Reaction Systems Made Simple
A Normal Form and a Classification Theorem

Luca Manzoni and Antonio E. Porreca

Dipartimento di Informatica, Sistemistica e Comunicazione
Università degli Studi di Milano-Bicocca
Viale Sarca 336/14, 20126 Milano, Italy
{luca.manzoni,porreca}@disco.unimib.it

Abstract. Reaction systems are models of computation inspired by the interactions between biochemical reactions. We define a notion of multi-step simulation among reaction systems and derive a classification with respect to the amount of resources (reactants and inhibitors) involved in the reactions. We prove that one reactant and one inhibitor per reaction are sufficient to simulate arbitrary systems. Finally, we show that the equivalence relation of mutual simulation induces exactly five linearly ordered classes of reaction systems.

1 Introduction

Reaction systems, introduced by Ehrenfeucht and Rozenberg [3, 4], are a formalised abstraction of biochemical processes in which the dynamics are discrete in both space and time and are described in terms of *reactions*. A reaction is modelled as a set of *reactants*, necessary for the reaction to take place, a set of *inhibitors*, whose presence blocks the reaction from occurring, and a set of *products*.

Reaction systems may be considered a qualitative model, as opposed to a quantitative one, as we only focus on the presence or absence of chemical species, and not on the precise amounts. In particular, multiple reactions having common reactants do not interfere; indeed, *all* reactions that are enabled at a certain time step happen simultaneously. Another feature of reaction systems which differentiates them from other biologically inspired computational models is the lack of permanency: the state of the system only consists of the products of the reactions that took place in the last time step, without preserving the entities that were not involved in any reaction.

Mathematically, a reaction systems defines a transition function (the *result* function) between states, i.e., sets of entities (chemical species), which completely describes the dynamics of the system. In many cases, the study of the properties of reaction systems involves the comparison of the result functions of different systems or classes of systems. A natural way to understand the modelling power of reaction systems is to consider their behaviour when the amount of resources (reactants and inhibitors per reactions) is limited. It was proved [2, 5] that there

exist infinite proper hierarchies of classes of result functions: by allowing more resources, more functions become definable by reaction systems.

While the analysis of result functions is a direct way to compare reaction systems, the classification it provides has a very high granularity. Requiring the equality of the whole dynamics can be restrictive for certain applications where we are interested in a higher-level view of the behaviour of the systems. As an analogy, consider a simulation between Turing machines: we are often not interested in a step-by-step correspondence of configurations, and we allow the simulation to be slower than the original machine. In a similar fashion, in this paper we define a notion of simulation in which the simulating system is allowed to use several steps to simulate a single step of the other system; auxiliary entities (analogous to an alphabet extension) may also be involved in the simulation. The resulting equivalence relation of mutual simulability is coarser than equality of result functions, but still captures the intuitive idea of "having the same behaviour".

This paper is structured as follows. In Section 2 we recall the definitions and notation related to reaction systems. In Section 3, we introduce the notion of k-simulation and prove that any reaction system can be k-simulated by using only one reactant and one inhibitor per reaction. In Section 4 we study reaction systems with no reactants or no inhibitors, and prove that exactly five linearly ordered equivalence classes exist. Finally, in Section 5 we discuss the results and provide some possible directions for further research.

2 Basic Notions

In this paper we denote sets by upper-case letters, reactions and atomic elements by lower-case letters, and reaction systems by calligraphic letters. Given a set X, we denote by 2^X the power set of X.

A reaction is formally defined as follows.

Definition 1. *Given a finite set S (the* background set*), a* reaction over S *is a triple of sets $a = (R_a, I_a, P_a) \in 2^S \times 2^S \times 2^S$. We call R_a the set of* reactants*, I_a the set of* inhibitors*, and P_a the set of* products*.*

Since we will show that one reactant and one inhibitor suffice to simulate any reaction system (see Theorem 1), in this paper we also admit empty reactant and inhibitor sets, as in the original definition [4], in order to investigate the expressivity of the resulting reactions and to prove that they are strictly weaker than reactions involving both kinds of resources.

Definition 2. *A* reaction system *is a pair $\mathcal{A} = (S, A)$ where S is a finite set and A a set of reactions over S.*[1]

A *state* of a reaction system $\mathcal{A} = (S, A)$ is any subset of S. The dynamics of a reaction systems are defined as follows.

[1] We may assume, without loss of generality, the existence of a countably infinite universe including every background set.

Definition 3. *Let $\mathcal{A} = (S, A)$ be a reaction system, $a = (R_a, I_a, P_a) \in A$, and $T \subseteq S$. We say that a is* enabled by T *iff $R_a \subseteq T$ and $I_a \cap T = \varnothing$.*

The result of a on T *is defined as*

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

The result of $\mathcal{A}$ on T *is defined as*

$$\mathrm{res}_{\mathcal{A}}(T) = \bigcup_{a \in A} \mathrm{res}_a(T).$$

The *state sequence* of a reaction system $\mathcal{A}$ with initial state T is given by successive iterations of the result function:

$$\big(\mathrm{res}^n_{\mathcal{A}}(T)\big)_{n \in \mathbb{N}} = \big(T, \mathrm{res}_{\mathcal{A}}(T), \mathrm{res}^2_{\mathcal{A}}(T), \dots\big).$$

Since the background set of a reaction system is finite, the state space is also finite; hence, every state sequence is ultimately periodic.

3 A Normal Form for Reaction Systems

We begin by observing that, for each reaction system, there exists another reaction system having the same result function (hence the same behaviour) but using only one product per reaction.

Proposition 1 (Brijder, Ehrenfeucht, Rozenberg [1]). *For each reaction system $\mathcal{A} = (S, A)$ there exists a reaction system $\mathcal{A}' = (S, A')$ over the same background set having reactions with at most one product per reaction and such that $\mathrm{res}_{\mathcal{A}}(T) = \mathrm{res}_{\mathcal{A}'}(T)$ for all $T \subseteq S$.* □

We classify reaction systems according to the maximum amount of reactants and inhibitors appearing in their reactions; the number of products is not used as a parameter due to the proposition above.

Definition 4. *For all $i, r \in \mathbb{N}$, we denote by $\mathcal{RS}(r, i)$ the class of reaction systems $\mathcal{A} = (S, A)$ such that, for all $(R_a, I_a, P_a) \in A$, we have $|R_a| \leq r$ and $|I_a| \leq i$. We also define the classes $\mathcal{RS}(\infty, i) = \bigcup_{r \in \mathbb{N}} \mathcal{RS}(r, i)$, $\mathcal{RS}(r, \infty) = \bigcup_{i \in \mathbb{N}} \mathcal{RS}(r, i)$, and $\mathcal{RS}(\infty, \infty) = \bigcup_{r,i \in \mathbb{N}} \mathcal{RS}(r, i)$.*

The classification into classes of the form $\mathcal{RS}(r, i)$ is exhaustive, and the class $\mathcal{RS}(\infty, \infty)$ contains all reaction systems.

In order to compare reaction systems with respect to their ability to generate state sequences, we define a notion of simulation less restrictive than equality of result functions: here, the simulating system may use several steps to simulate a single step of the original system. This is consistent with notions of simulation employed for many computational models (e.g., Turing machines), when we are not interested in the strict correspondence of every pair of configurations, but only in the overall behaviour of the two systems.

Definition 5 (k-simulation). *Let $\mathcal{A} = (S, A)$ and $\mathcal{A}' = (S', A')$, with $S \subseteq S'$, be reaction systems, and let $k \in \mathbb{N}$. We say that $\mathcal{A}'$ k-simulates $\mathcal{A}$ iff, for all $T \subseteq S$ and all $n \in \mathbb{N}$, we have*

$$\mathrm{res}^n_{\mathcal{A}}(T) = \mathrm{res}^{kn}_{\mathcal{A}'}(T) \cap S.$$

In other words, when considering the sequences of states of $\mathcal{A}$ and $\mathcal{A}'$ starting from T, the n-th state of $\mathcal{A}$ coincides with the (kn)-th state of $\mathcal{A}'$ with respect to the elements of S (some auxiliary elements of $S' - S$ may also occur).

We use the notion of k-simulation to define a relation on classes of reaction system.

Definition 6. *Let X and Y be classes of reaction systems, and let $k \in \mathbb{N}$. We define the binary relation $\preceq_k$ as follows: $X \preceq_k Y$ iff for all $\mathcal{A} \in X$ there exists a reaction system in Y that ℓ-simulates $\mathcal{A}$ for some $\ell \leq k$.*

We say that $X \preceq Y$ iff $X \preceq_k Y$ for some $k \in \mathbb{N}$. We write $X \approx_k Y$ if $X \preceq_k Y$ and $Y \preceq_k X$, and $X \approx Y$ for $X \preceq Y \wedge Y \preceq X$. Finally, the notation $X \prec Y$ is shorthand for $X \preceq Y \wedge Y \not\preceq X$.

Notice that $X \subseteq Y$ implies $X \preceq_1 Y$, since any reaction system is trivially 1-simulated by itself.

A k-simulation and an ℓ-simulation can be composed into a $(k\ell)$-simulation.

Lemma 1. *$X \preceq_k Y$ and $Y \preceq_\ell Z$ implies $X \preceq_{k\ell} Z$.* □

From this lemma, we immediately get the following result:

Proposition 2. *The relation $\preceq$ is a preorder. Hence, the relation $\approx$ is an equivalence relation.* □

We now begin analysing the relationships among the classes of reaction systems defined above. The first step is to show that a reaction system can always be simulated using only one reactant.

Lemma 2. *$\mathcal{RS}(r, i) \preceq_3 \mathcal{RS}(1, r)$ for all $r \geq 1, i \geq 1$.*

Proof. Let $\mathcal{A} = (S, A) \in \mathcal{RS}(r, i)$. Let $\mathcal{A}' = (S', A')$ be a reaction system with $S' = S \cup \hat{S} \cup \bar{S} \cup 2^S$, where $\hat{S} = \{\hat{s} : s \in S\}$ and $\bar{S} = \{\bar{s} : s \in S\}$. The power set of S is included in S' in order to encode subsets of S as atomic elements in $\mathcal{A}'$. The set A' contains, for each $s \in S$, the reactions

$$(\varnothing, \{s\}, \{\bar{s}\}) \qquad (\{s\}, \varnothing, \{\hat{s}\}). \tag{1}$$

Furthermore, for each reaction $a = (R_a, I_a, P_a) \in A$, with $R_a = \{x_1, \ldots, x_p\}$ and $I_a = \{y_1, \ldots, y_q\}$, the set A' contains the following reactions:

$$(\varnothing, \{\bar{x}_1, \ldots, \bar{x}_p\}, \{R_a\}) \tag{2}$$

$$(\{\hat{y}_1\}, \varnothing, \{I_a\}), \ldots, (\{\hat{y}_q\}, \varnothing, \{I_a\}) \tag{3}$$

$$(\{R_a\}, \{I_a\}, P_a) \tag{4}$$

In order to prove that $\mathcal{A}'$ 3-simulates $\mathcal{A}$, we show that the following statement holds: if n is a multiple of 3 (i.e., $n = 3m$ for some m), then

$$\mathrm{res}^n_{\mathcal{A}'}(T) \cap S = \mathrm{res}^{n/3}_{\mathcal{A}}(T); \tag{5}$$

if $n = 3m + 1$, then

$$\mathrm{res}^n_{\mathcal{A}'}(T) \cap (\hat{S} \cup \bar{S}) = \big\{\hat{y} : y \in \mathrm{res}^{(n-1)/3}_{\mathcal{A}}(T)\big\} \cup \big\{\bar{x} : x \notin \mathrm{res}^{(n-1)/3}_{\mathcal{A}}(T)\big\}, \tag{6}$$

and if $n = 3m + 2$, then

$$\begin{aligned}\mathrm{res}^n_{\mathcal{A}'}(T) \cap 2^S = {} & \big\{R_a : R_a \subseteq \mathrm{res}^{(n-2)/3}_{\mathcal{A}}(T)\big\} \cup \\ & \big\{I_a : I_a \cap \mathrm{res}^{(n-2)/3}_{\mathcal{A}}(T) \neq \varnothing\big\}.\end{aligned} \tag{7}$$

By induction on n: if $n = 0$, then (5) holds by definition.

If $n > 0$ has the form $3m + 1$, then by induction hypothesis we have

$$\mathrm{res}^{n-1}_{\mathcal{A}'}(T) \cap S = \mathrm{res}^{(n-1)/3}_{\mathcal{A}}(T).$$

Notice that the only reactions producing elements of $\hat{S}$ or $\bar{S}$ are those in (1), which, for every $s \in S$, produce $\hat{s}$ if $s \in \mathrm{res}^{n-1}_{\mathcal{A}'}(T)$, and $\bar{s}$ otherwise. As a consequence, statement (6) holds.

If $n > 0$ has the form $3m + 2$, then by induction hypothesis we have

$$\mathrm{res}^{n-1}_{\mathcal{A}'}(T) \cap (\hat{S} \cup \bar{S}) = \big\{\hat{y} : y \in \mathrm{res}^{(n-2)/3}_{\mathcal{A}}(T)\big\} \cup \big\{\bar{x} : x \notin \mathrm{res}^{(n-2)/3}_{\mathcal{A}}(T)\big\}.$$

The only reactions having elements of 2^S as products are (2) and (3): for every reaction $a \in A$, the set $\{R_a\}$ is produced iff for all $x \in R_a$ we have $\bar{x} \notin \mathrm{res}^{n-1}_{\mathcal{A}'}(T)$, which is equivalent (by induction hypothesis) to $x \in \mathrm{res}^{(n-2)/3}_{\mathcal{A}}(T)$. Furthermore, for every $a \in A$, the set $\{I_a\}$ is produced iff there exists at least one $y \in I_a$ such that $\hat{y} \in \mathrm{res}^{n-1}_{\mathcal{A}'}(T)$, which in turn means that $y \in \mathrm{res}^{(n-2)/3}_{\mathcal{A}}(T)$. Hence, statement (7) holds.

Finally, if $n > 0$ has the form $3m$, by induction hypothesis we have

$$\begin{aligned}\mathrm{res}^{n-1}_{\mathcal{A}'}(T) \cap 2^S = {} & \big\{R_a : R_a \subseteq \mathrm{res}^{(n-3)/3}_{\mathcal{A}}(T)\big\} \cup \\ & \big\{I_a : I_a \cap \mathrm{res}^{(n-3)/3}_{\mathcal{A}}(T) \neq \varnothing\big\}.\end{aligned}$$

The only reactions having products in S are of the form (4). For every reaction $a = (R_a, I_a, P_a) \in A$, the corresponding reaction $(\{R_a\}, \{I_a\}, P_a) \in A'$ is enabled in $\mathcal{A}'$ at time $n - 1$ iff a is enabled in $\mathcal{A}$ at time $\frac{n-3}{3} = \frac{n}{3} - 1$. Hence, the two reaction systems $\mathcal{A}$ and $\mathcal{A}'$ have the same state (ignoring symbols in $S' - S$) at time $\frac{n}{3}$ and n respectively, as required.

In particular, statement (5) holds for all n, i.e., $\mathcal{A}'$ 3-simulates $\mathcal{A}$. □

Now we show that the number of inhibitors can also be reduced to one.

Lemma 3. $\mathcal{RS}(r,i) \preceq_2 \mathcal{RS}(\max\{r,1\},1)$.

Proof. Let $\mathcal{A} = (S,A) \in \mathcal{RS}(r,i)$. Consider the reaction system $\mathcal{A}' = (S',A')$ with $S' = S \cup 2^S$ and A' containing, for each reaction $a = (R_a, I_a, P_a) \in A$ with $I_a = \{s_1, \ldots, s_n\}$, the following reactions:

$$(\{s_1\}, \varnothing, \{I_a\}), \ldots, (\{s_n\}, \varnothing, \{I_a\}) \qquad (R_a, \varnothing, \{R_a\}) \qquad (\{R_a\}, \{I_a\}, P_a).$$

In order to prove that $\mathcal{A}'$ 2-simulates $\mathcal{A}$, we can show by induction that

$$\operatorname{res}^n_{\mathcal{A}'}(T) = \operatorname{res}^{n/2}_{\mathcal{A}}(T) \tag{8}$$

if n is even, and

$$\operatorname{res}^n_{\mathcal{A}'}(T) = \left\{R_a : R_a \subseteq \operatorname{res}^{(n-1)/2}_{\mathcal{A}}(T)\right\} \cup \left\{I_a : I_a \cap \operatorname{res}^{(n-1)/2}_{\mathcal{A}}(T) \neq \varnothing\right\}$$

if n is odd; the conditions are easily seen to hold by considering the form of the reactions of $\mathcal{A}'$. The result then follows immediately from (8). □

By combining the previous two lemmata, we are finally able to show that one reactant and one inhibitor can simulate any reaction systems, thus providing a normal form into which every reaction system can be reduced.

Theorem 1 (Normal form). $\mathcal{RS}(\infty,\infty) \approx_6 \mathcal{RS}(1,1)$.

Proof. By definition we have $\mathcal{RS}(1,1) \subseteq \mathcal{RS}(\infty,\infty)$, thus it follows immediately that $\mathcal{RS}(1,1) \preceq_6 \mathcal{RS}(\infty,\infty)$. Conversely, if $r \geq 1$ and $i \geq 1$ we have $\mathcal{RS}(r,i) \preceq_3 \mathcal{RS}(1,r)$ by Lemma 2 and $\mathcal{RS}(1,r) \preceq_2 \mathcal{RS}(\max\{1,1\},1) = \mathcal{RS}(1,1)$ by Lemma 3, hence $\mathcal{RS}(i,r) \preceq_6 \mathcal{RS}(1,1)$ by transitivity (Lemma 1). Since $\mathcal{RS}(r,i) \preceq_1 \mathcal{RS}(r+1,i+1)$, the result holds even if $i = 0$ or $r = 0$. □

4 Classification of Reaction Systems

Having established a minimum amount of resources needed to simulate general reaction systems, we are interested in analysing the behaviour of weaker systems, i.e., with reactions involving no reactants or no inhibitors.

First of all, we show that the number of reactants in a reaction with no inhibitors can be halved, provided that their initial number is greater than two.

Lemma 4. $\mathcal{RS}(r,0) \preceq_2 \mathcal{RS}\left(\left\lceil \frac{r}{2} \right\rceil, 0\right)$ *for all* $r > 2$.

Proof. Let $\mathcal{A} = (S,A) \in \mathcal{RS}(r,0)$. We simulate $\mathcal{A}$ with another reaction system $\mathcal{A}' = (S',A')$, where $S' = S \cup 2^S$.

Each reaction $a = (R_a, \varnothing, P_a) \in A$ is simulated by at most three reactions in A'. We can write R_a as the union of (non necessarily distinct) sets R_1 and R_2 consisting of at most $\left\lceil \frac{r}{2} \right\rceil$ elements each. The reaction a is then simulated by

$$(R_1, \varnothing, \{R_1\}) \qquad (R_2, \varnothing, \{R_2\}) \qquad (\{R_1, R_2\}, \varnothing, P_a).$$

Notice that $\mathcal{A}' \in \mathcal{RS}\left(\left\lceil \frac{r}{2} \right\rceil, 0\right)$. In order to show that $\mathcal{A}'$ 2-simulates $\mathcal{A}$, we prove by induction on n that

– if n is even, then $\mathrm{res}^n_{\mathcal{A}'}(T) \cap S = \mathrm{res}^{n/2}_{\mathcal{A}}(T)$;
– if n is odd, then for all $a = (R_a, \varnothing, P_a) \in A$ we have $R_a \subseteq \mathrm{res}^{(n-1)/2}_{\mathcal{A}}(T)$ iff there exist (non necessarily distinct) sets $Q_1, Q_2 \subseteq S$ such that $R_a = Q_1 \cup Q_2$ and $Q_1, Q_2 \in \mathrm{res}^n_{\mathcal{A}'}(T)$.

If $n = 0$, then the condition clearly holds, since $\mathrm{res}^0_{\mathcal{A}'}(T) \cap S = T = \mathrm{res}^0_{\mathcal{A}}(T)$.

If $n > 0$ is odd, by induction hypothesis $\mathrm{res}^{n-1}_{\mathcal{A}'}(T) \cap S = \mathrm{res}^{(n-1)/2}_{\mathcal{A}}(T)$. Hence, for all $a \in A$, we have $R_a \subseteq \mathrm{res}^{(n-1)/2}_{\mathcal{A}}(T)$ iff $R_a \subseteq \mathrm{res}^{n-1}_{\mathcal{A}'}(T) \cap S$; this is equivalent to the existence of $Q_1, Q_2 \subseteq \mathrm{res}^{n-1}_{\mathcal{A}'}(T)$ such that $Q_1 \cup Q_2 = R_a$; in particular, by letting $Q_1 = R_1$ and $Q_2 = R_2$ as described above, we get $Q_1, Q_2 \in \mathrm{res}^n_{\mathcal{A}'}(T)$ by applying the reactions. Conversely, if there exist sets $Q_1, Q_2 \subseteq S$ such that $R_a = Q_1 \cup Q_2$ and $Q_1, Q_2 \in \mathrm{res}^n_{\mathcal{A}'}(T)$, these are necessarily produced by the two reactions $(Q_1, \varnothing, \{Q_1\})$ and $(Q_2, \varnothing, \{Q_2\})$, implying $Q_1, Q_2 \subseteq \mathrm{res}^{n-1}_{\mathcal{A}'}(T)$, that is $R_a \subseteq \mathrm{res}^{n-1}_{\mathcal{A}'}(T) \cap S = \mathrm{res}^{(n-1)/2}_{\mathcal{A}}(T)$.

Now assume $n > 0$ and even. Let $x \in \mathrm{res}^n_{\mathcal{A}'}(T) \cap S$. Then $x \in P_a$ for some reaction $a' = (\{R_1, R_2\}, \varnothing, P_a) \in A'$ enabled at time $n-1$, hence $R_1, R_2 \in \mathrm{res}^{n-1}_{\mathcal{A}'}(T)$. The reaction a' has a corresponding reaction $a = (R_a, \varnothing, P_a) \in A$ with $R_a = R_1 \cup R_2$. Then, by induction hypothesis, we have $R_a \subseteq \mathrm{res}^{(n-2)/2}_{\mathcal{A}}(T)$: reaction a is enabled in $\mathcal{A}$ at time $\frac{n-2}{2}$, producing x at time $\frac{n}{2}$. Conversely, let $x \in \mathrm{res}^{n/2}_{\mathcal{A}}(T)$. Then $x \in P_a$ for some reaction $a = (R_a, \varnothing, P_a) \in A$ and $R_a \subseteq \mathrm{res}^{(n-2)/2}_{\mathcal{A}}(T)$; by induction hypothesis then $R_a \subseteq \mathrm{res}^{n-2}_{\mathcal{A}'}(T) \cap S$. Since A' contains reactions of the forms $(R_1, \varnothing, \{R_1\})$, $(R_2, \varnothing, \{R_2\})$, and $(\{R_1, R_2\}, \varnothing, P_a)$ with $R_1 \cup R_2 = R_a$, the element x is produced in two steps, i.e., $x \in \mathrm{res}^n_{\mathcal{A}'}(T) \cap S$.

In particular, we have $\mathrm{res}^{2n}_{\mathcal{A}'}(T) \cap S = \mathrm{res}^n_{\mathcal{A}}(T)$ for all $n \in \mathbb{N}$. □

By iterating Lemma 4 the number of reactants can be reduced to two.

Proposition 3. $\mathcal{RS}(\infty, 0) \approx \mathcal{RS}(2, 0)$.

Proof. By definition we have $\mathcal{RS}(2, 0) \preceq_1 \mathcal{RS}(\infty, 0)$. By applying Lemma 4 repeatedly we obtain $\mathcal{RS}(r, 0) \preceq_r \mathcal{RS}(2, 0)$, implying $\mathcal{RS}(\infty, 0) \preceq \mathcal{RS}(2, 0)$. □

In a reaction system where only one reactant per reaction is allowed, each element appearing at a given time in the state of the system can be either traced back to a single element of the initial state, or it is always generated, independently of the initial state.

Lemma 5. *Let $\mathcal{A} = (S, A) \in \mathcal{RS}(1, 0)$. Then, for all $T \subseteq S$, for all $n \in \mathbb{N}$, for all $x \in \mathrm{res}^n_{\mathcal{A}}(T)$ either there exists $y \in T$ such that $x \in \mathrm{res}^n_{\mathcal{A}}(\{y\})$, or we have $x \in \mathrm{res}^n_{\mathcal{A}}(\varnothing)$.*

Proof. By induction on n. When $n = 0$ we have $\mathrm{res}^0_{\mathcal{A}}(T) = T$ and $x \in T$. Hence, it suffices to choose $y = x$.

Now assume $n > 0$ and let $x \in \mathrm{res}^n_{\mathcal{A}}(T)$. There are two cases: either x is generated by a reaction $(\varnothing, \varnothing, \{x\}) \in A$, or by a reaction $(\{z\}, \varnothing, \{x\}) \in A$ for some $z \in S$. In the first case, we have $x \in \mathrm{res}^n_{\mathcal{A}}(\varnothing)$. Otherwise, by induction hypothesis, there are two sub-cases:

– either $z \in \mathrm{res}_{\mathcal{A}}^{n-1}(\varnothing)$, and then $x \in \mathrm{res}_{\mathcal{A}}^{n}(\varnothing)$, or
– there exists $y \in T$ such that $z \in \mathrm{res}_{\mathcal{A}}^{n-1}(\{y\})$, and then $x \in \mathrm{res}_{\mathcal{A}}^{n}(\{y\})$. □

As a consequence, reaction systems where two reactants per reaction are allowed can produce more complex state sequences than those with only one reactant, since the generation of products may depend on the simultaneous presence of several reactants.

Proposition 4. $\mathcal{RS}(1,0) \prec \mathcal{RS}(2,0)$.

Proof. Clearly $\mathcal{RS}(1,0) \preceq_1 \mathcal{RS}(2,0)$.

Let $\mathcal{A} = (S, A) \in \mathcal{RS}(2,0)$ be defined by $S = \{x, y, z\}$ and the reaction $(\{x, y\}, \varnothing, \{z\})$. Suppose that $\mathcal{A}' \in \mathcal{RS}(1,0)$ k-simulates $\mathcal{A}$ for some k. We have $\mathrm{res}_{\mathcal{A}}(\{x, y\}) = \{z\}$, hence $\mathrm{res}_{\mathcal{A}'}^{k}(\{x, y\}) \cap S = \{z\}$. By Lemma 5, one of the two following conditions holds:

$$z \in \mathrm{res}_{\mathcal{A}'}^{k}(\varnothing) \tag{9}$$

$$z \in \mathrm{res}_{\mathcal{A}'}^{k}(\{w\}) \quad \text{for some } w \in \{x, y\}. \tag{10}$$

If (9) holds, then we have $z \in \mathrm{res}_{\mathcal{A}'}^{k}(\varnothing) \cap S \neq \mathrm{res}_{\mathcal{A}}(\varnothing) = \varnothing$, contradicting the fact that $\mathcal{A}'$ k-simulates $\mathcal{A}$. On the other hand, if (10) holds, we have $z \in \mathrm{res}_{\mathcal{A}'}^{k}(\{w\}) \cap S \neq \mathrm{res}_{\mathcal{A}}(\{w\}) = \varnothing$, once again a contradiction.

Therefore $\mathcal{A}$ cannot be k-simulated by any reaction system in $\mathcal{RS}(1,0)$. □

In the absence of inhibitors, at least one reactant is needed in order to obtain state sequences that actually depend on the initial state.

Proposition 5. $\mathcal{RS}(0,0) \prec \mathcal{RS}(1,0)$.

Proof. Clearly $\mathcal{RS}(0,0) \preceq_1 \mathcal{RS}(1,0)$.

We prove that there exists $\mathcal{A} = (A, S) \in \mathcal{RS}(1,0)$ such that no $\mathcal{A}' = (A', S') \in \mathcal{RS}(0,0)$ simulates $\mathcal{A}$. Observe that $\mathrm{res}_{\mathcal{A}'}^{n}(T) = \mathrm{res}_{\mathcal{A}'}(\varnothing)$ for all $T \subseteq S'$ and $n \geq 1$, i.e, the evolution of $\mathcal{A}'$ reaches a fixed point immediately after the first step, irrespective of the initial state. On the other hand, if $\mathcal{A}$ is defined by $S = \{x\}$ with the reaction $(\{x\}, \varnothing, \{x\})$, we have

$$\mathrm{res}_{\mathcal{A}}(\varnothing) = \varnothing \neq \{x\} = \mathrm{res}_{\mathcal{A}}(\{x\}).$$

Hence $\mathcal{RS}(1,0) \not\preceq \mathcal{RS}(0,0)$. □

Unlike reactants, any number of inhibitors can be simulated by a single one.

Proposition 6. $\mathcal{RS}(0,\infty) \approx_3 \mathcal{RS}(0,1)$.

Proof. Trivially, $\mathcal{RS}(0,1) \preceq_3 \mathcal{RS}(0,\infty)$ holds.

Let $\mathcal{A} = (S, A) \in \mathcal{RS}(0,\infty)$, and let $\mathcal{A}' = (S', A') \in \mathcal{RS}(0,1)$ with $S' = S \cup \bar{S} \cup 2^S$, where $\bar{S} = \{\bar{x} : x \in S\}$. For each $x \in S$, A' contains the reaction

$$(\varnothing, \{x\}, \{\bar{x}\}) \tag{11}$$

and, for each $a = (\varnothing, I_a, P_a)$ with $I_a = \{x_1, \ldots, x_p\}$, the reactions

$$(\varnothing, \{\bar{x}_1\}, \{I_a\}), \ldots, (\varnothing, \{\bar{x}_p\}, \{I_a\}) \tag{12}$$
$$(\varnothing, \{I_a\}, P_a). \tag{13}$$

We prove, by induction on n, that for all $T \subseteq S$ we have

$$\mathrm{res}^n_{\mathcal{A}'}(T) \cap S = \mathrm{res}^{n/3}_{\mathcal{A}}(T) \qquad \text{if } n = 3m; \tag{14}$$
$$\bar{x} \in \mathrm{res}^n_{\mathcal{A}'}(T) \iff x \notin \mathrm{res}^{(n-1)/3}_{\mathcal{A}}(T) \qquad \text{if } n = 3m+1; \tag{15}$$
$$I_a \in \mathrm{res}^n_{\mathcal{A}'}(T) \cap 2^S \iff I_a \cap \mathrm{res}^{(n-2)/3}_{\mathcal{A}}(T) \neq \varnothing \qquad \text{if } n = 3m+2. \tag{16}$$

For $n = 0$, we have $\mathrm{res}^0_{\mathcal{A}'}(T) \cap S = T = \mathrm{res}^0_{\mathcal{A}}(T)$.

If $n > 0$ is a multiple of 3, then by induction hypothesis

$$I_a \in \mathrm{res}^{n-1}_{\mathcal{A}'}(T) \cap 2^S \iff I_a \cap \mathrm{res}^{(n-3)/3}_{\mathcal{A}}(T) \neq \varnothing.$$

Notice that, if $X \in \mathrm{res}^{n-1}_{\mathcal{A}'}(T) \cap 2^S$, then necessarily $X = I_a$ for some $a \in A$, as the only reactions producing elements of 2^S have the form (12). For each reaction $a \in A$ we have a corresponding reaction a' of type (13), and a is inhibited at time $\frac{n-3}{3}$ in $\mathcal{A}$ iff $'$ is inhibited at time $n-1$ in $\mathcal{A}'$: statement (14) follows.

If $n > 0$ with $n = 3m+1$, by induction hypothesis we have

$$\mathrm{res}^{n-1}_{\mathcal{A}'}(T) \cap S = \mathrm{res}^{(n-1)/3}_{\mathcal{A}}(T).$$

We have $\bar{x} \in \mathrm{res}^n_{\mathcal{A}'}(T)$ iff the reaction $(\varnothing, \{x\}, \{\bar{x}\})$ was enabled at time $n-1$, that is $x \notin \mathrm{res}^{n-1}_{\mathcal{A}'}(T) \cap S = \mathrm{res}^{(n-1)/3}_{\mathcal{A}}(T)$ as required.

Finally, if $n > 0$ with $n = 3m+2$, by induction hypothesis

$$\bar{x} \in \mathrm{res}^{n-1}_{\mathcal{A}'}(T) \iff x \notin \mathrm{res}^{(n-2)/3}_{\mathcal{A}}(T).$$

Let $a \in A$. We have $I_a \in \mathrm{res}^n_{\mathcal{A}'}(T) \cap 2^S$ iff at least one of the reactions of the form (12) was enabled at time $n-1$. This means that there exists $x \in I_a$ such that $\bar{x} \notin \mathrm{res}^{n-1}_{\mathcal{A}'}(T)$ and $x \in \mathrm{res}^{(n-2)/3}_{\mathcal{A}}(T)$. Equivalently, $I_a \cap \mathrm{res}^{(n-2)/3}_{\mathcal{A}}(T) \neq \varnothing$. This proves (16).

The statement of the proposition immediately follows from (14). □

Perhaps surprisingly, reactants can also be simulated by a single inhibitor.

Lemma 6. $\mathcal{RS}(\infty, 0) \preceq_2 \mathcal{RS}(0, 1)$.

Proof. Let $\mathcal{A} = (S, A) \in \mathcal{RS}(\infty, 0)$. Let $\mathcal{A}' = (S', A') \in \mathcal{RS}(0, 1)$ with $S' = S \cup 2^S$ and having, for each reaction $a = (R_a, \varnothing, P_a)$ with $R_a = x_1, \ldots, x_p$, the following set of reactions:

$$(\varnothing, \{x_1\}, \{R_a\}), \ldots, (\varnothing, \{x_p\}, \{R_a\}) \tag{17}$$
$$(\varnothing, \{R_a\}, P_a). \tag{18}$$

Let $T \subseteq S$. We prove, by induction on n, that

$$\text{res}^n_{\mathcal{A}'}(T) \cap S = \text{res}^{n/2}_{\mathcal{A}}(T) \qquad \text{if } n \text{ is even} \tag{19}$$

$$R_a \in \text{res}^n_{\mathcal{A}'}(T) \iff R_a \not\subseteq \text{res}^{(n-1)/2}_{\mathcal{A}}(T) \qquad \text{if } n \text{ is odd.} \tag{20}$$

For $n = 0$ we have $\text{res}^0_{\mathcal{A}'}(T) \cap S = T = \text{res}^0_{\mathcal{A}}(T)$.

For even $n > 0$ we have, by induction hypothesis,

$$R_a \in \text{res}^{n-1}_{\mathcal{A}'}(T) \iff R_a \not\subseteq \text{res}^{(n-2)/2}_{\mathcal{A}}(T).$$

Notice that the only reactions in A' with products in S have the form (18), and they are enabled at time $n-1$ iff $R_a \subseteq \text{res}^{(n-2)/2}_{\mathcal{A}}(T)$, i.e., iff reaction a is enabled in $\mathcal{A}$ at time $\frac{n-2}{2}$. Condition (19) follows.

For odd $n > 0$, by induction hypothesis we have

$$\text{res}^{n-1}_{\mathcal{A}'}(T) \cap S = \text{res}^{(n-1)/2}_{\mathcal{A}}(T)$$

The only reactions of $\mathcal{A}'$ having products in 2^S have the form (17). The element R_a is produced iff there exists $x \in R_a$ with $x \notin \text{res}^{n-1}_{\mathcal{A}'}(T) \cap S$, i.e., iff reaction a is *not* enabled in $\mathcal{A}$ at time $\frac{n-1}{2}$, as in (20).

The statement of the lemma follows from (19). □

The following two lemmata provide a property of the result function of reaction systems without inhibitors and without reactants, respectively.

Lemma 7. *If $\mathcal{A} = (S, A) \in \mathcal{RS}(\infty, 0)$, then the function $\text{res}_{\mathcal{A}}$ is monotone, i.e., $T_1 \subseteq T_2$ implies $\text{res}_{\mathcal{A}}(T_1) \subseteq \text{res}_{\mathcal{A}}(T_2)$ for all $T_1, T_2 \subseteq S$. As a consequence, the function $\text{res}^n_{\mathcal{A}}$ is monotone for all $n \in \mathbb{N}$.*

Proof. For each reaction $a \in A$, if a is enabled by T_1 then it is also enabled by $T_2 \supseteq T_1$, as a has no inhibitors. Thus, $\text{res}_a(T_1) \subseteq \text{res}_a(T_2)$, and

$$\text{res}_{\mathcal{A}}(T_1) = \bigcup_{a \in A} \text{res}_a(T_1) \subseteq \bigcup_{a \in A} \text{res}_a(T_2) = \text{res}_{\mathcal{A}}(T_2).$$

The function $\text{res}^n_{\mathcal{A}}$ is monotone by induction on n. □

In a similar way, the next result can be proved.

Lemma 8. *If $\mathcal{A} = (S, A) \in \mathcal{RS}(0, \infty)$, then the function $\text{res}_{\mathcal{A}}$ is anti-monotone, i.e., $T_1 \subseteq T_2$ implies $\text{res}_{\mathcal{A}}(T_1) \supseteq \text{res}_{\mathcal{A}}(T_2)$ for all $T_1, T_2 \subseteq S$. As a consequence, the function $\text{res}^n_{\mathcal{A}}$ is anti-monotone for odd n, and monotone for even n.* □

These properties imply that reaction systems using one inhibitor exclusively can produce state sequences that no reaction system using only two reactants (or, by Proposition 3, any number of reactants) can generate.

Proposition 7. $\mathcal{RS}(2,0) \prec \mathcal{RS}(0,1)$.

Proof. By Lemma 6 we have $\mathcal{RS}(2,0) \preceq_2 \mathcal{RS}(0,1)$.

Let $\mathcal{A} = (S, A) \in \mathcal{RS}(0,1)$ be defined by $S = \{x\}$ and $(\varnothing, \{x\}, \{x\})$ as the only reaction. By Lemma 8, the function $\mathrm{res}_{\mathcal{A}}$ is anti-monotone (furthermore, it is not monotone as it is not the identity function). By Lemma 7, for any $\mathcal{A}' \in \mathcal{RS}(2,0)$ the function $\mathrm{res}^k_{\mathcal{A}'}$ is monotone for all $k \in \mathbb{N}$. Therefore, $\mathcal{A}'$ cannot k-simulate $\mathcal{A}$. □

Finally, we show that both reactants and inhibitors are needed in order to simulate arbitrary state sequences, thus proving the minimality of the normal form of Theorem 1.

Proposition 8. $\mathcal{RS}(0,1) \prec \mathcal{RS}(1,1)$.

Proof. Trivially, we have $\mathcal{RS}(0,1) \preceq_1 \mathcal{RS}(1,1)$.

Consider the reaction system $\mathcal{A} = (S, A) \in \mathcal{RS}(1,1)$ defined by $S = \{x, y\}$ and the reaction $(\{x\}, \{y\}, \{x, y\})$. We have

$$\mathrm{res}_{\mathcal{A}}(\varnothing) = \varnothing \qquad \mathrm{res}_{\mathcal{A}}(\{x\}) = \{x, y\} \qquad \mathrm{res}_{\mathcal{A}}(\{x, y\}) = \varnothing.$$

Hence, $\mathrm{res}_{\mathcal{A}}$ is neither monotone nor anti-monotone. No $\mathcal{A}' \in \mathcal{RS}(0,1)$ can simulate $\mathcal{A}$, since (by Lemma 8) the function $\mathrm{res}^k_{\mathcal{A}'}$ is monotone for even k and anti-monotone for odd k. □

All the results proved in this paper can be summarised by the following theorem, which provides a complete classification of reaction systems with respect to the number of reactants and inhibitors per reaction.

Theorem 2. *The relation $\preceq$ is a total preorder on the set of classes of reaction systems of the form $\mathcal{RS}(r,i)$. The classes are comparable according to the following diagram for all $r, i \geq 2$:*

$$\begin{array}{ccccccc}
\mathcal{RS}(0,0) & \prec & \mathcal{RS}(1,0) & \prec & \mathcal{RS}(2,0) & \prec & \mathcal{RS}(0,1) & \prec & \mathcal{RS}(1,1) \\
 & & & & \wr & & \wr & & \wr \\
 & & & & \mathcal{RS}(r,0) & \prec & \mathcal{RS}(0,i) & \prec & \mathcal{RS}(r,i) \\
 & & & & \wr & & \wr & & \wr \\
 & & & & \mathcal{RS}(\infty,0) & \prec & \mathcal{RS}(0,\infty) & \prec & \mathcal{RS}(\infty,\infty)
\end{array}$$

In particular, the relation $\approx$ induces exactly five equivalence classes. □

5 Further Remarks

After having introduced the notion of k-simulation, we have proved that every reaction system $\mathcal{A} = (S, A) \in \mathcal{RS}(r,i)$ can be simulated by using one reactant

and one inhibitor per reaction. We have then analysed reaction systems with no reactants or no inhibitors, showing that there exist a finite, linear hierarchy of non-equivalent classes of reaction systems.

The simulating reaction system $\mathcal{A}'$ has a linear slowdown, that is, simulating n steps of the original system is performed in kn steps, and k is usually independent of $|S|$, $|A|$, r, and i. The only exception is the simulation in $O(rn)$ steps of an $\mathcal{RS}(r, 0)$ reaction system by means of an $\mathcal{RS}(2, 0)$ reaction system. Furthermore, the size of $\mathcal{A}'$ can be always made polynomial with respect to the size of $\mathcal{A}$: even if we often include, for the sake of simplicity, the whole power set of S in the background set of $\mathcal{A}'$, only a polynomial number of elements (depending on $|S|$ and on the number of reactions of $\mathcal{A}$) actually appear as reactants, inhibitors, or products: hence, the remaining ones can be simply removed. The number of reactions of $\mathcal{A}'$ is also polynomial with respect to the size of $\mathcal{A}$.

Although in this paper we focused only on system were the input is all given in the initial state, the original definition [4] allows the system to receive further input (i.e., new elements to be inserted in the state of the system) at every step. It is possible to extend the definition of k-simulation to this case, and prove all the results above in the new setting.

5.1 Open Problems

An open problem involves the minimality of k in certain k-simulations described here. For instance, the k-simulation of Proposition 6 is provably minimal by anti-monotonicity (Lemma 8); similarly, in Lemma 6 the simulation cannot be performed in one step (by Lemmata 7 and 8). In the r-simulation of Proposition 3 the dependency on r is probably unavoidable. On the other hand, it is unknown whether there exist reaction systems necessarily requiring a 6-simulation in order to reduce reactants and inhibitors to one. Furthermore, most k-simulations in this paper employ auxiliary elements, and it seems unlikely that they can always be eliminated. Can we at least ensure that every (kn)-th state of the simulating system is identical to the n-th state of the simulated one?

References

1. Brijder, R., Ehrenfeucht, A., Rozenberg, G.: Reaction systems with duration. In: Kelemen, J., Kelemenová, A. (eds.) Computation, Cooperation, and Life, Lecture Notes in Computer Science, vol. 6610, pp. 191–202. Springer (2011)
2. Ehrenfeucht, A., Main, M., Rozenberg, G.: Functions defined by reaction systems. International Journal of Foundations of Computer Science 22(1), 167–168 (2011)
3. Ehrenfeucht, A., Rozenberg, G.: Basic notions of reaction systems. In: Calude, C.S., Calude, E., Dinneen, M.J. (eds.) Developments in Language Theory, 8th International Conference, DLT 2004, Lecture Notes in Computer Science, vol. 3340, pp. 27–29. Springer (2005)
4. Ehrenfeucht, A., Rozenberg, G.: Reaction systems. Fundamenta Informaticae 75, 263–280 (2007)
5. Salomaa, A.: Functions and sequences generated by reaction systems. Theoretical Computer Science 466, 87–96 (2012)