

Minimal and almost minimal reaction systems

Arto Salomaa

Published online: 11 May 2013
© Springer Science+Business Media Dordrecht 2013

Abstract In *reaction systems* introduced by Ehrenfeucht and Rozenberg the number of resources is, by definition, at least 2. If it is exactly 2, the system is referred to as *minimal*. We compare minimal reaction systems with *almost minimal* ones, where the number of resources equals 3. The difference turns out to be huge. While many central problems for minimal systems are of low polynomial complexity, the same problems in the almost minimal case are NP- or co-NP-complete. The situation resembles the difference between 2-SAT and 3-SAT, also from the point of view of techniques used. We also compare maximal sequence lengths obtainable in the two cases. We are concerned only with the most simple variant of reaction systems.

Keywords Reaction system · Reactant · Inhibitor · Inverse function · Sequence length · NP-completeness · 1-in-3 3SAT problem

1 Introduction

A formal model of *reaction systems* was introduced by Ehrenfeucht and Rozenberg (2007). Everything is defined within a fixed finite *background set* S . The sets of *reactants*, *inhibitors* and *products*, R , I and P , are nonempty subsets of S , the sets R and I being disjoint. A *reaction system* \mathcal{A} consists of finitely many such triples (R, I, P) , called *reactions*.

The original purpose was to model interactions between biochemical reactions. Ehrenfeucht and Rozenberg (2007) contains some of the original motivation and initial setup. Each reaction is characterized by its set of reactants, each of which has to be present for the reaction to take place, by its set of inhibitors, none of which is allowed to be present, and by its set of products, each of which will be present after a successful reaction. Thus, a single reaction is based on facilitation and inhibition.

Reaction systems provide a new kind of mechanism for generating functions and sequences over a finite set. The application of \mathcal{A} to a subset Y of S (to be explained in detail below) produces another subset Y' of S and, thus, we have a *function* from subsets of S to subsets of S . Iterating the function we get a *sequence* of subsets of S , also referred to as *states* of the sequence.

The model of reaction systems turns out to be suitable in different setups. Then also many variants of reaction systems and extensions to them were introduced. Brijder et al. (2011) constitutes a recent survey. However, in this paper we are concerned with the basic variant only.

The elements of the sets R and I are also referred to as *resources* of the reaction. Since R and I are nonempty and disjoint, the smallest possible cardinality of the resource set equals 2. In this paper we compare such *minimal* reaction systems with *almost minimal* ones, where the cardinality of the resource set equals 3. Some comparisons concern algorithmic complexity. We also investigate functional constructions and possibilities of obtaining long state sequences in case of minimal and almost minimal reaction systems.

The class of functions (from the set of subsets of S into itself) generated by minimal reaction systems was characterized in Ehrenfeucht et al. (2012b). General considerations concerning functions and sequences are contained in

A. Salomaa (✉)
Turku Centre for Computer Science, Joukahaisenkatu 3–5 B,
20520 Turku, Finland
e-mail: asalomaa@utu.fi

Ehrenfeucht et al. (2011), Rozenberg (2011), Salomaa (2012a, b).

The exposition in this paper is largely self-contained. In particular, the basic setup concerning reaction systems is given in Sect. 2 The presentation is focused on the needs of the next sections and, therefore, also some earlier results are quoted. Section 3 presents our main results concerning NP-completeness and co-NP-completeness of basic problems about almost minimal reaction systems. Thereby the basic reduction tool is the 1-in-3 3SAT problem, (Schaefer 1978). Intuitively, the difference between minimal and almost minimal reaction systems amounts to the difference between satisfiability problems 2-SAT and 3-SAT.

In the transition from minimal to almost minimal reaction systems, functional constructions become essentially richer. We construct in Sect. 4 sequences for almost minimal reaction systems that are exponentially longer than the longest known sequences for minimal reaction systems. Also certain functions not obtainable by minimal reaction systems can be obtained using almost minimal ones. However, a characterization of functions, resembling the one given in the minimal case in Ehrenfeucht et al. (2012b), seems hard to obtain in the almost minimal case.

In the final section we discuss further complexity issues and present some open problems.

2 Preliminaries and earlier results

All constructions dealing with a reaction system take place within a fixed finite set S , referred to as the *background set*. Thus, mathematically, we are working on functions and sequences over a finite set. This concerns the *basic variant* of reaction systems discussed in this paper. The reader may consult (Brijder et al. 2011) for more complicated variants.

We begin with the basic definition of a *reaction* and a *reaction system*.

Definition 1 A reaction *over* the finite background set S is a triple

$$\rho = (R, I, P),$$

where R, I and P are nonempty subsets of S such that R and I do not intersect. The three sets are referred as reactants, inhibitors and products, respectively. A reaction system \mathcal{A}_S over the background set S is a finite nonempty set

$$\mathcal{A}_S = \{\rho_j | 1 \leq j \leq k\},$$

of reactions over S .

In this paper S will always denote the background set. Observe that no assumptions are made concerning the relation of the set P to the sets R and I . Thus, P may be included in one of the latter two sets or contain elements

from both of them, as well as other elements of S . We will omit the index S from \mathcal{A}_S whenever S is understood.

We now come to the definitions dealing with functions and sequences. The *cardinality* of a finite set X is denoted by $\sharp X$. The *empty set* is denoted by \emptyset .

Definition 2 Consider a reaction $\rho = (R, I, P)$ over S and a subset Y of S . The reaction ρ is enabled with respect to Y (or for Y), in symbols $en_\rho(Y)$, if $R \subseteq Y$ and $I \cap Y = \emptyset$. If ρ is (resp. is not) enabled, then we define the result by

$$res_\rho(Y) = P(\text{resp. } = \emptyset).$$

For a reaction system $\mathcal{A} = \{\rho_j | 1 \leq j \leq k\}$, we define the result by

$$res_{\mathcal{A}}(Y) = \bigcup_{j=1}^k res_{\rho_j}(Y).$$

As an example, let \mathcal{A}_1 be the reaction system over the background set $\{1, 2, 3\}$, consisting of the three reactions

$$\begin{aligned} \rho_1 &= (\{1, 2\}, \{3\}, \{1, 3\}), & \rho_2 &= (\{2\}, \{3\}, \{2\}), \\ \rho_3 &= (\{1\}, \{2\}, \{1, 3\}). \end{aligned}$$

Consider $Y = \{1, 2\}$. Then $en_{\rho_1}(Y)$ and $en_{\rho_2}(Y)$, whereas $en_{\rho_3}(Y)$ does not hold. Consequently,

$$\begin{aligned} res_{\rho_1}(Y) &= \{1, 3\}, & res_{\rho_2}(Y) &= \{2\}, & res_{\rho_3}(Y) \\ &= \emptyset, & res_{\mathcal{A}_1}(Y) &= \{1, 2, 3\}. \end{aligned}$$

The following consequence of Definition 2 should be observed. Whenever an element is in a set Y , it is considered to be there always when needed. Thus, the element 2 of Y is not “consumed” in the application of the reaction ρ_1 but is also available for ρ_2 when $res_{\mathcal{A}_1}(Y)$ is computed. In this sense there is no “conflict” between ρ_1 and ρ_2 .

Elements in the set $R \cup I$ are also referred to as *resources*. Reaction systems are classified according to the maximal cardinalities of the sets of reactants and inhibitors, and also according to the cardinality of the set of resources.

Definition 3 A reaction system \mathcal{A} is a (k, l) reaction system if the conditions $\sharp R \leq k$ and $\sharp I \leq l$ are satisfied for every reaction (R, I, P) in \mathcal{A} . A reaction system is minimal (resp. almost minimal) if the cardinality of the set of resources equals 2 (resp. is at most 3).

Thus, minimal reaction systems are always (1, 1) reaction systems, whereas almost minimal reaction systems may contain, in addition, both (2, 1) and (1, 2) reactions.

Functions and *sequences* associated with reaction systems constitute a very interesting and natural way of handling operations over subsets of a finite set. Functions defined by reaction systems can be viewed as another description of the mapping *res*.

Definition 4 Consider a reaction system \mathcal{A} over S . Denote by S_1 (resp. S_2) the set of all nonempty (resp. proper nonempty) subsets of S . (Thus, the cardinalities of S_1 and S_2 are $2^{|S|} - 1$ and $2^{|S|} - 2$, respectively.) Consider the function $F_{\mathcal{A}}$ defined as follows, for $Y \subseteq S$. If $res_{\mathcal{A}}(Y) = Y' \neq \emptyset$, then $F_{\mathcal{A}}(Y) = Y'$. If $res_{\mathcal{A}}(Y) = \emptyset$, then $F_{\mathcal{A}}(Y)$ is undefined. The function $F_{\mathcal{A}}$ is termed total if it is defined for all elements of S_2 .

The range and domain of the function $\mathcal{F}_{\mathcal{A}}$ are subsets of S_1 and S_2 , respectively. The set S itself cannot belong to the domain because every reaction has to have at least one inhibitor.

For example, the function $F_{\mathcal{A}_1} = F_1$ defined by the reaction system \mathcal{A}_1 considered above satisfies

$$F_1\{1\} = \{1, 3\}, F_1\{2\} = \{2\}, F_1\{1, 2\} = \{1, 2, 3\}, F_1\{1, 3\} = \{1, 3\},$$

whereas the function is undefined for the arguments $\{3\}$ and $\{2, 3\}$. Hence, the function F_1 is not total. (We often omit unnecessary parentheses and write $F\{a\}$ instead of $F(\{a\})$).

The function $F_{\mathcal{A}}$ is total iff, for every proper nonempty subset Y of S , we have $en_{\rho}(Y)$, for at least one reaction ρ in \mathcal{A} . Thus, a total function $F_{\mathcal{A}}$ remains total if the product sets in some of the reactions in \mathcal{A} are changed. This matter was investigated further in Salomaa (2012), where the notion of a *core* of a reaction system was introduced.

Sequences generated by reaction systems can be viewed as iterations of the functions $F_{\mathcal{A}}$. If $\mathcal{F}_{\mathcal{A}}(Y) = Y'$, we use the notation

$$Y \Rightarrow_{\mathcal{A}} Y',$$

or simply $Y \Rightarrow Y'$ if \mathcal{A} is understood. If

$$F_{\mathcal{A}}(X_i) = X_{i+1}, \quad 0 \leq i \leq m - 1,$$

we write briefly

$$X_0 \Rightarrow X_1 \Rightarrow \dots \Rightarrow X_m$$

and call X_0, X_1, \dots, X_m states of a sequence of length m generated (or defined) by the reaction system \mathcal{A} . The sequence itself is referred to as a state sequence. The numbering of the states, or steps in a sequence, is as above: the step number 1 is obtained after one application of the function $F_{\mathcal{A}}$. To exclude trivial cases, we assume that X_0 is always a proper nonempty subset of S .

Since the number of subsets of S is finite, one of the following two alternatives always occurs for sequences

$$X_0 \Rightarrow X_1 \Rightarrow \dots \Rightarrow X_{m-1},$$

for large enough m .

1. $F_{\mathcal{A}}(X_{m-1}) = X_{m_1}$, for some $m_1 \leq m - 1$. In this case we say that the sequence has (or ends with) a cycle of length $m - m_1$.

2. The function $F_{\mathcal{A}}$ is undefined for the argument X_{m-1} , whereas it is defined for all arguments X_j , $j < m - 1$. In this case we write $X_m = \emptyset$ and say that the sequence is a terminating sequence of length m . Thus in this case $en_{\rho}(X_{m-1})$ holds for no reaction ρ in \mathcal{A} .

As an example, let \mathcal{A}_2 be a reaction system over the background set $\{1, 2, 3\}$, consisting of the four reactions

$$\rho_1 = (\{1\}, \{2\}, \{1\}), \quad \rho_2 = (\{1\}, \{3\}, \{2\}), \\ \rho_3 = (\{2\}, \{3\}, \{3\}), \quad \rho_4 = (\{3\}, \{1\}, \{1, 3\})$$

We obtain now a cycle

$$\{1\} \Rightarrow \{1, 2\} \Rightarrow \{2, 3\} \Rightarrow \{1, 3\} \Rightarrow \{1\}$$

of length 4, as well as another cycle

$$\{2\} \Rightarrow \{3\} \Rightarrow \{1, 3\} \Rightarrow \{1\} \Rightarrow \{1, 2\} \Rightarrow \{2, 3\} \Rightarrow \{1, 3\},$$

also of length 4. In the latter case we also have an initial part of length 2, and the whole sequence contains all proper nonempty subsets of the background set. Consequently, the function $F_{\mathcal{A}_2}$ is total.

Examples of long sequences and cycles will be presented in Sect. 4.

In the next section we will consider well-formed formulas, WFF's, Φ built from propositional variables x_1, x_2, \dots , by the use of connectives \sim, \vee, \wedge (negation, disjunction, conjunction). A truth-value assignment for such a formula Φ is a mapping of the set of variables occurring in Φ into the set $\{T, F\}$. For any given truth-value assignment, the truth-value assumed by Φ is then computed in the usual way using the truth-tables of the connectives. A WFF Φ is satisfiable if it assumes the value T for at least one truth-value assignment. The variables x_i and their negations $\sim x_i$ are referred to as literals. A WFF is in 3-conjunctive normal form if it is a conjunction of clauses, each of which is a disjunction of three literals.

For instance, Salomaa (in press), the following WFF in 3-conjunctive normal form

$$(\sim x_1 \vee \sim x_2 \vee \sim x_3) \wedge (x_1 \vee x_2 \vee \sim x_4) \wedge (\sim x_1 \vee x_2 \vee x_4) \\ \wedge (x_1 \vee x_2 \vee \sim x_5) \wedge (x_1 \vee \sim x_2 \vee \sim x_5) \wedge (x_1 \vee x_3 \vee x_4) \\ \wedge (\sim x_1 \vee x_3 \vee \sim x_5) \wedge (x_1 \vee \sim x_4 \vee x_5) \wedge (x_2 \vee \sim x_3 \vee x_4) \\ \wedge (x_3 \vee x_4 \vee x_5) \wedge (x_3 \vee \sim x_4 \vee x_5)$$

is satisfiable, with three satisfying assignments $FTTFF, TFTTF, TFTTT$ for the ordered set of variables $\{x_1, x_2, x_3, x_4, x_5\}$.

It is well known that the satisfiability problem of WFF's in 3-conjunctive normal form, briefly 3-SAT, is NP-complete. This fact was used in Salomaa (in press) to establish the NP-completeness and co-NP-completeness of many problems concerning reaction systems. Some of the

reaction systems used were (3,3) systems and thus far from almost minimal ones. Moreover, the problem *monotone 1-in-3 3SAT*, Schaefer (1978), is also a suitable reduction tool for problems concerning reaction systems. One considers WFF's in 3-conjunctive normal form, where each variable is *unnegated*. (Thus the negation \sim does not appear at all.) One looks for such truth-value assignments for the variables, where *exactly one variable in each clause* assumes the value T . For instance, in the sense of this requirement, the WFF

$$(x_1 \vee x_2 \vee x_3) \wedge (x_1 \vee x_2 \vee x_4) \wedge (x_1 \vee x_3 \vee x_4) \\ \wedge (x_2 \vee x_3 \vee x_4)$$

is not satisfiable, whereas the WFF obtained by removing any of the four clauses is satisfiable.

Also the monotone 1-in-3 3SAT problem is NP-complete, (Schaefer 1978). It is a very suitable tool for reductions involving almost minimal reaction systems. Then also the dummy letters needed for cases where all or none of the three letters in a clause will have the truth-value T , (Ehrenfeucht et al. 2012a; Rozenberg 2011; Salomaa in press), can be totally avoided.

3 NP- and co-NP-complete problems about almost minimal reaction systems

We begin with the definition of a “package” of reactions, useful for many constructions. Modifications of the package will be used in the sequel.

Definition 5 For the background set S and $s \in S$, we define the set of (1,1) reactions $U(S, s)$ by

$$U(S, s) = \{(\{r\}, \{i\}, \{s\}) \mid r, i \in S, r \neq i\}.$$

The presence of the set $U(S, s)$ as a subset of the set of reactions guarantees that the element s is always in the product set, no matter what the initial argument is. Indeed, the following lemma is obvious by the definitions.

Lemma 1 Assume that the set of reactions of a reaction system \mathcal{A}_S contains the set $U(S, s)$ as a subset, for some $s \in S$. Then the function $F_{\mathcal{A}_S}$ is total and, moreover, for any nonempty proper subset X of S ,

$$s \in F_{\mathcal{A}_S}(X).$$

We will now investigate NP-completeness and co-NP-completeness of some decision problems concerning almost minimal reaction systems. No corresponding results are known for minimal reaction systems. In fact, the problems investigated for minimal reaction systems have been shown to be in P. Our basic tool is the

NP-completeness of the monotone 1-in-3 3SAT problem. We begin with the following two problems.

Definition 6 Given a reaction system \mathcal{A} , an element D of the background set of \mathcal{A} , as well as an integer $n \geq 1$, we have to decide whether D occurs at the n th step in every sequence of \mathcal{A} consisting of at least n steps. This problem will be referred to as the convergence problem for reaction systems. The inverse function problem means the following. We are given a reaction system \mathcal{A} and a subset Y of the background set of \mathcal{A} . We have to decide whether or not the function $F_{\mathcal{A}}$ satisfies the equation

$$F_{\mathcal{A}}(X) = Y,$$

for some argument X .

The problems mentioned are natural in the general framework and motivations presented in connection with reaction systems, (Brijder et al. 2011; Rozenberg 2011). When dealing with a known set of reactions, it is often important to know whether a specific (maybe somehow desirable or undesirable) result is obtainable by the reactions, either directly or after a number of steps. The specification of n in the definition of the convergence problem is needed because we consider only sufficiently long sequences. Of course the step 0 is arbitrary in a sequence.

In view of the finiteness of the background set, it is obvious that problems such as the ones in Definition 6 are decidable. It is also clear that they are, depending on the formulation, in NP or in co-NP. In Salomaa (in press), we considered problems for arbitrary reaction systems. We will now prove that the problems in Definition 6 are, in fact, NP-complete or co-NP-complete for almost minimal reaction systems.

Consider a WFF Φ in 3-conjunctive normal form

$$(x_1 \vee y_1 \vee z_1) \wedge \cdots \wedge (x_m \vee y_m \vee z_m),$$

where the variables x_i, y_i, z_i , all of them unnegated, come from a set $U_k = \{u_1, \dots, u_k\}$ consisting of k variables. We say that a truth-value assignment for the variables u_1, \dots, u_k satisfies Φ if, for each i , $1 \leq i \leq m$, exactly one of the variables x_i, y_i, z_i assumes the value T . The WFF Φ is *satisfiable* if there exists a truth-value assignment satisfying it. Moreover, we say that a *subset* U of U_k satisfies Φ if the truth-value assignment assigning the value T (resp. F) to the variables in U (resp. $U_k - U$) satisfies Φ .

We now define a specific almost minimal reaction system associated with the WFF Φ .

Definition 7 Consider the WFF Φ as above, with m clauses and the set of variables U_k . The reaction system $\mathcal{A}(\Phi)$ associated with Φ has the background set $S_{\Phi} = U_k \cup \{A_1, \dots, A_m\}$ and the $3m$ reactions

$$(\{x_i\}, \{y_i, z_i\}, \{A_i\}), (\{y_i\}, \{x_i, z_i\}, \{A_i\}), (\{z_i\}, \{x_i, y_i\}, \{A_i\}), 1 \leq i \leq m.$$

Intuitively, the appearance of A_i in the product set indicates the “acceptance” of the i th clause, $1 \leq i \leq m$. Exactly one variable in the clause assumes the value T .

Lemma 2 *The equation*

$$F_{\mathcal{A}(\Phi)}(X) = \{A_1, \dots, A_m\}$$

holds for a subset X of the set U_k exactly in case X satisfies Φ . There is a subset Y of S_Φ such that

$$F_{\mathcal{A}(\Phi)}(Y) = \{A_1, \dots, A_m\}$$

exactly in case Φ is satisfiable.

Proof The first sentence follows by the definition of the reactions. The second sentence is a consequence of the observation that the presence of some elements A_i in the argument Z does not affect the function value $F_{\mathcal{A}(\Phi)}(Z)$. \square

Lemma 2 now implies the following result.

Theorem 1 *The inverse function problem is NP-complete for almost minimal reaction systems.*

Observe that the reaction systems used in the proof are, in fact, (1, 2) reaction systems.

To deal with the convergence problem, we need a modification of the reaction system $\mathcal{A}(\Phi)$, Definition 7. We use also the construction present in Lemma 1.

We consider still the WFF Φ with k variables and m clauses, as above. Also the notations U_k and S_Φ will be used in the same meaning as above. The set of reactions in the reaction system $\mathcal{A}(\Phi)$ is denoted by Λ .

Definition 8 The second reaction system $\mathcal{B}(\Phi)$, associated with the WFF Φ , has the background set

$$S_{\mathcal{B}(\Phi)} = S_\Phi \cup \{B, C\}.$$

The set of reactions consists of the reactions in Λ and $U(S_{\mathcal{B}(\Phi)}, B)$, as well as of the additional reactions

$$\rho_j = (\{B\}, \{A_j\}, \{C\}), \quad 1 \leq j \leq m.$$

We will now prove that the element C appears at the second step in every sequence of $\mathcal{B}(\Phi)$ exactly in case the WFF Φ is not satisfiable. Because the set $U(S_{\mathcal{B}(\Phi)}, B)$ is included in the set of reactions, the element B appears at every step in every sequence, after the beginning step 0. Consequently, every sequence has at least two steps. (Indeed, every sequence ends with a loop.)

Lemma 3 *Consider a sequence Z_0, Z_1, Z_2, \dots of the reaction system $\mathcal{B}(\Phi)$. Let $V_0 = Z_0 \cap U_k$. Then $C \in Z_2$ exactly in case V_0 does not satisfy Φ .*

Proof Considering the product sets in the reactions of $\mathcal{B}(\Phi)$, we conclude that, for every Z_0 ,

$$Z_1 \subseteq \{B, A_1, \dots, A_m, C\}, \quad B \in Z_1.$$

Apart from the reactions in $U(S_{\mathcal{B}(\Phi)}, B)$, the only reactions possibly enabled for Z_1 are the reactions ρ_j , $1 \leq j \leq m$. The product in the latter reactions is always $\{C\}$. There are two possibilities.

1. At least one reaction ρ_j is enabled for Z_1 . Then $Z_2 = \{B, C\}$.
2. No reaction ρ_j is enabled for Z_1 . Then $Z_2 = \{B\}$.

The first possibility occurs exactly in case V_0 does not satisfy Φ . Consequently, our lemma follows. \square

Lemma 3 shows that the only possibility for C not to occur in Z_2 is that

$$\{A_1, \dots, A_m\} \subseteq Z_1.$$

But this happens exactly in case V_0 satisfies Φ . Hence, we obtain the following result.

Theorem 2 *The convergence problem is co-NP-complete for almost minimal reaction systems.*

Observe again that Theorem 2 holds for (1, 2) reaction systems. In fact, the additional reactions in $\mathcal{B}(\Phi)$ are all (1, 1) reactions.

We have investigated above two typical problems, where the satisfiability or non-satisfiability of a WFF can be applied for complexity issues dealing with almost minimal reaction systems. Analogous results can be obtained for other similar problems, such as the ones investigated in Salomaa (in press). Thereby further modifications of the reaction systems associated with the WFF Φ , notably concerning the behavior of the additional elements of the background set, are needed.

When dealing with minimal reaction systems, the situation is essentially different. The use of techniques such as the ones above lead to 2-conjunctive normal forms and, hence, cannot imply NP-completeness. We hope to return to decision problems concerning minimal reaction systems in another contribution. Problems so far considered are in P. A particularly simple algorithm for *the totality problem*, that is, whether or not a given minimal reaction system defines a total function, was given in Salomaa (2012).

Is a given reaction system equivalent (that is, defines the same function) to a minimal one? The complexity of this question is a very interesting open problem.

4 Functions and long sequences in minimal and almost minimal reaction systems

Although the formal setup in reaction systems is very simple, still amazingly complex phenomena may emerge.

Even in the case of minimal reaction systems, no upper bound polynomial in terms of $\#S$ exists for the length of terminating sequences, (Ehrenfeucht et al. 2011; Salomaa 2012a, b). The best known lower bound is presented in the following theorem, (Salomaa 2012).

Theorem 3 *For every $i \geq 1$, there is a (1,1) reaction system having $4i$ elements in the background set and having a terminating state sequence of length v_i , where*

$$v_1 = 7, \quad v_2 = 26, \quad v_i = 3 \cdot 3^i + 9(3^{i-3} - 1)/2 + 2, \quad i \geq 3.$$

The availability of three resources in reactions (as is the case in almost minimal reaction systems) essentially enhances possibilities for constructions. It is easy to obtain terminating state sequences exponentially longer than the ones in Theorem 3. Consider first the (2, 1) reaction system with the background set $\{a, b, c\}$ and reactions

$$(\{a\}, \{b\}, \{b\}), (\{b\}, \{a\}, \{c\}), (\{c\}, \{a\}, \{a, b\}),$$

$$(\{a, b\}, \{c\}, \{a, c\}), (\{a, c\}, \{b\}, \{b, c\}).$$

We obtain the terminating state sequence

$$\{a\} \Rightarrow \{b\} \Rightarrow \{c\} \Rightarrow \{a, b\} \Rightarrow \{a, c\} \Rightarrow \{b, c\}$$

$$\Rightarrow \{a, b, c\} \Rightarrow \emptyset$$

of length 7. Indeed, this is a longest possible state sequence if the background set S has 3 elements: every subset of S is present. This reaction system \mathcal{A}_1 constitutes the basis of our inductive procedure.

We now assume inductively that, for $i \geq 1$, we have constructed a (2,1) reaction system \mathcal{A}_i having the background set S of cardinality $3i$, as well as a terminating state sequence

$$Z_0 \Rightarrow Z_1 \Rightarrow Z_2 \Rightarrow \dots \Rightarrow Z_{v_i-1} \Rightarrow \emptyset$$

of length v_i . We now construct, for $i + 1$, a (2, 1) reaction system \mathcal{A}_{i+1} having a background set of cardinality $3(i + 1)$ and a terminating state sequence of length $3v_i + 2$.

The background set of \mathcal{A}_{i+1} is $S \cup \{d, e, f\}$, where the elements d, e, f are not in S . We add f to the product set of each reaction in \mathcal{A}_i . Thus, if (R, I, P) is a reaction in \mathcal{A}_i , then the corresponding reaction in \mathcal{A}_{i+1} is $(R, I, P \cup \{f\})$. Moreover, the reaction system \mathcal{A}_{i+1} contains the following additional reactions:

$$(\{d, i\}, \{e\}, \{d\}), (\{d\}, \{f\}, Z_0 \cup \{e, f\}),$$

$$(\{e, i\}, \{d\}, \{e\}), (\{e\}, \{f\}, Z_0),$$

where i ranges over S . Clearly, \mathcal{A}_{i+1} is a (2,1) reaction system.

It is now easy to verify that the following sequence is a legitimate terminating state sequence of the reaction system \mathcal{A}_{i+1} .

$$Z_0 \cup \{d, f\} \Rightarrow Z_1 \cup \{d, f\} \Rightarrow \dots \Rightarrow Z_{v_i-1} \cup \{d, f\} \Rightarrow \{d\}$$

$$\Rightarrow Z_0 \cup \{e, f\} \Rightarrow Z_1 \cup \{e, f\} \Rightarrow \dots$$

$$\Rightarrow Z_{v_i-1} \cup \{e, f\} \Rightarrow \{e\} \Rightarrow Z_0 \Rightarrow Z_1 \cup \{f\}$$

$$\Rightarrow \dots \Rightarrow Z_{v_i-1} \cup \{f\} \Rightarrow \emptyset.$$

Clearly, the length of this sequence is $3v_i + 2$.

As an illustration, consider the reaction system \mathcal{A}_2 . We begin with \mathcal{A}_1 and conclude that \mathcal{A}_2 has the background set $\{a, b, c, d, e, f\}$ and reactions

$$(\{a\}, \{b\}, \{b, f\}), (\{b\}, \{a\}, \{c, f\}), (\{c\}, \{a\}, \{a, b, f\}),$$

$$(\{a, b\}, \{c\}, \{a, c, f\}), (\{a, c\}, \{b\}, \{b, c, f\}), (\{d, i\}, \{e\}, \{d\}),$$

$$(\{d\}, \{f\}, \{a, e, f\}), (\{e, i\}, \{d\}, \{e\}), (\{e\}, \{f\}, \{a\}),$$

where i ranges over the set $\{a, b, c\}$. We obtain now the following terminating state sequence of length 23.

$$\{a, d, f\} \Rightarrow \{b, d, f\} \Rightarrow \{c, d, f\} \Rightarrow \{a, b, d, f\} \Rightarrow \{a, c, d, f\}$$

$$\Rightarrow \{b, c, d, f\} \Rightarrow \{a, b, c, d, f\} \Rightarrow \{d\} \Rightarrow \{a, e, f\} \Rightarrow \{b, e, f\}$$

$$\Rightarrow \{c, e, f\} \Rightarrow \{a, b, e, f\} \Rightarrow \{a, c, e, f\} \Rightarrow \{b, c, e, f\}$$

$$\Rightarrow \{a, b, c, e, f\} \Rightarrow \{e\} \Rightarrow \{a\} \Rightarrow \{b, f\} \Rightarrow \{c, f\}$$

$$\Rightarrow \{a, b, f\} \Rightarrow \{a, c, f\} \Rightarrow \{b, c, f\} \Rightarrow \{a, b, c, f\} \Rightarrow \emptyset.$$

This means that we have the subsequent recurrence relation for the numbers v_i .

$$v_1 = 7, v_{i+1} = 3v_i + 2, \quad \text{for all } i \geq 1.$$

The solution of this recurrence is

$$v_i = (8/3)3^i - 1, \quad i \geq 1.$$

We have, thus, established the following lower bound for the length of state sequences of almost minimal reaction systems.

Theorem 4 *For every $i \geq 1$, there is an almost minimal reaction system having $3i$ elements in the background set and a terminating state sequence of length $(8/3)3^i - 1$.*

The proof above shows that Theorem 4 holds, in fact, for (2, 1) reaction systems. The bound obtained from Theorem 4 is exponentially better than the one obtained from Theorem 3, the orders of growth in the two cases being $x^{\#S}$ with $x = \sqrt[3]{3}$ and $x = \sqrt[4]{3}$ respectively. They are still far from the optimal bound $2^{\#S}$, obtainable if the number of resources is unlimited. An essentially better bound would result if the use of the “almost dummy” letter f could be eliminated.

It is natural to estimate lower bounds in terms of the cardinality of the background set rather than the cardinality of the set of reactions. Maximally inhibited reaction systems generate longest possible sequences and have only one reaction for each set of reactants.

In general, it is easy to obtain long cycles from long terminating sequences, (Ehrenfeucht et al. 2011; Salomaa

2012a). We do this now for almost minimal reaction systems, by a modification of the construction above.

Theorem 5 *For every $i \geq 1$, there is an almost minimal reaction system having $3i+2$ elements in the background set and having a cycle of length $(8/3)3^i$.*

Proof By Theorem 4 there is, for all $i \geq 1$, an almost minimal reaction system \mathcal{A}_i with a background set S of cardinality $3i$ and with a terminating state sequence

$$Z_0 \Rightarrow Z_1 \Rightarrow Z_2 \Rightarrow \dots \Rightarrow Z_{m-1} \Rightarrow \emptyset$$

of length $m = (8/3)3^i - 1$. We now modify \mathcal{A}_i to an almost minimal reaction system \mathcal{B}_i as follows.

The background set of \mathcal{B}_i is obtained by adding two new elements x and y to S . The set of reactions in \mathcal{B}_i is obtained by adding the element y to the product set of every reaction in \mathcal{A}_i and, moreover, taking the additional reactions

$$(\{x\}, \{y\}, Z_0 \cup \{x, y\}), (\{x, j\}, \{k\}, \{x\}), j, k \in S, k \neq j, x.$$

Thus, x is preserved in a sequence as long as there are elements of S present, whereas y is present as long as some reaction is enabled. Consequently, we obtain the cycle

$$\begin{aligned} Z_0 \cup \{x, y\} &\Rightarrow Z_1 \cup \{x, y\} \Rightarrow \dots \Rightarrow Z_{m-1} \cup \{x, y\} \\ &\Rightarrow \{x\} \Rightarrow Z_0 \cup \{x, y\} \end{aligned}$$

of length $m + 1 = (8/3)3^i$.

The following observation should be made concerning the basis of induction on i . If we follow the proof of Theorem 4 with $v_1 = 7$, then the above construction does not work at the first step. The reason is that then $Z_{m-1} \cup \{x, y\}$ is the entire background set, whence no continuation is possible. However, this is no problem because, for $i = 1$, we have now a 5-element background set. Then we immediately get the estimate $v_1 = 7$, without using the entire background set in the terminating sequence. We can easily get better estimates for the starting point, for instance, $v_1 = 21$. Then the recursion

$$v_1 = 21, \quad v_{i+1} = 3v_i + 3, \quad i \geq 1$$

gives the final result

$$(15 \cdot 3^i - 3)/2, i \geq 0.$$

Observe that the initial value v_1 affects only the coefficients α and β in $\alpha \cdot 3^i + \beta$. □

No exhaustive characterization of functions defined by almost minimal reaction systems is known. This was quoted in Ehrenfeucht et al. (2012b) as a challenging open problem. However, it is easy to construct classes of functions definable by almost minimal reaction systems but not definable by minimal reaction systems. We now present two typical examples of such functions. Thereby also the impossibility of using only one reactant (resp. only one

inhibitor) for certain purposes is illustrated. The examples are over the background set $\{a, b, c\}$.

Consider a function f_1 satisfying

$$f_1\{a\} = f_1\{b\} = \{a\}, f_1\{a, b\} = \{c\}.$$

(We are interested only in these values, the values of f_1 for other subsets of the background set are irrelevant.) Such a function f_1 is defined by the (2, 1) reaction system with the three reactions

$$(\{a\}, \{b\}, \{a\}), (\{b\}, \{a\}, \{a\}), (\{a, b\}, \{c\}, \{c\}).$$

A minimal reaction system cannot define such a function f_1 . Since every reaction has only one reactant, the function value for the argument $\{a, b\}$ must be contained in the union of the functions values for the arguments $\{a\}$ and $\{b\}$.

Consider next a function f_2 satisfying

$$f_2\{a, b\} = f_2\{a, c\} = \{b\}, f_2\{a\} = \{c\}.$$

Such a function f_2 is defined by the (1, 2) reaction system with the three reactions

$$(\{a\}, \{b\}, \{b\}), (\{a\}, \{c\}, \{b\}), (\{a\}, \{b, c\}, \{c\}).$$

The requirements for f_2 cannot be met if there is only one inhibitor in each reaction. The negative conclusions for f_1 and f_2 follow also by the general result quoted in the next section.

5 Further considerations and open problems

Maximally inhibited reaction systems, (Salomaa 2012a), give an easy method of defining an arbitrary function. We now quote the important result, (Ehrenfeucht et al. 2012b), characterizing the functions defined by minimal reaction systems. We need the following definition from (Ehrenfeucht et al. 2012b).

Definition 9 A function f (in the sense of Definition 4) is

- union-subadditive if $f(X \cup Y) \subseteq f(X) \cup f(Y)$,
- intersection-subadditive if $f(X \cap Y) \subseteq f(X) \cup f(Y)$,
- for all subsets X and Y of S .

A characterization result for functions defined by minimal reaction systems was given in (Ehrenfeucht et al. 2012b).

Theorem 6 *A function $F_{\mathcal{A}}$ defined by a reaction system \mathcal{A} is defined by a (1,1) reaction system exactly in case $F_{\mathcal{A}}$ is both union-subadditive and intersection-subadditive.*

The characterization given in Theorem 6 is exhaustive. However, since the two conditions needed seem computationally hard to test, it is not clear whether Theorem 6 simplifies the problem mentioned above: Is a given reaction systems equivalent to a minimal one?

We would like to present the following, somewhat vaguely stated,

Conjecture Complexity grows more in the transition from minimal to almost minimal reaction systems, than in the transition from almost minimal to unrestricted reaction systems.

Results in Sect. 3 support this conjecture. As regards the lower bounds for the lengths of terminating sequences and cycles, it seems likely that the bounds in Sect. 4 can be improved to justify the following conclusion. The ratio between the optimal bound 2^i and the lower bound x^i for almost minimal reaction systems is smaller than the ratio between x^i and the lower bound for minimal reaction systems.

Acknowledgements I want to thank the anonymous referees for many useful comments and suggestions.

References

- Brijder R, Ehrenfeucht A, Main M, Rozenberg G (2011) A tour of reaction systems. *Int J Found Comput Sci* 22:1499–1517
- Ehrenfeucht A, Kleijn J, Koutny M, Rozenberg G (2012) A natural computing approach to the functioning of living cells (in press)
- Ehrenfeucht A, Kleijn J, Koutny M, Rozenberg G (2012) Minimal reaction systems (in press)
- Ehrenfeucht A, Main M, Rozenberg G (2011) Functions defined by reaction systems. *Int J Found Comput Sci* 22:167–178
- Ehrenfeucht A, Rozenberg G (2007) Reaction systems. *Fundamenta Informaticae* 7:263–280
- Rozenberg G (2011) Reaction systems. A course of lectures given at the University of Turku in June
- Salomaa A (2012) On state sequences defined by reaction systems. *Springer Lect Notes Comput Sci* 7230:271–282
- Salomaa A (2012) Functions and sequences generated by reaction systems. *Theor Computer Sci* 466:87–96 doi:[10.1016/j.tcs.2012.07.022](https://doi.org/10.1016/j.tcs.2012.07.022)
- Salomaa A Functional constructions between reaction systems and propositional logic. *Int J Found Comput Sci* (in press)
- Schaefer TJ (1978) The complexity of satisfiability problems. *Proceedings of the 10th Annual ACM symposium on theory of computing*. San Diego, pp 216–226. doi:[10.1145/800133.804350](https://doi.org/10.1145/800133.804350)