International Journal of Approximate Reasoning 53 (2012) 1397–1414

Contents lists available at SciVerse ScienceDirect

International Journal of Approximate Reasoning

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

Efficient sensitivity analysis in hidden Markov models

Silja Renooij *

Department of Information and Computing Sciences, Utrecht University, P.O. Box 80.089, 3508 TB Utrecht, The Netherlands

ARTICLE INFO

Article history:
Available online 21 June 2012

Keywords:
Sensitivity analysis
Bayesian networks
Hidden Markov models
Sensitivity function

ABSTRACT

Sensitivity analysis in hidden Markov models (HMMs) is usually performed by means of a perturbation analysis where a small change is applied to the model parameters, upon which the output of interest is re-computed. Recently it was shown that a simple mathematical function describes the relation between HMM parameters and an output probability of interest; this result was established by representing the HMM as a (dynamic) Bayesian network. To determine this sensitivity function, it was suggested to employ existing Bayesian network algorithms. Up till now, however, no special purpose algorithms for establishing sensitivity functions for HMMs existed. In this paper we discuss the drawbacks of computing HMM sensitivity functions, building only upon existing algorithms. We then present a new and efficient algorithm, which is specially tailored for determining sensitivity functions in HMMs.

1. Introduction

Hidden Markov models (HMMs) are frequently applied statistical models for describing processes that evolve over time. Applications of hidden Markov models are found in areas such as speech recognition, machine translation and bioinformatics (see [1] for an overview). A stationary HMM with discrete variables can be represented by a simple dynamic Bayesian network [2,3]. This entails that various theoretical results and algorithms available for (dynamic) Bayesian networks can be straightforwardly applied to HMMs as well.

In this paper we focus on enhancing techniques for sensitivity analysis in HMMs, using results from research into sensitivity analysis of Bayesian networks. Our focus is on parameter sensitivity analysis, which is a standard technique for studying how the output of a model varies with variation of its parameters. In this case, the parameters are the (conditional) probabilities specified for the various variables in the model and the output is some posterior probability of interest. Motivated by the fact that the specified parameters are bound to be inaccurate to at least some degree, a parameter sensitivity analysis can serve multiple purposes [4–6]. In the construction phase of the model, sensitivity analysis can be used to identify those parameters for which an accurate assessment seems important. Subsequently, the results from sensitivity analysis can be used as a basis for parameter tuning, as well as for studying the robustness of the model output to changes in the parameters.

In the context of HMMs, sensitivity analysis is usually performed by means of a perturbation analysis where a small change is applied to the parameters, upon which the output of interest is re-computed [7,8]. Perturbation is the general approach to sensitivity analysis in mathematical models, and can be implemented in various ways [9]. A structured analysis of repeated perturbations is typically inefficient, yet necessary to arrive at reliable results [10].

Research into sensitivity analysis has shown that, in the context of Bayesian networks, a simple mathematical function exists that describes an output probability of interest as a function of one or more network parameters [11,12]. The benefit of having such a sensitivity function is that it captures the effects of *any* change in the parameters under consideration, and not just infinitesimal changes. Moreover, various algorithms have been designed that can establish these sensitivity functions

* Tel.: +31 30 2539266; fax: +31 30 2532804.
E-mail address: S.Renooij@uu.nl

0888-613X/$ - see front matter © 2012 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.ijar.2012.06.003

in a reasonably efficient way. Finally, knowledge of the general form of the function can be used, for example, to compute bounds on output probabilities without actually performing the sensitivity analysis [13–16].

Recently, it was demonstrated that the relation between model parameters and output probabilities in HMMs can also be described by simple mathematical functions, similar to Bayesian network sensitivity functions [17]. For determining these functions for HMMs, however, no algorithms exist. Rather, it was suggested to represent the HMM as a dynamic Bayesian network, unrolled for a fixed number of time slices, and to apply existing Bayesian network sensitivity analysis algorithms. In this paper we follow up on this suggestion by demonstrating how HMM sensitivity functions can be computed using Bayesian network sensitivity analysis algorithms. We argue that these methods are inefficient for the purpose of computing HMM sensitivity functions, due to the fact that the repetitive character of the HMM, with the same parameters occurring for each time step, is not exploited in the computation. We then introduce two new algorithms, that build on existing algorithms for HMMs in order to compute the constants of the HMM sensitivity function. In addition, we present a new algorithm that is specially tailored to the computation of sensitivity functions directly from HMMs. To the best of our knowledge, we present the first algorithms for determining an HMM sensitivity function that do not rely on a Bayesian network representation. We introduced the basic ideas behind the tailored algorithm in a previous paper [18]. The current paper extends this work, by applying it to a larger set of inference tasks and by presenting details of this new algorithm. Our tailored algorithm exploits the recursive properties of an HMM by basically performing the Forward–Backward HMM inference algorithm with additional bookkeeping.

This paper is organised as follows. In Section 2, we present some preliminaries concerning HMMs, Bayesian networks and sensitivity functions. In Section 3, we discuss how to compute HMM sensitivity functions using existing algorithms for establishing Bayesian network sensitivity functions. Various approaches that work directly on HMMs are introduced in Section 4, where details of our tailored algorithm are provided in Section 5. We conclude the paper with directions for future research in Section 6.

2. Preliminaries

In this section we present preliminaries concerning Bayesian networks, hidden Markov models, and sensitivity functions. Throughout this paper, variables will be denoted by capital letters, and their values by lower case.

2.1. Bayesian networks

A *Bayesian network* is a discrete, static statistical model for representing and reasoning about a domain of application. In essence, a Bayesian network is a concise representation of the joint probability distribution on the set of statistical variables relevant to the application domain [19,20]. A Bayesian network $\mathcal{B}$ combines an acyclic directed graph $G = (V_G, A_G)$, representing the statistical variables and their dependencies by means of nodes V_G and arcs A_G, with a set of conditional probability distributions $\Theta = \{p(V|\pi_V) | V \in V_G\}$ that describe the strengths of the various dependences between a node V and its immediate predecessors π_V in the graph. More formally, the Bayesian network defines the unique distribution

$$p(V_G) = \prod_{V \in V_G} p(V|\pi_V)$$

on V_G, that respects the probabilistic independences read from the digraph G by means of the d-separation criterion [20]. As such, the network provides for computing any prior or posterior probability over its variables. Computing probabilities from Bayesian networks, also known as inference, is in general NP-hard [21]. However, inference in a Bayesian network whose directed graph takes the form of a tree, where every node has at most one parent, requires a number of computations which is linear in the number of nodes [20].

A *dynamic Bayesian network* can cope with discrete-time evolving processes by repeating and connecting a Bayesian network for a number of time steps, or *time slices* [2]. The relations among the variables within a time slice are taken to be instantaneous, whereas the relationships across time slices are temporal.

2.2. Hidden Markov models

In this section we review the necessary background on hidden Markov models (HMMs), their relation to dynamic Bayesian networks and the recursive properties that underlie inference in HMMs.

2.2.1. Definition of an HMM

A hidden Markov model [22,23] consists of a discrete time Markov chain, repeating a single hidden variable X with a finite number of states. The chain is stationary, i.e., the probability of transitioning from one state to another is time-invariant. The state of the hidden variable in each time step can be indirectly observed by some memoryless test or sensor Y. The uncertainty in the discrete test or sensor output is captured by a set of observation probabilities, which are also time-invariant. Generalisations of HMMs with continuous variables exist, but are not considered here. More formally, an HMM is a statistical model $H = (X, Y, A, O, \Gamma)$, where