Sufficient conditions for reachability in automata networks with priorities

Maxime Folschette [a,b,*], Loïc Paulevé [c], Morgan Magnin [a], Olivier Roux [a]

[a] *LUNAM Université, École Centrale de Nantes, IRCCyN UMR CNRS 6597, Institut de Recherche en Communications et Cybernétique de Nantes, 1 rue de la Noë, B.P. 92101-44321 Nantes Cedex 3, France*
[b] *School of Electrical Engineering and Computer Science, University of Kassel, Germany*
[c] *CNRS, Laboratoire de Recherche en Informatique (LRI), Université Paris-Sud, CNRS UMR 8623, France*

ARTICLE INFO

ABSTRACT

In this paper, we develop a framework for an efficient under-approximation of the dynamics of Asynchronous Automata Networks (AANs). An AAN is an Automata Network with synchronised transitions between automata, where each transition changes the local state of exactly one automaton (but any number of synchronising local states are allowed). The work we propose here is based on static analysis by abstract interpretation, which allows to prove that reaching a state with a given property is possible, without the same computational cost of usual model checkers: the complexity is polynomial with the total number of local states and exponential with the number of local states within a single automaton. Furthermore, we address AANs with classes of priorities, and give an encoding into AANs without priorities, thus extending the application range of our under-approximation. Finally, we illustrate our method for the model checking of large-scale biological networks.

1. Introduction

Discrete modelling frameworks for biological networks is an active research field where formal methods have be proven very powerful [1–3]. Such a work started in the seventies, with the emergence of the notion of Boolean Network [4] and its use to represent biological phenomena [5]. It was later enriched in many directions and widely used to elucidate many biological questions. Among these questions, a major one is to understand precisely how biological systems evolve and behave; why and how they change their usual behaviour, etc. These questions are strongly linked to the (possible or inevitable) reachability of some states. The ultimate goal is to discover how it could be possible to prevent biological systems from reaching some pathological conditions.

Of course, such formal models on which analyses are performed are abstract representations of the actual studied systems. They are associated with parameters that have to be synthesised to give the most faithful representation of the real systems with their observed behaviours. As a matter of fact, the abstractions we get are more or less rough or accurate. Prevalent formal frameworks for such modelling activities are state-transition systems or process algebras. We developed a quite similar framework named the Process Hitting [6], consisting in a restriction of these frameworks where the evolution of a component is determined by the state of at most one other component that does not evolve. This is modelled by

actions of the form $X + Y \rightarrow X + Y'$, where X behaves like a catalyst molecule that "hits" another molecule Y and changes it into Y', without being changed itself. Assuming catalysts are always available, this can represent any biochemical system made of monomolecular reactions, and can also represent catalytic networks such as metabolic networks. Our motivation behind this framework was to design a model and analysis techniques adapted to biological modelling. These analyses avoid building the whole state space, which allows to tackle very large systems (that would have led to a huge number of states, hopelessly too huge to be analysed). They are based on the fact that most biological models have few levels of expression per component: in Boolean networks [4,5] there are only two levels per component, and in their multivalued equivalent [2], components rarely have more than four levels.

Besides, one further objective of our work is now to improve the accuracy of the description of the studied systems dynamics. The idea for this is to introduce timing features into models: we are interested in taking into account some knowledge about the relative length of some phenomena as it is a way to refute some models (or parameters) that are inconsistent with the observed dynamic behaviours. In this paper, we are dealing with these timing properties through priorities, that are based on the simple founding idea that actions with higher priority have to be processed before the ones with lower priority. Furthermore, due to the Process Hitting framework restrictions, multimolecular reactions were previously not immediately available, but one could simulate them with an encoding called "cooperative sort". That encoding however introduces extra reactions, that produce a temporal shift between the presence of the reactants and the playability of the reaction. This is where the priorities become useful, if not necessary: the extra reactions can for example be given "infinite speed" (highest priority) so that they do not affect the behaviour of "normal" (lower priority) reactions, including the multimolecular ones.

The approach used in this paper consists in considering a broader class of models, that we call Asynchronous Automata Network, and that allows to naturally model these cooperations by defining several requisites for a transition. Moreover, such automata networks are still compatible with the notion of priority, that can also be used to model different reaction rates in the model. Asynchronous Automata Networks (and, a fortiori, their restriction, the Process Hitting framework) can be considered as a subset of Communicating Finite State Machines or safe Petri nets [7]. Our work is thus related to such semantic ramifications of extending traditional process algebras with the concept of priority that allows for some transitions to be given precedence over others. We focus here on static priorities that allow to model time constraints such as reaction rates or delays between regulations, but can also model preemptions between evolution branches.

Until now, such a priority scheduling of the actions was not studied extensively in the different formal modelling frameworks dedicated to systems biology. Nevertheless, such an attempt has been carried out for Petri nets by F. Bause [8], and the concept of priority relations among the transitions of a network has also more recently been introduced by A.K. Wagler et al. [9,10] in order to allow modelling deterministic systems for biological applications. The concept of priority is rather straightforward in the approach of process algebras as it was shown by R. Cleaveland and M. Hennessy in [11] and their abstractions and equivalences were studied in [12]. It was later extended for applications in the field of systems biology by M. John et al. [13].

1.1. Contributions

In this paper, we develop a method that allows to efficiently compute under-approximations of the dynamics of Asynchronous Automata Networks (AANs), generalising a prior work carried out on Process Hitting [14]. Rather than using brute force or symbolic model checking techniques, our method focuses on static analysis by abstract interpretation. Our work aims at checking reachability properties that, for a given initial state s and the discrete expression level n of a given component x, have the form: "Starting from state s, is it possible to reach a state in which x is at level n?", thus answering either "True" (in which case the reachability is formally proven) or "Inconclusive" (which however does not stand for "False", that can be proven by other analysis tools, such as the over-approximation proposed in [14]). Moreover, the successive or simultaneous reachability of several local states can also be checked, and in the case of a "True" response, an execution path satisfying the property can be produced. Extensions of AANs where transitions are split in priority classes are addressed with an encoding in AANs without classes of priorities. Our work thus allows to efficiently analyse the dynamics of regulation networks, especially the widespread Logical Networks [15,2], which encompass variables with a limited number of discrete values alongside with evolution functions or focal parameters. To show the scalability of our method, we apply it to two large-scale biological models with around 100 components.

Section 2 presents the Asynchronous Automata Networks (AANs) without any notion of classes of priorities. In Section 3, we develop our under-approximation method allowing to efficiently compute reachability analyses; we also show how to extract a valid execution path if the response is positive, and propose two refinements in case one needs to check successive or simultaneous reachability properties. The framework of AANs with classes of priorities is given in Section 4, alongside with an encoding into AANs without classes of priorities (or, equivalently, with one class of priority). Finally, Section 5 provides a detailed example and two large-scale examples of the application of our method, and Section 6 gives a conclusion and a discussion about our work.

The main additions in this paper compared to [16] are Section 3.3 that allows the extraction of a concrete trace of execution when our static analysis method is conclusive, Section 3.4 that refines our approach in the case of a successive reachability in order to increase the conclusiveness, and Section 4 which states that any ANN with any number of classes of priorities can be represented into a simple ANN (or, equivalently, with only one class of priority), thus extending the scope