Noname manuscript No.
(will be inserted by the editor)

On the complexity of occurrence and convergence problems in reaction systems

Enrico Formenti · **Luca Manzoni** · **Antonio E. Porreca**

the date of receipt and acceptance should be inserted later

Abstract Reaction systems are a model of computation inspired by biochemical reactions introduced by Ehrenfeucht and Rozenberg. Two problems related to the dynamics (the evolution of the state with respect to time) of reaction systems, namely, the occurrence and the convergence problems, were recently investigated by Salomaa. In this paper, we prove that both problems are **PSPACE**-complete when the numerical parameter of the problems (i.e., the time step when a specified element must appear) is given as input. Moreover, they remain **PSPACE**-complete even for minimal reaction systems.

Keywords Reaction systems · Computational complexity · Discrete dynamical systems

1 Introduction

Reaction systems have been recently introduced as a formal model inspired to the reactions taking place in biochemical systems (Ehrenfeucht and Rozenberg 2007). Each reaction is described by means of a set of reactants, a set of inhibitors and a set of products. At each time step, the reactions transform a set of entities, generating a state sequence over time.

E. Formenti
Univ. Nice Sophia Antipolis, CNRS, I3S, UMR 7271, 06900 Sophia Antipolis, France
E-mail: enrico.formenti@unice.fr

Univ. Nice Sophia Antipolis, CNRS, I3S, UMR 7271, 06900 Sophia Antipolis, France
L. Manzoni
E-mail: luca.manzoni@i3s.unice.fr

A.E. Porreca
Dipartimento di Informatica, Sistemistica e Comunicazione, Università degli Studi di Milano-Bicocca, Viale Sarca 336/14, 20126 Milano, Italy
E-mail: porreca@disco.unimib.it

This note concerns some questions regarding the dynamics of reaction systems. In particular, it investigates the computational complexity of determining if a particular element appears at the m-th step in at least one state sequence (resp., all state sequences) consisting of at least m steps. Those two problems are called the *occurrence* and *convergence* problems and were studied by Salomaa (2013b; 2013a) when the parameter m is fixed. We prove that, when m is allowed to vary, the complexity of the decision problem increases from **coNP**-completeness or **NP**-completeness to **PSPACE**-completeness, even when the reaction systems are minimal, i.e., when sets of reactants and inhibitors for all reactions in the system are singletons. Furthermore, we prove that the complexity results provided by Salomaa (2013b, Theorem 2; 2013a, Corollary 1) extend to the case of minimal reaction systems.

The paper is structured as follows. In Section 2, the necessary basic notions are introduced. In Section 3 the main results are stated. The proofs are pretty technical and are given in a separate section (Section 4), which is organized in several subsections. We describe an implementation of finite cellular automata (4.1) and binary counters (4.2) using reaction systems. Section 4.3 shows the **PSPACE**-completeness of the two decision problems in the general case, and Section 4.4 provides a way to transport the complexity results to minimal reaction systems. Finally, Section 5 draws our conclusions.

2 Basic notions

This section recalls the concepts of reaction, reaction systems, and of the state sequence generated by a reaction system starting from a state. The definitions of the occurrence and convergence problems for reaction systems are also stated.

Definition 1 Given a finite set S, a reaction a over S is a triple (R,I,P) of *nonempty* subsets of S with $R \cap I = \varnothing$. The sets are called *reactants*, *inhibitors*, and *products*, respectively. The union of the reactants and inhibitors of a is the set of *resources* of a.

Notice that Definition 1 implies that a reaction has at least two resources. In this case the reaction is called *minimal*. If the number of resources is at most three, then the reaction is *almost minimal*.

Given a set $T \subseteq S$ and a reaction $a = (R,I,P)$, a is *enabled by* T iff $R \subseteq T$ (i.e., all reactants are present) and $I \cap T = \varnothing$ (i.e., there are no inhibitors).

Definition 2 A *reaction system* is a pair $\mathcal{A} = (S,A)$ where S is a nonempty finite set, called the *background set*, and A is a finite set of reactions over S.

A reaction system $\mathcal{A} = (S,A)$ is *minimal* (resp., *almost minimal*) if all its the reactions are minimal (resp., almost minimal). A reaction $a = (R,I,P) \in A$ defines a transition function $\mathrm{res}_a \colon 2^S \to 2^S$, usually called the *result function*, as follows:

$$\forall T \subseteq S \qquad \mathrm{res}_a(T) = \begin{cases} P & \text{if } a \text{ is enabled by } T \\ \varnothing & \text{otherwise.} \end{cases}$$

Similarly, the result function of a set A of reactions is defined as $\mathrm{res}_A(T) = \bigcup_{a \in A} \mathrm{res}_a(T)$. A reaction system $\mathcal{A} = (S,A)$ has $\mathrm{res}_{\mathcal{A}} = \mathrm{res}_A$ as its result function. Notice that, by the definition of reaction system, $\mathrm{res}_{\mathcal{A}}(S) = \mathrm{res}_{\mathcal{A}}(\varnothing) = \varnothing$. The *state sequence* starting at $T \subseteq S$ is

$$(T, \mathrm{res}_{\mathcal{A}}(T), \mathrm{res}^2_{\mathcal{A}}(T), \mathrm{res}^3_{\mathcal{A}}(T), \ldots).$$

The *length* of a state sequence is the number of nonempty positive iterates of $\mathrm{res}_{\mathcal{A}}$. A state sequence starting at $T \subseteq S$ is a *cycle* if $\mathrm{res}^t_{\mathcal{A}}(T) = T$ for some $t > 0$. The *period* of the cycle is the minimum such t.

3 Main results

Salomaa investigated two problems related to the dynamics of reaction systems, the occurrence and convergence problem (Salomaa 2013b,a).

Definition 3 (Occurrence problem) Given a reaction system $\mathcal{A} = (S,A)$, an element $s \in S$, and $m \in \mathbb{N}$, the decision problem of determining if s appears at the m-th step of at least one state sequence of $\mathcal{A}$ is called the *occurrence problem*.

Definition 4 (Convergence problem) Given a reaction system $\mathcal{A} = (S,A)$, an element $s \in S$, and $m \in \mathbb{N}$, the decision problem of determining if s appears at the m-th step of every state sequence of $\mathcal{A}$ consisting of at least m steps is called the *convergence problem*.

For some fixed values of the parameter m, the occurrence problem is **NP**-complete (Salomaa 2013b) and the convergence problem is **coNP**-complete (Salomaa 2013a).

The main results of this paper are given by the following theorems, stating that the two problems, when m is given as input, are **PSPACE**-complete and that all these complexity results generalize to minimal reaction systems.

Theorem 1 *The occurrence and the convergence problems for reaction systems are* **PSPACE***-complete.*

Theorem 2 *The occurrence and the convergence problems for minimal reaction systems are* **PSPACE***-complete.*

Furthermore, we show that, when m is fixed, both problems maintain their complexity in the minimal case.

Theorem 3 *The occurrence problem with fixed m for minimal reaction systems is* **NP***-complete for some $m \in \mathbb{N}$.*

Theorem 4 *The convergence problem with fixed m for minimal reaction systems is* **coNP***-complete for some $m \in \mathbb{N}$.*

4 Proofs and constructions

In order to prove the above-mentioned results, two auxiliary constructions are needed: the simulation of cellular automata and the implementation of a binary counter by means of reaction systems.

The main idea is to reduce a convenient version of the reachability problem for cellular automata, which is known to be **PSPACE**-complete (Sutner 1995), to the occurrence and convergence problems.

4.1 Cellular automata

A cellular automaton is a (finite or infinite) lattice of identical cells (finite state automata) whose state depends only on a fixed number of neighboring cells. More formally:

Definition 5 A *one-dimensional cellular automaton* $\mathscr{C}$ is a triple (Σ, r, λ) where Σ is a finite alphabet, $r \in \mathbb{N}$ is the *radius*, and $\lambda : \Sigma^{2r+1} \to \Sigma$ is called the *local rule*.

A *configuration* of a cellular automaton is a function $x \colon C \to \Sigma$ (i.e., $x \in \Sigma^C$) where C denotes the lattice on which the cellular automaton is defined. Infinite one-dimensional cellular automata are usually defined with $C = \mathbb{Z}$. Finite cellular automata, on the other hand, have $C = \{0, \ldots, n-1\}$ for some $n \in \mathbb{N}$. In this case we say that $\mathscr{C}$ has size n.

A cellular automaton induces a global transition function $\Lambda : \Sigma^C \to \Sigma^C$ defined as follows:

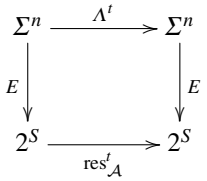
$$\forall x \in \Sigma^C \; \forall i \in C \qquad \Lambda(x)_i = \lambda(x_{i-r}, \ldots, x_i, \ldots, x_{i+r}).$$

Given a configuration $x \in \Sigma^C$, it is possible to obtain the configuration sequence of the CA by iterating the global transition function Λ. In the following, $\Lambda^t(x)$ will denote the configuration of the CA after $t \in \mathbb{N}$ time steps and starting from the initial configuration x.

As usual, for each configuration $x \in \Sigma^C$, the state $x(i)$ of the i-th cell is denoted by x_i. There is a minor technicality for finite cellular automata, since the local rule cannot be directly applied to cells near the boundary of the lattice. It is possible to have *fixed boundary conditions*, that is, each cell whose index is outside C has a fixed value $\sigma \in \Sigma$. Another common choice is to assume that the configuration is periodic (i.e., *periodic boundary conditions*) and each index is actually considered modulo the size of the automaton. The arguments presented in this paper can be applied to both options. However, for the sake of simplicity, we will consider only periodic boundary conditions. Hence, in the rest of the paper the indices are always assumed to be integers modulo n, where n is the size of the lattice.

Let $\hat{E}\colon \Sigma^n \times \{1,\dots,n\} \to \Sigma \times \{1,\dots,n\}$ be the mapping $(x,i) \mapsto (x_i,i)$. The encoding of a configuration $x \in \Sigma^n$ is defined as $E(x) = \{\hat{E}(x,i) \ : \ 1 \le i \le n\}$. Notice that E is injective.

Lemma 1 *For each finite cellular automaton $\mathscr{C} = (\Sigma, 1, \lambda)$ of length $n \in \mathbb{N}$, there exists a reaction system $\mathcal{A} = (S,A)$ with $S = \{(\sigma,i) \mid \sigma \in \Sigma,\ 1 \le i \le n\} \cup \{\blacksquare\}$ such that for all $x \in \Sigma^n$ and for all $t \in \mathbb{N}$ we have $\mathrm{res}^t_{\mathcal{A}}(E(x)) = E(\Lambda^t(x))$, i.e., the following diagram commutes:*



Proof The set A contains, for each $i \in \{1,\dots,n\}$ and for each $\alpha,\beta,\gamma \in \Sigma$, the reaction

$$\big(\{(\alpha,i-1),(\beta,i),(\gamma,i+1)\},\{\blacksquare\},\{(\lambda(\alpha,\beta,\gamma),i)\}\big), \qquad (1)$$

where $i-1$ and $i+1$ are to be considered operations modulo n.

Now let $x \in \Sigma^n$ be a configuration of $\mathscr{C}$. For every position $i \in \{1,\dots,n\}$, the encoding $E(x)$ contains $(x_{i-1},i-1)$, (x_i,i), $(x_{i+1},i+1)$, and no other entity with the same second component belongs to $E(x)$. Hence the reaction of type (1) having $\alpha = x_{i-1}$, $\beta = x_i$, $\gamma = x_{i+1}$ is enabled, producing the entity resulting from the application of the local function to the cell in position i, thus generating the encoding of $\Lambda(x)$. The statement of the lemma follows by induction on t. □

In other words, Lemma 1 says that, for any cellular automaton of the kind considered above, there exists a conjugate reaction system.

Notice that the number of entities and reactions used by the construction of Lemma 1 is polynomial with respect to the description of the cellular automaton given in input (there are $n \times |\Sigma| + 1$ entities and $n \times |\Sigma|^3$ reactions) and that the construction can be performed in polynomial time.

4.2 Counters

This section describes the construction of a binary counter by means of reaction systems. It is inspired by a similar construction by Ehrenfeucht and Rozenberg (2009).

Lemma 2 *For every $n \in \mathbb{N}$ there exists a reaction system $\mathcal{C}_k = (S,A)$ with $|S| = n+2$ implementing a binary counter with n bits overflowing at $k = 2^n$.*

Proof Let $S = \{b_0,\dots,b_{n-1},\blacktriangleright,\blacksquare\}$. We interpret any subset $B \subseteq \{b_0,\dots,b_{n-1}\}$ as an n-bit integer $m = \sum_{b_i \in B} 2^i$, i.e., the i-th bit of m is 1 iff $b_i \in B$. Let B_m be the subset representing the integer m. We shall design the reactions in A in order to obtain the following result function

$$\mathrm{res}_{\mathcal{C}_k}(B_m \cup \{\blacktriangleright\}) = B_{(m+1) \bmod k} \cup \{\blacktriangleright\} \qquad (2)$$

where by $x \bmod y$ we denote the remainder of the integer division of x by y. The result function gives $\varnothing$ on the sets that are not in the form $B_k \cup \{\blacktriangleright\}$. The reaction system can thus be seen as a binary counter overflowing at $k = 2^n$.

The set A contains, for each $i \in \{0,\dots,n-1\}$, the following reactions:

$$(\{b_i,\blacktriangleright\},\{b_j,\blacksquare\},\{b_i\}) \qquad \text{for } 0 \le j < i \qquad (3)$$

$$(\{b_0,\dots,b_{i-1},\blacktriangleright\},\{b_i,\blacksquare\},\{b_i\}) \qquad (4)$$

together with the reaction

$$(\{\blacktriangleright\},\{\blacksquare\},\{\blacktriangleright\}). \qquad (5)$$

If the current state contains $\blacktriangleright$ and does not contain $\blacksquare$, the reactions of type (3) preserve the bits set to 1 if there is a less significant bit set to 0, since they are not going to be modified by the increment operation. The reactions of type (4) set a currently null bit to 1 when all the less significant bits are 1. Notice that for the least significant bit this means that it will always be created if it is not present, i.e., the reaction of type (4) for this bit is $(\{\blacktriangleright\},\{b_0,\blacksquare\},\{b_0\})$. By construction, no reaction of type (3) or (4) is enabled when the current state of $\mathcal{C}_k$ is $B_{k-1} \cup \{\blacktriangleright\} = B_{2^n-1} \cup \{\blacktriangleright\}$, resetting the counter to $B_0 \cup \{\blacktriangleright\} = \{\blacktriangleright\}$. Finally, the reaction (5) preserves the entity $\blacktriangleright$.

Hence, the reaction system $\mathcal{C}_k$ defines the result function of Equation (2), and thus it has a cycle of period $k = 2^n$.

If $\blacktriangleright$ is missing, or $\blacksquare$ is present, the system reaches $\varnothing$ in one step, since $\blacktriangleright$ is a reactant of every reaction, and the entity $\blacksquare$ inhibits all of them. □

The construction of Lemma 2 can be performed in polynomial time with respect to the number of bits of the counter to be constructed. This construction will be used to force a sequence of states to terminate (i.e., to reach $\varnothing$) after a predefined amount of time.

4.3 The occurrence and the convergence problems

The reachability problem for one-dimensional finite cellular automata of radius 1 and global rule Λ, i.e., deciding if there exists $t \in \mathbb{N}$ such that $\Lambda^t(x) = y$ for two given configurations $x, y \in \Sigma^n$, is **PSPACE**-complete (Sutner 1995). This allows us to prove, by reduction, that the occurrence and convergence problems for reaction systems are **PSPACE**-complete when the parameter m is given as an input.

Proof (Theorem 1) First of all, notice that the occurrence and convergence problems are both in **PSPACE**, since they can be solved by simply checking every state sequence for m steps, and this can be done using only polynomial space.

In order to prove the **PSPACE**-hardness of the problems, we combine the previously described constructions of reaction systems simulating a cellular automaton (Lemma 1) and a binary counter (Lemma 2). The background set of the resulting reaction system will include the entities of both.

Let $\mathscr{C} = (\Sigma, 1, \lambda)$ be a one-dimensional cellular automaton of size n and let $x, y \in \Sigma^n$ be two configurations. Define a reaction system $\mathcal{A} = (S, A)$ with an entity $\heartsuit$ such that $\heartsuit$ is present at a certain time step m in all state sequences of length at least m iff y is reachable from x in $\mathscr{C}$; we are also going to ensure that there always exists a state sequence of length at least m.

Let $\mathcal{B} = (S', B)$ be the reaction system simulating $\mathscr{C}$ as in Lemma 1 and let $\mathcal{C} = (S'', C)$ be the reaction system simulating a binary counter of $p+1$ bits (with $2^p \geq |\Sigma^n|$) as in Lemma 2.

Define the reaction system $\mathcal{A} = (S, A)$ having the background set $S = S' \cup S'' \cup \{\blacktriangleright\!\!|, \blacktriangleleft, \blacktriangleright\!\!|, \heartsuit, \spadesuit\}$ and A containing the reactions of B and C, as well as the following ones; notice that (6)–(8) are reaction schemata, i.e., there is an instance of these reactions for each entity s ranging over the set on the right-hand side:

$$(\{\blacktriangleright\!\!|\}, \{s\}, \{\blacksquare, \blacktriangleleft\}) \quad \text{for } s \in S - \{\blacktriangleright\!\!|\} \tag{6}$$
$$(\{\blacktriangleleft\}, \{s\}, E(x) \cup \{\blacktriangleright\}) \quad \text{for } s \in S - \{\blacktriangleleft\} \tag{7}$$
$$(\{s\}, \{\blacktriangleright, \blacksquare\}, \{\blacksquare\}) \quad \text{for } s \in S - \{\blacktriangleright, \blacksquare, \heartsuit, \spadesuit\} \tag{8}$$
$$(E(y) \cup \{\blacktriangleright\}, \{\blacksquare\}, \{\blacktriangleright\!\!|\}) \tag{9}$$
$$(\{\blacktriangleright\!\!|\}, \{\blacksquare\}, \{\blacktriangleright\!\!|\}) \tag{10}$$
$$(\{b_p\}, \{\blacksquare\}, \{\blacksquare\}) \tag{11}$$
$$(\{\blacktriangleright\!\!|, \blacksquare\}, \{\blacktriangleleft\}, \{\heartsuit\}) \tag{12}$$
$$(\{\blacksquare\}, \{\blacktriangleright\!\!|, \blacktriangleleft\}, \{\spadesuit\}) \tag{13}$$

Intuitively, the interaction between the simulations of the cellular automaton and the binary counter is performed using $\blacksquare$, in the sense that the presence of $\blacksquare$ forces to stop both simulations at the same time. Furthermore, the absence of $\blacktriangleright$ indicates that the binary counter is not increasing, and hence that the simulation of the cellular automaton must be aborted.

The reactions of type (6) and (7) ensure that, when $\blacktriangleright\!\!|$ is present, all the elements of the existing configuration are discarded (using $\blacksquare$) and that a known state, containing $E(x)$ and $\blacktriangleright$, is reached in two steps by means of $\blacktriangleleft$. Reactions of type (8) ensure that if the counter is not increasing (i.e., $\blacktriangleright$ is absent), then the dynamics reaches $\varnothing$. Reaction (9) generates $\blacktriangleright\!\!|$ when the encoding of y is reached. The entity $\blacktriangleright\!\!|$ is preserved by reaction (10). The reaction (11) ensures that, when the timer has reached the value 2^p (or greater, which can only happen if it was already greater in the initial state), the state $\varnothing$ is reached by means of $\blacksquare$. When $\blacksquare$ is present, it inhibits all the reactions except (7), (12), and (13). In particular, exactly one between (12) and (13) is enabled, depending on the presence of $\blacktriangleright\!\!|$.

Let $T \subseteq S$ be a nonempty state of $\mathcal{A}$. Then, we distinguish the following cases:

- If $\blacktriangleright\!\!| \in T$ then $\mathrm{res}^2_{\mathcal{A}}(T) = E(x) \cup \{\blacktriangleright\}$. Hence we have $\mathrm{res}^{2^p+4}_{\mathcal{A}}(T) = \{\heartsuit\}$ if configuration y is reachable from configuration x, and $\mathrm{res}^{2^p+4}_{\mathcal{A}}(T) = \{\spadesuit\}$ otherwise. Notice that this case can only happen if $\blacktriangleright\!\!|$ belongs to the initial state, since no reaction produces $\blacktriangleright\!\!|$.
- Otherwise, if $\blacktriangleright \notin T$ or it is not produced by a reaction of type (7) in one step, then in three steps the state $\varnothing$ is reached.
- In all other cases the counter reaches a value greater than $2^p - 1$ in at most 2^p steps. Hence the dynamics reaches the state $\varnothing$ by the time step $2^p + 4$.

All the state sequences that are nonempty at time step $2^p + 4$ have the same state at time step 2, and contain $\heartsuit$ at time step $2^p + 4$ if y is reachable from x in $\mathscr{C}$ and $\spadesuit$ otherwise. Since this property holds for all sequences that reach time step $m = 2^p + 4$, and there is at least one such sequence (e.g., starting from the state $\{\blacktriangleright\!\!|\}$), both the occurrence and the convergence problems for reactions systems are **PSPACE**-complete. □

4.4 Conversion to minimal reaction systems

It is interesting to note that there is no difference in these problems with respect to the computational complexity in the unrestricted case and in the minimal case. To prove this result we need the following general theorem which shows that any reaction system can be "simulated" by a minimal one. This construction is inspired by the idea of k-simulation (Manzoni and Porreca 2013; Manzoni et al 2014).

Theorem 5 *For any reaction system $\mathcal{A} = (S,A)$ there exists a minimal reaction system $\mathcal{A}' = (S',A')$ with background set $S' = S \cup A \cup \{\triangleright_1, \triangleright_2, \diamond\}$ such that the following statements hold:*

1. *If $T \subseteq S$ then we have $\mathrm{res}^{2t}_{\mathcal{A}'}(T \cup \{\triangleright_1\}) \cap S = \mathrm{res}^{t}_{\mathcal{A}}(T)$ for all $t \in \mathbb{N}$.*
2. *If $B \subseteq A$ we have $\mathrm{res}_{\mathcal{A}'}(B \cup \{\triangleright_2\}) = T \cup \{\triangleright_1\}$, where $T = \bigcup\{P : (R,I,P) \in A - B\}$.*
3. *If $V \subseteq S'$ is neither of the form $T \cup \{\triangleright_1\}$ for some $T \subseteq S$, nor of the form $Q \cup \{\triangleright_2\}$ for some $Q \subseteq A$ (with T and Q nonempty), then $\mathrm{res}^{4}_{\mathcal{A}'}(V) = \varnothing$.*

Proof For each reaction $a = (R,I,P) \in A$, the set A' contains:

$$(\{\triangleright_1\}, \{s\}, \{a\}) \qquad \text{for each } s \in R \qquad (14)$$

$$(\{s\}, \{\diamond\}, \{a\}) \qquad \text{for each } s \in I \qquad (15)$$

$$(\{\triangleright_2\}, \{a\}, P) \qquad (16)$$

Furthermore A' contains the following reactions:

$$(\{s\}, \{\triangleright_2\}, \{\triangleright_2\}) \qquad \text{for each } s \in S \qquad (17)$$

$$(\{\triangleright_2\}, \{\triangleright_1\}, \{\triangleright_1\}) \qquad (18)$$

$$(\{s\}, \{\triangleright_1\}, \{\diamond\}) \qquad \text{for each } s \in S \qquad (19)$$

$$(\{a\}, \{\triangleright_2\}, \{\diamond\}) \qquad \text{for each } a \in A \qquad (20)$$

$$(\{\diamond\}, \{s\}, S') \qquad \text{for each } s \in S' - \{\diamond\} \qquad (21)$$

Let $T \subseteq S$ and $s \in \mathrm{res}^{2}_{\mathcal{A}'}(T \cup \{\triangleright_1\}) \cap S$. Since an entity $s \in S$ can only be produced by reactions of type (16) or (21), it suffices to notice that in $T \cup \{\triangleright_1\}$ no reaction of type (19) or (20) is enabled, the entity $\diamond$ is not produced, and no reaction of type (21) is enabled in the next step. Thus, s is necessarily produced by a reaction of type (16), corresponding to a reaction $a = (R,I,P) \in A$, and we have $a \notin \mathrm{res}_{\mathcal{A}'}(T \cup \{\triangleright_1\})$. By construction, this means that both $R \subseteq T \cup \{\triangleright_1\}$ and $I \cap (T \cup \{\triangleright_1\}) = \varnothing$ hold, hence $R \subseteq T$ and $I \cap T = \varnothing$. Thus a is enabled by T and $s \in \mathrm{res}_{\mathcal{A}}(T)$. Conversely, $s \in \mathrm{res}_{\mathcal{A}}(T)$ implies $s \in \mathrm{res}^{2}_{\mathcal{A}'}(T \cup \{\triangleright_1\}) \cap S$. Statement 1 of the theorem follows by induction on t.

If $B \subseteq A$, then the only reactions enabled by $B \cup \{\triangleright_2\}$ are of type (16) and (18). The former generate the products of the reactions not in A, while the latter produces $\triangleright_1$. This proves statement 2 of the theorem.

Now let V be a subset of S'. If V is empty, statement 3 follows immediately, hence suppose $V \neq \varnothing$. Then the following cases can arise:

- $\diamond \in V$. Then either $V = S'$ or at least one of the reactions of type (21) is enabled, producing S' in one step. In both cases $\varnothing$ is reached in at most two steps.
- $\{\triangleright_1, \triangleright_2\} \cap V = \varnothing$. In this case either $\diamond \in V$ or at least one of the reactions of type (19) and (20) is enabled, producing $\diamond$ in one time step.
- $\{\triangleright_1, \triangleright_2\} \subseteq V$. In this case the reactions (17) and (18) are not enabled, neither $\triangleright_1$ nor $\triangleright_2$ are preserved to the next time step, and we fall back to the previous case.
- $S \cap V \neq \varnothing$ and $A \cap V \neq \varnothing$. In this case if we only have exactly one between $\triangleright_1$ and $\triangleright_2$ (otherwise we are in some other case) at least one of the reactions of type (19) and (20) is enabled, producing $\diamond$ in one step, and we fall back to the previous case.

Since all possible cases are covered, the third statement of the theorem holds. □

Essentially, the reaction system $\mathcal{A}'$ of Theorem 5 simulates each step of the original reaction system $\mathcal{A}$ in two steps when starting from the same configuration (with an auxiliary entity adjoined) and terminates, i.e., reaches $\varnothing$, if $\mathcal{A}$ does.

We apply this minimization procedure to prove that the occurrence and convergence problems remain complete for **PSPACE** in the minimal case.

Proof (Theorem 2) The reaction system $\mathcal{A}$ of the proof of Theorem 1 produces $\heartsuit$ at step m in all state sequences iff configuration y of the simulated cellular automaton is reachable from x, and $\spadesuit$ otherwise. Let $\mathcal{A}'$ be the reaction system obtained by applying Theorem 5 to $\mathcal{A}$. Then, the reachability (resp., non-reachability) of y from x can be deduced by observing the presence of $\heartsuit$ (resp., $\spadesuit$) in all state sequences of $\mathcal{A}'$ after $2m$ steps. Hence, the occurrence and convergence problems remain **PSPACE**-complete for minimal reaction systems. □

Similarly, we can minimize an existing construction by Salomaa (2013a) and prove the **NP**-completeness of the occurrence problem when the parameter m is fixed.

Proof (Theorem 3) Given a Boolean formula Φ, it is possible to construct a reaction system $RS(\Phi)$ having an entity s (denoted by A in the original proof) that appears at time step 3 in a state sequence iff Φ is satisfiable (Salomaa 2013a, Theorem 2). By applying Theorem 5 to $RS(\Phi)$, we can construct a minimal reaction system where the entity s appears, under the same conditions as described above, at time step 6. This proves that the occurrence problem with $m = 6$ is **NP**-complete for minimal reaction systems. □

The construction used in the following proof is inspired by the proof of **coNP**-completeness of the problem of functional equality of reaction systems (Ehrenfeucht and Rozenberg 2007).

Proof (Theorem 4) We now reduce the validity problem for Boolean formulae (Papadimitriou 1993) to the convergence problem in order to prove its **coNP**-hardness. Let $\varphi = \varphi_1 \vee \cdots \vee \varphi_\ell$ be a Boolean formula in disjunctive normal form over the variables $V = \{x_1, \ldots, x_n\}$; for a given disjunctive term φ_j, let $\mathrm{pos}(\varphi_j)$ be the set of variables occurring

as positive literals in φ_j, and $\mathrm{neg}(\varphi_j)$ be the set of variables occurring as negated literals.

Let $\mathcal{A} = (S,A)$ be the minimal reaction system having $S = V \cup \{\varphi_1,\ldots,\varphi_\ell,\triangleright_1,\triangleright_2,\triangleright_3,\heartsuit\}$ and the following reactions:

$$(\{\triangleright_1\},\{x_i\},\{\varphi_j\}) \qquad \text{for } 1 \le j \le \ell \text{ and } x_i \in \mathrm{pos}(\varphi_j) \tag{22}$$

$$(\{x_i\},\{\triangleright_2\},\{\varphi_j\}) \qquad \text{for } 1 \le j \le \ell \text{ and } x_i \in \mathrm{neg}(\varphi_j) \tag{23}$$

$$(\{\triangleright_2\},\{\varphi_j\},\{\heartsuit\}) \qquad \text{for } 1 \le j \le \ell \tag{24}$$

$$(\{\triangleright_1\},\{\triangleright_2\},\{\triangleright_2\}) \tag{25}$$

$$(\{\triangleright_2\},\{\triangleright_3\},\{\triangleright_3\}) \tag{26}$$

We claim that $\heartsuit$ appears at the second step of every state sequence consisting of at least two steps if and only if φ is a tautology. Given $X \subseteq V$, the set X represents the truth assignment in which x_i is true iff $x_i \in X$.

For each $T \subseteq S$ such that $\triangleright_1 \in T$ and $\triangleright_2 \notin T$, we have that $\mathrm{res}_{\mathcal{A}}(T) = C \cup \{\triangleright_2\}$, where $C \subseteq \{\varphi_1,\ldots,\varphi_\ell\}$ contains exactly the clauses that are *not* satisfied by the assignment $T \cap V$ (reactions of type (22), (23), and (25)). In the next step we have $\mathrm{res}_{\mathcal{A}}(C \cup \{\triangleright_2\}) = \{\triangleright_3\}$ if $C = \{\varphi_1,\ldots,\varphi_\ell\}$, i.e., if $T \cap V$ does not satisfy φ, and $\mathrm{res}_{\mathcal{A}}(C \cup \{\triangleright_2\}) = \{\heartsuit,\triangleright_3\}$ otherwise (reactions of type (24) and (26)).

If $T \subset S$ does not contain $\triangleright_1$, or contains both $\triangleright_1$ and $\triangleright_2$, then we always have $\mathrm{res}^2_{\mathcal{A}}(T) = \varnothing$, as can be easily verified case by case.

Hence, all state sequences consisting of at least two steps contain $\heartsuit$ after $m = 2$ steps if and only if φ is a tautology. Since the mapping $\varphi \mapsto \mathcal{A}$ is computable in polynomial time, the **coNP**-hardness of the problem follows. □

5 Conclusions

In this paper we have proved that the occurrence and convergence problems for reaction systems are **PSPACE**-complete when the number of steps m is given as input. These complexity results can also be proved for minimal reaction systems by exploiting a rather general minimization method. Similarly, we have also shown that, for some fixed values of m, the occurrence problem is **NP**-complete and the convergence problem is **coNP**-complete for minimal reaction systems, extending the results already known for general reaction systems.

Hence, the occurrence and convergence problems do not highlight a difference in complexity of the dynamics of minimal, almost minimal, and general reaction systems. One interesting open question is to find naturally arising problems related to the dynamical behavior whose complexity separates these three classes of reaction systems.

Acknowledgements This work has been partially supported by the French National Research Agency project EMC (ANR-09-BLAN-0164).

References

Ehrenfeucht A, Rozenberg G (2007) Reaction systems. Fundamenta Informaticae 75:263–280, URL http://iospress.metapress.com/content/b86t11hryvwq69l0/

Ehrenfeucht A, Rozenberg G (2009) Introducing time in reaction systems. Theoretical Computer Science 410(4):310–322, URL http://dx.doi.org/10.1016/j.tcs.2008.09.043

Manzoni L, Porreca AE (2013) Reaction systems made simple: A normal form and a classification theorem. In: Mauri G, Dennunzio A, Manzoni L, Porreca AE (eds) Unconventional Computation and Natural Computation, 12th International Conference, UCNC 2013, Lecture Notes in Computer Science, vol 7956, Springer, pp 150–161, URL http://dx.doi.org/10.1007/978-3-642-39074-6_15

Manzoni L, Poças D, Porreca AE (2014) Simple reaction systems and their classification. International Journal of Foundations of Computer Science (to appear)

Papadimitriou CH (1993) Computational Complexity. Addison-Wesley

Salomaa A (2013a) Functional constructions between reaction systems and propositional logic. International Journal of Foundations of Computer Science 24(1):147–159, URL http://dx.doi.org/10.1142/S0129054113500044

Salomaa A (2013b) Minimal and almost minimal reaction systems. Natural Computing 12(3):369–376, URL http//dx.doi.org/10.1007/s11047-013-9372-y

Sutner K (1995) On the computational complexity of finite cellular automata. Journal of Computer and System Sciences 50(1):87–97, URL http://dx.doi.org/10.1006/jcss.1995.1009