PERGAMON

Neural Networks 13 (2000) 555–556

Neural Networks

www.elsevier.com/locate/neunet

Book review

Adaptive Resonance Theory Microchips: Circuit Design Techniques

T. Serrano-Gotarredona, B. Linares-Barranco, A.G. Andreou; Kluwer Academic Publishers, Boston, 1998, hardbound, ISBN: 0-792-38231-5, US$ 120.00

Adaptive Resonance Theory Microchips Circuit Design Techniques by Teresa Serrano-Gotarredona, Bernabé Linares-Barranco, and Andreas G. Andreou is a book that is well worth obtaining. The book combines an appropriately terse review of adaptive resonance theory (ART) algorithms, a deep look at circuitry comprising "ART chips", and a glimpse at potential applications for the chips. The text should be readable with moderate difficulty by engineers and scientists at the mezzanine graduate level. For computational neuroscientists, this book explains the basic obstacles faced in designing neural microchips and the steps taken to overcome these obstacles. For neural engineers, both beginning and advanced, this book covers the many details that make the designs work, which are often left out of journal papers.

Adaptive Resonance Theory is an established neural network theory developed by Gail Carpenter, Stephen Grossberg, and their collaborators at Boston University and the Center for Adaptive Systems. Most of the ART implementation literature to date has reported work on architectures and applications that have been developed as software for conventional sequential computers. However, the parallel nature of the theoretical ART architectures lends itself to a hardware implementation as a VLSI system.

This is the first book that addresses efficient VLSI design of ART systems. Unfortunately, none of the ART algorithms can be built explicitly on a chip without compromise. Herein lies the essence of the analog designer's challenge: the chip itself must be a merging of a simplified model that preserves salient aspects of the original theory, with the realities of what can be built in terms of working circuits. "VLSI-friendly" modifications to the original ART1 algorithms are elaborated, and side-by-side comparisons between the original and modified algorithms are presented. The structure of the book is as follows:

Chapter 1 reviews the salient feature of several ART designs, e.g. ART1, ART2, Fuzzy ART and ARTMAP. The authors provide what amounts to a "circuit designer's" approach to framing the volumes of text that exist in the formal literature. This introductory chapter is not about circuit design, but it puts the reader into a framework from which realities of ART circuitry can be faced. Example training sets are presented and then simulated using MATLAB. Much of the MATLAB code is included in an appendix and is also available on the web from http://www.imse.cnm.es/~bernabe.

Chapter 2 presents a variant of the original ART1 architecture. The variation is necessary because several components of the original ART architecture are not too " VLSI-friendly". Modifications to the original ART1 architecture are described in the context of how each change was necessary in order to implement the system in analog hardware. This is one of the book's strongest points as it allows readers insight into the authors' thought processes. The previous example training sets are presented to the modified ART1 architecture and are simulated using MATLAB. The chapter concludes by comparing the simulation results of the original and modified systems. Computational equivalency of the modified and original ART1 algorithms is proven in an appendix.

Chapter 3 begins the hardware part of the book. The beginning of the chapter is devoted to a high-level description of the modified ART1 microchip. In a manner that is easy to follow, the authors describe the five operations needed to implement the system. This is accomplished at a block diagram level while keeping the circuit details for later in the chapter. This allows the reader to examine each of the operations as it fits into the overall ART1 algorithm and also to see how the blocks (or operations) are connected together. From a high level description, the authors move into a description of each block (or operation) of the ART1 architecture. It is here that a non-engineer might have difficulty, as the authors provide extensive circuit schematics with very brief explanations. To those familiar with circuit design, the circuit schematics are easily understandable. Next, the chapter dwells on a topic for more advanced engineers as it discusses issues at the frontier of neural system engineering research. Those issues are reliability of circuits and connectivity, either on-chip or between different chips. Basically, the inter-chip connectivity problem is a result of not having an unlimited financial resource with which to build chips. Thus, academic researchers have had to devise a way to hook up multiple smaller chips in an effort to develop larger ART systems like those demonstrated on sequential computers. The later part of the chapter presents the test results from two sets of microchips based on the modified ART1 architecture. The first set of microchips has a small number of clustering categories, as it was a prototype chip. Circuitry tests as well as input/output test patterns are provided. Using the

PII: S0893-6080(00)00023-X

results of the first microchip set, the authors built a better and bigger system with more clustering categories. Again, circuitry tests as well as input/output test patterns are provided.

Chapter 4 is about winner-take-all (WTA) circuitry and its application to multi-chip capability. This is an advanced chapter about the circuitry used in the ART1 implementation described in Chapter 3. This chapter begins by describing the basics of WTA circuitry from an architecture level. This is followed by a more in-depth discussion of the circuit implementation of WTA than was allotted in Chapter 3. The authors again give the readers insight into their thinking during the design of the system by examining different types of current mirrors that could have been used in the system. Lastly, experimental circuit results from the WTA circuits are examined.

Chapter 5 is an invited chapter on Fuzzy-ART by Gert Cauwenberghs and his collaborators from Johns Hopkins University. Whereas the ART1 architecture was a binary-valued clustering-type system, the Fuzzy-ART is a continuous-valued, discrete time system. Again, the authors start from a high-level overview of the system by discussing the system with an equation and hardware block-level description. As for the circuit implementation, some of the discussions of the hardware blocks are fairly terse while in other cases the circuit schematics as well as VLSI layout are explained in great detail. The chapter concludes by giving microchip results about the minimum category distance.

Overall, we recommend this book. For the beginner interested in building neural systems like the ART architecture, this book can provide good insight into thought processes that are generally useful. For the advanced engineer interested in neural systems, this book offers many useful discussions on chip interconnectivity and how to put together a relatively large neural system. Data provided in an appendix on transistor mismatches are a highly unexpected and interesting find. Finally, for the computational neuroscientist interested in how hardware systems are built, this book can be read as an overview demonstrating the level of difficulty of the task.

A.E. Hubbard
T.A. Hinck
Boston University VLSI and Neural Systems Laboratory,
College of Engineering,
Boston, MA 02215, USA
E-mail address: aeh@vlsi.bu.edu