«LE MATEMATICHE»
Vol. XLIII (1988) - Fasc. I, pp. 125-194

A SURVEY OF COMPUTABLE SET THEORY

DOMENICO CANTONE (Catania, New York) (*)

This paper surveys various decidability results in set theory. In the first part, we focus on certain classes of unquantified set-theoretic formulae involving the relations $\in$ (membership), $=$ (equality), and the operators $\cap$ (intersection), $\cup$ (binary union), $\setminus$ (set difference), $\{\cdot\}$ (singleton), *pow* (powerset), *Un* (general union), η (choice operator), etc. The unquantified theory basic to the results presented is the so called Multilevel Syllogistic (MLS) which is the set of all quantifier-free formulae in the language $\in, =, \cap, \cup, \setminus$.

The second part of the paper covers the quantified case and, among others, we consider: a quantified version of MLS, the class of formulae quantified over sets in which the membership predicate is not allowed, the class of prenex set-theoretic formulae having three quantifiers and the class having a prefix of the form $\forall\forall\cdots\forall\exists$.

In the third part, some applications to domains different from pure set theory will be reviewed.

Some undecidability results will be discussed in the last section.

An appendix on Zermelo-Fraenkel set-theory concludes the paper.

(*) This work has been partially supported by ENI and ENIDATA within the AXL project.

1. Introduction.

In the last decade, the decision problem for various classes of set-theoretic formulae has been studied very intensively, as part of a joint project between New York University, the University of Catania, and Enidata-Bologna for the design and implementation of a set-theoretically based proof verifier (see [28]). The large body of results originated from such research has given rise to the novel field of *Computable Set Theory*, whose central problem is to find new decidable fragments of set theory.

In this paper, which has evolved from [7], we will survey the state-of-the-art of computable set theory (see [17] for a more extensive treatment).

The formalism of set theory is very well suited for an *all-purpose* theorem prover, mainly due to the great expressive power of the set-former construct

$$\{e(x_1, \ldots, x_n) : P(x_1, \ldots, x_n)\}$$

(where $e(x_1, \ldots, x_n)$ and $P(x_1, \ldots, x_n)$ stand respectively for a set-expression and a predicate) which allows an explicit instantiation of the objects crucial to predicate arguments. For instance, using this formalism it is possible, starting from the barest set-theoretic rudiments, to define cardinals, integers, rationals, reals, and complex numbers in entirely precise formal fashion, and to culminate in a full formal statement of, say, the Cauchy Integral Theorem, all within the space of no more than 150 lines (cf. [28]).

The size of the elementary deduction steps that a proof verifier system is able to carry out automatically (or at least to recognize as valid inferences in an interactive proof verification session) is directly proportional to the richness of its *inferential core*, and it is critical for any practical application. Indeed, if the size of the elementary steps is very small (as, for instance, in a resolution-based proof verifier) then any proof, even of very elementary theorems, becomes very long and requires entry of an overwhelming mass of tedious details.

The same observation applies even more forcefully to the much more pragmatic problem of program verification. Any hope of making such verification practical must presume that existing proof verifiers can be greatly improved, since in attempting to prove anything but very tiny programs correct one will convert them into substantial masses of mathematical statements, all of which must be verified formally if the correctness of the original program is to be established in any rigorously checked sense.

Computable set theory can serve to formalize the layer of details that are ordinarily missing in proofs, as found in the typical mathematical literature and therefore it could have important applications in the fields of program verification and automated deduction. For instance, the proof verifier envisaged in [28] will have among other fundamental components an *inferential core* comprised of an extensive (and extensible) collection of decision and semi-decision procedures for several fragments of set theory and other branches of mathematics too.

To be a little bit more formal, given a class of formulae $\mathcal{C}$ and a theory $\mathcal{T}$, a decision procedure for $(\mathcal{C}, \mathcal{T})$ is any algorithm which allows to mechanically establish for any given formula φ in $\mathcal{C}$ whether φ is in $\mathcal{T}$ or not, i.e., whether φ is valid in $\mathcal{T}$ or not. Dually, a decision procedure could be defined with respect to the satisfiability problem, and it is in this latter sense that the term decision procedure will be used in the present paper. For our purposes the theory $\mathcal{T}$ will be the Zermelo-Fraenkel theory of sets. In fact our considerations will take place in the von Neumann universe $\mathcal{V}$ of sets, but they can easily be formalized in ZF set theory.

Notice, in addition, that in most cases the decision procedures for the fragments of set theory considered here allow also to instantiate models. In other words, when a formula is declared satisfiable, the procedure actually returns a description of a collection of assignments satisfying it.

In the next section, we will consider some unquantified extensions of the basic theory MLS which includes the operators $\cup, \cap, \setminus$ and the relators = and $\in$. Some of the constructs considered are: *pow* (powerset), *Un* (unary union), $\bigcap$ (unary intersection), rank and cardinality related constructs, $\{\cdot\}$ (singleton), *Finite*, etc. No explicit quantification is allowed at this stage.

Languages admitting quantification will be reviewed in Section 3, where we will consider a class of formulae with restricted quantifiers and no alternations, as well as some classes of closed formulae whose quantifier prefixes are of a given type.

Section 4 will discuss some applications of the decision procedures presented in the first part of the paper to domains different from pure set theory.

Theoretical limitations to the problem of mechanizing set theory will be considered in Section 5.

Finally, an appendix on Zermelo-Fraenkel set-theory will conclude the paper.

2. Unquantified Theories.

It has been argued above that the richness of the inferential core of a proof verifier can allow to shorten drastically the length of the verification of proofs and/or the correctness of programs. Another motivation for undertaking the research of new decision algorithms is to prove a result analogous to the Herbrand's theorem of predicate calculus, but applying to set theory, which would be of great pragmatic importance (see [46, 56]).

THEOREM 2.1. (Herbrand) *There exists an automatic procedure P which, given an arbitrary formula of predicate calculus, produces an infinite sequence $g_1, g_2, g_3, \ldots$ of propositional formulae such that q is a theorem if and only if at least one of the g_i's is unsatisfiable. (For more details see, e.g., [1]).* ∎

It can be conjectured that the relationship of set theory to the class of unquantified formulae in the language including the operators $\cup$ (union), $\cap$ (intersection), $\setminus$ (set difference), $\{\cdot\}$ (singleton), *pow* (powerset), *Un* (general union), η (choice function), $\times$, (cartesian product) is analogous to the relationship of predicate calculus to propositional calculus. A preliminary step in finding such an analogue of the Herbrand theorem might therefore consist in proving the decidability of the theory sketched above. This section will review some partial results of this kind.

2.1. Multilevel Syllogistic.

In this subsection we illustrate a technique which, properly generalized, has been used to derive most of the decidability results in this field.

We begin by illustrating its application to a simple theory, the so-called Multilevel Syllogistic (MLS for brevity), and then we show how some extensions of MLS can be proved decidable much in the same way (cf. [36]; see also [55]).

The language MLS is considered first. (Note that here and in the following we will use the term «language» or «theory» in the broad sense of «set of formulae», not necessarily closed under logical deduction).

The language MLS is composed using

- variables: $x, y, z, \ldots$
- operators: $\cup, \cap, \setminus$
- relators: $\in, =$
- boolean connectives: $\&, \vee, \neg, \rightarrow, \leftrightarrow$.

The language MLS is then the set of unquantified formulae which can be built up from the above constituents by observing the usual syntactic rules:

$$(x \in (y \cup z) \,\&\, x \notin y) \rightarrow x \in z$$

is an example of a formula in the MLS language.

We will prove that MLS is decidable, i.e. that there exists an automatic procedure for deciding if any given formula in MLS is or is not satisfiable (in the standard von Neumann universe of set theory (see Appendix)). Such a procedure will be exhibited explicitly.

It is easily seen that we can limit ourselves to considering only those formulae Q of MLS which are conjunctions of literals of the following types:

$$(\in) \qquad x \in y$$

$$(=) \qquad x = y \cup z, \;\; x = y \setminus z$$

For instance, literals of type $x = y \cap z$ are equivalent to the atomic formula $x = y \setminus (y \setminus z)$, whereas the literal $x \notin y$ is equisatisfiable with the formula $x \in z \;\&\; z = z \setminus y$.

Before entering into the details of the proof, some terminology and definitions are needed. In order to make the discussion which follows clearer, it is convenient to assume that Q can be satisfied by an injective model M, i.e., a model which maps distinct variables into distinct sets. Let $V = \{y_1, y_2, \ldots, y_m\}$ be the set of all distinct variables occurring in Q. With Q we associate its *Venn diagram* $\mathcal{V}_Q^M$ with respect to the model M in the universe $U = \bigcup_{x \in V} (Mx \cup \{Mx\})$. Specifically, $\mathcal{V}_Q^M$ is the set of the equivalence classes with respect to the equivalence relation $\sim$ defined on U by:

$$s \sim t \text{ iff } s \in Mx \Leftrightarrow t \in Mx, \text{ for all variables } x \text{ occurring in } Q.$$

Let $\sigma_1, \sigma_2, \ldots, \sigma_n$ be the elements of $\mathcal{V}_Q^M$. Sets σ_i, for $i = 1, \ldots, n$, are called the *parts of the Venn diagram* $\mathcal{V}_Q^M$. Observe that every set σ_i is either fully contained in Mx or is disjoint from Mx, for every variable x occurring in Q. Therefore with each part σ_i of the Venn diagram $\mathcal{V}_Q^M$ of Q it is possible to associate a 0/1-valued function π_i

defined on V by putting

$$(1) \qquad \pi_i(x) = \begin{cases} 1 & \text{if } \sigma_i \subseteq Mx \\ 0 & \text{if } \sigma_i \cap Mx = \emptyset \end{cases}$$

The following two properties hold for each such function π_i:

(a) if the literal $x = y \cup z$ is in Q, then $\pi_i(x) = \pi_i(y) \vee \pi_i(z)$

(b) if the literal $x = y \backslash z$ is in Q, then $\pi_i(x) = \pi_i(y) \& \neg \pi_i(z)$.

DEFINITION 2.2. *Any 0/1-valued function which enjoys properties (a) and (b) above is called a place of Q.* ■

Since $\sigma_1 \cup \sigma_2 \cup \ldots \cup \sigma_n = U$, then for every variable x in Q there exists a unique part σ^x of the Venn diagram $\mathcal{V}_Q^M$ such that $Mx \in \sigma^x$. We call the corresponding place π^x the *place of Q at the variable x* It follows that

(c) $\pi^x(y) = 1$, if $x \in y$ occurs in Q.

Indeed, if $x \in y$ is in Q, then $Mx \in My$. But $Mx \in \sigma^x$, therefore by definition of the sets σ_i we have $\sigma^x \subseteq My$, and then $\pi^x(y) = 1$.

Note that sets σ_i and places π_i serve to describe the model M fully. Indeed, for every variable x in Q we have

$$Mx = \bigcup_{\pi_i(x)=1} \sigma_i.$$

These last observations constitute a first step towards reducing the decision problem for MLS to a *purely combinatorial* problem concerning the clauses of Q and the set of places π which actually appear in a model M of Q. We go on to show that it is possible to formulate decidable conditions concerning the set of places π and the clauses of Q which are both necessary and sufficient for Q to be satisfiable. Furthermore we will show that when these conditions are satisfied, canonical sets $\bar{\pi}$ can be associated to places π in such a way that the assignment $Mx = \bigcup_{\pi(x)=1} \bar{\pi}$ satisfies Q.

In order to find appropriate conditions on places and clauses of Q, it is convenient to keep in mind that places are the syntactic counterpart of the semantic notion of nonempty disjoint parts of a Venn diagram of Q. In particular, a place at a variable x (i.e., that place which satisfies properties (a), (b), and (c) above) is intended to represent that part of the Venn diagram which contains the «set» x.

Assume that a set $\Pi = \{\pi_1, \ldots, \pi_n\}$ of places of Q is given such that no two variables of Q are Π-equivalent, i.e. for all distinct x, y occurring in Q there is a place $\pi \in \Pi$ such that $\pi(x) \neq \pi(y)$. Suppose also that a place π^x at the variable x is associated with each variable x. Finally, let an ordering $<$ among variables occurring in Q be given (this ordering is supposed to correspond to the ideal rank ordering among the «set» x, for x in Q; for the definition of *rank* see Appendix).

A natural way to define an assignment M on the variables of Q is to associate pairwise disjoint nonempty sets $\bar{\pi}$ with places $\pi \in \Pi$. After such sets have been chosen, the definition $Mx = \bigcup_{\pi(x)=1} \bar{\pi}$, where x denotes any variable x in Q, plainly satisfies all literals of type (=) in Q. In order for M to satisfy the remaining literals in Q of type ($\in$) also, we only need to be sure that we have

$$Mx \in \bar{\pi}^x, \tag{2}$$

for every variable x in Q. Indeed, if $x \in y$ is a literal of Q, then by the mere definition of a place at the variable x (cf. (c) above), we have $\pi^x(y) = 1$. Therefore, by (2), $Mx \in \bar{\pi}^x \subseteq \bigcup_{\pi(y)=1} \bar{\pi} = My$, and $Mx \in My$.

This shows that formulation of conditions necessary and sufficient for construction of pairwise disjoint and nonempty sets $\bar{\pi}$ satisfying (2) solves the decidability problem for MLS.

The most natural way to associate a set $\bar{\pi}$ to a place π in such a way that the preceding conditions hold is as follows:

INSTANTIATION PROCEDURE

(A) FOR ALL π IN Π DO

$\bar{\pi} := \bar{\pi}^{(0)}$ (sets $\bar{\pi}^{(0)}$ should be chosen nonempty and paiwise disjoint, and in such a way that the subsequent $\in$-phase does not disrupt disjointness)

(B) FOR ALL VARIABLES x OCCURING IN Q, IN THE ORDER DEFINED BY $<$ DO

$$\bar{\pi}^x := \bar{\pi}^x \cup \left\{ \bigcup_{\pi(x)=1} \bar{\pi} \right\}$$

(C) FOR ALL VARIABLES x OCCURRING IN Q, PUT

$$Mx := \bigcup_{\pi(x)=1} \bar{\pi}.$$

We call loop (A) *initialization phase*, and loop (B) $\in$*-phase.*

For the moment let us focus our attention on the $\in$-phase (B) which is intended to force property (2) to hold. Once a variable x has been processed by performing the assignment

$$\bar{\pi}^x \to \bar{\pi}^x \cup \{Mx\},$$

to be sure that $\bigcup_{\pi(x)=1} \bar{\pi} = Mx \in \bar{\pi}^x$ holds at termination of the algorithm shown above, where $\bar{\pi}$ are the final values, it is enough to require that

(3) if $\pi^y(x) = 1$ then the variable y must precede x in the ordering $<$.

Indeed sets $\bar{\pi}^y$, with $y \geq x$, are the only ones which can change because of an $\in$-step during and after the processing of x (i.e. after the value $Mx = \bigcup_{\pi(x)=1} \bar{\pi}$ has been determined). Conversely, it is easy to see that property (3) is necessary for the satisfiability of Q; indeed, if the ordering $<$ is defined to be any linear ordering extending the partial ordering induced on variables by the rank ordering of the corresponding values Mx (where M is a given model of Q) then $x \leq y$ implies that *rank* $(Mx) \leq$*rank*(My), so that $My \in Mx$ is impossible. But since by definition $My \in \sigma^y$, we have by (1) $\pi^y(x) = 0$.

Referring to the initialization phase, it is easy to see that the choice $\{\{n+1, i\}\}$ for $\bar{\pi}_i^{(0)}$ (where $n = |\Pi|$) prevents disruption of disjointness during the $\in$-phase. Indeed, no Mx can ever equal a set $\{n+1, i\}$, since either *rank* $(Mx) = 0$ or *rank* $(Mx) > n+3$, whereas *rank* $(\{n+1, i\}) = n+2$, $i = 1, \ldots, n$. Furthermore, since these sets, called *individuals*, «mark» places to which they belong, in the sense that on termination of the instantiation algorithm we have

$$\{n+1, i\} \in Mx \text{ if and only if } \pi_i(x) = 1, \text{ for } i = 1, \ldots, n,$$

it follows easily that $Mx \neq My$, for all distinct variables x, y in Q.

This proves that condition (3) is both necessary and sufficient for satisfiability of a set of formulae of MLS.

Summing up, we have proved:

THEOREM 2.3. ([36]) *Let Q be a conjunction of literals of the types $(=, \in)$. Then Q is satisfiable if and only if there exist*

(i) a set Π of places of Q,

(ii) a correspondence $x \mapsto \pi^x$, where x ranges over the variable of Q, and π^x is a place at x,

(iii) an ordering $<$ of the variables in Q

such that

(a) no two distinct variables of Q are Π-equivalent,

(b) if $x \leq y$ then $\pi^y(x) = 0$. ■

Since places are 0/1-valued functions defined on the set $\{y_1, \ldots, y_m\}$ of the variables occurring in Q, it follows that the number of all possible places of Q is bounded by 2^m. Therefore satisfiability of the conditions appearing in the preceding theorem above is decidable, and we have

COROLLARY 2.4. *The theory MLS is decidable.* ■

Much the same technique described above allow to prove the decidability of the extension MLSS obtained from MLS when the *singleton* operator $\{s\}$ which maps a set s into the set whose only element is s is present.

By applying a straightforward normalization process as before, it is easily seen that we can limit ourselves to considering only those formulae Q of MLSS which are conjunctions of literals of the type $(=, \in)$ and literals of the type

$$(\{\cdot\}) \qquad x = \{y\}.$$

In order to satisfy a literal in Q of the type $x = \{y\}$ we must force $y \in x$, and also must ensure that the set x is given just one element. The first condition can be satisfied by adding the formula $y \in x$ to Q. To satisfy the second condition we need to impose the condition that there is exactly one place π for which $\pi(x) = 1$, and moreover that this place is precisely the place at y. It is clear that these conditions together with those listed in Theorem 2.3 are both necessary and sufficient for Q to be satisfiable. This simple argument gives:

THEOREM 2.5. ([36]) *The theory MLSS is decidable.* ■

We mention here that also the extension MLSΣ of MLS by one occurrence of the operator $\Sigma(x)$, where $\Sigma(x) = \{\{y\} : y \in x\}$, is decidable (cf. [15]). Note that MLSΣ properly extends MLSS. The proof of the decidability of MLSΣ involves considerations of various special orderings on places.

In the next sections we will consider other extensions of MLS in which similar considerations play an important role in proving the decidability.

The decision procedures given for MLS and MLSS are easily seen to be elementary recursive in the sense of Kalmar (cf. [47]), i.e. the length of the computation required to decide MLSS can be bounded by a quantity multiexponential in the length of the input formula. In fact, by refining the arguments above it can be shown that the satisfiability problem for MLSS in NP-complete (cf. [27]). Roughly speaking, the NP-completeness of the decidability problem for MLSS can be proved as follows. The NP-hardness is immediate, since the propositional calculus is linearly embeddable in MLSS. To prove the membership to the NP-class, one can show that every satisfiable formula Q of MLSS admits a model which can be described by a directed graph of size $O(|Q|^2)$, where $|Q|$ denotes the number of distinct variables in Q. Such a graph is called a *model graph of* Q.

More generally, it can be shown that all unquantified extensions of MLS to be discussed below are elementary recursive.

However, in Section 3 we will exhibit techniques for proving the decidability of particular quantified classes of formulae of set theory which do not imply at once the elementary recursiveness of the theories considered.

2.2. *Syllogistic schemes.*

In the previous section we have seen that a formula Q MLSS is satisfiable if and only if there exists a set O of finite and bounded objects (such as places, places at variables, etc.) satisfying a collection C of combinatorial conditions.

It has also been provided an instantiation procedure that given an MLSS formula Q and a set O (of certain objects) produces an assignment $M_O(\bar{\pi}_1^{(0)}, \ldots, \bar{\pi}_n^{(0)})$, depending on the initial choice for the sets $\bar{\pi}_1^{(0)}, \ldots, \bar{\pi}_n^{(0)}$ (we are assuming that $\Pi = \{\pi_1, \ldots, \pi_n\}$ is one of the objects in O). If O satisfies the conditions C, then there exists sets $\bar{\pi}_1^{(0)}, \ldots, \bar{\pi}_n^{(0)}$ (for instance, $\bar{\pi}_i^{(0)} = \{\{n+1, i\}\}$, $i = 1, \ldots, n$) such that $M_O(\bar{\pi}_1^{(0)}, \ldots, \bar{\pi}_n^{(0)})$ is a model for Q (observe that in general $M_O(\bar{\pi}_1^{(0)}, \ldots, \bar{\pi}_n^{(0)})$ needs not be a model for Q).

Given a satisfiable MLSS formula Q two questions arise at this point:

- Is any model of Q representable as $M_O(\bar{\pi}_1, \ldots, \bar{\pi}_n)$, for some sets $\bar{\pi}_1, \ldots, \bar{\pi}_n$?
- Are the classes of models relative to distinct sets of objects disjoint?

In [23], the set of objects O are formalized in the language

MLSSF (MLSS plus the predicate *Finite* (x)) in such a way that both questions have affirmative answer, at least in the case of the decision problem for MLSSF. The analogous problem for the extension of MLSSF with a choice function and the comparison predicate has been considered in [18, 19]; see also Section 2.9.

In [23], it is shown that for every set $X = \{x_1, \dots, x_m\}$ of variables, all possible assignments over the variables in X can be partitioned into finitely many classes, each of which characterized by a suitable formula called *syllogistic scheme*, such that every MLSSF formula involving only variables in X has the same truth value under all assignments in the same partition class.

To see how syllogistic schemes are obtained, consider m sets $a_1, \dots, a_m$. Put $i \sim j$ if $a_i = a_j$, for $i, j = 1, \dots, m$. Let $i_1, \dots, i_n$ be a set of $\sim$-representatives and for simplicity put $s_h = a_{i_h}$, $h = 1, \dots, n$. Then $s_i \neq s_j$, for every $i, j = 1, \dots, n$ such that $i \neq j$. Also put $i \mathcal{R} j$ whenever $s_i \in s_j$. Hence, each s_i can be rewrittens as

$$s_i = b_i \cup \{s_j : j \mathcal{R} i\}$$

for suitable sets b_i disjoint from $\{s_1, \dots, s_n\}$. Additionally, form all sets

$$c_H = \bigcap_{i \in H} b_i \setminus \bigcup_{j \notin H} b_j,$$

for all $\emptyset \neq H \subseteq \{1, \dots, n\}$, so that we have

$$s_i = \bigcup_{i \in H} c_H \cup \bigcup_{j \mathcal{R} i} \{s_j\}.$$

Finally, put $Z = \{H : c_H \neq \emptyset\}$ and $F = \{H \in Z : c_H \text{ is finite}\}$.

It turns out that the equivalence relation $\sim$, the acyclic relation $\mathcal{R}$, and the sets Z, F are sufficient to allow the evaluation of any MLSSF formula under any assignment which generates them.

More formally, let X be a set of variables, $\sim$ an equivalence relation over X whose equivalence classes are $\{x_{10}, x_{11}, \dots, x_{1L_1}\}, \dots, \{x_{n0}, x_{n1}, \dots, x_{nL_n}\}$. Put $s_i = x_{i0}$, for $i = 1, \dots, n$, and let $S = \{s_1, \dots, s_n\}$.

Additionally, let $\mathcal{R}$ be an acyclic relation over S. For each $\emptyset \neq H \subseteq \{1, \ldots, n\}$ introduce a new variable y_H and let Y be the collection of all these variables. Finally, let $F \subseteq Z \subseteq Y$ be such that for no distinct s_i, s_j in S

- $s \mathcal{R} s_i$ if and only if $s \mathcal{R} s_j$, for all s in S, and
- $i \in H$ if and only if $j \in H$, for all y_H in Z

can hold together.

Let $\sigma_{X \sim \mathcal{R} Y Z F}$ denote the formula

$$\&_{y_H \in Y} (y_H \neq \emptyset \;\&\; \&_{i \leq n} s_i \notin y_H \;\&\; \&_{y_{H'} \in Y \setminus \{y_H\}} y_H \cap y_{H'} = \emptyset)$$

$$\&\; \&_{i \leq n} \left(s_i = x_{i1} = \ldots = x_{iL_i} = \bigcup_{\substack{i \in H \\ y_H \in Z}} y_H \cup \bigcup_{j \mathcal{R} i} \{s_j\} \right)$$

$$\&\; \&_{y_H \in F} \mathit{Finite}\,(y_H) \;\&\; \&_{y_H \in Z \setminus F} \neg\, \mathit{Finite}\,(y_H).$$

$\sigma_{X \sim \mathcal{R} Y Z F}$ is called a *syllogistic scheme over* X *(relative to* $\sim, \mathcal{R}, Z, F$*).*

Let Σ_X denote the finite collection of all syllogistic schemes over X. Then the following properties hold.

THEOREM 2.6. ([23]).

(a) Irredundancy: Every syllogistic scheme in Σ_X *is satisfiable.*

(b) Mutual exclusion: For every two distinct syllogistic schemes $\sigma_1, \sigma_2 \in \Sigma_X$, $\sigma_1 \& \sigma_2$ *is unsatisfiable.*

(c) Exhaustivity: $(\forall_{x \in X} x)(\exists_{y \in Y} y) \left(\bigvee_{\sigma \in \Sigma_X} \sigma \right)$ *is valid.* ■

Moreover, if Q is an MLSSF conjunction with no compound terms, involving only variables in X, then

THEOREM 2.7. ([23])

(a) For every $\sigma \in \Sigma_X$*, either* $\sigma \rightarrow Q$ *or* $\sigma \rightarrow \neg Q$ *is valid. Moreover it is possible to establish algorithmically which of these two cases holds.*

(b) It is possible to determine algorithmically a subset Σ_Q of Σ_X such that the formula

$$(\forall_{x \in X} x) \left(Q \leftrightarrow (\exists_{y \in Y} y) \left(\bigvee_{\sigma \in \Sigma_Q} \sigma \right) \right)$$

is valid. ■

$\bigvee_{\sigma \in \Sigma_Q} \sigma$ is called a *syllogistic normal form of Q.*

Syllogistic normal forms can be easily generalized to any formula of MLSSF (not necessarily simple conjuctions). Notice that an MLSSF formula is satisfiable if and only if its syllogistic normal form is nonempty. However, the decision procedure suggested by this observation is utterly inefficient. In [27] it has been proved that in the absence of the singleton operator, every satisfiable formula of MLSF admits always a «minimum effort» syllogistic scheme, so that in trying to prove satisfiability of a given formula, one needs only to look for such minimal syllogistic schemes; see [27, 46, 38] for further details.

In the following sections we will review some other decidable extensions of MLS.

2.3. *MLS with cardinality constructs.*

Next we consider the extension of MLS with the cardinality operator $|\cdot|$. Such a theory involves two kinds of variables, namely set-variables and variables denoting cardinal numbers. Moreover the addition + and cardinal relations $<$ and $=$ are also allowed.

Much as before, it turns out that in order to get a decidability result for this theory, it is enough to consider conjunctions Q of literals of the form $(=, \in)$ togetehr with literal of the form

$$(|\cdot|) \qquad |x| = v,$$

$$(C) \qquad v_1 + v_2 = v_3, \;\; v_1 \leq v_2, \;\; v_1 = 0, \;\; v_1 = 1$$

where x is a set-variable and v, v_1, v_2, v_3 are cardinal variables.

Given a set $\Pi = \{\pi_1, \ldots, \pi_n\}$ of places of Q, we introduce a new cardinal variable v_π for each place π in Π. Then a system $C_{Q,\Pi}$ of cardinal linear equations and inequalities is obtained by adding to the cardinal literals in Q of type (C) the equations $\sum_{\pi(x)=1} v_\pi = v$, one for each literal in Q of the form $|x| = v$.

The instantiation algorithm described in the preceding section generalizes easily to MLS as extended to include a cardinal operator (at least for what concerns set variables). Indeed, once having instantiated all cardinal variables, all that is needed is to modify the initialization phase of the instantiation algorithm in such a way that the correct number of individuals goes to each set $\bar{\pi}$; this number is determined by the value of v_π and the number of variables x such that π is a place at x.

Thus Q is satisfiable if and only if the conditions of Theorem 2.3 can be satisfied with a set Π of places of Q such that the system of cardinal linear equations and inequalities $C_{Q,\Pi}$ has a cardinal solution. Since the solvability of such a system $C_{Q,\Pi}$ can be tested effectively (e.g. [53]), it follows that MLS extended by the above cardinality-related operations is decidable.

It is also known that purely additive integer arithmetic is decidable (cf. [50], [31]). Therefore much the same procedure implies the decidability of MLS theory extended by the cardinality-related operators that have been listed, but with the further assumption that cardinal variables can range only over nonnegative integers.

Observe that MLS with cardinals actually extends the simpler MLSS theory considered before. Indeed, the literal $x = \{y\}$ can be expressed by writing $y \in x \;\&\; |x| = 1$. Likewise, it can be expressed that a set is infinite. For instance, the formula $|y \cup \{y\}| = |y|$ is satisfiable if and only if the set y is infinite whereas $y = \emptyset \vee (x \in y \;\&\; |y \backslash \{x\}| < |y|)$ is satisfiable if and only if y is finite.

2.4. MLS with rank constructs.

In this subsection we will present some extensions of MLS with rank related constructs. We recall that the rank of a set is a measure of the depth of nesting of the constructor $\{\cdot\}$ within it. Thus, for instance, $rank(\emptyset) = 0$, $rank(\{\{\emptyset\}, \emptyset\}) = 2$, etc. (see Appendix). Intuitively, if $rank(s) < rank(t)$, then the set s must have been constructed before t in the universe of all sets. The relation $rank(s) < rank(t)$ will be abbreviated with $s < t$ and $<$ is called *rank comparison*. We will survey in some detail the extension of MLS with $<$ (see [15]).

To take into account the new relation $<$, the set of places $\Pi = \{\pi_1, \ldots, \pi_n\}$ relative to a given formula Q of MLS plus $<$ is given an ordering relation (which, for simplicity, coincides with the indexing order). The intuition behing $\pi_i < \pi_j$ (i.e., $i < j$), with $1 \leq i$, $j \leq n$, is that the set $\bar{\pi}_i$ (to be associated with the place π_i) must have rank not greater than the rank of the set $\bar{\pi}_j$.

In addition, it is necessary to locate the places in which there is a rank jump. This is given by an increasing sequence of integers

$$0 = r_0 < r_1 < \ldots < r_e = n,$$

with the understanding that

$$r_{a-1} < j_1, j_2 \leq r_a \Leftrightarrow rank(\bar{\pi}_{a-1}) < rank(\bar{\pi}_a).$$

It is also useful to introduce the map $R : \{1, \ldots, n\} \rightarrow \{1, \ldots, e\}$ such that

$$R(j) = a \text{ whenever } r_{a-1} < j \leq r_a.$$

Then it can be shown the following theorem.

THEOREM 2.8. ([15]) *A formula Q of MLS plus, $<, \leq$ is injectively satisfiable if and only if there exist:*

(i) a set $\Pi = \{\pi_1, \ldots, \pi_n\}$ of places of Q;

(ii) a mapping $x \mapsto \pi^x$ from the variables of Q into Π;

(iii) *a sequence of integers* $0 = r_0 < r_1 < \ldots < r_e = n$ *and a nondecreasing surjective map* $R : \{1, \ldots, n\} \rightarrow \{0, 1, \ldots, e\}$

such that:

(a) *no two variables in* Q *are* Π*-equivalent;*

(b) π^x *is a place at* x*, for all* x *in* Q*;*

(c) $r_{R(j)-1} < j \leq r_{R(j)}$*, for all* $j = 1, \ldots, n$*;*

(d) *if* $\pi_j(y) = 1$ *and* $\pi^y = \pi_l$*, then* $R(j) < R(l)$*;*

(e) *if* $y_{i_1} \leq y_{i_2}$ *[resp.* $y_{i_1} < y_{i_2}$*] is in* Q*, then*

$$\max\{R(t) : \pi_t(y_{i_1}) = 1\} \leq \max\{R(t) : \pi_t(y_{i_2}) = 1\}$$

$$[\textit{resp.} \max\{R(t) : \pi_t(y_{i_1}) = 1\} < \max\{R(t) : \pi_t(y_{i_2}) = 1\}]. \qquad \blacksquare$$

The decidability of MLS plus $<$ is an immediate consequence of the fact that the conditions of Theorem 2.8 are effectively verifiable.

The theory MLS plus $<$ has been variously generalized. In [15] a decision procedure is given for the extension of MLS in which terms of type $rank(x)$ can be freely combined together with the other set theoretic constructs. Since

$$s < t \quad \text{iff} \quad rank(s) \in rank(t),$$

MLS plus $<$ can be embedded in MLS plus *rank*; see also [13] for a quantified theory involving the operator $rank(x)$.

Also related to the notion of rank are the operators $pred_<$ and $pred_\leq$ defined respectively by

$$pred_<(s) = \{t : rank(t) < rank(s)\}$$

$$pred_\leq(s) = \{t : rank(t) \leq rank(s)\}$$

(see [54]).

[13] proves the decidability of the theory MLS plus $pred_<$. Since $s < t$ if and only if $s \in pred_<(t)$, this theory embodies MLS plus $<$. In

the same paper it is also proved the decidability of the theory MLS plus $<$ and singleton. Both such results are generalized and combined in [11] where it is shown that MLS with singleton, $pred_<$, $pred_\leq$, and the predicate *Finite* is decidable. It is also argued that decidability is not disrupted even if the additional predicate which says that a set is *hereditarily finite* (i.e., it has rank less than $\omega = \{0, 12, \ldots\}$) is allowed.

Notice that the presence of the singleton operator complicates matters quite a bit. Indeed, starting from the empty set (which is expressible in MLS by the formula $y_0 = y_0 \setminus y_0$) the singleton operator allows to build in a finite number of steps sets t of a given fixed rank. This implies that every set s whose rank is bounded in the formula by any set t of fixed rank can vary only in a finite collection of sets, a *priori* determinable.

The above discussion is formalized by the notion of *trapped place*. For instance, in the case of MLS plus $<, \leq, \{\cdot\}$ the following notion of *admissible set of trapped places* is introduced.

DEFINITION 2.9. *An* admissible set of trapped places *is any subset $\mathcal{T}$ of Π such that for all places π in Π and for all variables y, y' in Q, if*

- $\pi(y) = 1$,
- $\{\pi_j : \pi_j(y') = 1\} \subseteq \mathcal{T}$,
- *either $y = \{y'\}$ is in Q, or $y \leq y'$ is in Q, or $y < y'$ is in Q*

then $\pi \in \mathcal{T}$.

Given an admissible set of trapped places $\mathcal{T}$, a variable y occurring in Q is said to be trapped *(with respect to $\mathcal{T}$) if $\{\pi_j : \pi_j(y) = 1\} \subseteq \mathcal{T}$.*

Then Theorem 2.8 takes the form

THEOREM 2.10. ([13]) *Let Q be a normalized conjunction of MLSS plus $<, \leq$ containing the literal $y_0 = y_0 \setminus y_0$. Let $V = \{y_0, \ldots, y_m\}$ be the set of variables occurring in Q. Then Q is injectively satisfiable if and* *only if there exist:*

(i) *a set* $\Pi = \{\pi_0, \dots, \pi_n\}$ *of places of* Q *such that* π_0 *is a place at* y_0*;*

(ii) *a positive integer* $k \leq n$ *such that* $\mathcal{T} = \{\pi_0, \dots, \pi_k\}$ *is a set of admissible trapped places, and for which there exists a positive integer* $h \leq m$ *such that exactly the variables* $y_0, \dots, y_h$ *are trapped;*

(iii) *nonempty pairwise disjoint hereditarily finite sets* $\overline{\pi_j}$*,* $0 \leq j \leq k$*, of rank lower than* $m+1$ *such that the assignment* $\bar{M}y_i = \bigcup_{\pi_j(y_i)=1} \overline{\pi_j}$ *is an injective model for the subset of* Q *involving only trapped variables;*

(iv) *a mapping* $x \mapsto \pi^x$ *from* V *into* Π *(for simplicity, we define also a function* F*:* $\{0, \dots, m\} \to \{0, \dots, n\}$ *such that* $F(i) = j$ *if* $\pi^{y_i} = \pi_j$*);*

(v) *a sequence of integers:* $r_0 = k < r_1 < \dots < r_e = n$ *and a nondecreasing surjective function* $R : \{k+1, \dots, n\} \to \{1, \dots, e\}$

such that:

(a) *no two variables in* P *are* Π*-equivalent;*

(b) $\pi^{y_i} (= \pi_{F(i)})$ *is a place at* y_i*, for all* $0 \leq i \leq m$*;*

(c) *if* y_i *and* π_j *are trapped and* $\bar{M}y_i \in \overline{\pi_j}$ *then* $\pi^{y_i} = \pi_j$*;*

(d) *if* $j > k$ *(i.e. if* π_j *is nontrapped) then* $r_{R(j)-1} < j \leq r_{R(j)}$*;*

(e) *if* $\pi_j(y_i) = 1$ *then* $R(j) < R(F(i))$*.*
For all $i \in \{h+1, \dots, m\}$ *we put*

$$i^* = \max\{R(t) : \pi_t(y_i) = 1\}.$$

Then we have

(f) *if* $y_{i_1} \leq y_{i_2}$ *is in* Q *and* y_{i_1} *is nontrapped, then* $i_1^* \leq i_2^*$*;*

(g) *if* $y_{i_1} < y_{i_2}$ *is in* Q *and* y_{i_1} *is nontrapped, then* $i_1^* < i_2^*$*;*

(h) *if* $y_{i_1} = \{y_{i_2}\}$ *is in* Q *and* y_{i_2} *is nontrapped, then*

(h_1) $\pi^{y_{i_2}}(y_{i_1}) = 1$*;*

(h_2) if $\pi_j \neq \pi^{y_{i_2}}$*, then* $\pi_j(y_{i_1}) = 0$*,* $j \in \{0, \dots, n\}$*;*

(h_3) if $F(i) = F(i_2)$*, then* $i = i_2$*, for all* $i \in \{0, \dots, m\}$ *(i.e.,* $\pi^{y_{i_2}}$ *is a place only at the variable* y_{i_2}*);*

(h_4) $R(F(i_2)) = i_2^* + 1$. ■

Notice that sets $\bar{\pi}_t$ relative to trapped places π_t can not contain «individuals», but have to range over a finite collection of sets of «low» rank. This makes the search for a model very inefficient since one can not define all sets $\bar{\pi}$ in a uniform manner but has to exhaustively search the space of all possibilities.

We conclude this section by mentioning that the extension of MLS with singleton and both rank and cardinality comparison has also been proved decidable in [10]. The decision method combines the techniques outlined in Section 2.3 for cardinality constructs with those described in the present section.

2.5. MLS with the general union operator.

In this section we show how the ideas introduced in Section 1.1 can be be generalized in such a way as to establish the decidability of the unquantified theory MLSU obtained from MLS by also allowing an unrestricted number of occurrences of the general union operator Un, where $Un(s)$ stands for the union of all elements of s (see Appendix). As usual, we can limit ourselves to considering only formulae Q of MLSU which are conjunctions of literals having one of the types $(\in, =)$ (as in MLS) and literals of the type

$$(Un) \qquad u = Un(y).$$

Assume that a set $\Pi = \{\pi_1, \dots, \pi_n\}$ of places of Q is given, and also suppose that an ordering $<$ among variables of Q is fixed. The heuristic technique used in order to find conditions which are both necessary and sufficient for Q to be injectively satisfiable resembles that used in the MLS case. That is, we use an instantiation algorithm of the type described in Section 1.1 with an intialization phase and an $\in$-phase, and we look for conditions which guarantee the correct output from such an algorithm.

A first condition is suggested by the following observation. Assume that $\pi^x(y) = 1$ for variables x and y for which there is a clause $u = Un(y)$ in Q. Then since $Mx \in \bar{\pi}^x$, it follows, by $\pi^x(y) = 1$ above, that when the instantiation algorithm terminates, we must have $Mx \in My$. In order for $Mu = Un(My)$ to hold we must therefore have $Mx \subseteq Mu$, which is ensured as long as for every place $\pi \in \Pi$ such that $\pi(x) = 1$, we have $\pi(u) = 1$. Thus a first obviously necessary condition is:

(*) if for a variable x in Q we have $\pi^x(y) = 1$, with $u = Un(y)$ a clause in Q, then for every place $\pi \in \Pi$, $\pi(x) = 1$ implies $\pi(u) = 1$.

The remaining conditions necessary for satisfiability can be derived by closer analysis of the initialization- and $\in$-phases. A first problem si how to find values $\bar{\pi}^{(0)}$ such that the «temporary» assignment $Mx = \bigcup_{\pi(x)=1} \pi^{(0)}$ satisfies all conjuncts in Q which do not involve the membership relation $\in$. For this, we begin with a set of places in which individuals can be put without exercising any special care (except that certain rank restrictions must be satisfied). Then we proliferate individuals in such a way as to initialize the sets $\bar{\pi}$ appropriately. It turns out that all inclusion relationships $Un\left(\bigcup_{\pi(y)=1} \bar{\pi}\right) \subseteq \bigcup_{\pi(u)=1} \bar{\pi}$, where $u = Un(y)$ occurs in Q, can be forced by the first part of the initialization phase (*proper initialization*). A subsequent *stabilization subphase* then turns these inclusions into equalities. During the $\in$-phase, every insertion of a set Mx into $\bar{\pi}^x$ can disrupt inclusion relationships of the type $\bigcup_{\pi(u)=1} \bar{\pi} \subseteq. \left(\bigcup_{\pi(y)=1} \bar{\pi}\right)$, where $u = Un(y)$ occurs in Q (the reverses of these inclusions are maintained because of property (*)). To cure this problem, it is only necessary for each substep of the $\in$-phase to be followed by a stabilization subphase.

For the various stabilization phases that we have just sketched to function properly some relation among places is required. This suggests the following definition which associates a graph with the conjunction Q and its set Π of places.

DEFINITION 2.11. *Given a conjunction Q and a set Π of places of Q as above, the Ugraph G of Q, Π is the graph whose set of nodes is Π, plus one additional node Ω, and whose edges are as follows:*

(i) a directed edge connects π to Ω if and only if $\pi(y) = 0$ for every variable y for which $u = Un(y)$ is in Q (intuitively this means that clauses $u = Un(y)$ of Q tell us nothing about the set $Un(\bar{\pi})$, which allows the proper initialization phase to start with such places);

(ii) otherwise, a directed edge connects the place α to the place β if and only if $\beta(u) = 1$ for all clauses $u = Un(y)$ such that $\alpha(y) = 1$ (intuitively, the nodes β such that $\alpha \rightarrow \beta$ is an edge of G represent all the sets $\bar{\beta}$ in which elements of $Un(\bar{\pi})$ can appear. Indeed, let $y_{i_1}, \ldots, y_{i_k}$ be all variables y such that $u = Un(y)$ is in Q and $\alpha(y) = 1$, so that $\bar{\alpha} \subseteq My_{i_1} \cup \ldots My_{i_k}$. It follows that $Un(\bar{\alpha}) \subseteq Mu_{i_1} \cup \ldots \cup Mu_{i_k}$, i.e. $Un(\bar{\alpha})$ is contained in the union of all sets $\bar{\beta}$ such that $\alpha \rightarrow \beta$ is an edge of G). ■

In the graph just introduced we can distinguish three kinds of nodes. Those from which there is a directed path which reaches Ω are called *safe*. A node is called *trapped* if every sufficiently long path from it eventually reaches a node from which no edge branches off (*null* node). Finally a node is *cyclic* if it is neither safe nor trapped. Intuitively, trapped places are those places whose elements are subject to severe restrictions. Consider for example the following formula:

$$Un(x) = \emptyset \;\&\; Un(y) = x \;\&\; Un(z) = y \;\&\; x = \emptyset.$$

It is easy to see that in any set of places of the formula above, all places are trapped. It is also evident that variables x, y, z can assume just a few values; this «semantic» fact reflects the «syntactic» fact that we have justed noted. Trapped places are dealt with by observing that such places can be assigned only sets having rank at most one more than the maximum length of a longest path forward from each of them to a null node. Therefore only a finite number of possible choices must be checked in order to determine the value $\bar{\pi}$ to associate with such a π. On the other hand it turns out that each nontrapped place π can be assigned an infinite $\bar{\pi}$. This fact simplifies a lot the initialization phase, since the «individuals» which we place into $\bar{\pi}$ initially easily propagate along the Ugraph via singletons or pairs.

A rough description of the first initialization phase is as follows (for simplicity we only consider the case in which no trapped place exists). First of all infinitely many individuals are associated with every place π of Q such that $\pi \to \Omega$ is an edge of the graph G. Then any safe place can iteratively be given an infinite supply of elements by drawing elements from its descendants and forming their singletons. The same technique can also be used to initialize cyclic places, once we observe that the null node $\pi_\emptyset$ must lie on a cycle which can be given elements by successive formation of singletons of the empty set $\emptyset$ (which is assigned to $\bar{\pi}_\emptyset$) and that the null node must be reachable along edges of G from every other node (by the regularity axiom of set theory; see Appendix). This observation, which in substance gives us a second condition for the satisfiability of Q, guarantees that proper initialization can be accomplished successfully. Once this phase is completed, all literals of type (=) are correctly modeled; however for literals $u = Un(y)$ in Q all we can say is that

$$Un\left(\bigcup_{\pi(y)=1} \bar{\pi}\right) \subseteq \bigcup_{\pi(u)=1} \bar{\pi}.$$

To get equalities in place of these inclusions, the following stabilization phase is then performed. For each element p which has been put into $\bar{\pi}$ and for every clause $u = Un(y)$ such that $\pi(u) = 1$ (i.e., intuitively, $\bar{\pi} \subseteq Mu$), an element A is found such that after inserting A into $\bar{\pi}$ no inclusion of the type above is disrupted. Then the pair $\{p, A\}$ is inserted in a place β such that $\beta \to \pi$ is an edge of the Ugraph G. We refrain from stating the conditions which guarantee that such a stabilization phase can actually take place, since they are quite involved. The interested reader can find a complete description of them in [22].

As observed earlier, a stabilization process is also needed in the $\in$-phase. Overall, the instantiation algorithm for MLSU therefore has the form

$$\text{Initialization} \quad \begin{cases} \text{For every place } \pi \text{ let} \\ \bar{\pi} := \bar{\pi}^{(0)} \\ \text{Stabilize} \end{cases}$$

$$\in -\text{phase} \quad \begin{cases} \text{Following the order} < \text{of variables } x \text{ in } Q \text{ do} \\ \bar{\pi}^x := \bar{\pi}^x \cup \{Mx\} \left(\text{where } Mx = \bigcup_{\pi(x)=1} \bar{\pi} \right) \\ \text{Stabilize} \end{cases}$$

2.6. MLS with the powerset operator.

The decidability problem for MLS extended by the powerset operator *pow* (where $pow(s) = \{t | t \subseteq s\}$) can be easily reduced to the satisfiability problem for conjunctions Q of clauses in MLS and clauses of the type

$$(pow) \qquad p = pow(q).$$

In the preceding section it was convenient to introduce a graph structure among places of Q. In the present case we will see that what is needed instead is a relation between sets of places and places.

We begin with some general considerations on the powerset operator. Let $s_1, s_2, \ldots, s_n$ be nonempty disjoint sets. Then we have

$$pow(s_1 \cup s_2 \cup \ldots \cup s_n) = \bigcup_{A \subseteq \{s_1, s_2, \ldots, s_n\}} pow^*(A),$$

where $pow^*(A)$ stands for the set of those subsets of $\bigcup_{s \in A} s$ which have nonempty intersection with every element of A. The validity of the formula above can be easily verified by observing that every element on the right-hand side of the equality above is a subset of $s_1 \cup s_2 \cup s_2 \cup \ldots \cup s_n$. On the other hand, if t is an element of $pow(s_1 \cup s_2 \cup \ldots \cup s_n)$, then $t \in pow^*(A_t)$, where $A_t = \{s_i | s_i \cap t \neq \emptyset, i = 1, \ldots, n\}$. Hence, if $p = pow(q)$ is a powerset clause in Q, and $\alpha_1, \ldots, \alpha_l$ are places of Q such that $\alpha_1(q) = \ldots = \alpha_l(q) = 1$, there must exist places $\beta_1, \ldots, \beta_k$ such that $\beta_1(p) = \ldots = \beta_k(p) = 1$, and such that elements of $pow^*(\bar{\alpha}_1, \ldots, \bar{\alpha}_l)$ can lie in $\bar{\beta}_1 \cup \ldots \cup \bar{\beta}_k$ only. We indicate this relationship by writing $\{\alpha_1, \ldots, \alpha_l\} \to \beta_i$, for each $i = 1, \ldots, k$. (The symbolism used suggests the idea of a «flow» from the places $\alpha_1, \ldots, \alpha_l$ to the place β_i). In order to be more precise, we give the following exact definition.

DEFINITION 2.12. *A nonempty set* $\{\alpha_1, \ldots, \alpha_l\}$ *of places of* Q *is called a* Q*-node if there is a powerset clause* $p = pow(q)$ *in* Q *such that* $\alpha_j(q) = 1$, *for all* $j = 1, \ldots, l$.

If A *is a* Q*-node, then a place* β *is called a target of* A *if for every powerset clause* $p = pow(q)$ *we have* $\beta(p) = 1$ *if and only if* $\alpha(q) = 1$ *for all* $\alpha \in A$.

A place β *is called initial if it is not the target of any* Q*-node* A. *(Intuitively, initial places are those places which are not constrained by powerset clauses. It is then reasonable to start initialization from these places.)* ■

A first condition for Q to be satisfiable follows immediately from the following consideration. If $s = pow(t)$, then $u \in s$ if and only if $u \subseteq t$. In terms of places, this translates as follows:

«if $p = pow(q)$ is a clause in Q, then $\pi^x(p) = 1$ if and only if for every place π such that $\pi(x) = 1$, we have $\pi(q) = 1$.»

This condition assures that during the $\in$-phase, insertion of Mx in $\bar{\pi}^x$ will not disrupt any inclusion of the type $\bigcup_{\pi(p)=1} \bar{\pi} \subseteq pow\left(\bigcup_{\pi(q)=1} \bar{\pi}\right)$, for any powerset clause $p = pow(q)$ in Q.

In order to force equalities in place of the above inclusions, a stabilization phase is needed each time a new variable x of Q is processed during the $\in$-phase. Such a stabilization will proceed in a manner defined by certain special edges of the type $\{\alpha_1, \ldots, \alpha_l\} \rightarrow \beta$. (These special edges are those edges whose target is a «set» of maximum rank; this idea guarantees against circularity).

In this case stabilization steps are easy to describe; they just consist of assignments of the form

$$\bar{\beta} := \text{pow}^*(\bar{\alpha}_1, \ldots, \bar{\alpha}_l) \setminus \bigcup_{\{\alpha_1, \ldots, \alpha_l\} \rightarrow \gamma} \bar{\gamma},$$

where $\{\alpha_1, \ldots, \alpha_l\} \rightarrow \beta$ is a special edge.

The initialization phase can be described roughly as follows. Since initial places are not restricted by any powerset clause, we can initialize them freely using a sufficiently large number of individuals. Moreover the empty set can be assigned to the place $\pi_\emptyset$ (for simplicity, it is convenient to assume that $\emptyset$ is a variable occurring in Q which stands for the empty set). At this point proliferation of elements can start. This will continue until each place has been assigned at least one element. More specifically, for each Q-node $\{\alpha_1, \ldots, \alpha_l\}$, with $\bar{\alpha}_1, \ldots, \bar{\alpha}_l$ nonempty, elements in $pow^*(\bar{\alpha}_1, \ldots, \bar{\alpha}_l) \setminus \bigcup_{\{\alpha_1, \ldots, \alpha_l\} \rightarrow \gamma} \gamma$ are opportunely distributed among all its targets γ. [21] states conditions which ensure that the initialization and subsequent stabilization phases can execute properly.

Note, finally, that it can be proved that if m is the number of different variables occurring in Q, then Q is satisfiable if and only if Q has a model of rank at most $2^{2^{m+1}+m+2} + 1$.

It is interesting to contrast this last result with the fact that there are formulae of MLS theory extended by the general union operator which admit only infinite models. For example, the formula $x \neq \emptyset \;\&\; x \subseteq Un(x)$ is not finitely satisfiable, even though the assignment $Mx = \{\emptyset_0, \emptyset_1, \ldots\}$, where $\emptyset_0 = \emptyset$ and $\emptyset_{n+1} = \{\emptyset_n\}$, for $n = 0, 1, \ldots$, clearly satisfiest it.

We close this section by noticing that the extension with also the singleton operator has been shown decidable in [6,8]. Roughly speaking, it is shown that a formula Q of MLS plus powerset and singleton is injectively satisfiable if and only if a certain nondeterministic association procedure can produce a canonical model of Q in time bounded by a doubly exponential expression in the number m of variables occurring in Q.

To show the necessity of such condition, an existing model M^* of Q is used as an oracle to instantiate a computation of the association procedure. Again, matters are complicated by the presence of singleton clauses, which cause some of the places to be considered as trapped. Trapped places are handled by maintaining a one-one partial map from the elements of the canonical model under construction into the elements of the oracle model. Such a map is intended to guide a correct instantiation of the association algorithm.

At each step of a computation of the association algorithm, the rank of any set can increase at most by one. Therefore, it follows as a by-product that any formula Q of MLS plus powerset and singleton is injectively satisfiable if and only if it has a model of rank doubly exponential in the number of variables in Q.

2.7. MLS with map constructs.

Next we consider two approaches to extend MLS with map constructs. The first, adopted in [37,4], consists in introducing a new sort of variables denoting (multivalued) maps, whereas in the second one, used in [29], a unique sort of variables is employed, i.e. everything is a set (recall that maps can be represented as sets of ordered pairs).

The language considered in [37] admits in addition to the usual MLS constructs the following atomic formulae

$$\begin{array}{lll} x_{f_y} = f[y], & x_{Df} = Df, & x_{Rf} = Rf, \\ y_{f_x^{-1}} = f^{-1}[x], & \text{SINGLEVALUED}(f), & \text{ONE-ONE}(f) \end{array}$$

where x_{f_y}, x_{Df}, x_{R_f}, $y_{f_x^{-1}}$ stand for set variables, f denotes a map variable, D and R are respectively the domain and the range operators, and $f[y]$, $f^{-1}[x]$ denote respectively the direct and inverse image.

Notice that we have

$$y = f^{-1}[x] \leftrightarrow y \subseteq Df \;\&\; f[y] = x \cap Rf \;\&\; f[Df \backslash y] \cap x = \emptyset.$$

Thus literals of type $y = f^{-1}[x]$ can be easily eliminated. To eliminate also literals of type $x_{f_y} = f[y]$, $x_{Df} = Df$, $x_{Rf} = Rf$ from a given formula Q one can proceed by forming all sets of type $y_1^{i_1} \cap y_2^{i_2} \cap \ldots \cap y_k^{i_k}$, for each map variable f, with $i_j \in \{0, 1\}$, where $y_1, y_2, \ldots, y_k$ are all the variables appearing on the right-hand side of literals of the form $x = f[y]$ in Q, and where $y_l^{i_l}$ stands for either $y_l \cap x_{Df}$ if $i_l = 0$, or for $x_{Df} \backslash y_l$ if $i_l = 1$. Also, for each f one introduces the new variables $s_{i_1 i_2 \ldots i_k}$ denoting the term $f[y_1^{i_1} \cap y_2^{i_2} \cap \ldots \cap y_k^{i_k}]$ and adds the following formulae to Q:

$$y_1^{i_1} \cap y_2^{i_2} \cap \ldots \cap y_k^{i_k} = \emptyset \leftrightarrow s_{i_1 i_2 \ldots i_k} = \emptyset$$

$$x_{f_{y_i}} = \bigcup_{i_j = 1} s_{i_1 i_2 \ldots i_k}$$

$$x_{Rf} = \bigcup_{i_1, \ldots, i_k} s_{i_1 i_2 \ldots i_k}$$

and deletes all literals of type $x_{f_y} = f[y]$, $x_{Df} = Df$, $x_{Rf} = Rf$. If $\bar{Q}$ is the resulting formula, then one can prove the following theorem

THEOREM 2.13. ([37]) *Q is satisfiable if and only if $\bar{Q}$ satisfiable.* ■

It is also easy to eliminate from Q literals of the form SINGLEVALUED (f) and ONE-ONE (f) thus reducing the satisfiability problem for MLS with map constructs to that for MLS. Cardinality constructs of the type considered in 2.3 can be also combined with the two-sorted language above still maintaining decidability (cf. [37]).

In [29], the map related notions such as domain, inverse, singlevalued, etc. are permitted to apply to any variable. Specifically, the theory considered is the propositional closure of the set-theoretic clauses of the forms

$$d = Df,\ \text{PAIR-IN}(x, y, f),\ \text{INV}(f, g),\ \text{SINGLEVALUED}(f)$$

in addition to MLS clauses. The intended meaning of the operator D and predicates PAIR-IN, INV, SINGLEVALUED is as follows. For every set f, let pairs(f) be the collection of all ordered pairs $[p, q]$ present in f. Then Df denotes the set $\{x : [x, y] \in \text{pairs}(f)$, for some $y\}$; the predicate PAIR-IN(x, y, f) is true if and only if $[x, y] \in$ pairs(f); INV(f, g) is true if and only if $\{[p, q] : [q, p] \in \text{pairs}\ (f)\} =$ pairs(g); the predicate SINGLEVALUED(f) is true if and only if for all $[p, q]$, $[p', q']$ in pairs(f) if $p = p'$ then $q = q'$. Observe that according to such interpretations, a set f admits a whole class of inverses g, i.e., $(\text{pairs}(f))^{-1} \cup S$, for every set S such that pairs $(S) = \emptyset$.

The set representation of an ordered pair is not central for the decidability result in question. The one adopted in [29] is

$$[p, q] = \{\{0, \{p\}\}, \{2, \{q\}\}\},$$

where the integers 0 and 2 stand for the set-theoretic representation of the ordinals 0 and 2, i.e. $\emptyset$ and $\{\emptyset, \{\emptyset\}\}$ respectively.

Notice that in the two-sorted language of [37], constructs like $f_1 = f_2 \cup f_3$, $f_1 = f_2 \setminus f_3$, $x \in f_1$, $f_1 \in f_2$, PAIR-IN (x, y, f), etc., where f_1, f_2, f_3 are function variables and x, y are set variables, are forbidden, whereas they are allowed in [29].

Other constructs expressible in the language are the following:

- $r = Rf$, where $Rf = \{q : [p, q] \in \text{pairs}(f)$, for some $p\}$;
- RESTR(g, f, x), which denotes $(\forall p)(\forall q)([p, q] \in g \leftrightarrow ([p, q] \in f \wedge p \in x))$;
- $y = f[x]$, $y = f^{-1}[x]$;
- INJECTIVE(f), which denotes $(\forall p)(\forall p')(\forall q)([p, q] \in f \wedge [p', q] \in f \rightarrow p = p')$.

Notice that the term pairs(x), which allows to characterize maps in their standard representation (i.e. as sets of ordered pairs), is not expressible in the language under discussion. The main difficulty in permitting also occurrences of such terms is that their presence would allow to express the singleton operator too, and therefore the typical pathology of trapped places (see Section 2.4) would have to be taken into account. At this stage, it is not clear how to combine the techniques developed for trapped places with the quite intricate construction described in [29].

In [29] it is proved that a formula Q of MLS plus map constructs is injectively satisfiable if and only if there exist some maps and relations, whose domains and ranges are finite and effectively determinable from Q, satisfying some eleven conditions which for simplicity are omitted here. In addition it is shown that when such conditions are satisfied, then it is possible to describe a *construction process* which build a *skeletal* model of Q (or a *model graph*; cf. [49]).

A skeletal model of Q is a well-founded, quasi-extensional directed graph G associated with Q admitting a *representation* which in turn determines an injective model for Q. More precisely.

DEFINITION 2.14. *A directed graph $G = (N, E)$ with nodes N and edges E is said to be well-founded if it has no infinite descending chain.*

A directed graph $G = (N, E)$ is said to be quasi-extensional if for all $v_1, v_2 \in N$,

$$\{u \in N : u \Rightarrow v_1 \text{ is in } E\} = \{u \in N : u \Rightarrow v_2 \text{ is in } E\} \neq \emptyset$$

implies $v_1 = v_2$. ■

DEFINITION 2.15. *A function R defined on the set N of nodes of a well-founded graph $G = (N, E)$ and with values on a class of sets is called a representation of G if for all $v_1, v_2 \in N$*

(i) $R(v_1) = R(v_2)$ implies $v_1 = v_2$,

(ii) $R(v_1) \in R(v_2)$ if and only if $v_1 \Rightarrow v_2$ is in E. ■

Typically, a representation of a well-founded graph $G = (N, E)$ is constructed inductively starting with a map I from N into a universe of sets as follows

$$R(v) = \{R(u) : u \Rightarrow v \text{ is in } E\} \cup I(v). \tag{4}$$

The elements of $I(v)$ (when nonempty) are the so-called *individuals*.

To see how this approach is a generalization of the instantiation techniques discussed in the previous sections, we show how to build a skeletal model for a formula Q of MLS for which a set of places $\Pi = \{\pi_1, \dots, \pi_n\}$ and an ordering $<$ of the variables in Q satisfying conditions (a) and (b) of Theorem 2.3 are given.

We put

$$N = pow(\Pi) \cup \Pi_1,$$

with $\Pi \cap pow(\Pi) = \emptyset$ and $|\Pi_1| = |\Pi|$. Moreover, let $\iota : \Pi \longrightarrow \Pi_1$ be a one-one map. Then E consists of the following edges:

(a) the edge $\{\pi : \pi(x) = 1\} \Rightarrow A$ is in E for each variable x in Q and each $A \subseteq \Pi$ such that $\pi^x \in A$;

(b) the edge $\iota(\pi) \Rightarrow A$ is in E for each $\pi \in \Pi$ and $A \subseteq \Pi$ such that $\pi \in \Pi$.

It is then easily seen that $G = (N, E)$ is a well-founded quasi-extensional graph. Moreover, if we put

$$I(v) = \begin{cases} \{\{n+1, i\}\} & \text{if } v = v_i \\ \emptyset & \text{if } v \in N \setminus \Pi_1, \end{cases}$$

where $v_1, \dots, v_n$ is an enumeration of Π_1, it turns out that (4) defines a representation of the graph G. A model M for Q can then be obtained by putting

$$Mx = R(\{\pi : \pi(x) = 1\})$$

(see Section 2.1).

This approach, which is quite redundant for the simple MLS case, allows to set a clear distinction between the initialization phase and the stabilization phase or $\in$-phase. The initialization phase consists in choicing the map $I(v)$, whereas the stabilization phase is the core of the construction process which actually builds the internal structure of the skeletal model. Moreover, the construction process can be stated nondeterministically, since at the stage of defining the skeletal model one does not need to be specific on the nature of the elements involved during the stabilization phase.

2.8 MLS with the unary intersection.

Another important set-theoretic operator which has been considered very recently in the field of Computable Set Theory is the unary intersection $\bigcap$, defined by

$$\bigcap s = \{u : u \in t \text{ for all } t \in s\}.$$

In [12] it is shown that the theory MLSSRI obtained by extending MLS with singleton, rank comparison, and unary intersection is decidable.

To simplify the statement of the satisfiability conditions, we give the following definitions.

Let Q be a normalized conjunction of MLSSRI, i.e., a conjunction of literals of the following types

$$(=)\ y_i = y_j \cup y_k, y_i = y_j \setminus y_k$$

$$(\{\cdot\})\ y_i = \{y_i\}$$

$$(\leq, <)\ y_i \leq y_j, y_i < y_j$$

$$(\bigcap)\ y_i = \bigcap y_j,$$

whose distinct variables are $V = \{y_1, y_2, \dots, y_m\}$.

Let $\Pi = \{\pi_1, \dots, \pi_n\}$ be a set of places for Q.

DEFINITION 2.16. *We say that a place $\pi_i \in \Pi$ is a* singleton place *if there exist y_s, y_t such that $y_s = \{y_t\}$ is in Q and $\pi_i(y_s) = 1$. We denote by SING the set of singleton places.* ■

In the following we will identify variables y_s and places π_i with their indices s and i, respectively.

To each variable y_s we associate the set of places

$$\Pi(y_s) = \{i : \pi_i(y_s) = 1\}.$$

Also, given a map $F : \{1, 2, \dots, m\} \to \{1, 2, \dots, n\}$, to each place π_i we associate the set of variables

$$V(\pi_i) = \{s : F(s) = i\}.$$

Finally, given an increasing sequence of integers $r_0 = 0 < r_1 < \dots < r_l = n$, we introduce the map $R : \{1, \dots, n\} \to \{1, \dots, l\}$ defined by

$$R(i) = \min\{h : r_{h-1} < i \leq r_h\} \tag{5}$$

and also put

$$s^* = \max\{R(j) : j \in \Pi(y_s)\}, \tag{6}$$

for each $s \in V$.

The presence of the singleton operator lead us to introduce the following notion of trapped places.

DEFINITION 2.17. *An* admissible set of trapped places *(with respect to a set of places Π, a map F, and an increasing sequence of integers $r_0 = 0 < r_1 < \dots < r_l = n$) is any subset $\mathcal{T}$ of Π such that:*

(a) if $i \in \mathcal{T}$, then $i' \in \mathcal{T}$, for all $i' \in \{1, \dots, r_{R(i)}\}$; and

(b) if $i \in$SING and $\Pi(y_s) \subseteq \mathcal{T}$ for all $s \in V(\pi_i)$ then $i \in \mathcal{T}$.

A place π_i is said to be trapped if $i \in \mathcal{T}$. A variable y_s is said to be trapped if $\Pi(y_s) \subseteq \mathcal{T}$. ■

The logical weight of clauses of types $y_s = \bigcap y_t$ and of type $y_t = \{y_s\}$ is taken into consideration in the following definition.

DEFINITION 2.18. *The I-graph G of Q is the graph having Π as set of nodes and such that $\pi_i \to \pi_j$ is an edge of G whenever either*

(a) there is a literal $y_s = \bigcap y_t$ in Q such that $\pi_i(y_s) = 1$ and $\pi_j(y_t) = 1$, or

(b) there is a literal $y_t = \{y_s\}$ in Q such that $\pi_i(y_s) = 1$ and $\pi_j(y_t) = 1$. ■

The satisfiability conditions for MLSSRI formulae in normal form are stated in the following theorem.

THEOREM 2.19. ([12]) *Let Q be a normalized conjunction of MLSSRI. Let $V = \{y_1, \ldots, y_m\}$ be the collection of variables occurring in Q. Then Q is injectively satisfiable if and only if there exist a set of places $\Pi = \{\pi_1, \ldots, \pi_n\}$, a map $F : \{1, \ldots, m\} \to \{1, \ldots, n\}$, a sequence $r_0 = 0 < r_1 < \ldots < r_l = n$, and an admissible set of trapped places $\mathcal{T} = \{1, \ldots, r_{h_1}\} \subseteq \Pi$ such that:*

(1) no two variables in Q are Π-equivalent;

(2) there exist nonempty, pairwise disjoint hereditarily finite sets of rank not greater than h_1, $\bar{\pi}_1, \ldots, \bar{\pi}_{r_{h_1}}$, such that, defined $\overline{M}y_i = \bigcup_{\pi_s(y_i)=1} \bar{\pi}_s$, then $\overline{M}$ is a model for the clauses in Q involving only trapped variables and such that if $\overline{M}y_i \in \bar{\pi}_j$ then $F(i) = j$;

(3) if $\pi_j(y_s) = 1$ then $R(j) < R(F(s))$, where R is defined in (5);

(4) if $y_s = \{y_t\}$ is in Q then:

(4.1) $\Pi(y_s) = 1 = \{\pi_{F(t)}\}$,

(4.2) $F^{-1}(F(t)) = \{t\}$, and

(4.3) $R(F(t)) = t^* + 1$, where * is defined in (6);

(5) if $y_s \leq y_t$ [resp. $y_s < y_t$] is in Q then $s^* \leq t^*$ [resp. $s^* < t^*$];

(6) if $y_s = \bigcap y_t$ is in Q then

(6.1) $\Pi(y_t) \neq \emptyset$,

(6.2) if $\pi_i(y_s) = 1$ and $\pi_{F(u)}(y_t) = 1$ for some $u \in \{1, \ldots, m\}$ then $\pi_i(y_u) = 1$,

(6.3) *if* $\pi_i(y_s) = 1$ *and* $\pi_j(y_t) = 1$ *then* $R(i) < R(j)$,

(6.4) *if* $\Pi(y_t) \subseteq SING$, *then if* $\pi_i(y_u) = 1$ *for all* $0 \leq u \leq m$ *such that* $\pi_{F(u)}(y_t) = 1$, *then* $\pi_i(y_s) = 1$.

(7) *Let* G *be the I-graph of* Q. *Then, if* Q *contains the conjunct* $y_s = \bigcap y_t$ *and there is an* $i \in \Pi$ *such that* $\pi_i \rightarrow \pi_j$ *for all* $j \in \Pi(y_t)$ *then* $\pi_i(y_s) = 1$. ■

The structure of the I-graph is exploited during the instantiation of a model for Q, assuming that the hypotheses of Theorem 2.19 are all satisfied. Specifically, suppose that there exist Π, F, $r_0 = 0 < r_1 < \ldots < r_l = n$, and $\mathcal{T} = \{1, \ldots, r_{h_1}\}$ such that conditions (1)-(7) are satisfied.

Put

$$\gamma = \sum_{\pi_i \in \mathcal{T}} |\bar{\pi}_i| + m + 2n + 1$$

and let

$$I_i = \{0, 1 \ldots, \gamma + S(i) + 2\} \setminus \{i\}$$

$$\dot{I}_i = \{0, 1, \ldots, \gamma + S(i)\} \setminus \{i\}$$

$$\hat{I}_i = \{\{\dot{I}_i, z\} : z \in \dot{I}_i\}$$

where

$$S(i) = \begin{cases} R(i) & \text{if } i \leq n \\ R(i-n) & \text{otherwise.} \end{cases}$$

Then the following procedure instantiate a model for Q.

```
PROC Buil-model;
    FOR i := 1 TO n DO
       L_i := ∅
    END_FOR
    FOR i ∈ T DO
       FOR z ∈ π̄_i DO
          Propagate (z, π_i)
```

```
    END_FOR
  END_FOR
  FOR i := 1 TO r_{h_1} DO
    σ_i := π̄_i
  END_FOR
  FOR i := r_{h_1} + 1 TO n DO
    IF (i ∈ SING) THEN DO nothing
    ELSE IF (L(i) = φ) THEN
        σ_i := {Î_i, Î_{n+i}}
      ELSE
        J_i := Î_i ∪ L(i)
        J_{n+i} := Î_{n+i} ∪ L(i)
        σ_i := {J_i, J_{n+i}}
      END_IF
    END_IF
    FOR s ∈ V(π_i) DO
(*)       My_s = ⋃_{π_j(y_s)=1} σ_j
    END_FOR
    σ_i := σ_i ∪ {My_s : s ∈ V(π_i)}
    FOR z ∈ σ_i DO
      Propagate (z, π_i)
    END_FOR
  END_FOR
END_PROC
```

The code of the procedure *Propagate* is given below

```
PROC Propagate;
  FOR each j ∈ Π\(T ∪ SING) DO
    IF π_i → π_j is in G THEN
      L(j) := L(j) ∪ {z}
```

END_IF

END_FOR

END_PROC

Then, it turns out that the assignment M defined in line (*) is a model for Q. Moreover, it follows from such a construction that every satisfiable formula Q of MLSSRI always possesses a model whose rank is finite and bounded by a number depending solely on the length of Q and *a priori* determinable.

2.9. MLS with a choice operator η.

A choice operator η is a map from the von Neumann universe $\mathcal{V}$ of all sets into itself such that $\eta s \in s$, for every nonempty set s.

By no means such a property characterizes univocally a choice function η, thus, for decidability purposes, it is necessary to consider also additional conditions. For this reason we require that for every nonempty set s, ηs must be the *minimum* of s with respect to a given well-ordering $\prec$ of all sets. Furthermore, we require that for a well-ordering of all sets to be admissible, it must satisfy the following axioms:

1. $rank(x) < rank(y) \rightarrow x \prec y$ (monotonicity);
2. for all finite sets x, y,

$$x \prec y \leftrightarrow \max_{\prec}(x \backslash y) \prec \max_{\prec}(y \backslash x)$$

(antilexicographicity);

3. if x is finite, then $x \prec y$ for all infinite sets y such that $rank(y) = rank(x)$.

[46, 16] describe well-orderings of $\mathcal{V}$ satisfying axioms 1-3, thus showing their consistency.

Notice that axioms 1-3 are not powerful enough to characterize univocally a well-ordering of all sets, so that we have still an entire class of choice functions. This indefiniteness leads to various different satisfiability notions (cf. [46]).

Let C denote a privileged version of the choice operator η. Then a formula Q is said

- *weakly satisfiable*, if Q admits a K-model for some version K of η;
- *strongly satisfiable*, if Q admits a K-model for every version K of η;
- *uniformly satisfiable*, if Q admits a C-model which is also a K-model for every version K of η.

The results we will review here refer to the strong satisfiability notion, with respect to well-orderings of all sets satisfying axioms 1-3 (and possibly others).

The *finite* satisfiability problem for MLS extended with η has been solved in [35]. In the same paper, it is also shown that the ordinary satisfiability problem for conjunctions of unquantified formulae involving only $=$ and $\in$ (i.e., when the operators $\cup$ and $\setminus$ are dropped from the MLS constructs) is solvable. Both such results have been subsequently extended in [49] where occurrences of the singleton operator are allowed. The main result is the following reflection principle:

THEOREM 2.20. ([49]) *Given n, there is an effectively computable function $h(n)$ such that for every n-tuple $a_1, \ldots, a_n$ of finite sets, there are hereditarily finite sets $a_1^*, \ldots, a_n^*$ of rank less than $h(n)$ such that every propositional combination of atoms of type*

$$x = y \cup z, \; x = y \setminus z, \; x = \eta y, \; x \in y, \; x = \{y\}$$

with n free variables which is satisfied by $a_1, \ldots, a_n$ is also satisfied by $a_1^*, \ldots, a_n^*$. ■

Clearly, the preceding theorem implies the decidability of MLS plus η and singleton with respect to finite satisfiability.

In the same paper it is also proved another reflection result which, in view of Theorem 2.20, allows to establish the decidability of the *ordinary* satisfiability problem for propositional combinations of atoms of type

$$x = y,\ x \in y,\ x = \{y\},\ x = \eta y,\ x = \emptyset. \tag{7}$$

Specifically, it is proved.

THEOREM 2.21. ([49]) *Given an n-tuple* $a_1, \ldots, a_n$ *of sets, there are finite sets* $a'_1, \ldots, a'_n$ *such that every propositional combination of atoms of type (7) with n free variables which is satisfied by* $a_1, \ldots, a_n$ *is also satisfied by* $a'_1, \ldots, a'_n$. ■

It appears that axioms 1-3 are not sufficient to guarantee the solvability of the strong ordinary satisfiability problem for fragments of set theory which involve at least $\emptyset$, η, $\in$, $\subseteq$, $=$. For this reason, [33] and [18, 19] introduce new axioms which force the well-orderings to be «random», in the following sense.

For simplicity, we consider the case of the theory MLS plus singleton and choice function. Consider the formula

$$y \neq \emptyset \ \&\ \eta\{x, y\} \neq y \ \&\ x = y \cup \{\emptyset\} \tag{8}$$

(where $\eta\{x, y\} \neq y$ simply expresses $x \prec y$). By putting

4a. $y \cup \{\emptyset\} \prec y$, for every infinite set y such that $\emptyset \notin y$;

4b. $y \prec y \cup \{\emptyset\}$, for every set y such that $\emptyset \notin y$;

it is easy to see that the two groups of axioms 1-3,4a and 1-3,4b are both consistent. Therefore, we can conclude that there are well-orderings satisfying axioms 1-3 and in which (8) is satisfiable and well-orderings axioms 1-3 in which (8) is unsatisfiable. This amounts to say that MLS plus $\{\cdot\}$ and η is not decidable, at least with respect to the strong satisfiability problem, when only axioms 1-3 are assumed;

Much the same argument could be carried out by substituting in (8) the set $\{\emptyset\}$ by any fixed hereditarily finite set. Thus MLS plus $\{\cdot\}, \eta$ is also undecidable with respect to axioms 1-3,4a (and 1-3,4b too).

What is needed is an axiom which postulates the «randomness» of any admissible well-ordering, i.e. a property which falsifies any possible axiom of type 4a or 4b. From this it then would follow the strong satisfiability of any formula which is not «manifestly» contradictory, where a formula is manifestly contradictory if it contradicts one of the axioms of set theory or one of the axioms 1-3.

Such an approach has been pursued in [33] and [18, 19]. More precisely, [33] introduces the combinatorial notion of *ordered* $\in -\{\}-$*graph* and postulates that every ordered $\in -\{\}-$graph can be η-realized (we refrain from giving here the exact definitions of ordered $\in -\{\}-$graphs and η-realizations; the interested reader can refer to [33]). The axioms assumed on $\prec$ are only:

- $y \in x \rightarrow y \prec x$;
- $\prec$ is antilexicographic.

Then it proves that any formula Q of MLS plus $\{\cdot\}, \eta$ is satisfiable if and only if it is possible to effectively associate to Q certain finite objects (among which an ordered $\in -\{\}-$graph) satisfying suitable combinatorial conditions.

The papers [18, 19] are concerned with the decision problem for MLS with $\{\cdot\}, \eta$ and the predicates *Finite*(x) and rank comparison $x \leq y$ (standing for $rank(x) \leq rank(y)$). The axiom 3 is further strengthened as follows.

3'. If $rank(x) \leq rank(y)$, x is h-infinite, and y is k-infinite, then $x \prec y$ whenever $h < k$, where a set x is called q-infinite if q is the least integer for which there exists a sequence $x_q \in x_{q-1} \in \ldots \in x_1 \in x_0$ with $x_0 = x$ and x_q finite.

Then the notion of *syllogistic diagram* is introduced (extending that of ordered $\in -\{\}-$hraph), and again it is postulated that every syllogistic diagram is realizable (see [18, 19] for details). In addition, it is shown that any formula Q in the above language can be effectively decomposed in a formula of the type

$$(\exists x_1)\ldots(\exists x_n)(Q_1 \vee \ldots \vee Q_m), \tag{9}$$

where the variables $x_1, \ldots, x_n$ do not occur in Q and every Q_i involves only variables in Q in addition to $x_1, \ldots, x_n$, and such that

a. Q_i & Q_j is unsatisfiable, for every $i \neq j$, $i, j = 1, \ldots, m$;

b. Q_i is satisfiable, for every $i = 1, \ldots, m$;

c. Q is equisatisfiable with (9).

In other words, the formulae $Q_1, \ldots, Q_m$ allow to partition the models of Q and to describe them in the same language in which Q is written (see also [23,27]; cf. Section 2.2).

It would be very interesting to see whether there are alternative formulations of the «randomness» axiom which are elementarily expressible (like axioms 1-3,3′, for instance).

We close this section by mentioning a semi-decision result with respect to the *finite* satisfiability problem for a quite broad class of set-theoretic formulae involving the choice function η. In [16] it is shown that every *safe* formula in the language MLS extended with $\{\cdot\}$, *pow*, $\times$ (cartesian product), and η is finitely satisfiable if and only if it is hereditarily finitely satisfiable, where a formula is safe according to the following definition.

DEFINITION 2.22. *A* safe *formula is an unquantified formula in which every term t is either safe or of the form ηs, with s safe.*

Safe terms are recursively constructed as follows:

(I) $\emptyset$ is a safe term;

(II) each variable y is a safe term;

(III) if s, t are safe terms, then so are $(s \cap t)$, $(\eta s \cap t)$, $(t \cap \eta s)$, $(s \backslash t)$, $(s \backslash \eta t)$, $(s \cup t)$, $pow(s)$, $(s \times t)$;

(IV) if $s_0, s_1, \ldots, s_q$ are safe terms, and t_i is either s_i or ηs_i, $i = 0, 1, \ldots, q$, then

$$\{t_0, t_1, \ldots, t_q\} \text{ and } \eta\{s_0, s_1, \ldots, s_q\}$$

are safe terms. ■

As a by product, one obtains a semidecision procedure *à la* Herbrand for the finite satisfiability problem for safe formulae.

3. Quantified theories.

In this section we consider prenex formulae of the form $Q_1 Q_2 \ldots Q_n p$, where

(1) p is a boolean combination of literals of type $x \in y$ and $x = y$;

(2) either all Q_i are of the form $\exists y_i \in z_i$ or every Q_i is of the form $\forall y_i \in z_i$, i.e. all quantifiers are restricted and there is no quantifier alternation;

(3) no z_j is a y_i, for any $i, j = 1, \ldots, n$, i.e. nesting of bound variables if forbidden.

We refer to conjunctions of such formulae as $(\forall)_0$-*simple prenex formulae* (in short, $(\forall)_0$-s.p. formulae).

After considering formulae of this class, a further subsection is devoted to the classical Behmann theory, i.e. the family of arbitrarily quantified set-theoretic formulae not involving the membership relation $\in$ (see [2], and [51]). Finally three additional results concerning complete classes of formulae are reviewed (see [39], [40], and [55]).

3.1. The theory of $(\forall)_0$-s.p. formulae.

Let T be any unquantified theory for which there is a satisfiability algorithm. Let φ be a propositional combination of prenex formulae of type $Q_1 Q_2 \dots Q_n p$, where $n > 0$ and Q_i is either $(\forall x_i)$ or $(\exists x_i)$, and p is an unquantified formula of T. The following algorithm, based on elimination of quantifiers, is a semidecision procedure for this class of formulae (cf. [4]).

1. Bring φ into disjunctive normal form, i.e. rewrite it as a disjunction $\varphi_0 \vee \varphi_1 \vee \dots \vee \varphi_m$, where each φ_i is a conjunction of prenex formulae.
2. If for each $i = 0, 1, \dots, m$ the procedure TEST (ϕ_i) described below returns «unsatisfiable», then return «unsatisfiable». Otherwise, return «ambiguous».

The procedure TEST (ψ) applies to conjunctions ψ of prenex formulae with matrix in T.

Procedure TEST (ψ).

1. Let Φ be the set of all free variables and constants in ψ.
2. While there exists a formula of type $(\exists x) q(x)$ in ψ, replace it by $q(z_x)$, where z_x is a newly introduced variable, and put
$$\Phi := \Phi \cup \{z_x\}.$$
3. If no quantifier is left in ψ, then apply the satisfiability algorithm available for the quantifier-free theory T to which ψ belongs, and return the result of this test.
4. Otherwise, nondeterministically choose some universally quantified formula $(\forall x) q(x)$ in ψ and replace it by the conjunction $\&_{z \in \Phi} q(z)$.
5. Go to Step 2.

The above algorithm is sound but not complete in general. It becomes complete for certain restricted classes of formulae. As shown in [4], this is the case when the underlying quantifier-free theory T is the theory T_1 in the language consisting of individual variables and the relations $=, \in$, and the prenex constituents satisfy properties (2), (3) stated at the beginning of this section.

The same statement holds in the extended case in which the underlying theory T is the extension T_2 obtained from T_1 by also allowing:

(a) two new constants N and O which must be interpreted as the set of intergers and the set of all accessible ordinals, respectively;

(b) a new infinity $f, g, h, \ldots$ of map variables, and also a domain operator D.

A variety of constructs can be expressed in this theory. Some examples are:

$$x \subseteq y \equiv_{Def} (\forall z \in x)(z \in y)$$

$$x = \{y\} \equiv_{Def} y \in x \,\&\, (\forall z \in x)(z = y)$$

$$x \text{ is an (accessible) ordinal} \equiv_{Def} x \in O$$

$$x \text{ is a set of integers} \equiv_{Def} (\forall z \in x)(z \in N)$$

$$f = g \equiv_{Def} Df = Dg \,\&\, (\forall z \in Df)(f(z) = g(z))$$

$$f \text{ is a } t\text{-uple} \equiv_{Def} Df \in N$$

$$f \text{ is one-to-one} \equiv_{Def} (\forall x \in Df)(\forall y \in Df)$$

$$(x \neq y \rightarrow f(z) \neq f(y)).$$

Interesting elementary theorems which can be expressed in this decidable theory are:

- the successor of any ordinal is an ordinal;
- there are no ordinals between an ordinal and its successor;
- any two ordinals are comparable.

Finally, we mention that even in the presence of alternating quantifiers, a simple variant of the algorithm described above often teminates successfully. For example, the following statement

«If x is a limit ordinal, then $x = Un(x)$», (where a limit ordinal is an ordinal having no immediate predecessors) can be proved by applying the algorithm above with a bit of backtracking.

Conditions (1), (2), and (3) have been recently loosened in various ways still maintaining decidability.

A class of formulae in which restricted forms of nestings of variables are permitted has been studied in [49]. Following [49], we give

DEFINITION 3.1. *A* $(\forall)_0$*-prenex formula* φ *is in normal form if the matrix of each conjunct in* φ *is a disjunction of literals.*

Given a normal form $(\forall)_0$*-p. formula* φ *and a variable* z*, we say that* z *is nested in* φ *at level* l *if for some* $x, y, x_0, \ldots, x_{l-1}$ *the sequence of quantifiers*

$$(\forall x \in y)(\forall x_0 \in x) \ldots (\forall x_{l-1} \in x_{l-2})(\forall z \in x_{l-1})$$

occurs in a conjunct of φ. ■

We denote by $(\forall)^{-}_{0,l}$ the class of all $(\forall)_0$-p. formulae φ in normal form such that every nested variable in φ has level not greater than $l+1$ and can occur only in negative literals (of the matrix of the conjunct in which it is nested). $(\forall)^{-}_{0,l}$-formulae are also called nested negative $(\forall)_0$-p. formulae of level l.

The decidability of the class $(\forall)^{-}_{0,l}$ is an immediate consequence of the following strong reflection principle over the hereditarily finite sets for the class $(\forall)^{-}_{0,l}$.

THEOREM 3.2. ([49]) *Given* n *and* l*, it can be effectively determined an integer* $h(n, l)$ *such that for any* n*-tuple of sets* $a_1, \ldots, a_n$ *there are hereditarily finite sets* $a^*_1, \ldots, a^*_n$ *of rank less than* $h(n, l)$ *such that every* $(\forall)^{-}_{0,l}$*-formula satisfied by* $a_1, \ldots, a_n$ *is also satisfied by* $a^*_1, \ldots, a^*_n$.

Since sets of bounded finite rank are only finitely many, an exhaustive search among assignments of rank less than $h(n, l)$ allows to establish the satisfiability of any given nested negative formula of level l with n free variables.

Other generalizations of the decidability of the class of $(\forall)_0$-s.p. formulae consist in extending the language of the basic unquantified theory T_0.

In [13] it is shown that the *finite* satisfiability problem for $(\forall)_0$-s.p. formulae admitting also the rank operator is decidable. We mention only that yet another variant of trappedness is introduced to solve the decision problem for such a class of formulae.

Of the same flavor is the decidability result concerning $(\forall)_0$-s.p. formulae extended with the predicate *Finite* (x), contained in [9]. Such result, which plainly refers to ordinary satisfiability, has been obtained under the syntactic restrictions that the predicate *Finite* can apply only to free variables and that the relations $x = y$ and $x \neq y$ never applies to pairs of bound variables x, y.

The same paper describes also a procedure for the elimination of quantifiers from set-theoretic formulae, which, under certain conditions, allows to lift to the quantified case most of the decidability results for classes of quantifier-free set-theoretic formulae of the kind seen in Section 2.

Specifically, let $\mathcal{L}$ be a language for set theory which includes, among others, the constant $\emptyset$, the set operators $\cup, \cap, \setminus, \{\cdot\}$, and the set predicates $=, \in$.

DEFINITION 3.3. *A formula φ in $\mathcal{L}$ is called 0-flat if each quantified variable x in φ appears only within atoms of type*

$$x = t \quad \textit{or} \quad x \in t,$$

where all the variables occurring in t are free. ■

[9] describes how given a 0-flat formula φ of $\mathcal{L}$ one can effectively construct a quantifier-free formula ψ of $\mathcal{L}$ such that

$$\vdash_{ZF} (\phi \leftrightarrow \psi).$$

Therefore if the class of quantifier-free formulae in the language $\mathcal{L}$ is decidable, so is the class of quantified 0-flat formulae of $\mathcal{L}$.

We close this section by mentioning a very interesting open problem.

Is the class of $(\forall)_0$*-prenex formulae decidable if condition (3) which forbids nesting of variables is dropped?*

In Section 5 we will discuss some undecidability results concerning theories closely related to such a class of formuale.

3.2. Behmann theory and some extensions.

In a famous 1922 paper, Behmann exhibits an effective procedure for calculating the truth value of any formula quantified over sets and involving only Boolean connectives $\&, \vee, \neg, \rightarrow, \leftrightarrow$, set inclusion and inequality, the set cardinality operator $|\cdot|$, integer constants, and inequalities (cf. [2]). This result is extended in [51], which gives an algorithm that decides the combination of Behmann theory with Presburger arithmetic (PB-theory) (see [50], cf. also [31]).

Two kinds of variables can occur in PB-theory: set variables and integer variables. Set variables range in the class of all *finite* sets; integer variables range in the set N of natural integers. Union, intersection, and set difference are the only set operators allowed. Also allowed are a cardinality operator and ordinary arithmetic addition among integer variables. Besides equality, set inclusion is the only relation in PB-theory.

The decision method described in [51] for PB-theory is based on elimination of quantifiers. First of all it can be shown that one can limit oneself to considering only prenex formulae of the type $Q_1 Q_2 \ldots Q_n p$, where each Q_i is either $(\exists x)$ or $(\forall x)$, with x either a set variable or an integer variable, and where p is a quantifier-free conjunction of literals of PB-theory. Elimination of quantifiers over integer variables proceeds as usual via the introduction of suitable congruence relations. In order to eliminate quantifiers over set variables, for each part of the Venn diagram of the set variables in the formula, one introduces a new integer variable, which stands for the cardinality of the corresponding part of the Venn diagram. Then set relations involving the quantified set variables are eliminated by adding suitable arithmetic relations on the newly introduced integer variables. The reason why set variables can be eliminated in favor of integer variables lies on the fact that in PB-theory formulae of the type $x \in y$ or $x \notin y$ cannot be expressed. This implies that no set can have elements having a proper individuality, and consequently the main characteristic of a set in this context is its cardinality.

The same set of formulae is still decidable even if set variables are allowed to range over the full von Neumann universe and integer variables are interpreted more generally as cardinal numbers (TPB-theory). Much the same technique applies in this case, but here one relies on Tarski's result about the decidability of the theory of cardinal numbers with addition (cf. [53]) rather than on the decidability of Presburger arithmetic.

Note that [51] also shows how much the same algorithm can serve to instantiate existentially quantified variables, thus solving the problem of set theoretic instantiation in this special case; this is an important issue in mechanical theorem proving (cf. [52]).

An interesting example of a formula belonging to both PB-theory and TPB-theory is

$$(\forall x)(\forall y)[(x \subseteq y \ \& \ |x| = |y|) \to x = y)].$$

This formula, which is true as a PB-formula, is false in TPB-theory, since infinite sets can have equipollent proper subsets.

Further generalizations considered in [51] concern the instantiation problem in presence also of map variables f. More precisely, the map constructs admitted are of type $f[E]$, with E any valid set expression, and of type Df, which denotes the domain of f.

3.3. Three complete families of quantified formulae in set theory.

In the paper [30], Ehrenfeucht introduces the notion of an «N-game» played by two players with structures A_1 and A_2 for an unquantified predicate language L. An N-game consists in a sequence of N moves. In the i-th move, player 1 chooses an element in the structure A_{l_i}, with $l_i \in \{1,2\}$, and player 2 chooses an element in the structure A_{3-l_i}. At the end of the game $a_{11}, a_{12}, \ldots, a_{1N}$ have been picked, in that order, from A_1 and $a_{21}, a_{22}, \ldots, a_{2N}$ from A_2. Player 2 wins if after this game the correspondence

$$a_{11} \leftrightarrow a_{21}$$

$$a_{12} \leftrightarrow a_{22}$$

$$\ldots$$

$$a_{1n} \leftrightarrow a_{2N}$$

is an isomorphism of these sequences with respect to the relations defined in L. That is, player 2 wins it for every quantifier free formula $F(x_1, x_2, \ldots, x_N)$ in L, $F(a_{11}, a_{12}, \ldots, a_{1N})$ is true in A_1 if and only if $F(a_{21}, a_{22}, \ldots, a_{2N})$ is true in A_2. Ehrenfeucht's main theorem is

THEOREM 3.4. ([30]) *If there is a closed formula in L with N quantifiers which is satisfied by structure A and not by structure B, then player 1 has a winning strategy in an N-game.* ∎

An important application of Ehrenfeucht's theorem is the following

THEOREM 3.5. ([40]) *The language of all 3-quantifier closed prenex formulae in set theory involving both* $\in$ *and* $=$ *is decidable.* ∎

The proof proceeds by showing that in a 3-game played with any two models for set theory, player 2 has a winning strategy. By Ehrenfeucht's theorem this implies that every sentence with three quantifiers is either true or false in all models of set theory. But then it follows from Gödel's completeness theorem that for every such sentence φ, either φ or $\neg\varphi$ is a theorem. Since in an axiomatic theory, theorems can be enumerated, dovetailing gives a procedure for establishing whether or not a 3-quantifier sentence in set theory is a theorem. Observe that the decision algorithm suggested by the proof above is not elementary recursive.

Another result by Gogol is the following

THEOREM 3.6. ([39]) *All closed sentences of standard Zermelo-Fraenkel theory of the form* $(\forall x_1)(\forall x_2)\dots(\forall x_n)(\exists x_{n+1})\,S(x_1, x_2, \dots, x_{n+1})$, *where* $S(x_1, x_2, \dots, x_{n+1})$ *is a quantifier-free formula and may contain the predicates* $=$ *and* $\in$, *are decidable.* ∎

As before, in view of the preceding completeness result, it follows that the validity problem for the closed $\forall_n\exists$-formulae is solvable, by way of the enumeration algorithm.

Notice, also, that the satisfiability problem for MLSS (cf. Section 2.1) can be reduced to the validity problem for closed $\forall_n\exists$-formulae, as the following argument shows. Preliminarily, observe that the satisfiability problem for MLSS is immediately reducible to the satisfiability problem for conjunctions of atoms of the following three types

$$x = y \cup z, \;\; x = y \setminus z, \;\; x = \{y\}.$$

So, let $q_1 \,\&\dots\&\, q_n$ be such a conjunction, involving only the variables $x_1, \dots, x_m$. Let w be a variable not occurring in any of the q_i. Each conjunct q_i can be easily rewritten as $(\forall w)\; q_i'$, with q_i' involving only the variables in q_i, the variable w, the set relators $=, \in$, and the propositional connectives. For instance, $x = \{y\}$ can be rewritten as $(\forall w)\;(y \in x \;\&\; (w \in x \rightarrow w = y))$. Thus, by using the standard quantifier manipulation rules, it follows that $q_1 \,\&\dots\&\, q_n$ is equivalent to $(\forall w)\;(q_1' \,\&\dots\&\, q_n')$, so that it follows that $q_1 \,\&\dots\&\, q_n$ is satisfiable if and only if the formula

$$(\exists x_1)\dots(\exists x_m)\,(\forall w)\,(q_1' \,\&\dots\&\, q_n') \tag{10}$$

is true in some model of ZF. At this point, it is enough to observe that the negation of (10) is a closed $\forall_n\exists$-formula of the type considered in Gogol's theorem.

The previous considerations suggest an alternative decision method for MLSS, namely enumeration of ZF theorems. It must be stressed, however, that the decision procedure given in [36] (and furtherly optimized in [27]) is far more efficient.

Gogol's result [39] generalizes an earlier theorem of Ville's (cf. [55]) on the completeness of the purely existential set-theoretic formulae in the language $\mathcal{L}$ consisting of $\emptyset$, $\{\cdot\}$, $\cup$, $\in, =$, with respect to the theory of sets with the following axioms:

(I) axioms for equality;

(II) extensionality;

(III) $(\forall x)\ (x \notin \emptyset)$ (axiom for empty set);

(IV) $(\forall x)\ (\forall y)\ (\forall u)\ (u \in x \cup y \leftrightarrow u \in x \vee u \in y)$ (axiom for binary union);

(V) $(\forall x)\ (\forall u)\ (u \in \{x\} \leftrightarrow u = x)$ (axiom for singleton);

(VI_n) $\neg(\exists x_1)\dots(\exists x_n)(x_1 \in x_2 \,\&\, \dots \,\&\, x_{n-1} \in x_n \,\&\, x_n \in x_1)$, for each integer $n \geq 1$ (regularity axiom in the weak form).

Ville's completeness result is proved by using model-theoretic arguments.

4. Some applications.

The decidability results and proof techniques presented in the preceding sections have also been used to find other decidability results in domains different from pure set theory. In this section we will briefly review three decidable theories which have been shown decidable either by explicit reduction to already known decidable fragments of set theory or by employing techniques similar to the ones discussed previously. More specifically, we will consider a fragment of general topology (cf. [25, 26]), a theory of directed graph (cf. [44, 45]), and a theory of partially ordered sets with monotone functions (cf. [20]). In all three cases, explicit quantification is forbidden.

4.1. Topological syllogistic.

The basic constituents of the two-level multi-sorted language with functions $\mathcal{L}^{2,\infty}$ considered in [25, 26] are:

(a) for each n, the constants $0^{(n)}$ and $1^{(n)}$, designating respectively the empty set and the support of the n-th topological space;

(b) for each n, two denumerable sequences of variables:
- *individual variables* $x_0^{(n)}, x_1^{(n)}, \ldots, y_0^{(n)}, y_1^{(n)}, \ldots,$ etc., ranging over elements of $1^{(n)}$,
- *set variables* $X_0^{(n)} X_1^{(n)}, \ldots, Y_0^{(n)}, Y_1^{(n)}, \ldots,$ etc., ranging over subsets of $1^{(n)}$;

(c) the operators $^{-(n)}$ (topological closure), $'^{(n)}$ (set complementation), $\cup, \cap, \setminus$;

(d) for each pair (n, m) of natural numbers, two denumerable sequences of function variables:
- $f_1^{(n,m)}, f_2^{(n,m)}, \ldots,$ denoting *continuous maps* from $1^{(n)}$ into $1^{(m)}$,
- $g_1^{(n,m)}, g_2^{(n,m)}, \ldots,$ denoting *closed maps* from $1^{(n)}$ into $1^{(m)}$,

(e) the standard relators $=, \in, \subseteq$;

(f) the propositional connectives $\neg, \&, \vee, \rightarrow, \leftrightarrow$, and parentheses.

Terms (of sort n) of $\mathcal{L}^{2,\infty}$ are individual and set variables, constants, and compound terms of the form $T_1^{(n)} \cup T_2^{(n)}$, $T_1^{(n)} \cap T_2^{(n)}$, $T_1^{(n)} \setminus T_2^{(n)}$, $(T_1^{(n)})', \overline{T_1^{(n)}}$, $h^{(m,n)}(t^{(m)})$, $(f^{(n,m)})^{-1}[T_3^{(m)}]$, $g^{(m,n)}[T_3^{(m)}]$, where $T_1^{(n)}$, $T_2^{(n)}$, $T_3^{(m)}$ are set terms, $t^{(m)}$ is an individual term, and $h^{(m,n)}$ stands for a continuous or closed map variable. (Notice that the inverse image operator can be applied only to continuous map variables, whereas the direct image operator can be applied only to closed map variables.)

Finally, the formulae of $\mathcal{L}^{2,\infty}$ are propositional combination of atoms of the forms

$$t_1^{(n)} = t_2^{(n)}, \quad t_1^{(n)} \in T_1^{(n)},$$
$$T_1^{(n)} = T_2^{(n)}, \quad T_1^{(n)} \subseteq T_2^{(n)}.$$

The semantics of the language $\mathcal{L}^{2,\infty}$ is defined in the most natural way.

To each formula p of $\mathcal{L}^{2,\infty}$, we associate a graph $G_p = (N_p, E_p)$, where $N_p = \{i_0, i_1, \ldots, i_k\}$ is the set of all sorts of individual and set variables present in p and where the edge $i_l \Rightarrow i_h$ is in E_p if and only if p contains either some occurrence of a continuous map variable of sort (i_h, i_l), or some occurrence of a closed map variable of sort (i_l, i_h). Then p is said to be *acyclic* if its associated graph G_p is acyclic.

In [25, 26], it is shown that the topological satisfiability problem for acyclic formulae of $\mathcal{L}^{2,\infty}$ can be reduced to the ordinary satisfiability problem for a restricted subclass of formulae of MLS, namely those in which the longest chain of $\in, \notin$ has length 1 and such that no variable on the left-hand side of a membership relation can appear in terms of type $x \cup y$ or $x \backslash y$ (these are the so-called *two-level syllogistic* (2LS) formulae; cf. [34]). More specifically, it is exhibited a procedure which eliminates topological and map constructs from acyclic formulae p of $\mathcal{L}^{2,\infty}$ by suitably introducing some 2LS formulae which possibly involve new variables. Literals containing variables relative to topological spaces which correspond to the leaves of the graph G_p are processed first. New topological and set-theoretic constraints are propagated along the edges of G_p. Subsequently, topological spaces corresponding to nodes of increasing height are dealt with. The acyclicity of G_p assures that such a process will terminate. On termination, all topological and map constructs are eliminated and a formula p^{**} of 2LS is obtained such that p is topologically satisfiable if and only if p^{**} is ordinarily satisfiable.

The language $\mathcal{L}^{2,\infty}$ is quite expressive. In fact, most of the elementary concepts and properties that can be typically found in the first chapters of introductory textbooks on general topology are expressible as acyclic formulae. For instance,

- *Int* $(A) \equiv_{Def} (\bar{A}')$ *(the interior of A);*
- $\partial(A) \equiv_{Def} \bar{A} \setminus$ *Int* (A) *(the boundary of A);*
- *open* $(A) \equiv_{Def} A = Int(A)$ *(A is an open set);*
- *closed* $(A) \equiv_{Def} A = \bar{A}$ *(A is a closed set);*
- *open-domain* $(A) \equiv_{Def} A =$ *Int* $(\bar{A})$ *(A is an open domain; see [42]);*
- *open-domain* $(A)\,\&$ *open-domain* $(B) \rightarrow$ *open-domain* $(A \cap B)$;
- *open-domain* $(A)\,\&$ *open-domain* $(B)\,\&\neg$ *open-domain* $(A \cup B)$;
- *open-domain* $(A)\,\&$ *open-domain* $(B) \rightarrow (A \subseteq B \leftrightarrow \bar{A} \subseteq \bar{B})$;
- etc.

4.2. *Graph Theory.*

The language of directed graph theory considered in [44, 45] contains besides the usual propositional connectives

(a) vertex variables $v_1, v_2, \ldots$;

(b) edge variables $e_1, e_2, \ldots$;

(c) graph variables $G_1, G_2, \ldots$;

(d) the constant $\emptyset$ (denoting the empty graph);

(e) the set operators $\cup, \cap$;

(f) the unary operators *head*(e) and *tail*(e) (which return the head and tail vertices of the edge e, respectively);

(g) the constructors *edge*(v_1, v_2), *vertexgraph*(v), and *edgegraph*(e) (*edge*(v_1, v_2) returns the directed edge having the vertices v_1 and v_2 as head and tail, respectively; *vertexgraph*(v) returns the singleton graph having as unique element the vertex v; *edgegraph*(e) returns the graph whose only elements are the edge e and its endpoints).

Notice that explicit quantification is not allowed.

An example of a valid well-formed formula in the above language is

$$vertexgraph(v_1) \cup (G_1 \cap G_2) =$$

$$= edgegraph(edge(v_2, v_3)) \rightarrow v_1 \in G_1 \cap G_2 \,\&\, (v_1 = v_2 \vee v_1 = v_3).$$

Given a conjunction S of the theory of directed graphs, [45] describes a decision procedure roughly consisting of the following steps (initially it is assumed that no set constructor of type (g) occurs in S):

1. eliminate compound terms by using new variables of the appropriate type;
2. form the transitive closure of = among vertex variables;
3. infer all possible edge equalities $e_1 = e_2$ from conjuncts of type $v_1 = head(e_1) \,\&\, v_1 = head(e_2) \,\&\, v_2 = tail(e_1) \,\&\, v_2 = tail(e_2)$ occurring in S;
4. form the transitive closure of = among edge and graph variables;
5. infer all possible membership relations from clauses in S of type $G_1 = G_2 \cup G_3$ and $G_1 = G_2 \cap G_3$ and membership clauses in S;
6. check for explicit contradictions (of form $v_1 \neq v_1$, or $v_1 \in G_1 \,\&\, v_1 \notin G_1$, etc.);
7. for each vertex variable v seek a singleton model of the set of literals $\{G \neq \emptyset : v \in G \text{ occurs in } S\} \cup \{G = \emptyset : v \notin G \text{ occurs in } S\}$ and literals of type $G_1 = G_2 \cup G_3$ and $G_1 = G_2 \cap G_3$ in S;
8. for each edge variable e seek a singleton model of the set of literals $\{G \neq \emptyset : e \in G \text{ occurs in } S\} \cup \{G = \emptyset : e \notin G \text{ occurs in } S\}$ and literals of type $v = head(e)$, $v = tail(e)$, $G_1 = G_2 \cup G_3$, and $G_1 = G_2 \cap G_3$ in S;
9. for each clause of type $G_1 \neq G_2$ in S seek a singleton model for the conjunction of $G_1 \neq G_2$ with all clauses in S of type $G_1 = G_2 \cup G_3$, and $G_1 = G_2 \cap G_3$ in S;

10. declare S unsatisfiable if either an explicit contradiction is found or one of steps 7, 8, 9 fails.

It is to notice that step 6 corresponds to the search of a set of *places* of S, whereas steps 7, 8, 9 correspond to the search of *places of S at (vertex and edge) variables* (cf. Definition 2.2 and subsequent paragraph).

Edge and graph constructors of type (g) are dealt with much in the same way as the singleton operator in the MLS context (see the paragraph preceding Theorem 2.5). For instance, if $G = vertexgraph(v)$ occurs in S, then the conjunction

$$\&_{w \text{ occurs in } S}(w \notin G \vee w = v) \,\&\, \&_{e \text{ occurs in } S}(e \notin G)$$

is added to S, and special care is taken during the instantiation of G, in order that G will not get any spurious element (different from v). Analogously if $G = edgegraph(e)$ occurs in S.

4.3. Theory of partially ordered sets.

We close this section by reviewing one of the results contained in [20], namely the decidability problem for the unquantified theory POSMF of partially ordered sets with (monotone) functions.

The symbols of the language POSMF are

- individual variables $x, y, z, \ldots$;
- function variables $f, g, h, \ldots$;
- the relators $=, <$;
- the predicates $up(f)$, $down(f)$ (denoting respectively that f is monotone nondecreasing in the first case, and monotone nonincreasing in the second case);
- the propositional connectives.

The formulae of POSMF are the propositional combinations of atoms of the following five types

$$x < y,\ x = y,\ x = f(y),\ up(f),\ down(f),$$

where x, y stand for individual variables and f stands for a function variable.

In [20] it is assumed that the domain of any interpretation of POSMF must be a partially ordered set $(D, \leq)$ such that

(i) for every $t, u \in D$ there are $s, v \in D$ such that $s \leq t$, $s \leq u$, $t \leq v$, and $u \leq v$ hold;

(ii) $\leq$ is antisymmetric, i.e., if $s \leq t$ and $t \leq s$ then $s = t$.

(Notice that if we strengthen (i) by requiring that every two elements in D must have a least upper bound and a greatest lower bound, then D is a lattice).

The literal *up*(f) (resp. *down*(f)) is true in an interpretation M of POSMF with domain D if the function $Mf : D \to D$ is nondecreasing (resp. nonincreasing). The satisfiability problem for POSMF can then be easily reduced to the satisfiability problem for MLS by the following procedure.

Let Q be a conjunction of literals of type:

$$\begin{array}{lllll} x \leq y, & x \not\leq y, & x = y, & x \neq y, & x = f(y), \\ up(f), & down(f), & \neg up(f), & \neg down(f). & \end{array}$$

1. For every clause in Q of type $\neg up(f)$ (resp. $\neg down(f)$) introduce four new individual variables x_1, x_2, y_1, y_2 and add the formula

$$y_1 = f(x_1) \,\&\, y_2 = f(x_2) \,\&\, x_1 \neq x_2 \,\&\, x_1 \leq x_2 \,\&\, y_2 \leq y_1 \,\&\, y_1 \neq y_2$$

(resp. $y_1 = f(x_1) \,\&\, y_2 = f(x_2) \,\&\, x_1 \neq x_2 \,\&\, x_1 \leq x_2 \,\&\, y_1 \leq y_2 \,\&\, y_1 \neq y_2$).

2. For every pair of clauses of Q of the form $x = f(y)$, $x' = f(y')$, add the formula $y = y' \to x = x'$.
3. For every triple of clauses of Q of the form *up*(f) (resp. *down*(f)), $x = f(y)$, $x' = f(y')$, add the formula $y \leq y' \to x \leq x'$ (resp. $y \leq y' \to x' \leq x$).

4. For every pair of variables x, x' in Q add the clause

$$x \leq x' \,\&\, x' \leq x \rightarrow x = x'.$$

5. Drop from Q all clauses involving function variables.
6. Replace each occurrence of the relator $\leq$ by $\subseteq$, regarding all individual variables as set variables.

Let Q' be the formula so obtained. We have then the following result:

THEOREM 4.1. ([20]) *Q is satisfiable by a partyally ordered model enjoying properties (i) and (ii) above if and only if Q' is satisfiable by a set model.* ■

5. Undecidability results.

The problem of finding undecidable classes of set-theoretic formulae is also of great relevance to the field of Computable Set Theory, since every undecidability result better delineates the boundary between what can be mechanized and the undecidable.

It is well known that Zermelo-Fraenkel set theory is undecidable, since Peano's arithmetic can be immediately reduced to it (cf. [24]). In view of the still open problem concerining the decidability of the $(\forall)_0$-p. formulae when nested variables are allowed (see Section 3.1), it is interesting to consider other restricted classes of formulae in which limited alternations of quantifiers are permitted. This has been first done in [48], where it has been shown that the class of $(\forall\exists\forall)_0$-formulae has an unsolvable satisfiability problem. We recall that a formula is of type $(\forall\exists\forall)_0$ if it is the conjuntion of restricted prenex formulae with at most two quantifier alternations.

The result in [48] is based on Gödel's incompleteness theorem and holds for any *reasonable* set theory, such as one which contains the theory Th_0 consisting of the following four axioms:

1. Empty set: $(\exists y)(\forall x)(x \notin y)$. i.e. there exists an empty set.
2. Extensionality: $(\forall x)(\forall y)((\forall z)(z \in x \leftrightarrow z \in y) \leftrightarrow x = y))$, i.e. two sets coincide if and only if they have the same elements.
3. One Element Addition Principle: $(\forall x)(\forall y)(\exists z)(\forall w)(w \in z \leftrightarrow (w \in x \vee w = y))$, i.e; if x and y are set, so is $x \cup \{y\}$.
4. One Element Subtraction Principle: $(\forall x)(\forall y)(\exists z)(\forall w)(w \in z \leftrightarrow (w \in x \vee w \neq y))$, i.e. if x and y are sets, so is $x \setminus \{y\}$.

Because of the extreme simplicity of the theory Th_0, an alternative way to encode syntactic objects and proof-theory predicates has been developed in [48]. Matters are also furtherly complicated by the fact that only $(\forall\exists\forall)_0$-formulae can be used in the coding process.

Such an undecidability result has been subsequently improved in [14], at least with respect to the compexity of the class of formulae involved. Indeed, [14] proves that the class of $(\forall\exists)_0$-formulae is undecidable with respect to extensions of ZF set theory, by exhibiting an effective encoding procedure which allows to associate to any polynomial Diophantine equation D a $(\forall\exists)_0$-formulae φ_D such that D has integer solutions if and only if φ_D is satisfiable (say, in the standard model of ZF set theory). Thus the undecidability of the $(\forall\exists)_0$-formulae follows at once from the unsolvability of Hilbert's tenth problem (cf. [43]).

More precisely, from the unsolvability of Hilbert's tenth problem, it follows easily that there can be no algorithm to test whether a system of integral equations of the following types

$$\begin{cases} \xi = \eta \\ \xi = \eta + \zeta \\ \xi = \eta \cdot \zeta \\ \xi = k \end{cases}$$

(where ξ, η, ζ are integer variables and k is an integer constant) has a solution.

So, let $\Sigma = \{\Sigma_1, \ldots, \Sigma_n\}$ be a system of equations of the above types, in which no variable can have multiple occurrence within the same equation Σ_i, $i = 1, \ldots, n$. To each integer variable ξ in Σ we associate n distinct set variables $x^1_\xi, \ldots, x^n_\xi$. Then we put:

$$\varphi_0 \equiv_{Def} \&_{(\xi,i)\neq(\eta,j)} x^i_\xi \cap x^j_\eta = \emptyset \,\&\, \&_{\xi,i,j} |x^i_\xi| = |x^j_\xi|$$

$$\varphi_i \equiv_{Def} \begin{cases} |x^i_\xi| = |x^i_\eta| & \text{if } \Sigma_i \text{ is of type } \xi = \eta \\ |x^i_\xi| = |x^i_\eta \cup x^i_\zeta| & \text{if } \Sigma_i \text{ is of type } \xi = \eta + \zeta \\ |x^i_\xi| = |x^i_\eta \times x^i_\zeta| & \text{if } \Sigma_i \text{ is of type } \xi = \eta \cdot \zeta \\ |x^i_\xi| = |k| & \text{if } \Sigma_i \text{ is of type } \xi = k, \end{cases}$$

for $i = 1, \ldots, n$, where we are assuming that any constant k has the standard von Neumann recursive representation $\{0, 1, 2, \ldots\}$.

Finally, we put

$$\varphi_{n+1} \equiv_{Def} \&_\xi \; \mathit{Finite}\,(x^1_\xi).$$

Let $\varphi_\Sigma \equiv_{Def} \&_{i=0}^{n+1} \varphi_i$. Then it is plain that the system Σ has an integral solution if and only if the formula φ_Σ is satisfiable. Thus, the collection of formulae in the language of φ_Σ has an unsolvable satisfiability problem. It is not hard to see that φ_Σ is expressible in the language MLS extended by the cartesian product $\times$ and the cardinality comparison predicate only, since both the singleton operator and the predicate *Finite*(x) are expressible in this language. Therefore we have:

THEOREM 5.1. ([14]) *The satisfiability problem for MLS extended by the cartesian product and cardinality comparison is unsolvable.* ■

(Note in this connection that it is not known whether the theory MLS plus $\times$ is decidable).

By using $(\forall\exists)_0$-formulae, one can easily express all literals of MLSS plus $\times$. Moreover, any literal of the form $|x| \leq |y|$ is equisatisfiable with the formula *injective*(f, x, y) saying that there exists an injective map f form x into y, which can promptly be expressed by $(\forall\exists)_0$-formuale (see [14]).

It remains only to eliminate literals of type *Finite*(x). For this, the most immediate way would be through the formula

$x = \emptyset \vee (z \in x \,\&\, |x \backslash \{z\}| < |x|)$. But it is not clear whether the relation $|x| < |y|$ can be expressed by $(\forall\exists)_0$-formulae. For instance, the formula $(\forall f)$ $(\neg injective(f, x, y))$ has two quantifier alternations. Hence, if one wants to maintain low the complexity of the formulae involved in a set-theoretic formulation of Hilbert's tenth problem, an alternative formalization of literals *Finite*(x) must be devised.

The approach taken in [14] consists in expressing the set $\omega = \{0, 1, 2, \ldots\}$ and then use the simple formula $|x| \leq |z| \,\&\, z \in \omega$ to express the predicate *Finite*(x). Using $(\forall\exists)_0$-formulae, the set ω can be characterized as the unique infinite limit ordinal that can be partitioned into two sets Z_0, Z_1 such that there are maps $f : Z_0 \rightarrow Z_1$, $g : Z_1 \rightarrow Z_0$ for which

- $s \in f(s)$, for all $s \in Z_0$,
- $t \in g(t)$, for all $t \in Z_1$,
- $f[Z_0] = Z_1 \backslash \{\emptyset\}$, and
- $g[Z_1] = Z_0 \backslash \{\emptyset\}$.

It turns out that all such conditions can be formalized by $(\forall\exists)_0$-formulae, thereby showing that the class of $(\forall\exists)_0$-formulae has an unsolvable satisfiability problem.

In view of the above undecidability results, the decision problem for the class of $(\forall)_0$-formulae assumes greater importance, since its solution would settle the decision problem for the hierarchy of restricted set-theoretic formulae based on quantifier alternations.

A Appendix.

The results discussed in the preceding sections can be formalized in Zermelo-Fraenkel set theory, whose axioms are listed below (cf. [41]).

I. *Axiom of Extensionality.* If X and Y have the same elements, then $X = Y$.

II. *Axiom of Pairing.* For any a and b there exists a set $\{a, b\}$ that contains exactly a and b. (If $a = b$, then the set $\{a\}$ is called a singleton).

III. *Axiom Schema of Separation.* If φ is a predicate (with parameter p), then for any X and p there exists a set $Y = \{u \in X : \varphi(u, p)\}$ that contains all the $u \in X$ that have the property φ.

IV. *Axiom of Union.* For any X there exists a set $Y = Un(X)$, the union of all elements of X ($Un(X) = \{z : z \in w$ for some $w \in X\}$).

V. *Power Set Axiom.* For any X there exists a set $Y = pow(X)$, the set of all subsets of X ($pow(X) = \{Z : Z \subseteq X\}$).

VI. *Axiom of Infinity.* There exists an infinite set.

VII. *Axiom Scheme of Replacement.* If F is a function, then for every X there exists a set $Y = F[X] = \{F(x) : x \in X\}$.

VIII. *Axiom of Regularity.* Every nonempty set has an $\in$-minimal element. (In particular chains of the type $x_0 \in x_1 \in \ldots \in x_n \in x_0$ are not allowed).

IX. *Axiom of Choice.* Every family of nonempty sets has a choice function.

By using the aximos of separation and powerset, one can infer the existence of the cartesian product $X \times Y$ of two sets, X, Y, defined by

$$X \times Y = \{z \in pow(pow(X \cup Y)) : (\exists x \in X)(\exists y \in Y)(z = \{\{x\}, \{x,y\}\})\},$$

where the ordered pairs $[x,y]$ have been encoded following Kuratowski, i.e. $[x,y] = \{\{x\}, \{x,y\}\})\}$.

A.1. Ordinals and Cardinals.

A set T is *transitive* if $T \subseteq pow(T)$. A set is an *ordinal number* if it is transitive and well ordered by $\in$.

A few properties of ordinals are:

(a) $0 = \emptyset$ is an ordinal.

(b) If α is an ordinal and $\beta \in \alpha$, then β is an ordinal too.

(c) If α, β are ordinals and $\alpha \subset \beta$, then $\alpha \in \beta$.

(d) If α, β are ordinals, then either $\alpha \subseteq \beta$ or $\beta \subset \alpha$.

If $\alpha = \beta + 1$, then α is a *successor ordinal*, otherwise α is a *limit ordinal.*

An ordinal α is called a *cardinal number* if $|\alpha| \neq |\beta|$ for all $\beta < \alpha$, i.e. if the ordinal α is not equipollent to any smaller ordinal β. (We remind that $|X|$ denotes the class of all sets which can be put in a one-to-one correspondence with X).

The arithmetic of finite cardinal numbers is equivalent to the well known arithmetic of natural integers. Instead, if α and/or β are infinite cardinals, then $\alpha + \beta = \alpha \cdot \beta = \max(\alpha, \beta)$.

A.2. Transinite Induction and the von Neumann Universe.

THEOREM A.1. (Transfinite induction) *Let C be a class of ordinals and assume that*

(i) $O \in C$,

(ii) if $\alpha \in C$ *then* $\alpha + 1 \in C$,

(iii) if α *is a nonzero limit ordinal and* $\beta \in C$ *for all* $\beta < \alpha$, *then* $\alpha \in C$.

Then C *is the class of all ordinals* Ord.

Using transfinite induction we put

$$\mathcal{V}_0 = \emptyset$$

$$\mathcal{V}_{\alpha+1} = pow(\mathcal{V}_\alpha)$$

$$\mathcal{V}_\alpha = \bigcup_{\beta<\alpha} \mathcal{V}_\beta \text{ if } \alpha \text{ is a limit ordinal.}$$

Then the von Neumann universe is the class $\mathcal{V} = \bigcup_{\alpha \in \mathrm{Ord}} \mathcal{V}_\alpha$.

The axiom of regularity implies that every set x is in some $\mathcal{V}_\alpha$. Then we may define the *rank* of x:

$$rank(x) = \text{least } \alpha \text{ such that } x \in V_{\alpha+1}.$$

The following facts hold:

(a) $\mathcal{V}_\alpha = \{x : rank(x) < \alpha\}$.

(b) If $x \in y$, then $rank(x) < rank(y)$.

(c) $rank(\alpha) = \alpha$ if and only if α is an ordinal.

(d) $rank(x) = \sup\{rank(z) + 1 : z \in x\}$ (and therefore in particular $rank(\{x\}) = rank(x) + 1$).

(e) If $x \subseteq y$, then $rank(x) \leq rank(y)$.

REFERENCES

[1] Andrews P.B., *Theorem proving via general matings,* Journal of A.C.M., (1981) 193-214.

[2] Behmann H., *Beiträge zur Algebra der Logik insbesondere zum Entscheidungs Problem,* Math. Annalen, **86** (1922) 163-220.

[3] Breban M., Ferro A., *Decision procedures for elementary sublanguages of set theory. III. Restricted classes of formulas involving the powerset operator and the general set union operator,* Advances in Applied Mathematics, **5** (1984), 147-215.

[4] Breban M., Ferro A., Omodeo E.G., Schawrtz J.T., *Decision procedures for elementary sublanguages of set theory. II. Formulas involving restricted quantifiers, together with ordinal, integer, map, and domain notions,* Comm. Pure Appl. Math., **34** (1981) 177-195.

[5] Bledsoe W.W., *Splitting and reduction heuristics in automatic theorem proving,* Art. Int., **2** (1971), 55-77.

[6] Cantone D., *A decision procedure for a class of unquantified formulae of set theory involving the powerset and singleton operators,* PhD thesis, New York University - GSAS, Courant Institute of Mathematical Sciences, 251 Mercer St., NYC, NY, 10012, February 1987. Ph. D. thesis.

[7] Cantone D., *A survey of some decidable fragments of set theory,* Technical Report 38, University of Udine, 1988.

[8] Cantone D., *Decision procedures for elementary sublanguages of set theory. X. Multilevel syllogistic extended by the singleton and powerset operators,* Technical report 373, New York University, 1988. To appear in J. Automated Reasoning.

[9] Cantone D., Cutello V., *Some decidability results on quantified sublanguages of set theory,* Technical Report 378, New York University, 1988.

[10] Cantone D., Cutello V., *A decision procedure for set-theoretic formulae involving rank and cardinality comparison,* In Proceedings of the 3rd Italian Conference on Theoretical Computer Science, Mantova, 1989 (A. Bertoni, C. Böhm, P. Miglioli, Eds.), World Scientific, 150-163.

[11] Cantone D., Cutello V., *Decision procedures for elementary sublanguages of set theory. XV. Multilevel syllogistic extended by the predicate Finite and the operators singleton and* $pred_<$. To appear in J. Automated Reasoning.

[12] Cantone D., Cutello V., *Decision procedures for elementary sublanguages of Set Theory XVI. Multilevel syllogistic extended by singleton, rank comparison, and unary intersection,* Bulletiin of EATCS **39** (1989), 139-148.

[13] Cantone D., Cutello V., Ferro A., *Decision procedures for elementary sublanguages of set theory. XIV. Three languages involving rank related constructs,* In Proceedings of Int. Symposium on Symbolic and Algebraic Computation, Rome, Italy. Volume 358 of Lecture Notes in Computer Science, pages 407-422. Springer-Verlag, 1989.

[14] Cantone D., Cutello V., Policriti A., *Set-theoretic reductions of Hilbert's tenth problem,* Technical Report 449, New York University, 1989.

To appear in Proceedings of Computer Science Logic Workshop, Kaiserslautern 1989.

[15] Cantone D., Ferro A., Micale A., Sorace G., *Decision procedures for elementary sublanguages of set theory. IV. Formulae involving a rank operator or one occurrence of* $\Sigma(x) = \{\{y\}|y \in x\}$, Comm. Pure Appl. Math., **40** (1987), 37-77.

[16] Cantone D., Ferro A., Omodeo E.G., *Decision procedures for elementary sublanguages of set theory. VIII. A semidecision procedure for finite satisfiability of unquantified set-theoretic formulae,* Comm. Pure Appl. Math., **41** (1988), 105-120.

[17] Cantone D., Ferro A., Omodeo E.G., *Computable set theory,* Oxford University Press, Oxford, 1989.

[18] Cantone D., Ferro A., Omodeo E.G., Policriti A., *Sillogistica con operatore di scelta,* Technical Report 34, University of Udine, 1988.

[19] Cantone D., Ferro A., Omodeo E.G., Policriti A., *Scomposizione sillogistica disgiuntiva,* In P. Mello, editor, GULP - 89, Atti del quarto convegno nazionale sulla programmazione logica, (1989), 199-209.

[20] Cantone D., Ferro A., Omodeo E.G., Schwartz J.T., *Decision algorithms for some fragments of analysis and related areas,* Comm. Pure Appl. Math., **40** (1987), 281-300.

[21] Cantone D., Ferro A., Schwartz J.T., *Decision procedures for elementary sublanguages of set theory. VI. Multi-level syllogistic extended by the powerset operator,* Comm. Pure Appl. Math., **38** (1985), 549-571.

[22] Cantone D., Ferro A., Schwartz J.T., *Decision procedures for elementary sublanguages of set theory,* V. Multi-level syllogistic extended by the general union operator. J. Computer and System Sciences, **34** (1) (1987), 1-18.

[23] Cantone D., Ghelfo S., Omodeo E.G., *The automation of syllogistic. I. Syllogistic normal forms,* J. Symb. Comp., **6** (1) (1988), 83-98.

[24] Church A., *An unsolvable problem of elementary number theory,* American Journal of Mathematics, **58** (1936) 345-363.

[25] Cantone D., Omodeo E.G., *Topological syllogistic with continuous and closed functions,* Comm. Pure Appl. Math. **42** (1989), 1175-1188.

[26] Cantone D., Omodeo E.G., *On the decidability of formulae involving continuous and closed functions,* In Proceedings of the 11th Int. Joint Conf; on Art. Int., Detroit, Michigan USA, (1989), 425-430.

[27] Cantone D., Omodeo E.G., Policriti A., *The automation of syllogistic. II. optimization and complexity issues,* To appear in J. Automated Reasoning, (1989).

[28] Cantone D., Schwartz J.T., *A set-theoretically based proof verifier,* (1988). Manuscript.

[29] Cantone D., Schwartz J.T., *Decision procedures for elementary sublanguages of set theory, XI. Multilevel syllogistic extended by some elementary map constructs,* Technical Report 374, New York University, 1988. To appear in J. Automated Reasoning.

[30] Ehrenfeucht A., *An application of games to the completeness problem for formalized theories,* Fund. Math., XLIX:129-141, 1961.

[31] Enderton H.B., *A mathematical introduction to logic,* Academic Press, New York and London, 1972.

[32] Ferro A., *A note on the decidability of MLS extended by the powerset operator,* Comm. Pure Appl. Math., **38** (1985) 367-374.

[33] Ferro A., *Decision procedures for elementary sublanguages of set theory. XII. Multilevel syllogistic extended with singleton and choice operators,* To appear in J. Automated Reasoning.

[34] Ferro A., Omodeo E.G., *An efficient validity test for formulae in extensional two-level syllogistic,* Le Matematiche (Cantania, Italy), **33** (1978), 130-137.

[35] Ferro A., Omodeo E.G., *Decision procedures for elementary sublanguages of set theory. VII. Validity in set theory when a choice operator is present,* Comm. Pure Appl. Math., **40** (1987) 265-280.

[36] Ferro A., Omodeo E.G., Schwartz J.T., *Decision procedures for elementary sublanguages of set theory. I. Multi-level syllogistic and some extensions,* Comm. Pure Appl. Math., **33** (1980), 599-608.

[37] Ferro A., Omodeo E.G., Schwartz J.T., *Decision procedures for some fragments of set theory,* pages 88-96. Volume 87 of Lecture Notes in Computer Science, Springer-Verlag, 1980.

[38] Ghelfo S., Omodeo E.G., *Towards practical implementations of syllogistic,* In EUROCAL '85, European Conference on Computer Algebra - Proceedings Vol. 2, pages 40-48, Springer-Verlag, 1985.

[39] Gogol D., *The $\forall_n\exists$-completeness of Zermolo-Fraenkel set theory,* Zeitschr. f. math. Logik und grundlagen d. Math., 1978.

[40] Gogol D., *Sentences with three quantifiers are decidable in set theory,* Fundamenta Mathematicae, 102, (1979).

[41] Jech T.J., *Set theory,* Academic press, 1978.

[42] Kuratowski K., *Sur l'opération $\bar{A}$ de l'Analysis Situs,* Fund. Math., **3** (1922), 182-199.

[43] Matijasevič Y., *Enumerable sets are Diophantine sets,* Soviet Math. Doklady **11** (1970) 354-357.

[44] Moser L.E., *A decision procedures for unquantified formulas of graph theory,* pages 344-357. Volume 310 of Lecture Notes in Computer Science, Springer-Verlag, 1988.

[45] Moser L.E., *Decidability of formulas in graph theory,* Fundamenta Informaticae, **12** (2) (1989), 163-180.

[46] Omodeo E.G., *Decidability and proof procedures for set theory with a choice operator,* PhD Thesis, New York University, 1984.

[47] Peter R., *Recursive functions,* Academic Press, New York, 1967.

[48] Parlamento F., Policriti A., *Decision procedures for elementary sublanguages of set theory. IX. Unsolvability of the decision problem for a restricted subclass of the Δ_0-formulas in set theory,* Comm. Pure Appl. Math., **41** (1988), 221-251.

[49] Parlamento F., Policriti A., *Decision procedures for elementary sublanguages of set theory. XIII. Model graphs, reflection and decidability,* To appear in J. Automated Reasoning.

[50] Presburger M., *Über die Vollständigkeit eines gewissen Systems der Arithmetic ganzer Zahlen, in welchem die addition als einzige Operation hervortritt,* In Comptes-rendus du Premier Congrès des Mathematiciens des Pays Slaves, pages 192-201, 395, Warsaw, (1929).

[51] Schwartz J.T., *Instatiation and decision procedures for certain classes of quantified set-theoretic formulae,* Technical report 78-10, Institute for Computer Applications in Science and Engineering, NASA Langley Reseach Center, Hampton, Virginia, 1978.

[52] Schwartz J.T., *A survey of program proof technology,* Technical Report 1, New York University, 1978.

[53] Tarski A., *Ordinal algebras,* North-Holland, New York, 1956.

[54] Vaught R.L, *Set theory. An introduction,* Birkhauser, Boston, 1985.

[55] Ville F., *Décidabilité des formules existentielles en théorie des ensembles,* C.R. Acad. Sc., t. 272, Série A: (1971) 513-516, Paris.

[56] Wos L., *Automated reasoning. 33 basic research problems,* Prentice Hall, Englewood Cliff, New Jersey, 1988.

Courant Institute
of Mathematical Sciences,
New York University,
251 Mercer Street
New York, N.Y. 10012

Università di Catania
Dipartimento di Matematica
Viale A. Doria, 6
95125 Catania