

---

# DISTRIBUTED ASYNCHRONOUS GAMES WITH CAUSAL MEMORY ARE UNDECIDABLE

HUGO GIMBERT

LaBRI, CNRS, Université de Bordeaux, France  
*e-mail address:* hugo.gimbert@cnrs.fr

---

**ABSTRACT.** We show the undecidability of the controller synthesis problem when both the plant and the controllers are asynchronous automata and the controllers have causal memory.

## INTRODUCTION

The decidability of the distributed version of the Ramadge and Wonham control problem [12], where both the plant and the controllers are modeled as asynchronous automata [13, 1] and the controllers have causal memory has been an open problem for some time [3, 9, 7].

In this setting a controllable plant is distributed on several finite-state processes which interact asynchronously using shared actions. On every process, the local controller can choose to block some of the actions, called *controllable* actions, but it cannot block the *uncontrollable* actions from the environment. The choices of the local controllers are based on two sources of information. First the controller monitors the sequence of states and actions of the local process. This information is called the *local view* of the controller. Second when a shared action is played by several processes then all the controllers of these processes can exchange as much information as they want. In particular together they can compute their mutual view of the global execution: their *causal past*.

A controller is correct if it guarantees that every possible execution of the plant satisfies some specification. The controller synthesis problem is a decision problem which, given a plant as input, asks whether the system admits a correct controller. In case such a controller exists, the algorithm should compute one as well.

The difficulty of controller synthesis depends on several factors, including the size and architecture (pipeline, ring, ...) of the system, the information available to the controllers, and the specification.

Assuming that processes can exchange information upon synchronization and use their causal past to take decisions is one of the key aspects to get decidable synthesis problems [3]. In early work on distributed controller synthesis, for example in the setting of [10], the only source of information available to the controllers is their local view. In this setting, distributed synthesis is not decidable in general, except for very particular architectures

---

Received by the editors June 2021.

like the pipeline architecture. The paper [2] proposes information forks as a uniform notion explaining the (un)decidability results in distributed synthesis. The idea of using causal past as a second source of information appeared in [3].

We adopt a modern terminology and call the plant a *distributed game* and the controllers are *distributed strategies* in this game. A distributed strategy is a function that maps the causal past of processes to a subset of controllable actions. In the present paper we focus on the *termination condition*, which is satisfied when each process is guaranteed to terminate its computation in finite time, in a final state. A distributed strategy is winning if it guarantees the termination condition, whatever uncontrollable actions are chosen by the environment.

We are interested in the following algorithmic problem:

**DISTRIBUTED SYNTHESIS PROBLEM:** given a distributed game decide whether there exists a winning strategy.

There are four classes of plants for which the DISTRIBUTED SYNTHESIS PROBLEM has been shown decidable: when the dependency graph of actions is series-parallel [3]; when the processes are connectedly communicating [6]; when the dependency graph of processes is a tree [4, 8]; and for the class of decomposable games, which encompasses these three former decidability results [5].

**Our contribution.** This paper shows that the DISTRIBUTED SYNTHESIS PROBLEM for plants with six processes is undecidable. The proof is in two steps: reduce the POST CORRESPONDENCE PROBLEM to an intermediary decision problem called the BIPARTITE COLORING PROBLEM (Theorem 2.2) and then reduce the latter problem to the DISTRIBUTED SYNTHESIS PROBLEM (Theorem 3.1). The combinatorial details incoming from the initial PCP instance are abstracted away in the proof of Theorem 2.2 while the proof of Theorem 3.1 sheds light on the use of concurrency and causal memory to propagate local constraints.

**Organisation of the paper.** Section 1 defines the DISTRIBUTED SYNTHESIS PROBLEM. Section 2 defines the BIPARTITE COLORING PROBLEM and shows that the POST CORRESPONDENCE PROBLEM effectively reduces to the BIPARTITE COLORING PROBLEM. Section 3 effectively reduces the BIPARTITE COLORING PROBLEM to the DISTRIBUTED SYNTHESIS PROBLEM.

## 1. DISTRIBUTED ASYNCHRONOUS GAMES WITH CAUSAL MEMORY

The theory of Mazurkiewicz traces is very rich, for a thorough presentation see [1]. Here we only fix notations and recall the notions of traces, views, prime traces and parallel traces.

We fix an alphabet  $A$  and a symmetric and reflexive dependency relation  $D \subseteq A \times A$  and the corresponding independency relation  $\parallel \subseteq A \times A$  defined as  $\forall a, b \in A, (a \parallel b) \iff (a, b) \notin D$ . A *Mazurkiewicz trace* or, more simply, a *trace*, is an equivalence class for the smallest equivalence relation  $\equiv$  on  $A^*$  which commutes independent letters i.e. for every letters  $a, b$  and every words  $w_1, w_2$ ,

$$a \parallel b \implies w_1 a b w_2 \equiv w_1 b a w_2 .$$

The words in the equivalence class are the *linearizations* of the trace. The trace whose only linearization is the empty word is denoted  $\epsilon$ . All linearizations of a trace  $u$  have the same set of letters and length, denoted respectively  $\text{Alph}(u)$  and  $|u|$ . Given  $B \subseteq A$ , the set of traces such that  $\text{Alph}(u) \subseteq B$  is denoted  $r_{\equiv}^*$  in particular the set of all traces is  $A_{\equiv}^*$ .

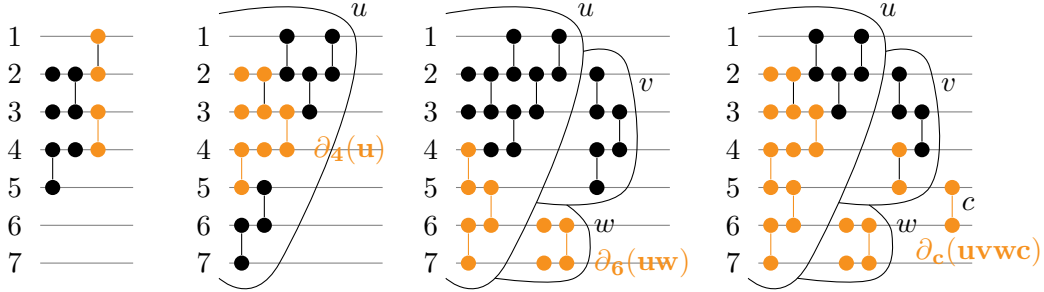


Figure 1: The set processes is  $\{1 \dots 7\}$ . A letter is identified with its domain. Here the domains are either singletons, represented by a single dot, or pairs of contiguous processes, represented by two dots connected with a vertical segment. The trace  $\{2\}\{3\}\{4, 5\}\{2, 3\}\{4\}\{1, 2\}\{3, 4\} = \{4, 5\}\{4\}\{2\}\{3\}\{2, 3\}\{3, 4\}\{1, 2\}$  is represented on the left-hand side. It has two maximal letters  $\{1, 2\}$  and  $\{3, 4\}$  thus is not prime. Center left: process 4 sees only its causal view  $\partial_4(u)$  (in yellow). Center right:  $uvw = uvw$  since  $\text{dom}(v) \cap \text{dom}(w) = \emptyset$ . Both  $uv$  and  $\partial_6(uw)$  (in yellow) are prime prefixes of  $uvw$  and they are parallel. Right:  $uv$  and  $\partial_c(uvwc)$  (in yellow) are parallel.

The concatenation on words naturally extends to traces. Given two traces  $u, v \in A_{\equiv}^*$ , the trace  $uv$  is the equivalence class of any word in  $uv$ . The prefix relation  $\sqsubseteq$  is defined by  $(u \sqsubseteq v \iff \exists w \in A_{\equiv}^*, uw = v)$  . and the suffix relation is defined similarly.

**Maxima, prime traces and parallel traces.** A letter  $a \in A$  is a *maximum* of a trace  $u$  if it is the last letter of one of the linearizations of  $u$ . A trace  $u \in \ell_{\equiv}^*$  is *prime* if it has a unique maximum, denoted  $\text{last}(u)$  and called the last letter of  $u$ . Two prime traces  $u$  and  $v$  are said to be *parallel* if

- neither  $u$  is a prefix of  $v$  nor  $v$  is a prefix of  $u$ ; and
- there is a trace  $w$  such that both  $u$  and  $v$  are prefixes of  $w$ . These notions are illustrated on Fig. 1.

**1.1. Asynchronous automata.** Asynchronous automata are to traces what finite automata are to finite words, as witnessed by Zielonka's theorem [13]. An asynchronous automaton is a collection of automata on finite words, whose transition tables do synchronize on certain actions.

**Definition 1.** An asynchronous automaton on alphabet  $A$  with processes  $\mathbb{P}$  is a tuple  $\mathcal{A} = ((A_p)_{p \in \mathbb{P}}, (Q_p)_{p \in \mathbb{P}}, (i_p)_{p \in \mathbb{P}}, (F_p)_{p \in \mathbb{P}}, \Delta)$  where:

- every process  $p \in \mathbb{P}$  has a set of actions  $A_p$ , a set of states  $Q_p$  and  $i_p \in Q_p$  is the initial state of  $p$  and  $F_p \subseteq Q_p$  its set of final states.
- $A = \bigcup_{p \in \mathbb{P}} A_p$ . For every letter  $a \in A$ , the domain of  $a$  is  $\text{dom}(a) = \{p \in \mathbb{P} \mid a \in A_p\}$  .
- $\Delta$  is a set of transitions of the form  $(a, (q_p, q'_p)_{p \in \text{dom}(a)})$  where  $a \in A$  and  $q_p, q'_p \in Q_p$ . Transitions are deterministic: for every  $a \in A$ , if  $\delta = (a, (q_p, q'_p)_{p \in \text{dom}(a)}) \in \Delta$  and  $\delta' = (a, (q_p, q''_p)_{p \in \text{dom}(a)}) \in \Delta$  then  $\delta = \delta'$  (hence  $\forall p \in \text{dom}(a), q'_p = q''_p$ ).

Such an automaton works asynchronously: each time a letter  $a$  is processed, the states of the processes in  $\text{dom}(a)$  are updated according to the corresponding transition, while the states of other processes do not change. This induces a natural commutation relation  $\mathbb{I}$  on  $A$ : two letters commute iff they have no process in common i.e.

$$(a \mathbb{I} b) \iff (\text{dom}(a) \cap \text{dom}(b) = \emptyset) .$$

The set of *plays* of the automaton  $\mathcal{A}$  is a set of traces denoted  $\text{plays}(\mathcal{A})$  and defined inductively, along with a mapping  $\text{state} : \text{plays}(\mathcal{A}) \rightarrow \prod_{p \in \mathbb{P}} Q_p$ .

- $\epsilon$  is a play and  $\text{state}(\epsilon) = (i_p)_{p \in \mathbb{P}}$ ,
- for every play  $u$  such that  $(\text{state}_p(u))_{p \in \mathbb{P}}$  is defined and  $(a, (\text{state}_p(u), q_p)_{p \in \text{dom}(a)})$  is a transition then  $ua$  is a play and  $\forall p \in \mathbb{P}, \text{state}_p(ua) = \begin{cases} \text{state}_p(u) & \text{if } p \notin \text{dom}(a), \\ q_p & \text{otherwise.} \end{cases}$

For every play  $u$ ,  $\text{state}(u)$  is called the *global state* of  $u$ . The inductive definition of  $\text{state}(u)$  is correct because it is invariant by commutation of independent letters of  $u$ .

**Counting actions of a process.** For every trace  $u$  we can count how many times a process  $p$  has played an action in  $u$ , which we denote  $|u|_p$ . Formally,  $|u|_p$  is first defined for words, as the length of the projection of  $u$  on  $A_p$ , which is invariant by commuting letters. The domain of a trace is defined as

$$\text{dom}(u) = \{p \in \mathbb{P} \mid |u|_p \neq 0\} .$$

**1.2. Strategies with causal memory.** Given an automaton  $\mathcal{A}$ , we want the processes to choose actions which guarantee that every play eventually terminates in a final state.

To take into account the fact that some actions are controllable by processes while some other actions are not, we assume that  $A$  is partitioned in

$$A = A_c \sqcup A_e$$

where  $A_c$  is the set of controllable actions and  $A_e$  the set of (uncontrollable) environment actions. Intuitively, processes cannot prevent their environment to play actions in  $A_e$ , while they can decide whether to block or allow any action in  $A_c$ .

We adopt a modern terminology and call the automaton  $\mathcal{A}$  together with the partition  $A = A_c \sqcup A_e$  a *distributed game*, or even more simply a *game*.

In a distributed game the processes play distributed strategies, which are individual plans of action for each process. The choice of actions by a process  $p$  is dynamic: at every step,  $p$  chooses a new set of controllable actions, depending on its information about the way the play is going on. This information is limited since processes cannot communicate together unless they synchronize on a common action. In that case however they exchange as much information about the play as they want. Finally, the information missing to a process is the set of actions which happened in parallel of its own actions. The information which remains is called the  $p$ -view of the play, it is illustrated on Fig. 1 and defined formally as follows.

**Definition 2 (Views).** For every set of processes  $\mathbb{Q} \subseteq \mathbb{P}$  and trace  $u$ , the  $\mathbb{Q}$ -view of  $u$ , denoted  $\partial_{\mathbb{Q}}(u)$ , is the unique trace such that  $u$  factorizes as  $u = \partial_{\mathbb{Q}}(u) \cdot v$  and  $v$  is the longest suffix of  $u$  such that  $\mathbb{Q} \cap \text{dom}(v) = \emptyset$ . In case  $\mathbb{Q}$  is a singleton  $\{p\}$  the view is denoted  $\partial_p(u)$  and is either empty or prime. For every letter  $a \in A$  we denote  $\partial_a(u) = \partial_{\text{dom}(a)}(u)$ .

Some useful properties of the  $\mathbb{Q}$ -view are:

$$\partial_{\mathbb{Q}}(uv) = \partial_{\mathbb{Q}'}(u) \partial_{\mathbb{Q}}(v) \text{ where } \mathbb{Q}' = \mathbb{Q} \cup \text{dom}(\partial_{\mathbb{Q}}(v)) \quad (1.1)$$

$$(\mathbb{Q} \subseteq \mathbb{Q}') \implies (\partial_{\mathbb{Q}}(u) \sqsubseteq \partial_{\mathbb{Q}'}(u)) . \quad (1.2)$$

We can now define what is a distributed strategy.

**Definition 3** (Distributed strategies, consistent and maximal plays). *Let  $G = (\mathcal{A}, A_c, A_e)$  be a distributed game. A strategy for process  $p$  in  $G$  is a mapping which associates with every play  $u$  a set of actions  $\sigma_p(u)$  such that:*

- all environment actions are allowed:  $A_e \subseteq \sigma_p(u)$ ,
- the decision depends only on the view of the process:  $\sigma_p(u) = \sigma_p(\partial_p(u))$ .

A distributed strategy is a tuple  $\sigma = (\sigma_p)_{p \in \mathbb{P}}$  where each  $\sigma_p$  is a strategy of process  $p$ . A play  $u = a_1 \cdots a_{|u|} \in \text{plays}(\mathcal{A})$  is consistent with  $\sigma$ , or equivalently is a  $\sigma$ -play if:

$$\forall i \in 1 \dots |u|, \forall p \in \text{dom}(a_i), a_i \in \sigma_p(a_1 \cdots a_{i-1}) .$$

A  $\sigma$ -play is maximal if it is not the strict prefix of another  $\sigma$ -play.

Note that a strategy is forced to allow every environment action to be executed at every moment. This may seem to be a huge strategic advantage for the environment. However depending on the current state, not every action can be effectively used in a transition because the transition function is not assumed to be total. So in general not every environment actions can actually occur in a play. In particular it may happen that a process enters a final state with no outgoing transition, where no uncontrollable action can happen.

**Winning games.** Our goal is to synthesize strategies which ensure that the game terminates and all processes are in a final state.

**Definition 4** (Winning strategy). *A strategy  $\sigma$  is winning if the set of  $\sigma$ -plays is finite and in every maximal  $\sigma$ -play  $u$ , every process is in a final state i.e.  $\forall p \in \mathbb{P}, \text{state}_p(u) \in F_p$  .*

The following decision problem is the central motivation for this work:

DISTRIBUTED SYNTHESIS PROBLEM: given a distributed game decide whether there exists a winning strategy.

Whether this problem is decidable has been an open problem for some time [3, 9, 7]. In this paper we show that this problem is undecidable.

## 2. THE BIPARTITE COLORING PROBLEM

The proof of the undecidability of the DISTRIBUTED SYNTHESIS PROBLEM makes use of an intermediary decision problem: In the sequel, we use the notation  $[n]$  for the integer interval

$$[n] = \{0, \dots, (n-1)\} .$$

**Definition 5** (Finite bipartite  $C$ -colorings). *Fix a finite set  $C$  called the set of colours. A finite bipartite  $C$ -coloring is a function  $f : [n] \times [m] \rightarrow C$  where  $n$  and  $m$  are positive integers. The initial and final colors of  $f$  are respectively  $f(0,0)$  and  $f(n-1, m-1)$ . We define three subsets of  $C^2$  called the patterns induced by  $f$ :*

- the squares of  $f$  are all pairs  $\{(f(x,y), f(x+1, y+1)) \mid (x,y) \in [n-1] \times [m-1]\}$ .
- the upper-triangles of  $f$  are all pairs  $\{(f(x,y), f(x+1, y)) \mid (x,y) \in [n-1] \times [m]\}$ .
- the lower-triangles of  $f$  are all pairs  $\{(f(x,y), f(x, y+1)) \mid (x,y) \in [n] \times [m-1]\}$ .

We often use the simpler term *coloring* to refer to a finite bipartite  $C$ -coloring.

The BIPARTITE COLORING PROBLEM asks whether there exists a coloring satisfying some constraints on the initial and final colors and on the induced patterns. A *coloring constraint* is given by two subsets  $C_i, C_f$  of  $C$  called respectively the sets of *allowed initial and final colors*, and three subsets  $S, UT, LT$  of  $C^2$ , called the *forbidden patterns*. A coloring  $f$  *satisfies* the constraint if its minimal color is in  $C_i$ , its maximal color is in  $C_f$  and none of the induced patterns of  $f$  is forbidden: no square belongs to  $S$ , no upper-triangle belongs to  $UT$  and no lower-triangle belongs to  $LT$ .

**Definition 2.1** (BIPARTITE COLORING PROBLEM). Given a finite set of colours  $C$  and a coloring constraint  $(C_i, C_f, S, UT, LT)$ , decide whether there exists a finite bipartite  $C$ -coloring satisfying the constraint.

As an exercise, set  $C = \{=, \neq\}$  and consider the coloring constraint given by  $C_i = \{=\}$ ,  $C_f = \{=, \neq\}$ ,  $UT = LT = \emptyset$  and  $S = \{(=, \neq), (-, =)\}$ . A coloring  $f : [n] \times [m] \rightarrow C$  satisfies this constraint iff the edges  $(x, y)$  colored by  $=$  are exactly those whose both coordinates are equal, i.e.  $x = y$ . Indeed, the square constraint enforces propagation of the  $=$  symbol in both directions: if an edge  $(x, y)$  is colored by  $=$ , then so shall be the edges  $(x - 1, y - 1)$  and  $(x + 1, y + 1)$  (provided these edges belong as well to  $[n] \times [m]$ ). Since the initial position  $(0, 0)$  can only be marked by  $=$  the color  $=$  appear in  $f$  exactly on the edges with both coordinates equal. If we remove  $\neq$  from  $C_f$ , this creates the extra constraint  $n = m$ .

Another example is  $C = \{0, +\}$  and the coloring constraint given by  $C_i = \{0\}$ ,  $C_f = \{0, +\}$ ,  $UT = \{(+, 0), (0, 0)\}$   $LT = \{(0, +)\}$  and  $S = \emptyset$ . A coloring  $f : [n] \times [m] \rightarrow C$  satisfies this constraint iff the edges  $(x, y)$  colored by  $0$  are exactly those whose first coordinate is  $x = 0$ . The upper-triangle constraint  $UT$  prevents the symbol  $0$  from appearing on any edge  $(x, y)$  such that  $x > 0$ . By induction, the lower-triangle constraint  $LT$  enforces the symbol  $0$  to appear on any edge  $(0, y)$ . One can extend this last example on the product alphabet  $\{0, +\}^2$  and enforce a coloring to "detect"  $0$  coordinates in both components  $x$  and  $y$ .

The BIPARTITE COLORING PROBLEM can encode problems far more complicated than these toy examples:

**Theorem 2.2.** *There is an effective reduction from the POST CORRESPONDENCE PROBLEM to the BIPARTITE COLORING PROBLEM.*

The proof of this theorem goes in two steps: first characterize a solution to POST CORRESPONDENCE PROBLEM with local rewriting rules and second express these local rewriting rules as coloring constraints. Remark that the reduction generates coloring constraints which allow only solutions  $f : [n] \times [m] \rightarrow C$  such that  $n = m$  (if any). On purpose we do not impose this constraint in the definition of the BIPARTITE COLORING PROBLEM.

*Proof of Theorem 2.2.* An instance of POST CORRESPONDENCE PROBLEM is a finite collection  $(u_i, v_i)_{i \in I}$  of tiles on an alphabet  $\Sigma$ . Each tile  $(u_i, v_i)$  is a pair of words on the alphabet  $\Sigma$ ,  $u_i$  is the top word and  $v_i$  is the bottom word. The problem is to determine whether there exists a finite sequence of indices in  $I$  such that the concatenation of the tiles produce the same two words on the top and the bottom. Duplicates are allowed: the same index  $i$  can appear several tiles in the sequence. In case such a sequence exists, the instance  $(u_i, v_i)_{i \in I}$  is said to have a solution. Checking whether an instance has a solution is known to be an undecidable problem [11].

The following result characterizes the existence of a solution to an instance of POST CORRESPONDENCE PROBLEM.

**Lemma 1** (Local characterization of a PCP solution). *Let  $(u_i, v_i)_{i \in I}$  be an instance of POST CORRESPONDENCE PROBLEM on the alphabet  $\Sigma$ . Let  $i_0, i_1, \dots, i_k$  and  $j_0, j_1, \dots, j_\ell$  be two finite sequences of indices and  $u = u_{i_0} \dots u_{i_k}$  and  $v = v_{j_0} \dots v_{j_\ell}$  the corresponding tiles concatenations. Denote  $X = [|u|] \times [|v|]$ .*

*The equality  $u = v$  holds if and only if there exists a subset **SameLength** of  $X$  with the following properties.*

- (1) **SameLength** contains  $(|u| - 1, |v| - 1)$ .
- (2) Let  $x = (x_u, x_v) \in \mathbf{SameLength}$  such that  $x_u > 0$  and  $x_v > 0$ . Then  $(x_u - 1, x_v - 1) \in \mathbf{SameLength}$ .
- (3)  $(0, 0)$  belongs to **SameLength** and this is the only element  $x = (x_u, x_v) \in \mathbf{SameLength}$  with either coordinate equal to 0.
- (4) For every  $x = (x_u, x_v) \in \mathbf{SameLength}$  the letter of index  $x_u$  in  $u$  is the same than the letter of index  $x_v$  in  $v$ .

*The equality  $i_0, i_1, \dots, i_k = j_0, j_1, \dots, j_\ell$  holds if and only if there exists a subset **SameTile** of  $X$  with the following properties:*

- (5) **SameTile** contains  $(|u| - 1, |v| - 1)$ .
- (6) Let  $x = (x_u, x_v) \in \mathbf{SameTile}$  such that  $x_u > 0$  and  $x_v > 0$ . Say that  $x$  starts a new tile in  $u$  if  $x_u$  is the first index of a tile in the factorisation  $u = u_{i_0} \dots u_{i_k}$  i.e. if there exists  $0 \leq m \leq k$  such that  $x_u = |u_{i_0} \dots u_{i_m}|$ . Say that  $x$  starts a new tile in  $v$  if the symmetrical condition holds for  $x_v$  and  $v = v_{j_0} \dots v_{j_\ell}$ . If  $x$  starts a new tile in both  $u$  and  $v$  then  $(x_u - 1, x_v - 1) \in \mathbf{SameTile}$ . If  $x$  does not start a new tile in  $u$  then  $(x_u - 1, x_v) \in \mathbf{SameTile}$ . If  $x$  does not start a new tile in  $v$  then  $(x_u, x_v - 1) \in \mathbf{SameTile}$ .
- (7)  $(0, 0)$  belongs to **SameTile**. For every  $(x_u, x_v) \in \mathbf{SameTile}$ ,  $(x_u < |u_{i_0}| \iff x_v < |v_{j_0}|)$ .
- (8) Let  $x = (x_u, x_v) \in \mathbf{SameTile}$  and denote  $i_u(x) \in I$  the index of the tile appearing at index  $x_u$  in the factorisation  $u = u_{i_0} \dots u_{i_k}$ . Define  $i_v(x) \in I$  symmetrically with respect to the factorisation  $v = v_{j_0} \dots v_{j_\ell}$ . Then  $i_u(x)$  and  $i_v(x)$  are equal.

*Proof.* The proof is elementary.

We prove the first part of the lemma: the characterisation of  $u = v$ . For the direct implication, assume  $u = v$ . Then the set  $\mathbf{SameLength} = \{(x, x) \mid 0 \leq x < |u| = |v|\}$  has properties 1. to 4. In the opposite direction, assume there exists  $\mathbf{SameLength} \subseteq X$  with properties 1. to 4. We show first that  $|u| = |v|$ . Assume w.l.o.g. that  $|u| = \min(|u|, |v|)$ . A simple induction shows that for every  $0 < i \leq \min(|u|, |v|)$ ,  $\mathbf{SameLength}$  contains  $(|u| - i, |v| - i)$ : the case  $i = 1$  follows from property 1. and the induction step from property 2. Then  $(0, |v| - |u|) \in \mathbf{SameLength}$  thus by property 3.  $|u| = |v|$ . Finally  $u = v$  follows by property 4.

We now prove the second part of the lemma: the characterisation of  $i_0, i_1, \dots, i_k = j_0, j_1, \dots, j_\ell$ . The proof relies on the subset  $A \subseteq X$  containing all pairs  $(x, y) \in X$  such that the number of tiles appearing *after* position  $x$  in  $u$  is the same that the number of tiles appearing *after* position  $y$  in  $v$ . Formally,

$$A = \{(x_u, x_v) \in X \mid \exists m \in 0, \dots, \min(k, \ell) \\ |u_{i_0} \dots u_{i_{k-m-1}}| \leq x_u < |u_{i_0} \dots u_{i_{k-m}}| \text{ and } |v_{j_0} \dots v_{j_{\ell-m-1}}| \leq x_v < |v_{j_0} \dots v_{j_{\ell-m}}|\}$$

For example,  $A$  contains  $(|u| - 1, |v| - 1)$ . And  $A$  contains  $(0, 0)$  iff  $k = \ell$ . Following its definition parametrized by  $m \in 0 \dots \min(k, \ell)$ ,  $A$  is partitioned in  $A_0, \dots, A_{\min(k, \ell)}$ .

For the direct implication, Assume  $i_0, i_1, \dots, i_k = j_0, j_1, \dots, j_\ell$  (hence in particular  $k = \ell$ ). Then the set **SameTile** =  $A$  clearly satisfies all properties (5) to (8).

Conversely, assume **SameTile**  $\subseteq X$  satisfies all properties (5) to (8), and show that  $i_0, i_1, \dots, i_k = j_0, j_1, \dots, j_\ell$ .

We show first that **SameTile** contains all elements in  $A$ . Assume by contradiction that some element  $x$  in  $A$  is missing from **SameTile**, and assume  $x$  is maximal for the lexicographic ordering in  $A \setminus \text{SameTile}$ . Let  $m$  such that  $x \in A_m$ . The set  $A_m$  is the cartesian product of two intervals, hence it has a maximum, namely  $x_m = (|u_{i_0} \dots u_{i_{k-m}}| - 1, |u_{i_0} \dots u_{i_{\ell-m}}| - 1)$ . There are two cases, depending whether  $x = x_m$  or not. Assume first  $x = x_m$ . According to property (5), the maximal element  $(|u| - 1, |v| - 1)$  in  $A$  belongs to **SameTile** thus  $x = x_m \neq (|u| - 1, |v| - 1)$ . Since  $(|u| - 1, |v| - 1) \in A_0$  then  $m > 0$ . As a consequence,  $(x_u + 1, x_v + 1)$  is the minimal element in  $A_{m-1}$  and by induction hypothesis,  $(x_u + 1, x_v + 1) \in \text{SameTile}$ . Moreover,  $(x_u + 1, x_v + 1)$  starts a new tile in both  $u$  and  $v$  and by property (6), we get  $x \in \text{SameTile}$  as well. The second case is when  $x$  is not maximum in  $A_m$ . Since  $A_m$  is the cartesian product of two intervals, either  $(x_u + 1, x_v)$  or  $(x_u, x_v + 1)$  is in  $A_m$  as well, and  $x$  does not start a new tile on the corresponding coordinate. We conclude that  $x \in \text{SameTile}$  using the induction hypothesis and property (6).

Now we show  $k = \ell$ . Assume w.l.o.g.  $k = \min(k, \ell)$ . Since **SameTile** contains all  $A$  then in particular **SameTile** contains  $(0, |v_{j_0} \dots v_{j_{\ell-k}}| - 1)$  (take  $m = k$  in the definition of  $A$ ). According to property (7),  $|v_{j_0} \dots v_{j_{\ell-k}}| \leq |v_{j_0}|$  thus  $k = \ell$ .

Using property (8), we conclude that  $i_0, i_1, \dots, i_k = j_0, j_1, \dots, j_\ell$ . □

We now define an instance of BIPARTITE COLORING PROBLEM encoding the constraints in Lemma 1. We start with defining the set of colors. Set  $Q_u$  be the finite subset of  $\Sigma \times I \times \mathbb{N} \times \{0, 1\}^2$  containing all  $(a, i, m, b_0, b_1)$  where  $m < |u_i|$  and  $a$  is the  $m$ -th letter of  $u_i$ . Given  $q = (a, i, m, b_0, b_1) \in Q_u$ ,  $a$  is called the letter of  $q$ ,  $i$  its tile index,  $m$  its tile position,  $b_0$  its initial letter flag and  $b_1$  its initial tile flag. The tile position is maximal in  $q$  if it is equal to  $|u_i| - 1$ . Define  $Q_v$  similarly with respect to the bottom parts of the tiles  $(v_i)_{i \in I}$ . Let  $C$  be the set of colors

$$Q_u \times Q_v \times \{\text{SameLength}, -\} \times \{\text{SameTile}, -\} .$$

We define a constraint coloring problem reflecting conditions (1) to (8) of the previous lemma. We use  $*$  as a wildcard character which can be replaced by any symbol.

First, we set constraints which enforce the  $Q_u$ -component of  $f(x_u, x_v)$  to depend only on  $x_u$  and the  $Q_v$ -component of  $f(x_u, x_v)$  to depend only on  $x_v$ . For that we include in the set of forbidden lower-triangle  $LT$  all pairs  $(q_u, *, *, *) (q'_u, *, *, *)$  with  $q_u \neq q'_u$ . A simple induction shows that if a coloring  $f : [n] \times [m] \rightarrow C$  does not induce such lower-triangles pattern, then there exists  $f_u : [n] \rightarrow Q_u$  such that for every  $x = (x_u, x_v) \in [n] \times [m]$ , the first component of  $f(x)$  is  $f_u(x_u)$ . Symmetrically, we add to the set of forbidden upper-triangles  $UT$  all pairs  $(*, q_v, *, *) (*, q'_v, *, *)$  with  $q_v \neq q'_v$  so that the  $Q_v$ -component of  $f$  is induced by  $f_v : [m] \rightarrow Q_v$ .

Denote  $u \in \Sigma^n$  to be the word of length  $n$  whose  $x_u$ -th letter is the first component of  $f_u(x_u) \in Q_u$ . Define  $v \in \Sigma^m$  symmetrically with respect to  $f_v$ .

Second, we want to ensure that  $u$  is a concatenation of upper tiles of the PCP instance, i.e. there exists  $i_0, i_1, \dots, i_k$  such that  $u = u_{i_0} \dots u_{i_k}$ . This regular constraint can be checked



by a deterministic automaton on  $Q_u$ , which we implement using the coloring constraints. First, we require that in every initial color  $(q, *, *, *)$ , the tile position in  $q$  is 0. Moreover, we put constraints on all upper-triangles  $(q, *, *, *)$ ,  $(q', *, *, *)$ : either the tile position is maximal in  $q$  and has value 0 in  $q'$ , or the tile index does not change between  $q$  and  $q'$  while the tile position is incremented of exactly one unit. Finally, we require that in every final color  $(q, *, *, *)$ , the tile position in  $q$  is maximal. With a symmetric construction, we ensure the symmetric constraint on  $v$ : there exists  $j_0, j_1, \dots, j_\ell$  such that  $v = v_{j_0} \cdots v_{j_\ell}$ .

Third, we want that the initial letter flag and the initial tile flag of  $f_u(x_u)$  indicate respectively whether or not  $x_u = 0$  and whether or not  $x_u < |u_{i_0}|$ . We require that all initial colors have both flags set to 1. We forbid any upper triangle  $(q_u, *, *, *)$ ,  $(q'_u, *, *, *)$  where  $q'_u$  has the initial letter flag set to 1. We forbid any upper triangle  $(q_u, *, *, *)$ ,  $(q'_u, *, *, *)$  where  $q'_u$  has the initial tile flag set to 1 unless  $q_u$  has also the initial tile flag set to 1 and the tile position in  $q_u$  is not maximal. A symmetric construction ensures that the initial letter and tile flags of  $f_v(u_v)$  indicate respectively whether or not  $u_v = 0$  and  $u_v < |v_{j_0}|$ .

This way we have a natural correspondence between finite bipartite coloring satisfying the above constraints and pairs of factorisations  $u = u_{i_0} \cdots u_{i_k}$  and  $v = v_{j_0} \cdots v_{j_\ell}$  like in Lemma 1. Using this correspondence, we can rephrase the eight conditions in Lemma 1 as eight coloring constraints on a set of colours  $C' \subseteq C$ .

- (1) Every final colour has type  $(*, *, \mathbf{SameLength}, *)$ .
- (2) Forbid all squares  $(*, *, -, *)$ ,  $(*, *, \mathbf{SameLength}, *)$ .
- (3) Every initial colour has type  $(q_u, q_v, \mathbf{SameLength}, *)$  with both initial letter flags in  $q_u$  and  $q_v$  set to 1.
- (4) Remove from the alphabet  $C$  any color  $(q_u, q_v, \mathbf{SameLength}, *)$  such that the letters in  $q_u$  and  $q_v$  (i.e. their projections on  $\Sigma$ ) are different.
- (5) Every final colour has type  $(*, *, *, \mathbf{SameTile})$ .
- (6) Forbid all squares  $(*, *, *, -)$ ,  $(q'_u, q'_v, *, \mathbf{SameTile})$  such that both tile positions in  $q'_u$  and  $q'_v$  are 0. Forbid all upper-triangles  $(*, *, *, -)$ ,  $(q'_u, *, *, \mathbf{SameTile})$  unless the tile position in  $q'_u$  is 0. Forbid all lower-triangles  $(*, *, *, -)$ ,  $(*, q'_v, *, \mathbf{SameTile})$  unless the tile position in  $q'_v$  is 0.
- (7) Every initial colour has type  $(q_u, q_v, *, \mathbf{SameTile})$  with both initial tile flags in  $q_u$  and  $q_v$  set to 1. Remove from the alphabet  $C$  any color  $(q_u, q_v, *, \mathbf{SameTile})$  such that the initial tile flags in  $q_u$  and  $q_v$  are different.
- (8) Remove from the alphabet  $C$  any color  $(q_u, q_v, *, \mathbf{SameTile})$  such that the tile indices in  $q_u$  and  $q_v$  are different.

Applying conditions, 4, 7 and 8, we get a set of colours  $C' \subset C$  on which the constraint coloring problem is defined. The initial condition  $C_i$  is defined by conditions 3 and 7: this is the set of colours in  $C'$  of type  $(q_u, q_v, \mathbf{SameLength}, \mathbf{SameTile})$  with both initial letter flags and initial tile flags set to 1 in both  $q_u$  and  $q_v$ . The final condition  $C_f$  is defined by conditions 1 and 5: any colour in  $C'$  of type  $(*, *, \mathbf{SameLength}, \mathbf{SameTile})$  is final. The set of forbidden patterns is obtained by taking the union of all forbidden squares, upper-triangles and lower-triangles mentioned above.

According to Lemma 1, there exists a coloring satisfies the constraints if and only if the PCP instance has a solution. This terminates the proof of Theorem 2.2.  $\square$

A direct consequence of Theorem 2.2 is that the BIPARTITE COLORING PROBLEM is undecidable.

### 3. UNDECIDABILITY OF THE DISTRIBUTED SYNTHESIS PROBLEM WITH 6 PROCESSES

The main result of this section is:

**Theorem 3.1.** *There is an effective reduction from the BIPARTITE COLORING PROBLEM to the DISTRIBUTED SYNTHESIS PROBLEM with six processes.*

We show how to effectively transform an instance  $(C_i, C_f, S, UT, LT)$  of the BIPARTITE COLORING PROBLEM on a set of colours  $C$  into a distributed game  $G$  such that:

**Lemma 3.2.** *There is a winning strategy in  $G$  if and only if there is a finite bipartite coloring satisfying the constraints  $(C_i, C_f, S, UT, LT)$ .*

In the rest of the section we describe the construction of the game  $G$  and then prove Lemma 3.2.

**3.1. Turning a coloring problem into a game.** The game  $G$  has six processes divided into two pools: the top pool  $T_0, T_1, T_2$  and the bottom pool  $B_0, B_1, B_2$ . The transitions of these processes are quite similar, thus we often use the notation  $X_\ell$  to designate any of the six processes, a generic notation where  $X \in \{T, B\}$  and  $\ell \in \{0, 1, 2\}$ . For every process  $X_\ell$  we denote  $\text{next}(X_\ell)$  the process  $X_{(\ell+1) \bmod 3}$  and we denote  $\text{prev}(X_\ell)$  the process  $X_{(\ell-1) \bmod 3}$ . For example,  $\text{next}(T_0) = T_1$  and  $\text{prev}(B_0) = B_2$ .

**Uninterrupted plays.** are those plays where only the following actions are used:

- There are controllable actions called *increments*, which synchronize two processes of the same pool. In each pool  $X \in \{T, B\}$  there are three increment actions  $\mathbb{I}_{X,0}, \mathbb{I}_{X,1}$  and  $\mathbb{I}_{X,2}$ . Every increment  $\mathbb{I}_{X,\ell}$  is controllable and is shared between processes  $X_\ell$  and  $\text{next}(X_\ell)$ .
- Every process  $X_\ell$  has also a controllable local action  $\text{END}_{X,\ell}$ .
- There is a controllable action  $\text{WIN}$  synchronizing all six processes.

The transitions are defined so that a maximal uninterrupted play look like the one on Figure 2: it is the parallel product of two plays  $(\mathbb{I}_{T,0}\mathbb{I}_{T,1}\mathbb{I}_{T,2})^n$  and  $(\mathbb{I}_{B,0}\mathbb{I}_{B,1}\mathbb{I}_{B,2})^m$  for some positive integers  $n, m$ , followed by the six local  $\text{END}$  actions and finally the global action  $\text{WIN}$ .

Since all the actions in an uninterrupted play are controllable, the number of rounds performed before playing actions  $\text{END}$  is determined by the players and is not influenced by the environment. Intuitively, when the number of rounds of the top and bottom pools are respectively  $n$  and  $m$ , the players claim to have a solution  $f : [n] \times [m] \rightarrow C$  to the constrained coloring problem.

In each pool  $X \in \{T, B\}$ , increments have to be played in a fixed order:

$$\mathbb{I}_{X,0}, \mathbb{I}_{X,1}, \mathbb{I}_{X,2}, \mathbb{I}_{X,0}, \mathbb{I}_{X,1}, \dots$$

Remark that no two increment actions of the same pool can commute, thus an interrupted play  $(\mathbb{I}_{X,0}\mathbb{I}_{X,1}\mathbb{I}_{X,2})^*$  in a pool  $X$  has a single linearization. Each subword  $\mathbb{I}_{X,0}, \mathbb{I}_{X,1}, \mathbb{I}_{X,2}$  is called a *round*. This pattern is enforced by maintaining in the state space of every process  $X_\ell$  a flag indicating whether or not the process is allowed to synchronize with  $\text{next}(X_\ell)$ . The flag is initially set to 1 for  $X_0$  and to 0 for  $X_1$  and  $X_2$ , and is toggled every time the process performs an increment.

The only way for processes to win is a transition on the global action  $\text{WIN}$ , in which case they immediately and definitively enter a final state, and no further transition can occur after that. A transition on  $\text{WIN}$  can be triggered once all six processes have performed a local transition on their  $\text{END}$  action.

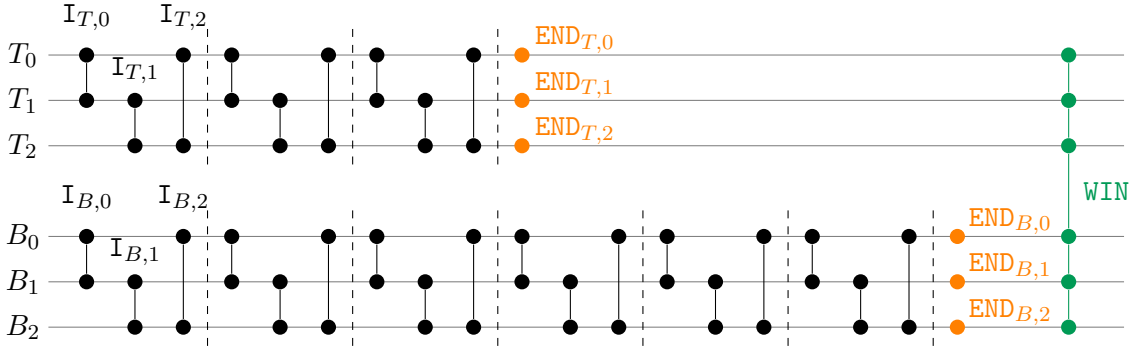


Figure 2: A maximal uninterrupted play where the top pool plays 3 rounds and the bottom pool plays 6 rounds. Rounds are separated by dashes. All actions are controllable.

The transition on the local action  $\text{END}_{X,\ell}$  for process  $X_\ell$  is available if and only if the process has played at least one increment action and has finished the current round. For  $X_0$  and  $X_2$  (resp. for  $X_1$ ) it means that the last increment was  $\text{I}_{X_2}$  (resp. was  $\text{I}_{X_1}$ ). This constraint can easily be encoded with a flag in the state space of the process.

A play stays uninterrupted, like the one on Figure 2, when all actions are controllable and the environment does not play at all, in which case the game is easy to win. But in general, a winning strategy should also react correctly to uncontrollable actions of the environment called *checks*.

**Interrupting plays by colour checks.** The environment has the ability to interrupt a play by triggering some uncontrollable *check actions*. This is represented on Figure 3. Intuitively, when the environment interrupts the play after  $x$  rounds of the top pool and  $y$  rounds of the bottom pool, the players are asked to pick the colour  $f(x, y) \in C$  of the edge  $(x, y)$  of the solution to the coloring problem. There are rules which enforce the players to define that way a coloring  $f : [n] \times [m] \rightarrow C$  which satisfies the constraints, where  $n$  and  $m$  are respectively the number of rounds of the top and bottom pools before playing END. If the strategy of the players makes them cheat or describe a coloring  $f$  which does not satisfy the constraints, the environment can trigger one or two checks which make the players lose.

For every index  $\ell \in \{0, 1, 2\}$ , there is an uncontrollable action  $\text{CHECK}_\ell$  which synchronizes the two processes  $T_\ell$  and  $B_\ell$ . A transition on action  $\text{CHECK}_\ell$  is possible if and only if the last action of both processes is either an increment or the action END. In particular, both processes should have played at least one increment. After the check, the only possible transitions rely on controllable actions  $\text{ANSWER}_{\ell,c}$  synchronizing  $T_\ell$  and  $B_\ell$  and indexed by colours  $c \in C$ . The answer is stored in the state space of the processes.

After that, there are only two possible outcomes: either the WIN or the LOSE action synchronizes all processes, and then no further transition is available. There is no extra condition needed to execute a global transition on WIN: as soon as at least one pair of processes has performed an answer the transition on WIN is available.

**Five ways to lose.** Considering the non-deterministic environment as adversarial, it prefers transitions on LOSE rather than those on WIN. Losing can occur in five different ways, corresponding to the five constraints in the definition of the BIPARTITE COLORING PROBLEM.

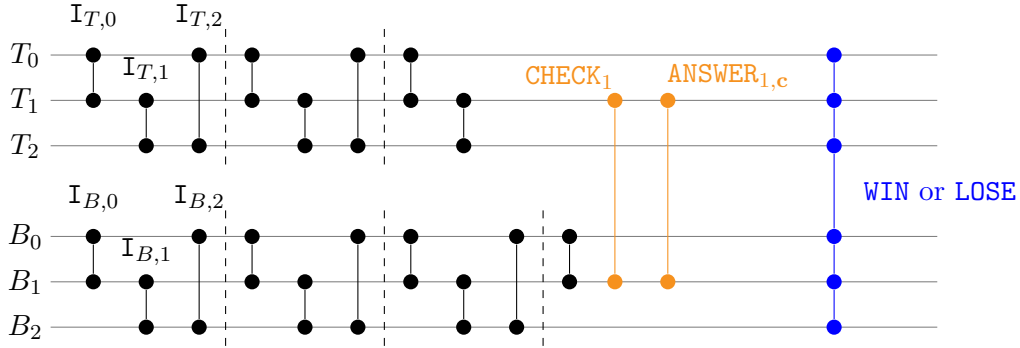


Figure 3: An uncontrollable check of processes  $T_1$  and  $B_1$  after 3 rounds of the top pool and 4 rounds of the bottom pool. Intuitively, the environment is asking to processes  $T_1$  and  $B_1$  "what is the value of  $f(3, 4)$ , where  $f : [n] \times [m] \rightarrow C$  is your solution to the constrained coloring problem?". The check is followed by the controllable action  $\text{ANSWER}_{1,c}$  of the same two processes, parametrized by a colour  $c \in C$ . This way the processes claim that  $f(3, 4) = c$ . The answer is followed by either the WIN or the LOSE action, which terminates the game.

The conditions for losing rely on the definition of the *round index* of a play  $u$  for a process  $X_\ell$ , defined as  $(h - 1)/2$  where  $h$  is the number of increments played by  $X_\ell$  in  $u$ . the round index is denoted  $R_{X_\ell}(u)$ . A transition on LOSE is possible in case one of the following conditions holds:

- (a) if both processes  $T_\ell$  and  $B_\ell$  are checked during round 0 and their answer is a colour  $c$  which is not initial (i.e.  $c \notin C_i$ ). The condition "being in round 0" can be stored in the state space of the processes, and used to allow or not this transition; or
- (b) if both processes  $T_\ell$  and  $B_\ell$  have played END and their answer is a colour  $c$  which is not final (i.e.  $c \notin C_f$ ).

There are three other ways to lose, which require that two checks occur in parallel, for two different pairs of processes indexed by  $\ell$  and  $\ell'$  respectively. This leads to two parallel answers corresponding respectively to colours  $c$  and  $c'$ .

The conditions for LOSE are based on a comparison of  $c$  and  $c'$  as well as the respective rounds of the processes. For each pool  $X \in \{T, B\}$  and every play  $u$  there are two possibilities. W.l.o.g., assume we have chosen  $\ell$  and  $\ell'$  so that  $\ell' < \ell$ . Either both processes  $X_\ell$  and  $X_{\ell'}$  are in the same round (meaning they have the same round index in  $u$ ) or process  $X_{\ell'}$  is one round ahead of  $X_\ell$  (meaning that  $R_{X_{\ell'}}(u) = 1 + R_{X_\ell}(u)$ ). No other case may occur because every uninterrupted play in pool  $X$  is a prefix of a word in  $(I_{X,0}I_{X,1}I_{X,2})^* \text{END}$  and  $\ell' < \ell$ .

A transition on LOSE is possible in case one of the following conditions holds:

- (c) in both pools both processes are in the same round but the answers are different (i.e.  $c \neq c'$ ); or
- (d) in both pools process  $X_{\ell'}$  is one round ahead of  $X_\ell$  and the pair of answers  $(c, c')$  is a forbidden square (i.e.  $(c, c') \in S$ ); or
- (e) in the top pool process  $T_{\ell'}$  is one round ahead of  $T_\ell$  while in the bottom pool both processes are in the same round and the pair of answers  $(c, c')$  is a forbidden upper-triangle (i.e.  $(c, c') \in UT$ ); or

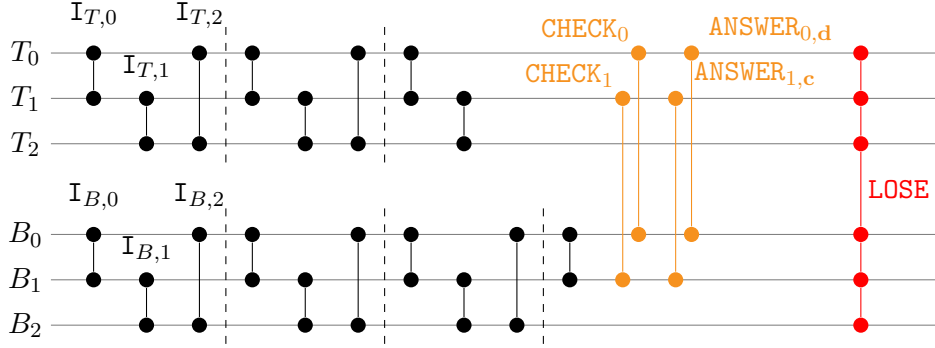


Figure 4: One of five ways to lose: two checks occur in parallel. In both pools, the two checks occur during the same round. The two pairs of processes give (in parallel as well) two different answers  $c \neq d$ . This allows the environment to trigger the LOSE action. Intuitively, the players have cheated since they gave two different answers to the question "what is the value of  $f(3, 4)$ " ?

f. in the top pool both processes are in the same round while in the bottom pool process  $B'_\ell$  is one round ahead of  $B_\ell$  and the pair of answers  $(c, c')$  is a forbidden lower-triangle (i.e.  $(c, c') \in LT$ ).

Remark that the condition "being in the same round" can be easily implemented in the transition table. For that, it is enough that each process keeps track in its state space of the number of its own increments, modulo 4. From that counter, one can derive the parity of the round index and compare the respective parities for the two processes  $X_\ell$  and  $X_{\ell'}$ .

Condition (c) is illustrated on Figure 4 while condition (d) is illustrated on Figure 5.

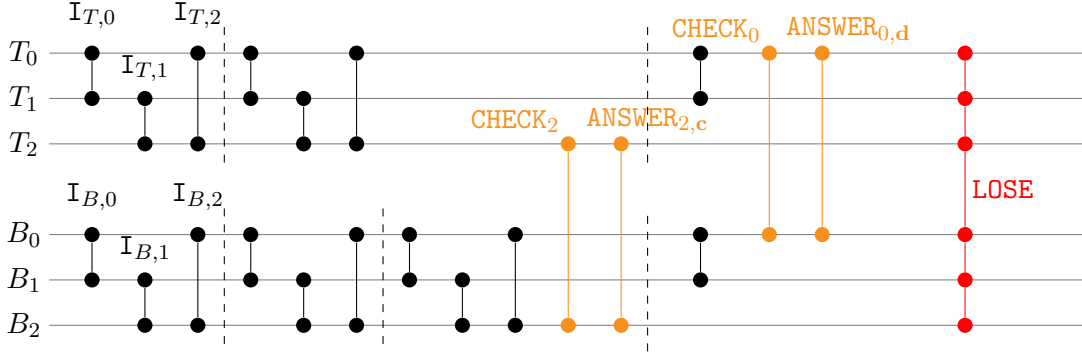


Figure 5: Another way to lose: two checks occur in parallel.  $CHECK_2$  occurs in round 2 of pool  $T$  and round 3 of pool  $B$  thus  $T_2, B_2$  are assumed to answer with  $d = f(2, 3)$ . In parallel,  $CHECK_0$  occurs in round 3 of pool  $T$  and round 4 of pool  $B$ . Hence  $T_0, B_0$  are assumed to answer with  $d = f(3, 4)$ . We assume here that the pairs of answers  $(c, d)$  is a forbidden square, which allows the environment to trigger the LOSE action.

**3.2. Proof of Lemma 3.2.** Lemma 3.2 is an equivalence. We start with the converse implication. Assume that there is a finite bipartite coloring  $f : [n] \times [m] \rightarrow C$  satisfying the constraints  $(C_i, C_f, S, UT, LT)$ . A winning strategy for processes of the top (resp. bottom) pool, as long as they are not interrupted by a check, consists in playing  $n$  rounds (resp.  $m$  rounds) of increments, followed by the END action. If the environment triggers a check on  $T_\ell$  and  $B_\ell$  after some play  $u$ , the processes answer  $f(x, y)$ , where  $x$  (resp.  $y$ ) is the round index of  $T_\ell$  (resp. of  $B_\ell$ ) in  $u$ . This information is available to both processes since they share the same causal past right after the check. Since  $f$  satisfies the constraint, the environment cannot trigger a transition on LOSE. Condition a) cannot occur because all the answers when both processes are in round 0 are equal to  $f(0, 0) \in C_i$ . Condition b) cannot occur because all the answers when both processes are in round  $(n, m)$  are equal to  $f(n, m) \in C_f$ . Condition c) cannot occur because the answers only depends on the rounds, independently of the identities of the processes and the number of increments they have performed in the current round. Condition d, e and f) cannot occur because no pattern induced by  $f$  is forbidden.

We now prove the direct implication of Lemma 3.2. Assume that processes have a winning strategy  $\sigma$ . The first step is to define a finite bipartite  $C$ -coloring  $f$ . Let  $n$  (resp.  $m$ ) be round index of  $T_1$  (resp. of  $B_1$ ) in the maximal uninterrupted play consistent with  $\sigma$  (which exists since  $\sigma$  is winning). For every  $(x, y) \in [n] \times [m]$  denote

$$\begin{aligned} u_{T,x} &= (\mathbb{I}_{T,0}\mathbb{I}_{T,1}\mathbb{I}_{T,2})^x \mathbb{I}_{T,0}\mathbb{I}_{T,1} \\ u_{B,y} &= (\mathbb{I}_{B,0}\mathbb{I}_{B,1}\mathbb{I}_{B,2})^y \mathbb{I}_{B,0}\mathbb{I}_{B,1} \\ u_{x,y} &\text{ the parallel product of } u_{T,x} \text{ and } u_{B,y} \end{aligned}$$

and let  $f(x, y)$  be the answer of  $T_1, B_1$  after a check on the play  $u_{x,y}$  i.e.

$$f(x, y) = c \text{ such that } \text{ANSWER}_{1,c} \in \sigma(u_{x,y}\text{CHECK}_1) .$$

This is well defined: since  $\sigma$  is winning it creates no deadlock in a non-final state, thus at least one answer is played by  $T_1, B_1$  after  $u_{x,y}\text{CHECK}_1$ .

We show that  $f$  satisfies the five constraints  $(C_i, C_f, S, UT, LT)$ .

**Constraint  $C_i$  on initial colors.** Since  $\sigma$  is winning, no transition on action LOSE is available after the check  $\sigma(u_{0,0}\text{CHECK}_1)$ . Thus according to condition a) above,  $f(0, 0)$  is an initial color.

**Constraint  $C_f$  on final colors.** We consider the checks

$$\begin{aligned} p_1 &= u_{n,m}\text{CHECK}_1 \\ p_2 &= u_{n,m}\mathbb{I}_{T,2}\mathbb{I}_{B,2}\text{CHECK}_2 \\ p_3 &= u_{n,m}\text{END}_{T,1}\text{END}_{B,1}\text{CHECK}_1 . \end{aligned}$$

By definition of  $f$ , the answer to  $p_1$  is  $f(n, m)$ . Denote respectively  $c_2$  and  $c_3$  the answers to  $p_2$  and  $p_3$ . Remark that  $p_1$  and  $p_2$  can occur in parallel, because processes  $T_1, B_1$  do not play in  $\mathbb{I}_{T,2}\mathbb{I}_{B,2}\text{CHECK}_2$ . For similar reasons, also  $p_2$  and  $p_3$  can occur in parallel. Moreover, in all these three plays, process  $T_1$  and  $T_2$  are in the same round, and processes  $B_1$  and  $B_2$  are also in the same round (remember that END actions are not accounted for when computing the round index). Since  $\sigma$  is winning, condition c) will neither be satisfied by the pair of parallel checks  $p_1, p_2$  nor by the pair  $p_2, p_3$ , thus  $f(n, m) = c_2 = c_3$ . Finally, after the

check  $p_3$ , both processes  $T_1, B_1$  are in the state **End** and they answer  $f(n, m)$ . According to condition b) above,  $f(n, m)$  is a final color.

**Constraint  $S$  on forbidden squares.** Let  $x, y \in [n - 1] \times [m - 1]$ . We consider four possible checks

$$\begin{aligned} p_1 &= u_{x,y}\text{CHECK}_1 \\ p_2 &= u_{x,y}\mathbb{I}_{T,2}\mathbb{I}_{B,2}\text{CHECK}_2 \\ p'_0 &= u_{x,y}\mathbb{I}_{T,2}\mathbb{I}_{B,2}\mathbb{I}_{T,0}\mathbb{I}_{B,0}\text{CHECK}_0 \\ p'_1 &= u_{x+1,y+1}\text{CHECK}_1 . \end{aligned}$$

By definition of  $f$ , the answers to  $p_1$  and  $p'_1$  are respectively  $f(x, y)$  and  $f(x + 1, y + 1)$ . Denote  $c$  the answer to  $p_2$  and  $d$  the answer to  $p'_0$ . The four checks are illustrated on Figure 6.

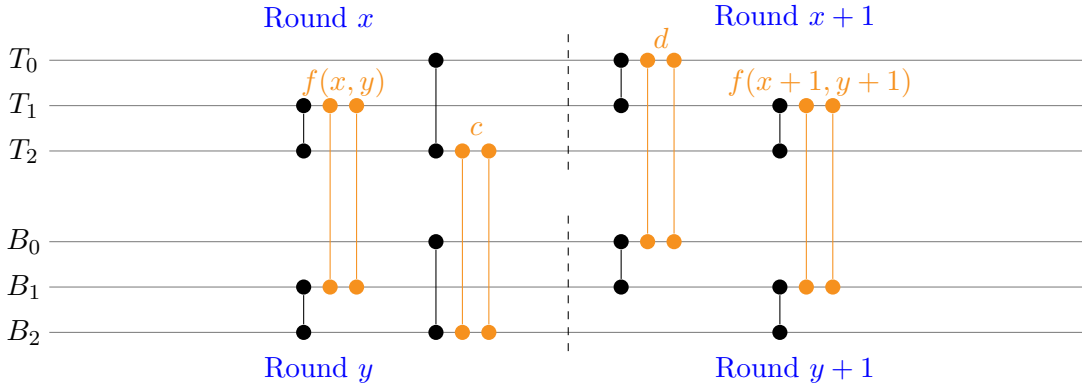


Figure 6: The four checks used to enforce the constraint on forbidden squares, and their answers. These four checks cannot happen altogether in the same play, but any pair of two successive checks can happen in parallel, in the same play.

The two checks  $p_1$  and  $p_2$  can occur in parallel during the same round, thus according to condition c),  $f(x, y) = c$ . The two checks  $p_2$  and  $p'_0$  can occur in parallel and the check  $p'_0$  is one round ahead of the check  $p_2$ , with respect to both pools. Since  $\sigma$  is winning, condition d) cannot happen hence  $(c, d)$  is not a forbidden square. The two checks  $p'_0$  and  $p'_1$  may occur in parallel during the same round, thus according to condition c),  $d = f(x + 1, y + 1)$ . Finally  $(f(x, y), f(x + 1, y + 1))$  is not a forbidden square.

**Constraint  $UT$  on forbidden upper-triangles.** Let  $x, y \in [n - 1] \times [m]$ . We consider the plays

$$\begin{aligned} p_1 &= u_{x,y}\text{CHECK}_1 \\ p_2 &= u_{x,y}\mathbb{I}_{T,2}\mathbb{I}_{B,2}\text{CHECK}_2 \\ p'_0 &= u_{x,y}\mathbb{I}_{T,2}\mathbb{I}_{B,2}\mathbb{I}_{T,0}\text{CHECK}_0 \\ p'_1 &= u_{x+1,y}\text{CHECK}_1 . \end{aligned}$$

By definition of  $f$ , the answers to  $p_1$  and  $p'_1$  are respectively  $f(x, y)$  and  $f(x, y + 1)$ . Denote  $c$  the answer to  $p_2$  and  $d$  the answer to  $p'_0$ . The four checks are illustrated on Figure 7.

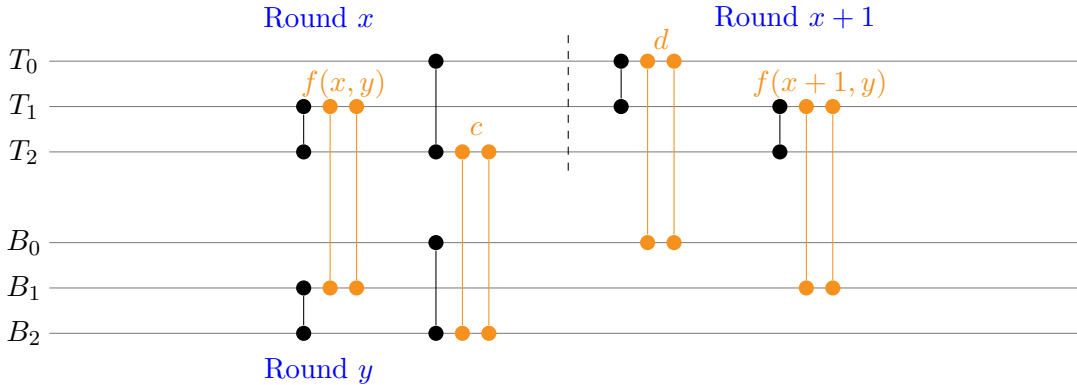


Figure 7: The four checks used to enforce the constraint on forbidden upper-triangles, and their answers. These four checks cannot happen altogether in the same play, but any pair of two successive checks can happen in parallel, in the same play.

The two checks  $p_1$  and  $p_2$  can occur in parallel during the same round thus according to c),  $f(x, y) = c$ . The two checks  $p_2$  and  $p'_0$  can occur in parallel, with process  $T_0$  one round ahead of  $T_1$  and processes  $B_0$  and  $B_1$  in the same round thus according to e),  $(c, d)$  is *not* a forbidden upper-triangle. The two checks  $p'_0$  and  $p'_1$  can occur in parallel during the same round thus according to c),  $d = f(x + 1, y)$ . Finally,  $(f(x, y), f(x + 1, y))$  is not a forbidden upper-triangle.

**Constraint  $LT$  on forbidden lower-triangles.** This case is symmetric with the  $UT$  case above.

Finally,  $f$  is a finite bipartite coloring satisfying all six constraints. This terminates the proof of the direct implication of Lemma 3.2.  $\square$

## CONCLUSION

We have proved that the DISTRIBUTED SYNTHESIS PROBLEM is undecidable for six processes, thus closing an open problem which was first raised in [3] and has afterwards been shown decidable in several special cases [6, 4, 5]. The proof relies on a reduction to the POST CORRESPONDENCE PROBLEM via an intermediary decision problem called the BIPARTITE COLORING PROBLEM. A direct reduction to the POST CORRESPONDENCE PROBLEM is of course also possible, but to our opinion it gives a less clear picture of the reasons for the undecidability. Finally, observe that the DISTRIBUTED SYNTHESIS PROBLEM is known to be decidable for four processes [5]. The case of five processes remains an open question.

## REFERENCES

- [1] Volker Diekert and Grzegorz Rozenberg. *The Book of Traces*. World Scientific, 1995. URL: <https://books.google.co.uk/books?id=vNFLOE2pjuAC>.
- [2] Bernd Finkbeiner and Sven Schewe. Uniform distributed synthesis. In *Logic in Computer Science, 2005. LICS 2005. Proceedings. 20th Annual IEEE Symposium on*, pages 321–330. IEEE, 2005.



- [3] Paul Gastin, Benjamin Lerman, and Marc Zeitoun. Distributed games with causal memory are decidable for series-parallel systems. In *FSTTCS 2004: Foundations of Software Technology and Theoretical Computer Science, 24th International Conference, Chennai, India, December 16-18, 2004, Proceedings*, pages 275–286, 2004. URL: [http://dx.doi.org/10.1007/978-3-540-30538-5\\_23](http://dx.doi.org/10.1007/978-3-540-30538-5_23), doi:10.1007/978-3-540-30538-5\_23.
- [4] Blaise Genest, Hugo Gimbert, Anca Muscholl, and Igor Walukiewicz. Asynchronous games over tree architectures. In *Automata, Languages, and Programming - 40th International Colloquium, ICALP 2013, Riga, Latvia, July 8-12, 2013, Proceedings, Part II*, pages 275–286, 2013. URL: [http://dx.doi.org/10.1007/978-3-642-39212-2\\_26](http://dx.doi.org/10.1007/978-3-642-39212-2_26), doi:10.1007/978-3-642-39212-2\_26.
- [5] Hugo Gimbert. On the control of asynchronous automata. In Satya V. Lokam and R. Ramanujam, editors, *37th IARCS Annual Conference on Foundations of Software Technology and Theoretical Computer Science, FSTTCS 2017, December 11-15, 2017, Kanpur, India*, volume 93 of *LIPICs*, pages 30:1–30:15. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2017. URL: <https://doi.org/10.4230/LIPICs.FSTTCS.2017.30>, doi:10.4230/LIPICs.FSTTCS.2017.30.
- [6] P. Madhusudan, P. S. Thiagarajan, and Shaofa Yang. The MSO theory of connectedly communicating processes. In *FSTTCS 2005: Foundations of Software Technology and Theoretical Computer Science, 25th International Conference, Hyderabad, India, December 15-18, 2005, Proceedings*, pages 201–212, 2005. URL: [http://dx.doi.org/10.1007/11590156\\_16](http://dx.doi.org/10.1007/11590156_16), doi:10.1007/11590156\_16.
- [7] Anca Muscholl. Automated synthesis of distributed controllers. In *Automata, Languages, and Programming - 42nd International Colloquium, ICALP 2015, Kyoto, Japan, July 6-10, 2015, Proceedings, Part II*, pages 11–27, 2015. URL: [http://dx.doi.org/10.1007/978-3-662-47666-6\\_2](http://dx.doi.org/10.1007/978-3-662-47666-6_2).
- [8] Anca Muscholl and Igor Walukiewicz. Distributed synthesis for acyclic architectures. In *34th International Conference on Foundation of Software Technology and Theoretical Computer Science, FSTTCS 2014, December 15-17, 2014, New Delhi, India*, pages 639–651, 2014. URL: <http://dx.doi.org/10.4230/LIPICs.FSTTCS.2014.639>.
- [9] Anca Muscholl, Igor Walukiewicz, and Marc Zeitoun. A look at the control of asynchronous automata. In M. Mukund K. Lodaya and eds. N. Kumar, editors, *Perspectives in Concurrency Theory*. Universities Press, CRC Press, 2009.
- [10] Amir Pnueli and Roni Rosner. Distributed reactive systems are hard to synthesize. In *Foundations of Computer Science, 1990. Proceedings., 31st Annual Symposium on*, pages 746–757. IEEE, 1990.
- [11] Emil L. Post. A variant of a recursively unsolvable problem. *Bull. Amer. Math. Soc.*, 52:264–268, 1946.
- [12] Peter JG Ramadge and W Murray Wonham. The control of discrete event systems. *Proceedings of the IEEE*, 77(1):81–98, 1989.
- [13] Wiesław Zielonka. Notes on finite asynchronous automata. *ITA*, 21(2):99–135, 1987.