

Modelling Stem Cells Lineages with Markov Trees

Victor Olariu, Daniel Coca, Stephen A. Billings,
and Visakan Kadirkamanathan

Department of Automatic Control and Systems Engineering
The University of Sheffield
UK

Abstract. A variational Bayesian EM with smoothed probabilities algorithm for hidden Markov trees (HMT) is proposed for incomplete tree structured data. The full posterior of the HMT parameters is determined and the underflow problems associated with previous algorithms are eliminated. Example results for the prediction of the types of cells in real stem cell lineage trees are presented.

1 Introduction

The existence of stem cells in inter convertible sub-states and the kinetics of the cells switching between the sub-states are observed using cell tracking and real time phenotype monitoring techniques. However, the available technologies are limited and the resulting stem cells lineage trees are incomplete. To confront this problem we use probabilistic techniques of analysis for cell lineage trees reconstruction based on observations gathered in real time (cells division rate) and combine this with particular surface antigen expression information gathered at the end of the set of cell divisions being monitored. We take as our starting point the Hidden Markov Models (HMM) which are used in various fields like speech recognition, financial time series prediction [7], natural language processing, ion channel kinetics [17] and general data compression [21]. They have played important roles in the modeling and analysis of biological sequences, in particular DNA, [26], [5] and they have proven to be useful tools for statistical signal and image processing.

Baum and colleagues developed the core theory of Hidden Markov Models [3]. In 1972 they proposed the forward-backward algorithm as an iterative technique for the Maximum Likelihood statistical estimation of probabilistic functions of Markov chains. Devijver demonstrated that the computation of joint likelihoods in Baum's algorithm could be converted to the computation of posterior probabilities [11]. The resulting algorithm was similar to Baum's except for the presence of a scaling factor suggested by Levinson et al. [22] which was robust to computational underflow. Further developments in HMMs has been done by MacKay [23], Beal and Ghahramani [4], Watanabe et al. [27], Ji et al. [19] in which they apply a variational Bayesian approach to these models.

Hidden Markov tree (HMT) models have been proposed by Crouse et al. for modelling the statistical dependencies between wavelet coefficients in signal processing [9]. They have been applied successfully to image de-noising and segmentation [8], [24], [6] to signal processing and classification [12], [14] and to tree structured data modelling [15], [5]. The forward-backward algorithm proposed by Baum was transposed to the Hidden Markov Trees context by Crouse et al. [9]. The resulting algorithm has been called the upward-downward algorithm but it suffered computational underflow problems as in Baum's algorithm. The upward downward recursions have been proposed by Ronen et al. for the E-step in maximum likelihood estimation of dependence tree models with missing observations [25]. The upward-downward algorithm was revisited by Durand et al. [14] who made changes to solve the computational underflow problems by adapting the ideas from Devijver's changes to the forward-backward algorithm. Romberg et al. proposed a Bayesian HMT model for image processing using wavelets [24] and later, Dasgupta and Carin developed the variational Bayesian hidden Markov tree model based on the model proposed by Crouse et al. with a similar application [10].

In this study we derive the variational Bayesian with smoothed probabilities implementation of Hidden Markov trees. We extend Durand's HMT framework [14] to variational Bayesian with the critical embodiment of prior probability distributions. Inclusion of prior probability distributions of a class of HMT models such as in the case of cell lineages is essential to avoid ill-posedness of the estimation problem. We demonstrate this through an application to modelling stem cell lineages using real data.

2 Hidden Markov Tree Model

A Hidden Markov Tree (HMT) model is composed of the observed random tree $X = x_1, \dots, x_N$ and hidden random tree $S = s_1, \dots, s_N$ which has the same indexing structure as the observed tree. S takes value in a set of k discrete states, which are referred as $1, \dots, k$. A distribution $P()$ satisfies the HMT property if and only if:

$$P(X, S) = P(s_1) \prod_{t \neq 1} P(s_t | s_{\rho(t)}) \prod_t P(x_t | s_t) \quad (1)$$

$\rho(t)$ represents the parent of node t , $\mathcal{C}(t)$ is the notation for the children of node t , X_t is the subtree rooted in t and $X_{1/t}$ represents the entire tree except for the subtree rooted in t .

The parameters describing the HMT model are similar to the HMM model parameters:

$$\begin{array}{ll} \pi_j = P(s_1 = j) & \text{is the initial hidden state prior} \\ P_{ij} = P(s_t = j | s_{\rho(t)} = i) & \text{is the transition probability matrix} \\ C_{jh} = P(x_t = h | s_t = j) & \text{is the emission probability matrix.} \end{array}$$

for $j = 1 \dots k$ where k is the number of possible discrete values of states.

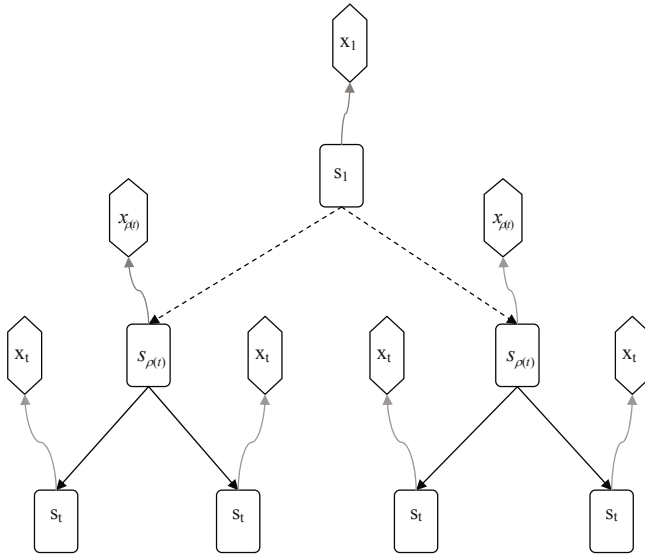


Fig. 1. Hidden Markov Tree (HMT) representation with observed nodes (x) and hidden nodes (s). The straight arrows show the state transitions, curly arrows show emission probabilities.

In the next sections we will refer to all parameters of the HMT model as θ .

$$\theta = [\pi, \text{vec}(P), \text{vec}(C)]^T$$

where $\text{vec}(\cdot)$ rearranges the matrix into a column vector.

3 Maximum Likelihood Estimation for Hidden Markov Tree Model

Considering that the S states are not observable, a popular approach to determine maximum likelihood estimates is the EM algorithm. For the E-step Crouse et al. realised a direct transposition to the Hidden Markov Tree (HMT) context of the forward-backward algorithm [9]. The result is the upward-downward algorithm which suffers from underflow problems [16]; [22]. In order to overcome this, Durand et al. [14] proposed the upward-downward algorithm for smoothed probabilities. Upward recursion for leaves of the tree:

$$\beta_t(j) = \frac{C_{jt}P(s_t = j)}{N_t} \tag{2}$$

where

$$N_t = P(x_t) = \sum_j C_{jt}P(s_t = j)$$

Upward recursion for non-leaves:

$$\beta_t(j) = \frac{\{\prod_{v \in \mathcal{C}(t)} \beta_{t,v}\} C_{jt} P(s_t = j)}{N_t} \tag{3}$$

where

$$N_t = \sum_j \{ \prod_{v \in \mathcal{C}(t)} \beta_{t,v} \} C_{jt} P(s_t = j)$$

and

$$\beta_{\rho(t),t}(j) = \sum_k \frac{\beta_t(k) P_{jk}}{P(s_t = k)}$$

Downward recursion:

$$\alpha_t(j) = \frac{1}{P(s_t = j)} \sum_i \frac{P_{ij} \beta_{\rho(t)}(i) \alpha_{\rho(t)}(i)}{\beta_{\rho(t),t}(i)} \tag{4}$$

At the M step, the maximisation of expectation of log-likelihood of the complete data $\ln \mathcal{Q}(X|\theta)$ is realised in order to re estimate the model parameters to be used in the next iteration.

$$\theta_{\tau+1} = \arg \max_{\theta} \ln \mathcal{Q}_{\tau}(X|\theta) \tag{5}$$

where $\theta_{\tau+1}$ represents the model parameters at iteration $\tau + 1$.

The α and β probabilities determined at the E-step are used to find the expression of $\ln \mathcal{Q}(X|\theta)$ as a function of parameter θ of the hidden Markov tree [25].

$$\begin{aligned} \ln \mathcal{Q}(X|\theta) = & \langle \sum_j P(s_1 = j|X, \theta) \ln P(s_1 = j) \\ & \sum_k \sum_j P(s_t = j|X, \theta) \ln P(s_t = j|s_{\rho(t)} = k) \rangle \end{aligned} \tag{6}$$

where the angled brackets $\langle \cdot \rangle$ denote the expectation of a conditional probability function with respect to the missing components.

Taking the derivative of $\ln \mathcal{Q}(X|\theta)$ and equating to zero gives the new parameters as shown:

$$\theta_t^{i,j} = \frac{1}{\lambda_{j,t}} \left(\sum_{v=\rho(leaf)} A_{\rho(t),v}^j + \sum_{v \neq \rho(leaf)} B_{t,v}^{i,j} \right) \tag{7}$$

where

$$A_{t,v}^j = \frac{\alpha_t(j) \beta_t(j)}{\sum_j \alpha_t(j) \beta_t(j)} \quad B_{t,v}^{i,j} = \frac{\alpha_{\rho(t)}(j) \beta_{\rho(t),t}(j) \beta_t(j) \theta_t^{i,j}}{\sum_j \alpha_t(j) \beta_t(j)}$$

The upwards-downwards algorithm mentioned above represents the base of the E step of the variational Bayesian EM algorithm with smoothed probabilities developed in the next section.

4 Variational Bayesian EM with Smoothed Probabilities Algorithm (VBEMS)

The ML approach for estimating the hidden Markov tree (HMT) model parameters produces just a single point estimate, at the same time ML tends to overfit the data. The solutions to these problems are given by the variational Bayesian approach proposed by [2]. This framework applied to HMTs is able to estimate approximate posterior distributions over hidden variables and parameters of the model. The computation is closely related to the computation of the EM algorithm which guarantees convergence. The algorithm proposed here for HMT uses the same strategy as the algorithm proposed for HMM by [23] and by [4]. The variational Bayesian framework in this study is adapted to HMT in a similar fashion to [10] but instead of using the simple upward-downward algorithm for the E-step we used the upward-downward method with smoothed probabilities. The aim is to determine the posterior probability distribution of the hidden variables S and parameters θ based on the set of observed variables X . For most model structures the exact deduction of hidden variables and parameters based on observed variables is not possible, therefore a variational distribution $q(S, \theta)$ which approximates the true posterior distribution must be obtained [4]. The log-marginal probability of the observed variables X can be decomposed as:

$$\ln P(X) = \mathcal{L}(q(S, \theta)) + KL(q(S, \theta) || P(S, \theta | X)) \tag{8}$$

$$\text{where } \mathcal{L}(q(S, \theta)) = \int d\theta \sum_S q(S, \theta) \ln \left\{ \frac{P(X, S | \theta)}{q(S, \theta)} \right\} \text{ and}$$

$$KL(q(S, \theta) || P(S, \theta | X)) = - \int d\theta \sum_S q(S, \theta) \ln \left\{ \frac{P(S, \theta | X)}{q(S, \theta)} \right\}$$

We consider $q(S, \theta)$ to be member of conjugate-exponential family and we seek the member of this family for which the KL divergence [20] between the variational posterior distribution approximation and the true posterior distribution is minimised. Minimising KL divergence with respect to $q(S, \theta)$ is equivalent to maximising the lower bound $\mathcal{L}(q(S, \theta))$. For achieving tractability we make the assumption that $P(S, \theta | X) \approx q(\theta)q(S)$ [4].

$$\begin{aligned} \mathcal{L}(q(\theta), q(S)) &= \int d\theta \sum_S q(\theta, S) \ln \frac{p(\theta)P(S, X | \theta)}{q(\theta, S)} \\ &= \int d\theta q(\theta) \left[\ln \frac{p(\theta)}{q(\theta)} + \sum_S q(S) \ln \frac{P(S, X | \theta)}{q(S)} \right] \end{aligned} \tag{9}$$

where $p(\theta)$ represents the prior distribution of the parameters and $q(S, \theta)$ are the variational posterior distributions. The prior distribution is restricted to the factorisation $p(\theta) = p(\pi)p(P_{ij})p(C_{jh})$. We chose the parameter priors over π , the rows of P_{ij} and the rows of C_{jh} to be Dirichlet distributions.

$$p(\pi) = \frac{\Gamma(U)}{\prod_j \Gamma(u_j)} \prod_j \pi_j^{u_j-1} \tag{10}$$

where $U = \sum_j u_j$ is the strength of the prior and hyperparameters u_j are subject to the constraint $u_j > 0$. The Dirichlet distributions have the advantage that they are conjugate to the complete-data likelihood terms, and they are appropriate for our model parameters which are probabilities, hence restricted to the interval $[0, 1]$. The variational posterior distributions have the same form as the priors with hyperparameters incremented by statistics of the observations and hidden states. At the E-step, the posterior distribution over the hidden nodes is computed by calculating the solution of $\delta\mathcal{L}(q)/\delta q(S) = 0$:

$$\begin{aligned} \ln q(S) &= s_1 \langle \ln \pi \rangle_{q(\pi)} + \sum_t s_{t-1} \langle \ln P_{ij} \rangle_{q(P_{ij})} s_t \\ &+ \sum_t s_t \langle \ln C_{jh} \rangle_{q(C_{jh})} x_t - \mathcal{Z} \end{aligned} \tag{11}$$

where \mathcal{Z} is a normalisation constant. The expression of the updated parameters is:

$$\begin{aligned} \hat{\theta} &= (\hat{\pi}, \hat{P}_{ij}, \hat{C}_{jh}) \\ &= (\exp \langle \ln \pi \rangle_{q(\pi)}, \exp \langle \ln P_{ij} \rangle_{q(P_{ij})}, \exp \langle \ln C_{jh} \rangle_{q(C_{jh})}) \end{aligned} \tag{12}$$

Based on the result $\int d\pi Dir(\pi; u) \ln \pi_j = \psi(u_j) - \psi(\sum_{j=1}^k u_j)$ where ψ is the digamma function, we calculate the expectation of the logarithm of the parameters under Dirichlet distributions:

$$\hat{\pi} = \exp[\psi(\omega_j^\pi) - \psi(\sum_{j=1}^k \omega_j^\pi)] \tag{13}$$

$$\hat{P}_{ij} = \exp[\psi(\omega_j^{P_{ij}}) - \psi(\sum_{j=1}^k \omega_j^{P_{ij}})] \tag{14}$$

$$\hat{C}_{jh} = \exp[\psi(\omega_j^{C_{jh}}) - \psi(\sum_{j=1}^k \omega_j^{C_{jh}})] \tag{15}$$

where k represents the number of possible discrete values of states. For the expectation step, we use the upwards-downwards with smoothed probabilities algorithm with the sub-normalised parameters where the normalisation constants change. This way the numerical stability is guaranteed and we are able to determine the β and α probabilities necessary for the maximisation step:

$$\beta_t(j) = \frac{\{\prod_{v \in \mathcal{C}(t)} \beta_{t,v}\} \hat{C}_{jt} P(S_t = j)}{N_u} \tag{16}$$

where

$$N_u = \sum_j \{ \prod_{v \in \mathcal{C}(t)} \beta_{t,v} \} \hat{C}_{jt} P(s_t = j)$$

and

$$\begin{aligned} \beta_{\rho(t),t}(i) &= \sum_k \frac{\beta_t(k) \hat{P}_{ik}}{P(s_t = k)} \\ \alpha_t(j) &= \frac{1}{P(s_t = j)} \sum_i \frac{\hat{P}_{ij} \beta_{\rho(t)}(i) \alpha_{\rho(t)}(i)}{\beta_{\rho(t),t}(i)} \end{aligned} \tag{17}$$

The M-step involves calculation of the variational posterior distribution of each parameter of the HMT model by solving $\delta \mathcal{L}(q) / \delta q(\theta) = 0$. They are Dirichlet distributions and they are functions of expected values which can be determined using the upward and downward probabilities from the E-step. The expressions are similar to the ones used in the original variational Bayesian algorithm for hidden Markov models, with the difference being in the expectations which are functions of the smoothed α and β probabilities. The M-step results in:

$$\begin{aligned} q(\pi) &= Dir(\pi_{1:k}; \omega_{1:k}^\pi) \\ \omega_j^\pi &= u_j^\pi + \langle s_1 = j \rangle_{q(S)} \end{aligned} \tag{18}$$

$$\begin{aligned} q(P_{ij}) &= \prod_{i=1}^k Dir(P_{i,1:k}; \omega_{i,1:k}^{P_{ij}}) \\ \omega_{ij'}^{P_{ij}} &= u_{ij'}^{P_{ij}} + \langle s_{\tau-1} = i s_\tau = j' \rangle_{q(S)} \end{aligned} \tag{19}$$

$$\begin{aligned} q(C_{jh}) &= \prod_{j=1}^k Dir(C_{j,1:p}; \omega_{j,1:p}^{C_{jh}}) \\ \omega_{jh}^{C_{jh}} &= u_h^{C_{jh}} + \langle s_\tau = j x_\tau = h \rangle_{q(S)} \end{aligned} \tag{20}$$

The variational posterior distribution has the same functional form of the Dirichlet distribution. The hyperparameters are equal to the sum between the strength of the prior distribution and statistics of the hidden state and observations which are functions of α and β determined at the E step.

5 Experimental Results

At this moment the technology of monitoring the cells' divisions is not able to determine the types of cells with respect to SSEA3 marker antigen at all levels of division in a stem cell lineage tree. SSEA3 is a cell surface antigen which is rapidly down-regulated as human Embryonic Stem (hES) cells differentiate [13]. The challenge for scientists is to reconstruct the cell lineage trees based on the observations gathered from experimental data. Frumkin et al. reconstructed cell lineage trees based on somatic microsatellites mutation rates [18].

In this study the reconstruction of the lineage trees is realised based on the observations of the SSEA3 expression level of the cells. The lineage trees were

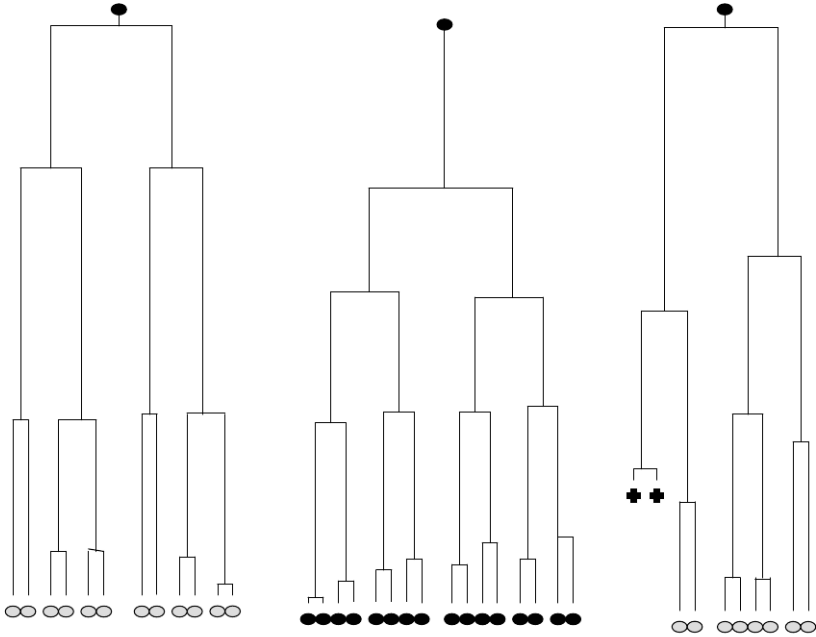


Fig. 2. Experimental stem cell lineage tree where light grey cells are positive definite, black cells are negative definite, the cross shape cells are dead cells

obtained from a purified population of $SSEA3^{Negative}$ NTERA2 stem cells. The pluripotent embryonal carcinoma (EC) cell line NTERA2 represent a human embryonic stem cells (hES) substitute [1]. The EC cells were subjected to time-lapse microscopy for seventy-two hours. After this time, the cells were sacrificed and examined by immunofluorescent labelling for SSEA3 expression. Cell division in relation to time was obtained from the time-lapse images and annotated in the form of lineage trees.

The outcome of the time-lapse experiment consists of a data set of 30 stem cell lineage trees in which the cells' expression of SSEA3 can only be observed at the leaf and the root levels as can be seen in Figure 2. In the experimental lineage trees used here the stem cells can be either $SSEA3^{Positive}$, $SSEA3^{Negative}$ or dead.

We estimated the hidden Markov tree model parameters using the variational-Bayesian approach with smoothed probabilities. The VBS-HMT model is applied to incomplete stem cells lineage tree data. The experimental data used in this study consists of 30 lineage trees, where just the type of cells at the start and at the end of each tree is known. The model developed here confronts the challenge of stem cell lineage tree reconstruction by determining the most likely state tree corresponding to the observed stem cell lineage tree. Using the proposed model we predicted the presence or absence of SSEA3 expression at the unobserved positions within the trees see Figure 3.

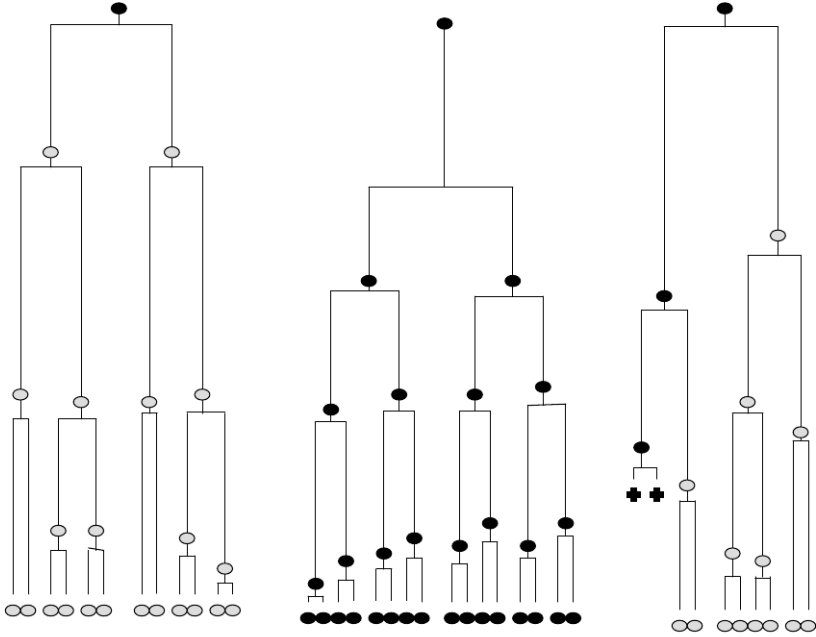


Fig. 3. Diagram representing complete stem cell lineage trees predicted by VBS-HMT model. The light grey cells are $SSEA3^{Positive}$ cells, black cells are $SSEA3^{Negative}$ cells, the cross shape cells are dead cells.

In several lineage trees our model predicted that $SSEA3^{Negative}$ cells have $SSEA3^{Positive}$ progeny. This suggests that NTERA2 stem cells could regain the SSEA3 expression i.e. the transition from $SSEA3^{Negative}$ to $SSEA3^{Positive}$ is possible. Our conclusion has been validated by the real stem cell experiment in which a percentage of the root cells which were $SSEA3^{Negative}$ stem cells produced only $SSEA3^{Positive}$ progeny.

6 Conclusion

In this paper we developed the variational Bayesian expectation maximisation with smoothed probabilities for hidden Markov trees model (VBS-HMT) and applied it to incomplete tree structured data. The model proved to be superior to the Maximum Likelihood approach and to the classical variational Bayesian method when tested on the prediction of the type of cells at each division level within a lineage tree as well as on the estimation of model parameters. We succeeded in confronting the underflow problems by combining the variational Bayesian method with the upwards-downwards algorithm with smoothed probabilities as an expectation step in the EM context. The resulting algorithm was demonstrated to have superior performance over the competing approaches and

was applied to the real stem cells lineage modelling problem. The VBS-HMT model provides the means to objectively predict a cell's phenotype from knowing the phenotype of the cells at the root and leaf level within the cell lineage tree. It is important to note that the proposed inference algorithm is able to predict novel behaviours based on incomplete data, which are not directly observable. These predictions can subsequently be validated by targeted experiments

Acknowledgments. The authors acknowledge that this work was supported by the Engineering and Physical Sciences Research Council (EPSRC).

References

1. Andrews, P.W.: Retinoic acid induces neuronal differentiation of a cloned human embryonal carcinoma cell line in vitro. *Dev. Biol.* 103, 285–293 (1984)
2. Attias, H.: A variational Bayesian framework for graphical models. In: *Advances in Neural Information Processing Systems*, vol. 12, pp. 209–215 (2000)
3. Baum, L.E., Petrie, T., Soules, G., Weiss, N.: A maximization technique occurring in the statistical analysis of probabilistic functions of Markov chains. *The Annals of Mathematical Statistics* 41(1), 164–171 (1970)
4. Beal, M., Ghahramani, Z.: The variational Bayesian EM algorithm for incomplete data: with application to scoring graphical model structures. *Bayesian Statistics 7*, 453–464 (2003)
5. Beerenwinkel, N., Drton, M.: A mutagenetic tree hidden Markov model for longitudinal clonal HIV sequence data. *Biostat.* 8(1), 53–71 (2007)
6. Bharadwaj, P., Carin, L.: Infrared-image classification using hidden Markov trees. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 24(10), 1394–1398 (2002)
7. Bulla, J., Bulla, I.: Stylized facts of financial time series and hidden semi-Markov models. *Computational Statistics Data Annals* 51(4), 2192–2209 (2006)
8. Choi, H., Baraniuk, R.G.: Multiscale image segmentation using wavelet-domain hidden Markov models. *IEEE Transactions on Image Processing* 10, 1309–1321 (2001)
9. Crouse, M., Nowak, R., Baraniuk, R.: Wavelet-based statistical signal processing using hidden Markov models. *IEEE Transactions on Signal Processing* (1997)
10. Dasgupta, N., Carin, L.: Texture analysis with variational hidden Markov trees. *IEEE Transactions on Signal Processing* 54(6), 2353–2356 (2006)
11. Devijver, P.A.: Baum's forward-backward algorithm revisited. *Pattern recognition Letters* 3, 369–373 (1985)
12. Diligenti, M., Frasconi, P., Gori, M.: Hidden Markov tree models for document image classification. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 25(4), 519–523 (2003)
13. Draper, J.S., Pigott, C., Thomson, J.A., Andrews, P.W.: Surface antigens of human embryonic stem cells: changes upon differentiation in culture. *Journal of Anatomy* 200, 249–258 (2002)
14. Durand, J.-B., Goncalves, P., Guedon, Y.: Computational methods for hidden Markov tree models—an application to wavelet trees. *IEEE Transactions on Signal Processing* 52(9), 2551–2560 (2004)

15. Durand, J.-B., Guédon, Y., Caraglio, Y., Costes, E.: Analysis of the plant architecture via tree-structured statistical models: The hidden Markov tree models. *New Phytologist* 166, 813–825 (2005)
16. Ephraim, Y., Merhav, N.: Hidden Markov processes. *IEEE Transaction on Information Theory* 48, 1518–1569 (2002)
17. Fredkin, D.R., Rice, J.A.: Fast evaluation of the likelihood of an HMM: Ion channel currents with filtering and colored noise. *IEEE Transactions on Signal Processing* 49, 625–633 (1997)
18. Frumkin, D., Wasserstrom, A., Kaplan, S., Feige, U., Shapiro, E.: Genomic variability within an organism exposes its cell lineage tree. *PLoS Computational Biology* 1, 382–394 (2005)
19. Ji, S., Krishnapuram, B., Carin, L.: Variational Bayes for continuous hidden Markov models and its application to active learning. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 28(4), 522–532 (2006)
20. Kullback, S., Leibler, R.A.: On information and sufficiency. *The Annals of Mathematical Statistics* 22(1), 79–86 (1951)
21. Lee, D.-S.: Substitution deciphering based on HMMs with applications to compressed document processing. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 24(12), 1661–1666 (2002)
22. Levinson, S.E., Rabiner, L.R., Sondhi, M.M.: An introduction to the application of the theory of probabilistic functions of a Markov process in automatic speech recognition. *Bell System Technology J.* 62, 1035–1074 (1983)
23. Mackay, D.J.C.: Ensemble learning for hidden Markov models (1997)
24. Romberg, J.K., Choi, H., Baraniuk, R.G.: Bayesian tree-structured image modeling using wavelet-domain hidden Markov models. *IEEE Transactions on Image Processing* 10, 1056–1068 (2001)
25. Ronen, O., Rohlicek, J., Ostendorf, M.: Parameter estimation of dependence tree models using the EM algorithm. *IEEE Signal Processing Letters* 2(8), 157–159 (1995)
26. Schliep, A., Costa, I.G., Steinhoff, C., Schnhuth, A.: Analyzing gene expression time-courses. *IEEE/ACM Transactions on Computational Biology and Bioinformatics* 2(3), 179–193 (2005)
27. Watanabe, S., Minami, Y., Nakamura, A., Ueda, N.: Variational Bayesian estimation and clustering for speech recognition. *IEEE Transactions on Speech Audio Process* 12, 365–381 (2004)