

DEFINITION THEORY AS BASIS FOR A CREATIVE PROBLEM SOLVER

H. Andráka, T. Gergely, I. Némethi
Hungarian Academy of Sciences
Budapest, Hungary

Abstract

In this paper the application of some deep theorems of mathematical logic is shown in the field of artificial intelligence. Namely, using some of the results of definition theory we give the mathematical base to systems for automatic designing. /SAD/. These systems are capable of solving constructive tasks of such kind that need some creativity from the psychological point of view. Above tasks contain the implicate description of the object to be constructed. First of all that unit is investigated at SAD which provides an explicit definition to the circumscribed object.

Introduction

One of the main directions in research of artificial intelligence is developing problem solving systems namely, systems for automatic designing /SAD/. Their practical importance is invaluable. These systems are capable to solve constructive tasks. A task is constructive if the unknown is some kind of an object of which characteristics are described in the conditions of the task. Two kinds of these are distinguished:

1. The objects to be constructed are defined explicitly:
 - a/ well-defined task
 - b/ incompletely defined task - here the conditions provide an incomplete description of the object
2. The objects to be constructed are defined implicitly.

In these tasks the objects are not named only certain expectations are given

about them.

Designing tasks appearing on the expectations of a non-professional customer belong to latter type. It can't be expected from him to give an explicit definition of a required program e.g. with the input-output relation. All he can do is to give some hints on his own expectations towards some "programlike" thing. Similar problems occur at decision making where information is implicitly connected with the question to be decided about.

A SAD capable of solving the second type constructive task, must consist of the following two basic components:

1. High-level problem defining unit which provides an explicit definition to the implicit object description
2. Solving unit which carries out the explicitly defined task

Mathematical logic and its model theory provides plenty of facilities in SAD research. In our present study we introduce the usefulness of definition theory, an intensively developing field of model-theory, from the point of view of SAD.

Basic definitions

The following triple form a language:
 $\langle \text{syntax, the set of possible worlds, validity} \rangle$, or formally $L = \langle F, M, F \rangle$.

A type t is a pair of functions, i.e.

$t = \langle t', t'' \rangle$ such that

1. $Rgt' \subseteq \omega \setminus \{0\}$ where $\omega = \{0, 1, 2, \dots\}$
2. $Rgt'' \subseteq \omega$
3. $Dot' \cap Dot'' = \emptyset$ where $\emptyset$ denotes the empty set.

Here Dot and Rgt are the domain and the range of t respectively. Dot' is the set of relation symbols and Dot'' is the set of function symbols.

In the followings we suppose that a t -type first-order language $L^t = \langle F^t, M^t, \models \rangle$

is given. Here M^t is the set of t-type structures. A t-type structure $\mathcal{A}$ is a function for which

1. $CR(\omega) \triangleq A$ is the universe of the structure
2. $CR(R) \subseteq {}^t A$ for each relation symbol $R \in Dot^t$
3. $CR(F): {}^t A \rightarrow A$ for each function symbol $F \in Dot^t$ and if $t^n(F) = 0$ then $CR(F) \in A$

Above are to be found in more details in [1]. Notations of common knowledge are also to be found there.

From now on when program is being discussed relation symbols will be used in describing the computer programs where such symbols may show what relation the input-output should have. This descriptive method provides a far more natural handling of the programs than the descriptions of programs by functions, since this approach is more close to the intuition of the non-programmer customers.

Intuitive description of SAD based on the definition theory

Let $P \in F^t$ be a set of first-order formulas which provides the knowledge of a discipline within that designing will occur. E.g. P provides the semantics of a programming language and the properties of different implemented programs.

The customer gives his requests with the help of a set of formulas Σ . This implicitly defines one or more relation symbols and/or function symbols which do not occur in $Dot^t \cup Dot^n$. In the followings without limiting generality, we suppose that Σ gives the implicit definition of only one relation symbol P . E.g. Σ gives the implicit definition of such a program of which input and output are in relation P . Let $\Sigma(P)$ denote the set of formulas defining the relation P implicitly.

Let $t_p \triangleq \langle t^1 U(p, n), t^1 \rangle$ be the extension of the type t and F^{t_p} be the syntax of the first-order language extended by relation P . Thus $\Sigma(P) \in F^{t_p}$.

To carry out the design of the required object we have to give its explicit description by a formula of F^t . So as to have the required program written in our programming language we have to find such a formula from F^t which defines P explicitly.

Let $P, P' \in Dot^t \cup Dot^n$ be two n-placed relation symbols and $\Sigma(P') \in F^{t_{P'}}$. We say that $\Sigma(P)$ defines P implicitly if

$$\Sigma(P) \cup \Sigma(P') \models \forall \bar{x}^{(n)} (P(\bar{x}^{(n)}) \leftrightarrow P'(\bar{x}^{(n)}))$$

where $\bar{x}^{(n)} \triangleq (x_1, \dots, x_n)$ and $\forall \bar{x}^{(n)} \forall x_1 \dots \forall x_n$.

We note that $\Sigma(P')$ is obtained from $\Sigma(P)$ by replacing P everywhere by P' .

We give an equivalent definition to this: Let $(\mathcal{A}, R) \triangleq CR(U\langle P, R \rangle)$ where $\mathcal{A} \in M^t$, $R \in A$. Given any models $(\mathcal{A}, R)$ and $(\mathcal{A}, R')$ for $\Sigma(P)$ then $R = R'$. This means that $\Sigma(P)$ implicitly defines P if for any model $\mathcal{A} \in M^t$ there is at most one n-placed relation R interpreting the relation symbol P such that $(\mathcal{A}, R) \models \Sigma(P)$.

Let $\varphi \in F^t$. If it has n free variables then we use the notation $\varphi[\bar{x}^{(n)}]$.

$\Sigma(P)$ explicitly defines the relation P if there is a formula $\varphi[\bar{x}^{(n)}] \in F^t$ for which

$$\Sigma(P) \models \forall \bar{x}^{(n)} (P(\bar{x}^{(n)}) \leftrightarrow \varphi[\bar{x}^{(n)}]).$$

Replacing P by φ in the set of formulas Σ everywhere we obtain $\Sigma(\varphi)$. For $\Sigma(\varphi)$ the following is true:

$$(*) \Sigma(\varphi) = \Sigma(P) \cup \{ \forall \bar{x}^{(n)} (P(\bar{x}^{(n)}) \leftrightarrow \varphi[\bar{x}^{(n)}]) \}$$

where $\equiv$ is the symbol of semantical equivalence.

What is the task of a high-level problem defining unit supposed to be at SAD? It has to find a definition $\theta \in F^t$ on the base of P knowledge to the requested expecta-

tion of the customer given by $\Sigma'(P)$ so that

$$\Gamma \cup \{ \forall \bar{x}^{(n)} (P(\bar{x}^{(n)}) \leftrightarrow \theta[\bar{x}^{(n)}]) \} \models \Sigma'(P) .$$

In other words using (*) such formula $\theta[\bar{x}^{(n)}] \in F^+$ has to be found for which $\Gamma \models \Sigma'(\theta)$.

This task results in the following questions:

1. Does a formula θ exist to Γ so $\Gamma \models \Sigma'(\theta)$. If such doesn't exist then could Γ be extended, let's say, to a Γ' ($\Gamma' \in F^+$) so as to have the required formula θ existing such that $\Gamma' \models \Sigma'(\theta)$.

This procedure can be done with the help of a system consisting of a theorem prover and of an inductive hypothesis generator. First it will examine the truth of $\Gamma \cup \{ \neg \Lambda \Sigma' \} \vdash \varphi \wedge \neg \varphi$ /here $\Lambda \Sigma'$ is obtained so that all the formulas of Σ' are linked with the "and" connective $\wedge$ /. If this isn't true then we examine whether $\Gamma' \cup \{ \neg \Lambda \Sigma' \} \vdash \varphi \wedge \neg \varphi$ is true. If this isn't so then we take another extension Γ'' etc.

We note that selecting Γ', Γ'' suppose an oriented inductivity.

The following problem belongs to here also. Is it true that all certain characteristic model $\mathcal{M} \in M^+$ of a set of formulas Γ becomes a model of $\Sigma'(\theta)$ too, i.e. is it true that $\mathcal{M} \models \Sigma'(\theta)$.

Let us suppose that the existence of $\theta \in F^+$ is proved or that taking the risk of a possible negative answer we suppose the existence of θ . In this case the following question appears.

2. How can we obtain the suitable formula θ from set of formulas $\Gamma \cup \Sigma'(P)$?

Here we show some of the possible ways of producing formula θ .

$$a/ \Gamma \cup \Sigma'(P) \models \forall \bar{x}^{(n)} (P(\bar{x}^{(n)}) \leftrightarrow \theta[\bar{x}^{(n)}])$$

$$b/ \Gamma \cup \Sigma'(P) \models \exists \bar{u}^{(m)} \forall \bar{x}^{(n)} (P(\bar{x}^{(n)}) \leftrightarrow \theta[\bar{x}^{(n)}] \bar{u}^{(m)})$$

i.e. the definition of P is parametrically given by the set of formulas $\Sigma'(P)$. Here $\bar{u}^{(m)} = (u_1, \dots, u_m)$, $\exists \bar{u}^{(m)} = \exists u_1 \dots \exists u_m$.

$$c/ \Gamma \cup \Sigma'(P) \models \bigvee_{1 \leq i \leq k} \forall \bar{x}^{(n)} (P(\bar{x}^{(n)}) \leftrightarrow \theta_i[\bar{x}^{(n)}])$$

i.e. $\Sigma'(P)$ defines P explicitly up to disjunction.

$$d/ \Gamma \cup \Sigma'(P) \models \bigvee_{1 \leq i \leq k} \exists \bar{u}^{(m)} \forall \bar{x}^{(n)} (P(\bar{x}^{(n)}) \leftrightarrow \theta_i[\bar{x}^{(n)}] \bar{u}^{(m)})$$

i.e. $\Sigma'(P)$ defines P explicitly up to parameters and disjunction.

It might happen that the set of formulas Γ has to be extended till Γ' as it is mentioned in 1. so as to define θ .

In that case if set of formulas $\Sigma'(P)$ is too weak then, similarly to the methods described in [2] we have to find such a formula θ for which

$$\Gamma \cup \{ \forall \bar{x}^{(n)} (P(\bar{x}^{(n)}) \leftrightarrow \theta[\bar{x}^{(n)}]) \} \models \Sigma' .$$

The set of formulas Γ can be extended here too if found necessary.

In that case if answer to question 1. is positive the following statement is true. Lemma: a/ if $\Gamma \cup \Sigma'(P) \models \forall \bar{x}^{(n)} (P(\bar{x}^{(n)}) \leftrightarrow \theta[\bar{x}^{(n)}])$ then $\Gamma \models \Sigma'(\theta)$

$$b/ \text{ if } \Gamma \cup \Sigma'(P) \models \bigvee_{1 \leq i \leq k} \forall \bar{x}^{(n)} (P(\bar{x}^{(n)}) \leftrightarrow \theta_i[\bar{x}^{(n)}])$$

then $\Gamma \models \Sigma'(\theta_1)$ or $\Gamma \models \Sigma'(\theta_2), \dots$, or $\Gamma \models \Sigma'(\theta_k)$.

We note here that we have to try the θ_i ($1 \leq i \leq k$) in b/ till the first formula where the statement stands for true.

If the answer to question 1. is negative then the knowledge within the disciplines defined by Γ is not enough for the explicit description of the required object.

On the basis of above a "high-level"

problem defining unit of SAD should operate the following way.

The basic knowledge of SAD is provided by set of formulas Γ . The customer gives his required object description by the help of set of formulas $\Sigma(P)$. As a first step the unit has to find an exact answer for the existence of θ , but since it is too complicated a task the following way is chosen. First the system controls whether $\Sigma(P)$ contradicts to knowledge Γ , i.e. it tries to deduce the identically false $(\varphi \wedge \neg \varphi)$ from $\Gamma \cup \Sigma$. If this doesn't succeed within a present time period then the system presupposes the existence of a formula and it will proceed onto the 2. task, i.e. producing θ .

Let us suppose that we succeeded in producing such a formula. It is followed by this trying:

$$\Gamma \vdash \Sigma(\theta)$$

If this is true then θ really becomes the requirements of the customer if not, then it may be supposed that the knowledge Γ of the SAD is not satisfactory for defining θ . Therefore Γ has to be extended till Γ' and above have to be repeated now for set of formulas Γ' . The system will go on with this either until it proves the impossibility of $\Sigma(P)$ on the basis of the extended set of formulas or, it succeeds to produce formula θ . Of course the system goes on with trying only for a fixed time. We note that the extension of set of formulas Γ need inductive logical means from the system.

Now we shall see that case when $\Sigma(P)$ defines P only up to the disjunction, that is when

$$\Gamma \cup \Sigma(P) = \bigvee_{1 \leq i \leq k} \forall \bar{x}^{(n)} (P(\bar{x}^{(n)}) \leftrightarrow \theta_i[\bar{x}^{(n)}])$$

The so obtained formulas $\theta_i[\bar{x}^{(n)}]$ ($1 \leq i \leq k$) have to be controlled one by one. So $\Gamma \vdash \Sigma(\theta_1)$ or $\Gamma \vdash \Sigma(\theta_2), \dots$, or $\Gamma \vdash \Sigma(\theta_k)$.

This control goes on until the first θ_i for which $\Gamma \vdash \Sigma(\theta_i)$. If neither θ_i satisfies above condition then it might be supposed that the knowledge Γ is not satisfactory. In this case the procedure goes on similarly, i.e. Γ is extended until Γ' , etc.

Useful theorems of definition theory

In this chapter we introduce those theorems of definition theory without proof which provide the explicit definition of P on the basis of $\Sigma(P)$ and Γ . Their proofs can be found in [1]. It is expected to obtain different types of theorems depending on the strength of $\Sigma(P)$. We begin with the theory containing the weakest conditions for $\Sigma(P)$.

If $\Sigma(P)$, Γ and a model $\mathcal{U}$ is given then the conditions of the theorems contain either that how many relations $R \subseteq A$ are there for which $(\mathcal{U}, R) \models \Sigma(P)$; or that how many such relations $R' \subseteq A$ are there to such a relation $R \subseteq A$ so as $(\mathcal{U}, R') \cong (\mathcal{U}, R)$

1. Theorem / Chang - Makkai Theorem /. If for every model $(\mathcal{U}, R)$ /for which $|A| > \omega /$ of $\Sigma \cup \Gamma$:

$$|\{R' : (\mathcal{U}, R) \cong (\mathcal{U}, R')\}| < 2^{|A|}$$

then there are a finite number of parametric formulas $\theta_1[\bar{x}^{(n)} \bar{v}^{(m)}], \theta_2[\bar{x}^{(n)} \bar{v}^{(m)}], \dots, \theta_k[\bar{x}^{(n)} \bar{v}^{(m)}]$ of F^+ such that

$$\Gamma \cup \Sigma(P) \models \bigvee_{1 \leq i \leq k} \exists \bar{v}^{(m)} \forall \bar{x}^{(n)} (P(\bar{x}^{(n)}) \leftrightarrow \theta_i[\bar{x}^{(n)} \bar{v}^{(m)}])$$

The theorem intuitively states if $\Sigma(P)$ circumscribes the relation P in some measure then there exists a parameter-vector $(v_1, \dots, v_m)$ and there are formulas $\theta_i[\bar{x}^{(n)} \bar{v}^{(m)}]$ ($1 \leq i \leq k$) of F^+ such that one of them gives the definition of P. In other words the set of formulas $\Sigma(P)$ defines P explicitly up to parameters and disjunction.

Theorem 2. If set of formulas $\Sigma(P)$ is such that to each model $\mathcal{M} \in M^t$ it is

$$|\{R : (\mathcal{M}, R) \models \Sigma\}| < 2^{|A|}$$

then there exists a finite number of first-order parametric formulas θ_i ($1 \leq i \leq k$) so that

$$\Gamma \cup \Sigma(P) \models \bigvee_{1 \leq i \leq k} \exists \bar{u}^{(m)} \forall \bar{x}^{(n)} (P(\bar{x}^{(n)}) \leftrightarrow \theta_i[\bar{x}^{(n)} \bar{u}^{(m)}])$$

The intuitive meaning of the theorem is as it follows: if the number of relations satisfying set of formulas $\Sigma(P)$ is less than the number of all possible relations then up to disjunction $\Sigma(P)$ parametrically defines relation P. The condition of the theorems claims that not all the possible relations should carry the characteristics described by the set of formulas $\Sigma(P)$.

Above theorems /Theorems 1. and 2./ are true also for that case when the number of the suitable relations is less than not $2^{|A|}$, but $|A|^+$ i.e. in this case there exists a finite number of first-order parametric formula and such a parameter vector that one of the formulas will give the definition of relation P by the suitable parameter vector.

Now let us see those cases when the possible number of relations satisfying $\Sigma(P)$ are finite in the models.

Theorem 3. If for every model $(\mathcal{M}, R) \in M^t$ of $\Sigma(P) \cup \Sigma$ it is true that

$$|\{R' : (\mathcal{M}, R) \equiv (\mathcal{M}, R')\}| < \omega$$

then there exists such a $k < \omega$ and there are such formulas $\sigma[\bar{u}^{(m)}]$, $\theta_i[\bar{x}^{(n)} \bar{u}^{(m)}]$ ($1 \leq i \leq k$) in F^t that

$$\Gamma \cup \Sigma(P) \models \exists \bar{u}^{(m)} (\sigma[\bar{u}^{(m)}] \wedge$$

$$\wedge \forall \bar{u}^{(m)} (\sigma[\bar{u}^{(m)}] \rightarrow$$

$$\rightarrow \bigvee_{1 \leq i \leq k} \forall \bar{x}^{(n)} (P(\bar{x}^{(n)}) \leftrightarrow \theta_i[\bar{x}^{(n)} \bar{u}^{(m)}])).$$

Theorem 4. If $\Sigma(P) \cup \Sigma$ is such that in every model $\mathcal{M} \in M^t$ it is $|\{R : (\mathcal{M}, R) \models \Sigma\}| < \omega$ then the statement of the previous theorem is true.

Intuitively the above theorems /Theorems 3. and 4./ state the following: if $\Sigma(P)$ is such that its required characteristics are fulfilled in every model by at least finite number of relations then there exists such a formula $\sigma \in F^t$ for calculating parameters $u_1, \dots, u_m$ and there exist formulas $\theta_1, \dots, \theta_k \in F^t$ out of which one defines relation P by the parameters determined by σ .

From the point of view of SAD this means that a theorem prover extended by inductive elements can prove, that

$$\Sigma(P) \cup \Gamma \vdash \exists \bar{u}^{(m)} \sigma[\bar{u}^{(m)}].$$

On the basis of this proof a zero-order termvector $\bar{t}^{(m)}$ must be selected so that $\Sigma \models \sigma[\bar{t}^{(m)}]$. After this it has to be proved, that

$$\Sigma \vdash \bigvee_{1 \leq i \leq k} \forall \bar{x}^{(n)} (P(\bar{x}^{(n)}) \leftrightarrow \theta_i[\bar{x}^{(n)} \bar{t}^{(m)}])).$$

Then on the basis of knowledge Γ , we select the suitable defining formula $\theta_i[\bar{x}^{(n)} \bar{t}^{(m)}]$.

Now we further restrict the requirements concerning set of formulas $\Sigma(P)$.

Theorem 5. If for each model $(\mathcal{M}, R)$ of $\Sigma(P) \cup \Sigma$ there exists such a finite $k \geq 1$, so

$$|\{R' : (\mathcal{M}, R) \equiv (\mathcal{M}, R')\}| \leq k$$

then there exist such formulas $\sigma_j[\bar{u}^{(m)}]$, $\theta_{ij}[\bar{x}^{(n)} \bar{u}^{(m)}]$ ($1 \leq j \leq r$, $1 \leq i \leq k$) in F^t that

$$\Sigma(P) \cup \Gamma \models \bigvee_{1 \leq j \leq r} \exists \bar{u}^{(m)} \sigma_j[\bar{u}^{(m)}] \wedge \forall \bar{u}^{(m)} (\sigma_j[\bar{u}^{(m)}] \rightarrow$$

$$\rightarrow \bigvee_{1 \leq i \leq k} \forall \bar{x}^{(n)} (P(\bar{x}^{(n)}) \leftrightarrow \theta_{ij}[\bar{x}^{(n)} \bar{u}^{(m)}])).$$

Theorem 6. /Kueker Theorem/: If $\Sigma(P) \cup \Sigma$ is such that for each model $\mathcal{M} \in M^t$ there exists a finite $k \geq 1$, so

$$|\{R : (\mathcal{M}, R) \models \Sigma\}| \leq k$$

then there exist such formulas $\sigma[\bar{u}^{(m)}]$

$\theta_i [\bar{x}^{(n)} \bar{y}^{(m)}] \quad (1 \leq i \leq k)$
in F^t such that
 $\Sigma(P) \cup \Gamma \models \exists \bar{y}^{(m)} \bigwedge [\bar{y}^{(m)}] \wedge \forall \bar{y}^{(m)} (\theta [\bar{y}^{(m)}] \rightarrow$
 $\rightarrow \forall \bar{x}^{(n)} (P(\bar{x}^{(n)}) \leftrightarrow \theta_i [\bar{x}^{(n)} \bar{y}^{(m)}])$
 $1 \leq i \leq k$

In these theorems similarly to Theorems 3. and 4. the formulas θ_i ($1 \leq i \leq r$) and the formula θ serve to define the parameter vector. The definition of relation $\sim$ is done also on the basis of those described after Theorem 4. There is a difference only when definition is done on the basis of Theorem 5, because here we have to try out the formulas not only according to i ($1 \leq i \leq k$) but also according to j ($1 \leq j \leq r$).

The conditions of Theorems 5. and 6. for $\Sigma(P)$ are so much stronger than those of Theorems 3. and 4. that now we claim the existence of such a finite k which is upper-bound of the number of suitable relations in each model.

The $\Sigma(P)$ is the strongest in that case if this conditions are satisfied in each model by at least one relation. Now we discuss those theorems which refer to this.

Theorem 7. /Svenonius' Theorem/: If for each model (U, R) of $\Sigma(P) \cup \Gamma$:

$$|\{R' : (U, R) \cong (U, R')\}| \leq 1$$

then there exists a finite $m < \omega$ and there exist such formulas $\theta_i [\bar{x}^{(n)}]$ ($1 \leq i \leq k$) in F^t so that,

$$\Sigma(P) \cup \Gamma \models \forall \bar{x}^{(n)} (P(\bar{x}^{(n)}) \leftrightarrow \bigvee_{1 \leq i \leq k} \theta_i [\bar{x}^{(n)}]).$$

Intuitively it means that if we take two extensions (U, R_1) and (U, R_2) of any model $(U, R) \models \Sigma(P)$ so that these become models of $\Gamma \cup \Sigma(P)$ and these are isomorphic then $R_1 = R_2$.

In this case the set of formulas $\Sigma(P)$ defines relation P up to disjunction.

Theorem 8. /Beth's Theorem/: If the set of formulas $\Sigma(P) \cup \Gamma$ is such that for each model $(U, R) \models M^t$:

$$|\{R : (U, R) \models \Sigma(P)\}| \leq 1$$

then there exists such a formula $\theta [\bar{x}^{(n)}]$ in F^t that

$$\Sigma(P) \cup \Gamma \models \forall \bar{x}^{(n)} (P(\bar{x}^{(n)}) \leftrightarrow \theta [\bar{x}^{(n)}]).$$

Intuitively if $\Sigma(P)$ is so strong that every model $(U, R) \models M^t$ can be extended to a model of $\Sigma \cup \Gamma$ by at the most one relation then $\Sigma(P)$ defines relation P explicitly.

Conclusion

As we could see from above the model theory provides mathematical bases suitable for the development of different kinds of SAD important in the practice. This is especially important because to construct implicitly described objects from psychological point of view is a task demanding creativity. The degree of creativity partly depends on the circumscription of the required object and partly on the development of the corresponding discipline. With the help of the theorems of different strength described in above we can obtain different SAD-s of different degree of creativity. So far we can see that the research of artificial intelligence requires the application of deep mathematical results of mathematical logic. To make SAD more effective we need the following problem to be solved: if $\Sigma(P) \subseteq F^t$ and $\Gamma \subseteq F^t$ are given then what conditions should $\Sigma(P)$ satisfy so as to have a formula $\theta \in F^t$ existing for which $\Gamma \models \Sigma(P)$.

References

1. C.C. Chang, M.J. Keisler, Model Theory, North Holland, 1973.
2. G.D. Plotkin, A further note on inductive generalization, Machine Intelligence 6, Editors B. Meltzer, D. Michine, University Press, Edinburgh, 1970.