Learning of Event-Recording Automata
Olga Grinchtein, Bengt Jonsson, Martin Leucker

To cite this version:

HAL Id: inria-00459696
https://hal.inria.fr/inria-00459696
Submitted on 24 Feb 2010

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
Learning of Event-Recording Automata

Olga Grinchtein\textsuperscript{a} Bengt Jonsson\textsuperscript{a} Martin Leucker\textsuperscript{b}

\textsuperscript{a}Department of Information Technology
Uppsala University
Uppsala, Sweden
\{olgag, bengt\}@it.uu.se

\textsuperscript{b}Institut für Informatik
TU München
München, Germany
leucker@in.tum.de

Abstract

In regular inference, a regular language is inferred from answers to a finite set of membership queries, each of which asks whether the language contains a certain word. One of the most well-known regular inference algorithms is the $L^*$ algorithm due to Dana Angluin. However, there are almost no extensions of these algorithms to the setting of timed systems. We extend Angluin’s algorithm for on-line learning of regular languages to the setting of timed systems. Since timed automata can freely use an arbitrary number of clocks, we restrict our attention to systems that can be described by deterministic event-recording automata (DERAs). We present three algorithms, $TL^*_\text{sg}$, $TL^*_\text{nsg}$ and $TL^*_s$, for inference of DERAs. In $TL^*_\text{sg}$ and $TL^*_\text{nsg}$, we further restrict event-recording automata to be event-deterministic in the sense that each state has at most one outgoing transition per action; learning such an automaton becomes significantly more tractable. The algorithm $TL^*_\text{nsg}$ builds on $TL^*_\text{sg}$, by attempts to construct a smaller (in number of locations) automaton. Finally, $TL^*_s$ is a learning algorithm for a full class of deterministic event-recording automata, which infers a so called simple DERA, which is similar in spirit to the region graph.

Key words: model inference, model learning, timed systems

1 Introduction

Research during the last decades have developed powerful techniques for using models of reactive systems in specification, automated verification (e.g., [CGP99]), test case generation (e.g., [FJJV97,SEG00]), implementation (e.g., [HLN+90]), and validation of reactive systems in telecommunication, embedded control, and related application areas. Typically, such models are assumed to be developed \textit{a priori} during the specification and design phases of system development.

Preprint submitted to Elsevier Science 21 February 2010
In practice, however, often no formal specification is available, or becomes outdated as the system evolves over time. One must then construct a model that describes the behavior of an existing system or implementation. In software verification, techniques are being developed for generating abstract models of software modules by static analysis of source code (e.g., [CDH+00, Hol00]). However, peripheral hardware components, library modules, or third-party software systems do not allow static analysis. In practice, such systems must be analyzed by observing their external behavior. In fact, techniques for constructing models by analysis of externally observable behavior (black-box techniques) can be used in many situations.

- To create models of hardware components, library modules, that are part of a larger system which, e.g., is to be formally verified or analyzed.
- For regression testing, a model of an earlier version of an implemented system can be used to create a good test suite and test oracle for testing subsequent versions. This has been demonstrated, e.g., by Hungar et al. [HHNS02, HNS03].
- Black-box techniques, such as adaptive model checking [GPY02], have been developed to check correctness properties, even when source code or formal models are not available.
- Tools that analyze the source code statically depend heavily on the implementation language used. Black-box techniques are easier to adapt to modules written in different languages.

The construction of models from observations of system behavior can be seen as a learning problem. For finite-state reactive systems, it means to construct a (deterministic) finite automaton from the answers to a finite set of membership queries, each of which asks whether a certain word is accepted by the automaton or not. There are several techniques (e.g., [Ang87, Gol67, KV94, RS93, BDG97]) which use essentially the same basic principles; they differ in how membership queries may be chosen and in exactly how an automaton is constructed from the answers. The techniques guarantee that a correct automaton will be constructed if “enough” information is obtained. In order to check this, Angluin and others also allow equivalence queries that ask whether a hypothesized automaton accepts the correct language; such a query is answered either by yes or by a counterexample on which the hypothesis and the correct language disagree. Techniques for learning finite automata have been successfully used for regression testing [HHNS02] and model checking [GPY02] of finite-state systems for which no model or source code is available.

In this paper, we extend the techniques for automata learning developed by Angluin and others to the setting of timed systems. One longer-term goal is to develop techniques for creating abstract timed models of hardware components, device drivers, etc. for analysis of timed reactive systems; there are many other analogous applications. It is not an easy challenge, and we will therefore
in this first work make some idealizing assumptions. We assume that a learning algorithm observes a system by checking whether certain actions can be performed at certain moments in time, and that the learner is able to control and record precisely the timing of the occurrence of each action. We consider systems that can be described by a timed automaton [AD94], i.e., a finite automaton equipped with clocks that constrain the possible absolute times of occurrences of actions. There are some properties of timed automata that make the design of learning algorithms difficult: the set of clocks is not known \textit{a priori}, and they cannot in general be determinized [AD94]. We therefore restrict consideration to a class of \textit{event-recording automata} [AFH99]. These are timed automata that, for every action \(a\), use a clock that records the time of the last occurrence of \(a\). Event-recording automata can be determinized, and are sufficiently expressive to model many interesting timed systems; for instance, they are as powerful as timed transition systems [HMP94,AFH99], another popular model for timed systems.

Although event-recording automata overcome some obstacles of timed automata, they still suffer from problems. One problem is that it is not clear how to generalize Nerode’s right congruence, another is that in general they do not have canonical forms. Therefore we work with classes of event-recording automata which have canonical forms and can be understood as finite automata over a symbolic alphabet.

We present three algorithms, \(TL^*_{sg}\), \(TL^*_{nsg}\) and \(TL^*_s\), for learning deterministic event-recording automata.

In algorithms \(TL^*_{sg}\) and \(TL^*_{nsg}\), we further restrict event-recording automata to be event-deterministic in the sense that each state has at most one outgoing transition per action (i.e., the automaton obtained by removing the clock constraints is deterministic). Under this restriction, timing constraints for the occurrence of an action depend only on the past sequence of actions, and not on their relative timing; learning such an automaton becomes significantly more tractable, and allows us to adapt the learning algorithm of Angluin to the timed setting.

\(TL^*_{sg}\) learns a so-called \textit{sharply guarded} event-deterministic event-recording automaton. We show that every deterministic event-recording automaton can be transformed into a unique sharply guarded one with at most double exponentially more locations. We show that if the size of the untimed alphabet is fixed, then the number of membership queries of \(TL^*_{sg}\) is polynomial in the size of the biggest constant appearing in guards, in the number \(n\) of locations of the sharply guarded event-deterministic event-recording automaton, in the size of the timed alphabet and in the length of the longest counterexample. The number of equivalence queries is at most \(n\).
The algorithm $TL^*_{nsg}$ addresses the problem of learning a smaller, not necessarily sharply guarded version of an event-deterministic event-recording automaton. It achieves this goal by unifying the queried information when it is “similar” which results in merging states in the constructed automaton. The number of needed queries exceeds those of $TL^*_{sg}$ in the worst case; however, in practice it can be expected that it behaves better than $TL^*_{sg}$.

$TL^*_s$ is a learning algorithm for the full class of deterministic event-recording automata. While we reuse the prosperous scheme developed in $TL^*_{sg}$, the details are different. We work out a characterization in terms of a (symbolic) regular language for the language of DERAs. Furthermore, we show that each symbolic word can be identified by a single timed word. Thus, one query in Angluin’s algorithm relates to a single timed query. $TL^*_s$ learns a so-called simple deterministic event-recording automaton. We show that every deterministic event-recording automaton can be transformed into a unique simple one with at most single exponentially more locations. Our transformation is based on ideas used to derive so-called region graphs. We show that the number of membership queries of $TL^*_s$ is polynomial in the size of the biggest constant appearing in guards, in the number $n$ of locations of the simple deterministic event-recording automaton, in the size of the untimed alphabet and in the length of the longest counterexample. The number of equivalence queries is at most $n$.

Related Work In another work [GJP06,Gri08], two of the authors of this paper have developed a completely different algorithm for learning deterministic event-recording automata. The algorithm differs from $TL^*_s$ in that the constructed automaton need not be a simple one. The transformation of an event-recording automaton to a corresponding simple automaton often increases its size significantly; a reason for this increase is that each transition is divided up into many “smaller” transitions, at least one transition for each possible integer value of the clocks of the automaton. The algorithm presented in [GJP06,Gri08] attempts to avoid this division of transitions by constructing an automaton, whose guards are “similar” to the guards of the event-recording automaton that is being learned. Main problems are then to synthesize “natural” guards on transitions, and to construct locations (control states). To address these problems, the algorithm does not re-use the structure of the $L^*$ algorithm; instead of the observation table used in $L^*$, it uses a new data structure, called timed decision tree, to organize results of queries. Its theoretical worst-case complexity is significantly higher than that of $TL^*_s$.

The only other work on learning of timed systems we are aware is by Verwer, de Weerdt and Witteveen [VdWW06], who present an algorithm for learning of timed automata with one clock which is reset at every transition. Their setting differs from ours in that the learner does not choose the words to be
used in membership queries, but is given a sample of accepted and unaccepted
timed words from which an automaton is to be constructed. The algorithm
constructs a prefix tree from the sample of timed words and then tries to merge
nodes of this tree pairwise to form an automaton. If the resulting automaton
does not agree with the sample then the last merge is undone and a new
merge is attempted. The algorithm does not construct timed automata in a
systematic way, and it is hard to generalize the algorithm to timed automata
with more than one clock.

Previous work on learning of infinite-state systems include the work by Berg
et al. [BJR08], who consider Mealy machines extended with state variables
that can assume values from a potentially unbounded domain. These values
can be passed as parameters in input and output messages, and can be used
in tests for equality between state variables and/or message parameters. Their
technique first uses the $L^*$ algorithm to generate a finite-state Mealy machine
for the case that the values are taken from a finite data domain. This finite-
state Mealy machine is then transformed into a symbolic form, representing
the desired infinite-state Mealy machine.

A problem related to automata learning (or regular inference) is that of confor-
mcance testing, where one is given an automaton specification of the intended
behavior of an implementation, and would like to derive a test suite which
checks that an implementation conforms to such a specification. In previous
work [BGJ+05], we showed that if a set of input words form a conformance
test suite for a finite state machine, then that state machine can be inferred
from this set of input words using automata learning techniques. Conversely, if
a finite state machine is uniquelly inferred from a set of input words which are
supplied in membership queries, then this set of input words forms a confor-
mcance test suite for the state machine. Springintveld et al. [SVD01] introduces
an algorithm which generates a conformance test suite for timed automata.
The algorithm constructs grid automata, which only contain states in which
every clock value is from the set of integer multiples of $2^{-n}$ for some sufficiently
large natural number $n$. Then the Vasilevskii-Chow algorithm [Cho78], [Vas73]
for generating conformance test suites for finite state machines, is applied to
the grid automaton to generate a conformance test suite for the timed au-
tomaton.

Several papers are concerned with finding a definition of timed languages
which is suitable as a basis for learning. There are several works that de-
fine determinizable classes of timed automata (e.g., [AFH99,SV96]) and right-
congruences of timed languages (e.g., [MP04,HRS98,Wil94]), motivated by
testing and verification.
Structure of Paper  The paper is structured as follows. After preliminaries in the next section, we define deterministic event-recording automata (DERAs) in Section 3. In Section 4 we describe the $L^*$ algorithm for learning DFAs. In Section 5 and Section 7, we present the algorithms $TL^*_{sg}$ and, respectively, $TL^*_{nsg}$ for learning event-deterministic DERAs (EDERAs). In Section 6 we describe the algorithm $TL^*_s$ suitable for learning general DERAs. Section 8 presents conclusions and directions for future research.

2 Preliminaries

We write $\mathbb{R}^{\geq 0}$ for the set of nonnegative real numbers, and $\mathbb{N}$ for the set of natural numbers. Let $\Sigma$ be a finite alphabet of size $|\Sigma|$. A timed word over $\Sigma$ is a finite sequence $w_t = (a_1, t_1)(a_2, t_2) \ldots (a_n, t_n)$ of symbols $a_i \in \Sigma$ that are paired with nonnegative real numbers $t_i$ such that the sequence $t_1t_2 \ldots t_n$ of time-stamps is nondecreasing. Each time-stamp $t_i$ denotes the time of occurrence of the symbol $a_i$, measured from some common “initial moment”. (1) We use $\lambda$ to denote the empty word. A timed language over $\Sigma$ is a set of timed words over $\Sigma$.

An event-recording automaton contains for every symbol $a \in \Sigma$ a clock $x_a$, called the event-recording clock of $a$. Intuitively, $x_a$ records the time elapsed since the last occurrence of the symbol $a$. We write $C_\Sigma$ for the set \{ $x_a$ $|$ $a \in \Sigma$ \} of event-recording clocks.

A clock valuation $\gamma$ is a mapping from $C_\Sigma$ to $\mathbb{R}^{\geq 0}$. For $a \in \Sigma$, we define $\gamma[x_a \mapsto 0]$ to be the clock valuation $\gamma'$ such that $\gamma'(x_a) = 0$ and $\gamma'(x_b) = \gamma(x_b)$ for all $b \neq a, b \in \Sigma$. For $t \in \mathbb{R}^{\geq 0}$, we define $\gamma + t$ to be the clock valuation $\gamma'$ such that $\gamma'(x_a) = \gamma(x_a) + t$ for all $a \in \Sigma$.

Throughout the paper, we will use an alternative, equivalent representation of timed words, namely clocked words. A clocked word $w_c$ is a sequence $w_c = (a_1, \gamma_1)(a_2, \gamma_2) \ldots (a_n, \gamma_n)$ of symbols $a_i \in \Sigma$ that are paired with clock valuations, which for all $a \in \Sigma$ satisfies

- $\gamma_1(x_a) = \gamma_1(x_b)$ for all $a, b \in \Sigma$, and
- $\gamma_i(x_a) = \gamma_{i-1}(x_a) + \gamma_i(x_{a_{i-1}})$ whenever $1 < i \leq n$ and $a \neq a_{i-1}$. The quantity $\gamma_i(x_{a_{i-1}})$ is the time elapsed between the occurrence of $a_{i-1}$ and $a_i$, and so can be arbitrary. (2)

Each timed word $w_t = (a_1, t_1)(a_2, t_2) \ldots (a_n, t_n)$ can be naturally transformed into a clocked word $CW(w_t) = (a_1, \gamma_1)(a_2, \gamma_2) \ldots (a_n, \gamma_n)$ where for each $i$ with $1 \leq i \leq n$, (2)

(1) TODO: This sentence was new/BJ

(2) TODO: This sentence was new/BJ
A clock constraint is a conjunction of atomic constraints of the form $x_a \sim n$, called a clock bound, or $x_a - x_b \sim n$, called a difference bound, for $x_a, x_b \in C_\Sigma$, $\sim \in \{<, \leq, \geq, >\}$, and $n \in \mathbb{N}$. A clock constraint is called non-strict if only $\sim \in \{\leq, \geq\}$ is used, and, similarly, it is called strict if only $\sim \in \{<, >\}$ is used. For example, $\varphi = x_a - x_b \geq 1 \land x_a - x_b \leq 7 \land x_a > 4 \land x_b > 1$ is a clock constraint, which is neither strict nor non-strict. We identify an empty conjunction with $true$.

We use $\gamma \models \varphi$ to denote that the clock valuation $\gamma$ satisfies the clock constraint $\varphi$, defined in the usual manner. A clock constraint $\varphi$ identifies a $|\Sigma|$-dimensional polyhedron $[\varphi] \subseteq (\mathbb{R}_{\geq 0})^{|\Sigma|}$ viz. the vectors of real numbers satisfying the constraint. In Figure 1(a), a clock constraint and the 2-dimensional polyhedron it identifies are shown.

For each clock constraint $\varphi$ there are in general several other clock constraints that are equivalent to $\varphi$ in the sense that they identify the same polyhedron. If $\varphi$ is satisfiable, there is among these a unique canonical clock constraint, denoted $\text{Can}(\varphi)$, obtained by closing $\varphi$ under all consequences of pairs of conjuncts in $\varphi$, i.e.,

- from two difference bounds, such as $x_a - x_b \leq 2$ and $x_b - x_c < 3$, we derive a new difference bound, viz. $x_a - x_c < 5$, and
- from a difference bound and a clock bound, such as $x_a - x_b \leq 2$ and $x_a \geq 3$, we derive a new clock bound, viz. $x_b \geq 1$,
- from an upper and a lower clock bound, such as $x_a \leq 3$ and $x_b > 2$, we derive a new difference bound, viz. $x_a - x_b < 1$,

until saturation, and thereafter keeping the tightest bounds for each clock and
each clock difference. If the canonical form contains inconsistent constraints, or requires some clock to be negative, then the clock constraint is unsatisfiable. The canonical form for an unsatisfiable clock constraint is defined to be false [Dil89].

Clock constraints satisfy an important closure property

**Proposition 1** For a clock constraint $\varphi$ and two clock valuations $\gamma$, $\gamma'$, if $\gamma \models \varphi$ and $\gamma' \models \varphi$, then $\min(\gamma, \gamma') \models \varphi$ and $\max(\gamma, \gamma') \models \varphi$, where $\min(\gamma, \gamma')$ is defined by $\min(\gamma, \gamma') = \min(\gamma(x_a), \gamma'(x_a))$ for all $a \in \Sigma$, and analogously for $\max(\gamma, \gamma')$ (see Figure 2).

**PROOF.** (* Here is an alternative proof/BJ *)

That $\min(\gamma, \gamma')$ satisfies a clock bound of form $x_a \sim n$ follows from the fact that $\min(\gamma, \gamma')(x_a)$ is either $\gamma(x_a)$ or $\gamma'(x_a)$, and that both $\gamma$ and $\gamma'$ satisfy $x_a \sim n$. To see that $\min(\gamma, \gamma')$ satisfies a difference bound of form $x_a - x_b \geq n$, assume that $\min(\gamma, \gamma')(x_a)$ is $\gamma(x_a)$. Then

$$\min(\gamma, \gamma')(x_a) - \min(\gamma, \gamma')(x_b) = \gamma(x_a) - \min(\gamma, \gamma')(x_b) \geq \gamma(x_a) - \gamma(x_b),$$

and by assumption $\gamma(x_a) - \gamma(x_b) \geq n$. The proof is analogous for difference bounds of form $x_a - x_b \leq n$ and for strict difference bounds. □

**PROOF.** Assume that a clock constraint $\varphi$ contains constraints of the form $k_1 \leq x_a \leq k_2$, $k_3 \leq x_b \leq k_4$ and $k_5 \leq x_a - x_b \leq k_6$. Since $\gamma \models \varphi$ and $\gamma' \models \varphi$, we have $k_1 \leq \min(\gamma, \gamma')(x_a) \leq k_2$, $k_1 \leq \max(\gamma, \gamma')(x_a) \leq k_2$, $k_3 \leq \min(\gamma, \gamma')(x_b) \leq k_4$ and $k_3 \leq \max(\gamma, \gamma')(x_b) \leq k_4$. It also follows that $k_5 \leq \gamma(x_a) - \gamma(x_b) \leq k_6$ and $k_5 \leq \gamma'(x_a) - \gamma'(x_b) \leq k_6$. If $\gamma(x_a) \leq \gamma'(x_a)$ and $\gamma(x_b) \leq \gamma'(x_b)$, then $k_5 \leq \min(\gamma, \gamma')(x_a) - \min(\gamma, \gamma')(x_b) \leq k_6$ and $k_5 \leq \max(\gamma, \gamma')(x_a) - \max(\gamma, \gamma')(x_b) \leq k_6$. Assume that $\gamma(x_a) \leq \gamma'(x_a)$ and $\gamma(x_b) \geq \gamma'(x_b)$ (proof for opposite case is similar). Then $k_5 \leq \gamma(x_a) - \gamma(x_b) \leq \min(\gamma, \gamma')(x_a) - \min(\gamma, \gamma')(x_b) \leq \gamma'(x_a) - \gamma'(x_b) \leq k_6$ and $k_5 \leq \gamma(x_a) - \gamma(x_b) \leq \max(\gamma, \gamma')(x_a) - \max(\gamma, \gamma')(x_b) \leq \gamma'(x_a) - \gamma'(x_b) \leq k_6$. □
A clock guard is a conjunction of atomic constraints of the form \( x_a \sim n \), for \( x_a \in C_S, \sim \in \{<, \leq, \geq, >\} \), and \( n \in \mathbb{N}, \) i.e., comparison between clocks is not permitted. A clock guard is called non-strict if only \( \sim \in \{\leq, \geq\} \) is used and strict if only \( \sim \in \{<, >\} \) is used. For example, \( x_a \geq 4 \land x_a \leq 8 \land x_b \geq 1 \land x_b \leq 3 \) is a clock guard, which is non-strict.

The set of clock guards is denoted by \( G_\Sigma \). A clock guard \( g \) identifies a \( |\Sigma| \)-dimensional hypercube \( [g] \subseteq (\mathbb{R}^{\geq 0})^{|\Sigma|} \). In Figure 3(a), two different clock guards \( g \) and \( g' \) and two 2-dimensional hypercubes—rectangles—they identify are shown: the bounded one for \( g \) and the partially unbounded one for \( g' \).

We use equalities in clock constraints and clock guards in the natural way, e.g., \( x_a = n \) denotes \( x_a \geq n \land x_a \leq n \). In timed automata, guards are restricted to being clock guards: one reason is that this is sufficient for many applications, another is allowing them to be general clock constraints would make analysis algorithms more complicated (see e.g., [BY03]). (3)

A guarded word is a sequence \( w_g = (a_1, g_1)(a_2, g_2)\ldots(a_n, g_n) \) of symbols \( a_i \in \Sigma \) that are paired with clock guards. For a clocked word \( w_c = (a_1, \gamma_1)(a_2, \gamma_2)\ldots(a_n, \gamma_n) \) we use \( w_c \models w_g \) to denote that \( \gamma_i \models g_i \) for \( 1 \leq i \leq n \). For a timed word \( w_t \) we use \( w_t \models w_g \) to denote that \( CW(w_t) \models w_g \). A guarded word \( w_g = (a_1, g_1)(a_2, g_2)\ldots(a_n, g_n) \) is called a guard refinement of \( a_1a_2\ldots a_n \), and \( a_1a_2\ldots a_n \) is called the word underlying \( w_g \). The word \( w \) underlying a timed word \( w_t \) is defined in a similar manner. A guarded word \( w_g = (a_1, g_1)(a_2, g_2)\ldots(a_n, g_n) \) is non-strict if \( g_i \) is non-strict for all \( 1 \leq i \leq n \). (4)

A clock constraint or a clock guard is \( K \)-bounded if it contains no constant larger than \( K \). A \( K \)-bounded simple clock guard is a clock guard whose conjuncts are only of the form \( x_a = n, n' < x_a < n' + 1 \) or \( x_a > K \), for \( 0 \leq n \leq K, 0 \leq n' \leq K - 1, x_a \in C_S \). In Figure 3(b), an example of a simple clock guard is shown. A \( K \)-bounded simple guarded word \( w_g \) is a sequence \( w_g = (a_1, g_1)(a_2, g_2)\ldots(a_n, g_n) \) of symbols \( a_i \in \Sigma \) that are paired with \( K \)-bounded simple clock guards.

The extracted guard from a clock constraint \( \varphi \), denoted \( guard(\varphi) \) is the conjunction of all clock bounds (i.e., conjuncts of form \( x_a \sim n \)) in \( Can(\varphi) \). In simple words, \( guard(\varphi) \) identifies the smallest hypercube surrounding the polyhedron identified by \( \varphi \). If \( \varphi \) is unsatisfiable, \( guard(\varphi) \) is defined as \( false \). This intuition immediately leads to the following proposition.

**Proposition 2** Let \( \varphi \) be a clock constraint and \( g \) be a clock guard. Then \( [\varphi \land g] = [\varphi] \) implies \( [guard(\varphi)] \subseteq [g] \). \( \square \)

For the developments to come, we define several operations on clock constraints \( \varphi \).
We define the *reset* of a clock $x_a$ in $\varphi$, denoted by $\varphi[x_a \mapsto 0]$, as $\text{Can}(\varphi')$, where $\varphi'$ is obtained from $\text{Can}(\varphi)$ by removing all conjuncts involving $x_a$, and adding the conjunct $x_a \leq 0$.

We define the *time elapsing* of $\varphi$, denoted by $\varphi \uparrow$, as $\text{Can}(\varphi')$, where $\varphi'$ is obtained from $\text{Can}(\varphi)$ by removing all upper bounds on clocks [DT98].

It is a standard result that these operations mirror the corresponding operations on clock valuations, in the sense that

- $\gamma' \models \varphi[x_a \mapsto 0]$ iff $\gamma' = \gamma[x_a \mapsto 0]$ for some $\gamma$ with $\gamma \models \varphi$, and
- $\gamma' \models \varphi \uparrow$ iff $\gamma' = \gamma + t$ for some $\gamma$ with $\gamma \models \varphi$ and $t \in \mathbb{R}^\geq 0$.

Following [DT98], we introduce the $K<$-approximation $\langle\langle \varphi \rangle\rangle^<_K$ of the clock constraint $\varphi$ as the constraint $\varphi'$ obtained from $\text{Can}(\varphi)$ by (5)

- removing all constraints of the form $x_a \sim n$ and $x_a - x_b \sim n$, whenever $\sim \in \{<, \leq\}$ and $n > K$, and
- replacing all constraints of the form $x_a \sim n$ and $x_a - x_b \sim n$ by $x_a > K$ and $x_a - x_b > K$, respectively, whenever $\sim \in \{>, \geq\}$ and $n > K$.

Note that the $K<$-approximation of a canonical clock constraint is in general not canonical.

We introduce the $K\leq$-approximation $\langle\langle g \rangle\rangle^\leq_K$ of the clock guard $g$ as the clock guard obtained by

- removing all constraints of form $x_a \leq n$ and $x_a < n$, whenever $n > K$, and
- replacing all constraints of form $x_a \geq n$ and $x_a > n$ by $x_a \geq K$, whenever $n > K$.

For a constraint $\varphi$ and guarded word $w_g$, we introduce the *strongest postcondition* of $w_g$ with respect to $\varphi$, denoted by $sp(\varphi, w_g)$. Postcondition computation is central in symbolic verification techniques for timed automata.
Fig. 4. $w_g = (a, true)(b, 2 \leq x_a \leq 4)(b, 1 \leq x_b \leq 3)$

[BDM+98, BLL+96], and can be done inductively as follows:

- $sp(\varphi, \lambda) = \varphi$,
- $sp(\varphi, w_g(a, g)) = ((sp(\varphi, w_g) \land g)[x_a \leftarrow 0])^\uparrow$.

We often omit the first argument in the postcondition, implicitly assuming it to be the initial constraint $\varphi_0 = \bigwedge_{a,b \in \Sigma} x_a = x_b$ (or $true$ if $\Sigma$ has only one symbol), i.e., $sp(w_g) = sp(\varphi_0, w_g)$. Intuitively, $sp(w_g)$ is the constraint on clock valuations that is induced by $w_g$ on any following occurrence of a clock valuation, i.e., $\gamma \models sp(w_g)$ if and only if there is a clocked word $w_c(a, \gamma)$ such that $w_c \models w_g$. We remark that the polyhedron identified by the strongest postcondition is a convex set [DT98], and that $sp(w_g)$ is non-strict if $w_g$ is non-strict.

In Figure 4, an example of the strongest postcondition for the guarded word $w_g = (a, true)(b, 2 \leq x_a \leq 4)(b, 1 \leq x_b \leq 3)$ is shown. Intuitively, taking $a$ resets clock $x_a$. The two subsequent $b$-actions can only be taken between $2 + 1$ and $4 + 3$ time units later and do not reset $x_a$. As $b$ is the last action taken in the word, there is no constraint on $x_b$.

For guarded word $w_g$, we also introduce the $K^\prec$-approximated postcondition $sp_K^\prec(w_g)$, defined by

- $sp_K^\prec(\lambda) = \bigwedge_{a,b \in \Sigma} x_a = x_b$
- $sp_K^\prec(w_g(a, g)) = \langle \langle sp(sp_K^\prec(w_g), (a, g)) \rangle \rangle^K$.

Given a natural number $K$, we define the region equivalence $\sim_K$ on the set of clock valuations by $\gamma \sim_K \gamma'$ if

- for all $x_a \in C_\Sigma$, either
  - $\gamma(x_a)$ and $\gamma'(x_a)$ are both greater than $K$, or
  - $\lfloor \gamma(x_a) \rfloor = \lfloor \gamma'(x_a) \rfloor$ and $\text{fract}(\gamma(x_a)) = 0$ iff $\text{fract}(\gamma'(x_a)) = 0$, and
- for all $x_a, x_b \in C_\Sigma$ with $\gamma(x_a) \leq K$ and $\gamma(x_b) \leq K$, $\text{fract}(\gamma(x_a)) \leq \text{fract}(\gamma(x_b))$ iff $\text{fract}(\gamma'(x_a)) \leq \text{fract}(\gamma'(x_b))$. 

---

11
A region is an equivalence class of clock valuations induced by \( \sim_K \). We denote by \([\gamma]_K \) the region of \( \gamma \). In Figure 1(b), an example of a region is depicted.

A clock constraint identifies the union of a set of regions, therefore region equivalence induces a natural equivalence on clock constraints. For two clock constraints \( \varphi \) and \( \varphi' \), define \( \varphi \approx_K \varphi' \), if for each clock valuation \( \gamma \) with \( \gamma \models \varphi \) there is a clock valuation \( \gamma' \) with \( \gamma' \models \varphi' \) such that \( \gamma \sim_K \gamma' \), and vice versa.

An important property of region equivalence is that it is preserved by reset and time elapsing operations. If \( \gamma \sim_K \gamma' \) then \( \gamma[x_a \mapsto 0] \sim_K \gamma'[x_a \mapsto 0] \), and for each \( t \in \mathbb{R}^{\geq 0} \) there is a \( t' \in \mathbb{R}^{\geq 0} \) such that \( \gamma + t \sim_K \gamma' + t' \) [Yov96]. If we combine this fact with the fact that the reset and time elapsing operations on constraints mirror the same operations on clock valuations, we infer that the relation \( \approx_K \) on clock constraints is preserved by reset and time elapsing. Thus, if \( \varphi \approx_K \varphi' \) then \( \varphi[x_a \mapsto 0] \approx_K \varphi'[x_a \mapsto 0] \), and \( \varphi \uparrow \approx_K \varphi' \uparrow \). The relation \( \approx_K \) is also preserved by approximation, i.e., \( \varphi \approx_K \langle \langle \varphi \rangle \rangle \). A corollary of these facts is the following lemma.

**Lemma 3** Let \( w_g \) be a guarded word. Then

\[
sp_K^\uparrow (w_g) \approx_K sp(w_g)
\]

**PROOF.** By induction on the length of \( w_g \). For the base case, where \( w_g \) is empty, the proof is immediate. For the inductive step, assume that \( sp_K^\uparrow (w_g) \approx_K sp(w_g) \) and consider the guarded word \( w_g(a, g) \). By definition, \( sp_K^\uparrow (w_g(a, g)) = \langle \langle sp_K^\uparrow (w_g), (a, g) \rangle \rangle \). From \( sp_K^\uparrow (w_g) \approx_K sp(w_g) \) we infer, using that \( \approx_K \) is preserved by reset, time elapsing, and conjunction with clock guards, that \( sp(sp_K^\uparrow (w_g), (a, g)) \approx_K sp(w_g(a, g)) \). Since \( \approx_K \) is preserved by approximation, we infer that \( \langle \langle sp(sp_K^\uparrow (w_g), (a, g)) \rangle \rangle \approx_K sp(sp_K^\uparrow (w_g), (a, g)) \approx_K sp(w_g(a, g)) \), i.e., that \( sp_K^\uparrow (w_g(a, g)) \approx_K sp(w_g(a, g)) \). □

For every satisfiable non-strict clock guard \( g \), we define its \( K \)-smallest corner, denoted by \( sc_K(g) \), as the set of clock valuations \( \gamma \) that satisfy \( \gamma(x_a) = n \) for all \( x_a \) such that the lower bound for \( x_a \) in \( g \) is of form \( n \leq x_a \) with \( n < K \), and that satisfy \( \gamma(x_a) \geq K \) whenever the lower bound for \( x_a \) in \( g \) is of form \( n \leq x_a \) with \( n \geq K \). (7) Similarly, we define the biggest corner of \( g \), denoted \( bc_K(g) \) as the set of valuations \( \gamma \) that are maximal in the dimensions where \( [g] \) has an upper bound and exceeds \( K \) in the others. In Figure 3(a), for an example, if \( K = 8 \), the biggest corner of guard \( g \) contains the only valuation \( \gamma \) with \( \gamma(x_a) = 8 \) and \( \gamma(x_b) = 3 \), while for \( g' \) if \( K = 2 \), the biggest corner contains all valuations \( \gamma \) with \( \gamma(x_a) = 2 \) and \( \gamma(x_b) > 2 \). (8)
3 Event-Recording Automata

In this section, we introduce event-recording automata, which are the subject of the learning algorithms in the paper. We also introduce the further restricted class of event-deterministic event-recording automata, which the algorithms $TL^*_{sg}$ and $TL^*_{nsg}$ are designed to learn. In the treatment, we will repeatedly make use of standard deterministic finite automata.

A deterministic finite automaton (DFA) $A = (\Gamma, Q, q_0, \delta, Q_f)$ over the alphabet $\Gamma$ consists of a finite set of states $Q$, and initial state $q_0$, a partial transition function $\delta : Q \times \Gamma \rightarrow Q$, and a set of final states $Q_f \subseteq Q$. A run of $A$ over the word $w = a_1a_2\ldots a_n$ is a finite sequence

$q_0 \xrightarrow{a_1} q_1 \xrightarrow{a_2} \cdots \xrightarrow{a_n} q_n$

of states $q_i \in Q$ such that $q_0$ is the initial state and $\delta(q_{i-1}, a_i)$ is defined for $1 \leq i \leq n$, with $\delta(q_{i-1}, a_i) = q_i$. In this case, we write $\delta(q_0, w) = q_n$, thereby extending the definition of $\delta$ to words in the natural way. The run is called accepting if $q_n \in Q_f$. The language $L(A)$ comprises all words $a_1a_2\ldots a_n$ over which an accepting run exists.

We are now ready to introduce the class of automata models whose objects we want to learn: deterministic event-recording automata.

Definition 4 An event-recording automaton (ERA) $D = (\Sigma, L, l_0, \delta, L_f)$ consists of a finite input alphabet $\Sigma$, a finite set $L$ of locations, an initial location $l_0 \in L$, a set $L_f$ of accepting locations, and a transition function $\delta : L \times \Sigma \times G_{\Sigma} \rightarrow 2^L$, which is a partial function with finite support that for each location, input symbol and guard potentially prescribes a set of target locations. An ERA is deterministic iff

- $\delta(l, a, g)$ is a singleton set whenever it is defined, and
- whenever $\delta(l, a, g_1)$ and $\delta(l, a, g_2)$ are both defined then $[g_1] \cap [g_2] = \emptyset$. \hfill $\square$

Thus, for a deterministic ERA, a location $l$ might have two different $a$ successors, which, however, have nonoverlapping guards. Due to the first restriction, we will consider $\delta$ to be of type $\delta : L \times \Sigma \times G_{\Sigma} \rightarrow L$, i.e., to map each triple in its domain to a single location rather than a set. An ERA is $K$-bounded if the guard $g$ is $K$-bounded whenever $\delta(l, a, g)$ is defined.

In this paper, we only consider deterministic ERAs, or DERAs for short, which is no significant restriction in terms of expressiveness as every ERA can be transformed into a DERA accepting the same language. For details, see [AFH99].

In order to define the language accepted by a DERA, we first understand it...
as a DFA, which accepts guarded words.

Given a DERA $D = (\Sigma, L, l_0, \delta, L_f)$, we define $\text{dfa}(D)$ to be the DFA $A_D = (\Gamma, L, l_0, \delta', L_f)$ over the alphabet $\Gamma = \Sigma \times G_\Sigma$ where $\delta' : L \times \Gamma \rightarrow L$ is defined by $\delta'(l, (a, g)) = \delta(l, a, g)$ if and only if $\delta(l, a, g)$ is defined, otherwise $\delta'(l, (a, g))$ is undefined. Note that $D$ and $\text{dfa}(D)$ have the same number of locations/states. Further, note that this mapping from DERAs over $\Sigma$ to DFAs over $\Sigma \times G_\Sigma$ is injective, meaning that for each DFA $A$ over $\Sigma \times G_\Sigma$, there is a unique (up to isomorphism) ERA over $\Sigma$, denoted $\text{era}(A)$, such that $\text{dfa}(\text{era}(A))$ is isomorphic to $A$.

The language $L(D)$ accepted by a DERA $D$ is defined to be the set of timed words $w_t$ such that $w_t \models w_g$ for some guarded word $w_g \in L(\text{dfa}(D))$. We call two DERAs $D_1$ and $D_2$ equivalent iff $L(D_1) = L(D_2)$, and denote this by $D_1 \equiv D_2$, or just $D_1 \equiv D_2$.

We introduce a restricted class of deterministic ERAs, which the algorithms $\text{TL}^*_s$ and $\text{TL}^*_n$ are designed to learn. The restriction is that each state has at most one outgoing transition per action. This means that timing constraints for the occurrence of an action depend only on the past sequence of actions, and not on their relative timing.

**Definition 5** An ERA $(\Sigma, L, l_0, \delta, L_f)$ is called event-deterministic (EDERA), if

- only non-strict guards are used,
- for every $l \in L$ and $a \in \Sigma$ there is at most one $g \in G_\Sigma$ such that $\delta(l, a, g)$ is defined, and
- every location is accepting. □

In case of an EDERA, its transition function $\delta : L \times \Sigma \times G_\Sigma \rightarrow L$ can be understood as two functions; $\eta : L \times \Sigma \rightarrow G_\Sigma$, which for a location and an input symbol prescribes a guard, and $\varrho : L \times \Sigma \rightarrow L$, which for a location and an input symbol prescribes a target location. Thus, we use also $D = (\Sigma, L, l_0, \varrho, \eta)$ for denoting an EDERA, where $L_f$ is omitted since $L_f = L$.

From the above definitions, we see that the language of an EDERA $D$ can be characterized by a prefix-closed set of guarded words $(a_1, g_1)(a_2, g_2)\ldots(a_n, g_n)$ in $L(\text{dfa}(D))$ such that each $a_1a_2\ldots a_n$ is a word underlying at most one $w_g \in L(\text{dfa}(D))$. (9) Thus, we can loosely say that $D$ imposes on each untimed word $a_1a_2\ldots a_n$ the timing constraints represented by the guards $g_1g_2\ldots g_n$.

**Example 6** The event-recording automaton shown in Figure 5 over the alphabet $\{a, b, c\}$ uses three event-recording clocks, $x_a$, $x_b$, and $x_c$. It is event deterministic, as all guards are non-strict and no location has two outgoing edges labelled with the same action. Location 0 is the initial location of the
automaton. The clock constraint $x_b \geq 3$ that is associated with the edge from location 1 to 4 ensures that the action $c$ can only be taken at least three time units after taking the transition from 0 to 1. This also implies that the time difference between the first $b$ and the subsequent $a$ is greater or equal to 3.

Fig. 5. An event-recording automaton

4 The $L^\ast$ algorithm for learning DFAs

In this section, we shortly review the $L^\ast$ algorithm, due to Angluin [Ang87] for learning a regular (untimed) language, $L(A) \subseteq \Gamma^\ast$, accepted by a minimal deterministic finite automaton (DFA) $A = (\Gamma, Q, q_0, \delta, Q_f)$. In this algorithm a so-called Learner, who initially knows nothing about $A$, is trying to learn $L(A)$ by asking queries to a Teacher, who knows $A$. There are two kinds of queries:

- A membership query consists in asking whether a string $w \in \Gamma^\ast$ is in $L(A)$.
- An equivalence query consists in asking whether a hypothesized DFA $H$ is correct, i.e., whether $L(H) = L(A)$. The Teacher will answer yes if $H$ is correct, or else supply a counterexample $w$, which is a word either in $L(A) \setminus L(H)$ or in $L(H) \setminus L(A)$.

The Learner maintains a prefix-closed set $U \subseteq \Gamma^\ast$ of prefixes, which are candidates for identifying states, and a suffix-closed set $V \subseteq \Gamma^\ast$ of suffixes, which are used to distinguish such states. The sets $U$ and $V$ are increased when needed during the algorithm. The Learner makes membership queries for all words in $(U \cup UT)V$, and organizes the results into a table $T$ which maps each $u \in (U \cup UT)$ to a mapping $T(u) : V \mapsto \{+, -\}$. The interpretation of $T$ is that for $u \in (U \cup UT)$ and $v \in V$ we have $T(u)(v) = +$ if $uv \in L(A)$ and $T(u)(v) = -$ if $uw \not\in L(A)$. In [Ang87], each function $T(u)$ is called a row. Thus two rows, $T(u)$ and $T(u')$, are equal, denoted $T(u) = T(u')$, if $T(u)(v) = T(u')(v)$ for all $v \in V$. Table $T$ is

- closed, if for each $u \in U$ and $a \in \Gamma$ there is a $u' \in U$ such that $T(ua) = T(u')$,
- consistent, if, for each $u, u' \in U$, $T(u) = T(u')$ implies $T(ua) = T(u'a)$.
If $T$ is not closed, we find $u' \in UT$ such that $T(u) \neq T(u')$ for all $u \in U$. Then we move $u'$ to $U$ and ask membership queries for every $u'av$ where $a \in \Gamma$ and $v \in V$. If $T$ is not consistent, we find $u, u' \in U$, $a \in \Gamma$ and $v \in V$ such that $T(u) = T(u')$ and $T(ua)(v) \neq T(u'a)(v)$. Then we add $av$ to $V$ and ask membership queries for every $u'av$ where $a \in \Gamma$ and $v \in V$. Checks whether $T$ is closed and consistent can be done in any ordering. When $T$ is closed and consistent the Learner constructs a hypothesized DFA $\mathcal{H} = (\Gamma, L, l_0, \delta, L_f)$, where

- $L = \{T(u) \mid u \in U\}$ is the set of distinct rows,
- $l_0$ is the row $T(\lambda)$,
- $\delta$ is defined by $\delta(T(u), a) = T(ua)$, and
- $L_f = \{T(u) \mid u \in U \text{ and } T(u)(\lambda) = \text{accepted}\}$ is the set of rows which are accepting without adding a suffix,

and submits $\mathcal{H}$ in an equivalence query. If the answer is yes, the learning procedure is completed. Otherwise the returned counterexample $w$ is processed by adding every prefix of $w$ (including $w$) to $U$, and subsequent membership queries are performed in order to make the table closed and consistent, after which a new hypothesized DFA is constructed, etc.

The $L^*$ algorithm constructs $\mathcal{A}$ after asking $O(kn^2m)$ membership queries and at most $n$ equivalence queries, where $n$ is the number of states in $\mathcal{A}$, $k$ is the size of the alphabet and $m$ is the length of the longest counterexample [Ang87]. The rough idea is that for each entry in the table $T$ a query is needed, and $O(knm)$ is the number of rows, $n$ is the number of columns.

A description of the $L^*$ algorithm is given as Algorithm 1 and Algorithm 2, using Java-style pseudocode. Since membership queries and equivalence queries can be implemented in different ways and also differ in timed and untimed settings, we introduce the interface $\text{Teacher}$ which contains two functions that are responsible for membership and equivalence queries (see Algorithm 1). Angluin’s algorithm is given as function $\text{Learner}$ of class $L^*$ (see lines 10–21 in Algorithm 2). The function $\text{Learner}$ first constructs an initial table by calling the function $\text{initialize}$ and then constructs hypothesized automata until the answer to an equivalence query is yes. Since each hypothesized automaton has to be constructed from a closed and consistent table, function $\text{Learner}$ checks these properties by calling functions $\text{isClosed}$ and $\text{isConsistent}$. If the table is not consistent, the function $\text{add_column}$ is called, which adds a distinguishing suffix to $V$. If the table is not closed, the function $\text{move_row}$ is called which moves the corresponding row $ua$ to $U$. When a hypothesized automaton is constructed, an equivalence query is performed and if a counterexample is obtained the function $\text{process_counterexample}$ is called.
Algorithm 1 Interface of Teacher

```java
1 interface Teacher {
2     Function membership_query(u)
3     Function equivalence_query(ℋ)
4 }
```

![Automaton](image)

(a) An automaton $\mathcal{A}$  
(b) An automaton $\mathcal{H}$

Fig. 6. An automaton to be learned and a hypothesized automaton

<table>
<thead>
<tr>
<th></th>
<th>$T_1$</th>
<th>$T_2$</th>
<th>$T_3$</th>
<th>$T_4$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\lambda$</td>
<td>$\lambda +$</td>
<td>$\lambda +$</td>
<td>$\lambda +$</td>
<td>$\lambda -$</td>
</tr>
<tr>
<td>$a$</td>
<td>$a -$</td>
<td>$ab -$</td>
<td>$ab -$</td>
<td>$a -$</td>
</tr>
<tr>
<td>$b$</td>
<td>$b +$</td>
<td>$aba +$</td>
<td>$aba +$</td>
<td>$aba +$</td>
</tr>
<tr>
<td>$aa$</td>
<td>$aa -$</td>
<td>$aa -$</td>
<td>$aa -$</td>
<td>$aa -$</td>
</tr>
<tr>
<td>$ab$</td>
<td>$abb -$</td>
<td>$abb -$</td>
<td>$abb -$</td>
<td>$abb -$</td>
</tr>
<tr>
<td>$aba$</td>
<td>$aba -$</td>
<td>$aba -$</td>
<td>$aba -$</td>
<td>$aba -$</td>
</tr>
<tr>
<td>$abab$</td>
<td>$abab -$</td>
<td>$abab -$</td>
<td>$abab -$</td>
<td>$abab -$</td>
</tr>
</tbody>
</table>

Fig. 7. Observation tables

Let us consider an example of the $L^*$ algorithm. Let $\mathcal{A}$ be the DFA shown in Figure 6(a). Initially, the Learner asks membership queries for $\lambda$, $a$ and $b$. The initial observation table $T_1$ is shown in Figure 7(a), where $U = V = \{\lambda\}$. This observation table is consistent, but not closed, since $T(a) \neq T(\lambda)$. The Learner moves the prefix $a$ to $U$ and then asks membership queries for $aa$ and $ab$ to construct the observation table $T_2$ shown in Figure 7(b). This observation table is closed and consistent. The Learner constructs the hypothesized automaton $\mathcal{H}$ shown in Figure 6(b) and asks an equivalence query to the Teacher. Assume that the Teacher replies with the counterexample $aba$, which is in $\mathcal{L}(\mathcal{A})$ but not accepted by $\mathcal{H}$. To process the counterexample $aba$, the Learner adds $ab$ and $aba$ to $U$ and asks membership queries for $abb$, $aba$ and $abab$ to construct the observation table $T_3$ shown in Figure 7(c). This observation table is not
Algorithm 2 Pseudo code for Angluin’s Learning Algorithm

class L*
Teacher teacher
Alphabet Γ

Constructor L*(t,Σ)
    teacher = t
    Γ = Σ

Function Learner()
    initialize (U, V, T)
    repeat
        while not isClosed((U, V, T)) or not isConsistent((U, V, T))
            if not isConsistent((U, V, T)) then add_column()
            if not isClosed((U, V, T)) then move_row()
        Construct hypothesized automaton H
        teacher.equivalence.query(H)
        if the answer to equivalence query is a counterexample u then
            process_counterexample(u)
        until the answer to equivalence query is ‘yes’ to the hypothesis H
    return H.

Function initialize(U, V, T)
    U := {λ}, V := {λ}
    T(λ)(λ) = teacher.membership_query(λ)
    for every a ∈ Γ
        T(a)(λ) = teacher.membership_query(a)

Function isClosed()
    if for each u ∈ U and a ∈ Γ there is u' ∈ U with T(ua) = T(u'a)
    return true
    else
        return false

Function isConsistent()
    if for each a ∈ Γ and u, u' ∈ U such that T(u) = T(u') we have T(ua) = T(u'a)
    return true
    else
        return false

Function add_column()
    Find a ∈ Γ, v ∈ V and u, u' ∈ U such that T(u) = T(u') and T(ua)(v) ≠ T(u'a)(v)
    Add av to V
    for every u ∈ U ∪ UT
        T(u)(av) = teacher.membership_query(uav)

Function move_row()
    Find u ∈ U, a ∈ Γ such that T(ua) ≠ T(u'a) for all u' ∈ U
    Move ua to U
    for every a' ∈ Γ and v ∈ V
        T(u'a')(v) = teacher.membership_query(u'a'v)

Function process_counterexample(u)
    Add every prefix u' of u to U
    for every a ∈ Γ, v ∈ V and prefix u' of u
        T(u')(v) = teacher.membership_query(u'v)
        T(u'a)(v) = teacher.membership_query(u'av)
consistent since $T(a) = T(ab)$ but $T(aa) \neq T(aba)$. Then the Learner adds $a$ to $E$ and asks membership queries for $ba$, $aaa$, $abba$, $abaaa$ and $ababa$ to construct the observation table $T_4$ shown in Figure 7(d). This observation table is closed and consistent. The Learner constructs the automaton shown in Figure 6(a) and asks an equivalence query to the Teacher. The Teacher replies yes and $L^*$ terminates.

As we will see later, the general scheme of this algorithm stays the same in our algorithms for learning timed languages. However, the initialization of the table, queries, closedness and consistency checks become different.

5 Learning Event-Deterministic ERAs

In this section, we present the algorithm $TL^*_sg$ for learning EDERAs, obtained by adapting the $L^*$ algorithm. A central idea in the $L^*$ algorithm is to let each state be identified by the words that reach it from the initial state (such words are called access strings in [BDG97]). States are congruent if, according to the queries submitted so far, the same continuations of their access strings are accepted. This idea is naturally based on the properties of Nerode’s right congruence (given a language $L$, two words $u, v \in \Sigma^*$ are equivalent if for all $w \in \Sigma^*$ we have $uw \in L$ iff $vw \in L$) which implies that there is a unique minimal DFA accepting $L$. In other words, for DFAs, every state can be characterized by the set of words accepted by the DFA when considering this state as an initial state, and every string leads to a state in a unique way.

For timed languages, it is not obvious how to generalize Nerode’s right congruence. In general there is no unique minimal DERA which is equivalent to a given DERA. As an example, consider Figure 5, assuming for a moment that the $c$-transition from location 7 to 5 is missing. Then the language of the automaton does not change when changing the transition from 1 into 4 to 1 into 5, although the language accepted from 4 is different from the one from 5. Furthermore, if we do not modify the automaton in Figure 5 we can reach location 4 by two guarded words: $(b, \text{true})(c, x_b \geq 3)$ as well as $(a, \text{true})(b, \text{true})$. Although they lead to the same location, they admit different continuations of event-clock words: action $a$ can be performed with $x_b = 2$ after $(a, \text{true})(b, \text{true})$ but not after $(b, \text{true})(c, x_b \geq 3)$. The complication is that each guarded word imposes a postcondition, which constrains the values of clocks that are possible at the occurrence of future actions.

Our approach to overcoming the problem that DERAs have no canonical form is to define a subclass of EDERAs which do have a canonical form, and which

1 See [MP04] for a study on right-congruences on timed languages.
furthermore can be understood as a DFA over $\Sigma \times G$ where $G$ is the set of clock guards. We can then use Angluin’s algorithm to learn this DFA, and thereafter interpret the result as an EDERA. In the next section, we define this canonical form, called sharply guarded EDERA, and prove that any EDERA can be transformed to this canonical form. We can therefore use Angluin’s algorithm to learn a DFA over $\Sigma \times G$. A problem is that membership queries will ask whether a guarded word is accepted by the DFA, whereas the EDERA to be learned answers only queries for timed words. We therefore extend the Learner in Angluin’s algorithm by an Assistant, whose role is to answer a membership query for a guarded word, posed by the Learner, by asking several membership queries for timed words to the (timed) Teacher. We describe the operation of the Assistant in Section 5.2. Thereafter, in Section 5.3 we present the complete algorithm for learning EDERAs.

5.1 Sharply Guarded EDERAs

Motivated by the previous discussion, in this section we define a class of EDERAs that admit a natural definition of right congruences.

**Definition 7** A $K$-bounded EDERA $D$ is sharply guarded if for all guarded words $w_g(a, g) \in L(dfa(D))$, we have that $sp^<_K(u_g) \land g$ is satisfiable and

$$g = \langle \bigwedge \{g' \in G_\Sigma \mid [sp^<_K(u_g) \land g'] = [sp^<_K(u_g) \land g] \} \rangle^<_K$$

\[ \square \]

Note that the conjunction is taken over all clock guards $g'$, i.e., also those that are not $K$-bounded. Figure 8 illustrates of Definition 7 by a postcondition $sp(u_g) = x_a - x_b \geq 2 \land x_a - x_b \leq 8$ together with a “sharp” guard $g_1 = x_a \geq 4 \land x_a \leq 12 \land x_b \geq 2 \land x_b \leq 4$ and a “non-sharp” guard $g_2 = x_a \geq 3 \land x_a \leq 15 \land x_b \geq 2 \land x_b \leq 4$. EDERA $D$ such that $u_g(a, g_2) \in L(dfa(D))$ and $K = 15$ is not sharply guarded, since $[sp^<_K(u_g) \land g_2] = [sp^<_K(u_g) \land g_1]$, where $[g_1] \subset [g_2] \ (11)$. As shown in Figure 8, there is no guard $g'$ such that $[g'] \subset [g_1]$ and $[sp(u_g) \land g'] = [sp(u_g) \land g_2]$. Intuitively, an EDERA $D$ is sharply guarded if the outgoing transitions from a location have guards which cannot be strengthened without changing the timing conditions under which the next symbol $a$ will be accepted. Thus the upper and lower bounds on clock values in a clock valuation $\gamma$ constraining the occurrence of $a$ do not depend implicitly on the postcondition of the previous sequence of transitions, taken by $D$. Thus we avoid the complications induced by postconditions described before Definition 7.

The use of approximation in Definition 7 does not affect the definition, as we will see in the following Lemma 8, but gives an a priori bound on the size of
sharply guarded automata that accept a given timed language.

The following lemma gives a simpler characterization of being sharply guarded.

**Lemma 8** Let $g$ be a non-strict $K$-bounded clock guard, and let $\varphi$ be a $K$-bounded clock constraint. Then the following clock guards are equal.

\begin{enumerate}  
\item \(\langle\langle \bigwedge \{ g' \in G_{\Sigma} \mid [\varphi \land g] = [\varphi \land g'] \}\rangle\rangle \leq K\),
\item \(\langle\langle \bigwedge \{ g' \in G_{\Sigma} \mid [\varphi \land g] \subseteq [g'] \}\rangle\rangle \leq K\),
\item \(\langle\langle \text{guard}(\varphi \land g) \rangle\rangle \leq K\).
\end{enumerate}

In the following, we will mostly use characterization c) when reasoning about sharply guarded EDERAs.

**PROOF.** We first prove that a) and b) are equal. We observe that \([\varphi \land g] \subseteq [g']\) implies \([\varphi \land g] \subseteq [\varphi \land g']\), and that \([\varphi \land g] = [\varphi \land g']\) implies \([\varphi \land g] \subseteq [g']\). It therefore suffices to prove that we get the same result when using \([\varphi \land g] \subseteq [\varphi \land g']\) as when using \([\varphi \land g] = [\varphi \land g']\) as the condition on $g'$ in the large conjunction. This follows by observing that for each guard $g'$ such that \([\varphi \land g] \subseteq [\varphi \land g']\), there is a guard $g''$ such that \([\varphi \land g] = [\varphi \land g'']\), namely $g'' = g \land g'$, and that the conjunction of these $g''$ is the same as the conjunction of all $g'$.

We then prove that b) and c) are equal. Since \([\varphi \land g] \subseteq [\text{Can}(\varphi \land g)] \subseteq [\text{guard}(\text{Can}(\varphi \land g))]\) we infer \([\varphi \land g] \subseteq [\text{guard}(\text{Can}(\varphi \land g))]\), from which it follows that b) is included in c). Conversely, for any guard $g'$ we have that \([\varphi \land g] \subseteq [g']\) implies that \([\text{guard}(\text{Can}(\varphi \land g))] \subseteq [g']\), from which the opposite inclusion follows. \(\square\)

Let us introduce the notation $\text{tightguard}_K(\varphi, g)$ for \(\langle\langle \text{guard}(\varphi \land g) \rangle\rangle \leq K\). Let us establish some basic facts about $\text{tightguard}_K(\varphi, g)$.

**Proposition 9** Let $g$ be a non-strict $K$-bounded clock guard, and let $\varphi$ be a $K$-bounded clock constraint. Then
\[(1) \; [\varphi \land g] = [\varphi \land \text{guard}(\varphi \land g)],\]
\[(2) \; \text{tightguard}_K(\varphi, g) = \text{tightguard}_K(\varphi, \text{tightguard}_K(\varphi, g)).\]

PROOF. Let \(\varphi\) and \(g\) be as in the statement of the proposition.

(1) follows from \([\text{guard}(\varphi \land g)] \subseteq [g]\) and \([\varphi \land g] \subseteq [\text{guard}(\varphi \land g)]\).

(2) By the definition of \(\text{tightguard}_K(\varphi, g)\), and using form \(b)\) and form \(c)\) of Lemma 8, we must prove

\[
\langle\langle \bigwedge \{g' \in G_S \mid [\varphi \land g] \subseteq [g']\} \rangle \rangle_K^< \\
= \langle\langle \bigwedge \{g' \in G_S \mid [\varphi \land \langle\langle \text{guard}(\varphi \land g) \rangle \rangle_K^< \subseteq [g']\} \rangle \rangle_K^< 

This follows by noting that the expression on the left-hand side of \(\subseteq\) is the same in both expressions, by property (1), since \(g\) is non-strict and \(K\)-bounded. \(\Box\)

The following lemma shows that sharply guarded EDERA can also be defined in terms of \(sp(w_g)\), or any other clock constraint \(\varphi\) such that \(\varphi \approx_K sp_K^< (w_g)\). The reason to define sharply guarded EDERA in terms of \(sp_K^< (w_g)\) is that in order to bound the size of sharply guarded EDERA, we need to use \(K^<\)-approximations of postconditions in the construction of a sharply guarded EDERA in Lemma 12.

Lemma 10 Let \(g\) be a non-strict \(K\)-bounded clock guard. Let \(\varphi\) and \(\varphi'\) be clock constraints such that \(\varphi \approx_K \varphi'\). Then

\[\text{tightguard}_K(\varphi, g) = \text{tightguard}_K(\varphi', g)\]

PROOF. A \(K\)-bounded clock guard is a union of a set of regions. Therefore, if \(g\) is a \(K\)-bounded clock guard, \(\varphi \approx_K \varphi'\) implies \(\varphi \land g \approx_K \varphi' \land g\), which implies \(\text{Can}(\varphi \land g) \approx_K \text{Can}(\varphi' \land g)\), which implies

\[
\langle\langle \text{guard}(\varphi \land g) \rangle \rangle_K^< = \langle\langle \text{guard}(\varphi' \land g) \rangle \rangle_K^<,

from which the lemma follows by the definition of \(\text{tightguard}_K(\varphi, g)\). \(\Box\)

By Definition 7, an EDERA is sharply guarded if the outgoing transitions from a location have guards which cannot be strengthened without changing the timing conditions under which the next symbol will be accepted. The following lemma shows that these guards can be characterized in terms of biggest and smallest corners.

22
Lemma 11 If \( w_g(a, g) \in \mathcal{L}(df\alpha(D)) \), where \( D \) is a \( K \)-bounded sharply guarded EDERA, then

(a) there is a timed word \( w_t \in \mathcal{L}(D) \) such that \( CW(w_t) = w_c(a, \gamma), w_c(a, \gamma) \models w_g(a, g) \) and \( \gamma \in bc_K(g) \).

(b) there is a timed word \( w_t \in \mathcal{L}(D) \) such that \( CW(w_t) = w_c(a, \gamma), w_c(a, \gamma) \models w_g(a, g) \) and \( \gamma \in sc_K(g) \).

**Proof.** We first prove (a). Since \( D \) is a sharply guarded EDERA then \( sp_K^c(w_g) \wedge g \) is satisfiable, hence \( sp(w_g) \wedge g \) is satisfiable (since \( sp(w_g) \approx_K sp_K^c(w_g) \) by Lemma 3 and \( g \) is \( K \)-bounded). The basic property of postconditions, that \( \gamma \models sp(w_g) \) if and only if there is a clocked word \( w_c(a, \gamma) \) such that \( w_c \models w_g \), implies that we must prove that there is a clock valuation \( \gamma \) such that \( \gamma \models sp(w_g) \) and \( \gamma \in bc_K(g) \). Let \( a \) be any action in \( \Sigma \). If \( x_a \leq n \) is a conjunct in \( g \) for \( n < K \), then by the definition of sharply guarded, using form (c) in Lemma 8, and the fact that \( sp(w_g) \) is non-strict, it follows that there is a clock valuation \( \gamma_a \) such that \( \gamma_a \models sp(w_g) \) and \( \gamma_a(x_a) = n \). Similarly, if there is no conjunct of form \( x_a \leq n \) in \( g \) for \( n \leq K \), then there is a clock valuation \( \gamma_a \) such that \( \gamma_a \models sp(w_g) \) and \( \gamma_a(x_a) > K \). Since this holds for any \( a \), by Proposition 1 it follows that there is a clock valuation \( \gamma \) such that \( \gamma \models sp(w_g) \) and \( \gamma \in bc_K(g) \).

The proof of (b) is analogous: we can infer that whenever \( x_a \geq n \) is a conjunct in \( g \) for \( n < K \), then there is a clock valuation \( \gamma_a \) such that \( \gamma_a \models sp(w_g) \) and \( \gamma_a(x_a) = n \). A slight difference occurs for the case where \( x_a \geq K \) is a conjunct in \( g \): here we use satisfiability of \( sp(w_g) \wedge g \) to infer that there is a clock valuation \( \gamma_a \) such that \( \gamma_a \models sp(w_g) \) and \( \gamma_a(x_a) \geq K \). Since this holds for any \( a \), by Proposition 1 it follows that there is a clock valuation \( \gamma \) such that \( \gamma \models sp(w_g) \) and \( \gamma \in sc_K(g) \). \( \square \)

Every EDERA can be transformed into an equivalent EDERA that is sharply guarded using the zone-graph construction [DT98].

Lemma 12 For every EDERA there is an equivalent EDERA that is sharply guarded.

**Proof.** Let the EDERA \( D = (\Sigma, L, l_0, \varphi, \eta) \) be \( K \)-bounded. We define an equivalent sharply guarded EDERA \( D' = (\Sigma, L', l_0', \varphi', \eta') \) based on the so-called zone automaton for \( D \). The set of locations of \( D' \) comprises pairs \((l, \varphi)\) where \( l \in L \) and \( \varphi \) is a \( K \)-bounded clock constraint. The intention is that \( \varphi \) is the \( K^\approx \)-approximated postcondition of any run from the initial location to \((l, \varphi)\). The initial location \( l_0' \) of \( D' \) is \((l_0, sp(\lambda))\). For any location \( l \in L \) and symbol \( a \) such that \( g(l, a) \) is defined and \( \varphi \wedge \eta(l, a) \) is satisfiable,
We set \( \varrho \) and \( \eta \) as in Proposition 9 and the fact that \( g \) is a \( K \)-bounded clock guard.

The property proven in the previous paragraph, together with the observation in property (2) in Proposition 9, that \( \eta'(l, \varphi, a) = \text{tightguard}_K(\varphi, \eta(l, a)) \), implies by characterization c) in Lemma 8 that \( D' \) is sharply guarded.

We then prove that \( D' \) is equivalent to \( D \) by induction. For every timed word \( w \), we need to show that \( w \in L(D) \) iff \( w \in L(D') \). Due to the one-to-one correspondence between timed words and clocked words we present the proof in terms of clocked words. For \( \lambda \), we start in \( l_0 \) in \( D \) and in \( (l_0, sp(\lambda)) \) in \( D' \). So \( \lambda \) is accepted by both automata.

From the second paragraph of this proof, it follows that whenever \( l = \delta(l_0, w_g) \) is defined, then there is a guarded word \( w'_g \) with the same underlying word as \( w_g \) such that \( \delta'(l'_0, w'_g) = (l, \varphi) \) and such that \( \varphi = sp^\omega_K(w_g) = sp^\omega_K(w'_g) \). Also the converse follows, i.e., that whenever \( (l, \varphi) = \delta'(l'_0, w'_g) \) is defined, then there is a guarded word \( w_g \) with the same underlying word as \( w'_g \), such that \( \delta(l_0, w_g) = l \) and such that \( \varphi = sp^\omega_K(w_g) = sp^\omega_K(w'_g) \). In both cases, it also follows for any alphabet symbol \( a \) that

\[
[sp^\omega_K(w_g) \land \eta(l, a)] = [sp^\omega_K(w_g) \land \eta'(l, \varphi, a)] .
\]

To prove that \( L(D) \subseteq L(D') \) we shall establish by induction over \( w_g \) that whenever \( w_c \models w_g \) where \( l = \delta(l_0, w_g) \) is defined, then there is a guarded word \( w'_g \) such that \( w_c \models w'_g \) and \( \delta'(l'_0, w'_g) = (l, \varphi) \) where \( \varphi = sp^\omega_K(w_g) = sp^\omega_K(w'_g) \). The base case \( w_g = \lambda \) is trivial. For the inductive step, let \( w_c(a, \gamma) \models w_g(a, g) \) where \( g = \eta(l, a) \). From the preceding paragraph, it follows that \( w_c(a, \gamma) \models w'_g(a, g') \) where \( g' = \eta'(l, \varphi, a) \). This concludes the inductive step.
We call two sharply guarded EDERAs equivalent if and only if their corresponding DFAs are equivalent.

The important property of sharply guarded EDERAs is that equivalence coincides with equivalence on the corresponding DFAs.

In Figure 9 is shown sharply-guarded EDERA $H'$, which is equivalent to EDERA $H$ shown in Figure 11(b). The initial location $(0, sp(\lambda))$ of $H'$ corresponds to the location $0$ in Figure 9. From the location $0$ in $H$ there are two transitions $a$ and $b$. Therefore we construct the transition $a$ from the location $(0, sp(\lambda))$ to the location $(0, \langle\langle sp(\lambda), (a, (a, x_a = 2 \land x_b = 2))\rangle\rangle^{K}_{\lambda}) = (0, x_a + 2 \leq x_b)$ (the location 1 in Figure 9) and the transition $b$ from the location $(0, sp(\lambda))$ to the location $(0, \langle\langle sp(\lambda), (b, (b, x_a \geq 1 \land x_b \geq 1))\rangle\rangle^{K}_{\lambda}) = (0, x_b + 1 \leq x_a)$ (the location 2 in Figure 9). Since $x_a + 2 \leq x_b \land x_a = 2 \land x_b = 2$ is not satisfiable, from the location $(0, x_a + 2 \leq x_b)$ we can construct only the transition $b$ which leads to the location $(0, \langle\langle (x_a + 2 \leq x_b, (b, (b, x_a \geq 1 \land x_b \geq 1))\rangle\rangle^{K}_{\lambda}) = (0, x_b + 1 \leq x_a)$ (location 2 in Figure 9). The guard for this transition is $tightguard^{K}_{\lambda}(x_a + 2 \leq x_b, x_a \geq 1 \land x_b \geq 1) = x_a \geq 1 \land x_b \geq 2$. From location $(0, x_b + 1 \leq x_a)$ we can construct only the transition $b$. The transition leads to the location $(0, \langle\langle (x_b + 1 \leq x_a, (b, (b, x_a \geq 1 \land x_b \geq 1))\rangle\rangle^{K}_{\lambda}) = (0, x_b + 2 \leq x_a)$ (the location 3 in Figure 9) and has the guard $tightguard^{K}_{\lambda}(x_b + 1 \leq x_a, x_a \geq 1 \land x_b \geq 1) = x_a \geq 2 \land x_b \geq 1$. The construction of $H'$ terminates after adding loop to the location $(0, x_b + 2 \leq x_a)$.

The important property of sharply guarded EDERAs is that equivalence coincides with equivalence on the corresponding DFAs.

**Definition 13** We call two sharply guarded EDERAs $D_1$ and $D_2$ dfa-equivalent, denoted by $D_1 \equiv_{dfa} D_2$, iff $dfa(D_1)$ and $dfa(D_2)$ accept the same language (in the sense of DFAs).

**Lemma 14** For two sharply guarded EDERAs $D_1$ and $D_2$, we have

$$D_1 \equiv_t D_2 \iff D_1 \equiv_{dfa} D_2$$

**Proof.** The direction from right to left follows immediately, since $L(D_t)$ is defined in terms of $L(dfa(D_1))$. To prove the other direction, assume that $D_1 \not\equiv_{dfa} D_2$. Then there is a shortest $w_g$ such that $w_g \in L(dfa(D_1))$ but $w_g \not\in L(dfa(D_2))$. Fig. 9. Sharply guarded automaton $H'$

The other inclusion $L(D') \subseteq L(D)$ follows in an analogous way.

\[ a[x_a = 2 \land x_b = 2] \quad b[x_a \geq 1 \land x_b \geq 2] \]

\[ 0 \quad b[x_a \geq 1 \land x_b \geq 1] \quad 2 \quad b[x_a \geq 2 \land x_b \geq 1] \quad 3 \]

\[ b[x_a \geq 2 \land x_b \geq 1] \]
By Lemma 11 this implies that there is a timed word \( w_t \) such that \( w_t \in \mathcal{L}(D_1) \) but \( w_t \notin \mathcal{L}(D_2) \), i.e., \( D_1 \not\equiv_t D_2 \). \( \square \)

We can now prove the central property of sharply guarded EDERAs.

**Theorem 1** For every EDERA there is a unique equivalent minimal sharply guarded EDERA (up to isomorphism).

**PROOF.** By Lemma 12, each EDERA \( D \) can be transformed into an equivalent EDERA \( D' \) that is sharply guarded. Let \( A_{\text{min}} \) be the unique minimal DFA which is equivalent to \( \text{dfa}(D') \) (up to isomorphism). Since (as was remarked after Definition 7) whether or not a EDERA is sharply guarded depends only on \( \mathcal{L}(\text{dfa}(D)) \), we have that \( D_{\text{min}} = \text{era}(A_{\text{min}}) \) is sharply guarded. By Lemma 14, \( D_{\text{min}} \) is the unique minimal sharply guarded EDERA (up to isomorphism) such that \( D_{\text{min}} \equiv D' \), i.e., such that \( D_{\text{min}} \equiv D \). \( \square \)

5.2 Learning guarded words

Angluin’s algorithm is designed to query (untimed) words rather than timed words. Before we can present the final learning algorithm for EDERAs, we must describe how the Assistant answers a membership query for a guarded word, posed by the Learner, by asking several membership queries for timed words to the (timed) Teacher.

To answer a membership query for a guarded word \( w_g \), the Assistant first extracts the untimed word \( w \) underlying \( w_g \). It thereafter determines the unique guard refinement \( w'_g \) of \( w \) that is accepted by \( A \) (if one exists) by posing several membership queries to the (timed) Teacher, in a way to be described below. Note that, as observed just before Example 6, each untimed word \( w \) has at most one guard refinement accepted by \( A \). Finally, the Assistant answers the query by yes iff \( w'_g \) equals \( w_g \).

The guard refinement of \( w \) accepted by \( A \) will be determined inductively, by learning the guard under which an action \( a \) is accepted, provided that a sequence \( u \) of actions has occurred so far. Letting \( u \) range over successively longer prefixes of \( w \), the Assistant can then learn the guard refinement \( w'_g \) of \( w \). Let \( u = a_1a_2\ldots a_n \), and assume that for \( i = 1, \ldots, n \), the Assistant has previously learned the guard \( g_i \) under which \( a_i \) is accepted, given that the sequence \( a_1 \ldots a_{i-1} \) has occurred so far. He can then compute the strongest postcondition \( sp(u_g) \), where \( u_g = (a_1, g_1) \ldots (a_n, g_n) \). The Assistant must now determine the strongest guard \( g_a \) such that \( a \) is accepted after \( u_g \) precisely when \( \varphi_a \equiv sp(u_g) \land g_a \) holds. Note that by Definition 5, there is a unique
strongest $g_a$ with this property. In the following, we assume that $sp(u_g)$ and $\varphi_a$ are both in canonical form.

The guard $g_a$ is determined by inquiring whether a set of clock valuations $\gamma_a$ satisfy $\varphi_a$. Without loss of generality, the Assistant works only with integer valuations. For each $\gamma_a$ that satisfies the postcondition $sp(u_g)$, he can make a membership query for some clocked word $w_c(a, \gamma_a)$, where $w_c$ satisfies the guarded word $u_g$, since such a guarded word $w_c(a, \gamma_a)$ exists precisely when $\gamma_a \models sp(u_g)$. In other words, he can ask the (timed) Teacher for every integer point in the polyhedron $[sp(u_g)]$ whether it is in $[\varphi_a]$. A typical situation for two clocks is depicted in Figure 10.

Let us now describe how clock valuations $\gamma_a$ are chosen in membership queries in order to learn the guard $g_a$ for $a$. As mentioned before, we assume that the Assistant knows the maximal constant $K$ that can appear in any guard. This means that if a clock valuation $\gamma$ with $\gamma(x_b) > K$ satisfies $g_a$, then clock $x_b$ has no upper bound in $g_a$. Thus, by Lemma 11, the guard $g_a$ can be uniquely determined by two clock valuations, one in its biggest corner $bc_K(g_a)$, and one in its smallest corner $sc_K(g_a)$.

Let us consider how to find a clock valuation in $bc_K(g_a)$. Suppose first that the Assistant knows some clock valuation $\gamma_a$ that satisfies $\varphi_a$. The Assistant will then repeatedly increase the clock values in $\gamma_a$ until $\gamma_a$ is in $bc_K(g_a)$. This is done as follows. At any point during this process, let $\text{Max}$ be the set of clocks, initially empty, for which the Assistant knows that they have reached a maximum, which is at most $K$, and let $\text{Above} K$ be the set of clocks which have become more than $K$, and let $\text{Unknown} = C_\Sigma \setminus (\text{Max} \cup \text{Above} K)$ be the clocks for which a maximum value is still searched. At each iteration, the Assistant finds the maximal $k \in \{1, \ldots, K + 1\}$ such that the valuation $\gamma_a$ can be changed by increasing all clocks in $\text{Unknown}$ by $k$, keeping the clocks in $\text{Max}$ unchanged, and finding suitable values for the clocks in $\text{Above} K$ such that $\gamma_a$ still satisfies $\varphi_a$. This can be done by binary search using at most $\log K$ queries. The Assistant then lets $\gamma_a$ be this new valuation. For all clocks $x_b$ with $\gamma_a(x_b) \geq K + 1$, the Assistant concludes that $x_b$ has no upper bound in $\varphi_a$. These clocks are moved over from $\text{Unknown}$ to $\text{Above} K$. If $\gamma_a(x_b) \leq K$ for
some clocks \( x_b \in Unknown \) then among these a clock (or several clocks) must be found that cannot be increased, which will be moved over from \( Unknown \) to \( Max \).

Let us examine how to find a clock \( x_b \) in \( Unknown \) that cannot be increased, i.e., such that \( \varphi_a \) implies the constraint \( x_b \leq \gamma_a(x_b) \). The idea is to increase each clock in turn by 1 and see whether the result still satisfies \( \varphi_a \). The particularity to handle is that it may be possible to increase a clock \( x_b \) only together with other clocks, since \( sp(u_g) \) must be satisfied (e.g., in Figure 10 we see that if \( x_a \) is incremented in \( \gamma_a \) then \( x_b \) must also be incremented to stay in \( sp(u_g) \)).

To define this in more detail, let us regard \( sp(u_g) \) and \( g_a \) as fixed, and define for each integer valuation \( \gamma_a \) such that \( \gamma_a \models \varphi_a \) the relation \( \leq^\gamma_a \) on the set \( C_\Sigma \) of clocks by

\[
x_b \leq^\gamma_a x_c \quad \text{if} \quad sp(u_g) \text{ implies } x_b - x_c \leq \gamma_a(x_b) - \gamma_a(x_c)
\]

(recall that we assume \( sp(u_g) \) to be canonical). Thus, \( x_b \leq^\gamma_a x_c \) means that in order to keep \( \gamma_a \models sp(u_g) \) while incrementing \( \gamma_a(x_b) \) by 1 we should also increment \( \gamma_a(x_c) \) by 1. For a valuation \( \gamma_a \), define \( dep_{\gamma_a}(x_b) \) as \( \{ x_c : x_b \leq^\gamma_a x_c \} \).

This means that in order to keep \( \gamma_a \models sp(u_g) \) while incrementing \( \gamma_a(x_b) \) by 1 we should also increment all clocks in \( dep_{\gamma_a}(x_b) \) by 1.

Assume that \( \gamma_a \models \varphi_a \). For a set \( C \) of clocks, define \( \gamma_a[C \oplus k] \) as \( \gamma_a(x_b) + k \) for \( x_b \in C \) and \( \gamma_a(x_b) \) otherwise. From the definition of \( dep_{\gamma_a} \), we infer that \( \gamma_a[dep_{\gamma_a}(x_b) \oplus 1] \models sp(u_g) \) for all \( x_b \) in \( Unknown \). We claim that for each clock \( x_b \) in \( Unknown \) for which \( \gamma_a[dep_{\gamma_a}(x_b) \oplus 1] \not\models \varphi_a \), the clock constraint \( \varphi_a \) must contain the conjunct \( x_b \leq \gamma_a(x_b) \). To see this, we first note that \( \gamma_a(x_b) \leq K \) and \( \gamma_a[dep_{\gamma_a}(x_b) \oplus 1] \not\models \varphi_a \) together imply that there must be some \( x_c \) in \( dep_{\gamma_a}(x_b) \) such that \( \varphi_a \) contains the conjunct \( x_c \leq \gamma_a(x_c) \). We also note that \( x_c \in dep_{\gamma_a}(x_b) \) means that \( sp(u_g) \) contains the conjunct \( x_b - x_c \leq \gamma_a(x_b) - \gamma_a(x_c) \). Hence, since \( \varphi_a \) is canonical, it contains the conjunct \( x_b \leq \gamma_a(x_b) \).

To update \( Max \), we thus move to \( Max \) all clocks such that \( \gamma_a[dep_{\gamma_a}(x_b) \oplus 1] \not\models \varphi_a \). As an optimization, we can sometimes avoid to make one query for each clock in \( Unknown \) increased by 1 by first analysing the structure of the graph whose nodes are the clocks in \( Unknown \), and whose edges are defined by the relation \( \leq^\gamma_a \). It is then sufficient to make at most one query for some clock in each strongly connected component of this graph, and use it as a query for each clock in the component.

After an iteration, another iteration is performed by finding a \( k \) to increase the clocks that remain in \( Unknown \), and thereafter finding out which of these have reached their upper bounds. When \( Unknown = \emptyset \), a valuation in \( bc_K(g_a) \) has been found and the algorithm terminates.
Thus, all in all, determining the upper bound of a guard $g_a$ needs at most $|C|\Sigma$ binary searches, since in every loop at least one clock is moved to $\text{Max}$. Each uses at most $\log K + |C|\Sigma$ membership queries. In an analogous way, we can find a minimal clock valuation that satisfies $\varphi_a$. The guard $g_a$ is given by the $K$-approximation of the guard that has the minimal clock valuation as smallest corner and the maximal clock valuation as biggest corner, which can easily be formulated given these two points. Thus, the Assistant needs at most $2|C|\Sigma(\log K + |C|\Sigma)$ membership queries to learn a guard $g_a$, if initially it knows a valuation which satisfies $\varphi_a$.

Suppose now that the Assistant does not know a clock valuation $\gamma_a$ that satisfies $\varphi_a$. In principle, $\varphi_a$ and therefore $g_a$ could specify exactly one valuation, meaning that the Assistant essentially might have to ask membership queries for all $\binom{|\Sigma|+K}{|\Sigma|}$ integer points that could be specified by $\varphi_a$. This is the number of non-increasing sequences of $|\Sigma| = |C|$ elements, where each element has values among 0 to $K$, since $sp(u_g)$ defines at least an ordering on the clocks.

Thus, the Assistant can answer a query for a guarded word $w_g$ using at most $|w|\binom{|\Sigma|+K}{|\Sigma|}$ (timed) membership queries.

### 5.3 Algorithm $TL^*_sg$

Let us now turn to the problem of learning a timed language $L(D)$ accepted by an EDERA $D$. We can assume without loss of generality that $D$ is the unique minimal and sharply guarded EDERA that exists due to Theorem 1. Then $D$ is uniquely determined by its symbolic language of $A = \text{dfa}(D)$, which is a regular (word) language. In this setting, we assume

- to know an upper bound $K$ on the constants occurring in guards of $D$,
- to have a Teacher who is able to answer two kinds of queries:
  - A membership query consists in asking whether a timed word $w_t$ over $\Sigma$ is in $L(D)$.
  - An equivalence query consists in asking whether a hypothesized EDERA $H$ is correct, i.e., whether $L(H) = L(D)$. The Teacher will answer yes if $H$ is correct, or else supply a counterexample $u$, either in $L(D) \setminus L(H)$ or in $L(H) \setminus L(D)$.

Based on the observations in Section 5.1, our solution is to learn $L(\text{dfa}(D))$, which is a regular language and can therefore be learned in principle using Angluin’s learning algorithm. However, Angluin’s algorithm is designed to query (untimed) words rather than timed words. Let us therefore extend the Learner in Angluin’s algorithm by an Assistant, whose role is to answer a membership query for a guarded word, posed by the Learner, by asking several
T = \{ (a, x_a = 2, x_b = 2) + (b, x_a \geq 1 \land x_b \geq 1) + \}

\begin{tabular}{|c|c|}
\hline
T & \lambda \\
\hline
\lambda & + \\
(a, x_a = 2, x_b = 2) & + \\
(b, x_a \geq 1 \land x_b \geq 1) & + \\
\hline
\end{tabular}

0

Table 11. Table \( T \), membership queries for timed words to the (timed) Teacher. This is described in Section 5.2. To complete the learning algorithm, we have to explain how the Assistant can answer equivalence queries to the Learner. Given a DFA \( H \), the Assistant can ask the (timed) Teacher, whether \( \text{era}(H) = D \). If so, the Assistant replies yes to the Learner. If not, the Teacher presents a timed word \( w_t \) that is in \( L(D) \) but not in \( L(\text{era}(H)) \) (or the other way round). For the word \( w \) underlying \( w_t \), we can obtain its guard refinement \( w_g \) as described in the previous paragraph. Then \( w_g \) is in \( L(\text{dfa}(D)) \) but not in \( L(H) \) (or the other way around). Thus, the Assistant can answer the equivalence query by \( w_g \) in this case.

At this point, we should remark that it can be the case that hypothesized automaton \( H \) which the algorithm constructs is not sharply guarded. This can happen if the observation table does not contain for each prefix \( u_g \) of the table and each symbol \( a \in \Sigma \) at least one column labeled by a suffix of form \( (a, g) \) such that \( u_g(a, g) \) is accepted. As an illustration, consider the hypothesized automaton \( H \) in Figure 11(b), constructed from the table \( T \) shown in Figure 11(a). Let us assume that \( K = 4 \) is given as an upper bound on constants in guards. The automaton \( H \) is non-sharply guarded, since after the guarded word \( (a, x_a = 2 \land x_b = 2)(b, x_a \geq 1 \land x_b \geq 1) \in L(\text{dfa}(H)) \) the postcondition implies \( x_b \geq 3 \), which means that after this guarded word, the guard on the following \( b \)-transition is not sharp. A so constructed non-sharply guarded automaton has always less locations than a corresponding sharply guarded automaton constructed from the same information.

We call the algorithm outlined in the section \( \mathcal{T}L_{sg} \). More specifically, the algorithm for learning sharply guarded EDERA is as Algorithm 2, but extended with the Assistant shown in Algorithm 3.

**Example 15** \(^{(12)}\) Let us illustrate the algorithm by showing how to learn the language of the automaton \( A \) depicted in Figure 12(a). Initially, the algorithm asks membership queries for \( \lambda \). It additionally asks membership queries to learn that \( (a, g) \) is accepted iff \( g = x_a \leq 1 \land x_b \leq 1 \) and \( (b, g) \) is accepted iff \( g = x_a = 0 \land x_b = 0 \). To follow the algorithm, we should also add rejected guarded words to the table: these are needed in order to find inconsistencies. In

\(^{(12)}\)TODO: I changed slightly the following paragraph, and moved the example here/BJ
Algorithm 3 Pseudo code for Assistant of $TL_{sg}^*$

```java
class $TL_{sg}^*$Assistant implements Teacher{
    Teacher timedteacher

Constructor $TL_{sg}^*$Assistant(teacher)
    timedteacher = teacher

Function equivalence_query($\mathcal{H}$)
    timedteacher equivalence_query($\mathcal{H}$)
    if the answer to equivalence query is a counterexample $w$
        Extract the word $w$ underlying $w$
        Learn guard refinement $w_g$ of $w$
    return $w_g$
    else
        return 'yes'

Function membership_query($w_g$)
    Extract underlying $w$ of $w_g$
    Learn guard refinement $w'_g$ of $w$
    if $w'_g = w_g$ then
        return 'yes'
    else
        return 'no'
}
```

Algorithm 4 Pseudo code for $TL_{sg}^*$

```java
class $TL_{sg}^*$ extends L*

Constructor $TL_{sg}^*$($timedteacher$, $\Sigma$, $K$)
    $\Gamma = TL_{sg}^*$Alphabet($\Sigma$, $K$)
    teacher = $TL_{sg}^*$Assistant($timedteacher$)
}
```

(a) Automaton $\mathcal{A}$

```
\begin{align*}
    a & \mid x_a \leq 1 \\
    0 & \\
    b & \mid x_a = 0 \land x_b = 0 \\
\end{align*}
```

(b) Automaton $\mathcal{H}_1$

```
\begin{align*}
    a & \mid x_a \leq 1 \\
    b & \mid x_a = 0 \land x_b = 0 \\
    b & \mid x_a = 0 \land x_b = 0 \\
\end{align*}
```

(c) Automaton $\mathcal{H}_2$

```
\begin{align*}
    a & \mid x_a \leq 1 \land x_b \leq 1 \\
    0 & \\
    b & \mid x_a = 0 \land x_b = 0 \\
\end{align*}
```

Fig. 12. EDERAs
this example, we need to have only \((a, x_a ≤ 1 ∧ x_b ≥ 0)\) in the table, and in order to keep the table as small as possible we do not add other rejected guarded words. This yields the initial observation table \(T_1\) shown in Figure 13(a). It is consistent and closed. Then the Learner constructs a hypothesized DERA \(\mathcal{H}_1\) shown in Figure 12(b) and submits \(\mathcal{H}_1\) in an equivalence query. Assume that the counterexample \((a, 1.0)(a, 1.5)\) is returned. It is accepted by \(\mathcal{A}\) but rejected by \(\mathcal{H}_1\). The algorithm processes the counterexample and produces the observation table \(T_2\) given in Figure 13(b), which is not consistent. Following Angluin’s algorithm we construct a closed and consistent table \(T_3\) shown in Figure 14. The sharply guarded EDERA \(\mathcal{H}_2\) visualized in Figure 12(c) corresponds to the observation table \(T_3\) and accepts the same language as \(\mathcal{A}\).

5.4 Complexity

In the \(L^*\) algorithm the number of membership queries is bounded by \(O(kn^2m)\), where \(n\) is the number of states, \(k\) is the size of the alphabet, and \(m\) is the length of the longest counterexample.

In our setting, a single membership query for a guarded word \(w_g\) might give...
rise to \(|w|(|\Sigma| + K)\) membership queries to the (timed) Teacher. The alphabet of the DFA \(\text{dfa}(D)\) is \(\Sigma \times G\). Thus, the query complexity of \(TL^*_sg\) for a sharply guarded EDERA with \(n\) locations is

\[
O\left( kn^2 ml \left( \frac{|\Sigma| + K}{|\Sigma|} \right) \right)
\]

where \(l\) is the length of the longest guarded word queried and \(k\) is the size of alphabet \(\Sigma \times G\). The longest queried word is bounded by \(O(m + n)\). If the Teacher always presents counterexamples of minimal length, then \(m\) is bounded by \(O(n)\). The number of equivalence queries remains at most \(n\). Note that, in general a (non-sharply guarded) EDERA \(D\) gives rise to a sharply guarded EDERA with double exponentially more locations. Since the number of regions is bounded by \(|\Sigma|!^2 (2|\Sigma| (2K + 2)|\Sigma|)\), the number of location in sharply guarded EDERA is bounded by \(n_1 2^{|\Sigma|} (2|\Sigma|)\), where \(n_1\) is the number of locations in EDERA. Thus, the query complexity of \(TL^*_sg\) for EDERA with \(n_1\) locations is

\[
O\left( kn_1^2 2^{|\Sigma|} ml \left( \frac{|\Sigma| + K}{|\Sigma|} \right) \right)
\]

6 Learning of DERA

Let us now turn our attention to learn the full class of deterministic event recording automata. The scheme for developing a learning algorithm is analogous to the scheme used for EDERAs in Section 5: we define a class of DERAs that admit a natural definition of right congruences, so that a DERA \(D\) in this class uniquely determines a language \(L(\text{dfa}(D))\). We show that each DERA can be transformed to this form. Then our solution is to learn \(L(\text{dfa}(D))\) using an assistant, whose role is to answer membership queries for guarded words by asking membership queries for timed words. In order to cope with the class of all DERAs, we need to find a different unique representation, and to change the task of the assistant.

6.1 Simple DERAs

**Definition 16** A \(K\)-bounded DERA \(D\) is simple if all its guards are simple and if whenever \(w_g(a, g)\) is a prefix of some word in \(L(\text{dfa}(D))\), then \(sp^<_K(w_g) \land g\) is satisfiable. \(\square\)

We remark that whether or not a DERA is simple depends only on \(L(\text{dfa}(D))\). A consequence of this definition is the following.
Lemma 17 If \( w_g \in \mathcal{L}(dfa(D)) \), where \( D \) is a simple DERA, then there is a timed word \( w_t \in \mathcal{L}(D) \) such that \( w_t \models w_g \).

**PROOF.** The lemma follows by induction from the fact that \( sp^R_K(w'_g) \land g' \) is satisfiable iff \( sp(w'_g) \land g' \) is satisfiable, whenever \( w'_g \) is a guarded word and \( g' \) is a \( K \)-bounded simple guard. \( \square \)

We prove an important property of simple guarded words.

Lemma 18 Let \( w_g = (a_1, g_1) \ldots (a_n, g_n) \) be a \( K \)-bounded simple guarded word, and let \( w_t = (a_1, \gamma_1) \ldots (a_n, \gamma_n) \) and \( w'_t = (a_1, \gamma'_1) \ldots (a_n, \gamma'_n) \) be two clocked words. If \( w_t \models w_g \) and \( w'_t \models w_g \) then \( \gamma_i \sim_K \gamma'_i \) for \( 1 \leq i \leq n \).

**PROOF.** Since \( \gamma_i \models g_i \) and \( \gamma'_i \models g_i \), then for all \( 1 \leq i \leq n \) and \( x_a \in C_\Sigma \) such that \( \gamma_i(x_a) \leq K \) and \( \gamma'_i(x_a) \leq K \) we have \( [\gamma_i(x_a)] = [\gamma'_i(x_a)] \), and \( \text{frac}(\gamma_i(x_a)) = 0 \) iff \( \text{frac}(\gamma'_i(x_a)) = 0 \). We prove by induction over \( i \) that for \( 1 \leq i \leq n \) we have \( \text{frac}(\gamma_i(x_{a_1})) - \text{frac}(\gamma'_i(x_{a_1})) \geq 0 \) iff \( \text{frac}(\gamma'_i(x_{a_1})) - \text{frac}(\gamma'_i(x_{a_k})) \geq 0 \) for all \( a_j, a_k \in \Sigma \), whenever \( \gamma_i(x_{a_1}) \leq K \) and \( \gamma'_i(x_{a_k}) \leq K \). For \( i = 1 \), this follows from \( \gamma_1(x_{a_1}) = \gamma_1(x_{a_k}) \) and \( \gamma'_1(x_{a_1}) = \gamma'_1(x_{a_k}) \). Assume that \( \gamma_i \sim_K \gamma'_i \), that \( \gamma_{i+1}(x_{a_1}) \leq K \), that \( \gamma_{i+1}(x_{a_k}) \leq K \), and that \( \text{frac}(\gamma_{i+1}(x_{a_1})) - \text{frac}(\gamma_{i+1}(x_{a_k})) \geq 0 \). From \( \gamma_i \sim_K \gamma'_i \), together with \( [\gamma_{i+1}(x_{a_1})] = [\gamma'_i(x_{a_1})] \) and \( [\gamma_{i+1}(x_{a_k})] = [\gamma'_i(x_{a_k})] \) we deduce \( \text{frac}(\gamma'_i(x_{a_1})) - \text{frac}(\gamma'_i(x_{a_k})) \geq 0 \). \( \square \)

Every DERA can be transformed into an equivalent DERA that is simple using the region-graph construction [Alu99].

Lemma 19 For every \( K \)-bounded DERA there is an equivalent \( K \)-bounded DERA that is simple.

**PROOF.** Let the DERA \( D = (\Sigma, L, l_0, L^f, \delta) \) be \( K \)-bounded. We define an equivalent simple DERA \( D' = (\Sigma, L', l'_0, L'^f, \delta') \) by adapting the region graph construction for \( D \).

The set of locations \( L' \) of \( D' \) comprises pairs \( (l, \varphi) \) where \( l \in L \) and \( \varphi \) is a \( K \)-bounded clock constraint. The intention is that \( \varphi \) is the \( K \)-approximated postcondition of any run from the initial location to \( (l, \varphi) \). The initial location \( l'_0 \) of \( D' \) is \( (l_0, sp(\lambda)) \). We also introduce the location \( l_e = (l_0, true) \) in \( L' \), where \( l_e \not\in L'^f \). For every symbol \( a \) and \( K \)-bounded simple guard \( g' \) for which there is a guard \( g \) such that \( \delta(l, a, g) \) is defined and \( g' \) implies \( g \), let \( \delta'(l, \varphi, a, g') \) be defined as \( (l', \varphi') \) where \( l' = \delta(l, a, g) \) and \( \varphi' = \langle (\varphi \land g')[x_a \mapsto 0] \uparrow \rangle \) \( K \)
We call two simple DERAs $D_1$ and $D_2$ dfa-equivalent, denoted $D_1 \equiv_{\text{dfa}} D_2$, if $\text{dfa}(D_1)$ and $\text{dfa}(D_2)$ accept the same language (in the sense of $K$).

The important property of simple DERAs is that equivalence coincides with equivalence on the corresponding DFAs.

**Definition 20** We call two simple DERAs $D_1$ and $D_2$ dfa-equivalent, denoted $D_1 \equiv_{\text{dfa}} D_2$, iff $\text{dfa}(D_1)$ and $\text{dfa}(D_2)$ accept the same language (in the sense
of DFAs).

Now, exactly as in Section 5, we get counterparts for Lemma 14 and Theorem 1.

**Lemma 21** For two simple DERAs $D_1$ and $D_2$, we have

$$D_1 \equiv_t D_2 \iff D_1 \equiv_{dfa} D_2$$

We can now prove the central property of simple DERAs.

**Theorem 2** For every DERA there is a unique equivalent minimal simple DERA (up to isomorphism).

**PROOF.** The proof is analogous to the proof of Theorem 1. \qed

### 6.2 Algorithm $TL^*_s$

Given a timed language that is accepted by a DERA $D$, we can assume without loss of generality that $D$ is the unique minimal and simple one that exists due to Theorem 2. Then $D$ is uniquely determined by its symbolic language of $A = dfa(D)$, which is a regular (word) language over $\Sigma \times G_s$, where $G_s$ is a set of simple clock guards. Thus, we can learn $A$ using Angluin’s algorithm and return $era(A)$. However, $\mathcal{L}(A)$ is a language over simple guarded words, but the Teacher in the timed setting is supposed to deal with timed words rather than guarded words. Then it can be the case that the Teacher answers yes to equivalence query for hypothesized automaton $\mathcal{H}$ and $\mathcal{H}$ is smaller than $A$.

Similar as in the previous section, we extend the Learner in Angluin’s algorithm by an Assistant, whose role is to answer a membership query for a simple guarded word, posed by the Learner, by asking membership query for timed words to the (timed) Teacher. Furthermore, it also has to answer equivalence queries, consulting the timed Teacher.

For a simple guarded word $w_g = (a_1, g_1) \ldots (a_n, g_n)$ each simple guard $g$ that extends $w_g$ together with an action $a$ defines exactly one region. Thus, if $w_g$ is accepted, it is enough to check $a$ in a single point in this region defined by $g$ and the postcondition of $w_g$. In other words, it suffices to check an arbitrary timed word $w_t \models w_g$ to check whether $w_g$ is in the symbolic language or not.

The number of successor regions that one region can have is $O(|\Sigma|K)$. Then the complexity of the algorithm for the simple DERA with $n$ locations is
$O(|\Sigma|^2 n^2 m K)$, where $m$ is the length of the longest counterexample.

**Algorithm 5** Pseudo code for Assistant of $TL_s^*$

```java
class TLsAssistant implements Teacher{
    Teacher timedteacher
    Constructor TLsAssistant(teacher)
        timedteacher = teacher
    Function equivalence_query(H)
        timedteacher.equivalence_query(H)
        if the answer to equivalence query is a counterexample $w_t$
            Construct simple guarded word $w_g$ such that $w_t \models w_g$
            return $w_g$
        else
            return 'yes'
    Function membership_query($w_g$)
        if there is no $w_t$ such that $w_t \models w_g$
            return 'no'
        else
            Choose $w_t$ such that $w_t \models w_g$
            timedteacher.membership_query($w_t$)
            if answer to membership query is 'no'
                return 'no'
            else
                return 'yes'
}
```

**Algorithm 6** Pseudo code for $TL_s^*$

```java
class TLs extends L*
    Constructor TLs(timedteacher, $\Sigma, K$)
        $\Gamma = TLsAlphabet(\Sigma, K)$
        teacher = TLsAssistant(timedteacher)
    }
```

**Example 22** Let us illustrate the algorithm by showing how to learn the language of the automaton $A$ depicted in Figure 15. Initially, the algorithm asks membership queries for $\lambda$, $(a, x_a = 0)$, $(a, 0 < x_a < 1)$, $(a, x_a = 1)$ and $(a, x_a > 1)$. This yields the initial observation table $T_1$ shown in Table 16(a). It is consistent but not closed, since $row((a, x_a = 0))$ is distinct from $row(\lambda)$. Following Angluin’s algorithm, we can construct a closed and consistent table $T_2$ shown in Table 16(b). Then the Learner constructs a hypothesized DERA $H_1$ shown in Figure 17 and submits $H_1$ in an equivalence query. Assume that the counterexample $(a, x_a = 0)(a, x_a = 0)(a, 0 < x_a < 1)$ is returned. It is rejected by $A$ but accepted by $H_1$. The algorithm processes the counterexample and finally produces the observation table $T_3$ given in Table 18. The automaton $H_2$ visualized in Figure 19 corresponds to the observation table $T_3$ and accepts the same language as $A$. 

Learning a sharply guarded EDERA allows to transfer Angluin’s setting to the timed world. However, in practice, one might be interested in a smaller non-sharply guarded EDERA rather than its sharply guarded version. In this section, we describe how to learn a usually smaller, non-sharply guarded version. The idea is to identify states whose futures are “similar”. While in the worst-case, more membership queries is needed, we hope that the algorithm converges faster in practice. In Figure 22 is shown EDERA $A_2$ constructed by the algorithm $TL_{sg}^*$. In Figure 21(a) is shown EDERA $A_1$, which accepts the same language as $A_2$, but has less locations than $A_2$. From $TL_{sg}^*$ point of view, locations 2 and 3 in $A_2$ are different, since $a$-transition from location 2 can be taken if $x_a = 1$ and $x_b = 3$, while from location 3 $a$-transition can not be taken if $x_a = 1$ and $x_b = 3$ (when we reach location 3, postcondition is
$$T_3$$

<table>
<thead>
<tr>
<th>$\lambda$</th>
<th>$(a, x_a = 0)$</th>
<th>$(a, x_a = 0)(a, x_a = 0)$</th>
<th>$(a, x_a = 0)(a, x_a = 0)(a, 0 &lt; x_a &lt; 1)$</th>
<th>$\lambda$</th>
<th>$(a, x_a &gt; 1)$</th>
<th>$(a, x_a = 0)(a, 0 &lt; x_a &lt; 1) \land (a, x_a = 0)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>+</td>
<td>-</td>
<td>1</td>
<td></td>
<td>+</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>-</td>
<td>+</td>
<td>+</td>
<td></td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>+</td>
<td>-</td>
<td>+</td>
<td></td>
<td>-</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>-</td>
<td>-</td>
<td>+</td>
<td></td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
</tbody>
</table>

Fig. 18. Table $T_3$

$$x_b - x_a = 3$$). However, we can merge location 2, 3 and 4 if we add $a$-transition from location 2 to location 2 with the guard $x_a = 1 \land x_b \leq 4$. In this section we show how to find guards which allow to merge locations, non-mergeable from $TL^*_a$ point of view. (13) We develop our ideas in the setting of learning non-sharply guarded EDERAs, but a similar study could be carried out to learn non-simple DERAs.

7.1 Learning based unification

Let us now define a relationship on guarded words, which will be used to merge states whose futures are “similar”, taking the postcondition into account.
Fig. 20. Not unifiable $PG = \{ (\varphi_1, (a, g_1)), (\varphi_2, (a, g_2)), (\varphi_3, (a, g_3)) \}$

Let $PG = \{ (\varphi_1, (a_1, g_{11}) \ldots (a_n, g_{1n})), \ldots, (\varphi_k, (a_1, g_{k1}) \ldots (a_n, g_{kn})) \}$ be a set of $k$ pairs of postconditions and guarded words with the same sequences of actions. We say that the guarded word $(a_1, \hat{g}_1) \ldots (a_n, \hat{g}_n)$ unifies $PG$ if for all $j \in \{1, \ldots, k\}$ and $i \in \{1, \ldots, n\}$

$$g_{ji} \land sp(\varphi_j, (a_1, g_{j1}) \ldots (a_{i-1}, g_{j(i-1)})) \equiv \hat{g}_i \land sp(\varphi_j, (a_1, \hat{g}_1) \ldots (a_{i-1}, \hat{g}_{i-1}))$$

Then, the set $PG$ is called unifiable and $(a_1, \hat{g}_1) \ldots (a_n, \hat{g}_n)$ is called a unifier. Intuitively, the guarded words with associated postconditions can be unified if there is a unifying, more liberal guarded word, which is equivalent to all guarded words in the context of the respective postconditions. Then, given a set of guarded words with postconditions among $\{ \varphi_1, \ldots, \varphi_k \}$, these guarded words can be considered to yield the same state, provided that the set of future guarded actions together with the respective postcondition is unifiable.

In the next example we show that if every pair in $PG$ is unifiable it does not follow that $PG$ is unifiable.
Example 23 Let

\[ g_1 = (x_a \geq 5 \land x_a \leq 7 \land x_b \geq 5 \land x_b \leq 7), \]
\[ \varphi_1 = (x_a = x_b), \]
\[ g_2 = (x_a \geq 7 \land x_a \leq 9 \land x_b \geq 3 \land x_b \leq 5), \]
\[ \varphi_2 = (x_b = x_a - 4), \]
\[ g_3 = (x_a \geq 9 \land x_a \leq 11 \land x_b \geq 1 \land x_b \leq 3), \]
\[ \varphi_3 = (x_b = x_a - 8), \]

see Figure 20(a). Let \( PG = \{(\varphi_1, (a, g_1)), (\varphi_2, (a, g_2)), (\varphi_3, (a, g_3))\} \) and

\[ g_4 = (x_a \geq 5 \land x_a \leq 9 \land x_b \geq 3 \land x_b \leq 7), \]
\[ g_5 = (x_a \geq 7 \land x_a \leq 11 \land x_b \geq 1 \land x_b \leq 5), \]
\[ g_6 = (x_a \geq 5 \land x_a \leq 11 \land x_b \geq 1 \land x_b \leq 7) \]

Then \((a, g_4)\) is the strongest unifier for \( \{\varphi_1, (a, g_1))\), see Figure 20(b), \((a, g_3)\) is the strongest unifier for \( \{\varphi_2, (a, g_2))\), see Figure 20(c) and \((a, g_6)\) is the strongest unifier for \( \{\varphi_1, (a, g_1))\), see Figure 20(d). Then the strongest possible unifier for \( PG \) should be \( g_6 \), but \( \varphi_2 \land g_6 \not\equiv \varphi_2 \land g_2 \). Then \( PG \) is not unifiable. \( \square \)

It is easy to check, whether \( PG \) is unifiable, using the property that the guards in \( PG \) are tight in the sense of Definition 7. The basic idea in each step is to take the weakest upper and lower bounds for each variable. Assume the guard \( g_{ji} \) is given by its upper and lower bounds:

\[ g_{ji} = \bigwedge_{a \in \Sigma} (x_a \leq c_{a,ji}^\leq \land x_a \geq c_{a,ji}^\geq) \]

For \( i = 1, \ldots, n \), define the candidate \( \hat{g}_i \) as

\[ \hat{g}_i = \bigwedge_{a \in \Sigma} \left(x_a \leq \max_j \{c_{a,ji}^\leq\}\right) \land \bigwedge_{a \in \Sigma} \left(x_a \geq \min_j \{c_{a,ji}^\geq\}\right) \]

and check whether the guarded word \( (a_1, \hat{g}_1) \ldots (a_n, \hat{g}_n) \) obtained in this way is indeed a unifier. We represent \textit{false} as the constraint \( \bigwedge_{a \in \Sigma} x_a \leq 0 \land x_a \geq K+1 \). It can be shown that if \( PG \) is unifiable, then this candidate is the strongest possible unifier.

The learning algorithm using the idea of unified states works similar as the one for EDERAs. However, we employ a slightly different observation table. Let \( \Gamma = \Sigma \times G_\Sigma \). Rows of the table are guarded words of a prefix-closed set \( U \subseteq \Gamma^* \). Column labels are untimed words from a set \( V \subseteq \Sigma^* \). The
entries of the table are sequences of guards describing under which values the
column label extends the row label. Thus, we define a **timed observation table**
\( T : U \cup UT \rightarrow (V \rightarrow G_2^*) \), where \( T(u_g)(v) = g_1 \ldots g_n \) implies \( |v| = n \). We
require the initial observation table to be defined over \( U = \{ \lambda \} \) and \( V = \Sigma \).

We define \( u_g \in U \cup UT \) and \( u'_g \in U \cup UT \) to be \( v \)-unifiable if \( v = a_1 \ldots a_n \in V \),
\( T(u_g)(v) = g_1 \ldots g_n \), \( T(u'_g)(v) = g'_1 \ldots g'_n \) and \( \{ (sp(u_g), (a_1, g_1) \ldots (a_n, g_n)), (sp(u'_g), (a_1, g'_1) \ldots (a_n, g'_n)) \} \) is unifiable. We define \( u_g \in U \cup UT \) and \( u'_g \in U \cup UT \) to be unifiable if for every \( v \in V \), \( u_g \) and \( u'_g \) are \( v \)-unifiable.

A timed observation table is **closed** if for every \( u_g \in UT \) there is \( u'_g \in U \) such that
\( u_g \) and \( u'_g \) are unifiable. A timed observation table is **consistent** if for all
\( u_g, u'_g \in U \) whenever \( u_g \) and \( u'_g \) unifiable, and \( u_g(a, g), u'_g(a, g') \in U \cup UT \) then
\( u_g(a, g) = u'_g(a, g') \) are unifiable.

A **merging** of the timed observation table \( T \) consists of a partition \( \Pi \) of the
-guarded words \( U \cup UT \), and an assignment of a clock guard \( CG(\pi, a) \) to each
block \( \pi \in \Pi \) and action \( a \in \Sigma \), such that for each block \( \pi \in \Pi \) we have

- for each suffix \( v = a_1 \ldots a_n \in V \), the set \( \{ (sp(u_g), (a_1, g_1) \ldots (a_n, g_n)) \mid u_g \in π, T(u_g)(v) = g_1 \ldots g_n \} \) is unifiable,
- if \( u_g, u'_g \in \pi \) and \( u_g(a, g), u'_g(a, g') \in U \cup UT \) then \( u_g(a, g) = u'_g(a, g') \) for
some block \( \pi' \in \Pi \), and
- \( (a, CG(\pi, a)) \) is the unifier for \( \{ (sp(u_g), (a, g')) \mid u_g \in \pi, T(u_g)(a) = g' \} \) for
each \( a \in \Sigma \).

Intuitively, a merging defines a grouping of rows into blocks, each of which
can potentially be understood as a state in a EDERA, together with a choice
of clock guard for each action and block, which can be understood as a guard
for the action in the EDERA. For each table there are in general several
possible mergings, but the number of mergings is bounded, since the number
of partitions is bounded, and since the number of possible unifiers \( GC(\pi, a) \)
is also bounded.

A **coarsest merging** of the timed observation table \( T \) is a merging with a
minimal number of blocks. From a closed and consistent timed table we can
get a lower bound on the number of blocks. It follows from Example 23 that
in order to construct coarsest merging we need to check whether all rows in
block are unifiable.

Given a merging \((\Pi, GC)\) of a closed and consistent timed observation table
\( T \), one can construct the EDERA \( \mathcal{H} = (\Sigma, L, l_0, g, \eta) \), where

- \( L = \Pi \) comprises the blocks of \( \Pi \) as locations,
- \( l_0 = \pi \in \Pi \) with \( \lambda \in \pi \) is the initial location,
- \( g(\pi, a) = \pi' \), if there are \( u \in \pi \) and \( g \) such that \( u \in U \), \( u(a, g) \in \pi' \) and
\[ T(u)((a, g)) \neq false, \text{ otherwise } \varrho(\pi, a) \text{ is undefined.} \]

- \( \eta \) is defined by \( \eta(\pi, a) = GC(\pi, a) \).

7.2 Algorithm TL*_{nsg}

The algorithm \( TL_{nsg}^* \) for learning (non-sharply guarded) EDERAs is as \( TL_{sg}^* \), except that the new notions of closed and consistent are used. If in Algorithm 2 check for closeness and consistency compares two rows in a table on equality, the algorithm \( TL_{nsg}^* \) checks whether rows are unifiable. Since entries of timed observation table are sequences of guards, we introduce function \( \text{learn\_guard} \), which learns guard refinement of some suffix of a word. (14) One further modification is that the hypothesis is constructed as described in the previous paragraph, using the computed merging. The rest of the algorithm remains unchanged (see Algorithm 8).

Algorithm 7 Pseudo code for Assistant of \( TL_{nsg}^* \)

```java
class TL_{nsg}^*Assistant extends TL_{sg}^*Assistant {
  Teacher timedteacher
  Constructor TL_{nsg}^*Assistant(teacher)
    timedteacher = teacher
  Function learn\_guard(u_g, w)
    Extract underlying u of u_g
    Learn guard refinement u_g(a_1, g_1) . . . (a_n, g_n) of uw
    return g_1 . . . g_n
}
```

Lemma 24 Let \( \mathcal{A} \) be a smallest EDERA equivalent to the system that is to be learned. Let \( |\mathcal{A}| \) be the number of locations in \( \mathcal{A} \). Then the algorithm \( TL_{nsg}^* \) terminates and constructs an EDERA \( \mathcal{A}' \) with \( |\mathcal{A}| \) locations, which is equivalent to \( \mathcal{A} \). (15)

**PROOF.** We first prove that every coarsest merging constructed from a timed observation table has at most \( |\mathcal{A}| \) blocks. Assume that the algorithm \( TL_{nsg}^* \) constructs the timed observation table \( T : U \cup UT \to (V \to G^*) \). Assume that \( u_g^1, . . . , u_g^n \) are all rows in \( U \cup UT \) such that \( v_g^1, . . . , v_g^n \) lead to the same location \( l \) in \( \mathcal{A} \) and for each \( 1 \leq i \leq n \), the word underlying \( v_g^i \) is equal to the word underlying \( u_g^i \). Since \( w_l \models v_g^i \) iff \( w_l \models v_g^i \), we infer \( sp(v_g^i) = sp(u_g^i) \) for each \( 1 \leq i \leq n \). Let \( a_1 . . . a_m \in V \) and \( T(u_g^i)(a_1 . . . a_m) = g_{i1} . . . g_{im} \) for each \( 1 \leq i \leq n \). Let \( v_g^1(a_1, g_1) . . . (a_m, g_m) \) lead to some location in \( \mathcal{A} \). Since \( w_l \models v_g^j(a_1, g_1) . . . (a_m, g_m) \) iff \( w_l \models v_g^j(a_1, g_1) . . . (a_m, g_m) \), we infer that for each \( 1 \leq i \leq n \) and \( 1 \leq j \leq m - 1 \) we have \( sp(v_g^i(a_1, g_1) . . . (a_j, g_j)) = \)
Algorithm 8 Pseudo code for $TL^*_{nsg}$

```java
class TL^*_{nsg} extends L^*{
    Alphabet Σ
    Constructor TL^*_{nsg}(timedteacher, Σ, K)
        Γ = TL^*_{nsg}Alphabet(Σ, K)
        this.Σ = Σ
        teacher = TL^*_{nsg}Assistant(timedteacher)
    Function initialize((U, V, T))
        U := {λ}, V := Σ
        for every a ∈ Σ
            g = teacher.learn_guard(λ, a)
            if g ≠ false
                Add (a, g) to UT
        for every u_g ∈ U ∪ UT and a ∈ Σ
            T(u_g)(a) = teacher.learn_guard(u_g, a)
    Function isClosed((U, V, T))
        if for each u_g ∈ UT there is u'_g ∈ U such that u_g and u'_g are unifiable then
            return true
        else
            return false
    Function isConsistent((U, V, T))
        if for each a ∈ Σ and u_g, u'_g ∈ U such that u_g(a, g), u'_g(a, g') ∈ U ∪ UT, and
        u_g and u'_g are unifiable we have u_g(a, g) and u'_g(a, g') are unifiable then
            return true
        else
            return false
    Function add_column()
        Find a ∈ Σ, v ∈ V, u_g, u'_g ∈ U and u_g(a, g), u'_g(a, g') ∈ U ∪ UT such that
        u_g and u'_g are unifiable, but u_g(a, g) and u'_g(a, g') are not v-unifiable
        Add av to V
        for every u_g ∈ U ∪ UT
            T(u_g)(av) = teacher.learn_guard(u_g, av)
    Function move_row()
        Find u_g ∈ UT such that for all u'_g ∈ U, u_g and u'_g are not unifiable
        Move u_g to U
        for every a ∈ Σ
            g = teacher.learn_guard(u_g, a)
            if g ≠ false
                Add u_g(a, g) to UT
                for every v ∈ V
                    T(u_g(a, g))(v) = teacher.learn_guard(u_g(a, g), v)
    Function process_counterexample(u_g)
        Add every prefix u'_g of u_g to UT.
        for every v ∈ V and prefix u'_g of u_g
            T(u'_g(v))(v) = teacher.learn_guard(u'_g, v)
        for every a ∈ Σ and prefix u'_g of u_g
            g = teacher.learn_guard(u'_g, a)
            if g ≠ false
                Add u'_g(a, g) to UT
                for every v ∈ V
                    T(u'_g(a, g))(v) = teacher.learn_guard(u'_g(a, g), v)
}
```

44
a \[ x_a = 1 \land x_b \leq 4 \]

\[ (a) \text{ Automaton } \mathcal{A}_1 \]

\[
\begin{array}{|c|c|}
\hline
T & a \\
\hline
u_1 & x_a = 1 \land x_b = 1 \\
u_2 & x_a = 1 \land x_b = 2 \\
u_3 & x_a = 1 \land x_b = 3 \\
u_4 & x_a = 1 \land x_b = 4 \\
u_5 & \text{false} \\
\hline
\end{array}
\]

(b) Table $T$

Fig. 21. A DERA to be learned and an observation table

$sp(u_g^i(a_1, g_{i1}) \ldots (a_j, g_{ij}))$ and both

$sp(v^i_g) \land g_1 \equiv sp(u^i_g) \land g^i_1$, and

$sp(v^i_g(a_1, g_{i1}) \ldots (a_j, g_{ij})) \land g_{j+1} \equiv sp(u^i_g(a_1, g_{i1}) \ldots (a_j, g_{ij})) \land g_{i(j+1)}$.

Then the set $\{(sp(u^i_g), (a_1, g_{i1}) \ldots (a_m, g_{im}))| 1 \leq i \leq n\}$ is unifiable. Let $\{a_1 \ldots a_k \in V\}$ We can use the same argument to show that for every $a \in \Sigma$, the set

$\{(sp(u^i_g(a, g_{i})), (a', g_{i}^{\prime}) \ldots (a_k, g_{i}^{\prime k}))| 1 \leq i \leq n, u^i_g(a, g_{i}) \in U \cup U\}$

is unifiable. Since $a_1 \ldots a_n$ and $a'_1 \ldots a'_k$ were chosen arbitrarily, we can conclude that there is merging $\Pi$ of $T$ such that $u^i_g, \ldots, u^n_g$ are in the same block. Thus a coarsest merging of $T$ can contain at most $|\mathcal{A}|$ blocks.

It follows that there can be at most $|\mathcal{A}|$ rows in $U$ such that no pair of them is unifiable. Then the number of calls to the function move row in Algorithm 8 is bounded, and hence a closed table can be constructed. If for every $u, u' \in U$, $u$ and $u'$ are not unifiable then the table is also consistent. Thus we need a bounded number of calls to the function add column to make the timed observation table consistent.

Since the number of blocks in a coarsest merging is bounded by $|\mathcal{A}|$, and the number of automata of the same size is bounded, then the algorithm $TL^*_nsg$ terminates and constructs an automaton $\mathcal{A}'$ with $|\mathcal{A}|$ locations. \(\square\)

Roughly, $TL^*_nsg$ can be understood as $TL^*_sg$ plus merging. Therefore, in the worst case, more steps and therefore queries are needed as in $TL^*_sg$. However, when a small non-sharply guarded EDERA represents a large sharply guarded EDERA, $TL^*_nsg$ will terminate using less queries. Therefore, a better performance can be expected in practice.

Example 25 Let us the algorithm $TL^*_nsg$ on a small example. Let the automaton $\mathcal{A}_1$ shown in Figure 21(a) be the EDERA to learn. After a number
Fig. 22. Automaton $A_2$

of queries of the algorithm $TL_{nsg}^*$, we obtain the observation table $T$ shown in Figure 21(b), where the guarded words $u_1 - u_5$ are defined by

$$
\begin{align*}
    u_1 &= \lambda \\
    u_2 &= (a, x_a = 1 \land x_b = 1) \\
    u_3 &= (a, x_a = 1 \land x_b = 1)(a, x_a = 1 \land x_b = 2) \\
    u_4 &= (a, x_a = 1 \land x_b = 1)(a, x_a = 1 \land x_b = 2)(a, x_a = 1 \land x_b = 3) \\
    u_5 &= u_4(a, x_a = 1 \land x_b = 4)
\end{align*}
$$

It turns out that all rows of $T$ are unifiable. Define $PG$ by

$$
PG = \{ \langle sp(u_1), (a, x_a = 1 \land x_b = 1) \rangle, \\
\langle sp(u_2), (a, x_a = 1 \land x_b = 2) \rangle, \\
\langle sp(u_3), (a, x_a = 1 \land x_b = 3) \rangle, \\
\langle sp(u_4), (a, x_a = 1 \land x_b = 4) \rangle, \\
\langle sp(u_5), (a, false) \rangle \} 
$$

where $false$ represents constraint $x_a \leq 0 \land x_a \geq 5$. It can be checked that the guarded word $(a, x_a = 1 \land x_b \leq 4)$ unifies $PG$. We will use the merging of the observation table $T$ as the partition which consists of exactly one block, and equipping the action $a$ with the guard $x_a = 1 \land x_b \leq 4$. The automaton obtained from this mergings is the automaton $A_1$ which consists of exactly one state. In contrast, the algorithm $TL_{nsg}^*$, which does not employ unification, would construct the sharply guarded EDERA $A_2$ shown in Figure 22. The automaton $A_2$ has 5 states, since table $T$ has 5 different rows. \qed

8 Conclusion

In this paper, we presented techniques for learning timed systems that can be represented as event-recording automata.

By considering the restricted class of event-deterministic automata, we can uniquely represent an automaton by a regular language of guarded words,
and the learning algorithm can identify states by access strings that are un-
timed sequences of actions. This allows us to adapt existing algorithms for
learning regular languages to the timed setting. The main additional work
is to learn the guards under which individual actions will be accepted. The
constructed automaton has the form of a zone graph, which, in general, can
be doubly exponentially larger than a minimal DERA representing the same
language, but for many practical systems the zone graph construction does
not lead to a severe explosion, as exploited by tools for timed automata verifi-
cation [BDM+98,BLL+96]. The resulting algorithm is called $TL^*_{sg}$. The query
complexity of $TL^*_{sg}$ for a sharply guarded EDERA with $n$ locations is

$$O\left(kn^2ml\left(\frac{|\Sigma| + K}{|\Sigma|}\right)\right)$$

where $l$ is the length of the longest guarded word queried and $k$ is the size of
alphabet $\Sigma \times G$. The query complexity of $TL^*_{sg}$ for EDERA with $n$ locations is

$$O\left(kn^22^{2|\Sigma|^{[\alpha]}(2K+2)^{[\beta]}}ml\left(\frac{|\Sigma| + K}{|\Sigma|}\right)\right)$$

$$(16)$$

We also introduced the algorithm $TL^*_{nsg}$ also for learning event-deterministic
automata, which simultaneously reduces the size of the automaton learned so
far, however, for the price of a larger worst-case complexity.

Without the restriction of event-determinism, the problem of learning guards
is significantly less tractable. We present the algorithm $TL^*_s$ that learns general
DERA. The drawback of the algorithm that it constructs a DERA in spirit
of a region graph. The query complexity of $TL^*_s$ for the simple DERA with $n$
locations is $O(|\Sigma|^2n^2mK)$. $$(17)$$ In verification for timed automata, it is a well-
known fact, that despite theoretically lower worst-case complexity, algorithms
based on region graphs perform less efficiently than algorithms based on zone
graphs.

Together with [GJP06], this paper describes initial efforts on learning algo-
rithms for timed systems. Future work has to show the benefit of these al-
gorithms in practical applications as well as to examine the most suitable
application area for each algorithm.

Acknowledgments  This work was supported in part by the European Re-
search Training Network GAMES, and Project 231167 CONNECT.
References


49


