Preface

This book arose from a pair of symposia on hidden Markov models (HMMs) that Kevin Vixie organized at the 2001 SIAM Conference on Applications of Dynamical Systems. At the end of the first symposium someone asked the speakers for a simple reference that explains the basic ideas and algorithms for applying HMMs. We were stumped. A group of the participants suggested writing this book to answer the question. I have aimed the book at readers who have backgrounds and interests typical of those who attended that conference. In my view, HMMs are discrete-state, discrete-time stochastic dynamical systems, and they are often used to approximate dynamical systems with continuous state spaces operating in continuous time. Thus, by using familiar analogies, it is easy to explain HMMs to readers who have studied dynamical systems.

The basic techniques were well developed in the 1970's for work on speech and language processing. Many in speech research learned about the techniques at a symposium held at the Institute for Defense Analysis in Princeton, NJ. J. D. Ferguson edited the proceedings [4], and copies were given to the attendees.[1] The volume was called *the blue book* by workers in the field. I was not part of that community, but I have a copy of the blue book. It explains the basic algorithms and illustrates them with simple concrete examples. I hope *this* book is as simple, useful, and clear.

Although there are other books and papers that are about HMMs exclusively or in part, I hope that readers find the following features of this present volume useful:

It is introductory. An undergraduate background in engineering, mathematics, or science that includes work in probability, linear algebra, and differential equations provides the prerequisites for most of the book. The exceptions are ideas from dynamical systems and information theory. In particular, I use the Gibbs inequality (see (2.53)) in developing the estimate maximize (EM) algorithm in Chapter 2. Although Chapter 5 deals with Lyapunov exponents and entropies, it is not a prerequisite for any other chapter.

Algorithms are explained and justified. I present enough of the theory behind the basic algorithms in Chapter 2 so that a reader can use it as a guide to developing his own variants.

I provide code implementing the algorithms and data for the examples. Although algorithms are given in pseudocode in the text, a working implementation of each of the

[1] The volume is available at a few libraries.

algorithms that I describe is available on the Web [56]. I have chosen to write the programs in the Python language [59] because it is easy to read and the interpreter is free software. I have written the programs to follow the descriptions and notation in the text. I provide data and scripts (makefiles, shell scripts, etc.) that make all of the figures in the text. On a GNU system, issuing "make book.pdf" from a command line compiles the software, runs the numerical experiments, makes the figures, and formats the entire book.

It uses analogies to dynamical systems. For example, I demonstrate the HMM training algorithm by applying it to data derived from the Lorenz system. The result, as Fig. 1.9 illustrates, is that the algorithm estimates a discrete-state generating mechanism that is an approximation to the state space of the original Lorenz system.

I illustrate with a practical example. In Chapter 6, I present an application to experimental measurements of electrocardiograms (ECGs).

Acknowledgments

In writing this book, I have used much that I learned from colleagues in Portland, OR. In particular, Todd Leen kept me informed about developments in the machine learning community. I have relied on the review of the ergodic theory of dynamical systems in Kevin Vixie's dissertation [27] for Chapter 5 of this book. Shari Matzner and Gerardo Lafferriere helped me understand the convergence properties of the EM algorithm. Also the following colleagues have given me constructive comments on drafts of the book: Katherine Backus, Patrick Campbell, Ralf Juengling, Shari Matzner, Kary Myers, Reid Porter, Cosma Shalizi, and Rudolph Van der Meer.

I thank the following people and organizations for providing the data that I used for examples:

Carl Otto Weiss for providing Tang's laser data [51],

PhysioNet for providing Penzel's ECG data [20] and George Moody's *WFDB* software [58],

Project Gutenberg for digitizing and distributing *A Book of Prefaces*, by H. L. Mencken.

I was fortunate to meet Karl Hegbloom in Portland. Karl contributes to several free software projects, and he helped me with the figures and software. Ralf Juengling also helped with the software. In addition to personal help with software, I relied on free software written by too many contributors to mention. Almost all of the tools I used are available in the Debian [55] and Ubuntu [61] distributions.

I acknowledge Portland State University for support of a sabbatical leave during which I started the book and for a computer on which I did much of the development. I thank Principal Investigator James Kamm for arranging to have the Los Alamos National Laboratory LDRD Office support work on Chapter 6 through project 20030037DR.