\partial_{\vec{f}}\, \phi$ is also in Φ .

(c) Φ is the mimimum family satisfying the above conditions (a) and (b).

From the above definition and the clasical result for a method of systhesizing an

automaton by taking derivatives of a regular expression, we have immediately the following proposition.

Proposition 2

Let Σ and Δ be alphabets, $W=\Sigma^*$ and $A=\Delta$.

We set $F=\{f_\sigma ; \sigma \in \Sigma\} \cup \{f_\varepsilon\}$ where f_σ's are in $\Pi_1(W)$, f_ε is in $\Pi_0(W)$, $D(f_\sigma)=\sigma\Sigma^*$, $f_\sigma(\sigma x)=x$ and $D(f_\varepsilon)=\{\varepsilon\}$. Let $H=\{id\} \cup A$, $E=\varnothing$ and $\mathfrak{F}=\{F\}$, where id is the identity function of $A \to A$ and the subset A of H is considered as $\Omega_0(A)$.

Then all mappings of $\Sigma^* \to \Delta$ induced by Moore-type sequential machines are inferable by derivatives with $(H, E, \mathfrak{F})$. Especially so are characteristic functions of all regular sets over Σ.

It should be noted that the above F is descending and an arbitrary partial function over Σ^* is differentiable by F.

Our first nontrivial result is the following theorem concerned with the inferability of gsm mappings.

Theorem 1

Let Σ, Δ be alphabets, $W=\Sigma^*$ and $A=\Delta^*$. We take the same E, F and $\mathfrak{F}$ as in the above proposition 2. For each $z \in \Delta^*$, let h_z denote a function in $\Omega_1(A)$ such that $h_z(y)=zy$ and set $H=\{h_z ; z \in \Delta^*\} \cup \{\varepsilon\}$, where ε is an empty word considered as an element of $\Omega_0(A)$.

Then all gsm mappings of $\Sigma^* \to \Delta^*$ are inferable by derivatives with $(H, E, \mathfrak{F})$.

[Proof] Let ω be a gsm mapping of $\Sigma^* \to \Delta^*$ computed by a gsm $S=(\Sigma, Q, q_0, \delta, \Delta, \lambda)$, where Q is a state set, q_0 is an initial state, δ is a state-transition function and λ is an output function. Supposing $Q=\{q_0, q_1, .. q_n\}$, we take functions $\omega_0, \omega_1, \ldots, \omega_n$ such that, for $i=0, 1, \ldots, n$, $\omega_i(x)=\widetilde{\lambda}(q_i, x)$, where the function $\widetilde{\lambda}$ of $Q \times \Sigma^* \to \Delta^*$ is

defined as $\widetilde{\lambda}(q,\varepsilon)=\varepsilon$ and $\widetilde{\lambda}(q,x\sigma)=\widetilde{\lambda}(q,x)\cdot\lambda(\widetilde{\delta}(q,x),\sigma)$ for $q\in Q$, $x\in\Sigma^*$ and $\sigma\in\Sigma$.

Then $\Gamma=(H;\omega_0,\omega_1,\ldots,\omega_n;F,F,\ldots,F)/\phi$ is a finite derivative closure and $\omega=\omega_0$. Here it should be noted that $\delta(q_i,\sigma)=q_i$ and $\lambda(q_i,\sigma)=z$ means that $\partial_{f_0}\omega_i\sqsubset h_z\circ\omega_j$.

We show an algorithm giving a halting finite derivative closure including given ω_0 specified by a finite number of inut-output samples of a gsm mapping.

```
begin YF={ω0} ; AF=∅;
  while YF≠∅ do
    begin take an ω in YF ;
      YF:=YF−{ω} ; AF:=AF∪{ω} ;
        for all f in F do
          begin take ∂fω ;
            search a z∈Δ* and an u∈AF∪YF such that
            ∂fω ⊏hz ∘u ;
            if there is no such a z and an u
            then YF:=YF∪{∂fω}
          end
    end
end
```

The variable AF stands for the set of partial functions already differentiated and the variable YF stands for the set of partial functions appearing in the computing process but not yet differentiated.

The haltness of this algorithm is obvious because of the finiteness of the domain of ω_0.

Let ω be a gsm mapping computed by a gsm with $n+1$ states.

If ω_0 is specified by all input-output samples of ω such that the length of

inputs is at most $n+2$ then the algorithm halts with $AF=\{\omega_0,\omega_1,\ldots,\omega_n\}$ and the finite derivative closure $\Gamma=(H;\omega_0,\omega_1,\ldots,\omega_n;F,F,\ldots,F)/\phi$ is a gsm computing w .

We have also the following corollary immediately.

Corollary 1

Let Σ,Δ, W, A, E, F, $\mathfrak{F}$ and h_z for $z\in\Delta^*$ are the same as in Theorem 1. Here we set $H=\{h_z;z\in\Delta^*\}\cup A$, where A is considered as $\Omega_0(A)$.

A function g of $\Sigma^*\rightarrow\Delta^*$ is inferable by derivatives with (H,E,F) if and only if there exist an gsm mapping g' and an element $y_0\in\Delta^*$ such that $g(x)=g'(x)\cdot y_0$ for every $x\in\Sigma^*$.

Next we show a theorem concerned with the inferability of linear context free languages.

For the theorem, we need the following definition.

Definition 5 (Strictly Deterministic Linear Context-Free Languages)

A linear context free grammar is said to be strictly deterministic if it satisfies the following conditions.

(a) Every productin rule is in a form $A\rightarrow\sigma B$, $A\rightarrow B\sigma$ or $A\rightarrow\varepsilon$, where A, B are nonterminals, σ is a terminal and ε is the empty word.

(b) Nonterminals are classified to two kinds. For a nonterminal A of one kind there exists no rule in a form $A\rightarrow B\sigma$, and for a nonterminal A of another kind there exists no rule in a form $A\rightarrow\sigma B$.

(c) For each nonterminal A and each terminal σ , there exists at most one B such that $A\rightarrow\sigma B$ or $A\rightarrow B\sigma$ is a rule.

Strictly deterministic linear context free languages are languages generated by strictly deterministic linear context-free grammar.

Theorem 2

Let W, H, E and f_ε be the same as in the proposition 2, where A is restricted to $\{true, false\}$. For each $\sigma \in \Sigma$, we define f_σ and g_σ in $\Pi_1(W)$ as $D(f_\sigma)=\sigma\Sigma^*$, $D(g_\sigma)=\Sigma^*\sigma$, $f_\sigma(\sigma x)=x$ and $g_\sigma(x\sigma)=x$.

Let $F^l=\{f_\sigma ; \sigma \in \Sigma\} \cup \{f_\varepsilon\}$ and $F^r=\{g_\sigma ; \sigma \in \Sigma\} \cup \{f_\varepsilon\}$, and set $\mathfrak{F}=\{F^l, F^r\}$.

Then the family of all characteristic functions of strictly deterministic linear context-free languages over Σ^* is inferable by derivatives with $(H, E, \mathfrak{F})$.

[proof] Let $G=(S,P,N,\Sigma)$ be a strictly deterministic linear context-free grammar, where S is a start symbol, P is a set of production rules and N is a set of nonterminals.

Suppose $N=\{A_0, A_1, \ldots, A_n\}$ where $A_0=S$ and take a language

$L_0, L_1, \ldots, L_n$ as $L_j=\{x \in \Sigma^* ; A_j \overset{*}{\Rightarrow} x\}$.

Let $\omega_0, \omega_1, \ldots, \omega_n$ be characteristic functions of $L_0, L_1, \ldots, L_n$ respectively.

Then clearly the system $\Gamma=(H ; \omega_0, \omega_1, \ldots, \omega_n ; F_0, F_1, \ldots, F_n)/E$ is a halting finite derivative closure, where $F_i=F^l$ if A_j has no rule in a form $A_j \rightarrow B\sigma$ and $F_i=F^r$ if A_j has no rule in a form $A_j \rightarrow \sigma B$.

Thus $L(G)=L_0$ is inferable by derivatives with $(H, E, \mathfrak{F})$.

Our third nontrivial result is a theorem which states that the family of all tree sets recognizable by deterministic tree automata of top-sown type is inferable by a finite number of positive or negative sample trees.

<u>Theorem 3</u>

Let Σ be a ranked alphabet, T_Σ be the set of all trees over Σ and $W=T_\Sigma$.

For each $\sigma \in \Sigma$, supposing $rank(\sigma)=k$, we define $\hat{\sigma} \in \Pi_k(W)$ as $D(\hat{\sigma})$ is the set of all trees in T_Σ having a root σ and

$\hat{\sigma}(\sigma<\tau_1, \ldots, \tau_k>)=(\tau_1, \ldots, \tau_k)$.

Note that if $rank(\sigma)=0$ then $d(\sigma)=0$ and $D(\hat{\sigma})=\{\sigma\}$.

We set $F=\{\hat{\sigma} ; \sigma \in \Sigma\}$ and $\mathfrak{F}=\{F\}$. Let $A=\{true, false\}$, $E=\varnothing$ and

$H=\Omega(A)$.

Then the family of all characteristic functions of tree sets over Σ recognizable by deterministic tree automata of top-down type is infererable by derivatives with $(H, E, \mathfrak{F})$.

[proof] Let M be a deterministic tree automaton of top-down type and Q be the set of states of M .

Supposing $Q=\{q_0, q_1, \ldots, q_n\}$, we take tree sets $L_0, L_1, \ldots, L_n$ where L_i is recognizable by M if M starts on the root of trees with the state q_i .

Let $\omega_0, \omega_1, \ldots \omega_n$ be characteristic functions of $L_0, L_1, \ldots, L_n$ respectively. Then $\Gamma=(H\,;\omega_0, \omega_1, \ldots, \omega_n\,; F, F, \ldots, F)/E$ is a halting finite derivative closure, and L_0 is the tree set recognizable by M . So, all tree sets recognizable by deterministic tree automaton of top-down type are inferable by derivatives with $(H, E, \mathfrak{F})$.

It should be noted that, in Γ , state-transition $(q_i, \sigma) \rightarrow (q_{j_1}, \ldots q_{j_k})$ corresponds to the formula $\partial\hat{\sigma}\omega_i \sqsubset h \circ (\omega_{j_1}, \ldots, \omega_{j_k})$ for a suitable Boolean function h .

§3. Finite Derivative Closure of Partial Functions of Multi-Arguments

Now we treat partial functions of multi-arguments. Of cause a partial function of multi-arguments is considered as a partial function of one vector arugument, and so it seems that the multi-arguments case is included in the unary case. But such a treatment of multi-arguments cannot success to construct a finite derivative closure for even simplest function of two arguments, for example, the 'append' in LISP functions.

So we need a new formulation for multi-arguments case.

<u>Definition 6</u> (Derivatives of Partial Functions of Multi-Arguments)

Let ω be in $\Omega_m(W, A)$ and $\vec{f}$ be in $\Pi_k(W)$.

We say that ω is differentiable by $\vec{f}$ w.r.t. the i -th argument if and only if, for arbitrary $x, y \in D(\vec{f})$, it holds that $\vec{f}(x)=\vec{f}(y)$ implies $\omega(x_1,\ldots x_{i-1}, x, x_{i+1},\ldots, x_m) = \omega(x_1,\ldots, x_{i-1}, y, x_{i+1},\ldots, x_m)$ for each $(x_1,\ldots, x_{i-1}, x_{i+1},\ldots, x_m) \in W^{m-1}$.

The derivative $\partial\vec{f}^{(i)}\omega$ of ω by $\vec{f}$ w.r.t. the i -th argument is defined as a partial function in $\Omega_{m-1+k}(W, A)$ such that $\partial\vec{f}^{(i)}\omega(x_1,\ldots, x_{i-1}, \vec{f}(x), x_{i+1},\ldots x_m) = \omega(x_1,\ldots, x_{i-1}, x, x_{i+1},\ldots, x_m)$, where the value of the left-hand side is defined when and only when the value of the right-hand side is defined.

We say that ω is differentiable by a cst F over W w.r.t. the i -th argument if and only if ω is differentiable by all $\vec{f}$'s in F w.r.t. the i -th argument.

<u>Definition 7</u> **(Finite Derivative Closure of Partial Functions of Multi-Arguments)**

Let H be a subfamily of $\Omega(A)$, E be a subfamily of $\Omega(W, A)$, $\omega_0, \omega_1,\ldots,\omega_n$ be in $\Omega(W, A)$, $F_0, F_1,\ldots, F_n$ be cst's over W and $i_0, i_1,\ldots, i_n$ be positive integers.

The system $\Gamma = (H; \omega_0, \omega_1,\ldots,\omega_n; F_0^{(i0)}, F_1^{(i1)},\ldots, F_n^{(in)})/E$ is called a finite derivative closure (including ω_0) if and only if, for each $j=0, 1,\ldots, n$, ω_j is differentiable by F_j w.r.t. the i_j -th argument and, supposing $\omega_j \in \Omega_m(W, A)$, $\vec{f} \in F_j$, $i=i_j$ and $k=d(\vec{f})$, there exist an $h \in \Omega_{m-1+k} \cup H$ and $\phi_1,\ldots,\phi_k \in \{\omega_0, \omega_1,\ldots,\omega_n\} \cup E$ such that

(1) if $k=0$ then $\partial\vec{f}^{(i)}\omega_j \sqsubset h$,

(2) if $k>1$ then, setting $z_l = \phi_l(x_1,\ldots, x_{i-1}, y, x_{i+1},\ldots, x_m)$ for $l=1,\ldots, k$, it holds that, when the value of $\partial\vec{f}^{(i)}\omega_j(x_1,\ldots, x_{i-1}, y_1,\ldots, y_k, x_{i+1},\ldots, x_m)$ is defined, it coincides with the value of $h(x_1,\ldots, x_{i-1}, z_1,\ldots, z_k, x_{i+1},\ldots, x_m)$. (The situation is represented by $\partial\vec{f}^{(i)}\omega_j(x_1,\ldots, x_{i-1}, y_1,\ldots, y_k, x_{i+1},\ldots, x_m) \sqsubset h(x_1,\ldots, x_{i-1}, z_1,\ldots, z_k, x_{i+1},\ldots, x_m)$.)

If all F_j 's are descending then Γ is said to be halting.

A halting finite derivative closure Γ defined as above is a recursive program computing functions $\tilde{\omega}_0, \tilde{\omega}_1, \ldots, \tilde{\omega}_n$ which are extensions of $\omega_0, \omega_1, \ldots, \omega_n$ respectively, if we know how to compute partial functions in H, E and $F_0, F_1, \ldots, F_n$.

In the multi-arguments case, inferability by derivatives with an environment is defined as the same as in the unary case.

But algorithms for constructing finite derivative closure including a given partial function specified by a finite number of input-output samples will be too complecated for a formal description.

Required algorithms will be able to be given as real big-size computer programs using many heuristics.

Here we show by a few simple examples that how such an algorithm should work.

We consider pure-LISP functions.

Let W_0 be a set of atoms, W_1 be the set of all lists constructed over W_0 , $W_1^+ = W_1 - \{\varepsilon\}$ where ε denotes the *nil* and $W = W_1 \cup W_0$.

Suppose that $A = W \cup \{\text{ integers}\} \cup \ldots$, $E \subset \Omega(W, A)$ and $H \subset \Omega(A)$, and that they contain required functions in the bellow examples.

We define c in $\Pi_2(W)$ as $D(c) = W_1^+$ and $c(x) = (car[x], cdr[x])$, λ in $\Pi_0(W)$ as $D(\lambda) = \{\varepsilon\}$ and t in $\Pi_0(W)$ as $D(t) = W_0$.

We set $F = \{c, \lambda, t\}$ and set $\mathfrak{F} = \{F\}$.

F is descending and all partial functions in $\Omega(W, A)$ which treat all atoms uniformly are differentiable by F w.r.t. an arbitrary argument.

In the following examples we concern with the inferability by derivatives with the environment $(H, E, \mathfrak{F})$.

<u>Example 1</u> (append)

Suppose $\omega = append$. ω has two arguments. Firstly we have $\partial_c^{(1)} \omega(car[x], cdr[x], y) = \omega(x, y) = cons[car[x]; \omega(cdr[x], y)]$, that is,

$\partial_c{}^{(1)}\omega(car[x], cdr[x], y) = h(e_1(car[x], y), \omega(cdr[x], y), y)$,

where h and e_1 are taken as $h(x, y, z) = cons[x; y]$ and $e_1(x, y) = x$.

Secondly we have $\partial_\lambda{}^{(1)}\omega(y) = \omega(\varepsilon, y) = id(y)$ where id is the identity function of $W \to W$.

Thirdly we have, $\partial_t{}^{(1)}\omega(y) = \omega(atom, y) = undefined$.

Thus, supposing $H \ni id$, h, '$undefined$' and $E \ni e_1$,

$\Gamma = (H; \omega; F^{(1)})/E$ is a halting finite derivative closure and hence the *append* is inferable by derivatives with $(H, E, \mathfrak{F})$.

<u>Example 2</u>

Let $\mathbb{N}$ be the set of all nonnegative integers and take a function ω of $W^2 \to \mathbb{N}$ such that the value of $\omega(x, y)$ is the product of the numbers of atoms except nil occruing in x and y .

We set $\omega_0 = \omega$.

Firstly we have

$\partial_\lambda{}^{(1)}\omega_0(y) = \omega_0(\varepsilon, y) = 0$, that is, $\partial_\lambda{}^{(1)}\omega_0 \sqsubset n$ where n is a function of $W \to \mathbb{N}$ such that $n(y) = 0$.

Secondly we have

$\partial_t{}^{(1)}\omega_0(y) = \omega_0(atom, y)$ and we set $\omega_t = \partial_t{}^{(1)}\omega_0$.

Thirdly we have

$\partial_c{}^{(1)}\omega_0(car[x], cdr[x], y) = \omega_0(car[x], y) + \omega_0(cdr[x], y)$, that is,

$\partial_c{}^{(1)}\omega_0(x_1, x_2, y) = h'(\omega_0(x_1, y), \omega_0(x_2, y))$ where h' is a function of $\mathbb{N}^2 \to \mathbb{N}$ such that $h'(l, m) = l + m$.

Finally we take derivatives of ω_t and we have

$\partial_\lambda\omega_t = \omega_t(\varepsilon) = \omega_0(atom, \varepsilon) = 0$,

$\partial_t\omega_t = \omega_t(a) = \omega_0(b, a) = 1$ where a, $b \in W_0$ and

$\partial_c\omega_t(car[x], cdr[x]) = \omega_t(x) = \omega_0(atom, x)$

$= \omega_0(atom, car[x]) + \omega_0(atom, cdr[x])$

$= h'(\omega_t(car[x]), \omega_t(cdr[y]))$, that is,

$\partial_\lambda\omega_t = 0$, $\partial_t\omega_t = 1$ and $\partial_c\omega_t \sqsubset h \circ (\omega_t, \omega_t)$.

Thus, supposing $H \ni h'$, 0 and $E \ni n$, $\Gamma = (H; \omega_0, \omega_t; F^{(1)}, F)/E$ is a halting finite derivative closure and hence the ω_0 is inferable by derivatives with $(H, E, \mathfrak{F})$.

References

[1] T.Nishizawa, 'On Finite Derivative Closures of Partial Functions over Well-Founded Sets' (in Japanese), Koukyue-roku (research report) of Research Institute of Mathematical Science, No. 625, Kyoto University, 1987.

[2] T.Nishizawa, 'Finite Derivative Closure of Partial Functions and A Method of Inference', Bull. of Information and Cybernetics, vol. 23, No. 1 and 2, Research Association of Statistical Science, 1988.

[3] T.Nishizawa, 'On Identifying Functions Using Derivatives of Partial Functions by Partial Functions' (in Japanese), technical report of WG of mathematical fondation in The Japanese Society for Artificial Intelligence, SIG-FAI-8904-8, 1990.