Synthesis of Optimal Strategies Using HyTech

Patricia Bouyer[1⋆], Franck Cassez[2⋆], Emmanuel Fleury[3], Kim G. Larsen[3]

[1] LSV, UMR 8643, CNRS & ENS de Cachan, France
Email: bouyer@lsv.ens-cachan.fr

[2] IRCCyN, UMR 6597, CNRS, France
Email: cassez@irccyn.ec-nantes.fr

[3] Computer Science Department
BRICS[†], Aalborg University, Denmark
Email: {fleury,kgl}@cs.auc.dk

Abstract. Priced timed (game) automata extend timed (game) automata with costs on both locations and transitions. The problem of synthesizing an optimal winning strategy for a priced timed game under some hypotheses has been shown decidable in [?]. In this paper, we present an algorithm for computing the optimal cost and for synthesizing an optimal strategy in case there exists one. We also describe the implementation of this algorithm with the tool HyTech and present an example.

1 Introduction

In recent years the application of model-checking techniques to scheduling problems has become an established line of research. Static scheduling problems with timing constraints may often be formulated as reachability problems on timed automata, viz. as the possibility of reaching a given goal state. Real-time model checking tools such as Kronos and Uppaal have been applied on a number of industrial and benchmark scheduling problems [?,?,?,?,?,?].

Often the scheduling strategy needs to take into account uncertainty with respect to the behavior of an environmental context. In such situations the scheduling problem becomes a dynamic (timed) game between the controller and the environment, where the objective for the controller is to find a *dynamic* strategy that will guarantee the game to end in a goal state [?,?,?].

A few years ago, the ability to consider quite general performance measures has been given. Priced extensions of timed automata have been introduced [?,?] where a cost c is associated with each location ℓ giving the cost of a unit of time spent in ℓ. Within this framework, it is possible to measure performance of runs and to give optimality criteria for reaching a given set of states.

In [?], we have combined the notions of games and prices and we have proved that, under some hypotheses, the optimal cost in priced timed game automata is computable and that optimal strategies can then be synthetized.

⋆ Work partially supported by ACI Cortos, a program of the French government.

† Basic Research in Computer Science (www.brics.dk).

In this paper, we present an algorithm for extracting optimal strategies in priced timed game automata. We also provide an implementation of the algorithm using the tool HyTech [?]. The outline of the paper is as follows: in section 2 we recall the definition of Priced Timed Game Automata and present an example; in section 3 we unveil an optimal cost computation method; in section 4 we detail the algorithm to synthesize the optimal strategies and we give some conclusions in section 6.

The HyTech files given in Fig. 4 and Fig. 5 and available at `http://www.lsv.ens-cachan.fr/aci-cortos/ptga/`. The detailed proofs of the theorems we refer to, as well as complementary definitions and explanations can be found in [?].

2 Priced Timed Games

Preliminaries. Let X be a finite set of real-valued variables called clocks. We denote $\mathcal{B}(X)$ the set of constraints φ generated by the grammar: $\varphi ::= x \sim k \mid \varphi \wedge \varphi$ where $k \in \mathbb{Z}$, $x, y \in X$ and $\sim \in \{<, \leq, =, >, \geq\}$. A *valuation* of the variables in X is a mapping from X to $\mathbb{R}_{\geq 0}$ (thus an element of $\mathbb{R}_{\geq 0}^X$). For a valuation v and a set $R \subseteq X$ we denote $v[R]$ the valuation that agrees with v on $X \setminus R$ and is zero on R. We denote $v + \delta$ for $\delta \in \mathbb{R}_{\geq 0}$ the valuation s.t. for all $x \in X$, $(v + \delta)(x) = v(x) + \delta$.

The (R)PTGA Model.

Definition 1 (RPTGA). *A* Priced Timed Game Automaton *(PTGA) G is a tuple* $(L, \ell_0, \mathsf{Act}, X, E, \mathsf{inv}, \mathsf{cost})$ *where: L is a finite set of* locations*; $\ell_0 \in L$ is the* initial *location;* $\mathsf{Act} = \mathsf{Act}_c \cup \mathsf{Act}_u$ *is the set of* actions *(partitioned into controllable and uncontrollable actions); X is a finite set of* real-valued clocks*; $E \subseteq L \times \mathcal{B}(X) \times \mathsf{Act} \times 2^X \times L$ is a finite set of* transitions*;* $\mathsf{inv} : L \longrightarrow \mathcal{B}(X)$ *associates to each location its* invariant*;* $\mathsf{cost} : L \cup E \longrightarrow \mathbb{N}$ *associates to each location a* cost rate *and to each discrete transition a* cost *value. We assume that PTGA are deterministic w.r.t. controllable actions (renaming). A* reachability *PTGA (RPTGA) is a PTGA with a distinguished set of locations* $\mathsf{Goal} \subseteq L$.

Runs, Costs of Runs. Let $G = (L, \ell_0, \mathsf{Act}, X, E, \mathsf{inv}, \mathsf{cost})$ be a RPTGA. A *configuration* of G is a pair (ℓ, v) in $L \times \mathbb{R}_{\geq 0}^X$. We denote Q the set of configurations of G. A *run* in G from (ℓ'_0, v_0) is a (finite or infinite) sequence of transitions $t_i = (\ell'_i, v_i) \xrightarrow{\alpha_i} (\ell'_{i+1}, v_{i+1})$ such that for every $i \geq 0$: (ℓ'_i, v_i) is a configuration of G; $\alpha_i \in \mathsf{Act} \cup \mathbb{R}_{>0}$; $\alpha_i \in \mathbb{R}_{>0}$ implies $\ell'_{i-1} = \ell'_i$ and $v_i = v_{i-1} + \alpha_i$; $\alpha_i \in \mathsf{Act}$ implies that there exists a transition $(\ell'_{i-1}, g, \alpha_i, Y, \ell'_i) \in E$ such that $v_{i-1} \models g$ and $v_i = v_{i-1}[Y]$. The *cost* of a transition t_i is given by $\mathsf{Cost}(t_i) = \alpha_i.\mathsf{cost}(\ell'_{i-1})$ if $\alpha_i \in \mathbb{R}_{>0}$ and $\mathsf{Cost}(t_i) = \mathsf{cost}((\ell'_{i-1}, g, \alpha_i, Y, \ell'_i))$ if $\alpha_i \in \mathsf{Act}$. A run ρ of G is *winning* if at least one of the states along ρ is in the set Goal. We note $\mathsf{Runs}(G)$ (resp. $\mathsf{WinRuns}(G)$) the set of (resp. winning) runs in G and $\mathsf{Runs}((\ell, v), G)$ (resp. $\mathsf{WinRuns}((\ell, v), G)$) the set of (resp. winning) runs in G starting in configuration

(ℓ, v). If ρ is a *finite* run we note $last(\rho) = (\ell'_n, v_n)$ and the *cost* of the run ρ is defined by: $\mathsf{Cost}(\rho) = \sum_{0 \le i \le n-1} \mathsf{Cost}(t_i)$.

Example 1. Consider the RPTGA in Fig. 1. Plain arrows represent controllable actions ($\mathsf{Act}_c = \{c_1, c_2\}$) whereas dashed arrows represent uncontrollable actions ($\mathsf{Act}_u = \{u\}$). Cost rates in locations ℓ_0, ℓ_2 and ℓ_3 are 5, 10 and 1 respectively. In ℓ_1 the environment may choose to move to either ℓ_2 or ℓ_3. However, due to the invariant $y = 0$ this choice must be made instantaneously.

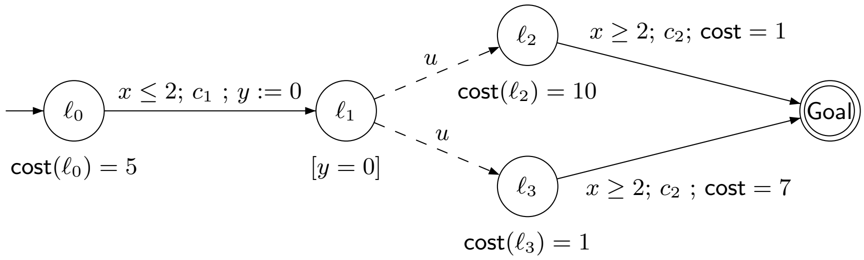

Fig. 1. A Reachability Priced Time Game Automaton $\mathcal{A}$

Strategies, Costs of Strategies.

Definition 2 (Strategy). *Let G be a (R)PTGA. A* strategy *f over G is a partial function from $\mathsf{Runs}(G)$ to $\mathsf{Act}_c \cup \{\lambda\}$.*

Definition 3 (Outcome). *Let $G = (L, \ell_0, \mathsf{Act}, X, E, \mathsf{inv}, \mathsf{cost})$ be a (R)PTGA and f a strategy over G. The* outcome *$\mathsf{Outcome}((\ell, v), f)$ of f from configuration (ℓ, v) in G is the subset of $\mathsf{Runs}((\ell, v), G)$ defined inductively by:*

- $(\ell, v) \in \mathsf{Outcome}((\ell, v), f)$,
- *if $\rho \in \mathsf{Outcome}((\ell, v), f)$ then $\rho' = \rho \xrightarrow{e} (\ell', v') \in \mathsf{Outcome}((\ell, v), f)$ if $\rho' \in \mathsf{Runs}((\ell, v), G)$ and one of the following three conditions hold:*
  1. $e \in \mathsf{Act}_u$,
  2. $e \in \mathsf{Act}_c$ *and* $e = f(\rho)$,
  3. $e \in \mathbb{R}_{\ge 0}$ *and* $\forall 0 \le e' < e, \exists(\ell'', v'') \in (L \times \mathbb{R}^X_{\ge 0})$ *s.t.* $last(\rho) \xrightarrow{e'} (\ell'', v'') \wedge f(\rho \xrightarrow{e'} (\ell'', v'')) = \lambda$.
- *an infinite run ρ is in $\in \mathsf{Outcome}((\ell, v), f)$ if all the finite prefixes of ρ are in $\mathsf{Outcome}((\ell, v), f)$.*

A strategy f over a RPTGA G is *winning* from (ℓ, v) whenever all maximal[1] runs in $\mathsf{Outcome}((\ell, v), f)$ are winning. We denote $\mathsf{WinStrat}((\ell, v), G)$ the set of winning strategies from (ℓ, v) in G. Let f be a winning strategy from configuration (ℓ, v). The *cost* of f from (ℓ, v) is defined by:

$$\mathsf{Cost}((\ell, v), f) = \sup\{\mathsf{Cost}(\rho) \mid \rho \in \mathsf{Outcome}((\ell, v), f)\}$$

[1] Roughly speaking a run is *maximal* if it can not be extended in the future by a controllable action (see [?] page 6, section 2.2); this point is discussed in the sequel in section 3.2.

Optimal Control Problems. Let $(\ell_0, \mathbf{0})$ denote the initial configuration of a RPTGA G. The three main problems we address in this paper are:

Optimal Cost Computation Problem: we want to compute the optimal cost one can expect in a RPTGA G from $(\ell_0, \mathbf{0})$, *i.e.* to compute

$$\mathsf{OptCost}((\ell_0, \mathbf{0}), G) = \inf\{\mathsf{Cost}((\ell_0, \mathbf{0}), f) \mid f \in \mathsf{WinStrat}((\ell_0, \mathbf{0}), G)\}$$

Optimal Strategy Existence Problem: we want to determine whether the optimal cost can actually be reached *i.e.* if there is an optimal strategy $f \in \mathsf{WinStrat}((\ell_0, \mathbf{0}), G)$ such that:

$$\mathsf{Cost}((\ell_0, \mathbf{0}), f) = \mathsf{OptCost}((\ell_0, \mathbf{0}), G)$$

Optimal Strategy Synthesis Problem: in case an optimal strategy exists we want to compute a witness.

Note that there are RPTGA for which no optimal strategy exists (see Example 2, Fig. 4, page 11 of [**?**]). In this case there is a family of strategies f_ε such that

$$|\mathsf{Cost}((\ell_0, \mathbf{0}), f_\varepsilon) - \mathsf{OptCost}((\ell_0, \mathbf{0}), G)| < \varepsilon$$

Thus another problem is, given ε, to compute such an f_ε strategy. This latter synthesis problem is not dealt with in this paper.

Example 2. We consider again Fig. 1. We want to compute an optimal strategy for the controller from the initial configuration. Obviously, once ℓ_2 or ℓ_3 has been reached the optimal strategy for the controller is to move to Goal asap (taking a c_2 action). The crucial (and only remaining) question is how long the controller should wait in ℓ_0 before taking the transition to ℓ_1 (doing c_1). Obviously, in order for the controller to win this duration must be no more than two time units. However, what is the optimal choice for the duration in the sense that the overall cost of reaching Goal is minimal? Denote by t the chosen delay in ℓ_0. Then $5t+10(2-t)+1$ is the minimal cost through ℓ_2 and $5t+(2-t)+7$ is the minimal cost through ℓ_3. As the environment chooses between these two transitions the best choice for the controller is to delay $t \leq 2$ such that $\max(21-5t, 9+4t)$ is minimum, which is $t = \frac{4}{3}$ giving a minimal cost of $14\frac{1}{3}$. In Fig. 2 we illustrate the optimal strategy for all states reachable from the initial state provided by our HyTech-implementation that will be described in section 3.4.

3 Optimal Cost Computation

In this section we show that computing the optimal cost for a RPTGA amounts to solving a simple[2] control problem on a linear hybrid game automaton [**?**]. As a consequence well-known algorithms [**?**,**?**] for computing winning states of reachability hybrid games enable us to compute the optimal cost of a RPTGA. We then show how to use the HyTech tool to implement the computation of the optimal cost for RPTGA.

[2] Without cost.

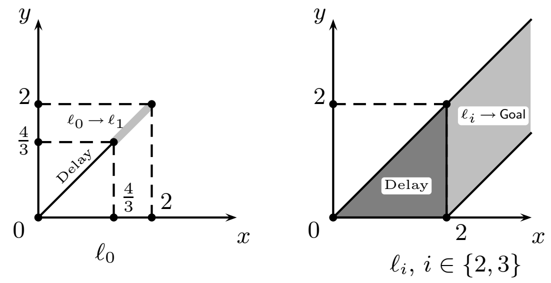

Fig. 2. *Optimal strategy for the RPTGA of Fig. 1. Optimal cost is* $14\frac{1}{3}$.

3.1 From Priced Timed Games to Linear Hybrid Games

Assume we want to compute the optimal cost of the RPTGA $\mathcal{A}$ given in Fig. 1. We translate this automaton into a linear hybrid game automaton (LHGA for short) $\mathcal{H}$ (see Fig. 3) where the cost function is encoded into a variable Cost of the LHGA. In $\mathcal{H}$ the variable Cost decreases with rate k in a location ℓ (*i.e.* $\frac{d\mathsf{Cost}}{dt} = -k$ in ℓ) if $\mathsf{Cost}(\ell) = k$ in $\mathcal{A}$. As for discrete transitions the variable Cost is updated by $\mathsf{Cost}' = \mathsf{Cost} - k$ in $\mathcal{H}$ if the corresponding transition's cost in $\mathcal{A}$ is k.

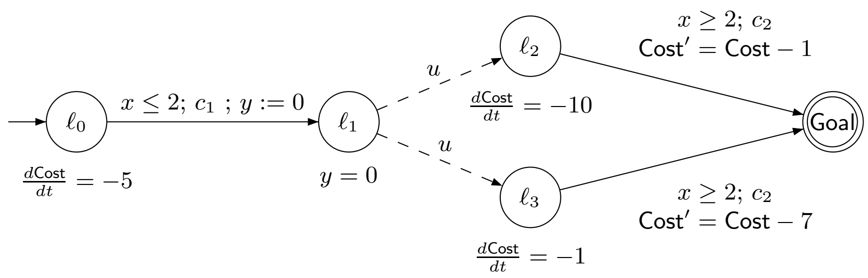

Fig. 3. The Linear Hybrid Game Automaton $\mathcal{H}$.

Let $\mathsf{CompWin}$ be a semi-algorithm (e.g. [?,?]) that computes the largest set of winning states for a reachability hybrid game. Using $\mathsf{CompWin}$ we can compute the largest set of winning states for $\mathcal{H}$ with the goal states given by $\mathsf{Goal} \wedge \mathsf{Cost} \geq 0$. The meaning of this new reachability game is that we want to win without having spent all the resources (Cost) we started with. Assume the corresponding (largest) set of winning states is denoted $\mathsf{CompWin}(\mathcal{H}, \mathsf{Goal} \wedge \mathsf{Cost} \geq 0)$. The meaning of the set $W = \mathsf{CompWin}(\mathcal{H}, \mathsf{Goal} \wedge \mathsf{Cost} \geq 0)$ is that in order to win one has to start in the region given by W and if one starts outside W the opponent has a strategy to win *i.e.* we loose. We can prove (see [?], Theorem 5 and Lemma 6) that the (largest) set of winning states is of the form $W = W' \wedge \mathsf{Cost} \succ k$ with $\succ \in \{\geq, >\}$ and $k \in \mathbb{Q}_{\geq 0}$ and W' is a set that constrains only the set of clocks X of the RPTGA (actually W' is the largest set of winning states of the game if we take the unpriced version with no cost, Theorem 4 of [?]). As a consequence k is the optimal cost: first note that we must start in W' to ensure we reach the Goal state; by definition of W if we start outside W

but in W' *i.e.* $\mathsf{Cost} = k_0$ with $k_0 < k$ (or $k_0 \leq k$) the opponent has a strategy to win and there are outcomes for which Cost will not be greater than or equal to 0 when we reach Goal. Thus computing $\mathsf{CompWin}(\mathcal{H}, \mathsf{Goal} \wedge \mathsf{Cost} \geq 0)$ is a semi-algorithm for computing the optimal cost of the RPTGA $\mathcal{A}$. Moreover we can decide (if $\mathsf{CompWin}$ terminates) whether there exists an optimal strategy or not: in case $W = W' \wedge \mathsf{Cost} > k$ there is no optimal strategy but a family of strategies f_ε (for $\varepsilon > 0$) with cost lower than $k + \varepsilon$. When $W = W' \wedge \mathsf{Cost} \geq k$ (and assuming $\mathsf{CompWin}$ terminates) we can compute an optimal strategy (this point is dealt with in the next section 4).

The formal definitions and proofs of this reduction are given in [**?**] (Definition 12, Lemma 5, Theorems 4 and 5, Corollaries 1 and 2).

3.2 The π Operator

The computation of the winning states (with $\mathsf{CompWin}$) is based on the definition of a *controllable predecessors* operator [**?**,**?**]. Let $G = (L, \ell_0, \mathsf{Act}, X, E, \mathsf{inv}, \mathsf{cost})$ be a RPTGA and Q its set of configurations of G. For a set $X \subseteq Q$ and $a \in \mathsf{Act}$ we define $\mathsf{Pred}^a(X) = \{q \in Q \mid q \xrightarrow{a} q', q' \in X\}$. The controllable and uncontrollable discrete predecessors of X are defined by $\mathsf{cPred}(X) = \bigcup_{c \in \mathsf{Act}_c} \mathsf{Pred}^c(X)$, respectively $\mathsf{uPred}(X) = \bigcup_{u \in \mathsf{Act}_u} \mathsf{Pred}^u(X)$. We also need a notion of *safe* timed predecessors of a set X w.r.t. a set Y. Intuitively a state q is in $\mathsf{Pred}_t(X, Y)$ if from q we can reach $q' \in X$ by time elapsing and along the path from q to q' we avoid Y. Formally this is defined by:

$$\mathsf{Pred}_t(X, Y) = \{q \in Q \mid \exists \delta \in \mathbb{R}_{\geq 0} \text{ s.t. } q \xrightarrow{\delta} q',\ q' \in X \ \wedge\ \mathsf{Post}_{[0,\delta]}(q) \subseteq \overline{Y}\} \quad (1)$$

where $\mathsf{Post}_{[0,\delta]}(q) = \{q' \in Q \mid \exists t \in [0, \delta] \mid q \xrightarrow{t} q'\}$. We are then able to define a *controllable predecessors* operator π as follows:

$$\pi(X) = \mathsf{Pred}_t\left(X \cup \mathsf{cPred}(X), \mathsf{uPred}(\overline{X})\right) \quad (2)$$

This definition of π captures the choice that uncontrollable actions cannot be used to win (this choice is made in [**?**] and in [**?**]). As a matter of fact there is no way to win in the RPTGA of Fig. 1 with this definition of π: ℓ_1 cannot be a winning state if we start iterating the computation of π from Goal as π only adds predecessors that can reach a winning state by a controllable transition. Another choice is possible: uncontrollable actions may be used to win if they are forced to happen. This second choice is rather involved when one wants to give a new definition of π in the general case. We adopt a position which is halfway between the previous two extremes: if an uncontrollable action is enabled from a state q where time cannot elapse and leads to a winning state q', and no uncontrollable transitions enabled at q can lead to a non-winning state, we declare q as winning. Assume the set of configurations of G where time cannot elapse is denoted $STOP$. Then a new definition of π where uncontrollable actions can be used to win is given by:

$$\pi'(X) = \mathsf{Pred}_t\left(X \cup \mathsf{cPred}(X) \cup (\mathsf{uPred}(X) \cap STOP), \mathsf{uPred}(\overline{X})\right) \quad (3)$$

Note that this choice does not change the results presented in [?]. In the example of Fig. 1, from location ℓ_1 only uncontrollable transitions are enabled, but they are bound to happen within a bounded amount of time (in this case as soon as we reach ℓ_1 because of the invariant $y = 0$). π' will add configuration $(\ell_1, x \geq 0 \wedge y = 0)$ to the set of winning states.

The semi-algorithm $\mathsf{CompWin}$ computes the least fixed point W of $\lambda X.\{X_0\} \cup \pi'(X)$ as the limit of an increasing sequence of sets of states W_i (starting from set $W_0 = X_0$) where $W_{i+1} = \pi'(W_i)$). If G is a RPTGA, the result of the computation of $\mathsf{CompWin}$ on the associated LHGA H starting from $\mathsf{Goal} \wedge \mathsf{Cost} \geq 0$ is $W = \mu X.\{\mathsf{Goal} \wedge \mathsf{Cost} \geq 0\} \cup \pi'(X)$. This result is also denoted $\mathsf{CompWin}(H, \mathsf{Goal} \wedge \mathsf{Cost} \geq 0)$ and gives the largest set of winning states.

3.3 Termination Issues

An important issue about the previous semi-algorithm $\mathsf{CompWin}$ is whether it terminates or not. We have identified a class of RPTGA for which $\mathsf{CompWin}$ terminates on the associated hybrid game.

Let G be a RPTGA satisfying:

- G is bounded, *i.e.* all clocks in G are bounded[3];
- the cost function of G is *strictly non-zeno*, *i.e.* there exists some $\kappa > 0$ such that the accumulated cost of every cycle in the region automaton associated with G is at least κ. Note that this condition can be checked. For more complete explanations, see [?].

Then the semi-algorithm $\mathsf{CompWin}(H, \mathsf{Goal} \wedge \mathsf{Cost} \geq 0)$ terminates (H is the hybrid game defined from G in the previous section). The formal statement and proof of this claim is given by Theorem 6 in [?]. We thus get:

Theorem 1. *Let G be a RPTGA satisfying the above-mentionned hypotheses (boundedness and strict non-zenoness of the cost). Then the optimal cost is computable for G.*

3.4 Implementation of CompWin in HyTech

HyTech [?,?] is a tool that implements "pre" and "post" operators for linear hybrid automata. Moreover it is possible to write programs that use these operators (and many others) on polyhedra in order to compute sets of states. The specification in HyTech of our LHGA $\mathcal{H}$ of Fig. 3 is given in Fig. 4, lines 7–25. We detail this specification in the sequel.

First note that the least fixed point of $\lambda X.\{\mathsf{Goal} \wedge \mathsf{Cost} \geq 0\} \cup \pi'(X)$ can be obtained equivalently using the operator π'' defined by:

$$\begin{aligned}\pi''(X) = \mathsf{Pred}_t(X, &\mathsf{uPred}(\overline{X})) \\ &\cup \mathsf{cPred}(X) \setminus \mathsf{uPred}(\overline{X}) \\ &\cup \big(\mathsf{uPred}(X) \cap \mathit{STOP}\big) \setminus \mathsf{uPred}(\overline{X})\end{aligned} \tag{4}$$

[3] This hypothesis is not a restriction, see [?].

For technical reasons (tagging of regions, see section 4), we prefer using this operator π'' instead of π' for computing the optimal cost.

Controllable and Uncontrollable Predecessors. HyTech provides the **pre** operator that computes at once the time predecessors and the discrete predecessors of a set of states. As we need to distinguish between time predecessors, discrete controllable predecessors and discrete uncontrollable predecessors, we use the following trick: in the HyTech source code of the LHA $\mathcal{H}$ we add two boolean variables u and c (Fig. 4, line 4) that are negated on each discrete uncontrollable (resp. controllable) transitions (Fig. 4, lines 11–24). In HyTech one can existentially quantify over a variable t by using the hide operator. Then the controllable predecessors can be computed by existentially quantifying over c and over a variable t that has rate[4] -1. We can express the cPred (and uPred) operator with existential quantifiers and two variables t and c as follows:

$$\mathsf{cPred}(X) = \{q \mid \ \exists c \in \mathsf{Act}_c \text{ s.t. } q \xrightarrow{c} q',\, q' \in X\} \tag{5}$$
$$= \{q \mid \exists t \mid \exists c \mid t = 0 \wedge c = 0 \wedge (q,t,c) \in \mathbf{pre}(X \wedge t = 0 \wedge c = 1)\} \tag{6}$$

where **pre** is the predecessor operator of HyTech.

We impose that the value of t stays unchanged to ensure that we just take discrete predecessors (Fig. 4, line 44). For uncontrollable predecessors we replace c by u (Fig. 4, line 46). Note that the computation of $STOP$ states (Fig. 4, line 37) can also be done using our extra variables t, c, u.

Safe Time Predecessors. The other operator $\mathsf{Pred}_t(Z, Y)$ is a bit more complicated. We just need to express it with existential quantification so that it is easy to compute it with HyTech. Also we assume we have time deterministic automata as in this case $\mathsf{Pred}_t(Z, Y)$ is rather simple (if we do not have time determinism a more complicated encoding must be done and we refer the reader to [?] for a detailed explanation.) From equation (1) we get:

$$\mathsf{Pred}_t(Z,Y) = \{q \mid \exists t \geq 0 \text{ s.t. } q \xrightarrow{t} q',\, q' \in Z \;\wedge\; \forall 0 \leq t_1 \leq t,\, q \xrightarrow{t} q'' \implies q'' \not\in Y\}$$
$$= \{q \mid \exists t \geq 0 \text{ s.t. } q \xrightarrow{t} q',\, q' \in Z \;\wedge\; \neg(\exists 0 \leq t_1 \leq t,\, q \xrightarrow{t} q'' \wedge q'' \in Y)\}$$

The latter formula can be encoded in HyTech using the hide operator (Fig. 4, lines 52–60) and two auxiliary variables t and t_1 that evolves at rate -1 (note that those variables are not part of the model but only used in existentially quantified formulas and they do not constrain the behavior of $\mathcal{H}$.) Finally if the computation of π'' terminates (Fig. 4, lines 40–64) the set `fix` contains all the winning states. It then suffices to compute the projection on `cost` in the initial state to obtain the optimal cost (Fig. 4, line 66).

Doing this we have solved the first two problems: computing the optimal cost and deciding whether there exists an optimal strategy.

[4] any rate different from 0 would also do but we need another variable t with rate -1 later on and use this one.

```
var
  x,y: clock;
  cost: analog; -- the cost variable
  c,u: discrete;    -- used to indicate that
  t,t1: analog;     -- used for existential quantification

automaton H
synclabs: ;
 initially  l0 & x=0 & y=0;

loc l0: while x>=0 & y>=0 wait {dcost=-5,dt=-1,dt1=-1}
        when x>=0 & x<=2 do {u'=u,c'=1-c,y'=0} goto l1;

loc l1: while y=0 wait {dcost=0,dt=-1,dt1=-1}
        when True do {u'=1-u,c'=c} goto l2;
        when True do {u'=1-u,c'=c} goto l3;

loc l3: while x>=0 & y>=0 wait {dcost=-10,dt=-1,dt1=-1}
        when x>=2 do {c'=1-c,u'=u,cost'=cost-1} goto Win;

loc l4: while x>=0 & y>=0 wait {dcost=-1,dt=-1,dt1=-1}
        when x>=2 do {c'=1-c,u'=u,cost'=cost-7} goto Win;

loc Win: while True wait {dcost=0,dt=-1,dt1=-1}
end

var init_reg,winning,fix,     -- sets of states
    STOP, -- set of STOP states from which time cannot elapse
    uPreX,uPrebarX,cPreX,X,Y,Z : region ;

-- first define the initial and winning regions
init_reg := loc[H]=l0 & x=0 ;
winning := loc[H]=Win & cost>=0;
-- fix is the fixpoint we want to compute i.e. the set of winning states W
fix := winning;
-- stopped states
STOP := ~(hide t,c,u in t>0 & c=0 & u=0 & pre(True & t=0 & c=0 & u=0) endhide) ;

-- compute the fixpoint of π''
X := iterate X  from winning using {
    -- uncontrollable predecessors of X̄: uPred(X̄)
    uPrebarX := hide t,u in t=0 & u=0 & pre(~X & u=1 & t=0) endhide;
    -- controllable predecessors of X: cPred(X)
    cPreX := hide t,c in t=0 & c=0 & pre(X & t=0 & c=1) endhide ;
    -- uncontrollable predecessors leading to winning states: uPred(X)
    uPreX := hide t,u in t=0 & u=0 & pre(X & u=1 & t=0) endhide;
    -- Z is the the first argument of π'' in the paper;
    -- Z = X ∪ cPred(X) ∪ (uPred(X) ∩ STOP)
    Z := (X | cPreX | (uPreX & STOP)) ;
    -- time predecessors of Z from which we can reach Z
    -- and avoid uPred(X̄) all along; X := Pred_t(Z, uPred(X̄))
    X := hide t in
           (hide c,u in t>=0 & c=0 & u=0 & pre(Z & t=0 & c=0 & u=0)
            endhide) &
           ~(hide t1 in
                (hide c,u in t1>=0 & t1<=t & c=0 & u=0 &
                       pre(uPrebarX & t1=0 & c=0 & u=0)
                 endhide)
             endhide)
         endhide;
    -- add the newly computed regions to the set of already
    -- computed region
    fix := fix | X ;
} ;
-- print the result
print omit all locations hide x,y in fix & init_reg endhide;
```

Fig. 4. Computation of the Optimal Cost.

4 Optimal Strategies Computation

In this section we show how to compute an optimal strategy when one exists. Then we give the HyTech implementation of this computation and discuss some properties of those strategies.

4.1 Strategy Synthesis For RPTGA

First we recall some basic properties of strategies for (unpriced) Timed Game Automata (TGA).

A strategy f is

- *state-based* whenever $\forall \rho, \rho' \in \mathsf{Runs}(G), \mathit{last}(\rho) = \mathit{last}(\rho')$ implies that $f(\rho) = f(\rho')$. State-based strategies are also called *memoryless* strategies in game theory [?,?];
- *polyhedral* if for all $a \in \mathsf{Act}_c \cup \{\lambda\}$, $f^{-1}(a)$ is a finite union of convex polyhedra for each location the game;
- *realizable*, whenever the following holds: for all $\rho \in \mathsf{Outcome}(q, f)$ s.t. f is defined on ρ and $f(\rho) = \lambda$, there exists some $\delta > 0$ such that for all $0 \leq t < \delta$, there exists q' with $\rho \xrightarrow{t} q' \in \mathsf{Outcome}(q, f)$ and $f(\rho \xrightarrow{t} q') = \lambda$.

Strategies which are not realizable are not interesting because they generate empty sets of outcomes. Nevertheless it is not clear from [?,?] how to extract strategies for reachability TGA and ensure their realizability[5]. This is why we have provided a secondary result in [?] for Linear Hybrid Games (Theorem 2 page 7) that can be rephrased in the context of RPTGA as:

Theorem 2 (Adapted from Theorem 2 of [?]). *Let G be a RPTGA. If the semi-algorithm* CompWin *terminates for the hybrid game associated with G (see section 3.1), then we can compute a winning strategy which is: polyhedral, realizable and stated-based.*

Let H be the LHGA associated to the RPTGA $G = (L, \ell_0, \mathsf{Act}, X, E, \mathsf{inv}, \mathsf{cost})$. A state of H is a triple (ℓ, v, c) where $\ell \in L, v \in \mathbb{R}_{\geq 0}^X$ and $c \geq 0$ (c is the value of the variable Cost of H). Thus if we synthetize a realizable winning state-based strategy f for H, we obtain a strategy that depends on the cost value. In case there is a winning strategy for H (see section 3.1) we can synthetize realizable state-based winning strategies for G (see [?], Corollary 2). This result is already satisfying but we would like to build strategies that are independent of the cost value *i.e.* in which there is no need for extra information to play the strategy on the original RPTGA G (this means we want to build a state-based strategy for the original RPTGA G.) To this extent we introduce the notion of *cost-independent* strategies.

Let $W = \mathsf{CompWin}(H, \mathsf{Goal} \wedge \mathsf{Cost} \geq 0)$ be the set of winning states of H. A state-based strategy f for H is *cost-independent* if $(\ell, v, c) \in W$ and $(\ell, v, c') \in W$

[5] See [?],page 5, Fig. 2 for an example of non realizable strategy.

imply $f(\ell, v, c) = f(\ell, v, c')$. Cost-independent strategies in H will then be used for having state-based strategies in G. Theorem 7 of [?] gives then sufficient conditions for the existence of a state-based, optimal, realizable strategy in G and, when back to the automaton reads as follows:

Theorem 3 (Adapted from Theorem 7 of [?]). *Let G be a RPTGA satisfying the following hypotheses:*

1. *G is bounded;*
2. *the cost function of G is strictly non-zeno;*
3. *constraints on controllable actions are non-strict (i.e. possible constraints are of the form $x \leq c$ and $x \geq c$);*
4. *constraints on uncontrollable actions are strict (i.e. possible constraints are of the form $x < c$ and $x > c$)*

Then we can compute a winning strategy for G which is: optimal, state-based and realizable.

Note that under the previous conditions we build a strategy which is *globally optimal* *i.e.* optimal for all winning states (and not only for the initial winning states). We now give an algorithm to extract such an optimal, state-based, realizable and winning strategy for a RPTGA G.

The example of Fig. 1 satisfies the assumptions of Theorem 3 and thus we can compute an optimal strategy for this model. Moreover, the strategy we obtain using HyTech is precisely the one we described in Fig. 2.

4.2 Synthesis of Optimal Strategies

First we recall that the set of winning states of H is computed iteratively using the functional π'' defined by equation (4). Note that $\mathsf{cPred}(X) \setminus \mathsf{uPred}(\overline{X})$ is equal to $\bigcup_{c \in \mathsf{Act}_c} \mathsf{Pred}^c(X) \setminus \mathsf{uPred}(\overline{X})$. In the sequel we need to compute the states that can let a strict positive delay elapse to define the strategy for the delay action. For a set X we denote $\mathsf{NonStop}(X)$ the set of states in X from which a strict positive delay can elapse and all the intermediary states lie in X *i.e.*

$$\mathsf{NonStop}(X) = \{q \in X \mid \exists t > 0 \mid q + t \in X \wedge \forall 0 \leq t' \leq t,\, q + t' \in X\} \qquad (7)$$

Tagged Sets. To synthetize strategies we compute iteratively a set of extended "tagged" states W^+ during the course of the computation of W (this follows from Theorem 2 and Lemma 6 of [?]). The tags will contain information about how a new set of winning states $W_{i+1} = \pi''(W_i)$ has been obtained.

We start with $W_0^+ = \emptyset$ and $W_0 = \mathsf{Goal} \wedge \mathsf{Cost} \geq 0$. Assuming W_i and W_i^+ are the sets obtained after i iterations of π'' we define W_{i+1}^+ as follows:

1. let $Y = W_{i+1} \setminus W_i$ where $W_{i+1} = \pi''(W_i)$;

2. for each $c \in \mathsf{Act}_c$, we define the tagged set $\Big(Y \cap \mathsf{cPred}^c(W_i)\Big)^{[c]}$ with the intended meaning: "$Y \cap \mathsf{cPred}^c(W_i)$ has been added to the set of winning states by a Pred^c and doing a c from this set leads to W_i";
3. define another tagged set $\Big(\mathsf{NonStop}(Y \cap \mathsf{Pred}_t(W_i))\Big)^{[\lambda]}$ with the intended meaning: "$\mathsf{NonStop}(Y \cap \mathsf{Pred}_t(W_i))$ has been added to the set of winning states by the (strictly positive) time predecessors operator and letting time elapse will lead to W_i";
4. define W_{i+1}^+ by :

$$W_{i+1}^+ = W_i^+ \cup \Big(\mathsf{NonStop}(Y \cap \mathsf{Pred}_t(W_i))\Big)^{[\lambda]} \cup \bigcup_{c \in \mathsf{Act}_c} \Big(Y \cap \mathsf{cPred}^c(W_i)\Big)^{[c]}$$

Computation of an Optimal Strategy. If $\mathsf{CompWin}$ terminates in j iterations we end up with $W^+ = W_j^+$. Note that by construction a state q of W may belong to several tagged sets $X_0^{[\lambda]}$, $X_1^{[c_1]}$, ..., $X_n^{[c_n]}$ (where $c_i \in \mathsf{Act}_c$ for each $i \in [1, n]$) of W^+. All the X_i's are of the form $X'_i \wedge \mathsf{Cost} \geq f_i$ where $X'_i \subseteq \{\ell_i\} \times \mathbb{R}_{\geq 0}^X$ for some location ℓ_i and $f_i : X'_i \to \mathbb{R}_{\geq 0}$ is a piecewise affine function (Lemma 6 of [**?**]) because the constraints on the guards of theorem 3 imply that each cost constraint is of the form $\mathsf{Cost} \geq f_i$. Thus the infimum of f_i is reachable and equal to the minimum of f_i.

Theorem 7 of [**?**] states that in this case an optimal state-based strategy f^* for q a winning state of G (no cost) will be obtained by taking the local optimal choice: let $m = \min_{i \in [0,n]} f_i(q)$ then defining $f^*(q) = c_i$ if $f_i(q) = m$ gives an optimal strategy.

As it can be the case that $f_i(q) = f_j(q) = m$ with $i \neq j$, we impose a total order $\sqsubset$ on the set of events in Act_c and define $f^*(q) = c_i$ where $i = \max\{j \mid f_j(q) = m\}$. To avoid realizability problems (see proof of Lemma 6 in [**?**]) on the boundary of a set X_i if $f_0(q) = m$ (which means that the optimal cost can be achieved by time elapsing) and $f_i(q) = m$ for some $i \in [1, n]$ we define $f^*(q) = c_i$. This can be easily defined in our setting by extending $\sqsubset$ to $\mathsf{Act}_c \cup \{\lambda\}$ and making λ the smallest element.

After these algorithmics explanations, we can summarize how we can synthesize an optimal, cost-independent strategy. We denote $W_{[c]}^+$ the set defined by:

$$W_{[c]}^+ = \bigcup_{S_i^{[c]} \in W^+} S_i \tag{8}$$

For each $c \in \mathsf{Act}_c \cup \{\lambda\}$, $W_{[c]}^+$ is a set of the form $X^c \wedge \mathsf{Cost} \geq h^c$ where h^c is a piecewise affine function on X^c (X^c is a union of convex polyhedra). Note that the constraint $\mathsf{Cost} \geq h^c$ is a polyhedron which constrains the Cost variable and the clocks. In what follows, a pair (q, α) will represent a state of H (α is the value of the Cost variable). For each winning state q of G, we want to compute the minimal cost for winning and which action we should do if we want to win with the optimal cost. Let us consider two actions $c_1,\ c_2 \in \mathsf{Act}_c \cup \{\lambda\}$. We denote $[c_1 \leq c_2]$ the set of winning states of G where it is better to do action c_1 than action c_2 ($h^{c_1}(q) \leq h^{c_2}(q)$). This set is defined by:

$$\begin{aligned}[c_1 \leq c_2] = \{q \in X^{c_1} \ \mid \exists \alpha_1 \,|\, (q,\alpha_1) \in W^+_{[c_1]} \ \mathit{and} \\ \forall \alpha_2 \,|\, (q,\alpha_2) \in W^+_{[c_2]},\ \alpha_1 \leq \alpha_2\}\end{aligned} \tag{9}$$

$$\begin{aligned}= \big\{q \in X^{c_1} \ \mid \exists \alpha_1 \,|\, (q,\alpha_1) \in W^+_{[c_1]} \wedge \\ \neg\big(\exists \alpha_2 \,|\, (q,\alpha_2) \in W^+_{[c_2]} \ \mathit{and}\ \alpha_2 < \alpha_1\big)\big\}\end{aligned} \tag{10}$$

Each set $[c_1 \leq c_2]$ is a polyhedral set. For each $c \in \mathsf{Act}_c \cup \{\lambda\}$ define

$$\mathsf{Opt}(c) = \bigcap_{c' \neq c} [c \leq c'] \tag{11}$$

$\mathsf{Opt}(c)$ is the set of states for which c is an action that gives the optimal cost. $W^* = \bigcup_{c \in \mathsf{Act}_c \cup \{\lambda\}} \mathsf{Opt}(c)$ is thus equal to the set of states on which we need to define the optimal strategy. Given the total order $\sqsubset$ on $\mathsf{Act}_c \cup \{\lambda\}$ with $\lambda \sqsubset c_1 \sqsubset \cdots \sqsubset c_n$, we can define an optimal strategy f^* as follows: for $i \in [0, n-1]$, let $B_i = \big(W^* \setminus (\cup_{k>i} B_k)\big) \cap \mathsf{Opt}(c_i)$ and $B_n = W^* \cap \mathsf{Opt}(c_n)$; define then $f^*(q) = c_i$ if $q \in B_i$. f^* is an optimal strategy that is (winning), state-based, realizable and polyhedral.

4.3 Implementation in HyTech

Controllable Tagged Sets. We first show how to compute tagged sets of states. Our HyTech encoding consists in adding a discrete variable a to the HyTech model of Fig. 4 and use it in the guards of controllable transitions: controllable action c_k of Fig. 3 corresponds to the guard $a = k$ in the HyTech model. The HyTech model of Fig. 4 is enriched as follows: we add the guard $a = 1$ to line 12, $a = 2$ to lines 19 and 22. In this way we achieve the tagging of controllable predecessors as now the computation of cPred (line 44 of Fig. 5) will compute a tagged region that will be a union of polyhedra with some $a = k$ constraints. Note that we also modify line 49 of Fig. 4 and replace it by line 25 in Fig. 5 where a is hidden from the new `cPreX` by `(hide a in cPreX endhide)` as a is not needed to compute the winning set of states.

New NonStop States. To compute $\mathsf{NonStop}(Y \cap \mathsf{Pred}_t(W_i))$ we use again our extra variables t, c, u and add the tag $a = 0$ to the result set. Lines 40–41 of Fig. 5 achieves this.

W^+ is stored in the region `fix_strat` in the HyTech code. To compute W^+_{i+1} we update `fix_strat` as described by line 43 in Fig. 5.

Computation of the Optimal Strategy. To compare the costs for each action and determine the optimal one we use the trick described in the previous subsection. Each tagged set gives the function h^{c_i} by the means of a constraint between the Cost variable and the rest of the state variables. To compute h^{c_i} we need to split the state space according to each action c_i: this is achieved by lines 49–51 where the state space that corresponds to h^{c_i} is stored in `ri`.

It remains to compute for each pair of actions (c_1, c_2) (c_i can be λ), the set $[c_1 \leq c_2]$ states. The encoding in HyTech of the formula given by equation (10) is quite straightforward using the `hide` operator that corresponds to existential quantification. The strategy is then computed has described at the end of the previous subsection by lines 59–68.

5 Experiments

Using a HyTech-code as described in this paper, we have done some more experiments. The most important example we have treated is a model of a mobile phone with two antennas trying to connect to a base station with an environment which can possibly jam some transmissions.

Description of the Mobile Phone Example. We consider a mobile phone with two antennas emitting on different channels. Making the initial connection with the base station takes 10 time units whatever antenna is in use. Statistically, a jam of the transmission (*e.g.* collision with another phone) may appear every 6 time units in the worst case. When a collision is observed, the antenna tries to transmit with a higher level of energy for a while (at least 5 time units for Antenna 1 and at least 2 time units for Antenna 2) and then can switch back to the lower consumption mode. Unfortunately, switching back to the low consumption mode requires more resources and forces to interrupt the other transmission (Antenna 1 resets variable y of Antenna 2 and vice-versa). The overall cost rate (consumption per time unit) for the mobile phone in a product state $s = (\mathsf{low_x}, \mathsf{high_y}, \mathsf{Y})$ is the sum of the rates of Antenna 1 and Antenna 2 (both are working) *i.e.* $1 + 20 = 21$ and $\mathsf{Cost}(s) = 21$ in our model. Once the connection with the base station is established (either $x \geq 10$ or $y \geq 10$) the message is delivered with an energy consumption depending on the antenna ($\mathsf{Cost} = 7$ for $\mathsf{Antenna\ 1}$ and $\mathsf{Cost} = 1$ for $\mathsf{Antenna\ 2}$). The aim is to connect the mobile phone with an energy consumption (cost) as low as possible whatever happens in the network (jam).

This system can be represented by a network of PTGAs (see Fig. 6) and the problem reduces to finding an optimal strategy for reaching one of the goal states $\mathsf{Goal_x}$ or $\mathsf{Goal_y}$. Note that our original model is a single PTGA and not a network of PTGAs, but networks of PTGA can be used as well because it does not add expressive power and it is simple to define the composed PTGA: in a global location (being a tuple of locations of simple PTGAs), the cost is simply the sum of the costs of all single locations composing it. *Idem* for a composed transition resulting from a synchronization: the cost of the synchronized transition is the sum of the costs of the two initial transitions. Of course, one has to pay attention that no controllable action can synchronize with an uncontrollable

```
-- same sets of variables as lines 1–5 in Fig. 4 plus some new vars:
var
  a: discrete;
  cost0,cost1,cost2: analog;

 Insert Here lines 7-25 of Fig. 4 with guard a = k when needed 
-- same set of variables here as lines 27–29 in Fig. 4 plus some new vars:
var
    fix_strat,nonstop,Y,
    r0,r1,r2,
    B0,B1,B2,
    inf_0_1,inf_0_2,inf_1_0,inf_1_2,inf_2_0,inf_2_1: region;

init_reg := loc[H]=l0 & x=0 ;
winning := loc[H]=Win & cost>=0;
fix := winning;
STOP := ~(hide t,c,u in t>0 & c=0 & u=0 & pre(True & t=0 & c=0 & u=0) endhide) ;

fix_strat := False; -- this is new and corresponds to W_0^+ = ∅

X := iterate X  from winning using {
    uPrebarX := hide t,u in t=0 & u=0 & pre(~X & u=1 & t=0) endhide;
    cPreX := hide t,c in t=0 & c=0 & pre(X & t=0 & c=1) endhide ;
    uPreX := hide t,u in t=0 & u=0 & pre(X & u=1 & t=0) endhide;
    Z := (X | (hide a in cPreX endhide) | (uPreX & ~uPrebarX & STOP)) ;
    X := hide t in
            (hide c,u in t>=0 & c=0 & u=0 & pre(Z & t=0 & c=0 & u=0)
             endhide) &
            ~(hide t1 in
                (hide c,u in t1>=0 & t1<=t & c=0 & u=0 &
                        pre(uPrebarX & t1=0 & c=0 & u=0)
                 endhide)
             endhide)
         endhide;

    Y := X & ~fix ; -- store the real new states in Y

    fix := fix | X ;
    -- computation of NonStop(X)
    nonstop := a=0 & Y &
        hide t,c,u in t>0 & c=0 & u=0 & pre(Y & t=0 & u=0 & c=0) endhide;
    -- computation of fix_strat
    fix_strat := fix_strat | (Y & cPreX) | nonstop ;
} ;
-- print the result as before
print omit all locations hide x,y in fix & init_reg endhide;

-- rename the cost fonction; then ri corresponds to h^{c_i}
r0 := hide a,cost in cost0=cost & fix_strat & a=0 endhide ;
r1 := hide a,cost in cost1=cost & fix_strat & a=1 endhide ;
r2 := hide a,cost in cost2=cost & fix_strat & a=2 endhide ;

-- compute the state space inf_i_j where h^{c_i}(q) ≤ h^{c_j}(q)
inf_0_1 := hide cost0 in r0 & ~(hide cost1 in r1 & cost1<cost0 endhide) endhide ;
...
inf_2_1 := hide cost2 in r2 & ~(hide cost1 in r1 & cost1<cost2 endhide) endhide ;

-- Output the result taking the best move according to the total order (Act_c ∪ {λ}, ⊏)
prints "Optimal Winning Strategy" ;
prints "do control from l3 or l4 to Win (a=2) on";
B2 := inf_2_0 & inf_2_1 ;
print B2 ;
prints "do control from l0 to l1 (a=1) on";
B1 := inf_1_0 & inf_1_2 & ~B2 ;
print B1 ;
prints "do wait (a=0) on";
B0 := inf_0_1 & inf_0_2 & ~B1 & ~B2 ;
print B0;
```

Fig. 5. Synthesis of Optimal Strategies.

action in order that we can define properly the nature, controllable or not, of the synchronization. In this example, see Fig. 6, $\mathsf{jam_x}$? (resp.$\mathsf{jam_y}$?) synchronizes with $\mathsf{jam_x}$! (resp. $\mathsf{jam_y}$!). The HyTech code of this example can be found in [?] and on the web page `http://www.lsv.ens-cachan.fr/aci-cortos/ptga/`.

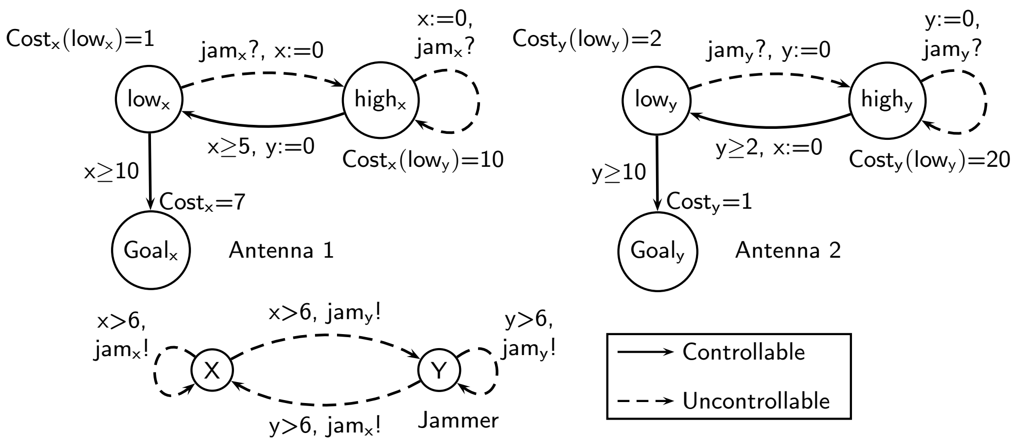

Fig. 6. *Mobile Phone Example.*

Results of our Experiments. We got that the optimal cost (lowest energy consumption) that can be ensured is 109. The optimal strategy is graphically represented on Fig. 7. The strategy is non-trivial and the actions to take depend on a complex partitioning of the clock space.

The computation took 828s on a 12" PowerBook G4 running Mac OS X.

6 Conclusion

In this paper we have described an algorithm to synthesize optimal strategies for a sub-class of priced timed game automata. The algorithm is based on the work described in [?] where we proved this problem was decidable (under some hypotheses we recall in this paper). Morever, we also provide an implementation of our algorithm in HyTech and use on small case-studies.

Our future work consists in extending the class of systems for which the algorithm we provided is correct and terminates. We would also like to extend this work to more general winning conditions (like safety conditions) and with other performance criteria (as for example the price per unit of time along infinite schedules).

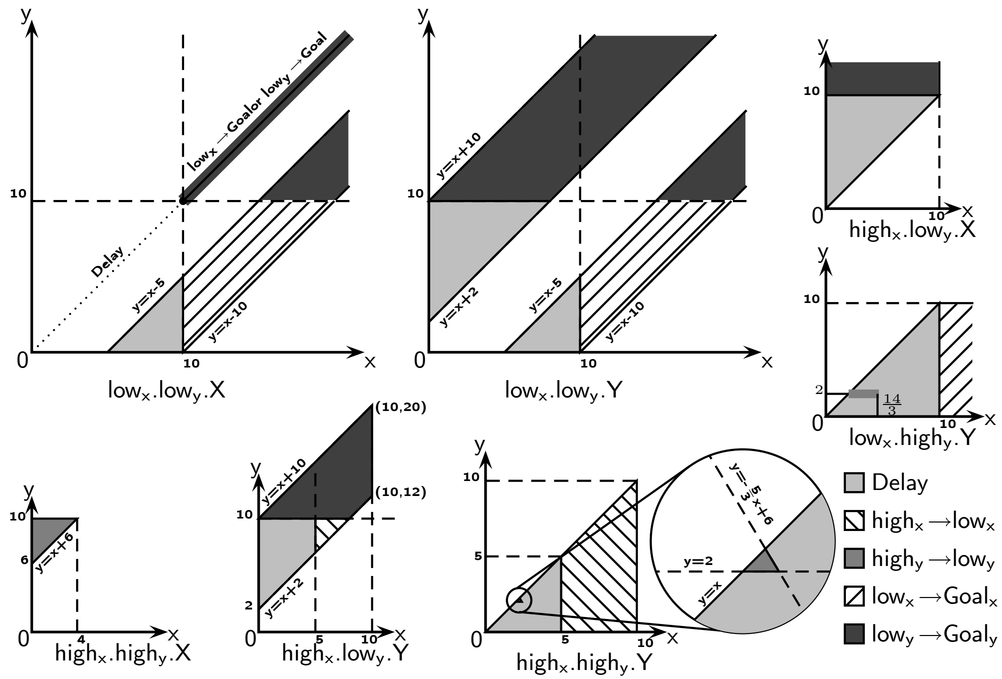

Fig. 7. *Strategy of the* mobile phone *example.*