Real-Time System =
Discrete System + Clock Variables*

Rajeev Alur
Thomas A. Henzinger**

TR 94-1403
January 1994

Department of Computer Science
Cornell University
Ithaca, NY 14853-7501

** Partially supported by the National Science Foundation under grant CCR-9200794 and by the United States Air Force Office of Scientific Research under contract F49620-93-1-0056.
Real-time System =
Discrete System + Clock Variables*

Rajeev Alur
AT&T Bell Laboratories
600 Mountain Avenue
Murray Hill, NJ 07974
alur@research.att.com

Thomas A. Henzinger†
Computer Science Department
Cornell University
Ithaca, NY 14853
tah@cs.cornell.edu

*How can we take a programming language off the shelf and upgrade it into a
real-time programming language? Programs such as device drivers and plant
controllers must explicitly refer and react to time. For this purpose, a variety
of language constructs—including delays, timeouts, and watchdogs—has been
put forward. We advocate an alternative answer, namely, to designate certain
program variables as clock variables. The value of a clock variable changes as
time advances. Timing constraints can be expressed, then, by conditions on
clock values. A single new language construct—the guarded wait statement—
suffices to enforce the timely progress of a program.

Our presentation proceeds in two steps. First we extend untimed systems
(Section 1) with clock variables (Section 2); then we introduce the guarded wait
statement (Section 3). The usage of clock variables and the guarded wait state-
ment is illustrated with real-time applications such as round-robin (timeout-
driven) scheduling, priority (interrupt-driven) scheduling, and embedded pro-
cess control (Section 4). Indeed, clock variables generalize naturally to variables
that measure environment parameters other than time (Section 5). In keeping
with an expository style, all references are clustered in bibliographic remarks
at the end of each section. We conclude by pointing to selected literature on
formal methods and support tools for our approach to real-time programming
(Section 6).

†A preliminary version of this paper appeared in the Proceedings of the First AMAST
Workshop on Real-time Systems, Iowa City, IA, 1993.

‡Partially supported by the National Science Foundation under grant CCR-9200794 and
by the United States Air Force Office of Scientific Research under contract F49620-93-1-0056.
1 Untimed Systems

1.1 States

At any time, the state of a system is determined by the values of a finite set of variables. For example, the state of a chess game is determined by the board positions of all pieces and a bit that indicates whether white or black gets to move next; the state of a circuit is determined by the boolean values of all wires and flip-flops; the state of a while program is determined by the value of the program counter and the values of all program variables (the state of a C program, by contrast, is determined by the values of all memory locations that contain the run-time environment of the program).

We use $V$ as the finite set of variables. We assume that all variables in $V$ are typed and use suggestive names like $m$ and $n$ for nonnegative integer variables and $x$ and $y$ for real-valued variables. A state is a function that maps every variable in $V$ to a value of the corresponding type.

A state predicate is a proposition that is either true or false for each state. For example, the proposition “The black king is attacked by the white queen” is either true or false for each state of a chess game; the proposition “All output wires are low” is either true or false for each state of a circuit; the proposition “$m = n + 2$” is either true or false for each state of a while program.

Formally, a state predicate is a formula of predicate logic whose free variables are taken from $V$. The truth value of the state predicate $\phi$ for the state $\sigma$ is obtained by interpreting each free variable $v$ in $\phi$ as $\sigma(v)$. For example, the state predicate $m = n + 2$ is true for a state $\sigma$ iff $\sigma(m) = \sigma(n) + 2$. Every state predicate $\phi$ defines, then, the set of states, denoted by $[\phi]$, for which $\phi$ is true.

1.2 Actions

The state of a system changes over time. We refer to the state changes of a system as actions. An action is a pair $(\sigma, \sigma')$ of states that consists of a source state $\sigma$ and a target state $\sigma'$. Intuitively, if a system is in the source state $\sigma$, then the action $(\sigma, \sigma')$ takes the system into the target state $\sigma'$. We say that an action is enabled in its source state and disabled in all other states. Two actions $(\sigma_1, \sigma'_1)$ and $(\sigma_2, \sigma'_2)$ are consecutive if the second action is enabled in the target state of the first action—i.e., if $\sigma'_1 = \sigma_2$. The action $(\sigma, \sigma')$ is a null action if $\sigma = \sigma'$.

Null actions arise in two situations. First, every state provides a system description that is necessarily abstract—i.e., given at a certain level of detail. Consequently, any action that modifies only detail that is not included in a state appears as a null action. Recall, for instance, that we represent the state of a while program by values for the program counter and the program variables. So the action “Load the value of $m$ into the accumulator,” which, at a level of greater detail, may be part of the execution of a while program, appears as
a null action, because it changes neither the value of the program counter nor the value of any program variable. Second, every system is embedded in an environment. Any environment action that does not modify the system state appears also as a null action.

An action predicate is a proposition that is either true or false for each action. For example, the proposition “A black piece is captured” is either true or false for each move of a chess game; the proposition “Some output wire goes from low to high” is either true or false for each clock cycle of a synchronous circuit; the proposition “The value of \( m \) is incremented” is either true or false for each execution step of a while program.

When describing an action, we use variables like \( v \in V \) to refer to the value of \( v \) in the source state of the action, and primed variables like \( v' \notin V \) to refer to the value of \( v \) in the target state of the action. The action predicate “The value of \( m \) is incremented,” for instance, will be written as \( \lnot m' = m + 1 \); that is, the new value of \( m \) is one greater than the old value of \( m \).

Formally, let \( V' \) be the set of primed variables whose unprimed versions occur in \( V \). An action predicate is a formula of predicate logic whose free variables are taken from \( V \) and \( V' \). The truth value of the action predicate \( \psi \) for the action \((\sigma, \sigma')\) is obtained by interpreting each free unprimed variable \( v \) in \( \psi \) as \( \sigma(v) \) and each free primed variable \( v' \) in \( \psi \) as \( \sigma'(v) \). For example, the action predicate \( m' = m + 1 \) is true for an action \((\sigma, \sigma')\) iff \( \sigma'(m) = \sigma(m) + 1 \). Every action predicate \( \psi \) defines, then, the set of actions, denoted by \( \llbracket \psi \rrbracket \), for which \( \psi \) is true.

1.3 Behaviors

We restrict ourselves to systems that are nondeterministic, discrete, and closed. While the actions of a probabilistic system occur with certain likelihoods, each action of a nondeterministic system is, in every state, either possible (enabled) or impossible (disabled). While a continuous system—such as the solar system—may perform infinitely many actions in a finite interval of time, the behavior of a discrete system results from a countable sequence of actions. While the behavior of an open system may depend on the environment of the system, every environment action of a closed system appears as a null action. We assume that the environment never terminates; the actions of a system that terminates after a finite number of actions are, therefore, followed by infinitely many null actions.

A behavior of a system, then, is a countably infinite sequence of states. Alternatively, a behavior \( \bar{\sigma} = \sigma_0, \sigma_1, \sigma_2, \sigma_3, \ldots \) can be viewed as the countably infinite sequence of actions \((\sigma_i, \sigma_{i+1})\), integer \( i \geq 0 \), with the property that any two neighboring actions are consecutive. For example, a legal sequence of moves, possibly interspersed with null actions, constitutes a behavior of the game of chess; a sequence of consecutive execution steps, possibly interspersed with null actions, constitutes a behavior of a given while program.

We use state and action predicates to describe behaviors. A state predicate
restricts the first state of a behavior: the state predicate $\phi$ is \textit{initially true} for the behavior $\bar{\sigma}$ if $\phi$ is true for the first state of $\bar{\sigma}$. An action predicate restricts all actions of a behavior except for environment actions: the action predicate $\psi$ is \textit{invariantly true} for the behavior $\bar{\sigma}$ if $\psi$ is true for all actions of $\bar{\sigma}$ that are not null actions. Consider, for instance, the behavior $\bar{\sigma}$ whose odd actions, starting from 0, increment the value of $m$ and whose even actions are null actions:

$$\bar{\sigma}_0(m) = 0, 1, 1, 2, 2, 3, 3, \ldots$$

The state predicate $m = 0$ is initially true for $\bar{\sigma}$ and the action predicate $m' = m + 1$ is invariantly true for $\bar{\sigma}$.

1.4 Systems

A system defines a set of possible behaviors. For example, the game of chess defines the initial board configuration and the set of all legal move sequences; a while program defines a set of execution-step sequences, one sequence for each combination of input values.

Formally, a system $S = (\phi_0, \psi_0)$ is a pair that consists of a state predicate $\phi_0$—the initial condition of $S$—and an action predicate $\psi_0$—the transition condition of $S$. The behavior $\bar{\sigma}$ is a behavior of the system $S$ if (1) the initial condition $\phi_0$ is initially true for $\bar{\sigma}$ and (2) the transition condition $\psi_0$ is invariantly true for $\bar{\sigma}$. Every system $S$ defines, then, the set of its behaviors, which is denoted by $[S]$.

Consider, for instance, the system $S_0$ that, starting from 0, increments the value of $m$ arbitrarily often. The system $S_0$ has the initial condition $m = 0$ and the transition condition $m' = m + 1$. Here are some behaviors of $S_0$:

$$\bar{\sigma}_1(m) = 0, 1, 2, 3, 4, 5, 6, \ldots$$
$$\bar{\sigma}_2(m) = 0, 1, 2, 2, 3, 3, 4, \ldots$$
$$\bar{\sigma}_3(m) = 0, 0, 0, 1, 2, 2, 3, 4, \ldots$$
$$\bar{\sigma}_4(m) = 0, 1, 2, 2, 2, 2, 2, 2, 2, \ldots$$

Not every set of behaviors can be defined by a system. First, a system cannot prevent environment actions: if, between any two actions, we add a null action to a behavior of the system $S$, then we obtain another behavior of $S$ (this property of systems is called \textit{stutter closure}). Second, the future evolution of a system is determined completely by the current state of the system and is independent of the past history of the system: if two behaviors $\bar{\sigma}_1$ and $\bar{\sigma}_2$ of the system $S$ share a state $\sigma$, then by composing the past of $\sigma$ according to $\bar{\sigma}_1$ and the future of $\sigma$ according to $\bar{\sigma}_2$ we obtain another behavior of $S$ (\textit{fusion closure}). Third, the operation of a system is characterized completely by local conditions on individual states and actions and there are no global conditions on the behaviors of a system: if all finite prefixes of a behavior $\bar{\sigma}$ can be extended
to behaviors of the system $S$, then $\bar{\sigma}$ itself is a behavior of $S$ (limit closure).\footnote{Conversely, if we admit infinitary formulas as state and action predicates, then every set of behaviors that is closed under stuttering, fusion, and limits is definable by a system.}

The three closure properties of systems have two important ramifications on the executability of systems. First, the behaviors of a system can be generated action by action: start with a state for which the initial condition is true and repeatedly choose either a null action or an enabled action for which the transition condition is true. Second, a system need not progress: when generating a behavior action by action, from any point on we may choose null actions only.

### 1.5 Programs

We use a guarded-command programming language to define systems. A \textit{guarded-command program} $P$ consists of a set of assignments, which defines the initial condition of $P$, followed by a set of guarded commands, which defines the transition condition of $P$. The program $P$ is executed in a stepwise fashion: first, starting from any state, execute the initial assignments to obtain an initial state of $P$; then continue to select, nondeterministically, either a null action or a guarded command whose guard is true. We write $\llbracket P \rrbracket$ for the set of execution sequences of the guarded-command program $P$.

Consider, for instance, the guarded-command program $\textit{NdUpDown}$ of Figure 1. The program $\textit{NdUpDown}$ starts from an initial state $\sigma$ with $\sigma(m) = 0$ and $\sigma(n) = 1$. The value of $n$ is incremented arbitrarily often before the program may change the value of $m$ to 1. Thereafter, the value of $n$ may be decremented repeatedly as long as it remains nonnegative.

The guarded-command program $\textit{NdUpDown}$ defines a system whose behaviors are the execution sequences of $\textit{NdUpDown}$. The initial condition of $\textit{NdUpDown}$ is given by the state predicate $\phi_0$ and the transition condition of

\begin{verbatim}
program NdUpDown:
  initially m := 0; n := 1;
  loop true → m := 1
    or m = 0 → n := n + 1
    or m = 1 ∧ n > 0 → n := n - 1
  end.
\end{verbatim}

Figure 1: Guarded-command program
$NdUpDown$ is given by the action predicate $\psi_o$:

\[
\phi_0 &= (m = 0 \land n = 1) \\
\psi_o &= (m' = 1)_m \lor \\
& (m = 0 \land n' = n + 1)_m \lor \\
& (m = 1 \land n > 0 \land n' = n - 1)_m
\]

The expression $(\psi)_F$, for an action predicate $\psi$ and a vector $\vec{v}$ of variables, is an abbreviation for the action predicate $\psi \land \vec{v}' = \vec{v}$; that is, the variables listed as a subscript remain unchanged. Each guarded command of the program $NdUpDown$ contributes a disjunct—called an action schema—to the transition condition. We leave it to the reader to translate an arbitrary guarded-command program $\mathcal{P}$ into an initial condition and a transition condition that generate the set $[\mathcal{P}]$ of behaviors.

**Bibliographic remarks.** Our system model is fairly standard, provided one accepts the interleaving view of concurrent activities (see, for example, the TLA approach of Lamport [Lam91] and the transition system approach of Manna and Pnueli [MP92]). As a word of caution, we point out that our actions are pairs of states rather than, as in TLA, sets of pairs of states. Our choice leads to a pleasant parallelism between state predicates and action predicates: a state predicate defines a set of states; an action predicate defines a set of actions. Stutter closure has been advocated by Lamport [Lam83]; limit closure (safety) was formally defined by Alpern, Demers, and Schneider [ADS86]; the relationship between systems and closure properties was studied by Emerson [Eme83]; the executability (machine closure) of systems, by Apt, Francez, and Katz [AFK88]. Guarded commands were introduced by Dijkstra [Dij75]; our guarded-command language is inspired by the UNITY language of Chandy and Misra [CM88].

2 Timed Systems 1: Safety

2.1 Clocks and delays

The state of a timed system is determined by two kinds of variables—system variables and clock variables. The value of each system variable is changed explicitly by the transition condition of the system; the value of each clock variable increases implicitly as time advances. Formally, we partition the finite set $V$ of variables into system variables and clock variables. The clock variables—or clocks—range over the nonnegative real numbers. We reserve the letters $x$, $y$, and $z$ for clocks. Given a state $\sigma$ and a nonnegative real number $\delta$, we write $\sigma + \delta$ for the state that maps all system variables $m$ to $\sigma(m)$ and all clocks $x$ to $\sigma(x) + \delta$.

The behavior of a timed system results from two kinds of actions—system actions and time actions. System actions update the values of the system variables and reset some of the clocks to 0; time actions do not modify the value
of any system variables and uniformly advance the values of all clocks. It is perhaps best to think of time actions as environment actions: the environment can advance time and, in doing so, increases the values of the clocks by the same amount; the system, on the other hand, has only one way of changing the value of a clock, namely, to restart it. In other words, timed systems are open systems whose environment actions appear as both time actions and null actions.

Formally, the action \((\sigma, \sigma')\) is a system action if for all clock variables \(x\), either \(\sigma'(x) = \sigma(x)\) or \(\sigma'(x) = 0\); the action \((\sigma, \sigma')\) is a time action—or delay—if there is a nonnegative real \(\delta\)—the duration of the delay—such that \(\sigma' = \sigma + \delta\). System actions have duration 0. Every null action is, by definition, both a system action and a delay of duration 0.

We henceforth write \((\sigma, \delta)\) for a delay with source state \(\sigma\) and duration \(\delta\). A state \(\sigma'\) is passed by the delay \((\sigma, \delta)\) if \(\sigma' = \sigma + \epsilon\) for some nonnegative real \(\epsilon < \delta\).

In particular, no state is passed by a null action, and while the source state \(\sigma\) is passed by a delay \((\sigma, \delta)\) with positive duration \(\delta > 0\), the target state \(\sigma + \delta\) is not passed by the delay \((\sigma, \delta)\).

A timed behavior of a system, then, is a countably infinite sequence of system actions and delays such that any two neighboring actions are consecutive. Given a timed behavior \(\tilde{\sigma}\), let \(\delta_i\) be the duration of the \(i\)-th action of \(\tilde{\sigma}\). The timed behavior \(\tilde{\sigma}\) diverges if the infinite sum \(\sum_{i \geq 0} \delta_i\) of durations diverges. A state \(\sigma\) occurs in the timed behavior \(\tilde{\sigma}\) if either \(\sigma\) is the source state of a system action of \(\tilde{\sigma}\), or \(\sigma\) is passed by a delay of \(\tilde{\sigma}\).

### 2.2 Clock constraints

State sets and action sets of timed systems are, as before, defined by formulas of predicate logic with free variables from \(V\) and \(V'\). We restrict, however, the way in which clocks may appear in such formulas to rule out timed systems that within a finite interval of time make infinitely many decisions on how to proceed. For example, we do not wish to consider system actions that are enabled iff the value of a certain clock is rational.

Let \((\sigma, \delta)\) be a delay, let \(\phi\) be a state predicate, and let \(\psi\) be an action predicate. The characteristic function of \(\phi\) maps each nonnegative real \(\epsilon < \delta\) to 1 if \(\phi\) is true for \(\sigma + \epsilon\), and otherwise to 0; the characteristic function of \(\psi\) maps \(\epsilon\) to 1 iff \(\psi\) is enabled in \(\sigma + \epsilon\). A state or action predicate varies finitely over the delay \((\sigma, \delta)\) if its characteristic function has finitely many discontinuities in the interval \((0, \delta]\). Abstractly, we restrict ourselves to state predicates and action predicates that vary finitely over all delays.

Concrete, we offer a syntax for the timed systems that base their decisions on comparisons of clock values with each other and with integer constants: a clock constraint is a formula of the form \(x \leq d, e \leq y, \text{ or } x + c \leq y + d\), where \(x\) and \(y\) are clocks, and \(c\) and \(d\) are nonnegative integers. Moreover, every system action leaves the value of a clock either unchanged or reset to 0: a reset condition is a formula of the form \(x' = x\) or \(x' = 0\), for a clock \(x\).
A timed state predicate, then, is a formula of predicate logic whose free variables are taken from $V$ such that all clocks occur in clock constraints only; a timed action predicate is a formula of predicate logic whose free variables are taken from $V$ and $V'$ such that all clocks and primed clocks occur in clock constraints and reset conditions only (in particular, all primed clocks of a timed action predicate occur in reset conditions only). While, for example, the formula $y \leq x < y+3 \land z = x+1$ is a timed state predicate, the formulas $x+y = 3$ and $2x \leq y$ are not. It is not difficult to check that every timed state predicate and every timed action predicate varies finitely over all delays.  

The timed action predicate $\psi$ is invariantly true for the timed behavior $\bar{\sigma}$ if $\psi$ is true for all actions of $\bar{\sigma}$ that are not delays. Consider, for instance, the timed behavior $\bar{\sigma}$ whose even actions increment the value of the system variable $m$ and restart the clock $x$, and whose odd actions are delays of duration 2:

$$\bar{\sigma}(m, x) = (0, 0), (0, 2), (1, 0), (1, 2), (2, 0), (2, 2), (3, 0), (3, 2), (4, 0), \ldots$$

The timed action predicate $x = 2 \land m' = m + 1 \land x' = 0$ are invariantly true for $\bar{\sigma}$.

### 2.3 Clock-constrained systems

A clock-constrained system $S = (\phi_0, \psi_0)$ is a pair that consists of a timed state predicate $\phi_0$—the initial condition of $S$—and a timed action predicate $\psi_0$—the transition condition of $S$. The timed behavior $\bar{\sigma}$ is a behavior of the clock-constrained system $S$ if (1) the initial condition of $S$ is initially true for $\bar{\sigma}$ and (2) the transition condition of $S$ is invariantly true for $\bar{\sigma}$. Every clock-constrained system $S$ defines, then, the set of its divergent behaviors, which is denoted by $[S]$.

Consider, for instance, the clock-constrained system $S_1$ that, starting from 0, increments the value of $m$ arbitrarily often such that the time difference between consecutive increment operations is at least 1 time unit and at most 2 time units. The system $S_1$ has the initial condition $m = 0 \land x = 0$ and the transition condition $1 \leq x \leq 2 \land m' = m + 1 \land x' = 0$. Here are some divergent behaviors of $S_1$:

$$\bar{\sigma}_1(m, x) = (0, 0), (0, 1), (1, 0), (1, 1), (2, 0), (2, 1), (3, 0), \ldots$$

$$\bar{\sigma}_2(m, x) = (0, 0), (0, 2), (1, 0), (1, 2), (2, 0), (2, 2), (3, 0), \ldots$$

$$\bar{\sigma}_3(m, x) = (0, 0), (0, 0.5), (0, 1.2), (1, 0), (1, 0), (1, 1.5), (2, 0), \ldots$$

$$\bar{\sigma}_4(m, x) = (0, 0), (0, 1), (0, 1.5), (0, 2), (0, 19), \ldots$$

---

2This is the case also for timed state and action predicates that are infinitary formulas, because the set $V$ of variables is finite. Conversely, there is an infinitary formula with addition of clocks that does not vary finitely over all delays ($x = 1 \lor 2x = 1 \lor 3x = 1 \lor \ldots$). Decidability considerations, too, suggest the restriction of clock constraints to successor and comparison operations on clocks.
Not every set of divergent timed behaviors can be defined by a clock-constrained system. First, the options of a clock-constrained system change only finitely often within a finite interval of time: the transition condition $\psi_b$ of $S$ varies finitely over all delays of behaviors of $S$ (finitely varying branching). Second, if we split a delay of a behavior of $S$ into two consecutive delays, or if we merge two consecutive delays of a behavior of $S$ into a single delay, then we obtain another behavior of $S$ (closure under timed stuttering; as a delay may be split into itself followed by a null action, closure under timed stuttering implies stutter closure). Third, the behaviors of $S$ are closed under fusion and, fourth, under divergent limits: if all finite prefixes of a divergent timed behavior $\tilde{\sigma}$ can be extended to behaviors of $S$, then $\tilde{\sigma}$ itself is a behavior of $S$ (as the infinite sequence of null actions is not a behavior of $S$, no clock-constrained system is closed under limits). Fifth, a clock-constrained system cannot prevent delays: if a finite prefix of a behavior of $S$ is extended by a delay, then we obtain again a prefix of a behavior of $S$ (closure under waiting).

Consequently, the divergent behaviors of a clock-constrained system can be generated action by action: (1) start with a state for which the initial condition is true and repeatedly choose either a delay or an enabled action for which the transition condition is true; (2) to generate divergent timed behaviors, make sure to choose appropriate delays—i.e., choose a sequence of delays whose sum diverges. Furthermore, a clock-constrained system need not progress: when generating a divergent behavior action by action, from any point on we may choose delays only.

2.4 Clock-constrained programs

A clock-constrained program $P$ consists of a set of assignments followed by a set of guarded commands such that (1) all assignments to clocks are of the form $x := 0$ and (2) all guards are timed state predicates. The program $P$ is executed in a stepwise fashion: first, starting from any state, execute the initial assignments to obtain an initial state of $P$; then continue to select, nondeterministically, either a delay or a guarded command whose guard is true. We write $[P]$ for the set of divergent execution sequences of the clock-constrained program $P$.

Consider, for instance, the clock-constrained program $CcUpDown$ of Figure 2. The program $CcUpDown$ contains a system variable, $n$, and two clocks, $x$ and $y$. The program starts from an initial state in which $n$ has the value 1 and both $x$ and $y$ have the value 0. Within the first 10 time units, the value of $n$ is incremented arbitrarily often as long as the time difference between consecutive increment operations is at least 1 time unit and at most 5 time units. Beginning at time 12, the value of $n$ may be decremented repeatedly, under the same timing constraints, as long as $n$ remains nonnegative.

The clock-constrained program $CcUpDown$ defines a clock-constrained system whose behaviors are the divergent execution sequences of $CcUpDown$. The initial condition of $CcUpDown$ is given by the timed state predicate $\phi_0$ and the
program CcUpDown:
declare x, y: clock;
initially n := 1; x := 0; y := 0;
loop
  x ≤ 10 ∧ 1 ≤ y ≤ 5 → n := n + 1; y := 0
  or x ≥ 12 ∧ 1 ≤ y ≤ 5 ∧ n > 0 → n := n - 1; y := 0
end.

Figure 2: Clock-constrained program

transition condition of CcUpDown is given by the timed action predicate ψₘ:

ϕ₀ = (n = 1 ∧ x = 0 ∧ y = 0)
ψₘ = (x ≤ 10 ∧ 1 ≤ y ≤ 5 ∧ n' = n + 1 ∧ y' = 0)ₓ ∨
     (x ≥ 12 ∧ 1 ≤ y ≤ 5 ∧ n > 0 ∧ n' = n - 1 ∧ y' = 0)ₓ

Each guarded command of the program CcUpDown contributes a disjunct—called a system action schema—to the transition condition. We leave it to the reader to translate an arbitrary clock-constrained program P into an initial condition and a transition condition that generate the set [P] of timed behaviors.

Bibliographic remarks. Clock variables—as we use them—were first introduced in temporal logic [AH89] (originally as so-called “freeze” quantifiers) and in finite automata [AD90]. Both papers realized the importance of the choice of clock constraints that are admitted by a system; this choice has immediate ramifications on semantic issues and verification problems. The dichotomy of system actions versus time actions, which had been advocated repeatedly by various proponents of the interleaving view of concurrency, was perhaps standardized by the real-time system model of timed transition systems [HMP91]. In particular, the timed transition system model has led to the study of closure under timed stuttering, divergent limits (divergence safety), and finitely varying branching [Hen91, Hen92, HNSY92].

3 Timed Systems 2: Progress

3.1 Delay predicates

Neither untimed systems nor clock-constrained systems need to progress—i.e., at any point a system may fail to perform another action while the environment keeps performing actions. To ensure the progress of an untimed system, one typically assumes fairness requirements on the system actions: a system action must not be enabled forever (or infinitely often) without being taken. To ensure the progress of a timed system, it will suffice to assume a fairness requirement
on the time actions: time must advance beyond any bound. While we have seen
that a restriction to divergent timed behaviors does not ensure progress per se,
we will use the divergence of time to enforce system actions by constraining
time actions. For this purpose, we introduce delay predicates. A delay predi-
cate typically enforces a system action by limiting the amount by which the
environment may advance time in a given state.

A delay predicate is a proposition that is either true or false for each de-
lay. For example, the proposition “Throughout the delay, the clock \( x \) shows
at most 5” is either true or false for a delay; we will write this proposition
as “\( x < 5 \),” using variables like \( x \) to refer to the value of \( x \) in all states that
are passed by the delay. The delay predicate \( x < 5 \), then, prevents time from
advancing past a state in which the value of the clock \( x \) is 5.

Formally, the syntax of delay predicates is identical to the syntax of timed
state predicates. The delay predicate \( \chi \) is true for a delay if \( \chi \), viewed as a
timed state predicate, is true for all states that are passed by the delay. For
example, the delay predicate \( x < 5 \) is true for a delay \( (\sigma, \delta) \) iff \( \sigma(x) + \delta \leq 5 \)
(since, by definition, the target state of a delay is not passed by the delay, the
delay predicates \( x < 5 \) and \( x \leq 5 \) are true for the same delays). Every delay
predicate \( \chi \) defines, then, the set \( [\chi] \) of delays for which \( \chi \) is true. In particular,
the set \( [\chi] \) contains all null actions and is closed under both the splitting and
the merging of delays; that is, a delay predicate \( \chi \) is true for two consecutive
delays \( (\sigma, \delta_1) \) and \( (\sigma + \delta_1, \delta_2) \) iff \( \chi \) is true for the combined delay \( (\sigma, \delta_1 + \delta_2) \).

While a timed action predicate restricts the system actions of a timed be-
havior, a delay predicate restricts the delays of a timed behavior. The delay
predicate \( \chi \) is invariantly true for the timed behavior \( \bar{\sigma} \) if \( \chi \) is true for all delays
of \( \bar{\sigma} \). Consider again the timed behavior \( \bar{\sigma} \) whose even actions increment the
value of the system variable \( m \) and restart the clock \( x \), and whose odd actions
are delays of duration 2:

\[
\bar{\sigma}(m, x) = (0, 0), (0, 2), (1, 0), (1, 2), (2, 0), (2, 2), (3, 0), (3, 2), (4, 0), \ldots
\]

The delay predicate \( x < 2 \) is invariantly true for \( \bar{\sigma} \); the delay predicate \( x \neq 1 \) is
not.

3.2 Real-time systems

A real-time system \( \mathcal{S} = (\phi_0, \psi, \chi_{\mathcal{E}}) \) is a triple that consists of a clock-con-
strained system \( (\phi_0, \psi) \) and a delay predicate \( \chi_{\mathcal{E}} \)—the environment condition
of \( \mathcal{S} \). The timed behavior \( \bar{\sigma} \) is a behavior of the real-time system \( \mathcal{S} \) if (1) \( \bar{\sigma} \) is a
behavior of the clock-constrained system that underlies \( \mathcal{S} \) and (2) the envi-
ronment condition of \( \mathcal{S} \) is invariantly true for \( \bar{\sigma} \). Every real-time system \( \mathcal{S} \) defines,
then, the set of its divergent behaviors, which is denoted by \( \llbracket \mathcal{S} \rrbracket \).

The transition condition of a real-time system asserts necessary conditions
on system actions and the environment condition asserts sufficient conditions on
system actions. For example, the following real-time system \( S_2 = (\phi_0, \psi_0, \chi_0) \) changes the value of \( m \) from 0 to 1 at time 3 at the earliest and at time 5 at the latest:

\[
\begin{align*}
\phi_0 &= (m = 0 \land x = 0) \\
\psi_0 &= (x \geq 3 \land m' = 1) \\
\chi_0 &= (m = 0 \land x < 5) \lor (m = 1)
\end{align*}
\]

While the requirement \( x \geq 3 \) of the transition condition ensures the lower time bound of 3—the system transition that changes the value of \( m \) may happen at or after time 3—the requirement \( x < 5 \) of the environment condition ensures the upper time bound 5—the transition must happen at or before time 5, because time cannot advance past 5 before the value of \( m \) has changed to 1.

Not every set of divergent timed behaviors can be defined by a real-time system. While all real-time systems exhibit finitely varying branching and are closed under timed stuttering, fusion, and divergent limits, they are not necessarily closed under waiting. This is because the environment condition can enforce the progress of a real-time system.

### 3.3 Real-time executability

There is a price to be paid, however, for progress: the divergent behaviors of a real-time system may no longer be generated action by action. By starting with a state for which the initial condition is true and repeatedly choosing either a delay that does not violate the environment condition or an enabled action for which the transition condition is true, we may be led into a situation from which time cannot diverge.

Consider, for example, the following variant \( S_2' \) of the real-time system \( S_2 \):

\[
\begin{align*}
\phi_0 &= (m = 0 \land x = 0) \\
\psi_0 &= (3 \leq x \leq 5 \land m' = 1) \\
\chi_0 &= (m = 0 \land x < 10) \lor (m = 1)
\end{align*}
\]

The real-time systems \( S_2 \) and \( S_2' \) have the same divergent behaviors—formally, \([S_2] = [S_2']\)—namely, all divergent timed behaviors in which the value of \( m \) changes from 0 to 1 at time 3 at the earliest and at time 5 at the latest. Only the real-time system \( S_2 \), however, can be executed in a stepwise fashion. Executing the real-time system \( S_2' \) we might start from the initial state by choosing a delay of duration, say, 8, which does not violate the environment condition. But there is no divergent continuation from the resulting state \( \sigma \) with \( \sigma(m) = 0 \) and \( \sigma(x) = 8 \): we have painted ourselves into a corner by having chosen an initial delay whose duration was greater than 5.

Real-time systems that do not exhibit this problem are called nonzeno: a real-timed system \( S \) is nonzeno if every finite prefix of a behavior of \( S \) is a
finite prefix of a divergent behavior of $S$. In particular, while the real-time system $S_2$ is nonzero, the real-time system $S_2$ is not. Nonzero-neness, then, is a property of a real-time system (syntax) rather than a property of a set of timed behaviors (semantics). Indeed, every set of timed behaviors that can be defined by a real-time system can also be defined by a nonzero real-time system, whose environment condition is suitably strengthened. For instance, the real-time system $S_2$ can be transformed into an equivalent nonzero real-time system by strengthening the environment condition to $(x < 5) \lor (m = 1)$.

3.4 Real-time programs

A guarded wait statement is an instruction of the form $\chi \rightarrow \text{wait}$, where the guard $\chi$ is a delay predicate. A guarded wait statement delays the program execution by an arbitrary amount of time, but at most until the guard, viewed as a timed state predicate, becomes false. The guarded wait statement $\chi \rightarrow \text{wait}$ results, then, in a delay that satisfies the delay predicate $\chi$. In particular, the guarded wait statement $\text{false} \rightarrow \text{wait}$ results in a null action, and the guarded wait statement $\text{true} \rightarrow \text{wait}$ results in a delay of arbitrary duration. We generally prefer the guard $x < 5$ over the equivalent delay predicate $x \leq 5$, perhaps because we like to think of the guarded wait statement $x < 5 \rightarrow \text{wait}$ as delaying execution as long as the value of $x$ is less than 5, and terminating as soon as the value of $x$ becomes 5.

A real-time program $P$ is a clock-constrained program that contains guarded wait statements in addition to guarded commands. The program $P$ is executed in a stepwise fashion: first, starting from any state, execute the initial assignments to obtain an initial state of $P$; then, continue to select, nondeterministically, either a guarded wait statement or a guarded command whose guard is true. Since it may not be possible to merge the consecutive execution of two different guarded wait statements into a single execution step, the resulting set $[P]$ of divergent execution sequences is not necessarily closed under timed stuttering. The set $[P]$ of behaviors of the real-time program $P$, therefore, is defined as the timed stutter closure of $[P]$—i.e., the smallest superset of $[P]$ that is closed under timed stuttering.

Consider, for instance, the real-time program $RtUpDown1$ of Figure 3. The program $RtUpDown1$ behaves like the clock-constrained program $CcUpDown$ except that the value of $n$ must be incremented or decremented (depending on how much time has elapsed since the program was started) at least once every 5 time units, until $n$ reaches the value 0.

The real-time program $RtUpDown1$ defines a real-time system that has the same behaviors. The initial condition of $RtUpDown1$ is given by the timed state predicate $\phi_0$, the transition condition of $RtUpDown1$ is given by the timed action predicate $\psi_\rightarrow$, and the environment condition of $RtUpDown1$ is given by
Program `RtUpDown1`:

```plaintext
program RtUpDown1:
    declare x, y: clock;
    initially n := 1; x := 0; y := 0;
    loop
        x <= 10 ∧ y >= 1 → n := n + 1; y := 0
        or n > 0 ∧ x >= 12 ∧ y >= 1 → n := n - 1; y := 0
        or n > 0 ∧ y < 5 → wait
        or n = 0 → wait
    end.
```

Figure 3: Real-time program

The delay predicate $\chi_0$:

- $\phi_0 = (n = 1 ∧ x = 0 ∧ y = 0)$
- $\psi_r = (x <= 10 ∧ y >= 1 ∧ n' = n + 1 ∧ y' = 0)_x ∨ (x >= 12 ∧ y >= 1 ∧ n > 0 ∧ n' = n - 1 ∧ y' = 0)_x$
- $\chi_0 = (n > 0 ∧ y < 5) ∨ (n = 0)$

Each guarded command of the program `RtUpDown1` contributes disjunct (a system action schema) to the transition condition, and each guarded wait statement of `RtUpDown1` contributes a disjunct—called a delay schema—to the environment condition. We leave it to the reader to translate an arbitrary real-time program $P$ into an initial condition, a transition condition, and an environment condition that define the set $[P]$ of timed behaviors. A clock-constrained program, in particular, is a real-time program with the guarded wait statement $true → wait$, which yields the environment condition $true$.

The execution strategy for real-time programs succeeds only if it cannot lead to a state from which there is no divergent continuation of actions. A real-time program $P$, then, is executable if it defines a nonzeno real-time system. The real-time program `RtUpDown1`, for example, is not executable. To see this, observe that the stepwise execution of `RtUpDown1` may lead to a state $\sigma$ with $\sigma(n) = 5$, $\sigma(x) = 11$, and $\sigma(y) = 5$, and there is no divergent sequence of actions continuing from $\sigma$. Figure 4 shows an executable real-time program that has the same behaviors as `RtUpDown1`. The environment condition of the executable program `RtUpDown2` is

- $\chi_0: (x < 10 ∧ y < 5) ∨ (x ≥ y + 7 ∧ y < 5) ∨ (n = 0)$.

The real-time program `RtUpDown2` terminates—i.e., the value of $n$ reaches 0 and execution continues with delays only—within at most 75 time units (to calculate this time bound on the termination of `RtUpDown2`, observe that the value of $n$ can be at most 12 when $x = 11$).
program $RtUpDown2$:
declare x, y: clock;
initially $n := 1; x := 0; y := 0$;
loop $x \leq 10 \land y \geq 1$ or $n > 0 \land x \geq 12 \land y \geq 1$ or $x < 10 \land y < 5$ or $x \geq y + 7 \land y < 5$ or $n = 0$
$\rightarrow n := n + 1; y := 0$ or $n := n - 1; y := 0$ or wait or wait or wait
end.

Figure 4: Executable real-time program

Bibliographic remarks. Nonzenoness (timed executability) was defined independently by several researchers [AL92, Hen92]; Abadi and Lamport coined the phrase. The first proposal for environment conditions as a means of ensuring the progress of a timed system advocated the use of program invariants [HINSY92]; that paper also presents an algorithm for transforming a real-time system into an equivalent nonzeno system. The real-time program $RtUpDown1$ is derived from an example, due to Pnueli [HMP91], that illustrates that an increase in the lower bound on the duration of a delay may lead to a decrease in the running time of a program (to see this, replace the lower bounds of 1 on the clock $y$ by 2; then $RtUpDown2$ will terminate within 50 time units).

4 Real-time Programming

4.1 Sequential real-time processes

Our choice of guarded commands as programming language has largely been instructionally motivated. In general, we may add clocks and guarded wait statements to any programming language. To illustrate this thesis, we now extend a language of while programs with clocks and guarded wait statements. As in synchronous programming languages, all instructions of a while program are executed immediately and instantaneously; all delays are indicated explicitly by guarded wait statements.

It is convenient to introduce abbreviations for several common applications of guarded wait statements. First, the instruction \textbf{wait} delays the program execution for an arbitrary amount of time; it stands for the guarded wait statement $true \rightarrow \textbf{wait}$. Second, the instruction \textbf{wait watching} $\phi$ delays the program execution for an arbitrary amount of time, but at most until the timed state predicate $\phi$ becomes true; it stands for the guarded wait statement $\neg \phi \rightarrow \textbf{wait}$. Third, the instruction \textbf{await} $\phi$ delays the program execu-
The program `RtUpDown3`:
```
program RtUpDown3:
  declare x, y: clock;
  n := 1; x := 0;
  wait [1, 5];
  while x ≤ 10 do
    n := n + 1; y := 0;
    repeat
      if x < y + 7
        then wait watching x ≥ 10 ∨ y ≥ 5
        else wait watching y ≥ 5
      fi
      until y ≥ 1
    od;
  repeat wait watching y ≥ 5 until x ≥ 12;
  while n > 0 do
    n := n - 1;
    wait [1, 5]
  od.
```

Figure 5: Real-time while program

...to run precisely until the timed state predicate $\phi$ becomes true; it stands for the program segment
```
repeat wait watching $\phi$ until $\phi$.
```

Fourth, the instruction `wait [c, d]` delays the program execution at least $c$ and at most $d$ time units, for nonnegative integers $c$ and $d$ with $c \leq d$; it stands for the program segment
```
z := 0; repeat wait watching $z \geq d$ until $z \geq c$,
```
where $z$ is a new clock. We write `wait c` short for `wait [c, c]`.

Now consider the real-time while program `RtUpDown3` of Figure 5. We present a formal semantics for real-time while programs such that the program `RtUpDown3` defines essentially the same real-time system as the guarded-command program `RtUpDown2` of Section 3.4.

We construct the real-time system that is defined by a while program $\mathcal{P}$ by encoding the control structure of $\mathcal{P}$ as a formula. For this purpose, we introduce a new system variable—the `program counter` $\text{pc}$—that ranges over the control locations of $\mathcal{P}$. If $\ell_0$ is the initial control location, then the initial condition of $\mathcal{P}$ is $\text{pc} = \ell_0$. An execution step of $\mathcal{P}$ is a move between control locations and corresponds to the execution of a guarded command (typically, an
assignment or a test) or the execution of a guarded wait statement. Each such execution step contributes a system action schema to the transition condition of \( P \); each guarded wait statement contributes, in addition, a delay schema to the environment condition of \( P \).

Consider, for example, the initial program segment

\[ \ell_0: n := 1; \quad \ell_1: x := 0 \]

of the while program \( RtUpDown3 \); it contributes the system action schema

\[ (pc = \ell_0 \land pc' = \ell_1 \land n' = 1) \]

to the transition condition of \( RtUpDown3 \). The program segment

\[ \ell_2: \textbf{wait watching } x \geq 10 \lor y \geq 5 \quad \ell_3: \textbf{until } y \geq 1 \]

contributes to the transition condition of \( RtUpDown3 \) the system action schema

\[ (pc = \ell_2 \land pc' = \ell_3) \]

and contributes to the environment condition of \( RtUpDown3 \) the delay schema

\[ (pc = \ell_2 \land x < 10 \land y < 5). \]

We leave it to the reader to define the transition condition of \( RtUpDown3 \) as the disjunction of all system action schemata and the environment condition of \( RtUpDown3 \) as the disjunction of all delay schemata. If we adjust the resulting timed behaviors by (1) projecting out the program counter and (2) merging certain consecutive states (for example, the consecutive assignments \( n := n + 1; \quad y := 0 \) of the while program \( RtUpDown3 \) constitute a single action of the guarded-command program \( RtUpDown2 \)), then we obtain precisely the real-time system that is defined by the guarded-command program \( RtUpDown2 \). It follows, in particular, that the real-time while program \( RtUpDown3 \) is executable.

### 4.2 Concurrent real-time processes

So far, we have discussed sequential processes and sequential real-time processes. We now illustrate that concurrent processes, too, accommodate naturally the introduction of timing constraints through clocks and guarded wait statements: while every concurrent while program defines a system, every concurrent real-time program defines a real-time system.

The actions of concurrent untimed processes are interleaved. Consider, for instance, the concurrent while program \( ParUp \) of Figure 6. The program \( ParUp \) consists of two independent processes. The "left" process repeatedly increases the value of the variable \( m \); the "right" process repeatedly increments the value
program ParUp:
  initially m := 0; n := 0;
  cobegin
    loop \ell_0: m := m + 1; \ell_1: m := m + 2 end
    ||
    loop k_0: n := n + 1 end
  coend.

Figure 6: Concurrent while program

of n. A behavior of ParUp interleaves, without regard to any fairness requirements, null actions with actions of either process: an action of the left process may be followed by any number of actions of the right process, and vice versa. Here are some behaviors of ParUp:

\[ \sigma_1(m, n) = (0, 0), (1, 0), (1, 1), (3, 1), (3, 2), (4, 2), (4, 3), \ldots \]
\[ \sigma_2(m, n) = (0, 0), (1, 0), (3, 0), (4, 0), (4, 0), (6, 0), (6, 1), (7, 1), \ldots \]
\[ \sigma_3(m, n) = (0, 0), (0, 1), (0, 2), (0, 3), (1, 3), (1, 4), (1, 5), \ldots \]
\[ \sigma_4(m, n) = (0, 0), (0, 1), (0, 1), (1, 1), (3, 1), (4, 1), (4, 2), \ldots \]

The behavior \( \sigma_1 \), for example, alternates the actions of both processes.

To define the semantics of the program ParUp formally, we use two program counters, left and right—one for each process (since the right process has a single control location, there is actually no need for the program counter right in this example). Each execution step of either process contributes an action schema to the transition condition of ParUp. The concurrent program ParUp defines, then, the system with the initial condition \( \phi_0 \) and the transition condition \( \psi_0 \):

\[ \phi_0 = (left = \ell_0 \land right = k_0 \land m = 0 \land n = 0) \]
\[ \psi_0 = (left = \ell_0 \land left' = \ell_1 \land m' = m + 1)_{right,n} \lor \\
   (left = \ell_1 \land left' = \ell_0 \land m' = m + 2)_{right,n} \lor \\
   (n' = n + 1)_{left, right,m} \]

The interleaving of the actions of concurrent untimed processes reflects an underlying assumption that the execution speeds of the processes are unknown, independent, possibly very different and even varying. Timing constraints relativize the assumption of speed independence by ruling out certain interleavings of the system actions of concurrent timed processes. This is because, while the system actions of concurrent timed processes are interleaved, their delays overlap.

Consider, for instance, the concurrent real-time program RtParUp of Figure 7. The real-time program RtParUp is obtained from the untimed program
program \textit{RtParUp}:
\begin{align*}
\text{initially } m &:= 0; \ n := 0; \\
\text{cobegin} \\
\text{loop wait 3; } m &:= m + 1; \text{ wait 1; } m := m + 2 \text{ end} \\
\| \\
\text{loop wait 7; } n &:= n + 1 \text{ end} \\
\text{coend.}
\end{align*}

Figure 7: Concurrent real-time program

\textit{ParUp} by adding delays between consecutive assignments. During the execution of \textit{RtParUp}, a delay of the left process will overlap with one or more delays of the right process, and vice versa. Also, a system action of the left process may interrupt a delay of the right process, and vice versa.

Figure 8 shows the program \textit{RtParUp} without abbreviations and annotated with control locations. Here are two behaviors of \textit{RtParUp}:
\begin{align*}
\bar{\sigma}_5(m, n, x, y) &= (0, 0, 0, 0), (0, 0, 3, 3), (1, 0, 3, 3), (1, 0, 0, 3), (1, 0, 1, 4), \\
& (3, 0, 1, 4), (3, 0, 0, 4), (3, 0, 3, 7), (4, 0, 3, 7), (4, 1, 3, 7), \ldots \\
\bar{\sigma}_6(m, n, x, y) &= (0, 0, 0, 0), (0, 0, 3, 3), (1, 0, 3, 3), (1, 0, 0, 3), (1, 0, 1, 4), \\
& (3, 0, 1, 4), (3, 0, 0, 4), (3, 0, 3, 7), (3, 1, 3, 7), (4, 1, 3, 7), \ldots
\end{align*}

In particular, the timing constraints of \textit{RtParUp} rule out the strict alternation of system actions from the left process and the right process; that is, no behavior of the real-time program \textit{RtParUp} corresponds to the behavior \bar{\sigma}_1 of the untimed program \textit{ParUp}.

When defining the semantics of \textit{RtParUp}, each execution step of either process contributes a system action schema to the transition condition of \textit{RtParUp}. In addition, each pair of guarded wait statements—one from the left process and the other one from the right process—contributes a delay schema to the environment condition of \textit{RtParUp}. The concurrent real-time program \textit{RtParUp} defines, then, the real-time system of Figure 9. This real-time system is the cartesian product of the two real-time systems that are defined by the component processes.

4.3 Embedded real-time processes

At last, we illustrate our approach to real-time programming by designing typical application programs such as schedulers and embedded controllers. A scheduler is a real-time process that runs concurrently to the tasks that are being scheduled; an embedded controller runs concurrently to continuous environment processes.
program RtParUp:
declare x, y: clock;
initially m := 0; n := 0;
cobegin
  loop
    ℓ₀: x := 0; repeat ℓ₁: wait watching x ≥ 3 ℓ₂: until x ≥ 3;
    ℓ₃: m := m + 1;
    ℓ₄: x := 0; repeat ℓ₅: wait watching x ≥ 1 ℓ₆: until x ≥ 1;
    ℓ₇: m := m + 2
  end
  ||
  loop
    k₀: y := 0; repeat k₁: wait watching y ≥ 7 k₂: until y ≥ 7;
    k₃: n := n + 1
  end
coend.

Figure 8: Annotated concurrent real-time program

The simple round-robin scheduler *RoundRobin* of Figure 10 schedules a task for exactly 10 time units, and then moves on to the next task. The more sophisticated round-robin scheduler *TtRoundRobin* of Figure 11 schedules a task either for 10 time units or until the task terminates, whatever happens first. The variable *i* indicates which task is currently running; the bit *done[i]* indicates if task *i* has terminated.

The priority scheduler *Priority* of Figure 12 schedules a task until it terminates or until an interrupt arrives from a task with higher priority, whatever happens first. The scheduler *Priority* is a real-time process even though it does not use a clock: the guarded wait statement ensures that *Priority* waits until a high-priority task sets the bit *interrupt*. If we wish to guarantee that no interrupt is lost, we must assume that no two high-priority tasks arrive simultaneously at precisely the same instant in real-numbered time; alternatively, we could provide a separate bit *interrupt[i]* for each task *i*.

The real-time process *GateController* of Figure 13 controls the gate at a railroad crossing. The gate is closed 5 time units after a train signals its approach by setting the bit *arrival*, and the gate is opened again as soon as the train signals its exit from the crossing by setting the bit *exit*. The more sophisticated controller *MtGateController* of Figure 14 accommodates multiple trains, perhaps on parallel tracks, in the crossing. This controller uses the variable *n* to count the number of trains that are currently in the crossing under the assumption that no two trains arrive simultaneously. The gate is opened only when all
\[ \phi_0 = (left = \ell_0 \land right = k_0 \land m = 0 \land n = 0) \]
\[ \psi_0 = (left = \ell_0 \land left' = \ell_1 \land x' = 0)_{right,m,n,y} \lor
  (left = \ell_1 \land left' = \ell_2)_{right,m,n,x,y} \lor
  (left = \ell_2 \land x < 3 \land left' = \ell_3)_{right,m,n,x,y} \lor
  (left = \ell_2 \land x \geq 3 \land left' = \ell_3)_{right,m,n,x,y} \lor
  (left = \ell_3 \land left' = \ell_4 \land m' = m + 1)_{right,n,x,y} \lor
  (left = \ell_4 \land left' = \ell_5 \land x' = 0)_{right,m,n,y} \lor
  (left = \ell_5 \land left' = \ell_6)_{right,m,n,x,y} \lor
  (left = \ell_6 \land x < 1 \land left' = \ell_7)_{right,m,n,x,y} \lor
  (left = \ell_6 \land x \geq 1 \land left' = \ell_7)_{right,m,n,x,y} \lor
  (left = \ell_7 \land left' = \ell_0 \land m' = m + 2)_{right,n,x,y} \lor
  (right = k_0 \land right' = k_1 \land y' = 0)_{left,m,n,x} \lor
  (right = k_1 \land right' = k_2)_{left,m,n,x,y} \lor
  (right = k_2 \land y < 7 \land right' = k_1)_{left,m,n,x,y} \lor
  (right = k_2 \land y \geq 7 \land right' = k_3)_{left,m,n,x,y} \lor
  (right = k_3 \land right' = k_0 \land n' = n + 1)_{left,m,x,y} \]
\[ \chi_0 = (left = \ell_1 \land right = k_1 \land x' < 3 \land y' < 7) \lor
  (left = \ell_5 \land right = k_1 \land x' < 1 \land y' < 7) \]

Figure 9: Real-time system \((\phi_0, \psi_0, \chi_0)\) defined by the program \(RtParUp\)

trains have left the crossing—i.e., when the value of \(n\) is 0.

**Bibliographic remarks.** The synchronicity assumption—that program instructions are instantaneous and time is introduced through the environment—is due to Berry, whose ESTEREL language inspired also the *wait watching* macro (for an introduction to synchronous programming languages, see the monograph by Halbwachs [Hal93]). The use of control-location labels to define the system actions of a while program is standard practice of the interleaving school for concurrency [Lam91, MP92]. Schedulers and railroad gate controllers pervade the real-time literature as examples.

## 5 Timed Systems 3: Continuous Environment

### 5.1 Drifting clocks

So far, we have assumed that all clocks are mathematically precise. In distributed systems, however, one is often confronted with physical clocks, which are necessarily imprecise. Typically all that is known about a process clock \(x\) is that the drift of \(x\) is bounded; that is, in any time interval of length 1 the clock
Figure 10: Round-robin scheduler

```
program RoundRobin:
  initially i := FirstTask;
  loop
    ResumeTask(i);
    wait 10;
    SuspendTask(i);
    i := NextTask(i)
  end.
```

Figure 11: Round-robin scheduler for terminating tasks

```
program TtRoundRobin:
  declare x: clock;
  initially i := FirstTask;
  loop
    ResumeTask(i);
    x := 0; await x ≥ 10 ∨ done[i];
    if ¬done[i] then SuspendTask(i) fi;
    i := NextTask(i)
  end.
```

may deviate from the actual time by ±ε, for a real number ε > 0. We indicate that the drift of a clock x is bounded by ε with the clock declaration declare x: clock drift ε (the condition drift 0 is suppressed as usual). For example, Figure 15 shows the railroad gate controller of the previous section (Figure 13), assuming that the drift of the controller clock x is bounded by ε = 0.1. This version of the controller may close the gate as early as 4.5 time units and as late as 5.5 time units after the train signals its approach.

The behavior of drifting clocks during delays is best described by differential constraints. For this purpose, let ˙v be the first derivative of the variable v. Then, if the drift of the clock x is bounded by 0.1, all states that are passed by delays satisfy the differential constraint 0.9 ≤ ˙x ≤ 1.1. The resulting system is called “hybrid,” because it consists of both discrete system variables that are updated by system actions and continuous environment variables—namely, drifting clocks—that are updated by the environment during delays.
program \textit{Priority}:
\begin{verbatim}
  initially interrupt := false;
  loop
    i := MaxPriorityTask;
    ResumeTask(i);
    await interrupt \lor done[i];
    if interrupt then interrupt := false fi;
    if \neg done[i] then SuspendTask(i) fi
  end.
\end{verbatim}

Figure 12: Priority scheduler

program \textit{GateController}:
\begin{verbatim}
  initially arrival := false; exit := false;
  loop
    await arrival; arrival := false;
    wait 5;
    CloseGate;
    await exit; exit := false;
    OpenGate
  end.
\end{verbatim}

Figure 13: Railroad gate controller

5.2 Hybrid systems

We now illustrate that clock variables can be generalized to model not only time, or drift-free physical clocks, but arbitrary continuous environment parameters such as temperature or position.

The behavior of a hybrid system results from system actions and hybrid delays. A \textit{hybrid delay} \((f, \delta)\) consists of a nonnegative real number \(\delta\)—the duration of the delay—and a differentiable function \(f\) from the closed interval \([0, \delta]\) of the real line to states (a function \(f\) from the reals to states is differentiable if for every variable \(v\), the projection \(f(v)\) is differentiable). The state \(f(0)\) is the source state of the hybrid delay \((f, \delta)\); the state \(f(\delta)\) is its target state. A state \(\sigma'\) is passed by the hybrid delay \((f, \delta)\) if \(\sigma' = f(\epsilon)\) for some nonnegative real \(\epsilon < \delta\). A \textit{hybrid behavior} is a countably infinite sequence of system actions and hybrid delays such that any two neighboring sequence elements are consecutive.

Hybrid delays are described by hybrid delay predicates. Let \(\bar{V}\) be the set of
program MGateController:
  declare x : clock;
  initially arrival := false; exit := false;
loop
  await arrival; arrival := false; n := 1;
  x := 0;
repeat
  wait watching arrival ∨ exit ∨ x ≥ 5;
  if arrival then arrival := false; n := n + 1 fi;
  if exit then exit := false; n := n − 1 fi
  until x ≥ 5;
CloseGate;
repeat
  wait watching arrival ∨ exit;
  if arrival then arrival := false; n := n + 1 fi;
  if exit then exit := false; n := n − 1 fi
  until n = 0;
OpenGate
end.

Figure 14: Railroad gate controller for multiple trains

dotted variables whose undotted versions occur in V. A hybrid delay predicate
is a formula of predicate logic whose free variables are taken from V and \( \bar{V} \).
The hybrid delay predicate \( \chi \) is true for a hybrid delay if \( \chi \) is true for all states
that are passed by the delay. The hybrid delay predicate \( \chi \) is invariantly true
for a hybrid behavior \( \bar{\sigma} \) if \( \chi \) is true for all hybrid delays of \( \bar{\sigma} \).

A hybrid system, then, is a real-time system whose environment condition
is a hybrid delay predicate. The hybrid behavior \( \bar{\sigma} \) is a behavior of the hybrid
system \( \mathcal{S} \) if (1) the initial condition of \( \mathcal{S} \) is initially true for \( \bar{\sigma} \) and (2) the
transition condition and the environment condition of \( \mathcal{S} \) are invariantly true
for \( \bar{\sigma} \). Suppose that \( \mathcal{S} \) is a hybrid system with the environment condition \( \chi_\mathcal{O} \).
A variable \( m \) is a (discrete) system variable of \( \mathcal{S} \) if the environment condition \( \chi_\mathcal{O} \)
implies \( \bar{m} = 0 \); that is, the variable \( m \) is not modified during delays. All other
variables in \( V \) are (continuous) environment variables. In particular, a variable \( x \)
is a clock if the environment condition \( \chi_\mathcal{O} \) implies \( \dot{x} = 1 \); the variable \( y \) is a clock
with bounded drift \( \epsilon > 0 \) if the environment condition \( \chi_\mathcal{O} \) implies \( 1 − \epsilon ≤ \dot{y} ≤ 1 + \epsilon \); the variable \( z \) is a stopwatch if the environment condition \( \chi_\mathcal{O} \) implies \( \dot{z} = 0 \) \( \lor (\dot{z} = 1) \).

Environment variables, however, are not limited to measuring time. Consider
the example of a thermostat that regulates the room temperature. Suppose that
program DcGateController:
    declare x : clock drift 0.1;
    initially arrival := false; exit := false;
    loop
        await arrival; arrival := false;
        x := 0; await x = 5;
        CloseGate;
        await exit; exit := false;
        OpenGate
    end.

Figure 15: Railroad gate controller with a drifting clock

the initial room temperature is 68 degrees and the heat is turned on. When the
temperature reaches 70 degrees, the heat is turned off, and the temperature $\theta$
decreases over time $t$ according to the exponential function $\theta(t) = 70e^{-ct}$,
where $c_0$ is a constant determined by the room; when the temperature falls to
65 degrees, the heat is turned on, and the temperature $\theta$ increases according to
the exponential function $\theta(t) = 65e^{-ct} + c_1(1 - e^{-ct})$, where $c_1$ is a constant
determined by the power of the heater. The resulting hybrid system is defined
by the initial condition $\phi_0$, the transition condition $\psi_\nu$, and the environment
condition $\chi_\delta$:

$\phi_0 = (\text{heat} = \text{on} \land \theta = 68)$
$\psi_\nu = (\text{heat} = \text{off} \land \theta \leq 65 \land \text{heat}' = \text{on})_\theta \lor$
(heat = on ∧ θ ≥ 70 ∧ heat' = off)$_\theta$
$\chi_\delta = (\text{heat} = \text{on} \land \dot{\theta} = c_0(c_1 - \theta) \land \theta < 70) \lor$
(heat = off ∧ $\dot{\theta} = -c_0\theta \land \theta > 65$

Here, the status of the heater is modeled by the discrete (boolean-valued) system
variable $\text{heat}$; the room temperature, by the continuous (real-valued) environment
variable $\theta$.

5.3 Hybrid programs

We define hybrid systems by while programs such as the Train process of Fig-
ure 16. Clock variables (possibly drifting) are declared, as before, as $\text{clock}$; all
other environment variables, as $\text{continuous}$. In particular, the environment
variable $d$ represents the distance, in meters, of the train from a railroad cross-
ing. The dotted variable $d$ represents, then, the speed of the train in meters
per time unit. All undeclared variables, such as the bits $\text{arrival}$ and $\text{exit}$, are
discrete by default.
Given a hybrid delay predicate $\chi$ and a vector $\vec{v}$ of variables, we write $(\chi)_{\vec{v}}$ for the hybrid delay predicate $\chi \land \vec{v} = 0$; that is, the variables listed as a subscript remain unchanged. Given a hybrid program, let $\vec{m}$ be the vector of discrete variables, let $\vec{c}$ be the vector of clocks, and let $\vec{\gamma}$ be the corresponding vector of drift bounds. The instruction $\mathtt{wait}[^{\chi}]$ $\mathtt{watching}$ $\phi$ maintains the hybrid delay predicate $\chi$ while delaying the program for an arbitrary amount of time, but at most until the timed state predicate $\phi$ becomes true; it stands for the guarded wait statement

$$(1 - \vec{\gamma} \leq \vec{\gamma} \leq 1 + \vec{\gamma} \land \chi \land \neg\mathcal{F}_m) \rightarrow \mathtt{wait}.$$ 

In our example, the speed of the train stays between 45 and 55 meters per time unit until the train is 1,000 meters from the crossing. At this point, the train signals its approach, by setting the bit $\mathtt{arrival}$, and slows down to 35 to 50 meters per time unit. After the train is 100 meters past the crossing, it leaves and may return along a cyclic track at any time, provided the length of the loop is at least 5,000 meters. The hybrid program $\mathtt{Train}$ defines, therefore, the hybrid system of Figure 17.

At last, we may wish to combine, using the parallel-composition operator $||$, the $\mathtt{Train}$ process with the gate controller $\textit{GtController}$ into a concurrent hybrid program. The corresponding hybrid system is defined, as usual, as the cartesian product of the component systems: the initial condition is obtained by taking the conjunction of the initial conditions of the components; the transition condition, by taking the disjunction of the transition conditions of the components; and the hybrid delay schemata of the environment condition result from
\[ \phi_0 = (pc = \ell_0) \]

\[ \psi_\triangledown = (pc = \ell_0 \land pc' = \ell_1)_{arrival, est, d} \lor \\
(pc = \ell_1 \land pc' = \ell_2 \land d' = 5000)_{arrival, est} \lor \\
(pc = \ell_2 \land pc' = \ell_3)_{arrival, est, d} \lor \\
(pc = \ell_3 \land d > 1000 \land pc' = \ell_2)_{arrival, est, d} \lor \\
(pc = \ell_3 \land d \leq 1000 \land pc' = \ell_4)_{arrival, est, d} \lor \\
(pc = \ell_4 \land pc' = \ell_5 \land arrival' = true)_{est, d} \lor \\
(pc = \ell_5 \land pc' = \ell_6)_{arrival, est, d} \lor \\
(pc = \ell_6 \land d > -100 \land pc' = \ell_5)_{arrival, est, d} \lor \\
(pc = \ell_6 \land d \leq -100 \land pc' = \ell_7)_{arrival, est, d} \lor \\
(pc = \ell_7 \land pc' = \ell_0 \land exit' = true)_{arrival, d} \]

\[ \chi_\square = (pc = \ell_0)_{arrival, est} \lor \\
(pc = \ell_2 \land -55 \leq d \leq -45 \land d \leq 1000)_{arrival, est} \lor \\
(pc = \ell_5 \land -50 \leq d \leq -35 \land d \leq -100)_{arrival, est} \]

Figure 17: Hybrid system \( (\phi_0, \psi_\triangledown, \chi_\square) \) defined by the program \textit{Train}

Taking all pairwise conjunctions of the component delay schemata. We leave the details to the ambitious reader.

**Bibliographic remarks.** The bounded-drift assumption underlies the clock synchronization problem for distributed systems (see, for example, the survey by Schneider [Sch87]). A recent workshop proceedings provides an excellent current overview of computer-science models for hybrid systems [Ner93]. The dichotomy of discrete system actions versus continuous environment activities (hybrid delays) was introduced by Manna, Maler, and Pnueli [MMP92]. The use of environment conditions (invariants) to model hybrid systems in general [ACHH93], and drifting clocks in particular [AHH93], was first worked out in an automata-theoretic framework. The thermostat example is due to Nicollin, Sifakis, and Yovine [NSY92].

# 6 Analysis of Timed Systems

Now that we have defined the formal semantics of real-time systems, we wish to develop both methods and tools for proving the correctness of real-time programs. The verification problem has been addressed by several papers in recent years; here we attempt only to direct the reader to some of this literature.
6.1 Specification

First, we need a specification language for stating the correctness requirements of a given real-time program. Process algebras, temporal logics, and automata are popular specification languages for untimed systems. Hence a variety of real-time extensions of these languages has been proposed and investigated, including the process algebras TCSP [RR88], ATP [NRSV90], and TCCS [Wan90]; the temporal logics RTL [JM86], TPTL [AH89], TCTL [ACD90], MTL [AH90, Koy90], and RTTL [Ost90]; as well as timed automata [AD90] and timed I/O automata [MMT91, LV92a] (see [NS91] for a survey of real-time process algebras; [AH92] for a survey of logic-based and automata-based real-time languages; and [dBHdRR92] for a variety of applications).

Particularly compatible with our approach are the temporal logics TPTL (linear time) and TCTL (branching time), as well as timed automata: TPTL and TCTL are obtained by adding clock variables, reset quantifiers, and clock constraints to temporal logic; timed automata are obtained by adding clocks, reset conditions, and clock constraints to finite automata. If the acceptance condition of a timed automaton is replaced by an environment condition, the resulting timed safety automaton [HNSY92] corresponds precisely to our notion of real-time system.

To illustrate the use of clocks for specification, let us specify a few properties of the real-time while program \textit{RtUpDown3} (Figure 5) using the temporal logic TPTL. First, consider the invariance formula

$$\varphi_1: \quad \Box (n \leq 12).$$

A (timed) behavior \(\tilde{\sigma}\) satisfies the formula \(\varphi_1\) iff the (timed) state predicate \(n \leq 12\) is true for all states that occur in \(\tilde{\sigma}\); a (real-time) program \(\mathcal{P}\) satisfies \(\varphi_1\) iff every behavior of \(\mathcal{P}\) satisfies \(\varphi_1\).

Second, we wish to specify the timing behavior of the program \textit{RtUpDown3}. Let \(x\) be a clock. The time-bounded response formula

$$\varphi_2: \quad x := 0. \Diamond (x \leq 75 \land n = 0)$$

asserts that the timed state predicate \(n = 0\) becomes true, in all behaviors of a real-time program, within at most 75 time units from the initial state. Dually, the time-bounded invariance formula

$$\varphi_3: \quad x := 0. \Box (x < 15 \rightarrow n \neq 0)$$

asserts that the value of \(n\) does not become 0, in all behaviors of a real-time program, before time 15. The reader may check that, indeed, the program \textit{RtUpDown3} satisfies all three TPTL-formulas \(\varphi_1\), \(\varphi_2\), and \(\varphi_3\). We conclude that the real-time program \textit{RtUpDown3} takes at least 15 time units and at most 75 time units to terminate.
6.2 Verification

Second, we need to verify that a given real-time program satisfies its specification. For this purpose, most methods for proving the correctness of untimed programs have been extended to timed programs, including assertion methods [Hoo92, SBM92], proof systems for temporal logics [HMP91, AL92], and simulation methods [LA90, LV92b]. None of these approaches, however, use clocks, and we would like to see a proof system for real-time programs with clocks.

On the other hand, clocks have been central to the finite-state approach to verification. A (real-time) system is finite-state if all (system) variables range over finite domains. The verification problem for untimed finite-state systems is decidable, and a variety of software tools have been developed and successfully applied—particularly to the verification of protocols and circuits. Real-time systems are more complex: even in the case of real-time finite-state systems, the clocks range over the infinite domain of the nonnegative real numbers. It is possible, however, to construct a finite quotient of the infinite state space that is adequate for solving the reachability problem for real-time finite-state systems. This observation led to verification algorithms for finite-state real-time systems [ACD90, AD90, AFH91].

The efficiency and scope of the finite-state approach to the verification of real-time systems has been improved along several directions. First, algorithms for constructing the minimal adequate quotient of an infinite state space have been developed [ACH92, LY92]. Second, to cope with the high computational complexity of analyzing clock constraints, an incremental approach has been implemented [AIKY92]: initially all clock constraints are ignored; then the clock constraints are added one by one, as needed to prove a given specification. For example, the real-time program $\text{RtUpDown}$ continues to satisfy the properties $\phi_1$ and $\phi_2$ without the clock constraint $x \geq 12$.

A third technique for improving both the efficiency and the scope of verification algorithms is based on symbolic computation [HNSY92]. Suppose, for instance, we wish to prove that a certain property is an invariant of a real-time system. For this purpose, we need to compute the set of states that appear along all behaviors of the system. The timed state predicate that characterizes this set of reachable states is the least fixpoint of an equation on timed state predicates. The fixpoint can be computed by symbolic execution of the system, which is guaranteed to terminate if the system is finite-state. The efficiency of the fixpoint computation depends on the representation of state predicates; in the case of untimed systems, binary decision diagrams have turned out to provide a cost effective data structure for representing state predicates [McM93].

A hybrid system is piecewise linear if all continuous variables are bounded by piecewise-linear functions. In particular, clocks and stopwatches are piecewise linear; drifting clocks are bounded by piecewise-linear functions; all other variables can be approximated by piecewise-linear envelopes. The symbolic ver-
IFICATION method applies also to the class of piecewise-linear hybrid systems, albeit it may not terminate [ACHH93]. Prototype implementations are under way both at Cornell [AHH93] and in Grenoble [NOSY93].

Acknowledgments. The authors thank Limor Fix, David Karr, Peter Kopke, and Fred Schneider for helpful comments.

References


