Performance Evaluation 73 (2014) 73–90

Contents lists available at ScienceDirect

Performance Evaluation

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

Approximate aggregation of Markovian models using alternating least squares

Peter Buchholz *, Jan Kriege

Department of Computer Science, TU Dortmund, Dortmund, Germany

ARTICLE INFO

Article history:
Available online 18 September 2013

Keywords:
Compositional modeling
Aggregation
Markov models
Non-negative least squares
Matrix factorization

ABSTRACT

State based analysis of Markovian models is faced with the problem of state space explosion. To handle huge state spaces often compositional modeling and aggregation of components are used. Exact aggregation resulting in exact transient or stationary results is only possible in some cases, when the Markov process is lumpable. Therefore approximate aggregation is often applied to reduce the state space. Several approximate aggregation methods exist which are usually based on heuristics.

This paper presents a new aggregation approach for Markovian components which computes aggregates that minimize the difference according to some algebraically defined function which describes the difference between the component and the aggregate. If the difference becomes zero, aggregation is exact, which means that component and aggregate are indistinguishable in the sense that transient and stationary results in any environment are identical. For the computation of aggregates, an alternating least squares approach is presented which tries to minimize the norm-wise difference between the original component and the aggregate. Algorithms to compute aggregates are also introduced and the quality of the approximation is evaluated by means of several examples.

1. Introduction

A major problem when analyzing Markovian models is the state space explosion which makes models intractable because of their size. Models resulting from real systems are often too large to be analyzed numerically. To deal with this problem, compositional modeling and the reduction of the state space resulting in smaller components that are combined into one complex model are applied. Compositional Markovian models appeared in the literature of the recent decades in various forms, for example as stochastic automata networks (SANs) [1], compositional variants of stochastic Petri Nets (SPNs) [2], stochastic process algebras [3], hierarchical Markovian models [4] or interactive Markov chains [5]. In contrast to many other variants, we consider here models with active output and passive input signals similar to probabilistic or stochastic variants of I/O automata [6].

In the past, several approaches have been proposed [2,5,3,1,7] that reduce the state space of components by finding a smaller component with equivalent behavior and substitute the original component by the smaller one in the compositional model. These approaches have in common that stochastic bisimulation [8,9], which is based on lumpability [10,11], is used to reduce the size of the state space. Since the overall state space grows combinatorially in the sizes of the state spaces of the components, substituting a component with m states by a component with $n < m$ states reduces the overall number of states by a factor n/m.

* Corresponding author. Tel.: +49 231 755 4746.
E-mail addresses: peter.buchholz@udo.edu (P. Buchholz), jan.kriege@udo.edu (J. Kriege).

0166-5316/$ – see front matter
http://dx.doi.org/10.1016/j.peva.2013.09.001

Approaches using bisimulation for state space reduction exploit the fact that bisimulation is a congruence according to model composition. In [12] a general equivalence definition for components of the same size has been proposed that goes beyond bisimulation and that has been extended to a definition of equivalence of components of different sizes in [13,14]. This equivalence definition, which includes bisimulation and lumpability as special cases, relates Markovian and non-Markovian representations (called Rational Arrival Processes [15]) and has been used in [16] to compute minimal representations for several components. However, the approach from [16] usually results in non-Markovian representations, which lack the intuitive representation as a Markov chain, and experiments show that the resulting models are often only slightly smaller than models resulting from lumpability. Therefore, we present an approach that uses approximate instead of exact state space reduction by solving non-negative least squares problems and restrict the model class to Markovian models by applying uniformization to the Markovian components.

This paper is an extension of [17]. In addition to [17], it presents a second equivalence relation as a base for approximate aggregation, it introduces extended algorithms to compute approximate aggregates, and presents the aggregation of several example models. The approach is a first step to exploit aggregation methods from linear systems theory [18] for Markov models.

In the following section we briefly review related work. Section 3 introduces the basic model class and exact aggregation of components. Approximate aggregation of Markovian components using a least squares optimization approach is presented in Section 4. Afterwards, in Section 5, algorithms for computing approximate aggregates are shown. In Section 6, we present some examples to show the possibilities of approximate state space reduction for different applications. The paper ends with the conclusions in Section 7.

2. Related work

Aggregation of Markov models has been a research subject for several decades. Apart from exact aggregation or state space reduction, methods based on lumpability, that have been mentioned above, approximate approaches, which combine aggregation and composition of components, are important. We briefly mention here some compositional methods for approximate aggregation without claiming to be exhaustive.

The first class of methods is based on the famous work of Courtois [19] on nearly completely decomposable (NCD) systems. If a component has a loose connection to its environment, which implies that the system observes the NCD property, then the component can be analyzed in isolation and substituted by a much simpler aggregate. The error is in this case proportional to the degree of coupling, which corresponds to the relation between rates of transitions inside the component and rates between the component and its environment. The NCD property has been exploited in various analysis approaches for Markov models [20].

Exact aggregation techniques based on lumpability have been extended to approximate aggregation by relaxing the required conditions which results in *nearly* or *quasi* lumpability [10,21]. Nearly/quasi lumpability has also been used in compositional aggregation [22,23].

A third class of aggregation techniques uses stochastic ordering. Related approaches compute stochastically larger models that can be exactly aggregated for example due to lumpability [24]. The resulting aggregated model is stochastically larger than the original model and allows one to compute bounds on certain result measures. Under specific conditions, aggregation based on stochastic ordering can also be applied in a compositional setting [25].

All mentioned approaches are based on very specific conditions that have to be observed in the state space. In this paper, we develop an approach that is more flexible by using a very general condition to define the behavioral difference of the original system and an aggregate. By minimizing this difference we can compute good aggregates.

3. Compositional Markovian models

We first introduce Markovian components, describe afterwards their composition and introduce two equivalence relations for Markovian components. The section captures results which have been published in [16] for a more general class of models, denoted as rational (arrival) processes. However, in contrast to [16] we restrict models to Markovian models and define the equivalence relation accordingly. To remain in the class of Markovian models, we apply uniformization to Markovian components which has not been done yet and allows us to do all computations with non-negative numbers.

3.1. Markov components

We begin with the definition of Markovian components with input and output transitions denoted as events.

Definition 1. $\mathcal{A} = (\pi_0, \mathbf{G}_0, \mathbf{G}_1, \ldots, \mathbf{G}_K, \mathbf{U}_{K+1}, \ldots, \mathbf{U}_{K+L})$ is a Markovian component of order n (i.e., with state space $\mathcal{S} = \{0, \ldots, n-1\}$) with input and output signals, where

- π_0 is an n-dimensional distribution (row) vector,
- $\mathbf{G}_0$ is an $n \times n$ matrix with non-negative elements outside the diagonal and row sum ≤ 0,