International Journal of Approximate Reasoning ••• (••••) •••–•••

Contents lists available at ScienceDirect

International Journal of Approximate Reasoning

www.elsevier.com/locate/ijar

Approximation enhancement for stochastic Bayesian inference [☆], [☆☆]

Joseph S. Friedman [a,b,*], Jacques Droulez [c], Pierre Bessière [c], Jorge Lobo [d], Damien Querlioz [a]

[a] Centre de Nanosciences et de Nanotechnologies, CNRS, Univ. Paris-Sud, Université Paris-Saclay, 220 rue André Ampère, 91405 Orsay, France
[b] Department of Electrical Engineering, The University of Texas at Dallas, 800 W. Campbell Rd., Richardson, TX 75080, USA
[c] Institut des Systèmes Intelligents et de Robotique, Université Pierre et Marie Curie, CNRS, 4 Place Jussieu, 75005 Paris, France
[d] Institute of Systems and Robotics, Department of Electrical and Computer Engineering, University of Coimbra, 3030-290 Coimbra, Portugal

ARTICLE INFO

Article history:
Received 15 May 2016
Received in revised form 12 March 2017
Accepted 14 March 2017
Available online xxxx

Keywords:
Stochastic computing
Muller C-element
Bayesian inference
Autocorrelation
Approximate inference

ABSTRACT

Advancements in autonomous robotic systems have been impeded by the lack of a specialized computational hardware that makes real-time decisions based on sensory inputs. We have developed a novel circuit structure that efficiently approximates naïve Bayesian inference with simple Muller C-elements. Using a stochastic computing paradigm, this system enables real-time approximate decision-making with an area-energy-delay product nearly one billion times smaller than a conventional general-purpose computer. In this paper, we propose several techniques to improve the approximation of Bayesian inference by reducing stochastic bitstream autocorrelation. We also evaluate the effectiveness of these techniques for various naïve inference tasks and discuss hardware considerations, concluding that these circuits enable approximate Bayesian inferences while retaining orders-of-magnitude hardware advantages compared to conventional general-purpose computers.

1. Introduction

The development of autonomous robotic systems requires fast, low-power circuits that enable real-time decision-making. Conventional approaches use general-purpose processors with decision algorithms implemented in software and mapped to conventional Boolean logic and arithmetic. Such systems have the ability to perform a wide range of tasks, but are not ideal for any task, including decision-making. For systems dedicated to a single task, specialized hardware can provide optimizations that lead to increased speed, reduced circuit area, and decreased energy consumption.

To autonomously reason and perform actions based on sensory information, Bayesian inference efficiently incorporate information from independent sources [1]. Bayesian inference has been suggested as a fundamental component of biological systems [2–4], and has been successfully applied to robotics and other sensory-motor systems [5]. In such systems,

☆ This work was supported by the EU collaborative FET Project BAMBI FP7-ICT-2013-C, project number 618024 and a public grant overseen by the French National Research Agency (ANR) as part of the "Investissements d'Avenir" program (Labex NanoSaclay, reference: ANR-10-LABX-0035).
☆☆ This paper is part of the Virtual special issue on Special Issue on Unconventional computing for Bayesian inference, edited by Jorge Lobo and João Filipe Ferreira.
* Corresponding author at: Department of Electrical Engineering, The University of Texas at Dallas, 800 W. Campbell Rd., Richardson, TX 75080, USA.
E-mail addresses: joseph.friedman@utdallas.edu (J.S. Friedman), jacques.droulez@college-de-france.fr (J. Droulez), pierre.bessiere@college-de-france.fr (P. Bessière), jlobo@isr.uc.pt (J. Lobo), damien.querlioz@u-psud.fr (D. Querlioz).

http://dx.doi.org/10.1016/j.ijar.2017.03.007
0888-613X/© 2017 Elsevier Inc. All rights reserved.

Table 1
Muller C-element truth table.

X	Y	Z
0	0	0
0	1	Z_{prev}
1	0	Z_{prev}
1	1	1

constantly-updated sensory information is provided to a reasoning circuit that makes decisions in response to new information resulting from a constantly-changing environment [6]. For such circuits, there are fundamental trade-offs between reaction time, circuit efficiency, and the level of approximation used for the inference and reasoning. Therefore, the design of systems specifically dedicated to Bayesian inference is an active research area, incorporating analog or digital circuits to enable systems that are more efficient than computers [7–11].

In this context, we recently proposed an especially compact and energy efficient computing system that provides approximate naïve Bayesian inference [12]. In this system, based on Muller C-elements used in a stochastic computing paradigm, the approximate probability of a particular event is calculated based on prior and evidence data. We demonstrated a nearly one-billion-fold improvement in the area-energy-delay product for approximate Bayesian inference, with exceptional robustness to hardware faults. However, we found that bitstream autocorrelation leads to an approximation of the inference, limiting this system to tasks that do not require exact computation.

In this paper, we explore techniques to improve the quality of the approximation by mitigating the bitstream autocorrelation. Circuits that reduce bitstream autocorrelation are presented that remove the inaccuracy at the cost of additional hardware components that reduce the system efficiency. We show that these autocorrelation mitigation techniques are effective in reducing bitstream autocorrelation, even for difficult inference tasks. Finally, we conclude that despite the hardware costs, autocorrelation mitigation permits the use of our stochastic C-element structure for approximations of Bayesian inference with massive efficiency improvements over conventional hardware systems.

2. Bayesian inference with stochastic C-elements

In this section, we summarize the findings of [12]. Approximate Bayesian inference can be performed with a simple cascade of C-elements, with stochastic bitstream inputs and outputs. This system provides several orders of magnitude improvement in computational efficiency, though bitstream autocorrelation limits the accuracy of the approximation.

2.1. Stochastic computing

Stochastic computing enables the efficient approximation of mathematical functions performed on streams of random binary values [13,14]. The value of a stochastic bitstream is the percentage of binary bits that are '1'. For example, "010011010" encodes $\frac{4}{9}$. The random nature of the bitstream prevents the encoding of exact values; the value is an approximation with an accuracy that increases with the bitstream length.

This bitstream encoding enables the efficient realization of various functions, notably multiplication with a single AND gate [14]. For example, consider two input bitstreams to an AND gate "0101010101" and "0000011111", which both encode a value of 0.5. If the bit-pairs are input sequentially into an AND gate, the resulting output is "0000010101", which has a value of 0.3. This is not the exact 0.25 result of the multiplication, but it is as precise as possible given that only ten bits are employed.

Correlations between bitstreams, and within a bitstream, can cause significant errors. For the case of an AND gate with both input bitstreams being "0101010101": the output is also "0101010101", which is equivalent to 0.5. As the bitstreams are correlated with each other, this AND-gate stochastic multiplication does not perform the desired function. Additionally, in cases of autocorrelation with bits interacting within a bitstream, this autocorrelation can impact the output bitstream to produce undesired results. This autocorrelation is the primary challenge that is addressed in this paper, and the resulting imprecision is considered as a tradeoff for decreased power consumption, area, and latency.

2.2. C-element inference

The small area of circuits developed for stochastic computing enables an especially efficient Bayesian inference function. In particular, a Muller C-element is a simple circuit composed of as few as eight transistors that performs the function described by Table 1 [15]. The output Z maintains its state Z_{prev} unless both inputs X and Y are opposite the current output state, in which case the output switches to the shared input value. For C-element input signals with no autocorrelation, the output probability is approximated by [12,16–18]

$$P(Z) = \frac{P(X)P(Y)}{P(X)P(Y) + (1 - P(X))(1 - P(Y))}. \tag{1}$$

Please cite this article in press as: J.S. Friedman et al., Approximation enhancement for stochastic Bayesian inference, Int. J. Approx. Reason. (2017), http://dx.doi.org/10.1016/j.ijar.2017.03.007