INDUCT IUE INFERENCE ON THE BASE OF FIXED POINT THEORY

T. 6ERGELY and ZS. SZRBO

Applied Logic Laboratory
P.O.B.146.Budapest 112.Hungary 1502.

ASTRACT

Inductive reasoning is an important way to generate knowledge from the propositions reflecting facts or directly from data We intend to extract new knowledge in the form of definitions given as fixed point equations. An appropriate fixed point theory is outlined in favour of our aim. This theory suggests the so called generative fixed point equation system as the form of hypothesis that we are looking for. The defining formulas in these equations use only bounded quantifiers but the relation to be defined may also negatively occur. An inductive inference method is presented as to find the hypothetical equation system from the increasing set of experimentel data such that its solutions would fit all the information

1.INTRODUCTION

Inductive reasoning is an important way to generate knowledge from the propositions reflecting fects e.g. propositions describing experimental results. The knowledge to be extracted from factual propositions or date is to provide e general characterization of these factual data. This characterization may be a law or a regularity which describes some basic interrelation among factual data (cf (Finn and Gergely 1984), (Gergely and Szabo 1986)). The process of discovering this regulerlty is but e leerning process which aims to acquisit new declarative knowledge (cf.(Cerbonell, Michalski and Mitchell 1983)). This new declarative knowledge may be also in the form of a general rule generated from examples that correspond to factual data (cf.(Ang1uin and Smith 1983)) Therefore the subject of this paper 1s connected with the field of machine learning and inductive inference which belong to the mainstream of Al research As it is known (e.g. from (Angluin and Smith 1983)) the majority of Inductive methods defines the regularity in question as e recursive function From the point of view of logic this means that these methods use the standard model of arithmetics as to model the reality and data obtained about the letter and they select one recursive function and identify the regularity with this function However this approach is not the best if we eim to derive new knowledge. Therefore we think ebout such on approach which approximate the regularity step by step without the preconception of recurslvlty. The selected recursive function 1s a new one within the standard model of arithmetics. To have it as a part of our knowledge it should be defined by the use of a formal language, i.e. the regularity should be defined by axioms obtained during the inductive procedure. These axioms already extend our knowledge supposed also to be axiomatical 1y given.

A definition is of the form $R(x)\leftrightarrow\varphi(x)$ where R is the new symbol to be defined and φ is the defining formula, if φ doesn't contein the new symbol R then the definition is seid to be explicit, otherwise it is implicit. The implicite definitions are the so called fixed point equations Depending on the form of the defining formula there exists en effective procedure to unfold the defined symbol It provides the latter with a fixed point of the equation which is but a new descriptive knowledge obtained from factual data We aim to develop an inductive method which builds up the implicit definition of the regularity or of the general rule in question in the form of a fixed point equation Of course, we do not expect to design a method which provides the searched fixed point equation at once, but we are looking for a procedure which approaches the definition of regularity step by step by discovering its different properties

This agrees with the fact that all fixed points (solutions) of a fixed point equation are of the same properties, e.g. they are symmetric The fixed point equation $R(x,y)R(y,x)$ is such an example that the solutions are but the symmetric relations If such data arrive which don't fit the definition of the property established earlier then the description is either to be modified or rejected Thus, step by step, we may find the constituent properties that the required fixed point equation can be built up from We aim to develop a method which reaches the required fixed point equation from the increasing set of factual data considered as elements of some fixed point and by combining the equations corresponding to these properties into the final equation. We require from the final equation that it fits all factual data even those which may appear sometimes In the future. If it is so then the hypothesis can be falsified. Note that the definition, which describes the regularity, is obtained in a certein moment of time on the base of the data arrived up to this moment. However the new data arrived after that moment may not fit the hypothesis, i.e. the letter will be falsified. Of course, we cannot expect a hypothesis in the definition form to be verified by using only experimental data. Verification is possible only w.r t. some knowledge. Since we are interested in the hypothesis generated on the base of pure experimental data we require felsiflcabllity. These formulas are just the universally quantified formulas. They can be transformed into a special kind of fixed point equation systems on the right hand side of which the defining formula may use only existential quantifiers beyond bounded quantifiers. So we are looking for the hypothesis in the form of implicit definition expressed as such a type of fixed point equation. Having a fixed point equation on effective way is required to find its solutions.

It Is known that if the defining formula contains the definable symbol only positively (i.e. not in the scope of negation) and it has only existential quantifiers then there is an effective method to obtain the least solution of the equation (see (Oergely and Ury 1983)). However, the restriction on the positive occurence of the definable relation symbol is too strict for us since negative statements ere also informative in inductive inferences. I.e. if e relation doesn't hold In a given situation then even this has information content for us. Therefore we need a fixed point theory which allows to work also with negation

We develop such a theory in the next section. The inductive inference method of building up appropriate fixed point equation system as hypothesis is given in Section 3. The defining formules in this equation system will use only bounded quantifiers. Note that this Is a subcase of the exlstentlally closed defining formulas. Taking into account the framework of recursive function theory we note thet oil primitive recursive relations can be defined by the formulas of the considered case.(see(0ergely and Szab6 1986)) For the majority of problems with practical interest the relations to be found can be defined by such a hypothetical implicit definition An example is given in Section 4.

2. QUASI EQUATION SYSTEMS

We use a first order language of a given alphabet (similarity type) σ and a given constructive model by which we interpret the factual data. The given similarity type is extended by new symbols which we correspond to the obtained solution of the hypothetical fixed point equation.

For the sake of simplicity, further on, we use the language of arithmetics which contains two binary relation symbols equality (=) and "less then" (<) and two constant symbols 0 and 1 and three binary function symbols +, · and exponantion ↑. The standard model $\mathfrak{N}$ of arithmetics is considered as the required constructive model where we denote the corresponding functions and relations similarly. We also allow the usage of other recursive functions and relations on natural numbers N (={0,1,2,...}). Let us see how new relation symbols can be introduced into the signature. Let $R_1,...,R_n$ denote new relation symbols of arity $k_1,...,k_n$, respectively. We write $\varphi(x)$ to indicate that the free variables of φ are among $x=x_1,...,x_m$. If we write φ without x then it means that it has no free variables.

A formula $\Theta(x,y)$ is called <u>finite bounding formula</u> w.r.t. x if for any fixed $b \in N^m$ the set $\{ a \mid \mathfrak{N} \models \Theta(x,y)[a,b] \}$ is finite. Let BF be a fixed recursive set of finite bounding formulas. We introduce shorthands

$\exists x[\Theta(x,y)]\varphi(x,y)$ for $\exists x(\Theta(x,y) \,\&\, \varphi(x,y))$

$\forall x[\Theta(x,y)]\varphi(x,y)$ for $\forall x(\Theta(x,y) \rightarrow \varphi(x,y))$

where we suppose that $\Theta(x,y) \in BF$, i.e. we bound the domain of quantification. Therefore further on we use the notations $\exists x[\Theta(x,y)]$ and $\forall x[\Theta(x,y)]$ as quantifiers and call them bounded quantifiers.

<u>DEFINITION 1.</u>

A, The set $\Delta_0(R_1,...,R_n)$ is the least set of formulas that

(i) contains atomic formulas among which relation symbols $R_1,...,R_n$ to be defined can also occur;

(ii) is closed under connective &, V, ¬;

(iii) is closed under bounded quantifiers $\exists x[\Theta(x,y)]$ and $\forall x[\Theta(x,y)]$ for any $\Theta(x,y) \in BF$.

B, The set $\Sigma_1(R_1,...,R_n)$ ($\Pi_1(R_1,...,R_n)$) is the least set of formulas that

(i) contains $\Delta_0(R_1,...,R_n)$ as subset;

(ii) is closed under connectives &, V;

(iii) is closed under bounded quantifiers;

(iv) is closed under existential (universal) quantifier.

A <u>total solution</u> of a formula $\eta \in \Pi_1(R_1,...,R_n)$ is said to be a set of interpretations $\varrho_1,...,\varrho_n$ of $R_1,...,R_n$ in $\mathfrak{N}$ for which $(\mathfrak{N},\varrho_1,...,\varrho_n) \models \eta$ holds. Let $TS(\eta)$ be the set of all total solution of η.

The <u>configuration space</u> of $R_1,...,R_n$ denoted as $CS(R_1,...,R_n)$ consist of all pairs (i,a), where i=1,...,n and $a \in N^{k_i}$. Note that $CS(R_1,...,R_n)$ provides all the possible interpretations of the relation symbols $R_1,...,R_n$ on N. Let

$WD^+(\eta) := \{(i,a) \in CS(R_1,...,R_n) \mid$ for all $(\varrho_1,...,\varrho_n) \in TS(\eta)$ $(\mathfrak{N},\varrho_1,...,\varrho_n) \models \eta \}$

$WD^-(\eta) := \{(i,a) \in CS(R_1,...,R_n) \mid$ for all $(\varrho_1,...,\varrho_n) \in TS(\eta)$ $(\mathfrak{N},\varrho_1,...,\varrho_n) \not\models \eta \}$

Let $WD(\eta) := WD^+(\eta) \cup WD^-(\eta)$ denote the <u>well-defined</u> part of $CS(R_1,...,R_n)$ defined by η.

THEOREM 1. There exist an algorithm which, for any formula $\eta \in \Pi_1(R_1,..,R_n)$, terminates if $TS(\eta)=\emptyset$, otherwise it enumerates $WD^+(\eta)$ and $WD^-(\eta)$.

Let $\varphi_i(z_i), \psi_i(z_i) \in \Sigma_1(R_1,..,R_n)$, where z_i are k_i long variable tuples.

$$\begin{array}{l} R_1(z_1) \leftarrow \varphi_1(z_1) \\ \neg R_1(z_1) \leftarrow \psi_1(z_1) \\ \vdots \\ R_n(z_n) \leftarrow \varphi_n(z_n) \\ \neg R_n(z_n) \leftarrow \psi_n(z_n) \end{array}$$

is said to be a quasi equation system. Throughout this chapter let Ξ denote a quasi equation system. We consider it as the conjunction of the implications, where each z_i is bounded by a universal quantifier. So $\Xi \in \Pi_1(R_1,..,R_n)$ since $\varphi \leftarrow \psi = \varphi \vee \neg\psi$ and $\neg\exists x\varphi = \forall x \neg\varphi$.

Let $\varphi(x) \in \Sigma_1(R_1,..,R_n)$ with x is a tuple and $a \in N^m$ be fixed. An element (j,b) of $CS(R_1,..,R_n)$ can be influencing for $\varphi[a]$ denoted as $(j,b) \in IN(\varphi(x),a)$ iff b is such an element of N^{k_j} that $R_j(b)$ takes part in the determination of the truthvalue of the formula $\varphi[a/x]$.

DEFINITION 2. For a given quasi equation system Ξ the dependency relation $\succ_\Xi$ over $CS(R_1,..,R_n)$ is defined as follows:
$(i,a) \succ_\Xi (j,b)$ iff $(j,b) \in IN(\varphi_i(z_i),a)$ or $(j,b) \in IN(\psi_i(z_i),a)$.

DEFINITION 3. The quasi equation system Ξ is reduced to an irreflexive partial ordering $\succ$ over $CS(R_1,..,R_n)$ iff for all $(i,a),(j,b) \in CS(R_1,..,R_n)$ if $(i,a) \succ_\Xi (j,b)$ then $(i,a) \succ (j,b)$.

The ordering $\succ$ allows us not to define the truthvalue of $R_i(a)$ while defining the truthvalue of $R_j(b)$, if $(i,a) \succ (j,b)$.

We also need that the formulas $\varphi_i(z_i)$ and $\psi_i(z_i)$ (i=1,..,n) of Ξ not be true at the same time i.e. $\varphi_i(z_i)$ and $\psi_i(z_i)$ are to be disjunct.

THEOREM 2. If a quasi equation system Ξ is reduced w.r.t. an irreflexive partial ordering and it satisfies the disjunction property, then $TS(\Xi) \neq \emptyset$.

THEOREM 3. There exists an ω-type well-ordering $\succ$ over $CS(R_1,..,R_n)$ which is explicitly definable by an Δ_0-formula and for which, for any formula $\eta \in \Pi_1(R_1,..,R_n)$, a quasi equation system Ξ can be constructed that is reduced w.r.t. this $\succ$ and it satisfies the disjunction condition, such that if $TS(\eta) \neq \emptyset$ then $TS(\eta)=TS(\Xi)$.

Let $Q_0(R_1,..,R_n)$ denote the set of quasi equation systems which are reduced w.r.t. some irreflexive partial ordering $\succ$ and the formulas $\varphi_i(z_i), \psi_i(z_i)$ of which belong to $\Delta_0(R_1,..,R_n)$. Since the implications $R_i(z_i) \leftarrow \varphi_i(z_i)$ and $\neg R_i(z_i) \leftarrow \neg\varphi_i(z_i)$ can be written in the form of $R_i(z_i) \leftrightarrow \varphi_i(z_i)$, it is called equivalent type equation. $EQ_0(R_1,..,R_n)$ denotes the set of those equation systems of $Q_0(R_1,..,R_n)$ which consist only of equivalent type equations and are reduced w.r.t. some well-founded ordering.

THEOREM 4. If $\Xi \in EQ_0(R_1,..,R_n)$ then it has one solution exactly, i.e. $card(TS(\Xi))=1$.

THEOREM 5. If $\Xi \in Q_0(R_1,..,R_n)$ and the ordering $\succ$ w.r.t. which it is reduced is recursive and well-founded, then $WD^+(\Xi)$ and $WD^-(\Xi)$ are recursive sets.

DEFINITION 4. An irreflexive well-founded ordering $\succ$ over $CS(R_1,..,R_n)$ is said to be natural iff for all i,j=1,..,n one of the following conditions holds:

(i) $(i,a_1,..,a_{k_i}) \succ (j,b_1,..,b_{k_j})$ for all $a_1,..,a_{k_i},b_1,..,b_{k_j} \in N$;

(ii) $(i,a_1,..,a_{k_i}) \succ (j,b_1,..,b_{k_j})$ if $a_{\gamma(i)} > b_{\gamma(j)}$ (where $\gamma(i) \in \{1,..,k_i\}$ for any i=1,..,n.)

If the ordering $\succ$ is natural then instead of $Q_0(R_1,\ldots,R_n)$ and $EQ_0(R_1,\ldots,R_n)$ we write $NQ_0(R_1,\ldots,R_n)$ and $NEQ_0(R_1,\ldots,R_n)$, respectively.

If we intend to determine the properties of the observed relations $R_1,\ldots,R_n$, then we also need some auxiliary relations $R_{n+1},\ldots,R_{n+d}$ defined by equivalent type equations, where the observed relation symbols cannot occure. Namely, we have the following, so called, generative quasi equation system :

$$\begin{array}{l} R_1 \leftarrow \varphi_1(R_1,\ldots,R_{n+d}) \\ \neg R_1 \leftarrow \psi_1(R_1,\ldots,R_{n+d}) \\ \vdots \\ R_n \leftarrow \varphi_n(R_1,\ldots,R_{n+d}) \\ \neg R_n \leftarrow \psi_n(R_1,\ldots,R_{n+d}) \\ R_{n+1} \leftrightarrow \varphi_{n+1}(R_{n+1},\ldots,R_{n+d}) \\ \vdots \\ R_{n+d} \leftrightarrow \varphi_{n+d}(R_{n+1},\ldots,R_{n+d}) \end{array}$$

DEFINITION 5. A generative quasi equation system is said to be natural (denoted their set as $NGQ_0(R_1,\ldots,R_n)$) if it is reduced w.r.t. an irreflexiv ordering $\succ$ for which $(i,a) \succ (j,b)$ iff $i,j \in \{1,\ldots,n\}$ and $(i,a) \succ_\omega (j,b)$;
or $i,j \in \{n+1,\ldots,n+d\}$ and $(i,a) \succ' (j,b)$;
or $i \in \{1,\ldots,n\}$ and $j \in \{n+1,\ldots,n+d\}$
for all $(i,a),(j,b) \in CS(R_1,\ldots,R_{n+d})$, where $\succ_\omega$ is an ω-type ordering on $CS(R_1,\ldots,R_n)$ and $\succ'$ is a natural ordering on $CS(R_{n+1},\ldots,R_{n+d})$.

THEOREM 6. For any $\Xi \in ENQ_0(R_1,\ldots,R_n)$ there exists a $\Xi^* \in NGQ_0(R_1,\ldots,R_m)$ such that their solutions for $R_1,\ldots,R_m$ are the same, where $m \leq n$.

3. INDUCTIVE INFERENCE ALGORITHM

Let us consider now the theory of inductive inference, especially how we discover properties from observations, we generate hypotheses in the form of natural generative quasi equation systems corresponding to the given set of observations. For simplicity's sake we give a method for the recognition of only one observed relation now. It can be generalized for more easily.

Let function $\chi: N^+ \to N^m$ be the data presentation i.e. χ describes a sequence of observations, where N^+ denotes the set of all positive integers. More precisely, if we want to recognize an m-ary relation R, then we get information $R[\chi_n]$ at the n-th step, where $R[\chi_n] := \{\langle\chi(1),R(\chi(1))\rangle;\ldots;\langle\chi(n),R(\chi(n))\rangle\}$.

Since the solution of the natural quasi equations for the auxiliary relations are unique, so they are enumerable, therefore instead of giving the defining equations for auxilary relations it is sufficient to refer to them by names of a fixed encoding. Let Φ be a function, enumerating all the formulas of $\Delta_0(R)$ with free variable z, and let $\Phi(i)$ denote the i-th element of the enumeration. However, if we have some knowledge in the form of quasi equations with $\Delta_0(R)$ formulas on the right hand side concerning the observed relation, then we take these formulas at the beginning of the enumeration. Of course, the enumerated formulas are composed of encoded names of the auxiliary relations and of the symbol R of the observed relation, where, if $R(t)$ appears in a formula, then t is bounded by z, in order to satisfy the condition of the reduction.

Let "." denote the concatenation of finite sequence. Let the generated hypothesis for information $R[\chi_n]$ be characterized by H_n which is a finite sequence of subhypotheses, more exactly $H_n = (a_1^n,\partial_1^n)\ldots(a_{k_n}^n,\partial_{k_n}^n)$, where $a_i^n \in N$ and $\partial_i^n \in \{0,1\}$. H_n corresponds to the following defining quasi equations:

$$R(z) \leftarrow \bigvee_{\substack{1\leq i\leq k_n \\ \text{and} \\ \partial_i^n=1}} \Big(\Phi(a_i^n) \,\&\, \big(\mathop{\&}_{\substack{1<j<i \\ \text{and} \\ \partial_j^n=0}} \neg\Phi(a_j^n)\big)\Big)$$

$$\neg R(z) \leftarrow \bigvee_{\substack{1\leq i\leq k_n \\ \text{and} \\ \partial_i^n=0}} \Big(\Phi(a_i^n) \,\&\, \big(\mathop{\&}_{\substack{1<j<i \\ \text{and} \\ \partial_j^n=1}} \neg\Phi(a_j^n)\big)\Big)$$

Let $\Xi(H_n)$ denote this quasi equation system to H_n. Here (a_i^n,∂_i^n) is a subhypothesis where a_i^n

refers to $\varphi(a_i^n)$ and ∂_i^n determines the equations where this formula occurs positively. The occurences of negated forms of these formulas guarantee the fulfilment of the criterion of disjunctiveness.

Let T and F denote "true" and "false" respectively.

We say that $H=(a_1,\partial_1)\ldots(a_k,\partial_k)$ <u>provides</u> y for x, if for all $g\in TF(\Xi(H))$ hold $g(x)=y$, where $y\in\{T,F\}$. Let $Fr(H,i)$ denote the left <u>fragment</u> of H, of lenght i, that is $(a_1,\partial_1)\ldots(a_i,\partial_i)$. If i is greater or equal to length of H, then let $Fr(H,i)=H$. Let "e" denote the empty H i.e. if length of H is 0.

Let $X_n:=\{\mathcal{X}(i) \mid 1\le i\le n\}$. We say that H <u>matches</u> $R[\mathcal{X}_n]$ on the set $M(H,R[\mathcal{X}_n])$ if $M(H,R[\mathcal{X}_n])=\{x\in X_n \mid H \text{ provides } R(x) \text{ for } x\}$; and H <u>does not match</u> $R[\mathcal{X}_n]$ on the set $\bar{M}(H,R[\mathcal{X}_n])$ if $\bar{M}(H,R[\mathcal{X}_n])=(X_n\cap WD(\Xi(H)))\setminus M(H,R[\mathcal{X}_n])$.

The generation of hypotheses occurs similarly to the original enumeration method introduced by Gold (Gold 1967). Informally:

0, We start with the empty hypothesis (or if we have some knowledge in the form of quasi equation system with formulas from $\Delta_0(R)$, then we start with this quasi equation system as hypothesis);

1, We do not change hypothesis if provides the result correctly for the next ovservation;

2, If we have no subhypothesis from H_{n-1} for the next observation $\mathcal{X}(n)$, then we take the first new index i for which $H_{n-1}.(i,\partial)$ is consistent with $R[\mathcal{X}_n]$ for $\partial=0$ or $\partial=1$ and attach (i,∂) to the end of H_{n-1};

3, If some of the subhypothesis of H_{n-1} contradict the next observation, then we delete them while the rest of H_n contradicts $R[\mathcal{X}_n]$, and we repeat this procedure until the resulted hypothesis will be consistent with all of the observations.

We want that the hypothesis H_n be consistent with the observations at every step i.e. if $g\in TS(\Xi(H_n))$ then $g(x)=R(x)$ for all $x\in X_n$. Furthermore, we would like that the sequence of the relation sets $TS(\Xi(H_n))$ converge to the relation to be recognized.

$\hat{C}(H,i)$ will denote the set of rejected or used indices at the i-th step and belong to H. For every possible hypothesis H and for any i let $\hat{C}(H,i)$ be empty at the beginning. This $\hat{C}$ infinite matrix will help the backtrack of the procedure.

0, If we have knowledge in the form

$$R(z)\leftarrow\varphi(z)$$
$$\neg R(z)\leftarrow\Psi(z)$$

then let $\varphi(0)=\varphi(z)$ and $\varphi(1)=\Psi(z)$ as we said. So

$$H_0=\begin{cases} e & \text{if we do not have knowledge;}\\ (0,1).(1,0) & \text{otherwise.}\end{cases}$$

For $n\ge1$:

1, If $M(H_{n-1},R[\mathcal{X}_n])=X_n$, then $H_n:=H_{n-1}$ and let $\hat{C}(H_n,n):=\hat{C}(H_{n-1},n-1)$.

2, If $\mathcal{X}(n)\notin WD(\Xi(H_{n-1}))$ then $H_n:=H_{n-1}.(a,\partial)$ where $a=\min\{i\in C_n \mid \exists\partial\in\{0,1\}$ such that $X_n=$
$=M(H_{n-1}.(i,\partial),R[\mathcal{X}_n])\}$
where $C_n=N^+\setminus\hat{C}(H_{n-1},n-1)$; and

$$\partial:=\begin{cases}1 & \text{if } X_n=M(H_{n-1}.(a,1),R[\mathcal{X}_n]);\\ 0 & \text{otherwise.}\end{cases}$$

In this case let $\hat{C}(H_n,n):=\hat{C}(H_{n-1},n-1)\cup\{a\}$ and let $\hat{C}(H_{n-1},n):=\hat{C}(H_{n-1},n-1)\cup$
$\cup\{\,i\in N^+ \mid i\le a\ \&\ \bar{M}(H_{n-1}.(i,0),R[\mathcal{X}_n])=\emptyset\ \&$
$\&\ \bar{M}(H_{n-1}.(i,1),R[\mathcal{X}_n])=\emptyset\,\}$.

3, If $\mathcal{X}(n)\in\bar{M}(H_{n-1},R[\mathcal{X}_n])$ then we cannot tell which subhypothesis is to be blamed. So we start out to find the wrong subhypothesis, beginning the search with the 1-st one. Therefore we prefer to save the subhypotheses generated earlier. In this way we'll obtain H_n^- from $H_{n-1}=(a_1^{n-1},\partial_1^{n-1})\ldots(a_k^{n-1},\partial_k^{n-1})$. More precisely:

$H_n^- := H_{n,k_{n-1}}^-$ where

$$H_{n,1}^- := \begin{cases} e & \text{if } \bar{M}((a_1^{n-1},\partial_1^{n-1}),R[\chi_n])\neq\emptyset; \\ (a_1^{n-1},\partial_1^{n-1}) & \text{otherwise;} \end{cases}$$

and for $1<i\leq k_{n-1}$

$$H_{n,i}^- := \begin{cases} H_{n,n-1}^- & \text{if } \bar{M}(H_{n,i-1}^-(a_i^{n-1},\partial_i^{n-1}),R[\chi_n]\neq\emptyset \\ H_{n,n-1}^-\cdot(a_i^{n-1},\partial_i^{n-1}) & \text{otherwise.} \end{cases}$$

Furthermore

$$C_{n,1}^- := \begin{cases} \emptyset & \text{if } H_{n,1}^- = e; \\ \{a_1^{n-1}\} & \text{otherwise,} \end{cases}$$

and for $1<i\leq k_{n-1}$:

$$C_{n,i}^- := \begin{cases} C_{n,i-1}^- & \text{if } H_{n,i}^- = H_{n,i-1}^-; \\ C_{n,i-1}^- \cup \{a_i^{n-1}\} & \text{otherwise.} \end{cases}$$

Here $C_{n,i}^-$ means the set of those indices whicl we reject during the generation of H_n^- up to the i-th step. Now we attach subhypotheses to enc H_n^- which provide new correct results. More exactly:

$H_n := H_{n,\xi}^+$ if $\xi=\{i\in N \mid M(H_{n,i}^+,R[\chi_n])=X_n\}$, where $H_{n,0}^+ := H_n^-$ and for $i>0$:

$H_{n,i}^+ := H_{n,i-1}^+\cdot(a,\partial)$, where

$a := \min\{ j\in C_{n,i} \mid \exists\, \partial\in\{0,1\}$ such that

$\bar{M}(H_{n,i-1}^+\cdot(j,\partial),R[\chi_n])=\emptyset$ and

$M(H_{n,i-1}^+,R[\chi_n])\subset M(H_{n,i-1}^+\cdot(j,\partial),R[\chi_n])$ }

where $C_{n,i} := N^+\setminus(C_{n,i}^- \cup \{x\in N^+ \mid \exists\, k\leq n$

$\exists\, j<\mathrm{length}(H_{n,i-1}^+)$

such that $x\in\hat{C}(Fr(H_{n,i-1}^+,j),k)\})$.

and let

$$\partial = \begin{cases} 1 & \text{if } M(H_{n,i-1}^+,R[\chi_n])\subset M(H_{n,i-1}^+\cdot(a,1),R[\chi_n]) \\ 0 & \text{otherwise} \end{cases}$$

and let $\hat{C}(H_{n,i-1}^+,n) := (N^+\setminus C_{n,i}) \cup$

$\cup\{x\in N^+ \mid x<a$ and $\bar{M}(H_{n,i-1}^+\cdot(x,0),R[\chi_n])\neq\emptyset$

and $\bar{M}(H_{n,i-1}^+\cdot(x,1),R[\chi_n])\neq\emptyset\}$.

In the end let $\hat{C}(H_n,n) := N^+\setminus C_{n,\xi}$.

4. AN EXAMPLE

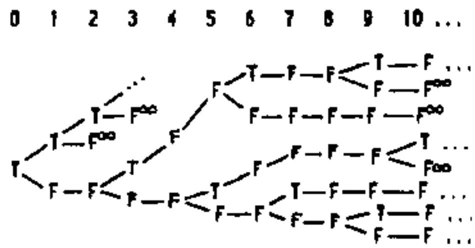
Let the data presentation χ be such that $\chi(1)=1$, $\chi(2)=0$, $\chi(3)=4$ and $\chi(4)=3$. In this order the results of the test of R are F,T,F,T i.e. we obtain $R(1)=F$, $R(0)=T$, $R(4)=F$ and $R(3)=T$. We have to recognize this relation R. Let us assume that we have knowledge about a property of R in the form of the following fixed point equation:

$$R(x) \leftrightarrow x=0 \vee \exists y[y\leq x](R(x-y)\ \&\ R(x+y))$$

This is an original positive existential fixed point equation with a bounded quantifier and is a good example to show that the original fixed point equations often are not informative enough. For example, it is difficult to decide whether $R(5)$ may be T when $R(6)=F$ for the above fixed point equation, while truthvalue of $R(3)$ influences $R(2)$ and $R(4)$ influence $R(3)$ e.t.c. However the above fixed point equation can be regarded as a universal statement with a bounded quantifier. Therefore it can be automatically transformed to the following quasi equation system Ξ reduced by usual ordering <:

$$R(z) \leftarrow z=0$$
$$\neg R(z) \leftarrow \exists x[0<2x\leq z](\neg R(z-x)\ \&\ R(z-2x))$$

This follows the natural way of thinking because this refers to the argument of R less than z. So all the solutions of Ξ can be generated as infinite paths of the following tree:



Therefore, one is able to decide whether a hypothesis is consistent with all the observations or not. It can be shown that $TS(\Xi)$ consists of all such relations which – as paths – are in one of the forms of the followings: T^{∞}, $T^n.F^{\infty}$, $(T.F^{2n})^{\infty}$ and $(T.F^{2n})^k.F^{\infty}$, where $k,n>0$. We remark that the original fixed point theory can determine the least fixed point only.

Let us see how our strategy can recognize relation R when $R(z)=\exists x(3x=z)$ for all z (this relation belongs to $TS(\Xi)$). Now we do not deal with the encoding of solutions of auxiliary relations of the natural generative quasi equations, because it would take the situation complicated and it is not necessary in this case. For simplicity's sake we assume that the enumerating function Φ enumerates formulas beginning with the following ones:

$\Phi(0) := (z=0)$, $\Phi(1) := \exists x[0<2x\leq z](\neg R(z-x)\ \&\ R(z-2x))$,

$\Phi(2) := (z>0)$, $\Phi(3) := (z=1)$, $\Phi(4) := (z>1)$,...

Formulas $\Phi(0)$ and $\Phi(1)$ are in our knowledge so $H_0 := (0,1).(1,0)$.

Let us denote the formulas $\Phi(0)$ and $\Phi(1)$ by φ and ψ, respectively.

The first information obtained is $R(1)=F$. In this case H_1 will be $(0,1).(1,0).(2,0)$ i.e

$$R(z) \leftarrow \varphi$$
$$\neg R(z) \leftarrow (z>0) \& \neg\varphi \quad \vee \Psi$$

Note $TS(\Xi(H_1))=\{g\}$, where $g(z) \leftrightarrow z=0$.

The next information obtained is $R(0)=T$. It matches the hypothesis H_1, therefore $H_2:=H_1$. The situation is the same in the case of information $R(4)=F$, so $H_3:=H_2$.

Observation $R(3)=T$ makes a change because H_3 does not match $R(3)=T$, therefore we have to reject subhypothesis $(2,0)$. Now $H_4^-:=(0,1).(1,0)$. Therefore $R(1)=F$, $R(4)=F$ and $R(3)=T$ are to be produced. $H_{4,1}^+:=(0,1).(1,0).(3,0)$ provides $R(1)=F$ and $R(4)=F$, but it does not provide $R(3)=T$, although it permits it i.e. $\Xi(H_{4,1}^+)$ does not define $R(3)$ well. $H_{4,2}^+:=(0,1).(1,0).(3,0).(4,1)$ will be correct for $R(3)=T$ too Therefore $H_4:=H_{4,2}^+$ i.e.:

$$R(z) \leftarrow \Psi \vee (z>1) \& \neg\Psi \& \neg(z=1)$$
$$\neg R(z) \leftarrow \Psi \vee (z=1) \& \neg\varphi$$

It would be ·etter if the strategy simplified the right hand side of the quasi equations e.g. it would write $(z>1)$ instead of $(x>1)\&(z=1)$

So $\Xi(H_4)$ is the following:

$$R(z) \leftarrow (z=0) \vee (z>1) \& \quad x[0<2x \leq z](R(z-x) \vee \neg R(z-2x))$$
$$\neg R(z) \leftarrow \exists x[0<2x \leq z](\neg R(z-x) \& R(z-2x)) \vee (z=1)$$

It is easy to see that the only solution of $\Xi(H_4)$ is just the relation to be recognized.

CONCLUSIIONI

We have given An inductive inference method which builds up the implicit definition of the regularity or of the general rule in question in the form of natural generative fixed point equation system with bounded quantifiers. It defines the regularity to be recognized as the solution of the hypothetical quasi equation system which can match in time the increasing set of experimental data The sequence of hypotheses generated by the developed inductive inference method will converge on a large class to the rule that 1s to be recognized. The so far described procedure can discover each primitive recursive relation provided the basic functions are Ackermann functions (cf.(Gergely and Szabo 1986)). Therefore the proposed inductive inference procedure is sufficient fx practical purposes.Ours is an asymptotic enumeration strategy end therefore it eppears to be better than the original enumeration strategies for being more stable. It is able to recognize even certain non-recursive relations in asymptotic sense (cf.(Szabo 1966)).

REFERENCES

1 Angluin, D. and C. Smith, Inductive Inference: Theory and methods. Comp. Surv. Vol.15, No.3,(1983),pp.237-269.

2 Carbonell, J. G., R. S. Michalski and T. M. Mitchell, Machine learning: a historical and methodological analisis. The AI Magazin, No,3. 1983, pp.69-79.

3 Finn, V. K. and T. Gergely, "Plausible inference + deduction" type problem solving in intellectual informational - computing systems. In Artificial Intelligence, IFAC Proceedings serie No.9, Pergamon Press, 1984, pp. 405-413.

4 Gergely T. and Zs. Szabó, Fixed point equations as hypotheses in inductive reasoning. Conf. on Artificial Intelligence and Inductive Inference held in Wendisch-Rietz, 1986.

5 Gergely T.and L. Ury, Adequate characterization of Hoare's logic. Semiotics and Informatics, Vol.22, 1983, pp.81-101.

6 Gold, E. M., Language identification in the limit,Information Control, Vol.10, (1967), pp.447-474.

7 Michalski, R. S., J. G. Carbonell and T. M. Mitschell, (eds), Machine Learning, an Artificial Intelligence Approach. Palo Alto, CA. Tioga Press, 1983.

8. Szabó, Zs.,Stratified inductive hypothesis generation. Conf. on Artifical Intelligence and Inductive Inference held in Wendisch-Rietz, 1986.