AI Education Matters: Teaching Hidden Markov Models

Todd W. Neller (Gettysburg College; tneller@gettysburg.edu)

DOI: 10.1145/3203247.3203252

Introduction

In this column, we share resources for learning about and teaching Hidden Markov Models (HMMs). HMMs find many important applications in temporal pattern recognition tasks such as speech/handwriting/gesture recognition and robot localization. In such domains, we may have a finite state machine model with known state transition probabilities, state output probabilities, and state outputs, but lack knowledge of the states generating such outputs. HMMs are useful in framing problems where external sequential evidence is used to derive underlying state information (e.g. intended words and gestures).

Video Introductions

While there are many videos online dedicated to the topic of HMMs, I'll highlight two here. Daphne Koller's 12-minute video"Template Models: Hidden Markov Models - Stanford University"[1] provides a brief application-focused overview of HMMs and can set a basic context and expectation for the value of further learning in this area.

A full 52-minute UBC lecture by Nando de Freitas, "undergraduate machine learning 9: Hidden Markov models - HMM"[2], is a much-recommended classroom introduction to the general topic of inference in probabilistic graphical models, with focus on HMM representation, prediction, and filtering.

Texts and Articles

Russell and Norvig's *Artificial Intelligence: a modern approach* (Russell & Norvig, 2009) gives a brief introduction to HMMs in §15.3, and learning HMMs as an application of the expectation-maximization (EM) algorithm is described in §20.3.3.

[1] https://youtu.be/mNSQ-prhgsw

[2] https://youtu.be/jY2E6ExLxaw

Of the many options to give focused attention to HMMs, consider *Speech and Language Processing* by Daniel Jurafsky and James H. Martin (Jurafsky & Martin, 2009). The draft third edition chapter 9 on HMMs is currently freely available from Jurafsky's website[3].

Also often recommended is Lawrence Rabiner's tutorial (Rabiner, 1990)[4].

Christopher Bishop's *Pattern Recognition and Machine Learning* (Bishop, 2006) §13.2 covers HMMs, maximum likelihood parameter estimation, the forward-backward algorithm, the sum-product algorithm, the Viterbi algorithm, and extensions to HMMs.

More text and article resources have been recommended in the StackExchange thread "Resources for learning Markov chain and hidden Markov models"[5].

Other Resources

Numerous MOOCs touch on HMMs. Udacity's "Intro to Artificial Intelligence" course by Peter Norvig and Sebastian Thrun has coverage of HMMs[6], as does "Artificial Intelligence - Probabalistic Models"[7]. In Coursera's Peking Univeristy Bioinformatics course, Ge Gao lectures on HMMs[8]. There is no shortage of online course materials for HMMs.

[3] https://web.stanford.edu/~jurafsky/slp3/9.pdf

[4] http://www.ece.ucsb.edu/Faculty/Rabiner/ece259/Reprints/tutorial%20on%20hmm%20and%20applications.pdf

[5] https://stats.stackexchange.com/questions/3294/resources-for-learning-markov-chain-and-hidden-markov-models

[6] https://www.udacity.com/course/intro-to-artificial-intelligence--cs271

[7] https://www.udacity.com/course/probabalistic-models--cx27

[8] https://www.coursera.org/learn/bioinformatics-pku/lecture/7pbUo/hidden-markov-model

Jason Eisner has written about his spreadsheet[9] for teaching the forward-backward algorithm (Eisner, 2002). More diverse resources, including toolkits and open-source algorithm implementations, may be sampled via the Quora question "What are some good resources for learning about Hidden Markov Models?"[10].

Model AI Assignments

Model AI Assignments are free, peer-reviewed assignment materials made available in order to advance AI education. Two Model AI Assignments to date experientially teach HMMs:

Sravana Reddy's recent Model AI Assignment "Implementing a Hidden Markov Model Toolkit"[11] is targeted to advanced undergraduates and beginning graduate students seeking an introduction to Natural Language Processing (NLP). In the assignment, students "implement a toolkit for Hidden Markov Models with discrete outputs, including (1) random sequence generation, (2) computing the marginal probability of a sequence with the forward and backward algorithms, (3) computing the best state sequence for an observation with the Viterbi algorithm, and (4) supervised and unsupervised maximum likelihood estimation of the model parameters from observations, using the Baum Welch algorithm for unsupervised learning."

John DeNero and Dan Klein's popular Model AI Assignment "The Pac-Man Projects" has a probabilistic tracking project, "Project #4: Ghostbusters"[12] in which "probabilistic inference in a hidden Markov model tracks the movement of hidden ghosts in the Pac-Man world. Students implement exact inference using the forward algorithm and approximate inference via particle filters."

[9] http://cs.jhu.edu/~jason/papers/eisner.hmm.xls

[10] https://www.quora.com/What-are-some-good-resources-for-learning-about-Hidden-Markov-Models

[11] http://modelai.gettysburg.edu/2017/hmm

[12] http://modelai.gettysburg.edu/2010/pacman/projects/tracking/busters.html

Your Favorite Resources?

If there are other resources you would recommend, we invite you to register with our wiki and add them to our collection at http://cs.gettysburg.edu/ai-matters/index.php/Resources.

References

Bishop, C. M. (2006). *Pattern recognition and machine learning (information science and statistics)*. Secaucus, NJ, USA: Springer-Verlag New York, Inc.

Eisner, J. (2002). An interactive spreadsheet for teaching the forward-backward algorithm. In *Proceedings of the acl-02 workshop on effective tools and methodologies for teaching natural language processing and computational linguistics - volume 1* (pp. 10–18). Stroudsburg, PA, USA: Association for Computational Linguistics. Retrieved from https://cs.jhu.edu/~jason/papers/eisner.tnlp02.pdf

Jurafsky, D., & Martin, J. H. (2009). *Speech and language processing* (2nd ed.). Upper Saddle River, NJ, USA: Prentice Hall.

Rabiner, L. R. (1990). A tutorial on hidden markov models and selected applications in speech recognition. In A. Waibel & K.-F. Lee (Eds.), *Readings in speech recognition* (pp. 267–296). San Francisco, CA, USA: Morgan Kaufmann.

Russell, S., & Norvig, P. (2009). *Artificial intelligence: A modern approach* (3rd ed.). Upper Saddle River, NJ, USA: Prentice Hall.

Todd W. Neller is a Professor of Computer Science at Gettysburg College. A game enthusiast, Neller researches game AI techniques and their uses in undergraduate education.