Theoretical Computer Science 623 (2016) 114–145

Contents lists available at ScienceDirect

Theoretical Computer Science

www.elsevier.com/locate/tcs

Investigating dynamic causalities in reaction systems

Roberto Barbuti, Roberta Gori *, Francesca Levi, Paolo Milazzo

Dipartimento di Informatica, Università di Pisa, Largo Pontecorvo 3, 56127 Pisa, Italy

ARTICLE INFO

Article history:
Received 7 April 2015
Received in revised form 15 October 2015
Accepted 21 November 2015
Available online 2 December 2015
Communicated by T. Yokomori

Keywords:
Reaction systems
Dynamic causalities
Abstract interpretation

ABSTRACT

Reaction systems are a qualitative formalism for modeling systems of biochemical reactions characterized by the *non-permanency* of the elements: molecules disappear if not produced by any enabled reaction. Moreover, reaction systems execute in an environment that provides new molecules at each step. Brijder, Ehrenfeucht and Rozenberg investigated *dynamic causalities* in reaction systems by introducing the idea of *predictors*. A predictor of a molecule s, for a given n, is the set of molecules to be observed in the environment in order to determine whether s is produced or not by the system at step n.

In this paper, we continue the investigation on dynamic causalities by defining an abstract interpretation framework containing three different notions of predictor: *Formula based predictors*, that is a propositional logic formula that precisely characterizes environments that lead to the production of s after n steps; *Multi-step based predictors*, that consist of n sets of molecules to be observed in the environment, one for each step; and *Set based predictors*, that are those proposed by Brijder, Ehrenfeucht and Rozenberg, and consist of a unique set of molecules to be observed in all steps.

For each kind of predictor we define an effective operator that allows predictors to be computed for any molecule s and number of steps n. The abstract interpretation framework allows us to compare the three notions of predictor in terms of precision, to relate the three defined operators and to compute minimal predictors. We also discuss a generalization of this approach that allows predictors to be defined independently of the value of n, and a tabling approach for the practical use of predictors on reaction systems models. As an application, we use predictors, generalization and tabling to give theoretical grounds to previously obtained results on a model of gene regulation.

1. Introduction

In the last decades, the mechanisms underlying the functioning of living cells have been the source of inspiration of many formalisms and notations in the field of Natural Computing [1,2]. Many of these formalisms are based on rewriting approaches. This is due to the similarity between rewrite rules and chemical reactions (that govern the functioning of living cells) and to the fact that rewriting approaches can be studied from the viewpoint of computing power with techniques that are typical of the theory of formal languages.

Among the proposed formalisms based on rewriting we mention Refs. [3–5]. Each of them deals with a different peculiar aspect of the function of living cells, that is exploited as the main computing feature. Membrane systems [3,4], for instance, exploit *maximal parallelism* in the application of rewrite rules and the hierarchical membrane structure of living cells. On the

* Corresponding author.
E-mail addresses: barbuti@di.unipi.it (R. Barbuti), gori@di.unipi.it (R. Gori), levifran@di.unipi.it (F. Levi), milazzo@di.unipi.it (P. Milazzo).

http://dx.doi.org/10.1016/j.tcs.2015.11.041
0304-3975/

other hand, Spiking Neural P Systems [5] exploit neuron signaling mechanisms. In addition to these, a number of variants of the membrane systems formalisms have been investigated [6–12].

Reaction systems [13,14] have been introduced by Ehrenfeucht and Rozenberg as a novel model for the description of biochemical processes driven by the interaction among reactions in living cells. Reaction systems 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 chemical 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 called the *threshold supply* assumption and characterizes reaction systems as a *qualitative* modeling formalism.

A reaction system evolves by means of the application of its set of 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. In particular, the next state consists only of the products of the reactions applied in the current step. This is one of the most original bio-inspired features of reaction systems that distinguishes it from the other formalisms mentioned above.

The overall behavior of a reaction system model is driven by the (set of) contextual elements which are received from the external environment at each step. Such elements join the current state of the system and, as the other objects in the system state, can enable or disable reactions. The computation of the next state of a reaction system is a deterministic procedure. However, since the contextual elements that can be received at each step can be any subset of the support set S, the overall system dynamics is non-deterministic.

Reaction systems have been used to model various features which are useful both for the modeling of computational devices and for the modeling of biological systems. For example, binary counters [13] form the basis for the inclusion of a notion of time [15]. In [16] an extension with duration of reaction systems is presented. Theoretical aspects of reaction systems have been studied in [17–23]. From the biological viewpoint, a model of the *lac* operon has been presented in [24].

In [18] Brijder, Ehrenfeucht and Rozenberg initiate an investigation of *causalities* in reaction systems, i.e. the ways that entities of a reaction system influence each other. Both static/structural causalities as well as dynamic causalities are discussed, 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. The concept of predictor is based on the idea that, in general, not all contextual elements are relevant for determining if s will be produced after n steps. Indeed, for given s and n, there might be a subset Q of S which is the part of S that it is essential to observe in contextual elements for predicting whether s will be produced after n steps or not. Such set Q can then be used to concentrate uniquely on the relevant part of the contextual elements received from the external environment at each step, ignoring all elements that are surely not involved in the production of s in n steps. If two different sequence of contextual sets become equal after dropping elements not in Q, we can be sure that they both determine either the presence or the absence of s after n steps. In other words, Q is a subset of S which is a cause for s to be *uniformly* either present or absent after n steps. Brijder et al. define such Q the predictor of s in n steps, since knowing the behavior of the system with a sequence of contextual element, allows us to predict the behavior of such system with any other sequence having the same sets of relevant contextual elements.

Following these ideas, predictors can be profitably used to decide whether s will appear or not after n steps, without executing the reaction system, by following an approach based on *tabling*. Given a reaction system and a predictor for it, a table can be constructed that contains one line for every possible sequence of contextual sets consisting only of elements of the predictor. Each line should indicate whether for such a sequence the symbol for which the predictor has been defined is actually produced after n steps or not. Then, when an observed system has to be evaluated, this should be done by considering the sequence of objects it receives from the context restricted to the elements of the predictor. The obtained sequence can then be used to access the previously constructed table and predict whether the object of interest will be produced or not.

We believe that the concept of predictor is very interesting and therefore deserves further and deeper investigation. In this paper we push forward the previous idea of a predictor by defining three different notions of predictor for Reaction Systems. The first notion is that of *formula based predictors*, namely predictors that consist in a propositional logic formula to be satisfied by contextual elements of a given reaction system. This is the most precise predictor that can be considered: satisfaction (or not) of the logic formula will discriminate the cases in which s will be produced after n steps from those in which it will not be produced.

The second notion of predictor we define is that of *multi-step based predictors*, namely predictors that consists of sets of objects to be observed in contextual elements at each step of execution of the reaction system. Predictors of this kind