ON THE DEFINITION OF CLASSES OF INTERPRETATIONS

Bruno COURCELLE

IRIA, Rocquencourt, 78150 Le Chesnay

France

Abstract : A class C of interpretations is *algebraic* if, roughly speaking, for every two recursive program schemes ϕ and ϕ', the equivalence of ϕ and ϕ' with respect to C can be proved by an induction on the length of computation [9] if it holds. Classes of interpretations can be defined by logical, and/or order theoretical conditions. We examine several cases of algebraicity (for classes defined by first-order conditions) and non-algebraicity.

1 - *Introduction*

The equivalence between programs is an essential concept in the mathematical theory of computation and programming, but very difficult to study for well-known theoretical reasons. Program schemes have been introduced to overcome these difficulties as much as possible.

A program P is splitted into a program scheme ϕ(i.e. a program where the domain of computation and the base functions are left unspecified) and an interpretation I (i.e. the specification of a domain D_I and a function $f_I = D_I^k \to D_I$ for each k-ary base function symbol). We denote by ϕ_I the function computed by ϕ under I i.e. the function computed by the program P.

We only consider in this paper recursive program schemes without assignments. The corresponding schemes have no interpreted base functions. The conditional operator *if... then...else...*is replaced by a 3-adic symbol h(...,...,...) which can be interpreted by an arbitrary 3-adic function.

The corresponding equivalence relation on schemes, namely $\phi \equiv \phi'$ iff $\phi_I = \phi'_I$ for every interpretation I, is very restrictive, and does not help very much for the study of interesting equivalences between real programs.
In order to get more concrete results, we use the notion of a *class of interpretations* C ; the associated equivalence between program schemes is then :

$$\phi \equiv_C \phi' \text{ iff } \phi_I = \phi'_I \text{ for every } I \in C.$$

A similar concept has been introduced by other authors. In [3] a system is described which uses pairs of program schemes which are equivalent *under conditions*, i.e. with respect to a certain class of interpretations. In [14] a program scheme is given with a first-order formula Φ and one only considers interpretations which validate Φ .
The order $\leq_I$ of an interpretation is extended to functions as usual and still denoted $\leq_I$. We also define $\phi \leq_C \phi'$ iff $\phi_I \leq_I \phi'_I$ for all $I \in C$.

Clearly $\phi \equiv_C \phi'$ if $\phi \leq_C \phi'$ and $\phi' \leq_C \phi$.

Let ϕ be a R.P.S (i.e. a recursive program scheme), let I be an interpretation and for each positive integer n let $\phi_I^{(n)}$ be the approximation of ϕ_I which can be computed with at most n nested recursive calls. This function can be defined by a finite term (written with base function symbols only) which we denote by $\phi^{(n)}$.

Let ϕ and ϕ' be R.P.S's, let C be a class of interpretations such that ,

(1) $\forall n \exists m \forall I \in C,\quad \phi_I^{(n)} \leq_I \phi_I'^{(m)}$

equivalently,

(2) $\forall n \exists m\ ,\ \phi^{(n)} \leq_C \phi'^{(m)}$;

then, since ϕ_I is the least upper bound of the $\phi_I^{(n)}$'s, we obtain that :

(3) $\forall I \in C,\quad \phi_I \leq_I \phi'_I$

or equivalently :

(4) $\quad \phi \leq_C \phi'$

This situation precisely occurs when (4) is provable by induction on the computation [9] . Many induction principles (including "Scott's induction principle") reduce to that one. An actual proof requires that, for every n,

$$\exists m\ ,\ \phi^{(n+1)} \leq_C \phi'^{(m)}$$

can be deduced from :

$$\exists m\ ,\ \phi^{(n)} \leq_C \phi'^{(m)}$$

which may be technically difficult to prove. But the main point for us is whether (2) holds or not.

A class C of interpretations is *algebraic* if for every two RPS's ϕ and ϕ', (4) implies (2). Hence, for an algebraic class of interpretations C, it is reasonable to search for proofs by induction on the computation.

In nonalgebraic classes, other proofs are required.

Example 1 : Let C be the class of interpretations of the form

$I = \langle D_I, \leq_I, \perp, p_I, f_I, h_I, a_I, b_I \rangle$ such that

1. D_I contains the truth values *true, false*
2. p_I is a continuous mapping : $D_I \to \{\perp, true, false\}$, and $\forall m,\ p_I(f_I^m(a_I)) \neq \perp$
3. $h_I(x,y,z)$ = *if* x *then* y *else* z
4. $\exists n \in \mathbb{N}$ such that $p_I(f_I^n(a_I))$ = *true*

(The reader is refered to the main text for the missing details).

Let Σ be the following RPS :

$$\begin{cases} \phi = \psi\ (a) \\ \psi(v_1) = h\ (p(v_1),\ b,\ \psi(f(v_1))). \end{cases}$$

One is easily convinced that $\phi \equiv_C b$, that $\phi^{(n)} \leq_C b$ holds (and can be proved by induction on n) but $b \leq_C \phi^{(n)}$ does not hold for any n. Hence $b \leq_C \phi$ cannot be proved by induction on the computation. An ad hoc proof by induction on the integer n of condition 2 above is possible but is not of a great theoretical interest.□

It is often possible to prove a valid equation $\phi_I = \phi'_I$ by an induction "on the domain". If $D_I = \bigcup_{n \in \mathbb{N}} D_I^{(n)}$, it suffices to prove by induction on n that for all $n \in \mathbb{N}$:

$$\forall d_1, \ldots, d_k \in D_I^{(n)},\ \phi_I(d_1,\ldots,d_k) = \phi'_I\ (d_1,\ldots,d_k).$$

This is called a *structural induction* in [2] .
But such a proof is only valid for one interpretation. On the other hand, proofs by induction on the computation, which only use *relations between base functions*, are valid for every interpretation which satisfies those relations hence are of more general interest.
Hence, in this approach, the relevant questions about a given family C of interpretations are the following ones :

(1) Is it algebraic ?
(2) If it is, find some characterization of $t \leq_C t'$ (for finite terms t and t')

In this paper, we relate the algebraicity of classes of interpretations with their definition by logical conditions, properties of programs etc...

Our two main results are the following ones :

(1) a first-order class of discrete interpretations is algebraic (theorem 2),
(2) the intersection of an algebraic class with a relational class of interpretations is not necessarly algebraic (theorem 5).

2 - *Recursive Program Schemes*

We recall quickly a few basic notions. (See [4] and [10] for more details).

Let F be a finite set of *function symbols* with arity ($f \in F$ has arity $\rho(f) \geq 0$) which contains a special symbol Ω of arity 0. Let $V_k = \{v_1, \ldots, v_k\}$ for $k \geq 1$ and $V = V_K$ (for some fixed large enough K)be the set of *variables*.

Let M(F,V) denote the set of finite well formed terms written on $F \cup V$ according to arities. This set is ordered by $\prec$, the least order relation such that :

(1) $\Omega \prec t$ for any $t \in M(F,V)$
(2) $f(t_1,\ldots,t_k) \prec f(t'_1,\ldots,t'_k)$ if $t_i \prec t'_i$ for $i = 1,\ldots,k$

A *recursive program scheme* is a set of equations or a context-free tree grammar of the form :

$$\Sigma \begin{cases} \phi_i(v_1,\dots,v_{k_i}) = t_i + \Omega \\ 1 \le i \le n,\ k_i = \rho(\phi_i)\ ,\ t_i \in M(F \cup \Phi, V_{k_i}) \end{cases}$$

$\Phi = \{\phi_1,\dots,\phi_n\}$ is the set of *function variables*.

Let $L(\Sigma,\phi_1) \subset M(F,V)$ be the tree language generated by Σ with axiom $\phi_1(v_1,\dots,v_{k_1})$.

Lemma 1 (Nivat [10]) : $L(\Sigma,\phi_1)$ *is directed w.r.t.* $\prec$.

In order to consider such systems as program schemes, we must define their semantics.

Definition 1 : An *interpretation* I is an object

$I = \langle D_I, \le_I, \bot, \langle f_I \rangle_{f \in F} \rangle$ (also called a *complete ordered F-magma*)

such that

1. $\le_I$ is a partial order with least element $\bot$ and $\bot = \Omega_I$.
2. for each k-ary $f \in F$, f_I is an increasing mapping : $D_I^k \to D_I$
3. Every directed subset Δ of D_I has a least upper bound $Sup(\Delta)$.
4. Each f_I is *continuous*, i.e. $f_I(Sup(\Delta_1),\dots, Sup(\Delta_k)) = Sup\ (f_I(\Delta_1\ ,\dots,\Delta_k))$ for directed $\Delta_1,\dots,\Delta_k \subset D_I$.

Every $t \in M(F,V)$ clearly defines a function $t_I : D_I^K \to D_I$ which is increasing and continuous.

The function computed by (Σ, ϕ_1) in I is $(\Sigma, \phi_1)_I$ also denoted by ϕ_{1I} if there is no ambiguity, defined by :

$$\phi_{1I}(d_1,\dots,d_K) = Sup\ \{ t_I(d_1,\dots,d_K) / t \in L(\Sigma, \phi_1)\}$$

This value does exist since the mapping $t \rightsquigarrow t_I(d_1,\dots,d_K)$ is increasing and by lemma 1.

This semantics coincides with the classical definition by fixed-points used for instance in Milner [8] . See [13],[10] for the proof.

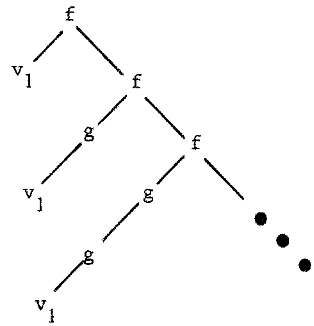
In fact we will only consider RPS's (Σ, ϕ) such that $\rho(\phi) = 0$. Hence ϕ_I is an element of D_I and not a function. This is not a loss of generality. Let (Σ, ϕ) with $\rho(\phi) = k$ and I be an interpretation. Add the new constants $c_1,\dots, c_k$ to F, a new 0-ary function symbol ϕ' to Φ, the equation $\phi' = \phi(c_1,\dots,c_k)$ to Σ. Instead of studying $\phi_I : D_I^k \to D_I$ we will study ϕ'_J where J ranges over $\{ \langle I, c_{1_J},\dots,c_{k_J} \rangle / c_{i_J} \in D_I \}$.

We denote by $M^\infty(F,V)$ the set of "infinite well formed" terms. They can be formally defined as least upper bounds of directed subsets of M(F,V) or, more intuitively, as infinite trees.

Example : The least upper bound of

$$\Delta = \{\Omega, f(v_1,\Omega),\ f(v_1,f(g(v_1),\Omega)),\dots\}$$

is



Hence, by lemma 1 there exists $T(\Sigma, \phi_1) = \text{Sup}\,(L(\Sigma, \phi_1))$ in $M^\infty(F,V)$. The semantics of (Σ, ϕ_1) is completely defined by $T(\Sigma, \phi_1)$. If $\rho(\phi_1) = 0$ then $T(\Sigma, \phi_1) \in M^\infty(F)$. For simplicity we will consider in this paper any element of $M^\infty(F)$ as defined by a R.P.S., i.e. we will give definitions relative to arbitrary elements of $M^\infty(F)$ and not to those which are associated with RPS's.

3 - Families of interpretations

We call *family* or *class of interpretations* a class of interpretations in the sense of set-theory.

Let C be a class of interpretations. For $T, T' \in M^\infty(F)$ we define $T \leq_C T'$ iff $T_I \leq_I T'_I$ for all $I \in C$, and $T \equiv_C T'$ iff $T \leq_C T'$ and $T' \leq_C T$. Hence (Σ,ϕ_1) and (Σ',ϕ'_1) are C-equivalent iff $T(\Sigma,\phi_1) \equiv_C T(\Sigma', \phi'_1)$.

A class C is *algebraic* (see [4] for a justification of the terminology) if for all $t \in M(F)$, and $T' = \sup_{n\geq 0}(t'_n) \in M^\infty(F)$ (where $t'_n \in M(F)$ for $n \geq 0$), $t \leq_C T'$ iff $t \leq_C t'_n$ for some n.

The class I of all interpretations is algebraic since $\leq_I$ is exactly $\prec$ (see [4] , [13]) More generally, let R be a possibly infinite subset of $M(F,V) \times M(F,V)$ and $C_R = \{I/I$ is an interpretation and $t_I \leq_I t'_I$ for all $(t,t') \in R\}$. For example, the class of interpretations such that some binary function is commutative is of that form. C_R is called a *relational class*.

Proposition 1 ([4]): *Every relational class is algebraic*

An interpretation I is *discrete* if $(D_I, \leq_I)$ satisfies the following :

$$\forall x, y \in D_I, x \leq_I y \iff x = \bot \text{ or } y = x.$$

Most concrete examples of R.P.S.'s use discrete interpretations. Many papers on recursive program schemes only consider discrete interpretations. See section 4 for a precise correspondance between these definitions and ours.

The family D of discrete interpretations is proved to be algebraic in [1]. We will extend this result in section 4.

Let us note that for $I \in D$, and $T \in M^\infty(F)$:

(i) either $t_I = \bot$, for all $t \prec T$, then $T_I = \bot$ (divergence)

(ii) or $t_I = d \neq \perp$ for some $t \prec T$ and $T_I = d$ (termination)

Proposition 2 : If C and C' are algebraic, $C \cup C'$ is algebraic.

Proof : Let $t \leq_{C \cup C'} T$ for $t \in M(F)$, $T \in M^{\infty}(F)$ then $t \leq_C T$ hence $t \leq_C t_1 \prec T$ for some $t_1 \in M(F)$ since C is algebraic. Similarly $t \leq_{C'} t_2 \prec T$. Let t_3 be the least upper bound of t_1 and t_2 in M(F). Then $t_3 \prec T$, $t_1 \leq_C t_3$. hence $t \leq_C t_3$. Similarly, $t \leq_{C'} t_3$ hence $t \leq_{C \cup C'} t_3$. □

Theorem 5 will show that $C \cap C'$ need not be algebraic. But theorem 4 defines a class F of algebraic families which is closed by intersection.

4 - First-order classes fo interpretations

The family I of all interpretations is the class of all models of a certain second-order theory. This is clear by the definition of an interpretation. In fact second - order theories are very powerful and define algebraic as well as nonalgebraic classes (see the end of section 5).

We will consider classes defined by *first-order conditions.*

Let A be a set of first-order closed formulas on a set of function symbols including F and a set of predicate symbols including ≤. Let us define :

Mod (A) to be the set of models of A (Schoenfield [12])

Int (A) = $I \cap$ *Mod* (A), the class of interpretations which are models of A.

Since I is not first-order definable, *Int* (A) need not be *Mod* (A') for any first-order theory A'.

Definition 2 : A class of interpretations C is *first-order* if C = *Mod*(A) for some first-order theory A.

Example : D is a first-order family. In fact, conditions 1 and 2 of definition 1 are first-order. Let A_0 be a (finite) set of first-order formulas expressing them. Let F_D be the following closed formula :

$$\forall x \ \forall y \ [x \leq y \Leftrightarrow x = \Omega \ \text{ or } \ x = y] ,$$

and $A_D = A_0 \cup \{F_D\}$. It is fairly easy to prove that conditions 3 and 4 of definition 1 are implied by A_D hence D = *Mod* (A_D).

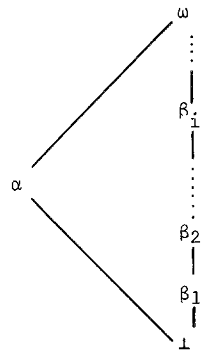
For a first-order theory A, let D(A) = $D \cap$ *Mod*(A). Then D(A) = *Mod* (A $\cup$ A_D) hence is a first-order class.

We state and prove two theorems ; the second one is positive and will have several interesting corollaries.

Theorem 1 : For a first-order theory A, *Int* (A) *is not algebraic in general.*

Theorem 2 : For a first-order theory A, D(A) *is algebraic.*

Proof of theorem 1 : Let $F = \{a, f\}$ with $\rho(a) = 0$ and $\rho(f) = 1$.
Let A consist of the following formula : $\forall x\ [f(x) = x \Rightarrow a \leq x]$
Let $T = \mathrm{Sup}\ \{f^n(\Omega)/n \geq 0\}$ be the least fixed point of $x = f(x)$ in $M^\infty(F)$ and $C = Int\ (A)$. Then $a \leq_C T$ follows from A since $f_I(T_I) \equiv T_I$ in every interpretation $I \in C$. But there is no $n \in \mathbb{N}$ such that $a \leq_C f^n(\Omega)$. To see that, it suffices to take I with $D_I = \{\ \bot, \alpha, \beta_1, \ldots, \beta_m, \ldots, \omega\}$ ordered as follows :



with $f_I(\bot) = f_I(\alpha) = \beta_1$, $f_I(\beta_i) = \beta_{i+1}$, $f_I(\omega) = \omega$ and $a_I = \alpha$. □

Proof of theorem 2 : This proof will be based on the compactness theorem for first-order logic.

Assume that $C = D(A) = Mod(A')$ is not algebraic, that $t' \in M(F)$ and $T = \mathrm{Sup}\ \{\ t_n/n = 0,1,\ldots\ \} \in M^\infty(F)$ with $t_0 \prec t_1 \prec t_2 \prec \cdots \prec t_n \prec \cdots$ such that :

(1) $t' \leq_C T$,

(2) $t' \not\leq_C t_n$ for each n.

By (2), each of $A' \cup \{\ \neg(t' \leq t_n)\}$ has a model $I_n \in C$. In fact, I_n is a model of $A' \cup \{\ \neg\ (t' \leq t_0),\ \neg(t' \leq t_1), \ldots, \neg\ (t' \leq t_n)\}$ since for each $i = 0, \ldots, n, t_i \prec t_n$; hence $t'_{I_n} \leq t_{iI_n}$ would imply $t'_{I_n} \leq t_{nI_n}$.

By the compactness theorem for first-order logic [12], $A' \cup\ \{\neg(t' \leq t_0), \neg\ (t' \leq t_1), \ldots, \neg(t' \leq t_n), \ldots\}$ has a model I. Hence $I \in Mod(A')$ $= D(A)$. Since I is discrete there exists n such that $T_I = [t_n]_I$. Hence $t'_I \leq [t_n]_I$ by hypothesis (1). But this contradicts the definition of I. □

As a first application, we consider classes of discrete interpretations with *interpreted conditionals.*

Let F contain a special triadic symbol h, special constants *true* and *false* and a set of special functions called *predicate* symbols (p, p', q, q', r, r' will be reserved to predicate symbols). The domain D_I of an interpretation I contains the truth values $true_I$ and $false_I$ which are assumed to be distinct and not ⊥. For simplicity we will omit the subscript I.

Let D_{cond} be the class of discrete interpretations satisfying the following conditions :

(1) $true \neq \Omega$ and $false \neq \Omega$ and $true \neq false$.

(2) $\forall x_1,\ldots,x_k\ [p(x_1,\ldots,x_k) = \Omega$ or $p(x_1,\ldots,x_k) = true$ or $p(x_1,\ldots,x_k) = false]$ for each k-ary predicate p.

(3) $\forall x, y, z, [x \neq true$ and $x \neq false \implies h(x,y,z) = \Omega]$

(4) $\forall x, y\ [h(true,x,y) = x$ and $h(false,x,y) = y]$

Then D_{cond} is clearly first-order and

Corollary 1 : D_{cond} is first-order and algebraic

In fact most papers on program schemes use only interpretations of D_{cond} with, possibly some extra conditions. Let us take [0] as an example. A typical R.P.S. is

$$\phi(x) = if\ p(x)\ then\ (if\ q(x)\ then\ f(x)\ else\ \phi(g(\phi(x))))\ else\ x$$

An interpretation is some $I = < D, p_I, q_I, f_I, g_I >$ where p_I, q_I are total functions : $D \to \{true, false\}$ and f_I, g_I are total functions : $D \to D$

We write the same scheme :

$\phi(x) = h(p(x), h(q(x), f(x), \phi(g(\phi(x)))),x)$ and take the discrete interpretation:

$$J = < D \cup \{ \bot, true, false\}, \bot, \leq, p_J, q_J, f_J, g_J, h_J >$$

where $p_J(\bot) = p_J(true) = p_J(false) = \bot$

$p_J(d) = p_I(d)$ if $d \in D$

and similarly for q_J, f_J, g_J. The function h_J is defined by (3) and (4) above. (The meaning of $\bot$ is : *undefined*). Clearly, $J \in D_{\text{cond}}$. The function ϕ_I(defined as in [0]) is related with ϕ_J by the following :

$\phi_J(d) = \phi_I(d)$ if $\phi_I(d)$ is defined

$\phi_J(d) = \bot$ if $d = \bot$ or $\phi_I(d)$ is undefined

We now consider discrete interpretations defined by relations.

Let $R \subset M(F,V) \times M(F,V)$ and C_R be the relational class associated with R as defined in section 3. Let $D(R) = D \cap C_R$. Then $D(R) = D(A_R)$ where A_R is the set of formulas

$\forall v_1,\ldots,v_K\ [t \leq t']$

for all $(t,t') \in R$.

Corollary 2 : For $R \subset M(F,V)^2$, $D(R) = D \cap C_R$ is algebraic.

This partially answers to a question of [6]. This result can be slightly extended.

Let B be the least family of classes of interpretations which contains the relational classes (i.e. the C_R's), the first-order classes of discrete interpretations (i.e. the $D(A)$'s) and is closed by finite union and finite intersection. Then

Theorem 3 : Every $C \in B$ is algebraic

Proof : Let $C \in B$. Then $C = \bigcup_{i=1}^{n} C_i$ where each C_i is the form :

$$C_{R_1} \cap C_{R_2} \cap \ldots \cap C_{R_k} \cap \quad D(A_1) \cap \ldots \cap D(A_\ell)$$

If $\ell = 0$ then $C_i = C_{R_1 \cup \ldots \cup R_k}$ is algebraic (proposition 1)

If $\ell \geq 1$ then $C_i = C_{R_1 \cup \ldots \cup R_k} \cap D(A_1 \cup \ldots \cup A_\ell)$

$= D(A_1 \cup \ldots \cup A_\ell \cup A_{R_1 \cup \ldots \cup R_k})$ and is algebraic by theorem 2. By proposition 2, C is algebraic. □

5 - Families defined by properties of programs

It has been proved in [4] that the family of all interpretations in which two given program schemes are equivalent is not algebraic.

We now consider the class of discrete interpretations in which a given program terminates. This class is not algebraic in general.

Let ϕ be a RPS, let $D_\phi = \{I \in D \ / \ \phi_I \neq \bot\}$ and $D_{cond,\phi} = \{ I \in D_{cond} / \ \phi_I \neq \bot \}$. Then

Theorem 4 : D_ϕ and $D_{cond,\phi}$ are not algebraic in general.

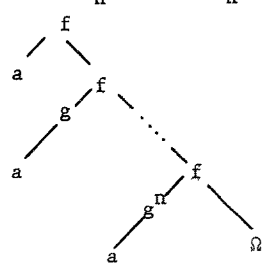
Proof : 1. Let $F = \{a, b, f, g\}$ with $\rho(a) = \rho(b) = 0$, $\rho(g) = 1$, $\rho(f) = 2$.

Let (Σ,ϕ) be the following RPS :

$$\begin{cases} \phi = f(\Omega, \psi(a)) \\ \psi(v_1) = f(v_1, \psi(g(v_1))) \end{cases}$$

Then $T = T(\Sigma,\phi) = \text{Sup} \{ t_n / n \in \mathbb{N} \}$ where $t_n = f(\Omega, t'_n)$ and

t'_n =



Let also $T' = \text{Sup} \{ t'_n / n \geq 0 \}$. Then $T = f(\Omega, T')$.

Let $C = D_\phi$ and $t = f(b,\Omega)$. We first prove that $t \leq_C T$. Let $I \in C$. If $t_I = \bot$ then $t_I \leq T_I$. If $t_I = d \neq \bot$ and $f_I(\bot,\bot) \neq \bot$ then $d = t_I = f_I(\bot,\bot) = T_I$.

If $t_I = d \neq \bot$ and $f_I(\bot,\bot) = \bot$ then $T_I = d' \neq \bot$ by definition of C.

Hence $d' = f_I(\perp,d'')$ where $d'' = T'_I \neq \perp$. Since I is discrete, $d = f(b_I,\perp) \neq \perp$ and $d' = f(\perp,d'') \neq \perp$ imply that $d = f(b_I,d'') = d'$; hence $t_I = T_I$.

We now prove that $t \not\leq_C t_n$ for each n. For $n \in \mathbb{N}$ let I be the discrete interpretation with domain $D = \{\perp,c_0,c_1,\ldots,c_{n+1}\}$ such that :

$$a_I = c_0$$

$$b_I = c_{n+1}$$

$$\begin{cases} g_I(c_i) = c_{i+1} & \text{for } i = 0, 1,\ldots, n \\ g_I(\perp) = \perp & \\ g_I(c_{n+1}) = c_{n+1} & \end{cases}$$

$$\begin{cases} f_I(\perp,d) = c_n & \text{if } d = c_n \\ \qquad\quad = \perp & \text{otherwise} \\ \left.\begin{array}{l} f_I(c_i,d) = c_n \quad \text{if } d = c_n \\ \qquad\quad = \perp \quad \text{otherwise} \end{array}\right\} & \text{for } i = 0,1,\ldots,n \\ f_I(c_{n+1},d) = c_n & \text{for all } d \in D \end{cases}$$

One easily checks that $I \in \mathcal{D}$ and that :

$$\begin{cases} [g^i(a)]_I = c_i & \text{for } i = 0,1,\ldots,n+1, \;, \\ [t'_{n+1}]_I = c_n \text{ and } T'_I = c_n \;, & \\ T_I = c_n \neq \perp \;, \text{ hence } I \in \mathcal{C} & \\ [t'_n]_I = \perp \text{ and } [t_n]_I = \perp & \\ t_I = f_I(c_{n+1},\perp) = c_n & \end{cases}$$

Hence this shows that $t \not\leq_C t_n$. Hence $\mathcal{D}_\phi$ is not algebraic

That $\mathcal{D}_{\mathrm{cond},\phi}$ is not algebraic follows from example 1 of the introduction. It is not difficult to see that $\mathcal{D}_{\mathrm{cond},\phi} \supseteq \mathcal{C} \cap \mathcal{D}$ (ϕ and $\mathcal{C}$ are defined in example 1). For every n, one can easily build an interpretation $I \in \mathcal{C} \cap \mathcal{D}$ such that $\phi_I^{(n)} = \perp, b_I \neq \perp$. Hence $b \leq_{\mathcal{C}'} \phi$ but $b \leq_{\mathcal{C}'} \phi^{(n)}$ for no $n \in \mathbb{N}$ where $\mathcal{C}' = \mathcal{D}_{\mathrm{cond},\phi}$. □

Remark : 1. Theorem 4 entails that $\mathcal{D}_\phi$ and $\mathcal{D}_{\mathrm{cond},\phi}$ are not $\mathcal{D}$(A) for any first-order theory A. This has already been proved by Kfoury and Park [7] (for $\mathcal{D}_{\mathrm{cond}}$ and flowcharts).

2. Powerful logical calculi such that 2^{nd}-order-logic, μ-calculus or $L\ \omega_1\omega$ (see Park [11]) allow to express properties of programs such that equivalence or termination. It is clear from theorem 4 that classes of interpretations which are definable by such calculi are not algebraic in general.

6 - Intersection of algebraic classes

It is known from [6] that the algebraicity of classes is not preserved under intersection. One could suspect that the algebraicity is preserved by intersection with a relational class C_R since this holds in two cases :

① $C_{R'} \cap C_R = C_{R \cup R'}$ (algebraic by proposition 1)

② $D \cap C_R = D(R)$ (algebraic by corollary 2)

This is not the case :

Theorem 5 : There exists an algebraic class of discrete interpretations C *and* $R = \{(t,t')\}$ *such that* $C \cap C_R$ *is not algebraic.*

Proof : let $F = \{f,g,a,b,\Omega\}$ with $\rho(f) = 2$, $\rho(g) = 1$, $\rho(a) = \rho(b) = 0$.
Let A be the set of first order formulas consisting of :

F_0 : $b \neq \Omega$

F_1 : $g(\Omega) = \Omega$

$F_{2,n}$: $g^n(a) \neq \Omega$ for $n = 0,1,2,...$

F_3 : $\forall x\ [f(\Omega,x) = \Omega]$

F_4 : $\forall x\ [x \neq \Omega \implies f(x,b) \neq \Omega]$

Let $D_0 = D(A)$ and $C = \{ I \in D_0 \ / \ [f(g^n(a),\Omega)]_I \neq \perp$ for some integer $n\}$

The proof uses the following lemma which will be proved later :

Lemma 2 : The order relations $\leq_C$ *and* $\leq_{D_0}$ *are the same*

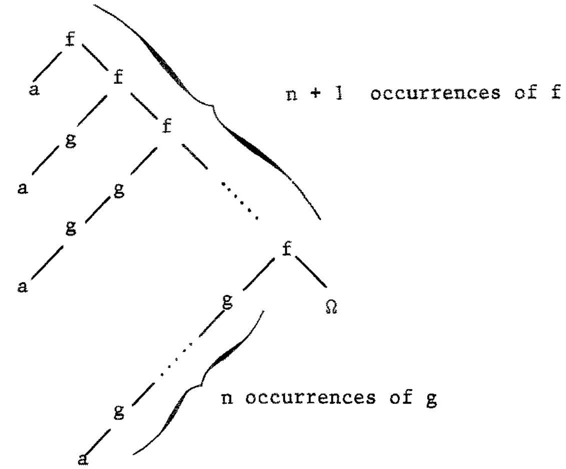
The class D_0 is algebraic by theorem 2 and so is C by lemma 2. Let $R = \{(f(v_1,b),b)\}$ and $C' = C \cap C_R$. Hence C' is the class of interpretations I of C satisfying :

R : $\forall d \in D_I\ ,\ f_I(d,b_I) \leq b_I$

which we will refer to in the sequel as condition R. We claim that C' is not algebraic.
Let $T = T(\Sigma,\phi)$ where Σ is following scheme :

$$\Sigma \begin{cases} \phi = \psi(a) \\ \psi(v_1) = f(v_1,\psi(g(v_1))) \end{cases}$$

Hence $T = \operatorname{Sup}_n(t_n)$ where t_n is



Technically, we define

$$t_{n,n} = f(g^n(a),\Omega) \quad , \quad \text{for } n = 0,1,2,\ldots \quad \text{and}$$

$$t_{n,m} = f(g^m(a),t_{n,m+1}) \quad \text{for } 0 \le m < n$$

Hence $t_n = t_{n,0}$.

Let $I \in C'$ and n such that $[t_{n,n}]_I \neq \bot$. Since I is discrete, $f(g^n(a),b)_I = [t_{n,n}]_I$ and the common value is b_I by condition R. By F_4 and F_2 $[t_{n,n-1}]_I \neq \bot$ and the value must be b_I by condition R again. By the same argument, we get that

$$b_I = [t_{n,n-1}]_I = [t_{n,n-2}]_I = \ldots = [t_{n,0}]_I$$

Hence $b \equiv_{C'} T$. If C' is algebraic, there exists an integer m such that $b \le_{C'} t_m$.
Let n > m and I be the following interpretation :

$I = \langle D,\le,\bot, f,g,\ldots\rangle$

$D = \{\bot,\alpha_0,\alpha_1,\alpha_2,\ldots,\alpha_n,\beta\rangle$

$a_I = \alpha_0$

$b_I = \beta$

$g_I(\bot) = g_I(\beta) = \bot$

$g_I(\alpha_i) = \alpha_{i+1} \quad$ if $0 \le i < n$

$g_I(\alpha_n) = \alpha_n$

$f_I(d,d') = \beta$ if $d = \alpha_n$ or ($d \neq \bot$ and $d' = \beta$) ,
$\quad = \bot$ otherwise

A straight forward verification shows that $I \in C'$. In particular $t_{n,n_I} = \beta$ but $[t_{m,m}]_I = \bot$. Hence $[t_{m,m-1}]_I = f(g^{m-1}(a), t_{m,m})_I = f_I(\alpha_{m-1}, \bot) = \bot$ and similarly $[t_{m,m-1}]_I = [t_{m,m-2}]_I = \ldots = [t_{m,0}]_I = \bot$. Hence $b_I \not\leq [t_m]_I$ and $b \not\leq_{C'} t_m$: this shows that C' is not algebraic.□

Before starting the proof of lemma 2, we describe a construction on interpretations which has been introduced in [1]. We redefine it in order to have a self-contained proof.

Let $I \in D_0$. We associate with I a discrete interpretation J which is defined as follows

$$D_J = \{<d,t> \in D_I \times M(F) \ / \ t_I = d\} ,$$

$$\bot_J = <\bot,\Omega> ,$$

$$a_J = <a_I,a> \quad ; \quad b_J = <b_I,b>$$

$$g_J(<d,t>) = <\bot,\Omega> \quad \text{if } g_I(d) = \bot$$
$$= <g_I(d) , g(t)> \quad \text{otherwise} ,$$

$$f_J(<d_1,t_1> , <d_2,t_2>) = <\bot,\Omega> \quad \text{if } f_I(d_1,d_2) = \bot ,$$
$$= <f(d_1,\bot) , f(t_1,\Omega)> \quad \text{if } f_I(d_1,\bot) \neq \bot$$
$$= <f(d_1,d_2) , f(t_1,t_2)> \quad \text{otherwise}$$

The interpretation J is in fact the *SD*-free interpretation associated with I of [1]. We denote it here by $F(I)$.

Lemma 3 : Let $t \in M(F)$

1. *There exists* $t' \in M(F)$ *such that :*
   (α) $t_J = <t_I,t'>$
   (β) $t_I = t'_I$
   (γ) $t' \prec t$
2. $t_J = \bot_J$ *if and only if* $t_I = \bot_I$,
3. $t_J = g^n(a)_J$ *if and only if* $t = g^n(a)$.

Proof : We prove (1) by induction on t. If $t \in \{\Omega,a,b\}$ then (α), (β) and (γ) follow from the definition of J. Assume now that $t = g(u)$. If $g_I(u_I) = \bot$, then $t_J = <\bot,\Omega>$, this case is trivial. If $g_I(u_I) = t_I \neq \bot$, then $u_I \neq \bot_I$ by condition F_1 and $u_J = <u_I,u'>$ with $u' \prec u$ and $u'_I = u_I$. Then $t_J = <g_I(u_I), g(u')>$ and (α), (β), (γ) are satisfied. If $t = f(u_1,u_2)$ the proof is similar.

(2) The only if direction is obvious. The other one follows from the definition of J and (1).

(3) If $t_J = g^n(a)_J = \langle g^n(a)_I, g^n(a)\rangle$ then $g^n(a) \prec t$ by (1) and this implies $g^n(a) = t$. □

Lemma 4 : The interpretation J *belongs to* $\mathcal{D}_0$. *If* $I \in \mathcal{C}$ *then* $J \in \mathcal{C}$

Proof : We only verify F_4. If $\langle d,u\rangle \in D_J - \{\langle\bot,\Omega\rangle\}$ then $f_I(d,b_I) = d' \neq \bot_I$ since $I \in \mathcal{D}_0$. Then $b_J = \langle b_I,b\rangle$ and $f_J(\langle d,u\rangle,b_J) = \langle d',f(u,\Omega)\rangle$ if $f_I(d,\bot) \neq \bot$ (hence $f_I(d,\bot) = d'$)

$= \langle d',f(u,b)\rangle$ otherwise

by our definition of J.

If $f(g^n(a),\Omega)_I = d \neq \bot_I$, then $f(g^n(a),\Omega)_J = \langle d,f(g^n(a),\Omega)\rangle \neq \langle\bot,\Omega\rangle$. Hence, if $I \in \mathcal{C}$ then $J \in \mathcal{C}$.□

Proof of lemma 2 : In order to prove that $\leq_{\mathcal{D}_0}$ and $\leq_{\mathcal{C}}$ coincide on $M^\infty(F)$, we need only prove for all $t \in M(F)$ and $T' \in M^\infty(F)$, $t \leq_{\mathcal{C}} T'$ iff $t \leq_{\mathcal{D}_0} T'$. Since $\mathcal{C} \subset \mathcal{D}_0$ the if direction is trivial. We will prove that if $t_I \not\leq T'_I$ for some $I \in \mathcal{D}_0$, then $t_J \not\leq T'_J$ for some $J \in \mathcal{C}$ which will complete the proof.

First case : $t_I = d \neq \bot$ and $T'_I = t'_I = d' \notin \{\bot,d\}$ for some $t' \prec T'$. Let n be a positive integer such that $g^n(a)$ is not a subtern of t or t'. Let $I' = \mathcal{F}(I)$ and $J = I(n)$ be the discrete interpretation defined as follows :

$$D_J = D_{I'}$$

$$\bot_J = \bot_{I'} \ ,\ a_J = a_{I'} \ ,\ b_J = b_{I'} \ ,\ g_J = g_{I'}$$

$$f_J(d_1,d_2) = f_{I'}(d_1,d_2) \text{ if } d_1 \neq g^n(a)_{I'}$$

$$= \delta \quad \text{if } d_1 = g^n(a)_{I'} \text{, where } \delta \text{ is some fixed element of } D_J - \{\langle\bot,\Omega\rangle\}$$

The verification that f_J is increasing w.r.t. the discrete ordering is left to the reader. It suffices to prove that J is a discrete interpretation . Furthermore, $J \in \mathcal{D}_0$ We only verify that F_4 is satisfied, the other verifications being similar.Let $d_1 \in D_J - \{\bot_J\}$. If $d_1 = g^n(a)_{I'}$ then $f_J(d_1,b_I) = \delta \neq \bot_J$. If $d_1 \neq g^n(a)_{I'}$, then $f_J(d_1,b_I) = f_{I'}(d_1,b_I) \neq \bot_J$ since $I' \in \mathcal{D}_0$. Clearly $f(g^n(a),\Omega)_J = f_J(g^n_J(a_J),\bot)$ $= f_J(g^n_{I'}(a_{I'}),\bot) = \delta \neq \bot_J$ hence $J \in \mathcal{C}$.

Finally we prove that $t_J \neq \bot$ and $t'_J \neq t_J$.

Let $u \in M(F)$ such that :

(1) for every subterm u' of u , $u'_{I'} \neq g^n(a)_{I'}$.

An easy inductive argument shows that $u_J = u_{I'}$. By lemma 3, $u'_{I'} = g^n(a)_{I'}$, if and only if $u' = g^n(a)$, hence, by our choice of n, $t_J = t_{I'}$ and $t'_J = t'_{I'}$. From the initial assumption that $t_I \neq t'_I$, we get $t_{I'} \neq t'_{I'}$ and $t_J \neq t'_J$. Similarly, $t_I \neq \perp$ implies $t_{I'} \neq \perp$ and $t_J \neq \perp$. The same holds for t'. Hence $t_J \not\leq T'_J$. Q.E.D.

Second case : Let $I \in D_0$ such that $t_I \neq \perp$ and $t'_I = \perp$ for all $t' \in M(F)$ such that $t' \prec T'$. We aim to define $J \in C$ such that $t_J \not\leq T'_J$. Choose an integer n such that $g^n(a)$ is not a subterm of t. Let I(n) be associated with I and n as in the first case. If $t'_{I(n)} = \perp$ for every $t' \prec T'$, then we can take J = I(n) and we are done since $t_{I(n)} \neq \perp$ as in the first case. Otherwise, if $t'_{I(n)} \neq \perp$ for some $t' \prec T'$, let us take $J = F(I(n))$. This interpretation belongs to C since $I(n) \in C$ and by lemma 4. Clearly, $t_J \neq \perp$ and $t'_J \neq \perp$. Let us prove that $t_J \neq t'_J$. Assume that $t_J = t'_J = \langle d,u \rangle$ for some $d \in D_{I(n)}$ and $u \in M(F)$. By lemma 3, $u \prec t$ and u has no subterm of the form $g^n(a)$ by our choice of n. Hence $u_{I(n)} = u_{I'}$. Also $u \prec t'$ and $t'_I = \perp$. This implies $u_I = \perp$ and $u_{I'} = \perp$ (since $I' = F(I)$). Hence $d = u_{I(n)} = \perp$ by lemma 3 again, but $d = t_{I(n)}$ and we get a contradiction since $t_{I(n)} \neq \perp$, by the proof of the first case. Hence we have proved that $t_J \not\leq T'_J$. □

7 - Conclusion

1. Most classes of interpretations which naturally occur when one considers recursive program schemes are of the form D(A), hence are algebraic by theorem 2. In order to perform proofs of valid relations $\phi \leq_{D(A)} \phi'$ by computation induction, one needs a characterization of $t \leq_{D(A)} t'$ for $t,t' \in M(F)$. In most cases, $t \leq_{D(A)} t'$ will be undecidable but a semi-decidability result is still of interest for proofs about programs.

2. To summarize, algebraic classes are essentially of the following types :

1) C_R (relational) ,

2) D(A) (first-order discrete). (*)

Most other definitions yield nonalgebraic classes.

References :

[0] E. Ashcroft, Z. Manna, A. Pnueli, "*Decidable properties of monadic functional schemas*", JACM Vol. 20 (1973) pp. 489-499.

(*) discrete can be replaced by bounded : increasing chains have a bounded length. This will be considered in a forthcoming paper.

[1] G. Berry, B. Courcelle, *"Program Equivalence and Canonical forms in stable discrete interpretations"*, Proc. of 3rd colloqium on Automata, Languages and Programming, Edinburgh, (1976). pp. 168-188.

[2] R. Burstall, *"Proving properties of programs by structural induction"*, Comp. Journal (12) pp. 41-48.

[3] R. Burstall, J. Darlington, *"A system which automatically improves programs"*, Proc. of 3rd Int. joint conf. artif. intelligence, S.R.I. (1973) pp. 537-542.

[4] B. Courcelle, M. Nivat, *"Algebraic Families of interpretations"*, 17th symp. on F.O.C.S., Houston (1976) (also available as Laboria Report n° 189).

[5] S.A. Greibach, *"Theory of program Structures : Schemes, Semantics, Verifications"* Lecture Notes in Computer Science Vol 36 (1975) Springer Verlag.

[6] I. Guessarian, *"Semantic equivalence of program schemes"*, Proc. of 3rd coll. on Aut., Languages and Programming, Edinburgh (1976), pp. 189-200.

[7] J. Kfoury, D. Park, *"On termination of program schemas"* Inf. and Control Vol 29 (1975) pp. 243-251.

[8] R. Milner, *"Models of LCF"*, Memo stanford. AIM-184 (1973).

[9] J.H. Morris, *"Another Recursion Induction Principle"*, CACM Vol 14, n°5 (1971).

[10] M. Nivat, *"On the interpretation of polyadic recursive schemes"*, symp. Mathematica Vol 15, Academic Press (1975).

[11] D. Park, *"Finiteness is μ-ineffable"*, Report n°3, Warwick University (1974).

[12] J.R. Schoenfield, *"Mathematical Logic"*, Addison- Wesley (1967).

[13] B. Courcelle, I. Guessarian, M. Nivat, *"The algebraic semantics of recursive program schemes"*, submitted for publication.

[14] A. Chandra, *"Generalized program schemas"* SIAM J. on computing Vol 5(1976) pp. 402-413.