Decidability of E*A-sentences in Membership Theories

Published
DOI:10.1002/malq.19960420105

Math. Log. Quart. 42 (1996) 41 - 58

Mathematical Logic
Quarterly

Decidability of $\exists^*\forall$-Sentences in Membership Theories[1)]

Eugenio G. Omodeo[a], Franco Parlamento[b], and Alberto Policriti[b]

[a] Department of Computer Science and Applications "R. M. Capocelli", University of Salerno, 84081 Baronissi, Italy[2)]

[b] Department of Mathematics and Computer Science, University of Udine, via Zanon 6, 33100 Udine, Italy[3)]

Abstract. The problem is addressed of establishing the satisfiability of prenex formulas involving a single universal quantifier, in diversified axiomatic set theories. A rather general decision method for solving this problem is illustrated through the treatment of membership theories of increasing strength, ending with a subtheory of Zermelo-Fraenkel which is already complete with respect to the $\exists^*\forall$ class of sentences. NP-hardness and NP-completeness results concerning the problems under study are achieved and a technique for restricting the universal quantifier is presented.

Mathematics Subject Classification: 03B25, 03E30.

Keywords: Decidability, Fragment of set theory, NP-complete, NP-hard.

1 Introduction

Investigations into the decision problem for "small axiomatic fragments of set theory", to use TARSKI's wording in [28], date back, at least, to [27] (see also the ensuing [8]), which stated the interpretability of ROBINSON's Arithmetic Q into the theory of sets consisting of the following axioms:

N	$(\forall x)(x \notin \emptyset)$	(*Null-set Axiom*),
W	$(\forall x)(\forall y)(\forall z)(x \in y \,\mathrm{w}\, z \leftrightarrow x \in y \vee x = z)$	(*With Axiom*),
E	$(\forall x)(\forall y)((\forall z)(z \in x \leftrightarrow z \in y) \rightarrow x = y)$	(*Extensionality Axiom*),

Weak membership theories like NWE (we will systematically denote a theory by naming its axioms) and decidable fragments of set theory as a whole, have recently attracted attention from the foundational point of view [18] as well as of the computer science community [6, 19, 11, 21]. Under the stimulus of foundational issues the interpretability of Q into NWE has been improved to obtain the interpretability of Q into NW (see [18]). The above-mentioned results imply the essential undecidability of the theories NWE and NW (cf. also [31]), but are of little or no help to assess precisely

1) This work has been supported by funds 40% and 60% MURST.

2) e-mail: omodeo@udsab.dia.unisa.it, omodeo@assi.dis.uniroma1.it

3) e-mail: {policrit,parlamen}@dimi.uniud.it

to which classes of sentences, classified according to syntactic criteria (e. g. quantificational prefix), the undecidability of such theories refers. Preliminary investigations in this sense have been carried out in [22, 5, 24, 3]. The objective of developing a theorem prover embodying set-theoretic notions, as well as efforts in the automated programming area (cf. [20]) or aimed at obtaining extensions of logic programming with sets (cf. [9, 10, 16]), have spurred us to investigate the theory NWL where the new axiom is

L $(\forall x)(\forall y)(\forall z)(z \in x \mathbin{l} y \leftrightarrow z \in x \wedge z \neq y)$ (*Less Axiom*).

In fact NWL is a "denominator" common to almost all of the theories for which decidability, unifiability, and constraint simplification results have been discovered to date.

This paper introduces a decision method which is then uniformly employed to solve the satisfiability problem for prenex formulas with one universal quantifier, in the language containing $=$ and $\in$, with respect to the theories NWL, NWLE, NWLR, and NWLER. Here the new axiom is

R $(\forall x)(x \neq \emptyset \rightarrow (\exists y)(y \in x \wedge (\forall z)(z \in y \rightarrow z \notin x)))$ (*Regularity Axiom*).

Furthermore it is shown that NWLER is complete with respect to existential closures of such formulas (i. e. $\exists^*\forall$-sentences); a completeness result which remarkably improves the analogous result relative to ZFC was established in [14]. Thus, since $\forall^*\exists$-sentences are dual to $\exists^*\forall$-sentences, the undecidability of $\forall^*\exists$ in pure quantificational logic (see [17]) vanishes when one comes to the theory NWLER.

As for the complexity of the proposed problems, it is shown in the last section that they are NP-hard w.r.t. all theories, even if strong restrictions are imposed on the syntactic form of the unquantified matrix. In particular, NP-hardness is proved also in the case in which the matrix is a conjunction of doubleton disjunctive clauses (2CNF). In the case of theories containing the Regularity Axiom, one proves NP-completeness through a reduction of $\exists^*\forall$-sentences to alike sentences where the universal quantifier is restricted, and by resorting then to an *ad hoc* satisfiability decision test drawn from [7].

The general decision method proposed here, as well as the technique for restricting the $\forall$-quantifier, has been successfully instantiated to a non-classical "kernel" set theory in [21] which systematically injures regularity as proposed in [1].

2 $\forall$-formulas

Definition 2.1. A *$\forall$-formula* is a prenex formula with just one universal quantifier whose matrix contains only $=$ and $\in$.

Under axiom N we can allow also $\emptyset$ to appear in the matrix of a $\forall$-formula, since any occurrence of $\emptyset$ can easily be eliminated. Any theory T extending NW proves the inequality of the numerals $\underline{n}$, inductively defined as follows: $\underline{0}$ is $\emptyset$, $\underline{n+1}$ is $\underline{n} \mathbin{w} \underline{n}$, i.e. for $m < n$, $\underline{n} \neq \underline{m}$, as is easily seen by induction. Thus every model of such theories is infinite. From that it follows that the satisfiability problem with respect to extensions of NW for the $\forall$-formulas is reducible to the problem of establishing whether they are satisfiable with distinct elements, all different from the interpretation of $\emptyset$, interpreting different variables; a problem which we will designate as the $1-1\neq\emptyset$*-satisfiability problem.*

Proposition 2.1. *For extensions of the theory* NW, *the satisfiability problem for $\forall$-formulas with matrix in $=$ and $\in$ is reducible to the $1-1\neq\emptyset$-satisfiability problem for $\forall$-formulas with matrix in $=$, $\in$ and $\emptyset$.*

In view of this proposition, the rest of the paper will deal almost entirely with the decision problem for $1-1\neq\emptyset$-satisfiability of $\forall$-formulas with matrix in $=$, $\in$ and $\emptyset$.

3 The decidability of $\forall$-formulas in NWL

Proposition 3.1. *If $\psi(x_1,\dots,x_n)$ is a $\forall$-formula, then $\psi(x_1,\dots,x_n)$ is $1-1\neq\emptyset$-satisfiable with respect to the theory* NWL *iff there exists a finite structure $\mathcal{G}=\langle G,c_0,c_1,\dots,c_n,R\rangle$, with $c_0,c_1,\dots,c_n$ in G and $R\subseteq G\times G$, such that*

1. *$c_i\neq c_j$ for all i,j with $0\leq i<j\leq n$,*
2. *c_0 has no R-predecessor in G,*
3. *for all $J\subseteq\{c_0,c_1,\dots,c_n\}$ and for all g $(0\leq g\leq n)$, there exists an element $k_{J,g}\in G\setminus\{c_0,c_1,\dots,c_n\}$ such that:*
   (a) *$\{c_i : c_i\,R\,k_{J,g}\}=J$,*
   (b) *$k_{J,g}\,R\,c_g$ if and only if $k_{J,g}\,R\,k_{J,g}$,*
   (c) *$\mathcal{G}\models\psi(c_1,\dots,c_n)$.*

Proof. ($\Leftarrow$). Assume $\psi(x_1,\dots,x_n)$ is $1-1\neq\emptyset$ satisfied in a model $\mathcal{M}$ of NWL by the n-tuple $c_1,\dots,c_n$ of elements of the support of $\mathcal{M}$ that we will denote by M. Let $c_0=\emptyset^M$. Obviously c_0 has no $\in^M$ predecessors in $\mathcal{M}$. Since $\mathcal{M}\models$ NW, there are natural numbers $h_0,\dots,h_n,h_{n+1}$ such that for $i=0,\dots,n$, $\underline{h_i^M}\neq c_i$ and $(\underline{h_{n+1}^M}\notin^M\underline{h_i^M})$. Given any $J\subseteq\{c_0,c_1\dots,c_n\}$ and any g, $0\leq g\leq n$, let $k_{J,g}$ be

$$(((c_g\,\mathbf{w}^M\underline{h_{n+1}^M}\cup J)\setminus(\{c_0,\dots,c_n\}\setminus J))\cup\{\underline{h_i^M}:\underline{h_i^M}\notin^M c_i\})\setminus\{\underline{h_i^M}:\underline{h_i^M}\in^M c_i\},$$

where we have used self explaining notations (for the operations of union and set-difference) as shorthand for iterated uses of the operations $\mathbf{w}^M$ and $\mathbf{l}^M$. Then it suffices to let

$$G=\{c_o,c_1,\dots,c_n\}\cup\{k_{J,g}:J\subseteq\{c_0,c_1,\dots,c_n\},\ 0\leq g\leq n\}$$

and $R\equiv\in^M\restriction G$, to obtain a finite structure fulfilling conditions 1. – 3.

Conditions 1 and 2 are obvious. For condition 3 notice that $\underline{h_i^M}\in^M k_{J,g}$ if and only if $\underline{h_i^M}\notin^M c_i$, and hence $k_{J,g}\notin\{c_0,\dots,c_n\}$. From $\underline{h_i^M}\neq c_i$ it follows that once all the elements in J have been added to c_g and all the elements in $\{c_0,c_1,\dots,c_n\}\setminus J$ have been removed, $k_{J,g}$ is possibly modified only by the insertion or removal of elements different from $c_0,\dots,c_n$, thus condition 3(a) holds. Since $\underline{h_{n+1}^M}$ is placed in $k_{J,g}$ to start with, and it is not subsequently removed, $\underline{h_{n+1}^M}\in^M k_{J,g}$. Therefore, since for all i, $0\leq i\leq n$, $(\underline{h_{n+1}^M}\notin^M\underline{h_i^M})$, we have that $k_{J,g}\neq\underline{h_i^M}$ for $0\leq i\leq n$. Hence if $k_{J,g}\in^M c_g$, then $k_{J,g}\in^M k_{J,g}$, since the elements which are removed from c_g to obtain $k_{J,g}$ are different from $k_{J,g}$. Similarly the only way $k_{J,g}$ can possibly be an $\in^M$ element of $k_{J,g}$ is to be an $\in^M$ element of c_g to start with. Thus condition 3(b) is also satisfied. Obviously $\mathcal{G} \models \psi(c_1, \ldots, c_n)$, since ψ is a universal formula.

($\Rightarrow$). Given $\mathcal{G}$, fulfilling the stated conditions, we show how to build a Herbrand model on $c_0, c_1, \ldots, c_n, \mathbf{w}, \mathbf{l}$ in which $c_1, \ldots, c_n$ satisfy ψ. Let $t_0, t_1, \ldots, t_n, \ldots, t_{i+1}, \ldots$ be an enumeration without repetitions of all terms in the Herbrand universe H such that t_0 is $\emptyset$, $t_1, \ldots, t_n$ are $c_1, \ldots, c_n$, respectively, and for any $i > n$ the term t_i is of the form $f(t_h, t_k)$ with $h, k < i$ (and f is either $\mathbf{w}$ or $\mathbf{l}$). We call *seed of t* the term inductively defined as follows:

$$\text{seed}(t) = \begin{cases} t & \text{if } t \text{ is a constant,} \\ \text{seed}(t_1) & \text{if } t = t_1 \,\mathbf{w}\, t_2 \text{ or } t = t_1 \,\mathbf{l}\, t_2. \end{cases}$$

Let $H_i = \{t_j \in H : j \leq n + i\}$. Thus H_0 is simply $\{c_0, c_1, \ldots, c_n\}$. Inductively on i, we define a binary relation R_i on H_i so that $R_j = R_i \restriction H_j$ for all $j < i$, and $\langle H_i, c_0, c_1, \ldots, c_n, R_i \rangle \models \psi(c_1, \ldots, c_n)$. The partial structure $\langle H_i, c_0, \mathbf{w}^i, \mathbf{l}^i, R_i \rangle$, where $\mathbf{w}^i$ and $\mathbf{l}^i$ are the (partial) canonical interpretations of $\mathbf{w}$ and $\mathbf{l}$, will also model NWL to the following extent:

(a) c_0 has no R_i-predecessor in H_i,

(b) if r, s, $r \,\mathbf{w}\, s \in H_i$, then

$$\text{for all } t \in H_i,\ (t \, R_i \, r \,\mathbf{w}\, s) \text{ iff } t \, R_i \, r \text{ or } t \equiv s,$$

(c) if r, s, $r \,\mathbf{l}\, s \in H_i$, then

$$\text{for all } t \in H_i,\ (t \, R_i \, r \,\mathbf{l}\, s) \text{ iff } t \, R_i \, r \text{ and } t \not\equiv s.$$

It is obvious that if the construction of such a sequence of partial structures succeeds, then $\langle H, c_0, \bigcup_{i\in\omega} \mathbf{w}^i, \bigcup_{i\in\omega} \mathbf{l}^i, \bigcup_{i\in\omega} R_i \rangle$ provides the desired model of NWL in which ψ is satisfied by $c_1, \ldots, c_n$.

R_0 is simply $R \restriction \{c_0, c_1, \ldots, c_n\}$ and $\mathbf{w}^0$, $\mathbf{l}^0$ are totally undefined.

Assume we have defined H_i and R_i, and so also $\mathbf{w}^i$ and $\mathbf{l}^i$. If $t_{i+1} \equiv r \,\mathbf{w}\, s$ and $r, s \in H_i$, then we let

$$J = \begin{cases} \{c_i : c_i R_i r\} & \text{if } s \notin \{c_0, c_1, \ldots, c_n\}, \\ \{c_i : c_i R_i r\} \cup \{s\} & \text{otherwise.} \end{cases}$$

Let c_g be the seed of t_{i+1}, and $k_{J,g}$ be an element in G satisfying conditions 1. – 3. R_{i+1} is obtained from R_i by adding the following pairs. For $t \in H_i$: if $t \equiv r \,\mathbf{w}\, s$ we add the pair (t, t_{i+1}) if $t \, R_i \, r$ or $t \equiv s$; if $t \equiv r \,\mathbf{l}\, s$ we add the pair (t, t_{i+1}) if $t \, R_i \, r$ and $t \not\equiv s$. Furthermore for $t \in H_{i+1}$ we add the pair (t_{i+1}, t) if $k_{J,g} \, R_i \, \text{seed}(t)$. It is easy to see that conditions (i) – (iii) are maintained. □

4 The decidability of $\forall$-formulas in NWLE

When the Extensionality Axiom is added to NWL it is obvious that the conditions the finite structure $\mathcal{G}$ must satisfy includes also the extensionality of R on $\{c_0, c_1, \ldots, c_n\}$, i.e., for $0 \leq i < j \leq n$ there must be $k \in G$ such that $k \, R \, c_i$ if and only if $k \not R \, c_j$.

The construction of a model of NWLE satisfying ψ is however much more involved. To obtain a model of NWLE once we have obtained a model of NWL according to Proposition 3.1, it suffices to make a quotient with respect to the minimal co-bisimulation of the model; however the elements satisfying ψ in the original structure may very well give rise to equivalence classes which do not satisfy ψ in the quotient structure. In the following proposition we will show that nevertheless it is possible to obtain a model of NWLE satisfying ψ as a quotient of the Herbrand universe, with respect to relations and operations which need to be defined inductively.

Proposition 4.1. *If $\psi(x_1,\ldots,x_n)$ is a $\forall$-formula, then $\psi(x_1,\ldots,x_n)$ is $1-1\neq\emptyset$-satisfiable with respect to the theory* NWLE *iff there exists a finite structure $\mathcal{G}=\langle G,c_0,c_1,\ldots,c_n,R\rangle$ such that*

1. *$c_i\neq c_j$ for all i,j with $0\leq i<j\leq n$,*
2. *c_0 has no R-predecessor in G,*
3. *for all $J\subseteq\{c_0,c_1,\ldots,c_n\}$ and for all g $(0\leq g\leq n)$, there exists an element $k_{J,g}\in G\setminus\{c_0,c_1,\ldots,c_n\}$ such that:*
   (a) $\{c_i : c_i\, R\, k_{J,g}\}=J$,
   (b) *$k_{J,g}\, R\, c_g$ if and only if $k_{J,g}\, R\, k_{J,g}$,*
   (c) $\mathcal{G}\models\psi(c_1,\ldots,c_n)$,
   (d) *R is extensional in G over $\{c_o,c_1,\ldots,c_n\}$.*

Proof. $(\Leftarrow)$. The proof is the same as for Proposition 3.1, with the addition to G of elements taken from the given model that witness the differences between the elements in $\{c_0,c_1,\ldots,c_n\}$. By [25] at most n elements need to be added to fulfill that requirement.

$(\Rightarrow)$. Given $\mathcal{G}$, fulfilling the stated conditions, let H and H_i be defined as in the proof of Proposition 3.1. Inductively on i, we will define below a partition $[\ ]_i : H_i \longrightarrow \mathrm{Pow}(H_i)$ of H_i, a binary relation R_i on $M_i=\{[t]_i : t\in H_i\}$, and two partial binary operations w^i and l^i on M_i, so that the following conditions are met:

(i) $\langle M_i,R_i\rangle\models\psi([c_1]_i,\ldots,[c_n]_i)$.

(ii) Let $\mathrm{Pred}_i : H_{i+1}\longrightarrow\mathrm{Pow}(M_i)$ be defined as follows:

$$\mathrm{Pred}_i(t)=\begin{cases}\{[s]_i : s\in H_i \text{ and } [s]_i\, R_i\, [t]_i\} & \text{if } t\in\{c_0,c_1,\ldots,c_n\},\\ \{[s]_i : s\in H_i \text{ and } [s]_i\, R_i\, [t']_i \text{ or } [s]_i=[t'']_i\} & \text{if } t=t'\,\mathbf{w}\,t'',\\ \{[s]_i : s\in H_i \text{ and } [s]_i\, R_i\, [t']_i \text{ and } [s]_i\neq[t'']_i\} & \text{if } t=t'\,\mathbf{l}\,t''.\end{cases}$$

Let, moreover, $\sim_i$ be the equivalence relation associated with $[\ ]_i$, that is $r\sim_i s$ iff $[r]_i=[s]_i$. Then the following must hold: for $r,s\in H_i$, $r\sim_i s$ iff $\mathrm{Pred}_i(r)=\mathrm{Pred}_i(s)$ and seed(r) is *trapped* with seed(s), in the sense that seed(r) and seed(s) have the same R-predecessors in $G\setminus\{c_0,c_1,\ldots,c_n\}$.

(iii) $\sim_i$ is a congruence with respect to the canonical interpretations of $\mathbf{w}$ and $\mathbf{l}$ in H_i, and w^i and l^i are the partial operations induced on H_i by them.

(iv) The axioms W and L are modelled in $\langle M_i,R_i,w^i,l^i\rangle$ in the usual weak sense that whenever $[r]_i\, w^i\, [s]_i$ is defined,

$$[t]_i\, R_i\, ([r]_i\, w^i\, [s]_i) \quad\text{iff}\quad [t]_i\, R_i\, [r]_i \text{ or } [t]_i=[s]_i;$$

and whenever $[r]_i\, l^i\, [s]_i$ is defined,

$$[t]_i\, R_i\, ([r]_i\, l^i\, [s]_i) \quad\text{iff}\quad [t]_i\, R_i\, [r]_i \text{ and } [t]_i\neq[s]_i.$$

Our inductive construction proceeds as follows:

Base case: We put $[c_h]_0 = \{c_h\}$, for $0 \leq h \leq n$. Furthermore for $0 \leq h, k \leq n$ we define $[c_h]_0 \, R_0 \, [c_k]_0$ iff $c_h \, R \, c_k$ and leave w^0 and l^0 be totally undefined. Since ψ is a universal formula and it is satisfied by $c_1, \ldots, c_n$ in $\mathcal{M}$, it is obvious that condition (i) holds. Conditions (ii) – (iv) are trivially satisfied.

Induction step: Assuming that $[\,]_i : H_i \longrightarrow \mathrm{Pow}(H_i)$, R_i, w^i and l^i have been defined and satisfy the conditions (i) – (iv), we show how to define the partition $[\,]_{i+1} : H_{i+1} \longrightarrow \mathrm{Pow}(H_{i+1})$, R_{i+1}, w^{i+1} and l^{i+1} so as to maintain conditions (i) – (iv) and to enforce the following:

(a) for all $t \in H_i$, $[t]_i \subseteq [t]_{i+1}$,

(b) $R_i = R_{i+1} \restriction M_i$,

(c) $w^i = w^{i+1} \restriction M_i \times M_i$,

(d) $l^i = l^i \restriction M_i \times M_i$.

To define the new partition on H_{i+1} we rely on the one already given on H_i and, either we enrich one of the classes $[r]_i$ by adding t_{i+1} to it or else we create a new class $[t_{i+1}]_{i+1} = \{t_{i+1}\}$.

The former is the case whenever for some $r \in H_i$ we have

(*) $\mathrm{Pred}_i(t_{i+1}) = \mathrm{Pred}_i(r)$ and $\mathrm{seed}(t_{i+1})$ is trapped with $\mathrm{seed}(r)$.

In that case we let $[r]_{i+1} = [r]_i \cup \{t_{i+1}\}$ for any r such that (*) holds, including t_{i+1} (i.e. $[t_{i+1}] = [r]_{i+1}$), and $[s]_{i+1} = [s]_i$ for any other term s in H_i. Since condition (ii) holds for $[\,]_i$ and R_i, it is clear that in this way all the classes in the partition of H_i are left unchanged except one which is increased by the addition of t_{i+1} to it. Putting, for $r, s \in H_i$,

$$[r]_{i+1} \, R_{i+1} \, [s]_{i+1} \quad \text{iff} \quad [r]_i \, R_i \, [s]_i$$

clearly defines R_{i+1} over M_{i+1}. It is rather straightforward to check that conditions (i) – (iv) are maintained.

If, on the other hand, there is no $r \in H_i$ such that (*) holds, we let

$$[s]_{i+1} = [s]_i \text{ for } s \in H_i, \text{ and } [t_{i+1}]_{i+1} = \{t_{i+1}\}.$$

R_{i+1} is obtained by extending R_i in the following way: Let

$$J_{t_{i+1}} = \{c_m : [c_m]_i \in \mathrm{Pred}_i(t_{i+1})\}$$

and $c_g = \mathrm{seed}(t_{i+1})$. Pick $k_g \in G \setminus \{c_0, c_1, \ldots, c_n\}$ such that

$$\{c_m : c_m \, R \, k_g\} = \begin{cases} J_{t_{i+1}} & \text{if } k_g \not R \, c_g, \\ J_{t_{i+1}} \cup \{k_g\} & \text{otherwise.} \end{cases}$$

Then put, for $r \in H_i$,

$$[r]_{i+1} \, R_{i+1} \, [t_{i+1}]_{i+1} \quad \text{iff} \quad [r]_i \in \mathrm{Pred}_i(t_{i+1}),$$

for $c_l \in \{c_0, c_1, \ldots, c_n\}$,

$$[t_{i+1}]_{i+1} \, R_{i+1} \, [c_l]_{i+1} \quad \text{iff} \quad k_g \, R \, c_l,$$

and for $t \in H_{i+1} \setminus \{c_0, c_1, \ldots, c_n\}$,

$$[t_{i+1}]_{i+1} \, R_{i+1} \, [t]_{i+1} \quad \text{iff} \quad k_g \, R \, \text{seed}(t).$$

Note that the last clause in the definition of R_{i+1} is independent of the representative t of the class $[t]_{i+1}$, since by induction hypothesis all terms in $[t]_{i+1}$ have seeds trapped with seed(t). In order to prove condition (i) let us first notice that $\langle M_i, R_i \rangle \models \psi([c_1]_i, \ldots, [c_n]_i)$ and $\mathcal{G} \models \psi(c_1, \ldots, c_n)$. Moreover notice that $[t_{i+1}]_{i+1}$ has, with respect to R_{i+1} and $[c_1]_{i+1}, \ldots, [c_n]_{i+1}$, the same relationship as k_g with respect to R and $c_1, \ldots, c_n$, from which it is clear that condition (i) holds for M_{i+1}, R_{i+1} as well, namely $\langle M_{i+1}, R_{i+1} \rangle \models \psi([c_1]_{i+1} \ldots, [c_n]_{i+1})$. It is also straightforward to check that the remaining conditions (ii) - (iv) are maintained. This ends our inductive construction.

Since conditions (i) - (iv) as well as conditions 1. - 3. in the statement of the proposition are maintained throughout the construction, we can build an interpretation of $\in$, $\emptyset$, w, l as follows: For all $t \in H$, let $[t] = \bigcup_{j \geq i} [t]_j$, where i is such that $t \equiv t_i$; moreover let $M = \{[t] : t \in H\}$ and $\emptyset^M = [c_0]$. Given $[t], [r] \in M$, let j be large enough so that $[t]_j, [r]_j \in M$ are defined; then we let: $[t] \in^M [r]$ iff $[t]_j \, R_j \, [r]_j$. Finally we let $[t] \, w^M \, [r] = [t \, \mathrm{w} \, r]$ and $[t] \, l^M \, [r] = [t \, \mathrm{l} \, r]$.

It is quite straightforward to see that $\mathcal{M} = \langle M, \in^M, \emptyset^M, w^M, l^M \rangle \models$ NWL, and furthermore that $\mathcal{M} \models \psi([c_1], \ldots, [c_n])$. But one is not guaranteed that $\mathcal{M} \models$ E.

To complete the proof, we just are to refine the above inductive construction so as to model the extensionality axiom. Our construction prescribes that two terms take different values when their seeds are not trapped, even if at the stage in which they are treated they have the same set of predecessors. The way to ensure that in the final model $\mathcal{M}$ they will have different $\in^M$-members is to force, as our construction goes on, an infinite difference between their seeds. This can be obtained by specializing the choice of the element k_g used as template in defining R_{i+1} when there is no $j \leq i$ such that $\text{Pred}_i(t_{i+1}) = \text{Pred}_j(t_j)$ and seed(t_{i+1}) is trapped with seed(t_j).

Let $\{k_1, \ldots, k_d\} \subseteq G \setminus \{c_0, \ldots, c_n\}$ be a differentiating set for the relation "c_i is trapped with c_j", namely if c_i is not trapped with c_j there is $1 \leq h \leq d$ such that $k_h \, R \, c_i$ if and only if $k_h \not R \, c_j$. By results in [25], d can be taken to be smaller than the number of equivalence classes induced on $\{c_0, \ldots, c_n\}$ by the relation *c_i is trapped with c_j*. As is is easy to see, the numerals give rise to different classes as our construction goes on; that is, if $m \neq m'$, then $[\underline{m}] \neq [\underline{m'}]$ and the $\in^M$-predecessors of $[\underline{m}]$ are precisely $[\underline{0}], [\underline{1}], \ldots, [\underline{m-1}]$. Furthermore for $m > |G|$ and $0 \leq i \leq n$, $[\underline{m}] \neq [c_i]$. To see this note that if there is $k \in G \setminus \{c_0, \ldots, c_n\}$ such that $k \, R \, c_i$, then seed$(\underline{m}) = c_0$ is not trapped with seed$(c_i) = c_i$, therefore $\underline{m}$ is not made equivalent with c_i when it is treated. On the other hand if the R-predecessors of c_i are among $c_0, \ldots, c_n$, then the only $\in^M$-members of $[c_i]$ are among $[c_0], \ldots, [c_n]$. So there are at most $n+1$ $\in^M$-members of $[c_i]$, while for $m > |G|$, $[\underline{m}]$ has more than $n+1$ $\in^M$-members, thus for any such m's, $[\underline{m}] \neq [c_i]$. It may happen however that $\underline{m}$ is not the first term to be encountered in the enumeration of H that gives rise to the class $[\underline{m}]$. On the other hand it is plain to check, by inspecting the structure $\langle \{c_0, \ldots, c_n\}, R \restriction \{c_0, \ldots, c_n\} \rangle$, whether a term t is the first one in the enumeration which originates the class $[\underline{m}]$. So we may assume that the definition of $\underline{m}$ is such that $\underline{m}$ is actually the first term in the enumeration of H which generates $[\underline{m}]$.

On the ground of the above observations we now build an infinite supply of terms that will differentiate non-trapped c_i's. For $1 \leq j \leq d$ and $p > 0$, let $h_p^j = j + p \cdot |G|$. If $k_j \not R k_j$ we determine a countably infinite sequence of terms d_p^j by letting

$$d_p^j = (\{c_m : c_m \, R \, k_j\} \setminus \{c_m : c_m \not R \, k_j\}) \cup \{\underline{h_p^j}\}.$$

Due to our assumption on the definition of the numerals, d_p^j is not forced to be equivalent to any preceding term and is amenable to be treated using k_j.

At this point, in order to guarantee that at the end of our construction any two non trapped c_i's will differ for infinitely many elements, as an additional feature of our construction we impose that when the term d_p^j is encountered, among the various possibilities to treat d_p^j, the one which uses k_j is chosen. The final outcome of this proviso is that the classes $[d_p^j]$ will turn out to be all distinct, as there will be infinitely many members of each class $[c_i]$ such that $k_j \, R \, c_i$.

The remaining case, namely when $k_j \, R \, k_j$, requires a bit more elaborated determination of the corresponding $[d_p^j]$'s. Let c_i be such that $k_j \, R \, c_i$. When the construction reaches the stage g_p in which $\underline{h_p^j}$ is taken into account, the following term d_p^j is determined:

$$d_p^j = \begin{cases} ((c_i \cup \{c_m : c_m \, R \, k_j\}) \setminus \{c_m : c_m \not R \, k_j\}) \setminus \{\underline{h_p^j}\} & \text{if } [\underline{h_p^j}]_{g_p} R_{g_p} [c_i]_{g_p}, \\ ((c_i \cup \{c_m : c_m \, R \, k_j\}) \setminus \{c_m : c_m \not R \, k_j\}) \cup \{\underline{h_p^j}\} & \text{otherwise.} \end{cases}$$

The proviso is taken that d_p^j be treated, when its turn comes, by choosing k_j.

We can now show that if this two further requirements concerning the way the terms d_p^j are treated are fulfilled, then the resulting structure $\mathcal{M}$ satisfies also the Extensionality Axiom.

Assume that $[t_h] \neq [t_k]$. Without loss of generality we may also assume that $h < k$.

If $\mathrm{Pred}_k(t_h) \neq \mathrm{Pred}_k(t_k)$, let r be such that $[r]_k \, R_k \, [t_h]_k$ and $[r]_k \not R_k \, [t_k]_k$. It is clear that $[r] \in^M [t_h]$ and $[r] \notin^M [t_k]$; therefore, $[t_h]$ and $[t_k]$ have different sets of $\in^M$-members.

On the other hand if $\mathrm{Pred}_k(t_h) = \mathrm{Pred}_k(t_k)$, then the seeds c_{t_h} and c_{t_k} of t_h and t_k are not trapped, otherwise $[t_h]_k = [t_k]_k$. In this case since $\mathcal{M} \models \mathrm{WL}$ the set of $\in^M$-members of $[t_h]$ and $[t_k]$ have a finite difference with respect to the set of $\in^M$-members of $[c_{t_h}]$ and $[c_{t_k}]$, respectively. Moreover, since c_{t_h} and c_{t_k} are not trapped, there is $k_j \in \{k_1, \ldots, k_d\}$ differentiating them in $\mathcal{G}$. We can assume without loss of generality that $k_j \, R \, c_{t_h}$ and $k_j \not R \, c_{t_k}$. Our construction ensures that $[c_{t_h}]$ receives all the infinitely many classes $[d_p^j]$ as $\in^M$-members and that none of them is an $\in^M$-member of $[c_{t_k}]$. Thus the differences between the sets of the $\in^M$-members of $[c_{t_k}]$ and of $[c_{t_h}]$ is infinite. It follows that the set of $\in^M$-members of $[t_h]$ is different from the set of $\in^M$-predecessors of $[t_k]$.

This argument takes care of all the cases except the one in which t_h and t_k are, say, c_h and c_k and different, but trapped. In this case we resort to the requirement that R is extensional on $\{c_0, c_1, \ldots, c_n\}$: among $c_0, c_1, \ldots, c_n$ there is a c_j such that $c_j \, R \, c_h$ iff $c_i \not R \, c_h$, and from that it follows that $[c_j]$ is an $\in^M$-member of $[c_h]$ iff $[c_j]$ is an $\in^M$-member of $[c_h]$. □

5 The decidability of $\forall$-formulas in NWLR

As opposed to the addition of the Extensionality Axiom to NWL, the addition of the Regularity Axiom makes the proof of the analog of Proposition 3.1 easier, and even a stronger result can be obtained.

Proposition 5.1. *If $\psi(x_1,\ldots,x_n)$ is a $\forall$-formula, then $\psi(x_1,\ldots,x_n)$ is $1-1\neq\emptyset$-satisfiable with respect to the theory* NWLR *iff there exists a finite structure $\mathcal{G}=\langle G,c_0,c_1,\ldots,c_n,R\rangle$ such that*

1. *$c_i\neq c_j$ for all i,j with $0\leq i<j\leq n$,*
2. *c_0 has no R-predecessor in G,*
3. *for all $J\subseteq\{c_0,c_1,\ldots,c_n\}$ there exists an element $k_J\in G\setminus\{c_0,c_1,\ldots,c_n\}$ such that:*
   (a) $\{c_i : c_i\, R\, k_J\}=J$,
   (b) *$k_J \not R c_j$ for all j with $1\leq j\leq n$,*
   (c) $\mathcal{G}\models\psi(c_1,\ldots,c_n)$,
   (d) *R is well-founded on G.*

Proof. ($\Leftarrow$). The proof is essentially the same as for Proposition 3.1 except that given $J\subseteq\{c_0,c_1,\ldots,c_n\}$ now it suffices to let

$$k_J=((J\mathrm{w}^M\underline{h^M_{n+1}}\cup\{\underline{h^M_i}:\underline{h^M_i}\notin^M c_i\})\setminus\{\underline{h^M_i}:\underline{h^M_i}\in^M c_i\})\cup\{\{c_1,\ldots,c_n\}\}.$$

($\Rightarrow$). It suffices to modify the construction in the corresponding part of the proof of Proposition 3.1 in the following way: In extending R_i to H_{i+1} simply *do not* add any pair of the form (t_{i+1},t), for $t\in H_i$. Since there is no R-cycle in G to start with, it is obvious that in this way no infinite descending chain with respect to $\bigcup_{i\in\omega}R_i$ is introduced in the Herbrand model eventually built, which is therefore a model of the Regularity Axiom. □

Even though in the above proposition we stated the decidability of the class of $\forall$-formulas, it is very easy to check that the previous proof works in fact for the decidability (with respect to NWLR) of the larger class of $\forall^*$-formulas. The class of the $\forall^*$-formulas corresponds to the well-known Bernays-Schönfinkel class of first-order logic (see [12]). It is interesting to notice that even though the decidability of the Bernays-Schönfinkel class is trivial, the decision problem for the $\forall$-formulas with respect to theories such as NWL or extensions of NWLE is still open.

In the context of NWLR, we can also state the following

Proposition 5.2. *There is a Herbrand model $\mathcal{M}$ such that if ψ is a $\forall$-formula satisfiable with respect to the theory* NWLR, *then ψ is satisfied in $\mathcal{M}$.*

Proof. For every satisfiable $\forall$-formula $\psi(x_1,\ldots,x_n)$, let

$$\langle G^\psi,c^\psi_0,c^\psi_1,\ldots,c^\psi_n,R^\psi\rangle$$

be a finite structure satisfying the conditions 1. – 3. of Proposition 5.1. Let H_0 be an infinite set of constants, with a superimposed well-founded binary relation R_0 such that for every $\forall$-formula $\psi(x_1,\ldots,x_n)$ a subgraph of R_0 isomorphic to $R^\psi\restriction\{c^\psi_0,\ldots,c^\psi_n\}$ can be found. It is straightforward to check that the Herbrand model built over H_0, $\mathbf{w}$, $\mathbf{l}$ with the interpretation of $\in$ induced, starting from R_0 by the axioms W and L fulfills the statement. □

6 The decidability of $\forall$-formulas in NWLER

As for the decidability problem we are concerned with, the effect of adding to NWLE the Regularity Axiom is the strongest possible, in the sense that it yields a theory which is complete with respect to existential closures of $\forall$-formulas. Hence our decision problem has a positive answer for every extension of NWLER, in particular for every classical theory of sets which assumes the Regularity Axiom. The simplifying effect of the addition of the Axiom of Regularity on the proof of Proposition 4.1 is also quite remarkable.

Proposition 6.1. *If $\psi(x_1,\ldots,x_n)$ is a $\forall$-formula, then $\psi(x_1,\ldots,x_n)$ is $1-1\neq\emptyset$-satisfiable with respect to the theory* NWLER *iff there exists a finite structure $\mathcal{G}=\langle G,c_0,c_1,\ldots,c_n,R\rangle$ such that*

1. *$c_i\neq c_j$ for all i,j with $0\leq i<j\leq n$,*
2. *c_0 has no R-predecessor in G,*
3. *for all $J\subseteq\{c_0,c_1,\ldots,c_n\}$ there exists an element $k_J\in G\setminus\{c_0,c_1,\ldots,c_n\}$ such that:*
   (a) *$\{c_i : c_i\,R\,k_J\}=J$,*
   (b) *$k_J\not R\,c_j$ for all j with $1\leq j\leq n$,*
   (c) *$\mathcal{G}\models\psi(c_1,\ldots,c_n)$,*
   (d) *R is extensional in G over $\{c_0,c_1,\ldots,c_n\}$,*
   (e) *R is well-founded on G.*

Furthermore NWLER *is complete with respect to existential closures of $\forall$-formulas.*

Proof. ($\Leftarrow$). The proof is the same as for Proposition 1.3.

($\Rightarrow$). Given a finite structure $\mathcal{G}$ satisfying the stated conditions, we show that ψ is satisfiable in the ordinary structure $\mathcal{HF}$ of the hereditarily finite sets with the obvious interpretation of $\emptyset$, $a\,\mathbf{w}^{HF}\,b=a\cup\{b\}$ and $a\,\mathbf{l}^{HF}\,b=a\setminus\{b\}$. Since $\langle G,R\rangle$ is well-founded we can define recursively a map $^*: G\longrightarrow HF$ satisfying the following conditions:

(i) for $i=1,\ldots,n$, $c_i^*=\{k^* : k\in G \text{ and } k\,R\,c_i\}$;

(ii) letting $G\setminus\{c_0,\ldots,c_n\}=\{k_1,\ldots,k_l\}$, for $i=1,\ldots,l$,
$$k_i^*=\{k^* : k\,R\,k_i\}\cup\{\{i,n+l+1\}\}.$$

Since $\text{rank}(\{i,n+l+1\})=n+l+2$, $\text{rank}(k_i^*)>n+l+3$. On the other hand, for $i=0,\ldots,n$, $\text{rank}(c_i^*)\leq n+l+1$, therefore for $k\in G$, $i=1,\ldots,l$, $k^*\neq\{i,n+1+l\}$. From this it follows that * is $1-1$ on $\{c_1,\ldots,c_n\}$ and in turn that * is an isomorphism between $\langle\{c_1,\ldots,c_n\},R\rangle$ and $\langle\{c_1^*,\ldots,c_n^*\},\in\rangle$. Since $\mathcal{G}\models\forall x\psi(c_1,\ldots,c_n,x)$ it is then obvious that if a is among $c_1^*,\ldots,c_n^*$, then $\mathcal{HF}\models\psi(c_1^*,\ldots,c_n^*,a)$.

If $a\in\mathcal{HF}\setminus\{c_1^*,\ldots,c_n^*\}$, we distinguish two cases:

1. For all i, $a\notin c_i^*$: Because of condition 3(c) on $\mathcal{G}$ it follows immediately that $\mathcal{HF}\models\psi(c_1^*,\ldots,c_n^*,a)$.

2. $a \in c_i^*$ for some $i = 1, \ldots, n$: Then by the defining conditions on *, there must be $k \in G$ such that $a = k^*$. From that it follows that a bears with respect to $\in$ and $\{c_1^*, \ldots, c_n^*\}$ the same relationship that k bears with respect to R and $\{c_1, \ldots, c_n\}$; therefore $\mathcal{HF} \models \psi(c_1^*, \ldots, c_n^*)$. Hence $\mathcal{HF} \models \forall x \psi(c_1^*, \ldots, c_n^*, x)$.

The completeness follows immediately since $\mathcal{HF}$ is isomorphically embedded as an $\in$-initial part into every model of NW. □

Remark. Let $\mathcal{G} = \langle G, c_0, \ldots, c_n, R\rangle$ be a finite structure of the kind considered so far, with $G = \{c_0, \ldots, c_n, c_{n+1}, \ldots, c_{n+l}\}$. The relevant properties of $\mathcal{G}$ ensuring that $\mathcal{G} \models \psi(c_1, \ldots, c_n)$, for a $\forall$-formula $\psi \equiv \forall x \varphi(x, x_1, \ldots, x_n)$, are all expressed by the following $\forall$-formula $\gamma_{\mathcal{G}}$, which we call the $\forall$-*diagram* of $c_1, \ldots, c_n$ in $\mathcal{G}$:

$$\forall x\,(\textstyle\bigwedge_{0\le i<j\le n}(x_i \neq x_j) \wedge \bigwedge_{i,j=0}^{n}(x_i \in_{ij} x_j) \wedge x \notin x_0$$
$$\wedge(\textstyle\bigwedge_{j=0}^{n}(x \neq x_j) \rightarrow \bigvee_{k=n+1}^{n+l}(x \in_{k,k} x \wedge \bigwedge_{i=0}^{n}(x \in_{k,i} x_i \wedge x_i \in_{i,k} x)))) .$$

Here $u \in_{ij} v$ stands for $u \in v$ if $c_i \, R \, c_j$, and for $u \notin v$ otherwise. Indeed if $\psi(x_1, \ldots, x_n)$ is satisfiable in $\mathcal{G}$ by $c_1, \ldots, c_n$, then $\gamma_{\mathcal{G}} \rightarrow \psi(x_1, \ldots, x_n)$ is logically valid. It is also clear that, given n, there are at most $2^{n^2+n} \cdot (2^{2^{2n+2}} - 1)$ inequivalent $\forall$-diagrams; moreover Propositions 3.1, 4.1, 5.1, and 6.1 provide an effective test to determine, for any of the theories we have taken into account, whether a $\forall$-diagram is satisfiable or not. Given a $\forall$-formula $\psi(x_1, \ldots, x_n)$ and a theory T among NWL, NWLE, NWLR, NWLER, if $\gamma_{G_1}, \ldots, \gamma_{G_p}$ are the $\forall$-diagrams with n free variables satisfiable with respect to T, we have that $\gamma_{G_1} \vee \cdots \vee \gamma_{G_p} \rightarrow \psi$ is logically valid. Conversely in every model of T in which ψ is satisfied by an n-tuple of elements, at least one of $\gamma_{G_1}, \ldots, \gamma_{G_p}$ is satisfied by the same n-tuple; therefore $\psi \rightarrow \gamma_{G_1} \vee \cdots \vee \gamma_{G_p}$ is a logical consequence of T. Therefore every $\forall$-formula is equivalent, with respect to T, to a disjunction of $\forall$-diagrams, which is empty just in case ψ is unsatisfiable with respect to T. For details on these issues, cf. [20].

7 Computational complexity

Let $\exists^*\forall$-CNF be the class of all prenex sentences in our membership language whose prefix has the $\exists^*\forall$ format and whose matrix is in conjunctive normal form. If each clause in the matrix has fewer than k disjuncts, then we speak of $\exists^*\forall$-kCNF.

In this section we show that the satisfiability problem for $\exists^*\forall$-CNF sentences is always NP-hard, and is in NP (and therefore is NP-complete) with respect to NWLR, NWLER or stronger theories.

We begin by showing the NP-hardness of two subclasses of the class $\exists^*\forall$-CNF.

Definition 7.1. We will say that a sentence Φ belongs to the class BT (Boolean Terms) if Φ is of the form $\exists x_1, \ldots, x_n \varphi(x_1, \ldots, x_n)$, where $\varphi(x_1, \ldots, x_n)$ is a conjunction of literals of one of the following forms:

$$x = y \cap z, \qquad x = y \cup z, \qquad x = \emptyset,$$

with x, y, and z variables.

Remark. A formula in BT is in fact a formula in the class $\exists^*\forall$-3CNF if the set-theoretic functional symbols appearing in Φ are eliminated via their (usual) definition in the language without functional symbols, and the formula obtained is brought to prenex conjunctive normal form. For example, the formula

$$\exists x, y, z(x = y \cup z \wedge x = y \cap z)$$

becomes

$$\exists x, y, z \forall w((w \in x \leftrightarrow w \in y \vee w \in z) \wedge (w \in x \leftrightarrow w \in y \wedge w \in z)),$$

which is easily seen to be equivalent to an $\exists^3\forall$-3CNF sentence.

Let 3CNFSAT be the satisfiability problem for propositional logic formulas in conjunctive normal form, where each clause is a disjunction of exactly 3 literals. We allow the logical constant $\perp$ (false) to occur as a literal.

Lemma 7.1. *3CNFSAT is polynomially reducible to the* BT *satisfiability problem with respect to* NWL, NWLE, NWLR, *and* NWLER.

Proof. Consider the following propositional formula in conjunctive normal form:

$$\varphi = (L_{1,1} \vee L_{1,2} \vee L_{1,3}) \wedge \cdots \wedge (L_{m,1} \vee L_{m,2} \vee L_{m,3}),$$

and let $P_1, \ldots, P_k$ be all the propositional variables appearing in φ. To define the formula φ^{set} to which φ will reduce, we will associate with each literal L that can appear in φ (namely P_i or $\neg P_i$ or $\perp$) a distinct set-theoretic variable denoted by $X(L)$. φ^{set} is defined as follows:

$$\begin{aligned}&\textstyle\bigwedge_{i=1}^{k}(X(P_i) \cap X(\neg P_i) = \emptyset) \wedge X(\perp) = \emptyset \\ &\wedge (X(L_{1,1}) \cup X(L_{1,2}) \cup X(L_{1,3})) \cap \cdots \cap (X(L_{m,1}) \cup X(L_{m,2}) \cup X(L_{m,3})) \neq \emptyset.\end{aligned}$$

The above φ^{set} can easily be formulated in BT (by recourse to additional variables to represent subterms); hence φ^{set} can be expressed as a $\forall$-3CNF formula.

We show that if φ is propositionally satisfiable, then φ^{set} is satisfiable within the model $\mathcal{HF}$ of hereditarily finite sets, which is a model of NWLER and hence of all of our theories.

If φ is propositionally satisfiable, then consider a truth assignment $\mathcal{V}$ which satisfies φ. For any propositional variable P_i, if $\mathcal{V}(P_i) = \top$ (true), then let

$$h(X(P_i)) = \{\emptyset\} \quad \text{and} \quad h(X(\neg P_i)) = \emptyset,$$

otherwise let

$$h(X(P_i)) = \emptyset \quad \text{and} \quad h(X(\neg P_i)) = \{\emptyset\}.$$

It is immediate to see that

$$\mathcal{HF} \models \varphi^{\mathrm{set}}\begin{bmatrix} X(P_1), \ldots, X(P_k), X(\neg P_1), \ldots, X(\neg P_k), X(\perp) \\ h(P_1), \ldots, h(P_k), h(\neg P_1), \ldots, h(\neg P_k), \emptyset \end{bmatrix}.$$

For the converse, assuming that $\mathcal{M}$ is a model of the weakest of our theories, namely NWL, in which φ^{set} is true when $m_1, \ldots, m_{2k}$ are substituted for the set variables $X(P_1), \ldots, X(P_k), X(\neg P_1), \ldots, X(\neg P_k)$, respectively, consider an element v in the domain of $\mathcal{M}$ such that

$$v \in^{\mathcal{M}} \delta^{\mathrm{set}}\begin{bmatrix} X(P_1), \ldots, X(P_k), X(\neg P_1), \ldots, X(\neg P_k), X(\perp) \\ m_1, \ldots, m_k, m_{k+1}, \ldots, m_{2k}, \emptyset^{\mathcal{M}} \end{bmatrix},$$

where δ^{set} denotes the term $\bigcap_{j=1}^{m}(X(L_{j,1}) \cup X(L_{j,2}) \cup X(L_{j,3}))$. Notice that such an element exists since the fact that φ^{set} is true in $\mathcal{M}$ guarantees that

$$\mathcal{M} \models (\delta^{set} \neq \emptyset)\begin{bmatrix} X(P_1), \ldots, X(P_k), X(\neg P_1), \ldots, X(\neg P_k), X(\perp) \\ m_1 \quad, \ldots, \quad m_k \quad, \ m_{k+1} \quad, \ldots, \quad m_{2k} \quad, \ \emptyset^M \end{bmatrix}.$$

It is now straightforward to check that the assignment $\mathcal{V}$ which sets P_i to $\top$ if and only if $v \in^M m_i$ will satisfy φ. □

Notice that the reduction technique that led us to the above result produced a $\forall$-3CNF formula involving no equality literals. By a more sophisticated coding that exploits equality along with membership, we will now obtain the following stronger reduction result.

Lemma 7.2. 3CNFSAT *is polynomially reducible to the satisfiability problem for the class* $\exists^*\forall$-2CNF *with respect to* NWL, NWLE, NWLR, *and* NWLER.

Proof. We maintain the same notation used in the preceding proof concerning the 3CNF input formula, the propositional atoms occurring in it, and the existential set variables to be exploited in the reduction. The formula φ^{set} has the quantificational prefix

$$\exists X(P_1), \ldots, X(P_k), X(\neg P_1), \ldots, X(\neg P_k), X(\perp)\ \forall V.$$

The matrix of φ^{set} consists of the following doubleton clauses, where $i = 1, \ldots, k$ and $j = 1, \ldots, m$:

(a) $V \notin X(\perp), \quad V \in X(P_i) \rightarrow V \notin X(\neg P_i)$,

(b) $X(\perp) \notin X(P_i) \rightarrow X(P_i) = X(\perp), \quad X(\perp) \notin X(\neg P_i) \rightarrow X(\neg P_i) = X(\perp)$,

(c) $X(L_{j,1}) \neq X(L_{j,2}) \vee X(L_{j,1}) \neq X(L_{j,3})$.

To obtain a set assignment h satisfying φ^{set} in $\mathcal{HF}$ from a truth assignment $\mathcal{V}$ satisfying φ, we put $h(X(\perp)) = \emptyset$ and furthermore, for any atom P_i,

$$h(X(P_i)) = \{\emptyset, \underline{i}\} \quad \text{and} \quad h(X(\neg P_i)) = \emptyset$$

if $\mathcal{V}(P_i) = \top$, and

$$h(X((P_i)) = \emptyset \quad \text{and} \quad h(X((\neg P_i)) = \{\emptyset, \underline{i}\},$$

otherwise.

It is straightforward to see that

$$\mathcal{HF} \models \varphi^{set}\begin{bmatrix} X(P_1), \ldots, X(P_k), X(\neg P_1), \ldots, X(\neg P_k), X(\perp) \\ h(P_1), \ldots, h(P_k), h(\neg P_1), \ldots, h(\neg P_k), \quad \emptyset \end{bmatrix}.$$

Conversely, to obtain a truth assignment $\mathcal{V}$ satisfying φ from an assignment h satisfying φ^{set} in any model $\mathcal{M}$ of NWL, we put for any atom P_i,

$$\mathcal{V}(P_i) = \top \quad \text{iff} \quad \emptyset^M \in^M h(X(P_i)).$$

Assuming, by contradiction, that $\mathcal{V}$ sets to $\perp$ each of $L_{j,1}$, $L_{j,2}$, and $L_{j,3}$ for some $j \in \{1, \ldots, m\}$, we argue as follows: From the clauses (a) we draw that

$$h(X(\perp)) \notin^M h(X(L_{j,b})) \text{ for } b = 1, 2, 3,$$

and from the clauses (b) we get $h(X(L_{j,b})) = \emptyset^M$, so that $h(X(L_{j,1})) = h(X(L_{j,2})) = h(X(L_{j,3}))$. This obviously conflicts with the clauses (c). □

Remark. An alternative method for reducing CNFSAT to the $\exists^*\forall$-CNF satisfiability problem can be found in [9], where φ gets translated into

$$\{\{X(P_1), X(\neg P_1)\}, \ldots, \{X(P_k), X(\neg P_k)\},\ \{X(L_1^1), \ldots, X(L_{n_1}^1), \emptyset\}, \ldots, \{X(L_1^m), \ldots, X(L_{n_m}^m), \emptyset\}\} = \{\{\emptyset, \{\emptyset\}\}\}.$$

Still other approaches can be found in [15] and in [29], [30].

Corollary 7.1. *Each of the satisfiability problems for the class* $\exists^*\forall$-2CNF *with respect to* NWL, NWLE, NWLR, NWLER *is* NP-*hard.*

Proof. The result follows directly from the NP-hardness of 3CNFSAT (cf. [13]) and the previous lemma. □

One might wonder whether the satisfiability problem for $\exists^*\forall$-1CNF is already NP-hard. The answer is presumably no, since from results in [4, 22] it easily follows that the problem at hand has deterministic polynomial time complexity.

We have seen in Sections 5 and 6 that the addition of the Regularity Axiom had a simplifying effect on our decidability proofs. This is the case also from the point of view of computational complexity, and in the following we prove the existence of algorithms for testing the satisfiability of formulas in the class $\exists^*\forall$-CNF with respect to NWLR and NWLER, which run in polynomial time on a non-deterministic Turing Machine. The above mentioned algorithms start with a preliminary reduction of the input formula to a formula containing *restricted* universal quantifiers only: consider a formula

$$\Psi \equiv \exists x_1, \ldots, x_n \forall y\, p,$$

with p in conjunctive normal form. We can assume that there are no occurrences of the constant symbol $\emptyset$ in p since any such occurrence could be eliminated by passing to the following (equisatisfiable) sentence

$$\exists x_0, x_1, \ldots, x_n \forall y (y \notin x_0 \wedge p[\emptyset / x_0]).$$

Let $D_1, \ldots, D_d$ be the conjuncts in p, where each of the D_i's is deleted if it contains a complementary pair of literals. In general D_i is assumed to have form

$$\ell_1^i \vee \cdots \vee \ell_{m_i}^i \vee \ell_{m_i+1}^i \vee \cdots \vee \ell_{m_i+s_i}^i,$$

where y does not appear in $\ell_1^i, \ldots, \ell_{m_i}^i$ and appears in $\ell_{m_i+1}^i, \ldots, \ell_{m_i+s_i}^i$. The case $m_i = 0$ corresponds to the case in which y is present in every one of the literals in D_i, whereas if there are no occurrences of y in D_i, then $s_i = 0$.

We can also assume that there are no literals of the form $z \neq z$, since such literals are always unsatisfiable and therefore they can be safely deleted from the clauses (declaring the entire formula Ψ to be unsatisfiable in case no other literal is left). At this point we can distribute the universal quantifier obtaining the following form for our formula:

$$\exists x_1, \ldots, x_n \bigwedge_{i=1}^{d} (\ell_1^i \vee \cdots \vee \ell_{m_i}^i \vee \forall y (\ell_{m_i+1}^i \vee \cdots \vee \ell_{m_i+s_i}^i)),$$

where any ℓ_h^i with $m_i + 1 \leq h \leq m_i + s_i$ is of one of the following forms:

$$x_j \in y, \quad y \in x_j, \quad x_j = y, \quad y = x_j, \quad x_j \notin y, \quad y \notin x_j, \quad x_j \neq y, \quad y \neq x_j.$$

In Proposition 5.1 we essentially showed that the following is a theorem of NWLR (and therefore of any stronger theory): For all $x_1, \ldots, x_N$ and for $1 \leq K \leq N$, there always exists a y different from and not belonging to any of the x_i's for $i \in \{1, \ldots, N\}$, and, moreover, having exactly $x_1, \ldots, x_K$ as only elements among $x_1, \ldots, x_N$; in formulas, for $1 \leq K \leq N$,

$$\begin{aligned} &\forall x_1, \ldots, x_N \left(\bigwedge\nolimits_{1 \leq i < j \leq N} x_i \neq x_j\right. \\ &\rightarrow \exists y \left(\bigwedge\nolimits_{i=1}^{K} (y \neq x_i \wedge y \notin x_i \wedge x_i \in y) \wedge \bigwedge\nolimits_{j=K+1}^{N} (y \neq x_j \wedge y \notin x_j \wedge x_j \notin y)\right). \end{aligned}$$

This implies that any formula of the form $\forall y\,(\ell^i_{m_i+1} \vee \cdots \vee \ell^i_{m_i+s_i})$ in which there is no ℓ^i_h of either the forms $y \neq x_j$, $x_j \neq y$, or $y \notin x_j$, can be rewritten as

$$\bigvee\nolimits_{(j_1,j_2) \in E_i} x_{j_1} = x_{j_2},$$

where $E_i = \{(j_1, j_2) :$ the literals $x_{j_1} \in y$ and $x_{j_2} \notin y$ are among $\ell^i_{m_i+1}, \ldots, \ell^i_{m_i+s_i}\}$. In other words, when no literals of the form $y \neq x_j, x_j \neq y$, or $y \notin x_j$ appear in D_i, the set of literals in D_i which contain the universally quantified variable either carries no information, being false at least for some element in the universe, or turns into a tautology if some pair of distinct existentially quantified variables are interpreted as representing the same set.

We perform the above substitution of the set of literals containing the universally quantified variable whenever possible, declaring Ψ to be unsatisfiable if there exists a conjunct D_i such that $m_i = 0$ and $E_i = \emptyset$.

Let us now consider a $\forall y\,(\ell^i_{m_i+1} \vee \cdots \vee \ell^i_{m_i+s_i})$ such that there exists an ℓ^i_h which is of the form $y \neq x_j$ or $x_j \neq y$. In case $s_i = 1$, we can simply remove such a part (which is unsatisfiable) from D_i and, as above, if $m_i = 0$, then we can declare Ψ to be unsatisfiable. Otherwise we substitute $\forall y\,(\ell^i_{m_i+1} \vee \cdots \vee \ell^i_{m_i+s_i})$ with

$$(\ell^i_{m_i+1} \vee \cdots \vee \ell^i_{h-1} \vee \ell^i_{h+1} \vee \cdots \vee \ell^i_{m_i+s_i})[x/x_j].$$

At this point any $\forall y\,(\ell^i_{m_i+1} \vee \cdots \vee \ell^i_{m_i+s_i})$ still remaining in our formula contains an ℓ^i_h which is of the form $y \notin x_j$, namely the formula is now in purely universal Δ_0-form and, in particular, belongs to the class of the so-called $(\forall)^1_0$-*simple prenex formulas*, introduced in [7] and there proved to have a satisfiability problem that can be solved in polynomial time on a suitable non-deterministic Turing Machine. Actually the underlying theory used in [7] was an unspecified axiomatization of naïve set theory (say ZF), but the argument works in the case of NWLER and can even be simplified for NWLR. Here is a short description of the decision algorithm for the theories NWLER and NWLR:

If a $(\forall)^1_0$-simple prenex formula is satisfied in a given model $\mathcal{M}$ of NWLER, then there exists a graph $\mathcal{G}$ (called *model-graph* of the formula in question) which has the following characteristics:

1. Nodes in $\mathcal{G}$ represent sets in $\mathcal{M}$ and, in particular, there is at most one node for each of the sets $m_1, \ldots, m_n$ used to interpret the existentially quantified variables;
2. for any pair of distinct sets m_i, m_j $(0 < i, j \leq n)$, there is a node in $\mathcal{G}$ representing a set which belongs to exactly only one between m_i and m_j;

3. there is an edge going from node a to node b if and only if the set represented by a is an element in $\mathcal{M}$ of the set represented by b;
4. $\mathcal{G}$ is acyclic.

It can also be shown, by induction on the number n of existentially quantified variables, that the total number of nodes in the smallest model-graph for a satisfiable formula is less than $2n$.

The decision test for the class of $(\forall)_0^1$-simple prenex formulas is based on the fact that the existence of a model-graph with the above characteristics is also a sufficient condition for the satisfiability of the given formula. As a matter of fact, it can be shown that in case a model-graph exists, the formula in question is true in the model $\mathcal{HF}$, and therefore in any model of NWLER, and this can be checked by simply inspecting the graph. The non-deterministic polynomial satisfiability algorithm will consist in guessing the model-graph and checking its adequacy against the formula. The reader is referred to [7] for the details.

Notice that the model-graph technique hinted at here is akin to the decision technique underlying Propositions 5.1 and 6.1. However here we can resort to a simpler (i. e. smaller) structure $\mathcal{G}$, that can be guessed in polynomial time, so taking advantage of the fact that universal quantifiers are restricted in the input formula.

The case of NWLR is treated in complete analogy, the only difference being that now the model-graph does not need to be *extensional*, namely condition 2 above does not need to hold, and therefore the only nodes that need to appear in the smallest model-graph are those associated to the sets used to interpret the existentially quantified variables. Given such a model-graph we can rebuild a model of NWLR by starting with n constants and closing with respect to the functional symbols w and l as we did in our constructions in the lemmas of the preceding section.

Remark. In order to obtain from the results in this section a proof of the NP-completeness of the entire class of $\exists^*\forall$-sentences, it would be necessary, for example, to have a polynomial-time reduction of $\exists^*\forall$-sentences to $\exists^*\forall$-CNF-sentences. Unfortunately, the classical reduction used in the context of propositional logic, based on the introduction of new propositional symbols, does not seem to generalize to this context.

Related work. Further studies along the lines presented here are currently under progress. Among them we mention the attempt to extend the results contained in this paper to the class of $\forall\forall$-formulas and the studies of the satisfiability problem in the context of (weak) set theories assuming a suitable form of *antifoundation* axiom (cf. [1]). The class of $\forall\forall$-formulas has been shown to be decidable with respect to the theory NWL in [23] by enhancing the method in this paper. By related (as well as by alternative) techniques, various satisfiability problems regarding non-well-founded membership theories have been considered and solved in [21, 2].

References

[1] Aczel, P., Non-Well-Founded Sets. CSLI Lecture Notes **14**, Stanford Univ., 1988.

[2] Bellé, D., and F. Parlamento, Decidability and completeness of open formulae in membership theories. Technical Report 25/93, Università di Udine, 1993.

[3] BELLÉ, D., and F. PARLAMENTO, Undecidability of weak membership theories. In: Proceedings of the International Conference on Logic and Algebra, Siena, April 1994.

[4] BREBAN, M., A. FERRO, E. G. OMODEO, and J. T. SCHWARTZ, Decision procedures for elementary sublanguages of set theory II. Formulas involving restricted quantifiers, together with ordinal, integer, map, and domain notions. Communications on Pure and Applied Mathematics **34** (1981), 177 - 195.

[5] CANTONE, D., V. CUTELLO, and A. POLICRITI, Decidability results for classes of purely universal formulas and quantifiers elimination in set theory. Le Matematiche **43** (1988), 303 - 336.

[6] CANTONE, D., A. FERRO, and E. G. OMODEO, Computable Set Theory. Vol. 1. Oxford University Press, London 1989.

[7] CANTONE, D., E. G. OMODEO, and A. POLICRITI, The automation of syllogistic II. Optimization and complexity issues. J. Automated Reasoning **6** (1990), 173 - 187.

[8] COLLINS, E. G., and J. D. HALPERN, On the interpretability of arithmetic in set theory. Notre Dame J. Formal Logic **11** (1970), 477 - 483.

[9] DOVIER, A., E. G. OMODEO, E. PONTELLI, and G.-F. ROSSI, Embedding finite sets in a logic programming language. In: Extensions of Logic Programming (E. LAMMA and P. MELLO, eds.), Lecture Notes in AI **660**, Springer-Verlag, Berlin-Heidelberg-New York 1992, pp. 150 - 167.

[10] DOVIER, A., E. G. OMODEO, E. PONTELLI, and G.-F. ROSSI, {log}: a language for programming in logic with finite sets. Technical Report Rap. 04. 93, Università di Roma La Sapienza, 1993.

[11] DOVIER, A., E. G. OMODEO, E. PONTELLI, and G.-F. ROSSI, {log}: a logic programming language with finite sets. In: Logic Programming, Proceedings of the Eight International Conference ICLP-91, MIT Press, Cambridge, Ma., 1991, pp. 111 - 124.

[12] DREBEN, B., and W. D. GOLDFARB, The Decision Problem. Solvable Classes of Quantificational Formulas. Addison-Wesley Co., Inc., Reading, Ma. 1979.

[13] GAREY, M. R., and D. S. JOHNSON, Computers and Intractability. A Guide to the Theory of NP-Completeness. W. H. Freeman and Company, New York 1979.

[14] GOGOL, D., The $\forall_n\exists$-completeness of Zermelo-Fraenkel set theory. Zeitschrift Math. Logik Grundlagen Math. **24** (1978), 289 - 290.

[15] HIBTI, M., NP-complétude du langage MLS. Technical report, Université de Franche-Comté, UFR des sciences et techniques, 1991.

[16] LEGROS, E., Tests de satisfiabilité dans un langage de programmation en logique avec contraintes ensemblistes. PhD thesis, Université de Franche-Comté, 1993.

[17] LEWIS, H. R., Unsolvable Classes of Quantificational Formulas. Addison-Wesley Co., Inc., Reading, Ma. 1979.

[18] MANCINI, A., and F. MONTAGNA, A minimal predicative set theory. Notre Dame J. Formal Logic **35** (1994), 186 - 203.

[19] OMODEO, E. G., F. PARLAMENTO, and A. POLICRITI, Truth tables for a combinatorial kernel of set theories. In: Proceedings of ECAI '90.

[20] OMODEO, E. G., F. PARLAMENTO, and A. POLICRITI, A derived algorithm for evaluating ε-expressions over abstract sets. J. Symbolic Computation **15** (1993), 673 - 704.

[21] OMODEO, E. G., and A. POLICRITI, Decision procedures for set/hyperset contexts. In: Design and Implementation of Symbolic Computation Systems (A. MIOLA, ed.). Lecture Notes in Computer Sciences **722**, Springer-Verlag, Berlin-Heidelberg-New York 1993, pp. 192 - 215.

[22] PARLAMENTO, F., and A. POLICRITI, Decision procedures for elementary sublanguages of set theory IX. Unsolvability of the decision problem for a restricted subclass of the Δ_0-formulae in set theory. Communications on Pure and Applied Mathematics **41** (1988), 221 - 251.

[23] PARLAMENTO, F., and A. POLICRITI, The satisfiability problem for the class $\exists^*\forall\forall$ in a set-theoretic setting. Technical Report 13/92, Università di Udine, October 1992.

[24] PARLAMENTO, F., and A. POLICRITI, Undecidability results for restricted universally quantified formulae of set theory. Communications on Pure and Applied Mathematics **46** (1993), 57 - 73.

[25] PARLAMENTO, F., A. POLICRITI, and K. P. S. B. Rao, Witnessing differences without redundancies. Technical Report 23/93, Università di Udine, 1993.

[26] POLICRITI, A., Completeness and decidability of the deducibility problem for some class of formulas of set theory. Le Matematiche **42** (1987), 49 - 66.

[27] SZMIELEW, W., and A. TARSKI, Mutual interpretability of some essentially undecidable theories. In: Proceedings of the International Congress of Mathematicians (Cambridge 1950), Vol. 1. Cambridge University Press, Cambridge 1952, p. 724.

[28] TARSKI, A., A. MOSTOWSKI, and R. M. ROBINSON, Undecidable Theories. North-Holland Publ. Comp., Amsterdam 1953.

[29] TETRUASHVILI, M. R., Complexity of the solution of the validity problem for an elementary sublanguage of quantifier-free set theory. Soobsh. Akad. Nauk. Gruzin. SSR **126** (1987), 245 - 247.

[30] TETRUASHVILI, M. R., Decision complexity of the theory MLSC. Soobsh. Akad. Nauk. Gruzin. SSR, **132** (1988), 465 - 468.

[31] VAUGHT, R. L., On a theorem of Cobham concerning undecidable theories. In: Proceedings of the 1960 International Congress on Logic, Methodology and Philosophy of Science (E. NAGEL, P. SUPPES, and A. TARSKI, eds.), Stanford University Press, Stanford 1962, pp. 14 - 25.

(Received: September 10, 1994)