Rev.R.Acad.Cien.Exact.Fis.Nat. (Esp)
Vol. 92, N.º 4, pp 299-306, 1998
Monográfico: Problemas complejos de decisión

A LOGIC AND COMPUTER ALGEBRA APPROACH TO A DECISION-MAKING PROBLEM IN MEDICINE[1]

(hidrogenación/dióxido de carbono/rodio soportado)

LUIS M. LAITA*, EUGENIO ROANES-LOZANO**, VÍCTOR MAOJO and A. DÍAZ***

* *Dept. Inteligencia Artificial,* Universidad Politécnica de Madrid. laita@fi.upm.es Académico. Correspondiente de la Real Academia de Ciencias.

** *Dept. Álgebra,* Universidad Complutense de Madrid. eroanes@eucmos.sim. ucm.es

*** *Dept. Inteligencia Artificial,* Universidad Politécnica de Madrid. vmaojo@infomed.dia.fi.upm.es

ABSTRACT

In this article, Expert Systems techniques, based on Logic and Computer Algebra, are applied to particular decisión-making procedures in Medicine. A result, relating tautological consequence and the ideal membership problem, stands at the basis of these techniques and allows and efficient implementation in CoCoA.

1. INTRODUCTION

This article addresses the application of Expert Systems based techniques (related to Logic and Computer Algebra) to the field of evaluation of decision-making criteria on the fitness of medical diagnosis. The article is complementary to (6).

The decision-making criteria on the fitness of medical diagnosis are usually collected in tabular format (4). Nevertheless, no efficient mechanical method for both evaluating the correctness and extracting new knowledge from the table exists. Such a limitation entails serious difficulties for the members of the medical profession who use the tables, because these tables sometimes contain more than a thousand information items.

In this article we describe an efficient computational method, based on logic and algebra, that automatically evaluates the correctness of those tables and extracts new information from the information implicitly contained in the tables. The method requires a prior translation of the tables into an Expert System based on many-valued and modal logic (that allows imprecise knowledge to be handled). As an illustration, a table drawn up by experts in coronary diseases (8) will be studied.

An expert system consists of a «knowledge base» and an «inference engine». The knowledge base consists, in the rule-based expert systems (denoted «RBES»), of «rules», «facts» and some other information items (of which the most important are the «integrity constraints»). The inference engine is a program oriented to the extraction of consequences from the knowledge base. An inference engine based on Computer Algebra and developed by the authors is used in this article.

The rules in the knowledge base are logical formulae that translate assertions (also called «knowledge items» in the rather inadequate Artificial Intelligence jargon). These assertions have the form «IF the conjunction of such and such presuppositions holds, THEN the disjunction of such and such conclusions also holds». The logical symbols that translate presuppositions and conclusions are called «literals».

The facts are simple logical formulae that translate simple knowledge. They are divided into «potential facts» and «given facts». Potential facts are all the literals that, belonging to the conjunction on the left-hand side of any rule, do not belong to the disjunction on the right-hand side of any rule. Given facts are all or some of the potential facts that an expert or a user have singled out for particular purposes.

Integrity constraints (referred to as IC) are logical formulae that translate the condition, assessed by the experts that have provided the information for constructing the

[1] This article is, partially, product of a project of DGES, Programa de Promoción General del Conocimiento, Ministerio de Educación y Cultura, Project PB96-0098-034, and project FIS 95/1952.

table, that two or more facts cannot hold at the same time. The negation (NIC) of each IC must be added to the RBES as new information.

1.0.1. Example. *Let us take a look at an elementary example of a RBES (based for the sake of simplicity on Boolean logic) that illustrates these concepts. Also for the sake of simplicity, the consequents consist of only one element, despite the fact that they generally are disjunctions of elements. As is known, $\neg$, $\wedge$, $\vee$ and $\rightarrow$ are named «connectives» in logic, respectively meaning, «no», «and», «or» and «implies».*

Rule 1. $A \wedge \neg B \rightarrow C$

Rule 2. $C \rightarrow D$

Rule 3. $D \rightarrow \neg E$

Rule 4. $F \rightarrow E$

Letters like A, or letters preceded by $\neg$, such as $\neg B$, are examples of literals.

Potential facts are A, $\neg B$ and F, and, as mentioned above, «fact» is any potential fact which it is of interest to single out. A rule «is fired» iff all the literals in the antecedent are facts (or «derived facts», which are described below in an example). Firing corresponds to the formal logic rule of «modus ponens».

In some cases, the consequent of one rule is a part of the antecedent of another rule, such as C in Rules 1 and 2. If A and $\neg B$ are facts, then by firing R1, C is obtained; and from C and the Rule 2, D is obtained. Thus the literals C and D are derived facts.

An integrity constraint (denoted IC) is, for example, $C \wedge \neg E$. We have thus new information to be added to the RBES, the negation NIC: ($C \wedge \neg E$) of the IC: $C \wedge \neg E$. Any other additional information is denoted as «ADDI».

Suppose that in a RBES formed only by the rules 1,2 and 3, to which the formula NIC is added, A and $\neg B$ are facts. In this case, the firing of the rules gives the IC, which contradicts the NIC.

If there were an RBES formed only by the rules 3 and 4 and the facts D and F, the firing of the rules would lead to the logical contradiction $E \wedge \neg E$

1.0.2. Definition. *A RBES will be said to be «inconsistent» with respect to some given facts iff a logical contradiction or an integrity constraint is obtained when its rules are fired. The inconsistency obtained from an integrity constraint IC becomes a logical inconsistency when the corresponding NICs are added to the RBES. An algebraic translation of inconsistency is given in 3.1.*

In the three-valued and modal logic-based RBES to be considered in this article, the literals are propositional variables, preceded or not by the connective $\neg$ or by two new connectives: $\Box$ (which means «it is necessary that») and $\Diamond$ (which means «it is possible that»), or by any admissible combination of these symbols. The rules have the following format:

$$\sigma_1 X[1] \wedge \sigma_2 X[2] \wedge \ldots \wedge \sigma_n X[n] \rightarrow \sigma_{n+1} X[n+1] \, V \ldots V \, \sigma_s X[s].$$

where the symbols «σ_i» represent $\neg$, $\Box$, $\Diamond$, any of the admissible combinations, or even the empty symbol. If, in addition, imprecise knowledge needs to be expressed about a rule, the symbols $\Box$, $\Diamond$, and their combinations, with or without $\neg$, must be written before the entire rule and/or its antecedent and/or its consequent.

In the logic on which the RBES to be addressed relies (Lukasiewicz's three-valued logic augmented with modal operators), if A is an IC, its NIC is not the simple negation of A, but the formula $\Box\neg$A.

The truth-values of this logic are determined by the functions:

$$H_{\neg}, H_{\Diamond}, H_{\Box} : \mathbb{Z}_3 \rightarrow \mathbb{Z}_3$$

$$H_{\vee}, H_{\wedge}, H_{\rightarrow} : \mathbb{Z}_3^2 \rightarrow \mathbb{Z}_3$$

defined in the tables below (0 represents «false», 1 represents «not-determined» and 2 represents «true».

$H_{\neg}$	
0	2
1	1
2	0

$H_{\Diamond}$	
0	0
1	2
2	2

$H_{\Box}$	
0	0
1	0
2	2

$H_{\vee}$	0	1	2
0	0	1	2
1	1	1	2
2	2	2	2

$H_{\wedge}$	0	1	2
0	0	0	0
1	0	1	1
2	0	1	2

$H_{\rightarrow}$	0	1	2
0	2	2	2
1	1	2	2
2	0	1	2

By means of the standard generalization, these functions can be used to define a truth-valuation v for any formula.

1.0.3. Definition. *A propositional formula Q_0 is a «tautological consequence» of the propositional formulae $Q_1, Q_2, \ldots, Q_m$ (denoted $\{Q_1, Q_2, \ldots, Q_m\} = Q_0$) iff for any truth valuation v, such that $v(Q_1) = v(Q_2) = \ldots = v(Q_m) =$ 2 (true), then $v(Q_0) = 2$ (true). In the general case of a p-valued logic, $p - 1$ substitutes 2 (see section 3).*

1.0.4. Definition. *The set $\{Q_1, Q_2, \ldots, Q_m\}$ is a «contradictory domain» iff for any formula Q of the language in which $Q_1, Q_2, \ldots, Q_m$ are expressed, $\{Q_1, Q_2, \ldots, Q_m\} =$*

Q. The name «contradictory domain» is justified bacause if any formula follows from $\{Q_1, Q_2, ..., Q_m\}$*, the formulae that translate contradiction will also follow.*

2. DESCRIPTION OF THE TABLE AND TRANSLATION TO A RBES

2.1. Description of the Table

A set of symptoms and other medical data (effort test positive, negative or not done, disease of one, two, or three blood vessels, LVEF value) was presented to a group of ten experts on coronary diseases. They were then asked about the pertinency of taking certain actions (revascularization, PTCA, CABG). LVEF means «Left Ventricle Ejection Fraction», CABG means «Coronary Artery Bypass Grafting» and PTCA means «Percutaneous Transluminal Coronary Angioplasty».

Both the data presented to the experts and their opinions were collected in a table containing 216 information items. Only some of these items, sufficient to understand the argument in the article, are transcribed below. This information will first be translated to rules of a RBES, to be later processed computationally.

A. Effort test positive

A.1. Left common trunk disease

A.1.1. Surgical risk low/moderate%

% LVEF (F)	Revascularization	PTCA	CABG
F > 50	**1:** $12345678^{1}*^{9}$ + *A*	**2:** $12345678^{1}*^{9}$	– + *A*
$50 \geq F > 30$	**5:** $12345678*^{10}$ + *A*	**6:** $12345678^{1}{}_{*}^{9}$	– + *A*
$30 \geq F > 20$	**9:** $12345678^{1}*^{9}$ + *A*	**10:** $1234567^{1}8^{1}*^{8}$	– + *A*

A.1.2. Surgical risk high

% LVEF (F)	Revascularization	PTCA	CABG
F > 50	**3:** $12345678^{4}*^{6}$ + *A*	**4:** $12345^{1}67^{1}8^{1}*^{7}$	– + *A*
$50 \geq F > 30$	**7:** $12345678^{3}*^{7}$ + *A*	**8:** $123456^{1}7^{1}8^{2}{}_{*}^{6}$	– + *A*
$30 \geq F > 20$	**11:** $1234567^{1}8^{2}*^{7}$ + *A*	**12:** $123456^{2}78^{2}*^{6}$	– + *A*

The experts were informed in the six cases **1, 2, 5, 6, 9, 10** that, for a certain patient, the data are: effort test positive, suffers from left common trunk disease, surgical risk is low/moderate and LVEF is in a given percentage bracket. In the cases **3, 4, 7, 8, 11, 12** the datum «surgical risk low/moderate» changes to «high», while the other data are unchanged. The experts were then asked about the pertinency of revascularization, PTCA and CABG.

Let us proceed with the transcription of the cases we have chosen in order to explain the meaning of the digits, letters and symbols used.

A.2. Three blood vessels disease

A.2.2. Surgical risk high%

% LVEF (F)	Revascularization	PTCA	CABG
$F \geq 50 > 30$	....	**20:** $123^{3}4^{2}5*78^{2}9^{3}$	+ + *D*

A.4. Two blood vessels disease, proximal anterior descendent not affected

A.4.2. Surgical risk high

% LVEF (F)	Revascularization	PTCA	CABG
$F \geq 50$	....	**40:** $1^{5}*2^{1}3^{3}45^{1}6789$	? – *A*

A number **(1, 2, ..., 216)** has been assigned to each row of digits, symbols +, –, *, and letters *A, D, I* of the table. The rules R1, R2,..., R216 of the RBES to which the table will be translated are numbered according to these numbers.

Each superscript expresses the number of experts that assigned the value, from 1 to 9, to the fitness of the action. For example, in **1**, one expert assigned the value 8 and nine experts the value 9 to the fitness of revascularization.

The symbol «*» indicates the median.

The information given by the superscripts is resumed in the table as follows. The symbol «+» means *it is unfit*, the symbol «–» means *it it is unfit*. The symbol «?» means *undecided fitness.*

The letter *A* means that there is *agreement* among the experts about the fitness, unfitness or undecided fitness of revascularization, PTCA or CABG. The letter *D* means *disagreement*. The letter *I* means *undecided agreement*. There is disagreement when the opinion of at least three experts is reflected as superscripts over the digits from 1 to 3 and the opinion of at least another three as superscripts over the digits from 7 to 9. There is agreement when there are no more than two opinions reflected outside an interval of [1,3], [4,6] or [7,9] that contains the median. In any other case, there is undecided agreement *I*.

2.2. Translation of the Table to a RBES

The relation between the four data and the corresponding list of digits and symbols in, for instance, **1** can be reinterpreted as the statement: «*IF* the effort test of a patient is positive *AND* he suffers from left common trunk disease *AND* his surgical risk is low/moderate *AND* his LVEF is over 50\%, *THEN* the experts have assessed that revascularization is fit»; moreover, «there is agreement (*A*) in this assessment». These *«IF-THEN»* assertions can be translated to logical formulae, which are precisely the rules of which the RBES will consist.

The first step to take is to assign a propositional variable, denoted *X*[*i*] (possibly preceded by ¬), to each datum and therapeutic action. All the variables are noted, even though not all of them will be used in the article.

- Surgical risk low/moderate: $\neg X[1]$, high: $X[1]$.
- Effort test positive: $X[2]$, negative: $\neg X[2]$, undetermined: $\Diamond\neg X[2]$ (translated as «it is possible that the effort proof is negative»).
- Common left trunk disease: $X[3]$, three vessels disease: $X[4]$, two vessels disease : $X[5]$, one vessel disease: $X[6]$.
- Anterior proximal vessel affected: $X[13]$, not affected: $\neg X[13]$.
- LVEF > 50%: $X[7]$, 50% $\geq$ LVEF > 30%: $X[8]$, 30% LVEF $\geq$ 20%: $X[9]$.
- Revascularization: $X[10]$, PTCA: $X[11]$, CABG: $X[12]$.

These variables combine to form rules, under the following conventions.

The symbols $\Box$ and $\Diamond$ do not precede the variables (or their negations) if they represent data. The reason is that the data should be taken into account only if they are presented under either a high or acceptable degree of certainty. A high certainty degree for, for instance, $X[3]$ is represented as $\Box X[3]$, and acceptable, as simply, $X[3]$. As in logic the implications $\Box X[3] \rightarrow X[3]$ and $X[3] \rightarrow X[3]$ both hold, if the data are given under either a high or an acceptable degree of certainty, the rules that contain them in their antecedent can be fired. But if given under a low certainty degree (represented by $\Diamond X[3]$ in our example), such as $\Diamond X[3] \not\rightarrow X[3]$, the rules that contain them in their antecedent cannot be fired.

The different possibilities are translated as follows (note that they are listed in decreasing order of fitness):

x	$+A$	$\rightarrow$	$\Box x$
x	$+I$	$\rightarrow$	x
x	$+D$	$\rightarrow$	$\Diamond x$
x ? (the three cases)		$\rightarrow$	*tautology*
x	$-D$	$\rightarrow$	$\Diamond\neg x$
x	$-I$	$\rightarrow$	$\neg x$
x	$-A$	$\rightarrow$	$\Box\neg x$

Unlike the approach taken in (6), we chose on this occasion to translate the information items that contain «?» by a trivially true implication. The reason is that we used the convention of not inferring any conclusion from rules where pertinency is not decided. This convention is lax compared with the one used in [6], but it is better suited for detecting only serious logical contradictions rather than those produced, for instance, by the concurrence of different experts opinions on the cases «?».

Under the translation used in (6), one conclusion was that some rules leading to contradiction should be changed (the suggestions that we made to the experts as a result were judged acceptable). Here, we tried an approach which is as close as possible to the original table. Nevertheless, contradictions arise in this approach too, as will be shown later.

For example, information items **1, 2, 3** and **4** are translated to RBES rules as follows.

R1: $\neg X[1] \wedge X[2] \wedge X[3] \wedge X[7] \rightarrow \Box X[10]$

R2: $\neg X[1] \wedge X[2] \wedge X[3] \wedge X[7] \rightarrow \Box\neg X[11] \wedge \Box X[12])$.

R3: $X[1] \wedge X[2] \wedge X[3] \wedge X[7] \rightarrow \Box\neg X[10]$

R2: $X[1] \wedge X[2] \wedge X[3] \wedge X[7] \rightarrow \Box\neg X[11] \wedge \Box X[12])$.

An example of a rule that expresses disagreement ($\Diamond$), is (see item **20** in the table):

R20: $\neg X[1] \wedge X[2] \wedge X[4] \wedge X[8] \rightarrow \Diamond X[11] \vee \Diamond X[12])$.

2.2.1. Observation. *Why use $\wedge$ in some rules and $\vee$ in others? Simply, for the convenience of expressing expert knowledge more adequately. For instance, there is no conflict in declaring that PTCA is appropriate and that CABG is not in R2:* $\Box\neg X[11] \wedge \Box X[12])$. *But in R20, we chose to write $\vee$ between two affirmations of pertinency (under disagreement) $\Diamond X[11]$ and $\Diamond X[12]$, because, even though the experts judge both actions (PTCA and CABG) appropriate, the external information that both cannot be simultaneously applied (this is expressed as* $\Box\neg X[11] \wedge X[12]$) *has to be taken into account. This formula is the NIC corresponding to the IC:* $X[11] \wedge X[12]$.

2.3. About the consistency of the RBES

The translation of the table into an RBES allows two important processes to be carried through: verify if the table contains anomalies and, once these have been corrected, extract new knowledge.

Let us consider, for instance, the rule that corresponds to item **40** of the table.

R40: $X[1] \wedge X[2] \wedge X[5] \wedge\neg X[13] \wedge X[7] \rightarrow$ *tautology* $\wedge \Box\neg X[12]$.

What would happen if a patient whose data are $X[1]$, $X[2]$ and $X[7]$, suffered at the same time from left common trunk disease ($X[3]$) and two vessels disease with anterior proximal descendent not affected $X[5] \wedge\neg X[13]$)?

In this situation, as both rules R4 and R40 can be fired, a simple visual logical examination of the consequents obtained shows a contradiction (rule 4} strongly rejects PTCA and strongly recommends CABG, whereas rule **40** is inconclusive about PTCA but strongly rejects CABG). Thus, there exists an anomaly in the table.

Let us analyze the different alternatives:

i) It is impossible for a patient to suffer from both diseases or it is possible that he suffers from both but they are not going to be treated simultaneously.

In such a case, the maximal consistent sets of facts must be carefully determined. Now, for instance, $x[7]$, $x[8]$ and $x[9]$ (% LVEF: > 50, >30 and ≤ 50, ≥ 20 and ≤ 30) are mutually exclusive. Similarly, the different diseases ($x[3]$, $x[4]$, $x[5] \wedge x[13]$, $x[5] \wedge \neg x[13]$) are also mutually exclusive.

ii) It is possible that a patient suffers from both diseases and they are going to be treated simultaneously.

a) There exists the possibility of treating one of the two diseases with one technique and the other with the other technique.

The same solution as in the previous paragraph would be adequate here.

b) It has to be decided which technique to apply: it is not possible to apply one technique to one disease and a different technique to the other disease.

This conflict has a simple but very interesting solution from the strictly logical viewpoint. The conflict vanishes if, for instance, rules R4 and R40 are substituted by NR4 and NR40:

NR4: $X[1] \wedge X[2] \wedge X[3] \wedge X[7] \rightarrow \Box\neg X[11] \wedge \Diamond X[12])$.

NR40: $X[1] \wedge X[2] \wedge X[5] \wedge X[13] \wedge X[7] \rightarrow$ *tautology* $\wedge \Diamond\neg X[12])$.

Clearly, an interaction with the experts is necessary in order to search for the best representation of knowledge from the medical viewpoint. In particular, it seems that the case in which two diseases are to be simultaneously treated was not considered by the experts who produced the table.

The logic-algebraic theory and its implementation in the CoCoA \cite{Capani} language that we outline below allow this type of contradictions to be detected automatically. Indeed, the above-mentioned contradiction (and many others) were found this way.

3. TAUTOLOGICAL CONSEQUENCE AND THE IDEAL MEMBERSHIP PROBLEM

In this section, a result that relates the concept of tautological consequence with the ideal membership problem in a polynomial ring, actually $\mathbb{Z}_p[x_1, x_2, ..., x_n]$, where p is any prime number (for example, $p = 3$ for the three-valued case) is described.

The numbers $0, 1, ..., p-1$, are considered as the truth values of a p-valued and modal logic. For instance, $p-1$ can represent the value «true», 0 can represent «false» and the reminding elements can represent intermediate truth values.

Let $X = \{X_1, X_2, ..., X_n\}$ and $C = \{c_1, c_2, ..., c_t\}$ be a set of propositional variables and a set of connectives, respectively . In the three-valued and modal case $C = \{c_1 = \neg, c_2 = \Diamond, c_3 = \Box, c_4 = \vee, c_5 = \wedge, c_6 = \rightarrow\}$. $P_C(X_1, X_2, ..., X_n)$ represents the set of well-formed propositional formulae from X and C. The letter Q (with or without subscripts) represents a generic element of $P_C(X_1, X_2, ..., X_n)$.

$\mathbb{Z}_p[x_1, x_2, ..., x_n]$ is the polynomial ring in the variables $x_1, x_2, ..., x_n$ with coefficients in $\mathbb{Z}_p$. The polynomial variables $x_1, x_2, ..., x_n$ correspond to the propositional variables $X_1, X_2, ..., X_n$ of X respectively. The letters q and r represent generic elements (generic classes of polynomials) of $\mathbb{Z}_p[x_1, x_2, ..., x_n]/I$, where I is the ideal: $< x_1^p - x_1, x_2^p - x_2, ..., x_n^p - x_n >$ (i.e., the ideal generated by the polynomials $x_1^p - x_1, x_2^p - x_2, ..., x_n^p - x_n$).

A class of polynomials is assigned to each formula of $P_C(X_1, X_2, ..., X_n)$, by firstly defining a function:

$$f_j : \left(\mathbb{Z}_p[x_1, x_2, ..., x_n]/I\right)^{s_j} \rightarrow \mathbb{Z}_p[x_1, x_2, ..., x_n]/I$$

where s_j is the «arity» of each connective c_j.

In the case of binary connectives (the unary case is simpler), f_j has the form:

$$f_j(q, r) = \sum_{i=0}^{p-1}\sum_{k=0}^{p-1} a_{i,k}\, q^i r^k + I \,;\; a_{i,k} \in \mathbb{Z}_p$$

and it passes through the points $\{(i, k, H_j(i, k)) : i, k \in \{0, 1, ..., p-1\}\}$, where the H_j are the functions of which $H_\vee$, $H_\wedge$, $H_\rightarrow$ mentioned in the introduction are particular cases.

A Maple V program, which is not transcribed here for the sake of brevity, determines (following the Lagrange interpolation method) the coefficients $a_{i,k}$. In particular, for Lukasiewicz's three-valued logic with modal operators (where 0, 1, 2 denote false, undetermined, true, respectively), it provides the following translations of the basic logical formulae to polynomial classes

$$f_\neg(q) = \left(2+2_q\right) + I$$

$$f_\vee(q, r) = \left(q^2r^2 + q^2r + qr^2 + 2qr + q + r\right) + I$$

$$f_\wedge(q, r) = \left(2q^2r^2 + 2q^2r + 2qr^2 + qr\right) + I$$

$$f_\rightarrow(q, r) = \left(2q^2r^2 + 2q^2r + 2qr^2 + qr + 2q + 2\right) + I$$

The function θ below, interacting with functions, f_j, translates any propositional formula to a polynomial class:

$$\theta : P_C(X_1, X_2, ..., X_n) \rightarrow \mathbb{Z}_p[x_1, x_2, ..., x_n] \,/\, I$$

defined recursively:

$$\theta(X_i) = x_i + I, \text{ for all i } = 1,..., \text{n}.$$

$$q(Q) = f_j\left(\theta(Q_1),..., \theta\left(Q_{s_j}\right)\right), \text{ if } Q \text{ is } c_j\left(Q_1, ..., Q_{s_j}\right).$$

The reduction of logical formulae to classes of polynomials allows the development of an algebraic theory leading to the following result, which we enunciate without proof.

3.0.1. Idea. *A formula Q_0 is a tautological consequence of other formulae $Q_1, ..., Q_m$, iff the polynomial that translates the negation of Q_0 belongs to the ideal generated by the polynomials that translate the negations of $Q_1,..., Q_m$.*

Formally:

3.0.2. Theorem. *Let $Q_0, Q_1, ..., Q_m \in P_C(X_1, X_2, ..., X_n)$. The following assertions are equivalent:*

(i) $\{Q_1, Q_2, ..., Q_m\} \models Q_0$

(ii) $f_\neg(\theta(Q_0)) \in \; < f_\neg(\theta(Q_1)), ..., f_\neg(\theta(Q_m)) >.$

The first proof of this theorem appeared in [1] and was improved in [3]. Another approach, taken by the authors of the present article from the viewpoint of Algebraic Geometry, can be found in [9].

It is well known that to check whether a polynomial belongs to an ideal, it has to be determined whether or not the «Normal Form» *(NF)* of the polynomial, modulo the ideal, is 0. Effective methods, such as Gröbner Bases', exist (see, for instance, [10]).

3.1. Application to the study of consistency

A set of propositional formulae $\{Q_1, Q_2, ..., Q_m\} \subseteq P_C(X_1, X_2, ..., X_n)$ is inconsistent when $\{Q_1, Q_2, ..., Q_m\}$ is a contradictory domain (see section 1). This, in algebraic terms, is translated into the ideal J, generated by the negations of $Q_1, Q_2, ..., Q_m$, being the whole ring, which happens if and only if $1 \in J$ in $\mathbb{Z}_p[x_1, x_2, ..., x_n]/I$. This last condition is checked by calculating whether or not the Gröbner basis of the ideal is $\{1\}$.

This method of verification of inconsistency is new [7] and differs substantially from the known approaches (a description of Expert Systems verification methods appears in [5]).

CoCoA cannot work in quotient rings so far. Thus, instead of checking whether the ideal J of $\mathbb{Z}_p[x_1, x_2, ..., x_n]/I$ is the whole ring, we will check whether the ideal $I + J$ of $\mathbb{Z}_p[x_1, x_2, ..., x_n]$ is the whole ring.

3.2. Application to the extraction of new knowledge

As a result of the above theorem, to check whether or not any given formula that translates a certain information item follows from the RBES, we simply have to check whether the polynomial translation of its negation belongs to the ideal generated by the polynomial translation of the negations of the facts, rules, etc. considered. As mentioned above, it is well known that this last condition holds if and only if the «normal form» of the polynomial modulo the ideal is 0.

4. IMPLEMENTATION IN CoCoA

CoCoA (Computations in Commutative Algebra) is a very efficient algebraic computation language, specialized in calculations of Gröbner bases in polynomial rings over finite characteristic fields[1]. An implementation in this language of the process of detecting inconsistencies in (a part of) the RBES described in the article is presented below.

CoCoA does not admit the definition of infix operators, so that it requires that logical formulae be written in prefix form. NEG, POS, NEC, OR1, AND1, IMP will denote $\neg$, $\Diamond$, $\Box$, $\vee$, $\wedge$, $\rightarrow$, respectively.

The polynomial ring A and the ideal I are declared first (see section 3).

A : : =Z (3) [x[1...13]];

USE A;

I := Ideal (x[1]^3 –x[1], ..., x[13]^3 –x[13];

The connectives are translated into polynomials of the quotient ring A/I as follows (see the form of the functions f_j in section 3). NF means «Normal Form» (OR1 and AND1, respectively, are used instead of OR and AND, because the latter are reserved words in CoCoA).

NEG (M) : = NF(2+2*M, I);

POS (M) : = NF(2*M^2, I);

NEC (M) : = NF(M^2+2*M, I);

OR1 (M, N) : = NF (M^2*N^2+M^2*N+M*N^2+2*M*N+M+N, I);

[1] More information about CoCoA can be obtained from cocoa@dima.unige.it or directly at the web page: http: // cocoa.dima.unige.it

AND1 (M, N) : = NF (2*M^2*N^2+2*M^2*N+2*M*N^2+M*N, I);

IMP (M, N) : = NF(2*M^2*N^2+2*M^2*N+2*M*N^2+M*N+2*M+2, I);

Subsequently the rules are written in prefix form, for example:

R8:=IMP (AND1(AND1(AND1(x[1],x[2]),x[3]),x[8]), NEC (AND1(NEC(NEG(x[11])),NEC(x[12]))));

R20:=IMP(AND1(AND1(AND1(x[1],x[2]),x[4]),x[8]), OR1(POS(x[11]),POS(x[12])));

and the potential facts of these rules are declared:

F2:=NEG(x[1]); F3:= x[2]; F4:= x[3]; F5:= x[4]; F8:= x[8];

In order to define the ideal generated by the polynomials that translate the negations of the facts and rules, it suffices to write down:

K:=Ideal(NEG(F2), NEG(F3), NEG(F4), NEG(F5), NEG(F8), NEG(R8), NEG(R20));

4.1. Detection of Inconsistency with CoCoA

Let us suppose that we have the case ii)b) of 2.3 (the most interesting case from the logical viewpoint).

In order to determine whether or not {F2, F3, F4, F6, F7, F10, R4, R40} is consistent, it is enough to calculate the Gröbner basis of the corresponding ideal:

R4:=IMP(AND1(AND1(AND1(x[1], x[2]), x[3]), x[7]), AND1(NEC(NEG(x[11])), NEC(x[12])));

R40:=IMP(AND1 (AND1 (AND1 (AND1(x[1], x[2]), x[5]), NEG(x[13])), x[7]), AND1 (2, NEC (NEG(x[12]))));

(constant 2 represents the tautology)

J4:=Ideal(NEG(F2), NEG(F3), NEG(F4), NEG(F6), NEG(F7), NEG(F10), NEG(R4),NEG(R40));
GBasis(I+J4);
[1]

and thus, there is inconsistency.

It is important to note that, really, the process of RBES verification is more complex than it seems to be. Initially, a lot of rules must be examined at the same time (in our case, we started with 48). If the program detects inconsistency, it is necessary to exactly locate the rules that produce it in order to debug the problem. For this purpose, we use a simple program, which we have named CONSIST, that takes the set of facts, NICs and ADDIs and adds rules one by one, checking the consistency of the system before adding the next one. The example shown here presents the simplest case: a direct incompatibility between two rules has been detected and located (there are also other problems).

The solution to the conflict proposed in 2.3 can be checked immediately with CoCoA:

NR4:=IMP(AND1 (AND1 (AND1 (x[1], x[2]), x[3]), x[7]), AND1(NEC(NEG(x[11])), POS(x[12])));

NR40:=IMP(AND1 (AND1 (AND1 (AND1 (x[1], x[2]), x[5]), NEG(x[13])), x[7]), AND1 (2, POS(NEG(x[12]))));

J5:=Ideal(NEG(F2), NEG(F3), NEG(F4), NEG(F6), NEG(F7), NEG(F10), NEG(NR4),NEG(NR40));}

GBasis(I+J5);}

is not [1]} and thus there is consistency.

4.2. Extraction of New Knowledge with CoCoA

Let us determine whether

$\Diamond\neg(\Box X[11] \vee \Box X[12])$

$\Diamond\neg X[11] \rightarrow \neg\Diamond X[12])$

$\Diamond\neg X[11] \wedge\neg X[12])$

follow from the rules R3, NR4, R7, R8, R11, R12, R20, R39, NR40, the facts F2, F3, F4, F5, F6, F7, F10 and the NIC:

J6:=Ideal(NEG(F2), NEG(F3), NEG(F4), NEG(F5), NEG(F6), NEG(F7), NEG(F10), NEG(R3), NEG(NR4), NEG(R7), NEG(R8), NEG(R11), NEG(R12), NEG(R20), NEG(R39), NEG(NR40), NEG(NIC));
GBasis(I+J6);

is not [1], thus there is no inconsistency. Therefore it makes sense to check if the above mentioned formulae follow from these rules, facts and the NIC (if there were inconsistency, they would trivially follow):

NF(POS(NEG(OR1(NEC(x[11]), POS(x[12])))), I+J6);
0

(thus the first formula follows)

NF(POS(NEG(IMP(POS(NEG(x[11])), POS(NEG(x[12])))), I+J6);
0

(thus the second formula follows)

NF(AND1(POS(NEG(x[11])), POS(NEG(x[12]))), I+J6);

is not 0 (thus the third formula does not follow).

The complete process of verification and extraction of knowledge takes, in this case, around ten seconds on a Pentium-based PC with 128 Mb of RAM.

5. CONCLUSION AND THANKS

The method described in the article illustrates the advisability of raising practical problems to the theoretical level (Formal Logic and Algebra in our case), at which means of simplifying reasoning and preparing efficient computer implementations are found. The method can be generalized to any type of information containing imprecise knowledge that could be translated into propositional many-valued and modal logics. So far, we have processed cases with as many as 200 propositional variables in three-valued logic using this method (the article only deals with 13).

We would like to thank P. Lázaro, K. Fitch, C. Benavente and A. Brunori for the information provided and their collaboration. We are especially grateful to Professor Sixto Ríos García for his invitation to submit this article to the «Revista de la Real Academia de Ciencias».

REFERENCES

1. Alonso, J.A.; Briales, E. & Riscos, A.: (1987) Preuve Automatique dans le Calcul Propositionnel et des Logiques Trivalentes. In Procs. Congress on Computational Geometry and Topology and Computation, Sevilla.
2. Capani, A. & Niesi, G. (1996) CoCoA User's Manual (v. 3.0b). Dept. of Mathematics, University of Genova.
3. Chazarain, J.; Riscos, A.; Alonso, J.A. & Briales, E. (1991) Multivalued Logic and Gröbner Bases with Applications to Modal Logic. Journal of Symbolic Computation 11; 181-194.
4. Field, M. & Lohr, M.K. (Eds.) (1992) Guidelines for Clinical Practice. From Development to Use. Washington D.C., National Academy Press.
5. Laita, L.M. & de Ledesma, L. (1997) Knowledge-Based Systems Verification. In Kent, A.; Williams, J.G. (Eds), Encyclopedia of Computer Science and Technology, pp. 253-280. New York, Marcel Dekker.
6. Laita, L.M.; Roanes Lozano, E. & Maojo, V. (1998) Inference and Verification in Medical Appropriateness Criteria. In: Calmet, J.; Plaza, J. (Eds.): Artificial Intelligence and Symbolic Computation, Procs. AISC'98 (LNAI-1476), pp. 183-194. Berlin, Springer-Verlag.
7. Roanes Lozano, E.; Laita, L.M. & Roanes Macías, E. (1995) Maple V in A.I.: The Boolean Algebra Associated to a KBS. Computer Algebra Nederland Nieuwsbrief. 14; 65-70.
8. Lázaro, P. & Fitch, K.: (1996) Criterios de uso apropiado para by-pass coronario. Unpublished report.
9. Roanes Lozano, E.; Laita, L.M. & Roanes Macías, E.: (1998) A Polynomial Model for Multivalued Logics with a Touch of Algebraic Geometry and Computer Algebra. Special issue «Non Standard Applications of CA». Mathematics and Computers in Simulation. 45 (1); 83-99.
10. Winkler, F. (1996) Polynomial Algorithms in Computer Algebra. Vienna, RISC-Linz Series, Springer-Verlag.