

LTL with the Freeze Quantifier and Register Automata

STÉPHANE DEMRI

LSV, CNRS & ENS Cachan & INRIA Futurs, France

and

RANKO LAZIĆ

Department of Computer Science, University of Warwick, UK

A data word is a sequence of pairs of a letter from a finite alphabet and an element from an infinite set, where the latter can only be compared for equality. To reason about data words, linear temporal logic is extended by the freeze quantifier, which stores the element at the current word position into a register, for equality comparisons deeper in the formula. By translations from the logic to alternating automata with registers and then to faulty counter automata whose counters may erroneously increase at any time, and from faulty and error-free counter automata to the logic, we obtain a complete complexity table for logical fragments defined by varying the set of temporal operators and the number of registers. In particular, the logic with future-time operators and 1 register is decidable but not primitive recursive over finite data words. Adding past-time operators or 1 more register, or switching to infinite data words, cause undecidability.

Categories and Subject Descriptors: F.4.1 [Mathematical Logic and Formal Languages]: Mathematical Logic—*Temporal logic*; F.4.1 [Mathematical Logic and Formal Languages]: Formal Languages—*Decision problems*; F.1.1 [Computation by Abstract Devices]: Models of Computation—*Automata*

General Terms: Algorithms, Verification

Additional Key Words and Phrases: Computational complexity, Expressiveness

1. INTRODUCTION

Context. Logics and automata for words and trees over finite alphabets are relatively well-understood. Motivated partly by the need for formal verification and synthesis of infinite-state systems, and the search for automated reasoning techniques for XML, there is an active and broad research programme on logics and automata for words and trees which have richer structure.

Segoufin’s recent survey [Segoufin 2006] summarises the substantial progress made on reasoning about data words and data trees. A data word is a word

This paper is a revised and extended version of [Demri and Lazić 2006].

The first author was supported by the ACI “Sécurité et Informatique” CORTOS.

The second author was supported by grants from the EPSRC (GR/S52759/01) and the Intel Corporation, and by ENS Cachan.

Permission to make digital/hard copy of all or part of this material without fee for personal or classroom use provided that the copies are not made or distributed for profit or commercial advantage, the ACM copyright/server notice, the title of the publication, and its date appear, and notice is given that copying is by permission of the ACM, Inc. To copy otherwise, to republish, to post on servers, or to redistribute to lists requires prior specific permission and/or a fee.

© 20YY ACM 1529-3785/20YY/0700-0001 \$5.00

over a finite alphabet, with an equivalence relation on word positions. Implicitly, every word position is labelled by an element (“datum”) from an infinite set (“data domain”), but since the infinite set is equipped only with the equality predicate, it suffices to know which word positions are labelled by equal data, and that is what the equivalence relation represents. Similarly, a data tree is a tree (countable, unranked and ordered) whose every node is labelled by a letter from a finite alphabet, with an equivalence relation on the set of its nodes.

First-order logic for data words was considered in [Bojańczyk et al. 2006], where variables range over word positions ($\{0, \dots, l-1\}$ or \mathbb{N}), there is a unary predicate for each letter from the finite alphabet, and there is a binary predicate $x \sim y$ for the equivalence relation representing equality of data labels. $\text{FO}^2(\sim, <, +1)$ denotes such a logic with two variables and binary predicates $x+1 = y$ and $x < y$. Over finite and over infinite data words, satisfiability for $\text{FO}^2(\sim, <, +1)$ was proved decidable and at least as hard as reachability for Petri nets. The latter problem is EXPSPACE -hard [Lipton 1976], but whether it is elementary has been open for many years. If the logic is extended by one more variable, $+1$ becomes expressible using $<$, but satisfiability was shown undecidable.

Words which contain data from domains with more than the equality predicate were proposed in [Bouajjani et al. 2007] as models of configurations of systems with unbounded control structures. Decidability of satisfiability was proved for the $\exists^*\forall^*$ fragment of a first-order logic over such words provided that the underlying logic on data is decidable.

Alternatively to first-order logic over data words, expressiveness and algorithmic properties of formalisms based on linear temporal logic were studied in [French 2003; Lisitsa and Potapov 2005; Demri et al. 2007; Lazić 2006; Demri et al. 2007; Demri et al. 2008]. LTL was extended by the freeze quantifier: \downarrow_r stores in register r the equivalence class of the current word position, and the atomic formula \uparrow_r in its scope is true at a word position iff the latter belongs to the equivalence class stored in r . Thus, data at different word positions can be compared for equality. Freeze quantification has also been considered in timed logics (cf. e.g. [Alur and Henzinger 1994]) and hybrid logics (cf. e.g. [Goranko 1996]), and Fitting has called for an investigation of effects of its addition to modal logics [Fitting 2002]. Let $\text{LTL}_n^\downarrow(\mathcal{O})$ denote LTL with the freeze quantifier, n registers, and temporal operators \mathcal{O} . Satisfiability over infinite data words was shown highly undecidable (i.e., Σ_1^1 -hard) for $\text{LTL}_2^\downarrow(\mathbf{X}, \mathbf{X}^{-1}, \mathbf{F}, \mathbf{F}^{-1})$ in [French 2003] (where \mathbf{X}^{-1} and \mathbf{F}^{-1} are the past-time versions of \mathbf{X} and \mathbf{F}) and for $\text{LTL}_2^\downarrow(\mathbf{X}, \mathbf{U})$ in [Lisitsa and Potapov 2005; Demri et al. 2007], so complexity (even decidability) of fragments with 1 register remained unknown. To obtain decidability, various structural restrictions were employed: flat formulae [Demri et al. 2007], Boolean combinations of safety formulae [Lazić 2006], and that the freeze quantifier is used only for expressing that the current datum occurs eventually in the future or past [Demri et al. 2007]. In [Demri et al. 2008], decidability was obtained by replacing satisfiability with model checking data words generated by deterministic one-counter automata.

A third approach to reasoning about data words are register automata [Kaminski and Francez 1994; Sakamoto and Ikeda 2000; Neven et al. 2004]. In addition to a finite number of control locations, such an automaton has a finite number of regis-

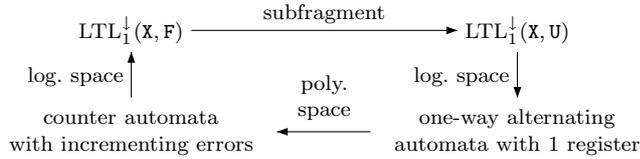


Fig. 1. A circle of translations which preserve language nonemptiness

ters which can store data for later equality comparisons. In pursuit of a satisfactory notion of regular languages of finite data words, nonemptiness was shown decidable for one-way nondeterministic register automata, but the class turned out not to be closed under complement, and the nonuniversality problem to be undecidable. However, for such automata \mathcal{A} and \mathcal{A}' , whether the language of \mathcal{A} contains the language of \mathcal{A}' was proved decidable provided \mathcal{A} has only 1 register (in the terminology of this paper). The subclass with 1 register was thus the best candidate found for defining regularity, but it is also not closed under complement. The case of infinite data words was not considered.

Contribution. The main technical achievement of the paper are translations as depicted in Figure 1. Over finite and over infinite data words, the translation from the logic to register automata preserves languages, and for the translations to and from counter automata, appropriate projections of data words are taken. For infinite data words, the register automata have the weak acceptance mechanism [Muller et al. 1986], which makes them closed under complement. Alternation is needed because there exist properties which are expressible in $LTL_1^{\downarrow}(X, F)$ (e.g., ‘no two word positions have equal data’) but not by any one-way nondeterministic register automaton. The counter automata are one-way, nondeterministic, accept infinite words by the Büchi mechanism, and are faulty in the sense that counters may erroneously increase at any time.

Using results in [Schnoebelen 2002; Mayr 2003; Ouaknine and Worrell 2006a; 2006b], we show that nonemptiness for the faulty counter automata is decidable and not primitive recursive over finite words, and Π_1^0 -complete over infinite words. Hence, the same bounds hold for $LTL_1^{\downarrow}(X, F)$ satisfiability, $LTL_1^{\downarrow}(X, U)$ satisfiability, and nonemptiness of one-way alternating automata with 1 register. The latter therefore provide an attractive notion of regular languages of finite data words, although the complexity of nonemptiness is very high.

The incrementing errors of counter automata correspond to restricted powers of future-time LTL and one-way alternating automata with 1 register. As soon as any of 1 more register, the F^{-1} temporal operator or backward automaton moves are added, even after replacing U by F and restricting to universal automata, decidability and Π_1^0 -membership break down: we obtain logarithmic-space translations from Minsky (i.e., error-free) counter automata, which result in Σ_1^0 -hardness over finite data words and Σ_1^1 -hardness over infinite data words. Together with the bounds via the faulty counter automata, that gives us a complete complexity table for fragments of LTL with the freeze quantifier defined by varying the set of temporal operators and the number of registers (see Figure 12).

Interestingly, similar results were reported in [Ouaknine and Worrell 2006a] for real-time metric temporal logic, by translations to and from alternating automata with 1 clock and machines with fifo channels. Indeed, computing a counter au-

tomaton with incrementing errors from a sentence of $LTL_1^\downarrow(\mathbf{X}, \mathbf{U})$ follows the same broad steps as may be employed to compute a channel machine with insertion errors from a sentence of (future-time) MTL, some of which were implicit already in the proof that whether the language of a one-way nondeterministic register automaton is contained in the language of such an automaton with 1 register is decidable over finite data words [Kaminski and Francez 1994, Appendix A]. However, there is no obvious translation from $LTL_1^\downarrow(\mathbf{X}, \mathbf{U})$ to MTL or alternating automata with 1 clock, and counters are less powerful than fifo channels. Also, the translations from counter automata to LTL with freeze in this paper and those from channel machines to MTL differ substantially.

The decidability of satisfiability for $LTL_1^\downarrow(\mathbf{X}, \mathbf{U})$ over finite data words makes it the competitor of $FO^2(\sim, <, +1)$. To clarify the relationship between the two logics, we extend the equiexpressiveness result in [Etessami et al. 2002] and show that $FO^2(\sim, <, +1, \dots, +m)$ is as expressive as $LTL_1^\downarrow(\mathbf{X}, \mathbf{X}^{-1}, \mathbf{F}, \mathbf{F}^{-1})$ with a restriction on freeze quantification. However, temporal sentences may be exponentially longer than equivalent first-order formulae.

Organisation. After setting up machinery in Section 2 and presenting the equiexpressiveness result for $FO^2(\sim, <, +1, \dots, +m)$ in Section 3, the core of the paper is Sections 4 and 5, which contain the results on complexity of satisfiability for fragments of LTL with the freeze quantifier and nonemptiness for classes of register automata. We conclude in Section 6.

2. PRELIMINARIES

After defining words with data, we introduce below logics and automata with which we shall work in the paper. To define acceptance by alternating automata, we recall a simple class of two-player games. This section also contains several results which will be used later. Their proofs are either relatively straightforward or heavily based on proofs in the literature.

2.1 Data Words

A *data word* σ over a finite alphabet Σ is a nonempty word $\text{str}(\sigma)$ over Σ together with an equivalence relation \sim^σ on its positions. We write $|\sigma|$ for the length of the word, $\sigma(i)$ for the letter at position i , and $[i]_{\sim^\sigma}$ for the class that contains i , where $0 \leq i < |\sigma|$. When σ is understood, we may write simply \sim instead of \sim^σ . We shall sometimes refer to classes of \sim as ‘data’.

Example 2.1. A data word of length 3 over $\{a, b\}$ is σ such that $\text{str}(\sigma) = aab$ and the classes of \sim^σ are $\{0, 2\}$ and $\{1\}$.

2.2 LTL over Data Words

Syntax. $LTL^\downarrow(\mathcal{O})$ will denote the linear temporal logic with the freeze quantifier and temporal operators in the set \mathcal{O} . Each formula is over a finite alphabet Σ . Atomic propositions a are elements of Σ , $\mathbf{0}$ ranges over \mathcal{O} , and r ranges over $\mathbb{N}_{>0}$.

$$\phi ::= a \mid \top \mid \neg\phi \mid \phi \wedge \phi \mid \mathbf{0}(\phi, \dots, \phi) \mid \downarrow_r \phi \mid \uparrow_r$$

An occurrence of \uparrow_r within the scope of some freeze quantification \downarrow_r is bound by it; otherwise, it is free. A sentence is a formula with no free occurrence of any \uparrow_r .

We consider temporal operators ‘next’ (X), ‘eventually’ (F), ‘until’ (U), and their past-time versions (X^{-1} , F^{-1} , U^{-1}). As $F\phi$ is equivalent to $\top U\phi$, F can be omitted from any set which contains U, and the same is true for F^{-1} and U^{-1} . As usual, we regard G (‘always’) and G^{-1} (‘past always’) as abbreviations for $\neg F\neg$ and $\neg F^{-1}\neg$.

Let $LTL_n^\downarrow(\mathcal{O})$ be the fragment with n registers, i.e. where $r \in \{1, \dots, n\}$.

Semantics. A register valuation v for a data word σ is a finite partial map from $\mathbb{N}_{>0}$ to the classes in σ , i.e. to $\{i : 0 \leq i < |\sigma|\} / \sim$. If $r \notin \text{dom}(v)$, then the atomic formula \uparrow_r will evaluate to false with respect to v . Such undefined register values will be used for initial automata states. We say that v is an n -register valuation iff $\text{dom}(v) \subseteq \{1, \dots, n\}$.

For a data word σ over a finite alphabet Σ , a position $0 \leq i < |\sigma|$, a register valuation v for σ , and a formula ϕ over Σ , writing $\sigma, i \models_v \phi$ will mean that ϕ is satisfied by σ at position i with respect to v . The satisfaction relation is defined as follows, where we omit the Boolean cases.

$$\begin{aligned} \sigma, i \models_v a &\stackrel{\text{def}}{\iff} \sigma(i) = a \\ \sigma, i \models_v X\phi &\stackrel{\text{def}}{\iff} i + 1 < |\sigma| \text{ and } \sigma, i + 1 \models_v \phi \\ \sigma, i \models_v X^{-1}\phi &\stackrel{\text{def}}{\iff} i - 1 \geq 0 \text{ and } \sigma, i - 1 \models_v \phi \\ \sigma, i \models_v \phi U \psi &\stackrel{\text{def}}{\iff} \text{for some } j \geq i, \sigma, j \models_v \psi \text{ and for all } i \leq j' < j, \sigma, j' \models_v \phi \\ \sigma, i \models_v \phi U^{-1} \psi &\stackrel{\text{def}}{\iff} \text{for some } j \leq i, \sigma, j \models_v \psi \text{ and for all } j < j' \leq i, \sigma, j' \models_v \phi \\ \sigma, i \models_v \downarrow_r \phi &\stackrel{\text{def}}{\iff} \sigma, i \models_{v[r \mapsto [i] \sim]} \phi \\ \sigma, i \models_v \uparrow_r &\stackrel{\text{def}}{\iff} r \in \text{dom}(v) \text{ and } i \in v(r) \end{aligned}$$

Example 2.2. Consider the sentence $\phi = G(a \Rightarrow \downarrow_1 X((G(a \Rightarrow \neg \uparrow_1)) \wedge (F(b \wedge \uparrow_1))))$ of $LTL_1^\downarrow(X, F)$, which is over the alphabet $\{a, b\}$. It states that no two letters a are in the same class, and that every letter a is followed by a letter b which is in the same class. Thus, for the data word σ in Example 2.1, we have $\sigma, 0 \not\models_\emptyset \phi$.

2.3 FO over Data Words

As defined in [Bojańczyk et al. 2006], $FO(\sim, <, +1, \dots, +m)$ denotes first-order logic over data words, in which variables range over word positions. We use variable names x_0, x_1, \dots . The predicates $x_i < x_j$ and $x_i = x_j + k$ are interpreted as expected. Each formula has an alphabet Σ , and it may contain unary predicates $P_a(x_i)$ which are satisfied by a data word iff the letter at position x_i is a . When we write $\phi(x_{i_1}, \dots, x_{i_N})$, it means that at most x_{i_1}, \dots, x_{i_N} occur free in ϕ .

$FO^n(\sim, <, +1, \dots, +m)$ is the fragment with n variables x_0, \dots, x_{n-1} .

Example 2.3. Let $\phi'(x_0)$ be the following formula of $FO^2(\sim, <)$, which states that, from position x_0 onwards, no two letters a are in the same class, and every letter a is followed by a letter b which is in the same class.

$$\begin{aligned} &\forall x_1 (\neg(x_1 < x_0) \wedge P_a(x_1) \Rightarrow \\ &\forall x_0 (x_1 < x_0 \wedge P_a(x_0) \Rightarrow \neg x_1 \sim x_0) \wedge \exists x_0 (x_1 < x_0 \wedge P_b(x_0) \wedge x_1 \sim x_0)) \end{aligned}$$

It is equivalent to the sentence ϕ from Example 2.2 in the sense that, for every data word σ over $\{a, b\}$ and $0 \leq i < |\sigma|$, we have $\sigma, i \models_\emptyset \phi$ iff $\sigma \models_{[x_0 \mapsto i]} \phi'(x_0)$.

2.4 Weak Games

The automata that will be introduced in the next section will be alternating and weak [Muller et al. 1986], so we shall use the following class of zero-sum two-player finitely branching games to define acceptance by such automata.

Games. A weak game G is a tuple $\langle P, P_1, P_2, \rightarrow, \rho \rangle$ such that:

- P is a set of all positions;
- P_1 and P_2 disjointly partition P into positions of players 1 and 2 (respectively);
- $\rightarrow \subseteq P \times P$ is a successor relation with respect to which every position has finitely many successors;
- $\rho : P \rightarrow \mathbb{N}$ specifies ranks so that, whenever $p \rightarrow p'$, we have $\rho(p) \geq \rho(p')$.

A play π of G is a sequence $p_0 p_1 \dots$ of positions of G such that $p_i \rightarrow p_{i+1}$ for each i . If π is infinite, let $\rho(\pi) = \rho(p_i)$, where i is such that $\rho(p_j) = \rho(p_i)$ for all $j > i$ (such an i necessarily exists).

We say that a play π of G is complete iff either it ends with a position without successors or it is infinite. For such π , we consider it winning for player 1 iff either it ends with a position of player 2 or it is infinite and $\rho(\pi)$ is even. The winning condition for player 2 is symmetric, with the opposite parity.

A strategy for player l from a position p of G is a tree $\tau \subseteq P^{<\omega}$ of finite plays of G such that:

- (i) $p \in \tau$ and it is the root;
- (ii) whenever $\pi \in \tau$ ends with a position p of player l which has at least one successor, it has a unique child;
- (iii) whenever $\pi \in \tau$ ends with a position p of the other player, it has all children $\pi p'$ with $p \rightarrow p'$.

We say that τ is positional iff the choices of successors in (ii) depend only on the ending positions p .

Now, a play by τ is either an element of τ or an infinite sequence whose every nonempty prefix is an element of τ . We say that τ is winning iff each complete play by τ is winning for player l .

Consistent Signature Assignments. Let G be a weak game as above.

A consistent signature assignment for G is a function α from some $W \subseteq P$ to \mathbb{N} such that the following are satisfied, where pairs of natural numbers are ordered lexicographically, i.e. $\langle n, m \rangle < \langle n', m' \rangle$ iff either $n < n'$, or $n = n'$ and $m < m'$.

- for every $p \in W \cap P_1$, there exists $p \rightarrow p'$ with $p' \in W$ and $\langle \rho(p'), \alpha(p') \rangle \leq \langle \rho(p), \alpha(p) \rangle$, where the inequality is strict if $\rho(p)$ is odd;
- for every $p \in W \cap P_2$ and every $p \rightarrow p'$, we have $p' \in W$ and $\langle \rho(p'), \alpha(p') \rangle \leq \langle \rho(p), \alpha(p) \rangle$, where the inequality is strict if $\rho(p)$ is odd.

Part (a) of the result below is straightforward, whereas part (b) is obtained by simplifying the proof of [Walukiewicz 2001, Lemma 10] which is for more general parity games.

THEOREM 2.4. *Suppose G is a weak game.*

- (a) For every consistent signature assignment α for G and every $p \in \text{dom}(\alpha)$, player 1 has a positional winning strategy from p .
- (b) There exists a consistent signature assignment α for G such that for every $p \notin \text{dom}(\alpha)$, player 2 has a positional winning strategy from p .

The following are two immediate corollaries:

- positional determinacy, i.e. that for every position p , one of the players has a positional winning strategy from p ;
- for every position p , there exists a consistent signature assignment which is defined for p iff player 1 has a positional winning strategy from p .

2.5 Register Automata

Corresponding to the addition of the freeze quantifier to LTL, finite automata can be extended by registers. We now define two-way alternating register automata over data words.

A state of such an automaton for a data word will consist of a word position, an automaton location and a register valuation. From it, according to the transition function, one of the following is performed:

- branching to another location depending on one of the following Boolean tests: whether the current letter equals a specified letter, whether the word position is the first or last, or whether the current datum equals the datum in a specified register;
- storing the current datum into a register;
- conjunctive or disjunctive branching to a pair of locations;
- acceptance or rejection;
- moving to the next or previous word position.

The automata will be weak in that each location will have a rank, which will not increase after any transition, and whose parities will be used to define acceptance.

Each location will also have a height, which will decrease after every transition which is not a move to another word position. The heights ensure that infinite progress cannot be made while remaining at the same word position. That constraint simplifies some proofs without reducing expressiveness.

Remark 2.5. In contrast to the formalisations of register automata in [Kaminski and Francez 1994; Sakamoto and Ikeda 2000; Neven et al. 2004], data stored in registers within an automaton state will not be required to be mutually distinct and to contain the datum from the previously visited word position.¹

¹That is a minor technical difference. It can be checked that, for every automaton with $n + 1$ registers in the sense of [Kaminski and Francez 1994; Sakamoto and Ikeda 2000; Neven et al. 2004], one can construct an equivalent automaton with $n + 1$ registers and an equivalent alternating automaton with n registers in the sense of this paper.

Automata. The set $\Delta(\Sigma, Q, n)$ of all transition formulae over a finite alphabet Σ , over a finite set Q of locations and with $n \in \mathbb{N}$ registers is defined below.

$$\begin{aligned} B(\Sigma, n) &= \{a, \mathbf{beg}, \mathbf{end}, \uparrow_r : a \in \Sigma, r \in \{1, \dots, n\}\} \\ \Delta(\Sigma, Q, n) &= \{q \not\leftarrow \beta \not\rightarrow q', \downarrow_r q, q \wedge q', q \vee q', \top, \perp, \mathbf{X}q, \overline{\mathbf{X}}q, \mathbf{X}^{-1}q, \overline{\mathbf{X}^{-1}}q : \\ &\quad \beta \in B(\Sigma, n), q, q' \in Q, r \in \{1, \dots, n\}\} \end{aligned}$$

We have that $\Delta(\Sigma, Q, n)$ is closed under the self-inverse operation of taking duals:

$$\begin{array}{lll} \overline{q \not\leftarrow \beta \not\rightarrow q'} = q \not\leftarrow \beta \not\rightarrow q' & \overline{q \wedge q'} = q \vee q' & \overline{\overline{\mathbf{X}}q} = \mathbf{X}q \\ \overline{\downarrow_r q} = \downarrow_r q & \overline{\top} = \perp & \overline{\mathbf{X}^{-1}q} = \overline{\mathbf{X}^{-1}}q \end{array}$$

The difference between transition formulae $\mathbf{X}q$ and their duals $\overline{\mathbf{X}}q$ is that the former will be rejecting and the latter accepting if there is no next word position, and similarly for $\mathbf{X}^{-1}q$ and $\overline{\mathbf{X}^{-1}}q$.

A *register automaton* \mathcal{A} is a tuple $\langle \Sigma, Q, q_I, n, \delta, \rho, \gamma \rangle$ as follows:

- Σ is a finite alphabet;
- Q is a finite set of locations;
- $q_I \in Q$ is the initial location;
- $n \in \mathbb{N}$ is the number of registers;
- $\delta : Q \rightarrow \Delta(\Sigma, Q, n)$ is a transition function;
- $\rho : Q \rightarrow \mathbb{N}$ specifies ranks and is such that, whenever q' occurs in $\delta(q)$, we have $\rho(q') \leq \rho(q)$;
- $\gamma : Q \rightarrow \mathbb{N}$ specifies heights and is such that, whenever $\delta(q)$ is of the form $q' \not\leftarrow \beta \not\rightarrow q'', \downarrow_r q', q' \wedge q''$ or $q' \vee q''$, we have $\gamma(q'), \gamma(q'') < \gamma(q)$.

We say that a register automaton is:

- one-way* iff no $\delta(q)$ is of the form $q' \not\leftarrow \mathbf{beg} \not\rightarrow q'', \mathbf{X}^{-1}q'$ or $\overline{\mathbf{X}^{-1}}q'$;
- nondeterministic* iff no $\delta(q)$ is of the form $q' \wedge q''$;
- universal* iff no $\delta(q)$ is of the form $q' \vee q''$;
- deterministic* iff it is both nondeterministic and universal.

For $d \in \{1, 2\}$ and $C \in \{A, N, U, D\}$, let $dCRA$ denote the class of all register automata with restrictions on directionality and control specified by d and C . Let $dCRA_n$ denote the subclass with n registers.

Acceptance Games. Let \mathcal{A} be a register automaton as above, and σ be a data word over Σ . The acceptance game of \mathcal{A} over σ is the weak game $G_{\mathcal{A}, \sigma} = \langle P, P_1, P_2, \rightarrow, \rho \rangle$ defined below. Player 1 (“automaton”) will be resolving the disjunctive branchings given by the transition function of \mathcal{A} , winning a finite play if it ends with an accepting state, and winning an infinite play if the limit location rank is even. Dually, player 2 (“pathfinder”) will be resolving the conjunctive branchings and winning at rejecting states or by odd limit ranks.

- P is the set of all states of \mathcal{A} for σ , which are triples $\langle i, q, v \rangle$ where $0 \leq i < |\sigma|$, $q \in Q$, and v is an n -register valuation for σ .

$\delta(q)$	owner of $\langle i, q, v \rangle$	successors of $\langle i, q, v \rangle$
$q' \not\prec \beta \succ q''$		$\{\langle i, q', v \rangle\}$, if $\sigma, i \models_v \beta$ $\{\langle i, q'', v \rangle\}$, if $\sigma, i \not\models_v \beta$
$\downarrow_r q'$		$\{\langle i, q', v[r \mapsto [i] \sim]\rangle\}$
$q' \wedge q''$	2	$\{\langle i, q', v \rangle, \langle i, q'', v \rangle\}$
\top	2	\emptyset
Xq'	1, if $i + 1 = \sigma $	$\{\langle i + 1, q', v \rangle\}$, if $i + 1 < \sigma $ \emptyset , if $i + 1 = \sigma $
$X^{-1}q'$	1, if $i = 0$	$\{\langle i - 1, q', v \rangle\}$, if $i > 0$ \emptyset , if $i = 0$

$$\sigma, i \models_v \mathbf{beg} \stackrel{\text{def}}{\Leftrightarrow} i = 0 \quad \sigma, i \models_v \mathbf{end} \stackrel{\text{def}}{\Leftrightarrow} i + 1 = |\sigma|$$

Fig. 2. Defining acceptance games

- The partition of P into P_1 and P_2 , and the successor relation, are given by the table in Figure 2. The ownership of states with unique successors has not been specified because it is irrelevant. The table omits dual transition formulae, which are treated by swapping the ownerships.
- For every $\langle i, q, v \rangle \in P$, $\rho(\langle i, q, v \rangle) = \rho(q)$.

Observe that every branching in $G_{\mathcal{A}, \sigma}$ is at most binary.

A run of \mathcal{A} over σ is a strategy τ in $G_{\mathcal{A}, \sigma}$ for player 1 from the initial state $\langle 0, q_I, \emptyset \rangle$. We say that τ is accepting iff it is winning, and that \mathcal{A} accepts σ iff \mathcal{A} has an accepting run over σ .

Example 2.6. Let \mathcal{A} be a register automaton with alphabet $\{a, b\}$ and 1 register, whose locations and transition function are shown in Figure 3, and such that the ranks of q_1 and q_7 are even but the rank of q_{11} is odd. It is straightforward to assign exact ranks and heights to the locations of \mathcal{A} so that the conditions in the definition of register automata are satisfied.

We have that \mathcal{A} is one-way, neither nondeterministic nor universal, and equivalent to the sentence ϕ from Example 2.2 in the sense that, for every data word σ over $\{a, b\}$, \mathcal{A} accepts σ iff $\sigma, 0 \models_\emptyset \phi$.

In particular, \mathcal{A} rejects the data word σ from Example 2.1. By positional determinacy (cf. Theorem 2.4) of the acceptance game $G_{\mathcal{A}, \sigma}$, player 2 (“pathfinder”) has a positional winning strategy from the initial state $\langle 0, q_1, \emptyset \rangle$. Such a strategy is shown in Figure 4, where $-$ and $\{1\}$ abbreviate register valuations \emptyset and $[1 \mapsto \{1\}]$ (respectively), sharp and oval frames indicate states belonging to players 1 and 2 (respectively), and states whose owner is irrelevant are not framed. The strategy is positional trivially, as no state is visited more than once. Essentially, the pathfinder challenges the automaton to find a letter b which follows the second letter a and is in the same class.

Closure Properties. We now consider closure of classes of register automata under complement, intersection and union.

THEOREM 2.7. (a) For each $d \in \{1, 2\}$, $dARA_n$ and $dDRA_n$ are closed under complement, and $dNRA_n$ is dual to $dURA_n$.

(b) For each $C \in \{A, N, U, D\}$, $1CRA$ is closed under intersection and union. For intersections of universal or alternating automata, and for unions of nondeter-

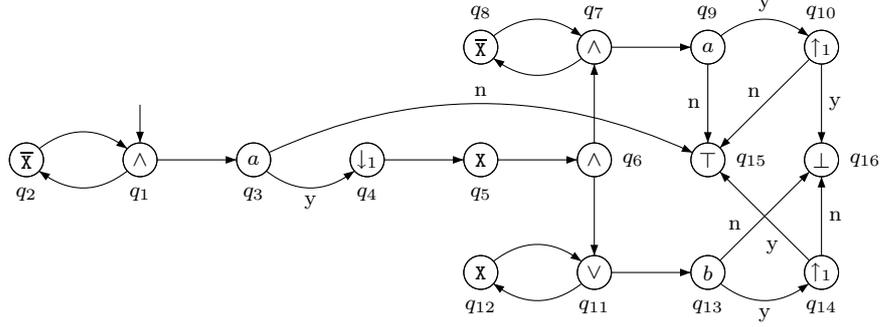


Fig. 3. A register automaton

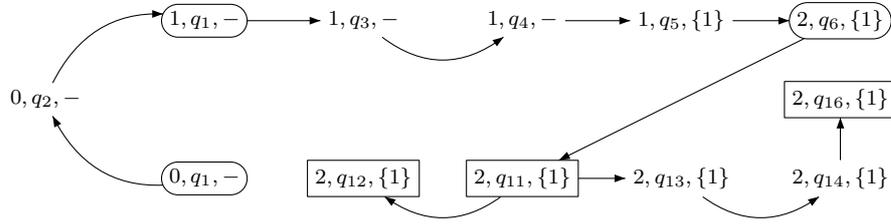


Fig. 4. A positional winning strategy

ministic or alternating automata, the maximum of the two numbers of registers suffices. Otherwise, their sum suffices.

- (c) 2URA is closed under intersection, 2NRA is closed under union, and 2ARA is closed under intersection and union. The maximum of the two numbers of registers suffices.

In each case, a required automaton is computable in logarithmic space.

PROOF. For a register automaton \mathcal{A} as above, its dual $\overline{\mathcal{A}} = \langle \Sigma, Q, q_I, n, \overline{\delta}, \overline{\rho}, \gamma \rangle$ is defined by $\overline{\delta}(q) = \delta(q)$ and $\overline{\rho}(q) = \rho(q) + 1$ for each $q \in Q$.

It suffices for (a) to show that, for every data word σ over Σ , \mathcal{A} accepts σ iff $\overline{\mathcal{A}}$ rejects σ . The latter is immediate by determinacy of weak games (cf. Theorem 2.4), and by observing that τ is a winning strategy in $G_{\mathcal{A}, \sigma}$ for player 1 from $\langle 0, q_I, \emptyset \rangle$ iff τ is a winning strategy in $G_{\overline{\mathcal{A}}, \sigma}$ for player 2 from $\langle 0, q_I, \emptyset \rangle$.

The nontrivial parts of (b) and (c) are the closures of 1NRA under intersection, 1URA under union, and 1DRA under intersection and union. By (a), we shall be done if, given $\mathcal{A}_1 = \langle \Sigma, Q_1, q_I^1, n_1, \delta_1, \rho_1, \gamma_1 \rangle$ and $\mathcal{A}_2 = \langle \Sigma, Q_2, q_I^2, n_2, \delta_2, \rho_2, \gamma_2 \rangle$ in 1NRA, we show how to compute in logarithmic space $\mathcal{A} = \langle \Sigma, Q, q_I, n_1 + n_2, \delta, \rho, \gamma \rangle$ in 1NRA which accepts a data word σ over Σ iff both \mathcal{A}_1 and \mathcal{A}_2 do, and such that \mathcal{A} is in 1DRA if both \mathcal{A}_1 and \mathcal{A}_2 are.

As in the proof of [Kaminski and Francez 1994, Theorem 3], \mathcal{A} is obtained by a product construction, so we only provide it. The locations of \mathcal{A} are pairs of locations of \mathcal{A}_1 and \mathcal{A}_2 , i.e. $Q = Q_1 \times Q_2$, and the initial location q_I is $\langle q_I^1, q_I^2 \rangle$. Transitions of \mathcal{A} will be of one of the following three kinds:

	$q'_2 \prec \beta_2 \triangleright q''_2, \downarrow_{r_2} q'_2, q'_2 \vee q''_2$	\top	\perp	Xq'_2	$\bar{X}q'_2$
$q'_1 \prec \beta_1 \triangleright q''_1,$ $\downarrow_{r_1} q'_1,$ $q'_1 \vee q''_1$	$\langle \delta_1(q_1), q_2 \rangle$	$\langle \delta_1(q_1), q_2 \rangle$	\perp	$\langle \delta_1(q_1), q_2 \rangle$	$\langle \delta_1(q_1), q_2 \rangle$
\top	$\langle q_1, \delta_2(q_2) \rangle$	\top	\perp	$\langle q_1, \delta_2(q_2) \rangle$	$\langle q_1, \delta_2(q_2) \rangle$
\perp	\perp	\perp	\perp	\perp	\perp
Xq'_1	$\langle q_1, \delta_2(q_2) \rangle$	$\langle \delta_1(q_1), q_2 \rangle$	\perp	$X\langle q'_1, q'_2 \rangle$	$X\langle q'_1, q'_2 \rangle$
$\bar{X}q'_1$	$\langle q_1, \delta_2(q_2) \rangle$	$\langle \delta_1(q_1), q_2 \rangle$	\perp	$X\langle q'_1, q'_2 \rangle$	$\bar{X}\langle q'_1, q'_2 \rangle$

 Fig. 5. Defining $\delta(\langle q_1, q_2 \rangle)$ from $\delta_1(q_1)$ (rows) and $\delta_2(q_2)$ (columns)

- a transition of one of \mathcal{A}_1 or \mathcal{A}_2 which does not change the word position;
- a transition of one of \mathcal{A}_1 or \mathcal{A}_2 which moves to the next word position, provided the other automaton has accepted;
- a pair of transitions of \mathcal{A}_1 and \mathcal{A}_2 which both move to the next word position.

For the former two kinds, we define $\langle \varphi_1, q_2 \rangle$ for a transition formula φ_1 of \mathcal{A}_1 and a location q_2 of \mathcal{A}_2 to be the transition formula of \mathcal{A} obtained by pairing with q_2 each location which occurs in φ_1 . Transition formulae $\langle q_1, \varphi_2 \rangle$ are defined similarly, where also occurrences of registers r_2 in φ_2 are replaced by $n_1 + r_2$, so e.g. $\langle q_1, q'_2 \prec \uparrow_{r_2} \triangleright q''_2 \rangle = \langle q_1, q'_2 \rangle \prec \uparrow_{n_1+r_2} \triangleright \langle q_1, q''_2 \rangle$. For each $\langle q_1, q_2 \rangle \in Q$, its transition formula is then defined in Figure 5, where the choice in the upper left-hand corner of $\langle \delta_1(q_1), q_2 \rangle$ instead of $\langle q_1, \delta_2(q_2) \rangle$ is arbitrary. The ranks are given by $\rho(\langle q_1, q_2 \rangle) = (\rho(q_1) + 1) \times (\rho(q_2) + 1) + 1$, which is even iff $\rho(q_1)$ and $\rho(q_2)$ are both even, and the heights by $\gamma(\langle q_1, q_2 \rangle) = \gamma(q_1) + \gamma(q_2)$. \square

2.6 Counter Automata

We define below two kinds of automata with counters, namely without errors and with incrementing errors, and then consider the complexity of deciding their non-emptiness, over finite and over infinite words.

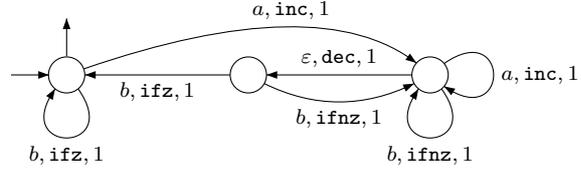
Automata. A *counter automaton* (CA) \mathcal{C} , with ε transitions and zero testing, is a tuple of the form $\langle \Sigma, Q, q_I, n, \delta, F \rangle$, where:

- Σ is a finite alphabet;
- Q is a finite set of locations;
- q_I is the initial location;
- $n \in \mathbb{N}$ is the number of counters;
- $\delta \subseteq Q \times (\Sigma \uplus \{\varepsilon\}) \times L \times Q$ is a transition relation over the instruction set $L = \{\text{inc}, \text{dec}, \text{ifz}\} \times \{1, \dots, n\}$;
- $F \subseteq Q$ is the set of accepting locations, such that $q' \notin F$ whenever $\langle q, \varepsilon, l, q' \rangle \in \delta$.

A state of \mathcal{C} is a pair $\langle q, v \rangle$ consisting of a location $q \in Q$ and a counter valuation $v : \{1, \dots, n\} \rightarrow \mathbb{N}$.

If \mathcal{C} is *Minsky* (i.e. without errors), its transitions are of the form $\langle q, v \rangle \xrightarrow{w, l} \langle q', v' \rangle$, which means that $\langle q, w, l, q' \rangle \in \delta$ and v' is obtained from v by performing instruction l in the standard manner, where $l = \langle \text{dec}, c \rangle$ requires $v(c) > 0$ and

Fig. 6. A counter automaton over finite words



$l = \langle \text{ifz}, c \rangle$ requires $v(c) = 0$. A run of \mathcal{C} is then a nonempty sequence of transitions $\langle q_0, v_0 \rangle \xrightarrow{w_0, l_0} \langle q_1, v_1 \rangle \xrightarrow{w_1, l_1} \dots$ where the initial state is given by $q_0 = q_I$ and $v_0(c) = 0$ for each c . We consider a finite run accepting iff it ends with an accepting location, and an infinite run accepting iff accepting locations occur infinitely often. We say that \mathcal{C} accepts a word w' over Σ iff there exists a run as above which is accepting and such that $w' = w_0 w_1 \dots$.

The other case we consider is when \mathcal{C} is *incrementing*, i.e. its counters may erroneously increase at any time. For counter valuations v and v_\dagger , we write $v \leq v_\dagger$ iff $v(c) \leq v_\dagger(c)$ for each c . Runs and acceptance of incrementing \mathcal{C} are defined in the same way as above, but using transitions of the form $\langle q, v \rangle \xrightarrow{w, l} \langle q', v' \rangle$, which means that there exist v_\dagger and v'_\dagger such that $v \leq v_\dagger$, $\langle q, v_\dagger \rangle \xrightarrow{w, l} \langle q', v'_\dagger \rangle$ and $v'_\dagger \leq v'$. When it is clear that we are considering an incrementing CA, we may write simply \longrightarrow instead of \longrightarrow_\dagger .

Example 2.8. Let $\mathcal{C}^{<\omega}$ be the 1-counter automaton with alphabet $\{a, b\}$ that is shown in Figure 6, and \mathcal{C}^ω the 2-counter automaton with the same alphabet that is given by Figure 7, where `ifnz` is used as syntactic sugar for a decrement succeeded by an ε increment. Since we shall consider acceptance of only finite words by $\mathcal{C}^{<\omega}$ and acceptance of only infinite words by \mathcal{C}^ω , their accepting locations are indicated in corresponding styles.

We have that $\mathcal{C}^{<\omega}$ (resp., \mathcal{C}^ω) accepts a finite (resp., infinite) word w iff every occurrence of a is followed by a separate occurrence of b , which is iff there exists a data word σ such that $\text{str}(\sigma) = w$ and which satisfies the LTL $^\dagger_1(\mathbf{X}, \mathbf{F})$ sentence ϕ from Example 2.2. That is the case regardless of whether the automata are regarded as Minsky or incrementing.

Counter automata $\mathcal{C}^{<\omega}$ and \mathcal{C}^ω were obtained from a register automaton \mathcal{A} as in Example 2.6, using the proof of Theorem 4.4 below.

Complexity of Nonemptiness. It turns out that CA with incrementing errors are easier to analyse than CA without errors.

THEOREM 2.9. (a) *For Minsky CA, nonemptiness is Σ_1^0 -complete over finite words, and Σ_1^1 -complete over infinite words.*

(b) *For incrementing CA, nonemptiness is decidable and not primitive recursive over finite words, and Π_1^0 -complete over infinite words.*

PROOF. For the finitary part of (a), we refer the reader to [Minsky 1967], and for the infinitary part, e.g. to [Alur and Henzinger 1994, Lemma 8]. Note that the lower bounds hold already with singleton alphabets, no ε transitions and 2 counters. Over finite words, those restrictions can be tightened by adding determinism, so that for each location, either there is one outgoing transition and it is an increment,

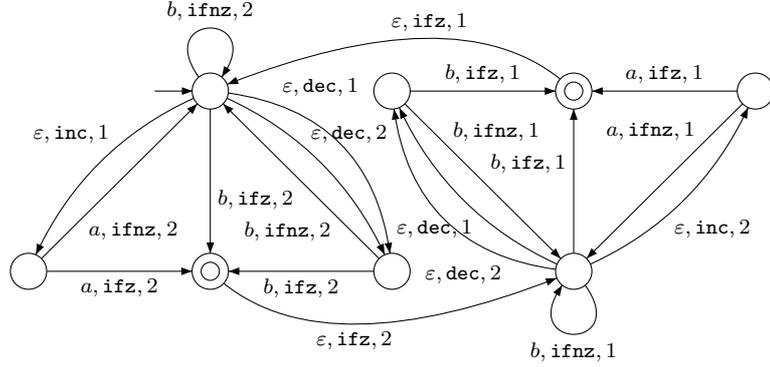


Fig. 7. A counter automaton over infinite words

or there are two outgoing transitions and they are a decrement and a zero test of the same counter.

To obtain the finitary part of (b), we observe that by reversing transition relations, there are logarithmic-space reductions between nonemptiness for incrementing CA over finite words and reachability for classic lossy counter machines [Mayr 2003]. The latter problem is indeed decidable [Mayr 2003, Theorem 6] and not primitive recursive [Schnoebelen 2002].

Incrementing CA can be seen as insertion channel machines with emptiness testing (ICMETs) [Ouaknine and Worrell 2006a] whose message sets are singletons, so Π_1^0 -membership of nonemptiness for incrementing CA over infinite words is a corollary of Π_1^0 -membership of the recurrent-state problem for ICMETs [Ouaknine and Worrell 2006b]. It can also be shown directly by considering the following procedure. Given an incrementing CA \mathcal{C} , compute a tree of all states which are reachable from the initial state. Allowing only incrementing errors which are decrements which do not alter the counter value makes the tree finitely branching. Along every branch, stop as soon as a state $\langle q, v \rangle$ is reached such that either q is accepting or some previous state on the branch is of the form $\langle q, v' \rangle$ with $v' \leq v$. By Dickson's Lemma [Dickson 1913], each branch is finite, so by König's Lemma, the computation terminates. The above is then repeated from each leaf whose location is accepting. It remains to observe that the procedure terminates iff \mathcal{C} does not have an accepting infinite run.

That nonemptiness for incrementing CA over infinite words is Π_1^0 -hard is obtained by verifying that the proof of Π_1^0 -hardness of the recurrent-state problem for ICMETs [Ouaknine and Worrell 2006a, Theorem 2] can be adapted to the more restrictive setting of incrementing CA. We reduce from emptiness over finite words for deterministic Minsky CA with singleton alphabets, no ϵ transitions and 2 counters. Given such an automaton \mathcal{C} , an incrementing CA $\hat{\mathcal{C}}$ with 5 counters C_1, C_2, C', D and D' , which performs the pseudo-code in Figure 8, and whose unique accepting location corresponds to the end of the **repeat** loop, is computable in logarithmic space. In the simulations of \mathcal{C} by $\hat{\mathcal{C}}$, counter D' prescribes how much “space” is allowed for the two counters and the number of steps of \mathcal{C} . As in the proof of [Ouaknine and Worrell 2006a, Theorem 2], we have that $\hat{\mathcal{C}}$ has an infinite accepting

```

repeat
{  $D' := D$ ;
  while  $D' > 0$ 
  { simulate  $\mathcal{C}$  using  $C_1$  and  $C_2$  as follows:
    - if  $\mathcal{C}$  accepts,  $\widehat{\mathcal{C}}$  stops
    - whenever  $C_1$  or  $C_2$  is decremented, increment  $D'$ 
    - whenever  $C_1$  or  $C_2$  is incremented, decrement  $D'$ 
    - after each step of  $\mathcal{C}$ , increment  $C'$  and decrement  $D'$ 
    - if  $D' = 0$ , exit the simulation;
       $D' = D' + C_1 + C_2 + C' - 1$ ;  $C_1, C_2, C' := 0$  };
   $D := D + 1$  }

```

Fig. 8. Pseudo-code for $\widehat{\mathcal{C}}$

run iff the unique infinite run of \mathcal{C} does not contain an accepting location. \square

2.7 Languages, Satisfiability and Nonemptiness

For a sentence ϕ of $LTL^\downarrow(\mathbf{X}, \mathbf{X}^{-1}, \mathbf{U}, \mathbf{U}^{-1})$ or $FO(\sim, <, +1, \dots, +m)$ with alphabet Σ , let $L^{<\omega}(\phi)$ (resp., $L^\omega(\phi)$) denote the language of all finite (resp., infinite) data words over Σ which satisfy ϕ . We say that ϕ is satisfiable over finite or infinite data words iff the corresponding language is nonempty.

Languages and nonemptiness of register automata and counter automata are defined analogously.

3. LTL WITH 1 REGISTER VERSUS FO WITH 2 VARIABLES

It was proved in [Etessami et al. 2002] that FO with 2 variables and predicates $<$ and $+1$ is as expressive as unary LTL (i.e., with operators \mathbf{X} , \mathbf{X}^{-1} , \mathbf{F} and \mathbf{F}^{-1}), but that in the worst case, the latter is exponentially less succinct. We now establish a similar equiexpressiveness result for $FO^2(\sim, <, +1, \dots, +m)$. First, we define the corresponding fragment of $LTL^\downarrow(\mathbf{X}, \mathbf{X}^{-1}, \mathbf{F}, \mathbf{F}^{-1})$, which has 1 register and in which scopes of the freeze quantifier are carefully restricted.

Suppose $m \in \mathbb{N}$. Let \mathcal{O}_m denote the following set of temporal operators:

$$\{\mathbf{X}, \mathbf{X}^{-1}, \dots, \mathbf{X}^m, \mathbf{X}^{-m}, \mathbf{X}^{m+1}\mathbf{F}, \mathbf{X}^{-(m+1)}\mathbf{F}^{-1}\}$$

where \mathbf{X}^k (resp., \mathbf{X}^{-k}) stands for k repetitions of \mathbf{X} (resp., \mathbf{X}^{-1}). We say that a formula of $LTL_1^\downarrow(\mathcal{O}_m)$ is *simple* iff each occurrence of a temporal operator is immediately preceded by \downarrow_1 (and there are no other occurrences of \downarrow_1).

A sentence ϕ of $LTL^\downarrow(\mathbf{X}, \mathbf{X}^{-1}, \mathbf{U}, \mathbf{U}^{-1})$ is said to be equivalent to a formula $\phi'(x)$ of $FO(\sim, <, +1, \dots, +m)$ iff they have the same alphabet Σ and, for every data word σ over Σ and $0 \leq i < |\sigma|$, we have $\sigma, i \models_\emptyset \phi \Leftrightarrow \sigma \models_{[x \mapsto i]} \phi'(x)$.

Example 3.1. It is straightforward to rewrite the sentence ϕ from Example 2.2 so that it belongs to simple $LTL_1^\downarrow(\mathcal{O}_0)$. Alternatively, that ϕ is equivalent to a sentence of simple $LTL_1^\downarrow(\mathcal{O}_0)$ is a consequence of the following theorem, since it was observed in Example 2.3 that ϕ is equivalent to the formula $\phi'(x_0)$ of $FO^2(\sim, <)$.

THEOREM 3.2. (a) For each sentence of simple $LTL_1^\downarrow(\mathcal{O}_m)$, an equivalent formula of $FO^2(\sim, <, +1, \dots, +m)$ is computable in logarithmic space.

(b) For each formula $\phi(x_j)$ of $FO^2(\sim, <, +1, \dots, +m)$, an equivalent sentence of simple $LTL_1^\downarrow(\mathcal{O}_m)$ is computable in polynomial space.

PROOF. The following notations will be convenient. Let $\mathbf{0}^0 = \downarrow_1$, $\mathbf{0}^k = \downarrow_1 \mathbf{X}^k$ for $k \in \{-m, \dots, -1, 1, \dots, m\}$, $\mathbf{0}^{m+1} = \downarrow_1 \mathbf{X}^{m+1} \mathbf{F}$, and $\mathbf{0}^{-(m+1)} = \downarrow_1 \mathbf{X}^{-(m+1)} \mathbf{F}^{-1}$. For $j \in \{0, 1\}$, let

$$\begin{aligned} \chi_0^j &\stackrel{\text{def}}{=} x_{1-j} = x_j \\ \chi_k^j &\stackrel{\text{def}}{=} x_{1-j} = x_j + k \quad (1 \leq k \leq m) \\ \chi_{-k}^j &\stackrel{\text{def}}{=} x_j = x_{1-j} + k \quad (1 \leq k \leq m) \\ \chi_{m+1}^j &\stackrel{\text{def}}{=} x_j < x_{1-j} \wedge \bigwedge_{1 \leq k \leq m} \neg x_{1-j} = x_j + k \\ \chi_{-(m+1)}^j &\stackrel{\text{def}}{=} x_{1-j} < x_j \wedge \bigwedge_{1 \leq k \leq m} \neg x_j = x_{1-j} + k \end{aligned}$$

(The equality predicate can be expressed using $<$.)

We have (a) by the following translations T_j which map simple $\text{LTL}_1^\downarrow(\mathcal{O}_m)$ formulae to $\text{FO}^2(\sim, <, +1, \dots, +m)$ formulae. Each sentence ϕ will be equivalent to $T_j(\phi)$ which will contain at most x_j free. The maps T_j are defined by structural recursion, by encoding the semantics of simple formulae into first-order logic, and by recycling variables (to use only two variables). The Boolean clauses are omitted.

$$T_j(a) \stackrel{\text{def}}{=} P_a(x_j) \quad T_j(\uparrow_1) \stackrel{\text{def}}{=} x_{1-j} \sim x_j \quad T_j(\mathbf{0}^k \psi) \stackrel{\text{def}}{=} \exists x_{1-j} (\chi_k^j \wedge T_{1-j}(\psi))$$

For (b), we proceed by adapting the proof of [Etesami et al. 2002, Theorem 1]. We define recursively translations T'_j from $\text{FO}^2(\sim, <, +1, \dots, +m)$ formulae $\phi(x_j)$ to equivalent simple $\text{LTL}_1^\downarrow(\mathcal{O}_m)$ sentences. The cases of Boolean operators and one-variable atomic formulae are straightforward. The remaining case is when $\phi(x_j)$ is of the form

$$\exists x_{1-j} \beta(\alpha_1(x_0, x_1), \dots, \alpha_L(x_0, x_1), \xi_1(x_j), \dots, \xi_N(x_j), \zeta_1(x_{1-j}), \dots, \zeta_M(x_{1-j}))$$

where β is a Boolean formula, and each $\alpha_i(x_0, x_1)$ is a $\sim, <$ or $+k$ atomic formula. Now, for each $-(m+1) \leq k \leq m+1$ and $b \in \{\top, \perp\}$, let $\alpha_i^{k,b}$ denote the truth value of $\alpha_i(x_0, x_1)$ under assumptions χ_k^j and $x_j \sim x_{1-j} \Leftrightarrow b$. Also, for each $X \subseteq \{1, \dots, N\}$, let $\xi_i^X = \top$ if $i \in X$, and $\xi_i^X = \perp$ otherwise. $T'_j(\phi(x_j))$ is then computed as

$$\bigvee_{-(m+1) \leq k \leq m+1} \bigvee_{b \in \{\top, \perp\}} \bigvee_{X \subseteq \{1, \dots, N\}} \left(\bigwedge_{i \in \{1, \dots, N\}} T'_j(\xi_i(x_j)) \Leftrightarrow \xi_i^X \right) \wedge \mathbf{0}^k((\uparrow_1 \Leftrightarrow b) \wedge \beta(\alpha_1^{k,b}, \dots, \alpha_L^{k,b}, \xi_1^X, \dots, \xi_N^X, T'_{1-j}(\zeta_1(x_{1-j})), \dots, T'_{1-j}(\zeta_M(x_{1-j}))))$$

The size of the equivalent simple $\text{LTL}_1^\downarrow(\mathcal{O}_m)$ formula is exponential in $|\phi|$, because the length of the stack of recursive calls is linear and the generalised conjunctions and disjunctions have at most exponentially many arguments. For the same reasons, polynomial space is sufficient for the computation. \square

It was shown in [Bojańczyk et al. 2006] that, over finite data words, satisfiability for $\text{FO}^2(\sim, <, +1)$ is reducible in doubly exponential time to reachability for Petri nets, and that there is a polynomial-time reduction in the reverse direction. Reachability for Petri nets is known to be decidable (cf. e.g. [Kosaraju 1982]) and EXPSPACE -hard [Lipton 1976]. Two extensions of the decidability of satisfiability for $\text{FO}^2(\sim, <, +1)$ over finite data words were also obtained in [Bojańczyk et al. 2006]: for $\text{FO}^2(\sim, <, +1, \dots, +m)$, and over infinite data words. By the following corollary of Theorem 3.2, those results have immediate consequences for complexity of satisfiability problems for simple fragments of $\text{LTL}_1^\downarrow(\mathcal{O}_m)$.

COROLLARY 3.3. *Over finite and over infinite data words, satisfiability for the simple fragment of $LTL_1^{\downarrow}(\mathcal{O}_m)$ is reducible in logarithmic space to satisfiability for $FO^2(\sim, <, +1, \dots, +m)$, and there is a polynomial-space reduction in the reverse direction.*

4. UPPER COMPLEXITY BOUNDS

A number of upper bounds on complexity of satisfiability for fragments of LTL with the freeze quantifier and complexity of nonemptiness for classes of register automata will be obtained below. The former will be corollaries of the latter, by the following result which shows that logical sentences are easily translatable to equivalent automata. Note that, in contrast to classical automata, alternating register automata are more expressive than nondeterministic and universal, and two-way register automata are more expressive than one-way [Kaminski and Francez 1994; Neven et al. 2004]. As a specific example, the “nonces property” that no two word positions are in the same class is expressible in $LTL_1^{\downarrow}(\mathbf{X}, \mathbf{F})$ (cf. Example 2.2) and by an automaton in $1ARA_1$, but not by any automaton in $1NRA$.

THEOREM 4.1. *For each sentence ϕ of $LTL_n^{\downarrow}(\mathbf{X}, \mathbf{X}^{-1}, \mathbf{U}, \mathbf{U}^{-1})$, an automaton \mathcal{A}_ϕ in $2ARA_n$ with the same alphabet and such that $L^{<\omega}(\phi) = L^{<\omega}(\mathcal{A}_\phi)$ and $L^\omega(\phi) = L^\omega(\mathcal{A}_\phi)$ is computable in logarithmic space. If ϕ is in $LTL_n^{\downarrow}(\mathbf{X}, \mathbf{U})$, \mathcal{A}_ϕ is in $1ARA_n$.*

PROOF. The translation is a simple extension of the classical one from LTL to alternating automata (cf. e.g. [Vardi 1996]).

We can assume that ϕ is in negation normal form, where we write \bar{a} , \perp , \vee , $\bar{0}$ and \uparrow_r for the duals of a , \top , \wedge , 0 in $\{\mathbf{X}, \mathbf{X}^{-1}, \mathbf{U}, \mathbf{U}^{-1}\}$ and \uparrow_r .

Let $\text{cl}(\phi)$ be the set of all subformulae of ϕ , together with \top , \perp , and all subformulae of $\psi \wedge \mathbf{X}(\psi \mathbf{U} \chi)$, $\psi \wedge \mathbf{X}^{-1}(\psi \mathbf{U}^{-1} \chi)$, $\psi \vee \bar{\mathbf{X}}(\psi \bar{\mathbf{U}} \chi)$ or $\psi \vee \bar{\mathbf{X}}^{-1}(\psi \bar{\mathbf{U}}^{-1} \chi)$ for each subformula of ϕ of the form $\psi \mathbf{U} \chi$, $\psi \mathbf{U}^{-1} \chi$, $\psi \bar{\mathbf{U}} \chi$ or $\psi \bar{\mathbf{U}}^{-1} \chi$ (respectively).

To define $\mathcal{A}_\phi = \langle \Sigma, Q, q_I, n, \delta, \rho, \gamma \rangle$, let $Q = \{q_\psi : \psi \in \text{cl}(\phi)\}$ and $q_I = q_\phi$.

The transition function is given below, where we omit dual cases:

$$\begin{array}{lll} \delta(q_a) = q_\top \not\leftarrow a \not\rightarrow q_\perp & \delta(q_\top) = \top & \delta(q_{\psi \wedge \chi}) = q_\psi \wedge q_\chi \\ \delta(q_{\uparrow_r}) = q_\top \not\leftarrow \uparrow_r \not\rightarrow q_\perp & \delta(q_{\mathbf{X}\psi}) = \mathbf{X}q_\psi & \delta(q_{\psi \mathbf{U} \chi}) = q_\chi \vee q_{\psi \wedge \mathbf{X}(\psi \mathbf{U} \chi)} \\ \delta(q_{\downarrow_r \psi}) = \downarrow_r q_\psi & \delta(q_{\mathbf{X}^{-1}\psi}) = \mathbf{X}^{-1}q_\psi & \delta(q_{\psi \mathbf{U}^{-1} \chi}) = q_\chi \vee q_{\psi \wedge \mathbf{X}^{-1}(\psi \mathbf{U}^{-1} \chi)} \end{array}$$

The ranks are defined so that every $q_{\chi \mathbf{U} \chi'}$ has odd rank and every $q_{\chi \bar{\mathbf{U}} \chi'}$ has even rank. For example, $\rho(q_\psi) = 2|\psi|$ unless ψ is of the form $\chi \mathbf{U} \chi'$, in which case $\rho(q_\psi) = 2|\psi| + 1$.

The heights may be defined as $\gamma(q_\psi) = |\psi|$ unless ψ is of the form $\chi \mathbf{U} \chi'$, $\chi \mathbf{U}^{-1} \chi'$, $\chi \bar{\mathbf{U}} \chi'$ or $\chi \bar{\mathbf{U}}^{-1} \chi'$, in which case $\gamma(q_{\chi \mathbf{U} \chi'}) = |\chi' \vee (\chi \wedge \mathbf{X}(\chi \mathbf{U} \chi'))|$ etc.

It remains to show the equalities between the languages of ϕ and \mathcal{A}_ϕ , so suppose σ is a data word over Σ . By a straightforward induction on $\psi \in \text{cl}(\phi)$, it holds that, for every $0 \leq i < |\sigma|$ and n -register valuation v for σ , we have $\sigma, i \models_v \psi$ iff player 1 has a winning strategy from state $\langle i, q_\psi, v \rangle$ in game $G_{\mathcal{A}_\phi, \sigma}$. In particular, $\sigma, 0 \models_\emptyset \phi$ iff \mathcal{A}_ϕ accepts σ . \square

The following basic upper bounds should be compared with the lower bounds in Theorem 5.4.

THEOREM 4.2. *Over finite data words, satisfiability for $LTL^\downarrow(\mathbf{x}, \mathbf{x}^{-1}, \mathbf{u}, \mathbf{u}^{-1})$ and nonemptiness for 2ARA are in Σ_1^0 . Over infinite data words, satisfiability for $LTL^\downarrow(\mathbf{x}, \mathbf{x}^{-1}, \mathbf{u}, \mathbf{u}^{-1})$ and nonemptiness for 2ARA are in Σ_1^1 , and nonemptiness for 2NRA is in Σ_2^0 .*

PROOF. By Theorem 4.1, it suffices to consider the register automata nonemptiness problems.

That nonemptiness for 2ARA is in Σ_1^0 over finite data words and in Σ_1^1 over infinite data words are straightforward consequences of Theorem 2.4.

Suppose \mathcal{A} is in 2NRA. Because of nondeterminism, \mathcal{A} accepts a data word iff there exists a complete play from the initial state in the acceptance game which is winning for player 1. By König’s Lemma, we have that \mathcal{A} accepts an infinite data word iff there exists $j \in \mathbb{N}$ such that for each $k \in \mathbb{N}$:

- (*) there exist a data word σ of length $k + 1$ and a play $\pi = p_0 p_1 \dots$ of length at most k from the initial state in $G_{\mathcal{A}, \sigma}$ such that either π is winning for player 1, or π is of length k and $\rho(p_{j'})$ is even for all $j' \geq j$.

The Σ_2^0 -membership follows by observing that (*) is decidable. \square

It was shown in [Sakamoto and Ikeda 2000] that nonemptiness for one-way nondeterministic register automata over finite data words is in NP, and that the problem is NP-hard already for deterministic automata. However, due to the technical differences noted in Remark 2.5, the complexity for the notion of register automata in this paper turns out to be PSPACE-complete, over infinite data words as well. The proof below of the PSPACE-memberships will also prepare us for establishing Theorem 4.4. The hardness results are in Theorem 5.1.

We remark that, for infinite data words, it is straightforward to extend Theorems 4.3 and 4.4 to register automata with Büchi acceptance, without affecting the complexity bounds.

THEOREM 4.3. *The following hold over finite and over infinite data words:*

- nonemptiness for 1NRA is in PSPACE;
- for every fixed $n \in \mathbb{N}$, nonemptiness for $1NRA_n$ is in NLOGSPACE.

PROOF. Suppose $\mathcal{A} = \langle \Sigma, Q, q_I, n, \delta, \rho, \gamma \rangle$ is in 1NRA.

Let $H_{\mathcal{A}}$ be the set of all “abstract states” of the form $\langle a, ee, R, q, E \rangle$ where $a \in \Sigma$, $ee \in \{\top, \perp\}$, R is either \emptyset or a class of E , $q \in Q$ and E is an equivalence relation on a subset of $\{1, \dots, n\}$. For a data word σ (over Σ), let $\alpha_{\mathcal{A}, \sigma}$ be the following mapping from states of \mathcal{A} (for σ) to elements of $H_{\mathcal{A}}$:

$$\alpha_{\mathcal{A}, \sigma}(\langle i, q, v \rangle) = \langle \sigma(i), i + 1 = |\sigma|, \{r : v(r) = [i]_{\sim}\}, q, \{ \langle r, r' \rangle : r, r' \in \text{dom}(v) \text{ and } v(r) = v(r') \} \rangle$$

We define a relation \rightarrow on $H_{\mathcal{A}}$ by $h \rightarrow h'$ iff there exist a data word σ and states p and p' of \mathcal{A} such that $\alpha_{\mathcal{A}, \sigma}(p) = h$, $\alpha_{\mathcal{A}, \sigma}(p') = h'$ and $p \rightarrow p'$. We say that $h \in H_{\mathcal{A}}$ is *initial* (resp., *winning*) iff for some (equivalently, for every) data word σ and state p of \mathcal{A} such that $\alpha_{\mathcal{A}, \sigma}(p) = h$, we have that p is initial (resp., has no successors and belongs to player 2).

Because of nondeterminism, \mathcal{A} accepts a data word σ iff there exists a complete play from the initial state in $G_{\mathcal{A}, \sigma}$ which is winning for player 1. Since \mathcal{A} is one-way,

for every sequence $h_0 \rightarrow h_1 \rightarrow \dots$ in $H_{\mathcal{A}}$ with h_0 initial, there exist a data word σ and a play $p_0 p_1 \dots$ from the initial state in $G_{\mathcal{A},\sigma}$ such that $\alpha_{\mathcal{A},\sigma}(p_j) = h_j$ for all j . Consequently:

- $L^{<\omega}(\mathcal{A})$ is nonempty iff there exists a sequence $h_0 \rightarrow h_1 \rightarrow \dots h_k$ in $H_{\mathcal{A}}$ such that h_0 is initial and h_k is winning;
- $L^{\omega}(\mathcal{A})$ is nonempty iff:
 - either there exists a sequence $h_0 \rightarrow h_1 \rightarrow \dots h_k$ in $H_{\mathcal{A}}$ such that h_0 is initial, h_k is winning, and the second component of h_k is \perp ,
 - or there exists a sequence $h_0 \rightarrow h_1 \rightarrow \dots h_k \rightarrow h_{k+1} \rightarrow \dots h_{k'}$ in $H_{\mathcal{A}}$ such that h_0 is initial, $h_k = h_{k'}$, and the rank of the location in h_k and $h_{k'}$ is even.

It remains to observe that, for storing an abstract state and for checking the successor relation and the initial and winning properties on abstract states, space which is logarithmic in $|\Sigma|$ and $|Q|$ and polynomial in n suffices. \square

In terms of the definitions in this paper, it was established in [Kaminski and Francez 1994, Appendix A] that containment of the language of an automaton in 1NRA in the language of an automaton in 1NRA₁ is decidable over finite data words. In particular, nonemptiness for 1URA₁ (see Theorem 2.7) over finite data words is decidable. We now prove the main result of this section, which shows that decidability in fact holds for 1ARA₁, and therefore also for $LTL_1^{\downarrow}(X, U)$ satisfiability. The two problems over infinite data words are shown to be co-r.e. The proof is by reductions to nonemptiness of incrementing counter automata, which will provide the first half of a correspondence between languages of incrementing CA and sentences of future-time fragments of LTL with 1 register (see Corollary 5.3).

Using the developments in the proof of Theorem 4.3, it is straightforward to extend the argument below to obtain the same upper bounds for containment of the language of an automaton in 1NRA in the language of an automaton in 1ARA₁, or in the language of a sentence in $LTL_1^{\downarrow}(X, U)$. Over infinite data words, extending further to one-way nondeterministic register automata with Büchi acceptance requires no extra work.

Theorem 5.4 will show that decidability and Π_1^0 -membership break down as soon as any of 1 more register, past-time operators or backward moves are added.

THEOREM 4.4. *Satisfiability for $LTL_1^{\downarrow}(X, U)$ and nonemptiness for 1ARA₁ are decidable over finite data words, and in Π_1^0 over infinite data words.*

PROOF. By Theorems 4.1 and 2.9 (b), it suffices to show that, given \mathcal{A} in 1ARA₁, incrementing CA $\mathcal{C}_{\mathcal{A}}^{<\omega}$ and $\mathcal{C}_{\mathcal{A}}^{\omega}$ are computable such that

$$L^{<\omega}(\mathcal{C}_{\mathcal{A}}^{<\omega}) = \{\text{str}(\sigma) : \sigma \in L^{<\omega}(\mathcal{A})\} \quad L^{\omega}(\mathcal{C}_{\mathcal{A}}^{\omega}) = \{\text{str}(\sigma) : \sigma \in L^{\omega}(\mathcal{A})\}$$

In both cases, the proof will consist of the following steps:

- replace the two-player acceptance games for \mathcal{A} by one-player games whose positions are built from sets of states of \mathcal{A} (\mathcal{A} cannot in general be translated to an automaton in 1NRA: see the remarks which precede Theorem 4.1), and whose successors are “big step” in the sense that they correspond to following strategies for the automaton until first moves to the next word position;

- combine the one-player acceptance games with searching for a data word to be accepted, resulting in a one-player nonemptiness game for \mathcal{A} ;
- show how to construct a CA which guesses and checks a winning play in the nonemptiness game;
- show that allowing incrementing errors in computations of the CA does not increase its language (such errors in an accepting computation will amount to introducing superfluous states of \mathcal{A} from which winning strategies for the automaton are then found).

First, we consider computing $\mathcal{C}_{\mathcal{A}}^{<\omega}$. Let $\mathcal{A} = \langle \Sigma, Q, q_I, 1, \delta, \rho, \gamma \rangle$.

To define a big-step successor relation between sets of states, for a state p and a set of states P' of \mathcal{A} for a data word σ over Σ , let us write $p \Rightarrow P'$ iff there exists a strategy τ for player 1 from p in game $G_{\mathcal{A}, \sigma}$ such that:

- each complete play $\pi \in \tau$ which contains no move to another word position is winning for player 1;
- P' is the set of all targets of first moves to another word position in plays of τ .

For sets of states $P \neq \emptyset$ and P' , we write $P \Rightarrow P'$ iff there exists a map $p \mapsto P'_p$ on P such that $p \Rightarrow P'_p$ for each $p \in P$ and $P' = \bigcup_{p \in P} P'_p$. Since \mathcal{A} is one-way, if the first component of every state in P is i and $P \Rightarrow P'$, then the first component of every state in P' is $i + 1$. We call a set of states *unipositional* iff its members have the same first components.

By positional determinacy of weak games (see Theorem 2.4), the decreasing heights discipline of register automata, and König's Lemma, we have:

- (I) for every finite data word σ over Σ , \mathcal{A} accepts σ iff there exists a sequence P_1, \dots, P_{k-1} of sets of states of \mathcal{A} for σ such that $\{\langle 0, q_I, \emptyset \rangle\} \Rightarrow P_1 \Rightarrow \dots \Rightarrow P_{k-1} \Rightarrow \emptyset$.

Now, let $H_{\mathcal{A}}$ consist of \emptyset and all “abstract sets” of the form $\langle a, ee, Q_-, Q_\emptyset, \sharp \rangle$ where $a \in \Sigma$, $ee \in \{\top, \perp\}$, $Q_-, Q_\emptyset \subseteq Q$, $\sharp : \mathcal{P}(Q) \setminus \{\emptyset\} \rightarrow \mathbb{N}$, and either $Q_- \neq \emptyset$ or $Q_\emptyset \neq \emptyset$ or $\sharp(Q_\dagger) > 0$ for some Q_\dagger . We define a mapping $\alpha_{\mathcal{A}, \sigma}$ from unipositional sets of states of \mathcal{A} for σ to elements of $H_{\mathcal{A}}$ as follows: $\alpha_{\mathcal{A}, \sigma}(\emptyset) = \emptyset$, and for nonempty P whose members' first component is i ,

$$\alpha_{\mathcal{A}, \sigma}(P) = \langle \sigma(i), i + 1 = |\sigma|, \{q : \langle i, q, [1 \mapsto [i]_{\sim}] \rangle \in P\}, \{q : \langle i, q, \emptyset \rangle \in P\}, \\ Q_\dagger \mapsto |\{D \neq [i]_{\sim} : \{q : \langle i, q, [1 \mapsto D] \rangle \in P\} = Q_\dagger\}| \rangle$$

In particular, the last component of the abstract set $\alpha_{\mathcal{A}, \sigma}(P)$ maps each nonempty $Q_\dagger \subseteq Q$ to the number of distinct data D which are not the class of i and such that the set of all q with $\langle i, q, [1 \mapsto D] \rangle \in P$ equals Q_\dagger .

As the first step to defining a big-step successor relation between abstract sets, for $a \in \Sigma$, $ee, uu \in \{\top, \perp\}$ and $q \in Q$, let $\langle\langle a, ee, uu, q \rangle\rangle$ be the set of pairs of sets of locations that is defined in Figure 9 by recursion over the height of q . (Observe that $Q'_{\neq} = \emptyset$ whenever $\langle Q'_{\neq}, Q'_{=} \rangle \in \langle\langle a, ee, \top, q \rangle\rangle$.) Those sets satisfy:

- (II) for every data word σ over Σ , state $\langle i, q, v \rangle$ and set of states P' of \mathcal{A} for σ , we have $\langle i, q, v \rangle \Rightarrow P'$ iff there exists

$$\langle Q'_{\neq}, Q'_{=} \rangle \in \langle\langle \sigma(i), i + 1 = |\sigma|, v = [1 \mapsto [i]_{\sim}], q \rangle\rangle$$

such that $P' = \{\langle i + 1, q', v \rangle : q' \in Q'_{\neq}\} \cup \{\langle i + 1, q', [1 \mapsto [i]_{\sim}] \rangle : q' \in Q'_{=}\}$.

$\delta(q)$	$\langle\langle a, ee, uu, q \rangle\rangle$
$q' \not\prec \beta \succ q''$	$\langle\langle a, ee, uu, q' \rangle\rangle$, if $a, ee, uu \models \beta$ $\langle\langle a, ee, uu, q'' \rangle\rangle$, if $a, ee, uu \not\models \beta$
$\downarrow_1 q'$	$\langle\langle a, ee, \top, q' \rangle\rangle$
$q' \wedge q''$	$\{\langle Q'_\neq \cup Q''_\neq, Q'_= \cup Q''_= \rangle : \langle Q'_\neq, Q'_= \rangle \in \langle\langle a, ee, uu, q' \rangle\rangle, \langle Q''_\neq, Q''_= \rangle \in \langle\langle a, ee, uu, q'' \rangle\rangle\}$
$q' \vee q''$	$\langle\langle a, ee, uu, q' \rangle\rangle \cup \langle\langle a, ee, uu, q'' \rangle\rangle$
\top	$\{\langle \emptyset, \emptyset \rangle\}$
\perp	\emptyset
$\mathbb{X}q'$	$\{\langle \{q'\}, \emptyset \rangle\}$, if $ee = \perp$ and $uu = \perp$ $\{\langle \emptyset, \{q'\} \rangle\}$, if $ee = \perp$ and $uu = \top$ \emptyset , if $ee = \top$
$\bar{\mathbb{X}}q'$	$\{\langle \{q'\}, \emptyset \rangle\}$, if $ee = \perp$ and $uu = \perp$ $\{\langle \emptyset, \{q'\} \rangle\}$, if $ee = \perp$ and $uu = \top$ $\{\langle \emptyset, \emptyset \rangle\}$, if $ee = \top$

$$a, ee, uu \models a' \stackrel{\text{def}}{\Leftrightarrow} a = a' \quad a, ee, uu \models \text{end} \stackrel{\text{def}}{\Leftrightarrow} ee = \top \quad a, ee, uu \models \uparrow_1 \stackrel{\text{def}}{\Leftrightarrow} uu = \top$$

Fig. 9. Defining abstract big-step successors

The following notations will be useful: given a map $f : X \rightarrow \mathcal{P}(Y_1) \times \mathcal{P}(Y_2)$, let

$$\bigcup_1 f = \bigcup \{Z_1 : \langle Z_1, Z_2 \rangle \in f(X)\} \quad \bigcup_2 f = \bigcup \{Z_2 : \langle Z_1, Z_2 \rangle \in f(X)\}$$

For $h, h' \in H_{\mathcal{A}}$, we write $h \Rightarrow h'$ iff h is of the form $\langle a, ee, Q_=:, Q_\emptyset, \# \rangle$ and there exist maps $q \in Q_=: \mapsto f_=(q) \in \langle\langle a, ee, \top, q \rangle\rangle$, $q \in Q_\emptyset \mapsto f_\emptyset(q) \in \langle\langle a, ee, \perp, q \rangle\rangle$ and $q \in Q_\dagger \mapsto f_{Q_\dagger, j}(q) \in \langle\langle a, ee, \perp, q \rangle\rangle$ for each nonempty $Q_\dagger \subseteq Q$ and $j \in \{1, \dots, \#(Q_\dagger)\}$ such that:

- either $h' = \emptyset$, $\bigcup_1 f_\emptyset = \emptyset$, and $\#'(Q'_\dagger) = 0$ for all Q'_\dagger ,
- or h' is of the form $\langle a', ee', \emptyset, \bigcup_1 f_\emptyset, \#' \rangle$,
- or h' is of the form $\langle a', ee', Q'_=:, \bigcup_1 f_\emptyset, \#'[Q'_=: \mapsto \#'(Q'_=:) - 1] \rangle$,

where, for each nonempty $Q'_\dagger \subseteq Q$, $\#'(Q'_\dagger)$ is defined as

$$|\{\langle Q'_\dagger, j \rangle : \bigcup_1 f_{Q'_\dagger, j} = Q'_\dagger\}| + \begin{cases} 1, & \text{if } \bigcup_2 f_=: \cup \bigcup_2 f_\emptyset \cup \bigcup_{Q'_\dagger, j} \bigcup_2 f_{Q'_\dagger, j} = Q'_\dagger \\ 0, & \text{otherwise} \end{cases}$$

From (II), it follows that:

- (III) for every data word σ over Σ , and unipositional sets P and P' of states of \mathcal{A} for σ , we have $P \Rightarrow P'$ iff $\alpha_{\mathcal{A}, \sigma}(P) \Rightarrow \alpha_{\mathcal{A}, \sigma}(P')$.

By (I) and (III), it suffices to compute an incrementing CA $\mathcal{C}_{\mathcal{A}}^{<\omega}$ for which:

- (IV) $\mathcal{C}_{\mathcal{A}}^{<\omega}$ accepts $a_0 \dots a_{l-1} \in \Sigma^{<\omega}$ iff there exists a sequence $h_0 \Rightarrow \dots \Rightarrow h_{k-1} \Rightarrow \emptyset$ of elements of $H_{\mathcal{A}}$ such that:
- h_0 is of the form $\langle a_0, ee, \emptyset, \{q_I\}, Q_\dagger \mapsto 0 \rangle$;
 - for each $0 < i < k$, the first component of h_i is a_i ;
 - $k = l$ if the second component of h_{k-1} is \top .

$\mathcal{C}_{\mathcal{A}}^{<\omega}$ is constructed so that it guesses and checks a sequence $h_0 \Rightarrow \dots \Rightarrow h_{k-1} \Rightarrow \emptyset$ as in (IV), storing at most two consecutive members in any state. To store an abstract set $\langle a, ee, Q_=:, Q_\emptyset, \# \rangle$, locations of $\mathcal{C}_{\mathcal{A}}^{<\omega}$ are used for the first four components, and $\#$ is stored by means of $2^{|Q|} - 1$ counters c_{Q_\dagger} for $\emptyset \neq Q_\dagger \subseteq Q$. $\mathcal{C}_{\mathcal{A}}^{<\omega}$ also has $4^{|Q|}$ auxiliary counters $c_{Q'_\neq, Q'_=}$ for $Q'_\neq, Q'_= \subseteq Q$. The nontrivial part of $\mathcal{C}_{\mathcal{A}}^{<\omega}$ is, given a,

```

for all  $\emptyset \neq Q_{\dagger} \subseteq Q$ 
{ while  $c_{Q_{\dagger}} > 0$ 
  { choose a map  $q \in Q_{\dagger} \mapsto f(q) \in \langle\langle a, ee, \perp, q \rangle\rangle$ ;
     $\text{dec}(c_{Q_{\dagger}})$ ;  $\text{inc}(c_{\bigcup_1 f, \bigcup_2 f})$  } };
choose a map  $q \in Q_{=} \mapsto f_{=}(q) \in \langle\langle a, ee, \top, q \rangle\rangle$ ;
choose a map  $q \in Q_{\emptyset} \mapsto f_{\emptyset}(q) \in \langle\langle a, ee, \perp, q \rangle\rangle$ ;
 $Q_{\ddagger} := \bigcup_2 f_{=} \cup \bigcup_2 f_{\emptyset}$ ;
for all  $Q'_{\neq}, Q'_{=} \subseteq Q$ 
{ if  $c_{Q'_{\neq}, Q'_{=}} > 0$  then  $Q_{\ddagger} := Q_{\ddagger} \cup Q'_{=}$  };
if  $Q_{\ddagger} \neq \emptyset$  then  $\text{inc}(c_{Q_{\ddagger}})$ ;
for all  $Q'_{\neq}, Q'_{=} \subseteq Q$ 
{ while  $c_{Q'_{\neq}, Q'_{=}} > 0$ 
  {  $\text{dec}(c_{Q'_{\neq}, Q'_{=}})$ ; if  $Q'_{\neq} \neq \emptyset$  then  $\text{inc}(c_{Q'_{\neq}})$  } }

```

Fig. 10. Computing an abstract big-step successor

ee , $Q_{=}$, Q_{\emptyset} , and $\#$ which is stored by the counters $c_{Q_{\dagger}}$, to guess maps $f_{=}$, f_{\emptyset} and $f_{Q_{\dagger}, j}$, and set the counters $c_{Q_{\dagger}}$ so that they store $\#'$ as defined above. Figure 10 contains pseudo-code for that computation, which is by ε transitions. It assumes that each $c_{Q'_{\neq}, Q'_{=}}$ is zero at the beginning, and ensures that the same holds at the end. The choices of maps are nondeterministic. If a map cannot be chosen because a corresponding set $\langle\langle a, ee, uu, q \rangle\rangle$ is empty, the computation blocks.

Suppose $a_0 \dots a_{l-1} \in \Sigma^{<\omega}$. If there exists a sequence $h_0 \Rightarrow \dots h_{k-1} \Rightarrow \emptyset$ as in (IV), the construction of $\mathcal{C}_{\mathcal{A}}^{<\omega}$ ensures that it accepts $a_0 \dots a_{l-1}$ by a run without incrementing errors. For the reverse direction in (IV), let $\langle a, ee, Q_{=}, Q_{\emptyset}, \# \rangle \sqsubseteq \langle a', ee', Q'_{=}, Q'_{\emptyset}, \#' \rangle$ mean that $a = a'$, $ee = ee'$, $Q_{=} \subseteq Q'_{=}$, $Q_{\emptyset} \subseteq Q'_{\emptyset}$ and there exists an injective

$$\iota : \{ \langle Q_{\dagger}, j \rangle : j \in \{1, \dots, \#(Q_{\dagger})\} \} \rightarrow \{ \langle Q'_{\dagger}, j' \rangle : j' \in \{1, \dots, \#'(Q'_{\dagger})\} \}$$

for which $Q_{\dagger} \subseteq Q'_{\dagger}$ whenever $\iota(\langle Q_{\dagger}, j \rangle) = \langle Q'_{\dagger}, j' \rangle$. Also, let $\emptyset \sqsubseteq h'$ for all $h' \in H_{\mathcal{A}}$. If $\mathcal{C}_{\mathcal{A}}^{<\omega}$ accepts $a_0 \dots a_{l-1}$ (by a run possibly with incrementing errors), we have that there exist $h'_0, \dots, h'_{k-1} \in H_{\mathcal{A}}$ such that:

- h'_0 is of the form $\langle a_0, ee, \emptyset, \{q_I\}, Q_{\dagger} \mapsto 0 \rangle$;
- for each $0 < i < k$, the second component of h'_{i-1} is \perp , the first component of h'_i is a_i , and $h'_{i-1} \Rightarrow h_i$ for some $h_i \sqsubseteq h'_i$;
- $h'_{k-1} \Rightarrow \emptyset$, and $k = l$ if the second component of h'_{k-1} is \top .

It remains to observe that \sqsubseteq is transitive, and downwards compatible with \Rightarrow , i.e. whenever $\emptyset \neq h \sqsubseteq h'$ and $h' \Rightarrow h'_*$, there exists h_* such that $h \Rightarrow h_*$ and $h_* \sqsubseteq h'_*$.

The computation of $\mathcal{C}_{\mathcal{A}}^{\omega}$ follows the same pattern, except that the construction in the proof of [Miyano and Hayashi 1984, Theorem 5.1] is used to replace existence of winning strategies in two-player weak games by existence of sequences of pairs of sets which satisfy a Büchi condition. Specifically, for sets $P \neq \emptyset$ and P' of states of \mathcal{A} for a data word σ over Σ , and subsets P_b of P and P'_b of P' which consist only of states with odd ranks, we write $\langle P, P_b \rangle \Rightarrow \langle P', P'_b \rangle$ iff there exists a map $p \mapsto P'_p$ on P such that $p \Rightarrow P'_p$ for each $p \in P$, $P' = \bigcup_{p \in P} P'_p$, $P'_b = P'_b$ if $P'_b \neq \emptyset$, and $P'_b = \{p' \in P' : \rho(p') \text{ is odd}\}$ if $P'_b = \emptyset$, where

$$P'_b = \{p' : \text{for some } p \in P_b, p' \in P'_p \text{ and } \rho(p') = \rho(p)\}$$

When $P'_b = \emptyset$ for some such map $p \mapsto P'_p$ on P , we say that P'_b is *fresh*. Instead of (I) above, we have:

- (V) for every infinite data word σ over Σ , \mathcal{A} accepts σ iff there exists a sequence $\langle P_0, P_{0,b} \rangle \Rightarrow \langle P_1, P_{1,b} \rangle \Rightarrow \dots$ of pairs of sets of states of \mathcal{A} for σ such that $P_0 = \{\langle 0, q_I, \emptyset \rangle\}$, $P_{0,b} = P_0$ if $\rho(q_I)$ is odd, $P_{0,b} = \emptyset$ if $\rho(q_I)$ is even, and either the sequence ends with $\langle \emptyset, \emptyset \rangle$ or $P_{i,b}$ is fresh for infinitely many i .

Finally, we remark that $\mathcal{C}_A^{<\omega}$ and \mathcal{C}_A^ω are computable in polynomial space. The pseudo-code in Figure 10 can be implemented so that at most one component of the maps $q \in Q_\dagger \mapsto f(q) \in \langle\langle a, ee, \perp, q \rangle\rangle$, $q \in Q_- \mapsto f_-(q) \in \langle\langle a, ee, \top, q \rangle\rangle$ and $q \in Q_\emptyset \mapsto f_\emptyset(q) \in \langle\langle a, ee, \perp, q \rangle\rangle$ is stored in any state (by means of its location). The definition in Figure 9 provides a nondeterministic algorithm which, given $a \in \Sigma$, $ee, uu \in \{\top, \perp\}$, $q \in Q$ and $Q'_\neq, Q'_= \subseteq Q$, checks whether $\langle Q'_\neq, Q'_= \rangle \in \langle\langle a, ee, uu, q \rangle\rangle$ in space polynomial in the size of \mathcal{A} . \square

5. LOWER COMPLEXITY BOUNDS

To warm up, we show that the upper bounds in Theorem 4.3 are tight already for deterministic automata, and in the case of the NLOGSPACE-memberships, already with no registers.

THEOREM 5.1. *The following hold over finite and over infinite data words:*

- (a) *nonemptiness for 1DRA is PSPACE-hard;*
 (b) *nonemptiness for 1DRA₀ is NLOGSPACE-hard.*

PROOF. Part (b) is an immediate consequence of NLOGSPACE-hardness of non-emptiness for classical DFA.

For (a), we reduce from the halting problem for Turing machines with binary alphabets and linearly bounded tapes. Precisely, we consider Turing machines $\mathcal{M} = \langle Q, q_I, \delta \rangle$ such that Q is a finite set of locations, q_I is the initial location, and $\delta : Q \times \{0, 1\} \rightarrow Q \times \{0, 1\} \times \{-1, 1\}$ is a transition function. A state of \mathcal{M} is a triple $\langle q, i, w \rangle$ where $q \in Q$ is the machine location, $0 \leq i < |\mathcal{M}|$ is the head position, and $w \in \{0, 1\}^{|\mathcal{M}|}$ is the tape contents. Let $\delta(q, w(i)) = \langle q', b, j \rangle$. If $0 \leq i + j < |\mathcal{M}|$, the state $\langle q', i + j, w[i \mapsto b] \rangle$ is the unique successor of $\langle q, i, w \rangle$. Otherwise, $\langle q, i, w \rangle$ has no successors. The following problem is PSPACE-hard: given \mathcal{M} as above, to decide whether the computation from the initial state $\langle q_I, 0, 00\dots 0 \rangle$ reaches a state with no successor.

We encode a computation $\langle q_0, i_0, w_0 \rangle \langle q_1, i_1, w_1 \rangle \dots$ of \mathcal{M} by the following data word over the alphabet $Q \uplus \{-\}$. Its underlying word is

$$-- a_0^0 a_1^0 \dots a_{|\mathcal{M}|-1}^0 a_0^1 a_1^1 \dots a_{|\mathcal{M}|-1}^1 \dots$$

where $a_l^k = q_k$ if $l = i_k$, and $a_l^k = -$ otherwise. There are two equivalence classes: $0 \not\sim 1$, $2 + |\mathcal{M}| \times k + l \sim 0$ if $w_k(l) = 0$, and $2 + |\mathcal{M}| \times k + l \sim 1$ if $w_k(l) = 1$.

It is straightforward to construct, in space logarithmic in $|\mathcal{M}|$, an automaton $\mathcal{A}_\mathcal{M}$ in 1DRA with alphabet $Q \uplus \{-\}$ which accepts a data word iff it has a prefix that encodes a computation of \mathcal{M} from the initial state to a state with no successor. $\mathcal{A}_\mathcal{M}$ has $2 + |\mathcal{M}|$ registers r_0, r_1 , and r'_l for $0 \leq l < |\mathcal{M}|$. It stores $[0]_\sim$ into r_0 and $[1]_\sim$ into r_1 , and checks that $0 \not\sim 1$ and the initial state is encoded correctly. Whenever

$\mathcal{A}_{\mathcal{M}}$ moves to a word position $2 + |\mathcal{M}| \times (k + 1)$, it has kept q_k and i_k using its location and has stored $[2 + |\mathcal{M}| \times k + l]_{\sim}$ in r'_l for each $0 \leq l < |\mathcal{M}|$. If $\langle q_k, i_k, w_k \rangle$ has no successor, $\mathcal{A}_{\mathcal{M}}$ accepts. Otherwise, it checks that positions $2 + |\mathcal{M}| \times (k + 1)$, \dots , $2 + |\mathcal{M}| \times (k + 2) - 1$ encode the successor state, simultaneously updates r'_0 , $r'_1, \dots, r'_{|\mathcal{M}|-1}$, and repeats. \square

Satisfiability for $\text{LTL}_1^{\downarrow}(\mathbf{X}, \mathbf{U})$ and nonemptiness for 1ARA_1 were shown in Theorem 4.4 to be decidable over finite data words, and in Π_1^0 over infinite data words. We now establish their non-primitive recursiveness in the finitary case, and Π_1^0 -hardness in the infinitary case. In fact, we have those lower bounds even for the unary logical fragment and universal automata.

THEOREM 5.2. *Satisfiability for $\text{LTL}_1^{\downarrow}(\mathbf{X}, \mathbf{F})$ and nonemptiness for 1URA_1 are not primitive recursive over finite data words, and Π_1^0 -hard over infinite data words.*

PROOF. By Theorem 2.9 (b), it suffices to show that, given an incrementing CA $\mathcal{C} = \langle \Sigma, Q, q_I, n, \delta, F \rangle$, sentences $\phi_{\mathcal{C}}^{\leq \omega}$ and $\phi_{\mathcal{C}}^{\omega}$ of $\text{LTL}_1^{\downarrow}(\mathbf{X}, \mathbf{F})$ and automata $\mathcal{A}_{\mathcal{C}}^{\leq \omega}$ and $\mathcal{A}_{\mathcal{C}}^{\omega}$ in 1URA_1 are computable in logarithmic space, such that their alphabet is $\widehat{\Sigma} = Q \times (\Sigma \cup \{\varepsilon\}) \times L \times Q$ where $L = \{\text{inc}, \text{dec}, \text{ifz}\} \times \{1, \dots, n\}$, and

$$L^{\alpha}(\mathcal{C}) = \{\bar{\sigma} : \sigma \in L^{\alpha}(\phi_{\mathcal{C}}^{\alpha})\} = \{\bar{\sigma} : \sigma \in L^{\alpha}(\mathcal{A}_{\mathcal{C}}^{\alpha})\}$$

for $\alpha \in \{\leq \omega, \omega\}$, where $\bar{\sigma} = w_0 w_1 \dots$ if $\text{str}(\sigma) = \langle q_0, w_0, l_0, q'_0 \rangle \langle q_1, w_1, l_1, q'_1 \rangle \dots$.

To ensure that a data word over $\widehat{\Sigma}$ encodes a run of \mathcal{C} , we constrain its equivalence relation. Firstly, there must not be two $\langle \text{inc}, c \rangle$ transitions or two $\langle \text{dec}, c \rangle$ transitions (with the same c) in the same class. For an $\langle \text{ifz}, c \rangle$ transition to be correct, whenever it is preceded by $\langle \text{inc}, c \rangle$, there must be an intermediate $\langle \text{dec}, c \rangle$ in the same class. Incrementing errors may occur because a $\langle \text{dec}, c \rangle$ transition may be preceded by no $\langle \text{inc}, c \rangle$ in the same class. Such a $\langle \text{dec}, c \rangle$ transition corresponds to a faulty decrement which leaves c unchanged. It is easy to check that, for every run of \mathcal{C} , there exists a run which differs at most in counter values and whose only incrementing errors are such faulty decrements.

More precisely, \mathcal{C} accepts a finite word w over Σ iff $w = \bar{\sigma}$ for some finite data word σ over $\widehat{\Sigma}$ satisfying the following, where $\text{str}(\sigma) = \langle q_0, w_0, l_0, q'_0 \rangle \langle q_1, w_1, l_1, q'_1 \rangle \dots$:

- (1) for each i , $\langle q_i, w_i, l_i, q'_i \rangle \in \delta$;
- (2) $q_0 = q_I$, and for each $i > 0$, $q'_{i-1} = q_i$;
- (3) for the maximum i , $q'_i \in F$;
- (4) there are no c and $i < j$ such that $l_i = l_j = \langle \text{inc}, c \rangle$ and $i \sim^{\sigma} j$;
- (5) there are no c and $i < j$ such that $l_i = l_j = \langle \text{dec}, c \rangle$ and $i \sim^{\sigma} j$;
- (6) for all c and i such that $l_i = \langle \text{inc}, c \rangle$, it is not the case that, there is $j > i$ with $l_j = \langle \text{ifz}, c \rangle$ but there is no $k > i$ with $l_k = \langle \text{dec}, c \rangle$ and $i \sim^{\sigma} k$;
- (7) there are no c and $i < j < k$ such that $l_i = \langle \text{inc}, c \rangle$, $l_j = \langle \text{ifz}, c \rangle$, $l_k = \langle \text{dec}, c \rangle$ and $i \sim^{\sigma} k$.

$\phi_{\mathcal{C}}^{\leq \omega}$ is constructed to express the conjunction of (1)–(7). (1)–(3) are straightforward. Among (4)–(7), the most interesting is (7), and the rest can be expressed

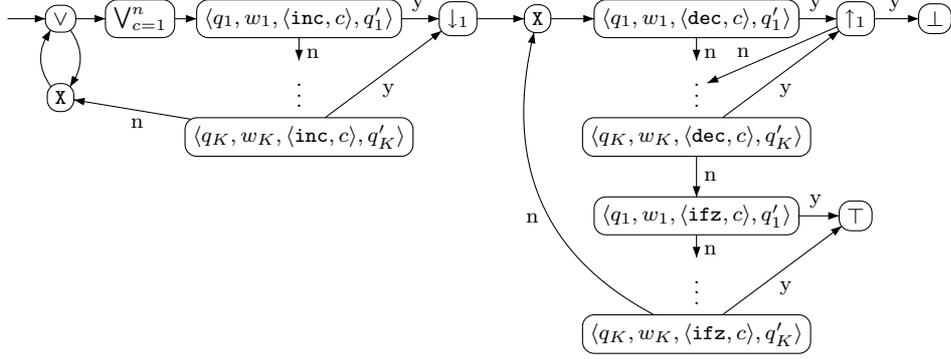


Fig. 11. Recognising a wrong zero test

similarly. Observe how (6) and (7) were formulated to avoid using the U operator. The following sentence expresses (7):

$$\neg \bigvee_{c=1}^n \mathbf{F} \left(\left(\bigvee_{q,w,q'} \langle q, w, \langle \text{inc}, c \rangle, q' \rangle \right) \wedge \downarrow_1 \mathbf{X} \mathbf{F} \left(\left(\bigvee_{q,w,q'} \langle q, w, \langle \text{ifz}, c \rangle, q' \rangle \right) \wedge \mathbf{X} \mathbf{F} \left(\left(\bigvee_{q,w,q'} \langle q, w, \langle \text{dec}, c \rangle, q' \rangle \right) \wedge \uparrow_1 \right) \right) \right)$$

For $\mathcal{A}_C^{\leq \omega}$, it is sufficient by Theorem 2.7, for each of (1)–(7), to compute in logarithmic space an automaton in 1NRA_1 which accepts a finite data word over $\widehat{\Sigma}$ iff it fails the condition. In fact, (6) and (7) can be treated together by checking that some $\langle \text{inc}, c \rangle$ instruction is followed by no occurrence of $\langle \text{dec}, c \rangle$ with the same datum until $\langle \text{ifz}, c \rangle$ occurs, and this automaton is the most interesting. It is shown in Figure 11, where $\langle q_1, w_1, q'_1 \rangle, \dots, \langle q_K, w_K, q'_K \rangle$ enumerates $Q \times (\Sigma \cup \{\varepsilon\}) \times Q$.

In the infinitary case, we replace (3) by:

(3') for infinitely many $i, q_i \in F$.

A sentence of $\text{LTL}_1^\downarrow(\mathbf{X}, \mathbf{F})$ which expresses (3') is $\text{GF} \bigvee_{q \in F, w, l, q'} \langle q, w, l, q' \rangle$. To express the negation of (3') in 1NRA_1 , the automaton guesses i and checks that $q_j \notin F$ for each $j \geq i$. \square

From the proofs of Theorems 4.4 and 5.2, and by observing that incrementing CA are closed under homomorphisms, we have the following characterisation of languages of incrementing CA in terms of languages of future-time LTL with 1 register. Remarkably, it is not affected by restricting to the unary logical fragment.

COROLLARY 5.3. *For both $\alpha \in \{\omega, \omega\}$ and every finite alphabet Σ , we have:*

$$\begin{aligned} & \{L^\alpha(\mathcal{C}) : \mathcal{C} \text{ is an incrementing CA with alphabet } \Sigma\} \\ &= \{ \{f(\text{str}(\sigma)) \in \Sigma^\alpha : \sigma \in L^\alpha(\phi)\} : \\ & \Sigma' \xrightarrow{f} \Sigma \cup \{\varepsilon\}, \phi \text{ is a sentence of } \text{LTL}_1^\downarrow(\mathbf{X}, \mathbf{F}) \text{ with alphabet } \Sigma' \} \\ &= \{ \{f(\text{str}(\sigma)) \in \Sigma^\alpha : \sigma \in L^\alpha(\phi)\} : \\ & \Sigma' \xrightarrow{f} \Sigma \cup \{\varepsilon\}, \phi \text{ is a sentence of } \text{LTL}_1^\downarrow(\mathbf{X}, \mathbf{U}) \text{ with alphabet } \Sigma' \} \end{aligned}$$

Our final result shows that the problems in Theorem 4.4 become Σ_1^0 -hard in the finitary case and Σ_1^1 -hard in the infinitary case as soon as any of 1 more register, the F^{-1} temporal operator or backward automaton moves are added, even after restricting to the unary logical fragment and universal automata. The result should also be compared with Theorem 4.2.

The theorem below improves [French 2003, Corollary 1] and [Demri et al. 2007, Theorem 3], which showed Σ_1^1 -hardness of the infinitary satisfiability problems for $LTL_2^\downarrow(\mathbf{X}, \mathbf{X}^{-1}, \mathbf{F}, \mathbf{F}^{-1})$ and $LTL_2^\downarrow(\mathbf{X}, \mathbf{U})$. Also, recalling Theorem 2.7, it implies [Neven et al. 2004, Theorem 5.1] where finitary nonuniversality for 1NRA was shown undecidable. Undecidability of finitary nonemptiness for 2DRA₁ was shown in [David 2004, Section 7.3], using a different encoding.

THEOREM 5.4. *Over finite (resp., infinite) data words, we have that satisfiability for $LTL_1^\downarrow(\mathbf{X}, \mathbf{F}, \mathbf{F}^{-1})$, nonemptiness for 2DRA₁ (resp., 2URA₁), satisfiability for $LTL_2^\downarrow(\mathbf{X}, \mathbf{F})$ and nonemptiness for 1URA₂ are Σ_1^0 -hard (resp., Σ_1^1 -hard).*

PROOF. By Theorem 2.9 (a), it is sufficient to show that, given a Minsky CA $\mathcal{C} = \langle \Sigma, Q, q_I, n, \delta, F \rangle$, sentences $\phi_{\mathcal{C}}^{<\omega}$ and $\phi_{\mathcal{C}}^\omega$ of $LTL_1^\downarrow(\mathbf{X}, \mathbf{F}, \mathbf{F}^{-1})$, automata $\mathcal{A}_{\mathcal{C}}^{<\omega}$ in 2DRA₁ and $\mathcal{A}_{\mathcal{C}}^\omega$ in 2URA₁, sentences $\psi_{\mathcal{C}}^{<\omega}$ and $\psi_{\mathcal{C}}^\omega$ of $LTL_2^\downarrow(\mathbf{X}, \mathbf{F})$ and automata $\mathcal{B}_{\mathcal{C}}^{<\omega}$ and $\mathcal{B}_{\mathcal{C}}^\omega$ in 1URA₂ are computable in logarithmic space, such that their alphabet is $\widehat{\Sigma} = Q \times (\Sigma \cup \{\varepsilon\}) \times L \times Q$ where $L = \{\mathbf{inc}, \mathbf{dec}, \mathbf{ifz}\} \times \{1, \dots, n\}$, and

$$\begin{aligned} L^\alpha(\mathcal{C}) &= \{\bar{\sigma} : \sigma \in L^\alpha(\phi_{\mathcal{C}}^\alpha)\} = \{\bar{\sigma} : \sigma \in L^\alpha(\mathcal{A}_{\mathcal{C}}^\alpha)\} \\ &= \{\bar{\sigma} : \sigma \in L^\alpha(\psi_{\mathcal{C}}^\alpha)\} = \{\bar{\sigma} : \sigma \in L^\alpha(\mathcal{B}_{\mathcal{C}}^\alpha)\} \end{aligned}$$

for $\alpha \in \{<\omega, \omega\}$, where $\bar{\sigma} = w_0 w_1 \dots$ if $\text{str}(\sigma) = \langle q_0, w_0, l_0, q'_0 \rangle \langle q_1, w_1, l_1, q'_1 \rangle \dots$.

To ensure that a data word over $\widehat{\Sigma}$ corresponds to a run of \mathcal{C} , we constrain its equivalence relation as was done for incrementing CA in the proof of Theorem 5.2, and in addition require that each $\langle \mathbf{dec}, c \rangle$ transition be preceded by some $\langle \mathbf{inc}, c \rangle$ in the same class, which eliminates the possibility of faulty decrements.

More precisely, \mathcal{C} accepts a finite (resp., infinite) word w over Σ iff $w = \bar{\sigma}$ for some finite (resp., infinite) data word σ over $\widehat{\Sigma}$ which satisfies (1)–(7) (resp., (1), (2), (3') and (4)–(7)) given in the proof of Theorem 5.2, and

(8) whenever $l_i = \langle \mathbf{dec}, c \rangle$, there exists $j < i$ such that $l_j = \langle \mathbf{inc}, c \rangle$ and $i \sim^\sigma j$,

where $\text{str}(\sigma) = \langle q_0, w_0, l_0, q'_0 \rangle \langle q_1, w_1, l_1, q'_1 \rangle \dots$.

To compute $\phi_{\mathcal{C}}^{<\omega}$ and $\phi_{\mathcal{C}}^\omega$, (8) is expressible in $LTL_1^\downarrow(\mathbf{X}, \mathbf{F}, \mathbf{F}^{-1})$ as:

$$\bigwedge_{c=1}^n \mathbf{G} \left(\left(\bigvee_{q,w,q'} \langle q, w, \langle \mathbf{dec}, c \rangle, q' \rangle \right) \Rightarrow \downarrow_1 \mathbf{F}^{-1} \left(\left(\bigvee_{q,w,q'} \langle q, w, \langle \mathbf{inc}, c \rangle, q' \rangle \right) \wedge \uparrow_1 \right) \right)$$

$\mathcal{A}_{\mathcal{C}}^{<\omega}$ is constructed to check (1)–(8) as follows:

- if the current transition, and the previous one (if any), fail (1) or (2), $\mathcal{A}_{\mathcal{C}}^{<\omega}$ rejects;
- if the current transition fails (3), $\mathcal{A}_{\mathcal{C}}^{<\omega}$ rejects;
- if the current instruction is $\langle \mathbf{inc}, c \rangle$, $\mathcal{A}_{\mathcal{C}}^{<\omega}$ stores the current class in the register, and then scans σ forwards and rejects if it finds an $\langle \mathbf{inc}, c \rangle$ in the same class, or an $\langle \mathbf{ifz}, c \rangle$ before a $\langle \mathbf{dec}, c \rangle$ in the same class, but otherwise returns;

- if the current instruction is $\langle \mathbf{dec}, c \rangle$, $\mathcal{A}_C^{\leq \omega}$ stores the current class in the register, and then scans σ backwards and rejects if it finds a $\langle \mathbf{dec}, c \rangle$ in the same class, or no $\langle \mathbf{inc}, c \rangle$ in the same class, but otherwise returns;
- if there is a next transition, $\mathcal{A}_C^{\leq \omega}$ repeats the above for it, but otherwise accepts.

\mathcal{A}_C^ω is constructed by extending the construction of \mathcal{A}_C^ω in the proof of Theorem 5.2 by expressing the negation of (8) by an automaton in 2NRA_1 : it guesses a position with a $\langle \mathbf{dec}, c \rangle$ instruction, and checks that it is not preceded by an $\langle \mathbf{inc}, c \rangle$ in the same class.

For the logical fragments and automata classes with 2 registers, we use a different encoding of runs of \mathcal{C} by data words, similar to the modelling in [Lisitsa and Potapov 2005, Section 4] of Minsky machines by systems of pebbles. Let $\tilde{\Sigma} = \widehat{\Sigma} \cup (\{\mathbf{hi}, \mathbf{lo}\} \times \{1, \dots, n\})$. For a data word σ over $\tilde{\Sigma}$, let $\bar{\sigma} = \sigma \upharpoonright \widehat{\Sigma}$.

Each transition $\langle q, w, l, q' \rangle$ in a run of \mathcal{C} is encoded by a block whose sequence of letters is $\langle \mathbf{hi}, 1 \rangle \langle \mathbf{lo}, 1 \rangle \cdots \langle \mathbf{hi}, n \rangle \langle \mathbf{lo}, n \rangle \langle q, w, l, q' \rangle$. For each counter c , two occurrences of $\langle \mathbf{hi}, c \rangle$ are in the same class iff there is no occurrence of $\langle \mathbf{inc}, c \rangle$ between them, which gives a sequence of classes D_0, D_1, \dots . Each occurrence of $\langle \mathbf{lo}, c \rangle$ is in some class D_i . If prior to a transition of \mathcal{C} , a counter c has value m , that is encoded by occurrences of $\langle \mathbf{lo}, c \rangle$ and $\langle \mathbf{hi}, c \rangle$ in the corresponding block being in some classes D_i and D_{i+m} .

More precisely, we have that \mathcal{C} accepts a finite (resp., infinite) word w over Σ iff $w = \bar{\sigma}$ for some finite (resp., infinite) data word σ over $\tilde{\Sigma}$ which satisfies:

- (i) $\text{str}(\sigma)$ is a sequence of blocks $\langle \mathbf{hi}, 1 \rangle \langle \mathbf{lo}, 1 \rangle \cdots \langle \mathbf{hi}, n \rangle \langle \mathbf{lo}, n \rangle \langle q, w, l, q' \rangle$;
- (ii) each $\langle q, w, l, q' \rangle$ is in δ ;
- (iii) for the first $\langle q, w, l, q' \rangle$, $q = q_I$, and for each $\langle q, w, l, q' \rangle$ and $\langle q'', w', l', q''' \rangle$ which are consecutive, $q' = q''$;
- (iv) for the last $\langle q, w, l, q' \rangle$, $q' \in F$ (resp., infinitely often $q \in F$);
- (v) in the initial block, for each c , $\langle \mathbf{hi}, c \rangle$ and $\langle \mathbf{lo}, c \rangle$ are in the same class;
- (vi) in each block immediately after an $\langle \mathbf{inc}, c \rangle$ one, $\langle \mathbf{hi}, c \rangle$ is not in the same class as any preceding $\langle \mathbf{hi}, c \rangle$, and $\langle \mathbf{lo}, c \rangle$ is in the same class as the previous $\langle \mathbf{lo}, c \rangle$;
- (vii) in each $\langle \mathbf{dec}, c \rangle$ block, $\langle \mathbf{hi}, c \rangle$ and $\langle \mathbf{lo}, c \rangle$ are not in the same class;
- (viii) in each block immediately after a $\langle \mathbf{dec}, c \rangle$ block B , $\langle \mathbf{hi}, c \rangle$ is in the same class as the previous $\langle \mathbf{hi}, c \rangle$, and $\langle \mathbf{lo}, c \rangle$ is in the same class as $\langle \mathbf{hi}, c \rangle$ in the block immediately after the last block containing $\langle \mathbf{hi}, c \rangle$ which is in the same class as $\langle \mathbf{lo}, c \rangle$ in B ;
- (ix) in each $\langle \mathbf{ifz}, c \rangle$ block, $\langle \mathbf{hi}, c \rangle$ and $\langle \mathbf{lo}, c \rangle$ are in the same class.

For $\psi_C^{\leq \omega}$ and ψ_C^ω , each of (i)–(ix) is expressed in $\text{LTL}_2^\downarrow(\mathbf{X}, \mathbf{F})$. In fact, (viii) naturally splits into two halves, and the second half is the most involved among (i)–(ix):

$$\mathbf{G} \bigwedge_{c=1}^n \left(\langle \mathbf{hi}, c \rangle \Rightarrow \downarrow_1 \mathbf{X}^{2n+1} \left(\neg \uparrow_1 \Rightarrow \downarrow_2 \mathbf{G} \left(\langle \mathbf{lo}, c \rangle \wedge \uparrow_1 \wedge (\mathbf{X}^{2(n-c)+1} \bigvee_{q,w,q'} \langle q, w, \langle \mathbf{dec}, c \rangle, q' \rangle) \Rightarrow \mathbf{X}^{2n+1} \uparrow_2 \right) \right) \right)$$

It remains by Theorem 2.7, for each of (i)–(ix), to compute in logarithmic space an automaton in 1NRA_2 which accepts a finite (resp., infinite) data word σ over

registers	finite data words		infinite data words	
	1	2	1	2
X, F	R, not PR	Σ_1^0 -complete	Π_1^0 -complete	Σ_1^1 -complete
X, U	R, not PR	Σ_1^0 -complete	Π_1^0 -complete	Σ_1^1 -complete
X, F, F^{-1}	Σ_1^0 -complete	Σ_1^0 -complete	Σ_1^1 -complete	Σ_1^1 -complete

Fig. 12. Complexity of satisfiability for fragments of LTL with the freeze quantifier

$\tilde{\Sigma}$ iff it fails the condition. For the second half of (viii), which is again the most involved, the automaton guesses a position with a $\langle \text{hi}, c \rangle$ letter, checks that the position $2n + 1$ steps forwards (which is the next occurrence of $\langle \text{hi}, c \rangle$) is not in the same class, guesses a subsequent position with the $\langle \text{lo}, c \rangle$ letter which is in the same class as the first $\langle \text{hi}, c \rangle$ position and whose block ends with the $\langle \text{dec}, c \rangle$ instruction, and checks that the position $2n + 1$ steps forwards (which is the next occurrence of $\langle \text{lo}, c \rangle$) is not in the same class as the second $\langle \text{hi}, c \rangle$ position. \square

6. CONCLUSION

By Theorems 4.2, 4.4, 5.2 and 5.4, we have the results on complexity of satisfiability shown in Figure 12, where ‘R, not PR’ means decidable and not primitive recursive. The entries not in bold follow from [French 2003, Corollary 1] and [Demri et al. 2007, Theorem 3].

The results on complexity of nonemptiness for register automata in Sections 4 and 5, except Σ_2^0 -membership of infinitary nonemptiness for 2NRA, are depicted in Figure 13. The edges indicate the syntactic inclusions between automata classes.

ACKNOWLEDGMENTS

We are grateful to Claire David, Massimo Franceschet, Marcin Jurdziński, Anca Muscholl, David Nowak, Joël Ouaknine, Philippe Schnoebelen and Luc Segoufin for helpful discussions.

REFERENCES

- ALUR, R. AND HENZINGER, T. 1994. A really temporal logic. *J. ACM* 41, 1, 181–204.
- BOJAŃCZYK, M., MUSCHOLL, A., SCHWENTICK, T., SEGOUFIN, L., AND DAVID, C. 2006. Two-variable logic on words with data. In *21th IEEE Symp. on Logic in Comput. Sci. (LICS)*. IEEE Comput. Soc., 7–16.
- BOUAJJANI, A., HABERMEHL, P., JURSKI, Y., AND SIGHIREANU, M. 2007. Rewriting systems with data. In *Fundamentals of Comput. Theory (FCT), 16th Int. Symp. Lect. Notes Comput. Sci.*, vol. 4639. Springer, 1–22.
- DAVID, C. 2004. Mots et données infinies. M.S. thesis, Laboratoire d’Informatique Algorithmique: Fondements et Applications, Paris.
- DEMRI, S., D’SOUZA, D., AND GASCON, R. 2007. A decidable temporal logic of repeating values. In *Logical Found. of Comput. Sci. (LFCS), Int. Symp. Lect. Notes Comput. Sci.*, vol. 4514. Springer, 180–194.
- DEMRI, S. AND LAZIĆ, R. 2006. LTL with the freeze quantifier and register automata. In *21th IEEE Symp. on Logic in Comput. Sci. (LICS)*. IEEE Comput. Soc., 17–26.
- DEMRI, S., LAZIĆ, R., AND NOWAK, D. 2007. On the freeze quantifier in Constraint LTL: Decidability and complexity. *Inf. Comput.* 205, 1, 2–24.
- DEMRI, S., LAZIĆ, R., AND SANGNIER, A. 2008. Model checking freeze LTL over one-counter
- ACM Transactions on Computational Logic, Vol. V, No. N, M 20YY.

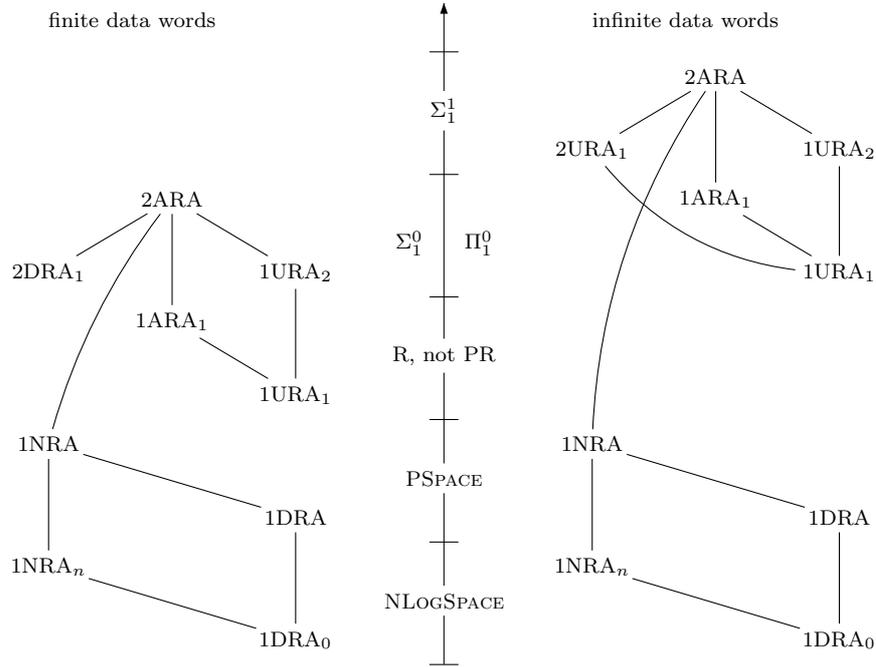


Fig. 13. Complexity of nonemptiness for classes of register automata

automata. In *Found. of Softw. Sci. and Comput. Struct. (FoSSaCS), 11th Int. Conf. Lect. Notes Comput. Sci.* Springer. To appear.

DICKSON, L. 1913. Finiteness of the odd perfect and primitive abundant numbers with distinct factors. *Amer. J. Math.* 35, 413–422.

ETESSAMI, K., VARDI, M., AND WILKE, T. 2002. First-order logic with two variables and unary temporal logic. *Inf. Comput.* 179, 2, 279–295.

FITTING, M. 2002. Modal logic between propositional and first-order. *J. Logic and Comput.* 12, 6, 1017–1026.

FRENCH, T. 2003. Quantified propositional temporal logic with repeating states. In *10th Int. Symp. on Temporal Representation and Reasoning / 4th Int. Conf. on Temporal Logic (TIME-ICTL)*. IEEE Comput. Soc., 155–165.

GORANKO, V. 1996. Hierarchies of modal and temporal logics with references pointers. *J. Logic, Lang., and Inf.* 5, 1–24.

KAMINSKI, M. AND FRANCEZ, N. 1994. Finite-memory automata. *Theor. Comput. Sci.* 134, 2, 329–363.

KOSARAJU, S. R. 1982. Decidability of reachability in vector addition systems (preliminary version). In *14th Ann. ACM Symp. on Theory of Comput. (STOC)*. ACM, 267–281.

LAZIĆ, R. 2006. Safely freezing LTL. In *FSTTCS: Found. of Softw. Technology and Theor. Comput. Sci., 26th Int. Conf. Lect. Notes Comput. Sci.*, vol. 4337. Springer, 381–392.

LIPTON, R. J. 1976. The reachability problem requires exponential space. Tech. Rep. 62, Yale University.

LISITSA, A. AND POTAPOV, I. 2005. Temporal logic with predicate λ -abstraction. In *12th Int. Symp. on Temporal Representation and Reasoning (TIME)*. IEEE Comput. Soc., 147–155.

MAYR, R. 2003. Undecidable problems in unreliable computations. *Theor. Comput. Sci.* 297, 1–3, 337–354.

- MINSKY, M. 1967. *Computation, Finite and Infinite Machines*. Prentice Hall.
- MIYANO, S. AND HAYASHI, T. 1984. Alternating finite automata on ω -words. *Theor. Comput. Sci.* 32, 321–330.
- MULLER, D. E., SAOUDI, A., AND SCHUPP, P. E. 1986. Alternating automata, the weak monadic theory of the tree, and its complexity. In *Automata, Lang. and Program., 13th Int. Coll. (ICALP)*. Lect. Notes Comput. Sci., vol. 226. Springer, 275–283.
- NEVEN, F., SCHWENTICK, T., AND VIANU, V. 2004. Finite state machines for strings over infinite alphabets. *ACM Trans. On Comp. Logic* 5, 3, 403–435.
- OUAKNINE, J. AND WORRELL, J. 2006a. On metric temporal logic and faulty Turing machines. In *Found. of Softw. Sci. and Comput. Struct. (FoSSaCS), 9th Int. Conf.* Lect. Notes Comput. Sci., vol. 3921. Springer, 217–230.
- OUAKNINE, J. AND WORRELL, J. 2006b. Personal communication.
- SAKAMOTO, H. AND IKEDA, D. 2000. Intractability of decision problems for finite-memory automata. *Theor. Comput. Sci.* 231, 2, 297–308.
- SCHNOEBELEN, P. 2002. Verifying lossy channel systems has nonprimitive recursive complexity. *Inf. Process. Lett.* 83, 5, 251–261.
- SEGOUFIN, L. 2006. Automata and logics for words and trees over an infinite alphabet. In *Comput. Sci. Logic (CSL), 20th Int. Works.* Lect. Notes Comput. Sci., vol. 4207. Springer, 41–57.
- VARDI, M. Y. 1996. An automata-theoretic approach to linear temporal logic. In *Banff Higher Order Works*. Lect. Notes Comput. Sci., vol. 1043. Springer, 238–266.
- WALUKIEWICZ, I. 2001. Pushdown processes: Games and model-checking. *Inf. Comput.* 164, 2, 234–263.

Received October 2006; revised February 2008; accepted April 2008