Studying Opacity of Reaction Systems through Formula Based Predictors

Roberta Gori[1], Damas Gruska[2], and Paolo Milazzo[1]

[1] Dipartimento di Informatica, Università di Pisa, Italy
[2] Department of Applied Informatics, Faculty of Mathematics, Physics and Informatics, Comenius University in Bratislava, Slovak Republic

Abstract. *Reaction systems* are a qualitative formalism for modeling systems of biochemical reactions characterized by the *non-permanency* of the elements: objects (representing molecules) disappear if not produced by any enabled reaction. Reaction systems execute in an environment that provides new objects at each step. Causality properties of reaction systems can be studied by using notions of formula based predictor. In this context, we define a notion of opacity that can be used to study information flow properties for reaction systems. Objects will be partitioned into high level (invisible) and low level (visible) ones. Opacity ensures that the presence (or absence) of high level objects cannot be guessed observing the low level objects only. Such a property is shown to be decidable and computable by exploiting the algorithms for minimal formula based predictors.

1 Introduction

Reaction systems [11, 7] were introduced by Ehrenfeucht and Rozenberg and are based on two opposite mechanisms, namely *facilitation* and *inhibition*. Facilitation means that a reaction can occur only if all its reactants are present, while inhibition means that the reaction cannot occur if any of its inhibitors is present. A rewrite rule of a reaction system (called *reaction*) is hence a triple (R, I, P), where R, I and P are sets of objects representing reactants, inhibitors and products of the modeled biochemical reaction. A *reaction system* is represented by a set of reactions having such a form, together with a (finite) support set S containing all of the objects that can appear in a reaction. The state of a reaction system consists of a finite set of objects, describing the biological entities that are present in the real system being modeled. In particular, the presence of an object in the state expresses the fact that the corresponding biological entity, in the real system being modeled, is present in a number of copies as high as needed. This is the *threshold supply* assumption and characterizes reaction systems as a *qualitative* modeling formalism.

A reaction system evolves by means of the application of its reactions. A reaction is applicable if its reactants are present and its inhibitor are not present in the current state of the system. The threshold supply assumption ensures that the application of different reactions never compete for their reactants, and hence all the applicable reactions in a step are always applied. The result of the application of a set of reactions results in the introduction of all of their products in the next state of the system. Reaction systems assume the *non permanency* of the elements, namely unused elements are never carried over to the next state.

The behavior of a reaction system model is driven by the (set of) contextual elements which are provided by the external environment at each step. Such elements join the current state of the system and can enable or disable reactions. The computation of the next state of a reaction system is a deterministic procedure. Consequently, if the contextual elements provided to the system at each step are know, then the whole execution of the system is determined. On the other hand, if the contextual elements provided at each step are not known, the description of the overall system dynamics becomes non deterministic.

Addressing causal relationships among the actions performed by a system is a very relevant issue in system biology (see e.g. [13, 5, 6]). Brijder, Ehrenfeucht and Rozenberg initiated an investigation of *causalities* in reaction systems [8], i.e. the ways that entities of a reaction system influence each other, introducing the idea of *predictor*. Assume that one is interested in a particular object $s \in S$ and in knowing if that object s will be present after n steps of execution of the reaction system. Since the only source of non-determinism are the contextual elements received at each step, knowing which objects will be received at each step can allow the creation of s after n steps to be predicted. In [2] the study of causal dependencies was enhanced by introducing the new notion of *formula based predictor*. A formula based predictor consists in a propositional logic formula to be satisfied by the sequence of (sets of) elements provided by the environment. *Satisfaction of the logic formula precisely discriminates the cases in which s will be produced after n steps from those in which it will not.* Minimal formula based predictors exist (for a given object s and n) and can be effectively computed. However, traditional predictors do not assume any knowledge on elements provided by the environment, even if often the sequences of sets of objects provided by the environment follow specific patterns. For this reasons, in [1, 3], the notion of formula based predictors was specialized in order to take into account the subset of context sequences that the environment could actually provide. Also in this case, minimal formulas exist and can be effectively computed.

In this paper we study reaction systems with two types of objects. Visible ones, so called low level (L) objects and high level (H) objects which are invisible for an external observer. We investigate how much information on the presence (or absence) of H-objects can an observer obtain just by observing the presence of L-objects. This problem, called information flow (see [12]), is widely studied in security theory and has found many formalizations. Among them, *opacity* is one of the most universal ([9, 10]). Here, we reformulate opacity for reaction systems and use formula based predictors in order to characterize the causal dependences between low level and high level objects. This characterization of opacity properties in terms of dynamic causalities (i.e., predictors) provides effective and efficient methods to prove information flow properties in reactions systems.

2 Reaction Systems

In this section we recall the basic definition of reaction systems [11, 7]. Let S be a finite set of symbols, called objects. A *reaction* is formally a triple (R, I, P) with $R, I, P \subseteq S$, composed of *reactants* R, *inhibitors* I, and *products* P. We assume reactants and inhibitors to be disjoint ($R \cap I = \emptyset$), otherwise the reaction would never be applicable. Reactants and inhibitors $R \cup I$ of a reaction are collectively called *resources* of such a reaction. The set of all possible reactions over a set S is denoted by $\mathrm{rac}(S)$. Finally, a *reaction system* is a

pair $\mathcal{A} = (S, A)$, with S being the finite background set, and $A \subseteq \text{rac}(S)$ being its set of reactions.

The state of a reaction system is described by a set of objects. Let $a = (R_a, I_a, P_a)$ be a reaction and T a set of objects. The result $\text{res}_a(T)$ of the application of a to T is either P_a, if T separates R_a from I_a (i.e. $R_a \subseteq T$ and $I_a \cap T = \emptyset$), or the empty set $\emptyset$ otherwise. The application of multiple reactions at the same time occurs without any competition for the used reactants (threshold supply assumption). Therefore, each reaction for which no inhibitor is present in the current state can be applied, and the result of application of multiple reactions is cumulative. Formally, given a reaction system $\mathcal{A} = (S, A)$, the result of application of $\mathcal{A}$ to a set $T \subseteq S$ is defined as $\text{res}_{\mathcal{A}}(T) = \text{res}_A(T) = \bigcup_{a \in A} \text{res}_a(T)$.

The dynamics of a reaction system is driven by the *contextual* objects, namely the objects which are supplied to the system by the external environment at each step. An important characteristic of reaction systems is the assumption about the *non-permanency* of objects. Under such an assumption the objects carried over to the next step are only those produced by reactions.

Formally, the dynamics of a reaction system $\mathcal{A} = (S, A)$ is defined as an *interactive process* $\pi = (\gamma, \delta)$, with γ and δ being finite sequences of sets of objects called the *context sequence* and the *result sequence*, respectively. The sequences are of the form $\gamma = C_0, C_1, \ldots, C_n$ and $\delta = D_0, D_1, \ldots, D_n$ for some $n \geq 1$, with $C_i, D_i \subseteq S$, and $D_0 = \emptyset$. Each set D_i, for $i \geq 1$, in the result sequence is obtained from the application of reactions A to a state composed of both the results of the previous step D_{i-1} and the objects C_{i-1} from the context; formally $D_i = res_{\mathcal{A}}(C_{i-1} \cup D_{i-1})$ for all $1 \leq i \leq n$. Finally, the *state sequence* of π is defined as the sequence $W_0, W_1, \ldots, W_n$, where $W_i = C_i \cup D_i$ for all $1 \leq i \leq n$. In the following we say that $\gamma = C_0, C_1, \ldots, C_n$ is a n-step context sequence.

3 Preliminaries on Predicate Logic

In order to describe conditions (causes) on the presence and absence of objects that lead to a given product, we use objects of reaction systems as propositional symbols of formulas. Formally, we introduce the set F_S of propositional formulas on S defined in the standard way: $S \cup \{true, false\} \subseteq F_S$ and $\neg f_1, f_1 \vee f_2, f_1 \wedge f_2 \in F_S$ if $f_1, f_2 \in F_S$.

The propositional formulas F_S are interpreted with respect to subsets of the objects $C \subseteq S$. Intuitively, $s \in C$ denotes the presence of element s and therefore the truth of the corresponding propositional symbol. The complete definition of the satisfaction is as follows.

Definition 1. *Let $C \subseteq S$ for a set of objects S. Given a propositional formula $f \in F_S$, the satisfaction relation $C \models f$ is inductively defined as follows:*

$C \models s$ *iff* $s \in C$, $\quad C \models true$,
$C \models \neg f'$ *iff* $C \not\models f'$, $\quad C \models f_1 \vee f_2$ *iff either* $C \models f_1$ *or* $C \models f_2$,
$C \models f_1 \wedge f_2$ *iff* $C \models f_1$ *and* $C \models f_2$.

In the following $\equiv_l$ stands for the logical equivalence on propositional formulas F_S. Moreover, given a formula $f \in F_S$ we use $atom(f)$ to denote the set of propositional symbols that appear in f and $simpl(f)$ to denote the simplified version of f. The simplified version of a formula is obtained by applying the standard formula simplification procedure of

propositional logic converting a formula to Conjunctive Normal Form. We recall that for any formula $f \in F_S$ the simplified formula $simpl(f)$ is equivalent to f, it is minimal with respect to the number of propositional symbols and unique up to commutativity and associativity. Thus, we have $f \equiv_l simpl(f)$ and $atom(simpl(f)) \subseteq atom(f)$ and there exists no formula f' such that $f' \equiv_l f$ and $atom(f') \subset atom(simpl(f))$.

The causes of an object in a reaction system are defined by a propositional formula on the set of objects S. First of all we define the *applicability predicate* of a reaction a as a propositional logic formula on S describing the requirements for applicability of a, namely that all reactants have to be present and inhibitors have to be absent. This is represented by the conjunction of all atomic formulas representing reactants and the negations of all atomic formulas representing inhibitors of the considered reaction.

Definition 2. *Let $a = (R, I, P)$ be a reaction with $R, I, P \subseteq S$ for a set of objects S. The* applicability predicate *of a, denoted by $ap(a)$, is defined as follows: $ap(a) = \left(\bigwedge_{s_r \in R} s_r\right) \wedge \left(\bigwedge_{s_i \in I} \neg s_i\right)$.*

The *causal predicate* of a given object s is a propositional formula on S representing the conditions for the production of s in one step, namely that at least one reaction having s as a product has to be applicable.

Definition 3. *Let $\mathcal{A} = (S, A)$ be a r.s. and $s \in S$. The* causal predicate *of s in $\mathcal{A}$, denoted by $cause(s, \mathcal{A})$ (or $cause(s)$, when $\mathcal{A}$ is clear from the context), is defined as follows*[3]*: $cause(s, \mathcal{A}) = \bigvee_{\{(R,I,P) \in A | s \in P\}} ap\left((R, I, P)\right)$.*

We introduce a simple reaction system as running example.

Example 1. Let $\mathcal{A}_1 = (\{A, \ldots, G\}, \{a_1, a_2, a_3\})$ be a reaction system with

$$a_1 = (\{A\}, \{\}, \{B\}) \qquad a_2 = (\{C, D\}, \{\}, \{E, F\}) \qquad a_3 = (\{G\}, \{B\}, \{E\}) \ .$$

The *applicability predicates* of the reactions are $ap(a_1) = A$, $ap(a_2) = C \wedge D$ and $ap(a_3) = G \wedge \neg B$. Thus, the *causal predicates* of the objects are

$$cause(A) = cause(C) = cause(D) = cause(G) = false,$$
$$cause(B) = A,\ cause(F) = C \wedge D,\ cause(E) = (G \wedge \neg B) \vee (C \wedge D).$$

Note that $cause(A) = false$ given that A cannot be produced by any reaction. An analogous reasoning holds for objects C, D and G.

4 Formula Based Predictors and Specialized Formula Based Predictors

We introduce the notion of *formula based predictor*, originally presented in [2]. A formula based predictor for an object s at step $n+1$ is a propositional formula satisfied exactly by the context sequences leading to the production of s at step $n+1$. Minimal formula based predictors can be calculated in an effective way.

[3] We assume that $cause(s) = false$ if there is no $(R, I, P) \in A$ such that $s \in P$.

Given a set of objects S, we consider a corresponding set of *labelled objects* $S \times \mathbb{N}$. For the sake of legibility, we denote $(s, i) \in S \times \mathbb{N}$ simply as s_i and we introduce $S^n = \bigcup_{i=0}^{n} S_i$ where $S_i = \{s_i \mid s \in S\}$. Propositional formulas on labelled objects S^n describe properties of n-step context sequences. The set of propositional formulas on S^n, denoted by F_{S^n}, is defined analogously to the set F_S (presented in Sect. 3) by replacing S with S^n. Similarly, the set F_{S_i} can be defined by replacing S with S_i. Given a formula $f \in F_S$, a corresponding formula $\mathit{labelled}(f, i) \in F_{S_i}$ can be obtained by replacing each $s \in S$ in f with $s_i \in S_i$.

A labelled object s_i represents the presence (or the absence, if negated) of object s in the i-th element C_i of the n-step context sequence $\gamma = C_0, C_1, \ldots C_n$. This interpretation leads to the following definition of satisfaction relation for propositional formulas on context sequences.

Definition 4. *Let $\gamma = C_0, C_1, \ldots C_n$ be a n-step context sequence and $f \in F_{S^n}$ a propositional formula. The satisfaction relation $\gamma \models f$ is defined as*

$$\{s_i \mid s \in C_i, 0 \leq i \leq n\} \models f\,.$$

As an example, let us consider the context sequence $\gamma = C_0, C_1$ where $C_0 = \{A, C\}$ and $C_1 = \{B\}$. We have that γ satisfies the formula $A_0 \wedge B_1$ (i.e. $\gamma \models A_0 \wedge B_1$) while γ does not satisfy the formula $A_0 \wedge (\neg B_1 \vee C_1)$ (i.e. $\gamma \not\models A_0 \wedge (\neg B_1 \vee C_1)$).

The latter notion of satisfaction allows us to define formula based predictor.

Definition 5 (Formula based Predictor). *Let $\mathcal{A} = (S, A)$ be a reaction system, $s \in S$ and $f \in F_{S^n}$ a propositional formula. We say that f* f-predicts *s* in $n+1$ steps *if for any n-step context sequence $\gamma = C_0, \ldots, C_n$*

$$\gamma \models f \Leftrightarrow s \in D_{n+1}$$

where $\delta = D_0, \ldots, D_n$ is the result sequence corresponding to γ and $D_{n+1} = res_{\mathcal{A}}(C_n \cup D_n)$.

Note that if formula f f-predicts s in $n+1$ steps and if $f' \equiv_l f$ then also f' f-predicts s in $n+1$. More specifically, we are interested in the formulas that f-predict s in $n+1$ and contain the minimal numbers of propositional symbols, so that their satisfiability can easily be verified. This is formalised by the following approximation order on F_{S^n}.

Definition 6 (Approximation Order). *Given $f_1, f_2 \in F_{S^n}$ we say that $f_1 \sqsubseteq_f f_2$ if and only if $f_1 \equiv_l f_2$ and $\mathit{atom}(f_1) \subseteq \mathit{atom}(f_2)$.*

It can be shown that there exists a *unique equivalence class* of formula based predictors for s in $n+1$ steps that is minimal with respect to the order $\sqsubseteq_f$.

We now define an operator `fbp` that allows formula based predictors to be effectively computed.

Definition 7. *Let $\mathcal{A} = (S, A)$ be a r.s. and $s \in S$. We define a function $\mathtt{fbp} : S \times \mathbb{N} \to F_{S^n}$ as follows: $\mathtt{fbp}(s, n) = \mathtt{fbs}(\mathit{cause}(s), n)$, where the auxiliary function $\mathtt{fbs} : F_S \times \mathbb{N} \to F_{S^n}$ is recursively defined as follows:*

$$\begin{array}{ll}
\mathtt{fbs}(s, 0) = s_0 & \mathtt{fbs}(s, i) = s_i \vee \mathtt{fbs}(\mathit{cause}(s), i-1) \quad \text{if } i > 0 \\
\mathtt{fbs}((f'), i) = (\mathtt{fbs}(f', i)) & \mathtt{fbs}(f_1 \vee f_2, i) = \mathtt{fbs}(f_1, i) \vee \mathtt{fbs}(f_2, i) \\
\mathtt{fbs}(\neg f', i) = \neg \mathtt{fbs}(f', i) & \mathtt{fbs}(f_1 \wedge f_2, i) = \mathtt{fbs}(f_1, i) \wedge \mathtt{fbs}(f_2, i) \\
\mathtt{fbs}(\mathit{true}, i) = \mathit{true} & \mathtt{fbs}(\mathit{false}, i) = \mathit{false}
\end{array}$$

The function `fbp` gives a formula based predictor that, in general, may not be minimal with respect to the approximation order $\sqsubseteq_f$. Therefore, the calculation of a minimal formula based predictor requires the application of a standard simplification procedure, denoted $simpl()$, to the obtained logic formula.

Theorem 1. *Let $\mathcal{A} = (S, A)$ be a r.s.. For any object $s \in S$,*

- $\texttt{fbp}(s, n)$ f-predicts s *in* $n + 1$ *steps;*
- $simpl(\texttt{fbp}(s, n))$ f-predicts s *in* $n + 1$ *steps and is* minimal *w.r.t.* $\sqsubseteq_f$.

Example 2. Let us consider again the reaction system $\mathcal{A}_1$ of Ex. 1. We are interested in the production of E after 4 steps. Hence, we calculate the logic formula that *f-predicts E in* 4 *steps* applying the function `fbp`:

$$\begin{aligned}\texttt{fbp}(E,3) &= \texttt{fbs}\big((G \wedge \neg B) \vee (C \wedge D), 3\big)\\ &= \big(\texttt{fbs}(G,3) \wedge \neg\texttt{fbs}(B,3)\big) \vee \big(\texttt{fbs}(C,3) \wedge \texttt{fbs}(D,3)\big)\\ &= \big((G_3) \wedge \neg(B_3 \vee \texttt{fbs}(A,2))\big) \vee (C_3 \wedge D_3)\\ &= \big(G_3 \wedge \neg B_3 \wedge \neg A_2\big) \vee (C_3 \wedge D_3)\big)\end{aligned}$$

A context sequence satisfies $\texttt{fbp}(E, 3)$ iff the execution of the reaction system leads to the production of object E after 4 steps. Furthermore, in this case the obtained formula is also minimal w.r.t. $\sqsubseteq_f$, since $simpl(\texttt{fbp}(E, 3)) = \texttt{fbp}(E, 3)$.

There might be cases where we are interested only in sets of context sequences sharing some common properties. If we have some knowledge on the class of environments we are interested in, we can compute a specialized predictor that precisely characterize the causal dependences for the environments of interest.

In [1, 3, 4], Barbuti et al. proposed *specialized formula based predictors*. A specialized formula based predictor is a propositional logical formula that predicts the production of an object after a given number of steps, by considering only the subset of the context sequences that already satisfy the properties we know to hold for the environments of interest. The properties on the behaviour of the environment can be expressed by temporal logic formulas on context sequences. In the logic, propositional formulas describe the properties of single contexts (i.e. the symbols that can/cannot appear in an element of a context sequence). Hence, such formulas play the role of state formulas in traditional temporal logics. Temporal properties are expressed by variants of the usual *next* and *until* operators, and by derived *eventually* and *globally* operators.

In this paper we consider a standard form for the properties on the behaviour of the environment. In particular, let f_I and f_C be logic formulas each expressed as a conjunctions of literals, namely basic propositional symbols (atoms) possibly negated. We assume that the environment provides a set of initial objects satisfying formula f_I and that at all subsequent steps it provides a set of objects always satisfying the conjunctive formula f_C.

Example 3. Let $S = \{A, B, C\}$ be the set of objects of a reaction system. With $f_I = A \wedge \neg B$ we express that the first element C_0 of every possible context sequence contains for sure A, does not contain B and may contain C or not. Similarly, with $f_C = \neg A$ we express that all the other elements $C_1, C_2, \ldots$ of every possible context sequence does not contain A, but may contain B or C.

Definition 8 (Specialized Formula based Predictor). *Let $\mathcal{A} = (S, A)$ be a reaction system, $s \in S$ an object and $f \in F_{S^n}$ a propositional formula. Given the conjunctions $f_I, f_C \in F_S$, we say that f* f-predicts *s in $n+1$ steps with respect to f_I and f_C iff for any n-step context sequence $\gamma = C_0, \ldots, C_n$ such that $\gamma \models labelled(f_I, 0) \wedge \bigwedge_{1 \leq i \leq n} labelled(f_C, i)$, we have that*

$$\gamma \models f \Leftrightarrow s \in D_{n+1}$$

where $\delta = D_0, \ldots, D_n$ is the result sequence corresponding to γ and $D_{n+1} = res_{\mathcal{A}}(C_n \cup D_n)$.

It should be clear that any formula f that f-predicts s in $n+1$ steps also f-predicts s in $n+1$ steps with respect to any possible pair of formulas f_I and f_C. However, we are interested in the minimal formula that f-predicts s in $n+1$ steps with respect to any conjunctions f_I and f_C. Theorem 2 provides a method to compute the minimal specialized predictor for f_I and f_C. First let us formally define the simplification procedure of a boolean formula f_2 with respect to a conjunction of literals f_1.

Definition 9. *Let $f_1, f_2 \in F_S$ and f_1 be a conjunction of literals. With $Simpl(f_1, f_2)$ we indicate the standard procedure that simplifies a formula f_2 according to the truth values assigned to the boolean variables by the conjunction f_I.*

The following result gives a method to compute the minimal specialized predictor for f_I and f_C.

Theorem 2. *Let $\mathcal{A} = (S, A)$ be a reaction system, $s \in S$ and $f \in F_S$ be a formula based predictor of s in $n+1$ steps.*

Given the conjunctions $f_I, f_C \in F_S$,

$$Simpl((labelled(f_I, 0) \wedge \bigwedge_{1 \leq i \leq n} labelled(f_C, i)), f)$$

is the minimal formula that f-predicts s in $n+1$ steps with respect to f_I and f_C

Consequently, the following property holds, and it provides the method for the computation of minimal specialized predictors.

Corollary 1. *Let $\mathcal{A} = (S, A)$ be a reaction system, $s \in S$. Given the conjunctions $f_I, f_C \in F_S$,*

$$Simpl((labelled(f_I, 0) \wedge \bigwedge_{1 \leq i \leq n} labelled(f_C, i)), \texttt{fbp}(s, n))$$

is the minimal formula that f-predicts s in $n+1$ steps with respect to f_I and f_C

5 Information flow

Let us consider a reaction system $\mathcal{A} = (S, A)$. We assume an external observer of this system who can detect or see only some of its objects, but who wants to obtain also information on objects invisible to her/him. To formalize this situation we borrow techniques developed for reasoning about systems security. Namely, we employ information flow based security

(see [12]). It is based on an idea that systems are secure if by observing public behaviour an intruder cannot learn its private activities. Translated to the presented formalism, this means that we could express which information on invisible objects can be revealed by observing only the visible ones.

Suppose that all objects from S are divided into two groups, namely public (low level) objects L and private (high level) objects H. It is assumed that $L \cup H = S$ and $L \cap H = \emptyset$. We assume that an observer can see only L-objects, i.e. objects from L, but wants to know something about H-objects.

We introduce an equivalence on sets of objects and on contexts. Two sets of objects A, B are equivalent with respect to the set M if they contain the same objects apart from those in M. Formally, $A \equiv_M B$ iff $A \setminus M = B \setminus M$. This can be applied to reaction system contexts: we write $\gamma_1 \equiv_M \gamma_2$ and $\delta_1 \equiv_M \delta_2$, respectively. To formalize information flow between L-objects and H-objects we exploit a concept known as *opacity* (see [14] for an overview paper).

5.1 Opacity

Let us consider a reaction system $\mathcal{A} = (S, A)$ and $H \subset S$.

Definition 10. *We say that $\mathcal{A}$ is opaque with respect to H iff for every two contexts γ, γ' such that $\gamma \equiv_H \gamma'$, we have $\delta \equiv_H \delta'$.*

This says that H-objects have no influence on $S \backslash H$ objects, hence we can learn nothing about their presence by looking to the low level objects $S \backslash H$.

Since formula based predictors express all causal dependences of an object from all the other objects of the reaction system, we can use it to be sure that a reaction system $\mathcal{A}$ is opaque.

Theorem 3. *A reaction system $\mathcal{A}$ is opaque with respect to a set of high level objects H iff every* minimal *predictor of an object in $S \backslash H$ in 1 steps does not contain any H object.*

The following result is a consequence of the previous theorem and of Theorem 1.

.

Corollary 2. *$\mathcal{A}$ is opaque with respect to H iff*

$$\forall L \in S \backslash H, \{A |\ A_0 \in atom(simpl(\mathtt{fbp}(L, 0)))\} \cap H = \emptyset.$$

This gives us an easy method to verify if a reaction system is opaque with respect to a set of high level objects H. Therefore we can state the following claim.

Proposition 1. *The property of a reaction system $\mathcal{A}$ to be opaque with respect to a set of high level objects H is decidable.*

Example 4. Let $\mathcal{A}_2 = (\{A, \ldots, D\}, \{a_1, a_2\})$ be a reaction system with

$$a_1 = (\{A, B, C\}, \{\}, \{DA\}) \qquad a_2 = (\{B, C\}, \{A\}, \{D\})$$

and consider $H = \{A\}$. Note that $\mathcal{A}_2$ is opaque even if an high level object appears in reactions producing a low level object. Indeed, every context sequence that contains both B and

C will produce D in the next step regardless of the presence of the high level object A. This can be also proved by considering $simpl(\texttt{fbp}(B,0)) = simpl(\texttt{fbp}(C,0)) = simpl(false) = false$ and

$$\begin{aligned} simpl(\texttt{fbp}(D,0)) &= simpl\big(\texttt{fbs}\big(((A \wedge B \wedge C) \vee (B \wedge C \wedge \neg A)),0\big)\big) \\ &= simpl((A_0 \wedge B_0 \wedge C_0) \vee (B_0 \wedge C_0 \wedge \neg A_0)) = \\ &= B_0 \wedge C_0 \end{aligned}$$

Since $\forall L \in S\backslash H.\{A|\ A_0 \in atom(simpl(\texttt{fbp}(L,0))\} \cap H = \emptyset$, we can conclude that $\mathcal{A}_2$ is opaque with respect to H. Assume now we add to the previous reaction system the following two reactions:

$$a_3 = (\{A\},\{\},\{B\}) \qquad a_4 = (\{C\},\{A\},\{BA\})$$

Now the reaction system is no longer opaque since the high level object A is now relevant for the production of B. Imagine a context sequence that at a certain step provides object A and C and another one that at the same step provides only object C. The former will not produce B in the next step while the latter will. This can be proved by considering

$$\begin{aligned} simpl(\texttt{fbp}(B,0)) &= simpl\big(\texttt{fbs}\big(((A) \vee (C \wedge \neg A)),0\big)\big) \\ &= simpl(A_0 \vee (C_0 \wedge \neg A_0)) = \\ &= A_0 \vee (C_0 \wedge \neg A_0) \end{aligned}$$

Hence, in this case, $\exists B \in S\backslash H.\{A|\ A_0 \in atom(simpl(\texttt{fbp}(B,0))\} \cap H \neq \emptyset$, therefore, by Corollary 2, we can state that the reaction system is not opaque.

As we showed in the previous examples, opacity is a very strong property because it does not allow any flow of information from object in H to objects in $S\backslash H$. Therefore, it could be useful to consider weaker notions of opacity.

5.2 Opacity with respect to sets of context sequences

For a reaction system, being opaque with respect to H and to *every possible context sequence* is a very strong constraint. In the following we give a notion of opacity by restricting our attention to sets of context sequences satisfying some properties. These are context sequences whose initial set satisfies a logic formula f_I and whose all subsequent sets satisfy another formula f_C, with both f_I and f_C expressed as conjunctions of logical literals. Formulas f_I and f_C should not constrain all the high level objects, otherwise the concept of opacity would become meaningless. Formally, we assume that $((atom(f_I) \cup atom(f_C)) \cap H) \subset H$, that is, at least an element of H is left free in the environment.

Definition 11. *We say that $\mathcal{A}$ is opaque with respect to H and context sequences satisfying f_I and f_C iff for every two contexts sequences γ and γ' such that $\gamma \models labelled(f_I,0) \wedge \bigwedge_{1\leq i\leq n} labelled(f_C,i)$, $\gamma' \models labelled(f_I,0) \wedge \bigwedge_{1\leq i\leq n} labelled(f_C,i)$ and $\gamma \equiv_H \gamma'$, we have $\delta \equiv_H \delta'$.*

Theorem 4. *$\mathcal{A}$ is opaque with respect to H and context sequences satisfying f_I and f_C iff $\forall L \in S\backslash H$, the following hold*

1. *every* minimal *specialized predictor of an object in* $S\backslash H$ *in* 1 *step with respect to* f_I *does not contain any* H *object,*
2. *every* minimal *specialized predictor of an object in* $S\backslash H$ *in* 1 *step with respect to* f_C *does not contain any* H *object.*

The following corollary follows from the previous theorem and from Corollary 1.

Corollary 3. *$\mathcal{A}$ is opaque with respect to H and context sequences satisfying f_I and f_C iff $\forall L \in S\backslash H$, the following holds*

- $\{A|\ A_0 \in atom(Simpl(labelled(f_I, 0), \texttt{fbp}(L, 0)))\} \cap H = \emptyset$, *and*
- $\{A|\ A_0 \in atom(Simpl(labelled(f_C, 0), \texttt{fbp}(L, 0)))\} \cap H = \emptyset$.

This gives us a method to verify if a reaction system is opaque with respect to a set of high level objects H and context sequences satisfying f_I and f_C.

Proposition 2. *The property for a reaction system $\mathcal{A}$ to be opaque with respect to a set of high level objects H and context sequences satisfying f_I and f_C is decidable.*

Example 5. Let $\mathcal{A}_3 = (\{A, \ldots, D\}, \{a_1, a_2\})$ be a reaction system with

$$a_1 = (\{A, B\}, \{\}, \{C, A\}) \qquad a_2 = (\{B, D\}, \{A\}, \{C\})$$

and consider $H = \{A\}$. Let us first consider an unconstrained environment (namely $f_I = f_C = true$). In this case (that corresponds to the situation considered in Section 5.1) the reaction system turns out to be not opaque. Indeed, we have $simpl(\texttt{fbp}(C, 0)) = simpl((A_0 \wedge B_0) \vee (B_0 \wedge D_0 \wedge \neg A_0)) = B_0 \wedge (A_0 \vee D_0)$ that shows a dependence of C from A.

Let us now consider $f_I = f_C = D$, namely the constraint on behavior of the environment is that it always provides D. In this case, according to the definition, we have to simplify the formula $\texttt{fbp}(C, 0)$ under the assumption D_0, that allows us to obtain $(A_0 \wedge B_0) \vee (B_0 \wedge \neg A_0)$, that is equivalent to B_0. The predictor in this case shows no dependence of C from A and hence the system is opaque.

5.3 Opacity at a given step with respect to sets of contexts sequences

We now give a third notion of opacity, again by restricting our attention to sets of context sequences satisfying some properties. As before, these are context sequences whose initial set satisfies f_I and whose subsequent sets satisfy f_C.

The notion of opacity we are going to define is bounded in the number of steps. A reaction system is opaque in this case if the low level objects that are present in the system at step n do not depend on the high level object received in the previous steps by the context.

Definition 12. *We say that $\mathcal{A}$ is opaque with respect to H and context sequences satisfying f_I and f_C at the n-th step iff for every two contexts sequences γ and γ' such that $|\gamma| = |\gamma'| = n - 1$, $\gamma \models labelled(f_I, 0) \wedge \bigwedge_{1 \leq i \leq n} labelled(f_C, i)$, $\gamma' \models labelled(f_I, 0) \wedge \bigwedge_{1 \leq i \leq n} labelled(f_C, i)$ and $\gamma \equiv_H \gamma'$, we have $D_n \equiv_H D'_n$.*

This notion of opacity ensures that the low level objects in D_n are independent from the high level object received in the previous steps. Nothing is said about the low level objects in D_i with either $i < n$ or $i > n$. A notion of bounded opacity stating that if $\gamma \equiv_H \gamma'$, it holds $\delta \equiv_H \delta'$ with $|\gamma| = |\gamma'| = n - 1$ and $|\delta| = |\delta'| = n$ could be trivially defined as an extension of the notion just introduced.

As in the previous cases, this third notion of opacity can be formulated in terms of formula based predictors.

Theorem 5. *$\mathcal{A}$ is opaque with respect to H and context sequences satisfying f_I and f_C at the n-th step iff $\forall L \in S\backslash H$ every* minimal *specialized predictor of an object in $S\backslash H$ in n step with respect to f_I and f_C does not contain any H object.*

Using Corollary 1, we have an effective way to verify this opacity property.

Corollary 4. *$\mathcal{A}$ is opaque with respect to H and context sequences satisfying f_I and f_C at the n-th step iff $\forall L \in S\backslash H$,*

$$\{A|\ A_i \in atom(Simpl(labelled(f_I, 0) \wedge \bigwedge_{1 \leq i \leq n} labelled(f_C, i), \texttt{fbp}(L, n)))\} \cap H = \emptyset.$$

Example 6. Let $\mathcal{A}_4 = (\{A, B\}, \{a_1, a_2\})$ be a reaction system with

$$a_1 = (\{A\}, \{B\}, \{B\}) \qquad a_2 = (\{B\}, \{\}, \{A\})$$

and consider $H = \{A\}$.

Moreover, let us consider $f_I = \neg B$ and $f_C = \neg A \wedge \neg B$, namely the environment is allowed only to (possibly) provide an object A at the very beginning of the reaction system execution.

It is rather obvious that the system is not opaque according to the notions of opacity given in the previous sections, since the presence of B in the system reveals that A was provided by the environment at the beginning.

It is also easy to see that the dynamics of the system causes the B object, if produced, to continuously appear and disappear from the system. As a consequence, an observer able to look at the system state only at even steps (e.g at step 2 or at step 4) would not see the B object independently of whether A was provided at the beginning or not. Another observer looking at the system state at odd steps (e.g. at step 1 or at step 3) would instead be able to determine whether A was provided or not.

As an example, we can assess opacity at step 2 by computing predictor $\texttt{fbp}(B, 1)$. (Recall that $\texttt{fbp}(s, n)$ gives the predictor for s at step $n + 1$.)

$$\texttt{fbp}(B, 1) = (A_1 \vee B_0) \wedge \neg(B_1 \vee (A_0 \wedge \neg B_0))$$

Formulas f_I and f_C allow us to simplify the predictor into $false$, that does not contain A among its atoms.

We can also assess that the system is not opaque at step 3 by computing predictor $\texttt{fbp}(B, 2)$.

$$\texttt{fbp}(B, 2) = (A_1 \vee B_2 \vee (A_0 \wedge \neg B_0)) \wedge \neg(B_1 \vee ((A_1 \vee B_0) \wedge \neg(B_1 \vee (A_0 \wedge \neg B_0))))$$

This time, formulas f_I and f_C allow us to simplify the predictor into A that shows that the system is not opaque.

6 Conclusions

We considered three notions of opacity for reaction systems that allow information flow properties to be assessed. Moreover, we showed that opacity properties can be computed by resorting to the algorithms for the computation of formula based predictors. As future work, we plan to complete the study of opacity notions by considering some weaker variants. The basic definition requires that H-objects have no influence on L-objects at all, but this could be too strong. It might be sufficient to require that there always exists a context with different H-objects, which would lead to the same L-objects (i.e., universal quantification is replaced by the existential one). Moreover, here we studied the so called state based opacity, where states (objects) are of interest. In the so called language based opacity, rules could be divided into public and private ones, and an observer of states could try to deduce whether also private rules have been performed in a given context.

References

1. Barbuti, R., Gori, R., Levi, F., Milazzo, P.: Specialized predictor for reaction systems with context properties. In: Proc. of the 24th Int. Workshop on Concurrency, Specification and Programming, CS&P 2015. pp. 31–43 (2015)
2. Barbuti, R., Gori, R., Levi, F., Milazzo, P.: Investigating dynamic causalities in reaction systems. Theoretical Computer Science 623, 114–145 (2016)
3. Barbuti, R., Gori, R., Levi, F., Milazzo, P.: Specialized predictor for reaction systems with context properties. Fundamenta Informaticae 147(2-3), 173–191 (2016)
4. Barbuti, R., Gori, R., Levi, F., Milazzo, P.: Generalized contexts for reaction systems: definition and study of dynamic causalities. Acta Informatica. (2017), doi:10.1007/s00236-017-0296-3
5. Bodei, C., Gori, R., Levi, F.: An analysis for causal properties of membrane interactions. Electr. Notes Theor. Comput. Sci. 299, 15–31 (2013)
6. Bodei, C., Gori, R., Levi, F.: Causal static analysis for brane calculi. Theor. Comput. Sci. 587, 73–103 (2015)
7. Brijder, R., Ehrenfeucht, A., Main, M.G., Rozenberg, G.: A Tour of reaction Systems. Int. J. Found. Comput. Sci. 22(7), 1499–1517 (2011)
8. Brijder, R., Ehrenfeucht, A., Rozenberg, G.: A Note on Causalities in Reaction Systems. ECEASST 30 (2010)
9. Bryans, J., Koutny, M., Ryan, P.: Modelling non-deducibility using petri nets. In: Proc. of the 2nd International Workshop on Security Issues with Petri Nets and other Computational Models (2004)
10. Bryans, J.W., Koutny, M., Mazaré, L., Ryan, P.Y.: Opacity generalised to transition systems. International Journal of Information Security 7(6), 421–435 (2008)
11. Ehrenfeucht, A., Rozenberg, G.: Reaction Systems. Fundam. Inform. 75(1-4), 263–280 (2007)
12. Goguen, J.A., Meseguer, J.: Security policies and security models. Proc. of IEEE Symposium on Security and Privacy (1982)
13. Gori, R., Levi, F.: Abstract interpretation based verification of temporal properties for bioambients. Inf. Comput. 208(8), 869–921 (2010)
14. Jacob, J., Lesage, J.J., Faure, J.M.: Overview of discrete event systems opacity: Models, validation, and quantification. Annual Reviews in Control 41 (2016)