Reliability Engineering and System Safety 139 (2015) 179–187

Contents lists available at ScienceDirect

Reliability Engineering and System Safety

journal homepage: www.elsevier.com/locate/ress

Automated generation of partial Markov chain from high level descriptions

P.-A. Brameret [a,*], A. Rauzy [b], J.-M. Roussel [a]

[a] Lurpa, Ens Cachan, Univ Paris-Sud, F-94235 Cachan, France
[b] Chaire Blériot-Fabre, LGI École Centrale de Paris, Grande voie des vignes, 92295 Châtenay-Malabry cedex, France

ARTICLE INFO

Article history:
Received 19 August 2014
Received in revised form
12 February 2015
Accepted 22 February 2015
Available online 3 March 2015

Keywords:
Model Based Safety Assessment
Markov chains
State space build
AltaRica

ABSTRACT

We propose an algorithm to generate partial Markov chains from high level implicit descriptions, namely AltaRica models. This algorithm relies on two components. First, a variation on Dijkstra's algorithm to compute shortest paths in a graph. Second, the definition of a notion of distance to select which states must be kept and which can be safely discarded.

The proposed method solves two problems at once. First, it avoids a manual construction of Markov chains, which is both tedious and error prone. Second, up the price of acceptable approximations, it makes it possible to push back dramatically the exponential blow-up of the size of the resulting chains.

We report experimental results that show the efficiency of the proposed approach.

1. Introduction

Markov chains are pervasive in Probabilistic Safety Analyses. They make it possible to assess performance indicators for systems with complex control structures such as cold spare units, or systems with limited number of resources. However, they suffer from the exponential blow-up of the number of states and transitions. This drawback has two aspects. First, the manual construction of Markov chains is both tedious and error prone. Second, assessment of large Markov chains is very resource consuming.

A way to solve the first problem consists in generating Markov chains from higher level descriptions, typically Generalized Stochastic Petri Nets [1] or AltaRica models [2]. These descriptions represent the state space in an implicit way. To obtain the Markov chain, a space exploration algorithm is used: starting from the initial state, states and transitions are progressively added to the resulting chain, until no more state or transition can be added.

However, only some of the many states of a very large Markov chain are relevant to the calculation of reliability indicators. The odds of reaching them is very low. Therefore, they have almost no influence on the calculated quantities and can be safely ignored. The same idea is behind algorithms that consider failure sequences in turn, while keeping only probable enough sequences; see e.g. [3–5]. What we propose here is rather to generate a relevant fraction of the whole Markov chain. Technically, the idea is to explore the underlying state graph at a bounded depth, i.e. to keep states (and transitions between these states) that are at the shortest distance from the initial state. Our algorithm relies on two components:

- An efficient way to explore the underlying graph in order to avoid revisiting states. To do so, we apply a variation of Dijkstra's algorithm to determine on-the-fly shortest paths in a graph [6].
- A suitable notion of distance which is basically the probability of the path and that is used as an indicator of relevance for states.

The combination of these two components proves extremely efficient. We present here examples for which a partial chain, whose size is a tiny fraction of the complete chain, makes it possible to approximate system unreliability with a relative error less than 0.25%.

It is not possible to guarantee a priori the quality of the approximation (to get a "probably approximately correct" result according to Valiant's scheme for approximation algorithms [7]). However, we show that it is possible to calculate a posteriori an upper bound of the probability of discarded states. This bound provides the analyst with a means to assess the accuracy of the approximation.

The method we propose in this paper is a contribution to the so-called Model-Based Safety Analyses: it makes Markov chains an effective tool to assess large high level models. This tool is of paramount interest for systems that show dependencies amongst

* Corresponding author.
E-mail addresses: pierre-antoine.brameret@lurpa.ens-cachan.fr (P.-A. Brameret), Antoine.Rauzy@ecp.fr (A. Rauzy), jean-marc.roussel@lurpa.ens-cachan.fr (J.-M. Roussel).

http://dx.doi.org/10.1016/j.ress.2015.02.009
0951-8320/

failures, i.e. systems for which combinatorial representations (such as Fault Trees) are not suitable.

The remainder of this paper is organized as follows. Section 2 introduces the context of the present work, and discusses related works. Section 3 presents the algorithm. Section 4 discusses issues regarding the practical implementation of the algorithm and the accuracy of the approximation. Finally, Section 5 presents experimental results.

2. Problem statement

2.1. Context

Classical formalisms used in safety analyses, such as Fault Trees and Markov chains, are well mastered by analysts. Moreover, they provide a good tradeoff between the expressiveness of the modeling formalism and the efficiency of assessment algorithms. They stand however at a low level. As a consequence, there is a significant distance between the specifications of the system under study and the safety models of this system. This distance is both error prone and a source of inefficiency in the modeling process. Not only are models difficult to share amongst stakeholders but any change in the specifications may require a tedious review of safety models.

Hence the idea is to describe systems with high level modeling formalisms and to compile these high level descriptions into lower level ones, typically Fault Trees and Markov chains, for which efficient assessment algorithms exist. AltaRica 3.0 is such a high level formalism (see e.g. [8]).

The semantics of AltaRica 3.0 is defined in terms of Guarded Transition Systems [9]. Prior to most of any assessment, including compilation into Markov chains, AltaRica 3.0 models are flattened into Guarded Transition Systems as illustrated Fig. 1 which gives an overview of the AltaRica 3.0 project.

As defined in [8], a Guarded Transition System (GTS for short) is a quintuple $\langle V, E, T, A, \iota \rangle$, where

- $V = S \uplus F$ is a set of variables, divided into two disjoint subsets: the subset S of state variables and the subset F of flow variables.
- E is a set of events.
- T is a set of transitions. A transition is a triple $\langle e, G, P \rangle$, denoted as $e : G \to P$, where $e \in E$ is an event, G is a guard, i.e. a Boolean formula built over V, and P is an instruction built over V, called the action of the transition. The action modifies only state variables.
- A is an assertion, i.e. an instruction built over V. The assertion modifies only flow variables.
- ι is the initial assignment of variables of V.

In a GTS, states of the system are represented by variable assignments. A transition $e : G \to P$ is said to be fireable in a given state σ if its guard G is satisfied in this state, i.e. if $G(\sigma) = \texttt{true}$. The firing of that transition transforms the state σ into the state $\sigma' = A(P(\sigma))$, i.e. σ' is obtained from σ by applying successively the action of the transition and the assertion.

Guarded Transition Systems are implicit representations of labeled Kripke structures, i.e. of graphs whose nodes are labeled by variable assignments and whose edges are labeled by events. The so-called reachability graph $\Gamma = \langle \Sigma, \Theta \rangle$ of a GTS $\langle V, E, T, A, \iota \rangle$ is the smallest Kripke structure such that

- $\iota \in \Sigma$.
- If $\sigma \in \Sigma$, $e : G \to P$ is a transition of T and $G(\sigma) = \texttt{true}$ (the transition is fireable in σ), then $\sigma' = A(P(\sigma)) \in \Sigma$ and $e : \sigma \to \sigma' \in \Theta$.

If exponential distributions are associated with events of E, the Kripke structure $\Gamma = \langle \Sigma, \Theta \rangle$ can be interpreted as a Continuous Time Homogeneous Markov Chain (for sake of brevity we shall just write Markov Chain in the remainder of the paper). The reliability indicators (such as system unavailability) can be defined by associating a reward (a real number) with each state of the chain.

The reachability graph may be very large, even for small GTS. Assume for instance that we model a system made of n independent, repairable components. Then, the number of variables of V is n, the number of transitions of T is $2 \times n$, but the number of states of Σ is 2^n and the number of transitions of Θ is $n \times 2^n$. Even when the components of the system are not fully independent, safety models tend to show the same picture, i.e. a number of states which is exponential in the number of components (or the variables in the GTS) and a number of transitions which is a small multiple of the number of states.

The idea is thus to generate (still starting from the initial state and applying the above principle) only a fraction of the Kripke structure,

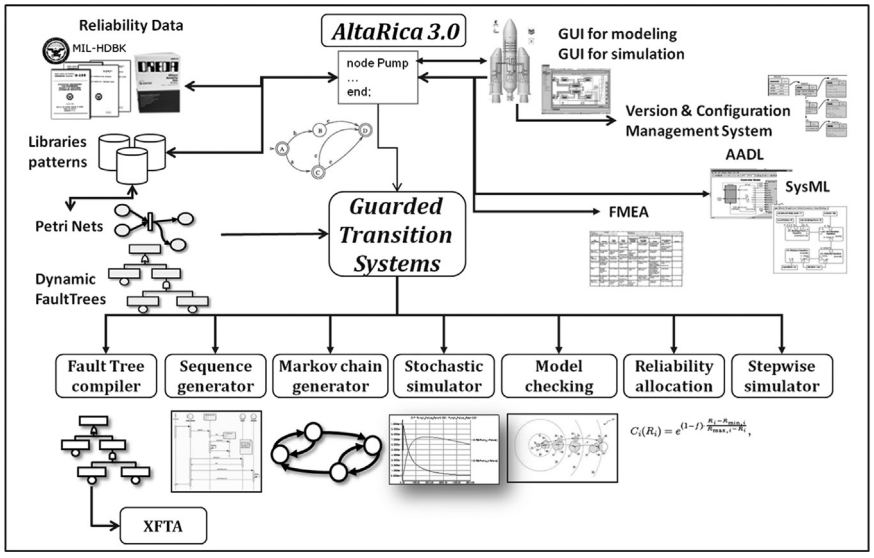

Fig. 1. Overview of the AltaRica 3.0 project.