New Complexity Results for Some Linear Counting Problems Using Minimal Solutions to Linear Diophantine Equations

(Extended Abstract)

Gaoyan Xie, Cheng Li and Zhe Dang*

School of Electrical Engineering and Computer Science
Washington State University
Pullman, WA 99164, USA

Abstract. The linear reachability problem is to decide whether there is an execution path in a given finite state transition system such that the counts of labels on the path satisfy a given linear constraint. Using results on minimal solutions (in nonnegative integers) for linear Diophantine systems, we obtain new complexity results for the problem, as well as for other linear counting problems of finite state transition systems and timed automata. In contrast to previously known results, the complexity bounds obtained in this paper are polynomial in the size of the transition system in consideration, when the linear constraint is fixed.

1 Introduction

Model-checking [7, 22] is a technique that automatically verifies a finite state transition system against a temporal property usually specified in, e.g., Computation Tree Logic (CTL) [7] or Linear Temporal Logic (LTL) [20], by exhaustively exploring the finite state space of the system. The usefulness of model-checking has been demonstrated by several successful model-checkers (e.g., SMV [17], SPIN [15], BMC [3]) which have been used to test/verify industrial-level hardware/software systems with significant sizes.

Although both CTL and LTL are expressive, many temporal properties are out of their scope. For instance, event counting is a fundamental concept to specify some important fairness properties. As a motivating example, we consider the design (depicted as a finite state transition system $\mathcal{A}$ in Figure 1) of a process scheduler. The scheduler schedules two kinds of processes: P_r and P_w according to some scheduling strategy. A transition with label P_r (resp. P_w) is taken when the scheduler chooses a P_r (resp. P_w) process to run. It is required that the design shall satisfy some fairness properties; e.g., starting from state s_0, whenever s_0 is reached, the number of P_r processes scheduled is greater than or equal to the number of P_w processes scheduled and less than or equal to twice the number of P_w processes scheduled. To ensure that the design meets the requirement, we need check whether for any path p that starts from and ends with s_0, the linear constraint, $\#_{P_w}(p) \leq \#_{P_r}(p) \leq 2\#_{P_w}(p)$, is satisfied, where $\#_{P_w}(p)$

* Corresponding author (zdang@eecs.wsu.edu).

(resp. $\#_{P_r}(p)$) stands for the count of labels P_w (resp. P_r) on path p. Notice that this property is nonregular [6] and, since the counts could go unbounded, the property is not expressible in CTL or LTL.

In general, by considering its negation, the property can be formulated as a *linear reachability problem* as follows.

- **Given:** A finite state transition system $\mathcal{A}$ with labels $a_1, \cdots, a_k$, two designated states s_{init} and s_{final}, and a linear constraint $U(x_1, ..., x_k)$.
- **Question:** Is there a path p of $\mathcal{A}$ from s_{init} to s_{final} such that p satisfies U (i.e., $U(\#_{a_1}(p), ..., \#_{a_k}(p))$ holds)?

The reachability problem is decidable. To see this, one can treat $\mathcal{A}$ to be a finite automaton with initial state s_{init} and final state s_{final}. $L(\mathcal{A})$ is the regular language over alphabet $\{a_1, \cdots, a_k\}$ accepted by $\mathcal{A}$. A naive decision procedure consists of the following three steps: *(i).* Compute a regular expression for $L(\mathcal{A})$, *(ii).* Calculate the semilinear set of the regular expression defined by a Presburger formula R [19], and *(iii).* Check the satisfiability of the Presburger formula $R \wedge U$. Unfortunately, the time complexity of this procedure is at least $O(2^{|S|})$, where $|S|$ is the number of states in $\mathcal{A}$, even when k is fixed. This is because the size of the regular expression, in worst cases, is exponential in $|S|$ [16].

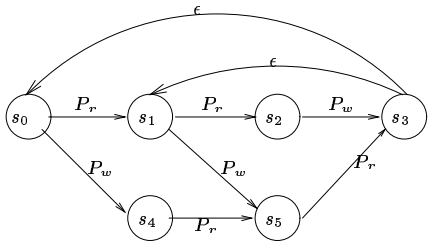

Fig. 1. An example of a scheduler

In this paper, we present a new algorithm solving the linear reachability problem. This algorithm is completely different from the naive one. In the algorithm, we estimate a bound B (called a bounding box) from $\mathcal{A}$ and U such that, the **Question**-part is true iff the truth is witnessed by some p on which the count $\#_{a_i}(p)$ for each label a_i is bounded by B. Interestingly, after a complex loop analysis, estimating a bounding box B is reduced to a number theory problem: finding nonnegative minimal solutions to linear Diophantine systems. There has been much research on this latter problem for homogeneous/inhomogeneous systems with (nonnegative) integer solutions [4, 5, 13, 21]. Suppose that U is in a disjunctive normal form over linear equations/inequalities. Using the Borosh-Flahive-Treybig bound in [4], we are able to show that, in worst cases, when $|S|$ is $\gg$ the size (which will be made clear later) of U, the bounding box B is bounded by $O(|S|^{k+L+3})$, where L is the maximal number of conjunctions in a single disjunctive term of U. Using this bounding box, one can easily show that the linear

reachability problem is solvable in time

$$O(|S|^{2k(k+L+3)+2}), \tag{1}$$

when $|S| \gg k$ and the size of U. In particular, when k and U are fixed, the complexity is polynomial in $|S|$. This is in contrast to the complexity of the naive algorithms that is exponential in the state number $|S|$.

Our complexity result in (1) for the linear reachability problem can be further used to obtain complexity bounds (which were unknown) for some other linear counting problems involving linear constraints over counts. For instance, in contrast to the reachability problem, we consider a *linear liveness problem* as follows. For an ω-path π of $\mathcal{A}$, we say that π is U-i.o. (infinitely often) at state s' if there are infinitely many prefixes p of π such that p ends at s' and satisfies U. The liveness problem is to decide, given two states s and s', whether $\mathcal{A}$ has an ω-path π that starts from s and is U-i.o. at s'. The application issues of the problem can be found in [11]. In particular, among other results in the same paper, it is shown that the liveness problem is decidable. However, the time complexity is unknown. In this paper, we are able to use (1) to establish a complexity bound for the liveness problem, which is similar to the bound given in (1).

We also consider the linear reachability problem when $\mathcal{A}$ is ordered; i.e., on any path p from s_{init} to s_{final}, each label a_j is after all the a_i's, whenever $i < j$. For this restricted model of $\mathcal{A}$, we obtain a smaller complexity bound, $O(|S|^{4k-1})$, than (1) for the reachability problem, using the Pottier bound [21] (the Borosh-Flahive-Treybig bound is not applicable here) for nonnegative minimal solutions to linear Diophantine systems. Interestingly, this restricted model and the complexity bound can be used to establish a new complexity result for timed automata [1]. We first consider discrete timed automata where clocks take integral values. The *linear reachability problem for discrete timed automata* is defined as follows:

- **Given:** A discrete timed automaton $\mathcal{D}$, two designated states s_{init} and s_{final}, and two linear constraints U and U' over k variables.
- **Question:** are there clock values $v_1, \cdots, v_k, v'_1, \cdots, v'_k$ such that $\langle s_{\text{init}}, v_1, \cdots, v_k \rangle$ reaches $\langle s_{\text{final}}, v'_1, \cdots, v'_k \rangle$ and both $U(v_1, \cdots, v_k)$ and $U'(v'_1, \cdots, v'_k)$ hold?

Though many temporal verification problems involving linear constraints over clocks are known to be undecidable [1, 2, 12], the linear reachability problem is decidable for timed automata (even when the clocks are dense) [8–10]. However, an upper bound for the worst case complexity is unknown. Using the result for the linear reachability problem when $\mathcal{A}$ is ordered, we can establish that the linear reachability problem for discrete timed automata is solvable in time $O(|S|^{8k})$ when $|S| \gg k$, the sizes of U and U', and the maximal absolute value of all the constants appearing in the clock constraints of $\mathcal{A}$. This result can be generalized to timed automata with dense clocks using the pattern technique in [9].

2 Preliminaries

Let $\mathbf{N}$ be the set of nonnegative integers and k be a positive integer. A *finite state transition system* can be defined as

$$\mathcal{A} = \langle S, \Sigma, E \rangle, \tag{2}$$

where S is a finite set of *states*, $\Sigma = \{a_1, \cdots, a_k\}$ is a set of *labels*, $E \subseteq S \times (\Sigma \cup \{\epsilon\}) \times S$ is a set of *transitions*. When $E \subseteq S \times \{\epsilon\} \times S$, the system is called a finite state machine. A *path* p of $\mathcal{A}$ is a finite sequence of transitions in the form of

$$(s_0 \tau_0 s_1)...(s_i \tau_i s_{i+1})..(s_{n-1} \tau_{n-1} s_n) \tag{3}$$

for some n such that for each $0 \leq i < n$, $(s_i \tau_i s_{i+1}) \in E$. Path p is a *simple cycle* if $s_0, ..., s_{n-1}$ are distinct and $s_0 = s_n$. Path p is a *simple path* if $s_0, ..., s_{n-1}, s_n$ are all distinct. For any path p of $\mathcal{A}$, let $\#(p)$ denote the k-ary vector $(\#_{a_1}(p), ..., \#_{a_k}(p))$, where each $\#_{a_i}(p)$ stands for the number of label a_i's occurrences on p, $1 \leq i \leq k$.

Let $x_1, \cdots, x_k$ be nonnegative integer variables. An *atomic linear constraint* is in the form of

$$b_1 x_1 + \cdots + b_k x_k \sim b \tag{4}$$

where $\sim \in \{=, \geq\}$, $b_1, ..., b_k$ and b are integers. When $\sim$ is $=$ (resp. $\geq$), the constraint is called an *equation* (resp. *inequality*). The constraint is *made homogeneous* if one makes $b = 0$ in the constraint. A *linear constraint* U is a Boolean combination of atomic linear constraints (using $\wedge, \vee, \neg, \rightarrow$). Without loss of generality, throughout this paper, we assume that the linear constraint U is always written as a disjunction $U_1 \vee \cdots \vee U_m$, for some m, of conjunctions of atomic linear constraints. When $m = 1$, U is called a *conjunctive* linear constraint. U is *made homogeneous* if each atomic linear constraint in U is made homogeneous; we use $U^{\hom}$ to denote the result. In particular, a conjunctive linear constraint U is a *linear Diophantine equation system* if each atomic linear constraint in U is an equation.

Suppose that U is a conjunctive linear constraint, which contains e equations and $l - e$ inequalities. One may write U into $\mathbf{Bx} \sim \mathbf{b}$, where $\sim \in \{=, \geq\}^l$, $\mathbf{B}$ (l by k) and $\mathbf{b}$ (l by 1) are matrices of integers, and $\mathbf{x}$ is the column of variables $x_1, \cdots, x_k$. As usual, $(\mathbf{B}, \mathbf{b})$ is called the augmented matrix of U, and $\mathbf{B}$ is called the coefficient matrix of U. We use $||\mathbf{B}||_{1,\infty}$ to denote $\max_i\{\sum_j |b_{ij}|\}$ (b_{ij} is the element of row i and column j in $\mathbf{B}$) and use $||\mathbf{b}||_\infty$ to denote the maximum of the absolute values of all the elements in $\mathbf{b}$. Assume r is the rank of $(\mathbf{B}, \mathbf{b})$, and Γ_1 (resp. Γ_2) is the maximum of the absolute values of all the $r \times r$ minors of $\mathbf{B}$ (resp. $(\mathbf{B}, \mathbf{b})$).

When U is a linear Diophantine equation system (i.e., $e = l$), for any given tuples $(v_1, \cdots, v_k)$ and $(v'_1, \cdots, v'_k)$ in $\mathbf{N}^k$, we say $(v_1, \cdots, v_k) \leq (v'_1, \cdots, v'_k)$ if $v_i \leq v'_i$ for all $1 \leq i \leq k$. We say $(v_1, \cdots, v_k) < (v'_1, \cdots, v'_k)$ if $(v_1, \cdots, v_k) \leq (v'_1, \cdots, v'_k)$ and $v_i < v'_i$ for some $1 \leq i \leq k$. A tuple $(v'_1, \cdots, v'_k)$ is a *minimal solution* to U if $(v'_1, \cdots, v'_k)$ is a solution to U but any $(v_1, \cdots, v_k)$ with $(0, \cdots, 0) < (v_1, \cdots, v_k) < (v'_1, \cdots, v'_k)$ is not. Clearly, there are only finitely many minimal solutions to U. It has been an active research area to estimate a bound for minimal solutions, and the following Borosh-Flahive-Treybig bound [4] is needed in this paper.

Theorem 1. *(Borosh-Flahive-Treybig bound) A linear Diophantine equation system U has solutions in nonnegative integers iff it has a solution $(x_1, \cdots, x_k)$ in nonnegative integers, such that r unknowns are bounded by Γ_1 and $k - r$ unknowns are bounded by $(\max(k, l) - r + 1)\Gamma_2$.*

The Borosh-Flahive-Treybig bound gives a bound for *one* of the minimal solutions in nonnegative integers to the inhomogeneous system U. In contrast, the following Pottier

bound gives an upper bound for *all* of the "minimal solutions" to a conjunctive U (which is not necessarily a linear equation system); this result can be simply obtained from Corollary 1 in [21].

Theorem 2. *(Pottier bound) For any conjunctive linear constraint U that contains e equations and $l - e$ inequalities, there are two finite sets S and $S^{\mathrm{hom}} = \{\mathbf{v}_1, \cdots, \mathbf{v}_q\}$, for some q, of vectors in $\mathbf{N}^k$ such that*

- *each element in S (resp. S^{hom}) is a solution to U (resp. U^{hom}),*
- *For any $\mathbf{v} \in \mathbf{N}^k$, $\mathbf{v}$ is a solution to U iff there are $t_1, \cdots, t_q \in \mathbf{N}$, $\mathbf{v} = \mathbf{v}_0 + t_1\mathbf{v}_1 + \cdots + t_q\mathbf{v}_q$ for some $\mathbf{v}_0 \in S$,*
- *each component of all the vectors in $S \cup S^{\mathrm{hom}}$ is bounded by $(2 + ||\mathbf{B}||_{1,\infty} + ||\mathbf{b}||_\infty)^{k+l+e}$.*

Therefore, for a conjunctive linear constraint U, every solution can be represented as the sum of a small solution and a nonnegative linear combination of small solutions to U^{hom} (clearly, the inverse is also true). Here, "small" means that the solutions are bounded by the Pottier bound.

When U is a linear constraint (i.e., $m \geq 1$), the Pottier bound of U is defined to be the maximal of all the bounds obtained from Theorem 2 for each conjunctive linear constraint in U.

An inequality can be translated into an equation by introducing a *slack* variable (e.g., $x_1 - 2x_2 \geq 3$ into $x_1 - 2x_2 - u = 3$ where u, a new variable on $\mathbf{N}$, is the slack variable). So if U is a conjunctive linear constraint (in which there are e equations and $l - e$ inequalities) over $x_1, \cdots, x_k$, we may write U into an equation system $U(x_1, \cdots, x_k, x_{k+1}, x_{k+l-e})$ with l equations, where x_{k+1}, x_{k+l-e} are the slack variables.

In the next section, we will derive a bounding box B for the linear reachability problem such that its **Question**-part is true iff the truth is witnessed by a p satisfying $\#_{a_i}(p) \leq B$ for each $1 \leq i \leq k$. From this B, the time complexity for solving the linear reachability problem can be easily obtained.

3 A Bounding Box B for the Linear Reachability Problem

Let $\mathcal{A}$ be a finite state transition system specified in (2). A set of k-ary nonnegative integer vectors Q is a *small linear set* (wrt the given $\mathcal{A}$) if Q is in the form of

$$\{\boldsymbol{e}_0 + \sum_{1 \leq j \leq r} X_j \boldsymbol{e}_j : \text{ each } X_j \geq 0\}, \tag{5}$$

where nonnegative integer r satisfies

$$r \leq |S|^k, \tag{6}$$

k-ary nonnegative integer vectors $\boldsymbol{e}_0, ..., \boldsymbol{e}_r$ satisfy

$$\boldsymbol{e}_0 \leq |S|^2 \cdot \mathbf{1}, \tag{7}$$

and for each $j = 1, ..., r$,

$$\boldsymbol{e}_j \leq |S| \cdot \mathbf{1}, \tag{8}$$

where $\mathbf{1}$ stands for the identity vector. Q is a *small semilinear set* if it is a union of finitely many small linear sets.

Recall that the linear reachability problem is to decide whether there exists a path p in $\mathcal{A}$ from s_{init} to s_{final} such that p satisfies a given linear constraint $U(x_1, ..., x_k)$. Let $\mathcal{P}$ be all paths of $\mathcal{A}$ from s_{init} to s_{final}. We use $\#(\mathcal{P})$ to denote the set of k-ary nonnegative integer vectors $\{\#(p) : p \in \mathcal{P}\}$. Using a complex loop analysis technique by reorganizing simple loops on a path, one can show that $\#(\mathcal{P})$ is a small semilinear set.

Lemma 1. *$\#(\mathcal{P})$ is a small semilinear set. That is, it can be represented as, for some t,* [1]

$$\#(\mathcal{P}) = \bigcup_{1 \leq i \leq t} Q_i \tag{9}$$

where each Q_i is a small linear set in the form of (5).

Remark 1. One might have already noticed that, (9) and (5) appear nothing new, since they simply rephrase the known fact that $\#(\mathcal{P})$ defines a semilinear set [19]. However, the bounds for the coefficients shown in (6,7,8) are new.

Now let's turn to the property formula U. Recall that U is written as a disjunction of m conjunctive linear constraints

$$U = \bigvee_{1 \leq i \leq m} U_i. \tag{10}$$

Fix any $1 \leq i \leq m$. Suppose that U_i contains l atomic linear constraints. After adding (at most l) slack variables $y_1, \cdots, y_l$, U_i can be written into the following form:

$$\begin{cases} b_{11}x_1 + ... + b_{1k}x_k + g_1y_1 \ = \ b_1 \\ \qquad \vdots \\ b_{l1}x_1 + ... + b_{lk}x_k + g_ly_l \ = \ b_l \end{cases} \tag{11}$$

where the b's and g's are integers (each g is -1 or 0). Let $\mathbf{B}$ be the coefficient matrix for variables $x_1, \cdots, x_k$ and $\mathbf{b}$ be the column of $b_1, \cdots, b_l$ in (11). Define $w_1 = ||\mathbf{B}||_{1,\infty}$ and $w_2 = ||\mathbf{b}||_\infty$. We may assume $w_1 > 0$ (otherwise let $w_1 = 1$). In the sequel, we shall use the following notions: W_1 (the maximum of all the values w_1 among all U_i's), W_2 (the maximum of all the values w_2 among all U_i's), and L (the maximum of all the values l among all U_i's).

Due to the disjunctive representations of (10) and (9), we can consider only one conjunction of U in the form of (11) and only one linear set in the form of (5). That is, by substituting $\boldsymbol{x} = (x_1, ..., x_k)$ in (11) with the expression in (5): $\boldsymbol{x} = \boldsymbol{e}_0 +$

[1] Note that though t may be large, it is irrelevant here.

$\sum_{1 \le j \le r} X_j \boldsymbol{e}_j$, the equation system (11) is transformed into the following equation system with unknowns $X_1, ..., X_r$ and $y_1, ..., y_l$:

$$\begin{cases} h_{11}X_1 + ... + h_{1r}X_r + g_1 y_1 &= d'_1 \\ \quad\vdots & \\ h_{l1}X_1 + ... + h_{lr}X_r + g_l y_l &= d'_l. \end{cases} \tag{12}$$

Hence, the linear reachability problem is reduced to finding a nonnegative integer solution to (12). Using (7) and (8), a simple calculation reveals that, in (12), all of the h's are bounded by $|S|W_1$ and all of the $d'_1, ..., d'_l$ are bounded by $|S|^2 W_1 + W_2$.

We use Γ_1 to denote the maximum of the absolute values of all the $t \times t$, $1 \le t \le l$, minors of the coefficient matrix for system (12) and Γ_2 to denote that of the augmented matrix. Using the above mentioned bounds for the coefficients and constants in (12), one can conclude that

$$\Gamma_1 \le (|S|W_1)^l l! \text{ and } \Gamma_2 \le (|S|^2 W_1 + W_2)(|S|W_1)^{l-1} l!. \tag{13}$$

A direct application of the Borosh-Flahive-Treybig bound in Theorem 1 shows that system (12) has solutions in nonnegative integers iff the system has a solution $(X_1, ..., X_r, y_1, ..., y_l)$ in nonnegative integers, among which r unknowns are bounded by Γ_1 and l unknowns are bounded by $(r+1)\Gamma_2$ (here, without loss of generality, we assume the worst case that the rank of coefficient matrix of (12) is l). Applying the bounds Γ_1 and $(r+1)\Gamma_2$ to X_j in (5) and using (7) and (8), the linear reachability problem is further reduced to the problem of finding a $p \in \mathcal{P}$ satisfying:

$$\#(p) \le (|S|^2 + (r-l)|S|\Gamma_1 + l|S|(r+1)\Gamma_2) \cdot \mathbf{1} \tag{14}$$

and $U(\#_{a_1}(p), ..., \#_{a_k}(p))$. Noticing that $l \le L$, and $r \le |S|^k$ according to (6), we apply the bounds of Γ_1 and Γ_2 in (13) to (14) and define a bounding box

$$B = (|S|^{k+2} W_1 + L|S|(|S|^k + 1)(|S|^2 W_1 + W_2))(|S|W_1)^{L-1} L! + |S|^2. \tag{15}$$

Hence,

Theorem 3. *Given a finite state transition system $\mathcal{A}$, two states $s_{\text{init}}, s_{\text{final}} \in S$, and a linear constraint $U(x_1, ..., x_k)$, the following two items are equivalent:*

- *There is a path p of $\mathcal{A}$ from s_{init} to s_{final} satisfying U,*
- *The above item is true for some p further satisfying $\#(p) \le B \cdot \mathbf{1}$, where B is defined in (15).*

Notice that B in (15) is independent of m in (10). Also, when the number of states $|S|$ in $\mathcal{A}$ is much larger than k and the metrics of U; i.e., $|S| \gg k, W_1, W_2, L$, the bounding box is in the order of

$$B = O(|S|^{k+L+3}). \tag{16}$$

In this case, one can easily show that the linear reachability problem is solvable in time

$$O(|S|^{2k(k+L+3)+2}). \tag{17}$$

4 The Linear Liveness Problem

An ω-path π of $\mathcal{A}$ is an infinite sequence such that each prefix is a path of $\mathcal{A}$. Let s and s' be any two designated states of $\mathcal{A}$. We say that π is U-i.o. (infinitely often) at s' if there are infinitely many prefixes p from s to s' of $\mathcal{A}$ such that p satisfies U (i.e., $U(\#_{a_1}(p), \cdots, \#_{a_k}(p))$ holds). The *linear liveness problem* can be formulated as follows:

- **Given:** A finite state transition system $\mathcal{A}$ in (2), two designated states s and s', and a linear constraint $U(x_1, ..., x_k)$.
- **Question:** Is there an ω-path π that starts from s and is U-i.o. at s'?

In [11], this problem is shown decidable. However, the time complexity is unknown. In this section, we reduce the liveness problem to a linear reachability problem.

Recall that U is in the form of (10), $U = \vee_{1 \leq i \leq m} U_i$, and U_i^{hom} is the result of making U_i homogeneous. One key observation is as follows. The **Question**-part in the linear liveness problem is true iff, for some $1 \leq i \leq m$, *(a).* there is a path of $\mathcal{A}$ from s to s' satisfying U_i, and, *(b).* there is a path of $\mathcal{A}$ from s' to s' satisfying U_i^{hom}. A proof of this observation can be followed from [11] using the pigeon-hole principle (noticing that each atomic linear constraint in U_i is in the form of (4) where $\sim \in \{=, \geq\}$). Both items are equivalent to the linear reachability problem for $\mathcal{A}$ concerning U_i and U_i^{hom}, respectively. By trying out all of the m number of U_i's and U_i^{hom}'s, and using Theorem 3 and (17), we conclude that:

Theorem 4. *The linear liveness problem is solvable in time shown in (17), when* $|S| \gg m, k, W_1, W_2, L$.

5 Ordered Finite State Transition Systems

Let $\mathcal{A}$ be a finite state transition system in (2). Suppose that an order of labels $a_1, \cdots, a_k$ is fixed, say $a_1 < \cdots < a_k$. $\mathcal{A}$ is *ordered* if, on any path p from s_{init} to s_{final}, each label a_i appears before each label a_j, whenever $i < j$. In this case, $\mathcal{A}$ behaves as follows: reading a_1's for 0 or more times, then reading a_2's for 0 or more times, and so on. For this restricted version of $\mathcal{A}$, we will derive a better complexity bound than (17) for the linear reachability problem.

Lemma 2. *The linear reachability problem for ordered* $\mathcal{A}$ *is solvable in time*

$$O(m \cdot |S|^{4k-2} \cdot P^{2k}), \tag{18}$$

where P *is the Pottier bound for* U *(i.e., the maximum of the Pottier bounds for all* U_i*'s). Furthermore, since* P *is independent of* $|S|$*, the linear reachability problem for ordered* $\mathcal{A}$ *is solvable in time*

$$O(|S|^{4k-1}), \tag{19}$$

when $|S| \gg m, k, P$.

Interestingly, this model of $\mathcal{A}$ and the complexity bound can be used to obtain a complexity bound for timed automata in the next section.

6 The Linear Reachability Problem for Timed Automata

A timed automaton [1] is a finite state machine augmented with a number of clocks. All the clocks progress synchronously with rate 1, except when a clock is reset to 0 at some transition. We first consider discrete timed automata where clocks take integral values. Formally, a *discrete timed automaton* $\mathcal{D}$ is a tuple

$$\langle S, \{x_1, \cdots, x_k\}, E\rangle,$$

where S is a finite set of *(control) states*, $x_1, \cdots, x_k$ are *clocks* taking values in $\mathbf{N}$, and E is a finite set of *edges* or *transitions*. Each edge $\langle s, \lambda, l, s'\rangle$ denotes a transition from state s to state s' with *enabling condition* l in the form of clock regions (i.e., $x\#c, x-y\#c$, where x, y are clocks, $\#$ denotes $\leq, \geq$, or $=$, and c is an integer) and a clock reset set $\lambda \subseteq \{1, \cdots, k\}$. Sometimes, we also write the edge as $s \rightarrow^l_\lambda s'$, or simply $s \rightarrow_\lambda s'$ when l is $true$. Without loss of generality, we assume that $|\lambda| \leq 1$. That is, each transition resets at most one clock. (Resetting several clocks can be simulated by resetting one by one.) When $\lambda = \emptyset$, the edge is called a progress transition. Otherwise, it is a reset transition. $\mathcal{D}$ is *static* if the enabling condition on each edge is simply $true$.

The semantics of $\mathcal{D}$ is defined as follows. A *configuration* is a tuple of a control state and clock values. Let $\langle s, v_1, \cdots, v_k\rangle$ and $\langle s', v'_1, \cdots, v'_k\rangle$ be two configurations. $\langle s, v_1, \cdots, v_k\rangle \rightarrow \langle s', v'_1, \cdots, v'_k\rangle$ denotes a *one-step transition* satisfying all of the following conditions:

- There is an edge $\langle s, \lambda, l, s'\rangle$ in $\mathcal{A}$,
- The enabling condition of the edge is satisfied; i.e., $l(v_1, \cdots, v_k)$ is true,
- If $\lambda = \emptyset$ (i.e., a progress transition), then every clock progresses by one time unit; i.e., $v'_i = v_i + 1$, $1 \leq i \leq k$, *(iv)*. If for some j, $\lambda = \{j\}$ (i.e., a reset transition), then x_j resets to 0 and all the other clocks do not change; i.e., $v'_j = 0$ and $v'_i = v_i$ for each $1 \leq i \neq j \leq k$.

We say that $\langle s, v_1, \cdots, v_k\rangle$ *reaches* $\langle s', v'_1, \cdots, v'_k\rangle$ if

$$\langle s, v_1, \cdots, v_k\rangle \rightarrow^* \langle s', v'_1, \cdots, v'_k\rangle,$$

where $\rightarrow^*$ is the transitive closure of $\rightarrow$.

The *linear reachability problem for discrete timed automata* is defined as follows.

- **Given:** A discrete timed automaton $\mathcal{D}$, two designated states s_{init} and s_{final}, and two linear constraints U and U' over k variables.
- **Question:** are there clock values $v_1, \cdots, v_k, v'_1, \cdots, v'_k$ such that $\langle s_{\text{init}}, v_1, \cdots, v_k\rangle$ reaches $\langle s_{\text{final}}, v'_1, \cdots, v'_k\rangle$ and both $U(v_1, \cdots, v_k)$ and $U'(v'_1, \cdots, v'_k)$ hold?

The problem is decidable, even when clocks are dense. Its decidability proofs and application examples can be found in [8, 10, 9]. However, as we mentioned earlier, the time complexity for the problem is unknown. Using (18), we will obtain a complexity bound in this section.

Without loss of generality, we assume that both U and U' in the linear reachability problem for timed automata are a disjunction of m conjunctive linear constraints.

Each conjunctive linear constraint contains at most L atomic linear constraints among which there are at most E equations. Similar to Section 3, we use W_1 (resp. W_2) to represent the maximal value $||\mathbf{B}||_{1,\infty}$ (resp. $||\mathbf{b}||_\infty$) of all the conjunctive linear constraints $\mathbf{Bx} \sim \mathbf{b}$ in U and U'. The complexity of the linear reachability problem will be measured on, among others, L, E, m, W_1, W_2, $|S|$, and k.

We first consider a simpler case when $\mathcal{D}$ is static. Before we proceed further, more definitions are needed. A *reset order* τ is a sequence $\lambda_1, \cdots, \lambda_n$, for some $1 \leq n \leq k$, where each λ_i contains exactly one element in $\{1, \cdots, k\}$, and all of the λ_i's are pairwisely disjoint. Let $\lambda_0 = \{1, \cdots, k\} - \cup_{1 \leq i \leq n} \lambda_i$. An execution path of $\mathcal{D}$ is *of reset order* τ if every clock in λ_0 does not reset on p, and for rest of the clocks, their last resets are in this order: $x_{i_1}, \cdots, x_{i_n}$, with $\lambda_1 = \{i_1\}, \cdots, \lambda_n = \{i_n\}$. For the instance of the linear reachability problem of static $\mathcal{D}$, we consider the **Question**-part witnessed by an execution path that is of any fixed reset order τ (there are only finitely many reset orders). From this instance and the given τ, we will construct an ordered finite state transition system $\mathcal{A}^\tau$ and a linear constraint U^τ. Then, we reduce the linear reachability problem of $\mathcal{D}$ to the linear reachability problem of $\mathcal{A}^\tau$. The key idea behind the construction is as follows. Suppose $\lambda_0 \neq \emptyset$. The execution path can then be partitioned into $n+1$ segments separated by the n last resets given in τ. We use $y_0, y_1, \cdots, y_n$ to denote the number of progress transitions made on each segment respectively. Suppose that the path starts with clock values $z_1, \cdots, z_k$ and ends with clock values $x_1, \cdots, x_k$. Observe that each x_i can be represented as a sum on (some of) $z_1, \cdots, z_k$ and $y_0, y_1, \cdots, y_n$. The case when $\lambda_0 = \emptyset$ is similar. Following this line, one can show:

Lemma 3. *The linear reachability problem for static discrete timed automata $\mathcal{D}$ is solvable in time*

$$O(k! \cdot m^2 \cdot (k + (k+1) \cdot |S|)^{8k-2} \cdot (2 + k \cdot W_1 + W_2)^{(2k+2L+2E) \cdot 4k}). \tag{20}$$

Hence, when $|S| \gg k, m, W_1, W_2$, the linear reachability problem for static $\mathcal{D}$ is solvable in time

$$O(|S|^{8k-1}). \tag{21}$$

Now, we consider the case when $\mathcal{D}$ is not necessarily static. Let C be one plus the maximal absolute value of all the constants appearing in enabling conditions in $\mathcal{D}$. We use T to denote the result of (20) after replacing $|S|$ with $(1+2C)^{k^2+k} \cdot |S|$, L with $L+k$, E with $E+k$, W_1 with $\max(W_1, 2)$, W_2 with $\max(W_2, C)$. From [12, 10], one can construct a static $\mathcal{D}'$ with two designated states s'_{init} and s'_{final} and with at most $(1+2C)^{k^2+k} \cdot |S|$ number of states to simulate $\mathcal{D}$ faithfully. From this result, one can show, using Lemma 3,

Theorem 5. *The linear reachability problem for discrete timed automata is solvable in time*

$$O(k! \cdot (1+C)^k \cdot T). \tag{22}$$

Again, when $|S|$ and C are $\gg$ the size of U and U', the time complexity of linear reachability problem for discrete timed automata is

$$O(|S|^{8k-1} \cdot C^{k+(k^2+k)\cdot(8k-2)+(6k+2L+2E)\cdot 4k}). \tag{23}$$

Remark 2. Using the techniques in Section 4, one can obtain a complexity bound similar to (23) for the linear liveness problem [12] for discrete timed automata. Also the complexity in (23) is more sensitive to C than to $|S|$.

We turn now to the case when $\mathcal{D}$ is a timed automaton with k dense clocks. One can similarly formulate the semantics and the linear reachability problem for $\mathcal{D}$ (e.g., see [9]). With the pattern technique presented in [9], it is easy to show the following. From $\mathcal{D}$ and U, U', one can construct a discrete timed automaton $\mathcal{D}'$ with k discrete clocks and two linear constraints W, W' such that the linear reachability problem of timed automaton $\mathcal{D}$ concerning U, U' is equivalent to the linear reachability problem of discrete timed automaton $\mathcal{D}'$ concerning W, W'. In addition, the number of states in $\mathcal{D}'$ is $O(2^{6(k+1)^2} \cdot |S|)$, where S is the state set in $\mathcal{D}$. (There are at most $2^{6(k+1)^2}$ *patterns* [9].) Furthermore, W and W' only depend on U, U' and k (independent of $\mathcal{D}$). Hence, we have the following conclusion:

Theorem 6. *The linear reachability problem for timed automata with dense clocks can still be solvable in time shown in (23), when $|S|, C \gg k$ and the size of U.*

7 Conclusions

We obtained a number of new complexity results for various linear counting problems (reachability and liveness) for (ordered) finite state transition systems and timed automata. At the heart of the proofs, we used some known results in estimating the upper bound for minimal solutions (in nonnegative integers) for linear Diophantine systems. In particular, when all the parameters (such as the number of labels/clocks, the largest constant C in a timed automaton, the size of the linear constraint to be verified, etc) except the number of states $|S|$ of the underlying transition system are considered constants, all of the complexity bounds obtained in this paper is polynomial in $|S|$. This is, as we mentioned in Section 1, in contrast to the exponential bounds that were previously known. In practice, a requirement specification (e.g., the U in a linear counting problem) is usually small and simple [14]. In this sense, our results are meaningful, since the large size of $|S|$ is usually the dominant factor in efficiently solving these verification problems.

The counts of labels in a finite state transition system can be regarded as monotonic counters. Therefore, our result also provides a technique to verify safety properties for a special class of counter machines M with monotonic counters: starting from a state s, whether it is always true that the counter values in M satisfy a given linear constraint whenever M reaches s'. In the future, we will study whether the techniques in this paper can be generalized to handle the case when M is further augmented with an unrestricted counter (i.e., can be incremented, decremented, and tested against 0) or even a pushdown stack. Additionally, it is also desirable to study whether our techniques can be further generalized to the reachability problem of some classes of Petri nets [18].

The authors would like to thank P. San Pietro and the anonymous referees for many valuable suggestions.

References

1. R. Alur and D. L. Dill. A theory of timed automata. *Theoretical Computer Science*, 126(2):183–235, April 1994.
2. R. Alur and T. A. Henzinger. A really temporal logic. *Journal of the ACM*, 41(1):181–204, January 1994.
3. A. Biere, A. Cimatti, E. Clarke, and Y. Zhu. Symbolic model checking without BDDs. In *TACAS'99*, volume 1579 of *LNCS*, pages 193–207. Springer-Verlag, 1999.
4. I. Borosh, M. Flahive, and B. Treybig. Small solutions of linear diophantine equations. *Discrete Mathematics*, 58:215–220, 1986.
5. I. Borosh and B. Treybig. Bounds on positive integral solutions of linear diophantine equations. *Proceedings of the American Mathematical Society*, 55:299–304, 1976.
6. A. Bouajjani, R. Echahed, and P. Habermehl. On the verification problem of nonregular properties for nonregular processes. In *LICS'95*, pages 123–133. IEEE CS Press, 1995.
7. E. M. Clarke, E. A. Emerson, and A. P. Sistla. Automatic verification of finite-state concurrent systems using temporal logic specifications. *TOPLAS*, 8(2):244–263, 1986.
8. H. Comon and Y. Jurski. Timed automata and the theory of real numbers. In *CONCUR'99*, volume 1664 of *LNCS*, pages 242–257. Springer, 1999.
9. Zhe Dang. Binary reachability analysis of pushdown timed automata with dense clocks. In *CAV'01*, volume 2102 of *LNCS*, pages 506–517. Springer, 2001.
10. Zhe Dang, O. H. Ibarra, T. Bultan, R. A. Kemmerer, and J. Su. Binary reachability analysis of discrete pushdown timed automata. In *CAV'00*, volume 1855 of *LNCS*, pages 69–84. Springer, 2000.
11. Zhe Dang, O. H. Ibarra, and P. San Pietro. Liveness Verification of Reversal-bounded Multicounter Machines with a Free Counter. In *FSTTCS'01*, volume 2245 of *LNCS*, pages 132–143. Springer, 2001.
12. Zhe Dang, P. San Pietro, and R. A. Kemmerer. On Presburger Liveness of Discrete Timed Automata. In *STACS'01*, volume 2010 of *LNCS*, pages 132–143. Springer, 2001.
13. E. Domenjoud. Solving systems of linear diophantine equations: an algebraic approach. In *MFCS'91*, volume 520 of *LNCS*, pages 141–150. Springer-Verlag, 1991.
14. Matthew B. Dwyer, George S. Avrunin, and James C. Corbett. Patterns in property specifications for finite-state verification. In *ICSE'99*, pages 411–421. ACM Press, 1999.
15. G. J. Holzmann. The model checker SPIN. *TSE*, 23(5):279–295, May 1997.
16. J. Hopcroft and J. Ullman. *Introduction to Automata theory, Languages, and Computation*. Addison-Wesley Publishing Company, 1979.
17. K.L. McMillan. *Symbolic Model Checking*. Kluwer Academic Publishers, Norwell Massachusetts, 1993.
18. Tadao Murata. Petri nets: Properties, analysis and applications. *Proceedings of the IEEE*, 77(4):541–580, 1989.
19. R. Parikh. On context-free languages. *JACM*, 13:570–581, 1966.
20. A. Pnueli. The temporal logic of programs. In *FOCS'77*, pages 46–57. IEEE CS Press, 1977.
21. L. Pottier. Minimal solutions of linear diophantine equations: Bounds and algorithms. In *Rewriting Techniques and Applications*, volume 488 of *LNCS*, pages 162–173. Springer-Verlag, 1991.
22. M. Y. Vardi and P. Wolper. An automata-theoretic approach to automatic program verification (preliminary report). In *LICS'86*, pages 332–344. IEEE CS Press, 1986.