Mach Learn
DOI 10.1007/s10994-013-5410-3

A comparison of collapsed Bayesian methods for probabilistic finite automata

Chihiro Shibata · Ryo Yoshinaka

Received: 11 December 2012 / Accepted: 14 August 2013
© The Author(s) 2013

Abstract This paper describes several collapsed Bayesian methods, which work by first marginalizing out transition probabilities, for inferring several kinds of probabilistic finite automata. The methods include collapsed Gibbs sampling (CGS) and collapsed variational Bayes, as well as two new methods. Their targets range over general probabilistic finite automata, hidden Markov models, probabilistic deterministic finite automata, and variable-length grams. We implement and compare these algorithms over the data sets from the Probabilistic Automata Learning Competition (PAutomaC), which are generated by various types of automata. We report that the CGS-based algorithm designed to target general probabilistic finite automata performed the best for any types of data.

Keywords Collapsed Gibbs sampling · Variational Bayesian methods · State-merging algorithms

1 Introduction

Since Hidden Markov Models (HMMs) are implemented in many applications, many inference methods for them have thus far been proposed and refined. It is difficult to find transition probabilities that maximize the generation probability of training samples. It is also intractable to marginalize out state transition probabilities and simultaneously sum them with respect to hidden variables. Therefore, some approximation and/or searching-local-optima technique is required. The Expectation Maximization (EM) algorithm, called Baum-Welch, is the most well-known classic method that is used as a statistical method for

Editors: Jeffrey Heinz, Colin de la Higuera, and Tim Oates.

C. Shibata (✉)
School of Computer Science, Tokyo University of Technology, Tokyo, Japan
e-mail: shibatachh@stf.teu.ac.jp

R. Yoshinaka
Graduate School of Informatics, Kyoto University, Kyoto, Japan
e-mail: ry@i.kyoto-u.ac.jp

Published online: 03 October 2013