Strategic Explorative Policies for Hidden Markov Model Learning

[1]Suneel P, [2]Dr Renu Jain
[1]Research Scholar, UIET, CSJMU, Kanpur.
[2]Professor, Department of Computer Science Engineering, UIET, CSJMU, Kanpur.

ABSTRACT: *Many computational problems arises in the areas of signal processing, pattern recognition, timetabling, digital data network design, network monitoring, bandwidth prediction and graph searching, which create a large search space. Most of the time, the search space is exponential in size and our objective is to reduce the search time in exponentially sized space. In searches where it is required to traverse the entire search space exhaustively, the time complexity is exponential, However, by exploiting certain structural properties of search space in some of the problems, it may be possible to have an efficient search. In this work, we have tried to explore different randomized search algorithms to reduce the search space of HMM learning problem. But often we are unable to find good structural properties in search spaces and reduce the search space. May of the above problems modeled as HMM learning problem and objective of HMM learning problem is to find a finite state automaton which gives best optimal value for the observed output.*

1. INTRODUCTION:

An HMM learning problem is considered difficult to solve since it requires creating best models for real applications by optimally adapting model parameters to observed training data. There is no lmown way to analytically solve the HMM learning problem. The popular algorithm for solving the learning problem is the Baum-Welch algorithm which is based on the expectation-maximization paradigm. I-However, the algorithm suffers from two disadvantages *via.*, only local maxima is found in general, and there is no control on running time. This algorithm iteratively updates the current best value for P (O| \), but there is no guarantee on the rate of improvement. The algorithm stops when it *reaches* a model which has no better neighboring model, that is, when a local maximum is reached.

Up to this point, different sign preparing models and calculations have been utilized in organic succession investigation, among which the concealed Markov models (HMMs) have been particularly well known. Gee are notable for their viability in displaying the connections between's nearby images, areas, or occasions, and they have been widely utilized in different fields, particularly in discourse acknowledgment [1] and computerized correspondence. Thinking about the exceptional accomplishment of HMMs in designing, it is nothing unexpected that a wide scope of issues in natural succession examination have additionally profited by them. For instance, HMMs and their variations have been utilized in quality forecast [2], pairwise and numerous arrangement [3, 4], base-calling , displaying DNA sequencing mistakes [6], protein auxiliary structure expectation [7], ncRNA recognizable proof [8], RNA basic arrangement [9], quickening of RNA collapsing and arrangement quick noncoding RNA explanation and numerous others.

A concealed Markov model is a measurable model that can be utilized to depict the advancement of noticeable occasions that rely upon interior variables, which are not legitimately detectable. We consider the watched occasion an 'image' and the undetectable factor fundamental the perception a 'state'. A HMM comprises of two stochastic cycles, to be specific, an imperceptible cycle of shrouded states and a noticeable cycle of perceptible images. The shrouded states structure a Markov chain, and the likelihood conveyance of the watched image relies upon the basic state. Thus, a HMM is additionally called a doubly-installed stochastic cycle [1].

Demonstrating perceptions in these two layers, one obvious and the other imperceptible, is exceptionally helpful, since numerous true issues manage characterizing crude perceptions into various classifications, or class marks, that are more significant to us. For instance, let us consider the discourse acknowledgment issue, for which HMMs have been broadly utilized for quite a few years [1]. In discourse acknowledgment, we are keen on foreseeing the articulated word from a recorded discourse signal. For this reason, the discourse recognizer attempts to discover the arrangement of phonemes (expresses) that offered ascend to the genuine articulated sound (perceptions). Since there can be a huge variety in the genuine elocution, the first phonemes (and at last, the expressed word) can't be straightforwardly watched, and should be anticipated.

2. LITERATURE SURVEY:

2.1 L. R. Bahal et al. [2] described a method for estimating the parameter of HMM (HMM learning) for speech recognition. Parameter values was chosen to maximize the Mutual Information [MMI] between the acousbc observation sequence and the corresponding word sequence. The results were found more accurate as compared to Maximum Likelihood [ML] estimation.

2.2 Jie Yang et al. [3] used HMM learning to allow a robot to learn human skills in certain tasks and to improve motions performance and the model was used by tele-robotics. They formulated the learning problem as a multi dimensional HMM and developed a test bed for a variety of skill learning applications. Based on “the most likely performance” criterion, the best action sequence was selected from all previously measured action data by modeling the skill as an HMM.

2.3 Kwong et al. [4] employed Genetic algorithm (GA) for optimization of HMM topology ln this work they applied a new training method based on GA and Baum-Welch algorithms to obtain an HMM model with optimized number of states in the HMM model and its parameters. It found more optimized results as compared to HMM trained by Baum-Welch Method.

Vasiliki L. et al. [61 '• troduced a new machine learning framework is introduced in this paper, based on the hidden Markov model (HMM), designed to provide scheduling in dynamic wireless push systems. A novel learning framework was presented in this paper for dynamic wireless push systems. The scheme applied involved hidden Markov models (HMMs) and aimed at supporting the system with accurate predictions as regards the selection of the most desirable client demanded, in order to operate efficiently in dynamic environments.

Raleigh Smith [7] and Baldi and Biunak [21) reviewed the use of HMMs in bioinformatics with application to gene finding in human DNA. It defines three main groups of problems in computational biology for which HMMs have been proven especially useful. First, HMMs can be used for multiple alignments of DNA sequences, which is a difficult task to perform using a naive dynamic programming approach. Second, the structure of trained HMMs can uncover patterns in biological data. Such patterns have been used to discover periodicities within specific regions of the data and to help predict regions in sequences prone to forming specific structures. Tlurd is the large set of classification problems. HMM based approaches have been applied to structure prediction the problem of classifying each nucleotide according to which structure it belongs. HMMs have also been used in protein profiling to discriminate between different protein families and predict a new protein's family or subfamily. HMM-based approaches have also been successful wiien applied to the problem of gene finding in DNA. This is the problem of classifying DNA bases according to which, type ofjob they perform during transcription.

HMM learning has been applied for prediction and analysis of stock market by Hassan and Baikunth Nath [19] and also by Behrooz et a1. [17].

In this work, we have modeled HMM learning as a discrete optimization problem and discrete optimization problem is solved using randomized search techniques. Randomized search techniques [RST] [22] find solution in polynomial time which may or may not be optimal.

A random search algorithm [22] refers to an algorithm that uses some kind of randomness or probability (typically in the form of a pseudo- random number generator) in the definition of the method, and in the literature, may be called a Monte Carlo method or a stochastic algorithm. Search algorithms, in general consists of systematically walking through the search space of possible solutions until an acceptable solution is found.Randomized search algorithms are those local search algorithms where at least at one step decisions are made on the basis of some probability measure while following a systematic work. The term meta-heuristic is also commonly associated with random search algorithms. RST fall in a class of algorithms which have characteristics such as uniqueness, less complicacy, easy to design, fast and produce the solution which is optimal or close to optimal and give better results. The understanding of execution of these algorithms and their capabilities and limitations lead to the design of better heuristics. The main advantage of usin8 the se algorithms is to solve a problem which works very well in the specific directions of experiments. In addition to this, search space is also reduced. A disadvantage of these methods is that they are currently customized to each specific problem largely through trial and error. A common experience is that random search algorithms perform well and are “robust" in the sense that they give useful information quickly for ill-smictured global optimization problems.

Rennera and Ekait [24] have discussed basics of Genetic Algorithm (GA) in computer aided design, and several advanced genetic algorithms that have proved to be efficient in solving difficult design problems. Evolutionary methods such as genetic algorithms have increasingly been applied in engineering in the past decade. Basically, genetic algorithms have been considered as tools for optimization and parameter training in engineering. Genetic algorithms constitute a class of search methods especially suited for solving complex optimization problems.

Neumann and Wegener [25] applied evolutionary algorithms to problems whose structure is not well understood, as well as to the problems in combinatorial optimization.

Andres Diaz et al. [27] applied a Tabu search approach for solving a difficult forest harvesting machine location problem.

Numerical results indicated that the heuristic approach was very attractive and led to better solubons than those provided by state-of-the-art integer programming codes in limited computation times.

3. RANDOMIZED SEARCH TECHNIQUES:

We discuss different randomized search techniques in details. In general, randomized algorithms are those algorithms which make use of a random number generator and some decisions are taken on the basis of these numbers. However, some search algorithms make use of probabilistic measure based on some objective function and these algorithms are called randomized search algorithms. These are mainly used for solving opamization problem

In computer science, local search [221 is a meta-heuristic method for solving computationally hard optimization problems. Local search is used for those problems which can be formulated as finding a solu0on maximizing a criterion among a number of candidate solutions. Local search algorithms move horn solution to solution in the space of candidate solutions (search space) by applying local changes, until a soluoon deemed optimal is found or a time bound is elapsed. Local search algorithms are widely applied to numerous bard computational problems, including problems film computer science (particularly artificial intelligence), mathematics, operations research, engineering, and bioinformatics.

4. OPTIMIZATION

Optimization is very important in Science, Industry and many other walks of life [1]. Train scheduling, timetabling, network design, shape optimization, VLSI design, etc. are all examples of optimization problems. An Optimization problems P can be stated as a set of three factors (S, C, F) where S is the search space defined over finite set of n variables x„ i 1 „.n. These variables may have discrete, continuous or mixed (discrete and continuous) domains and corresponding problems of optimization are called discrete, continuous or mixed optimization. C is the set of constraints among the variables used in the problem. F is the set of objective functions that requires solution.

The goal of the objective function is to find a solution s belonging to S such that HS) >= f (s) (in case we want to maximize the objecti e function). f(S) <— f (s) (in case we want to minimize the objective function).

4.1 Modeling HMM Learning Problem as a Discrete Optimization Problem:

We have modeled the HMM learning problem as a discrete optimization problem and then solved it using randomized search algorithms. We have implemented algorithms Metropolis

(MA) and Simulated Annealing (SA) and taking a data set for a casino example. \Ve tested it for a number of observation sequences and compared the results obtained from SA and NtA.

Ln a;j denotes the probability of moving from the i* state to the j“ state, Rd bij d0notes the probability of emitting the j° symbol while at the i° state. For each '. lmw . , aq} denotes the transition probabilities from state i, which is a distribution on the set of states. Such a distribution can be discretized by assuming that each aj is a multiple of l/k, for some fixed k. clearly the distribution (p,, ...,J } can then be modeled as a distribution of k indistinguishable balls into n distinct cells. As there are “’Ct ways of distributing k balls in n cells, we have therefore, under the discretization assumed, 'Ct different transition probability distributions for the state i.

For discretizing emission probabilities, let us assume t b;j is a multiple of l/l, for some fixed non-zero positive integer 1. -ollowing reasoning similar to the above, we see that there can be, “‘*'C, different

probability distributions Z that can act as the symbol emission probability distribution for a given state. Here, m is the number of symbols

As there are n states, the n transition probability distributions can be chosen in ('Ct)° ways. Also, the n syiribol emission probability distributions can be chosen then in (^{1} *^{I}C,)° ways.

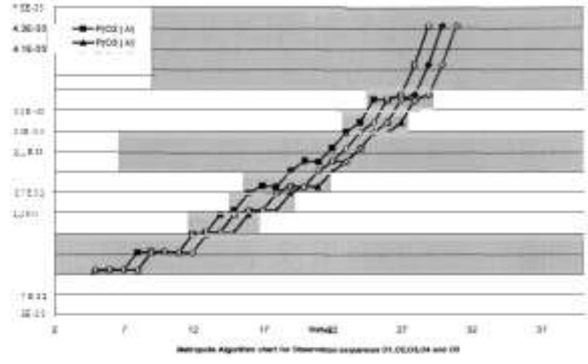
5. IMPLEMENTATION AND RESULTS:

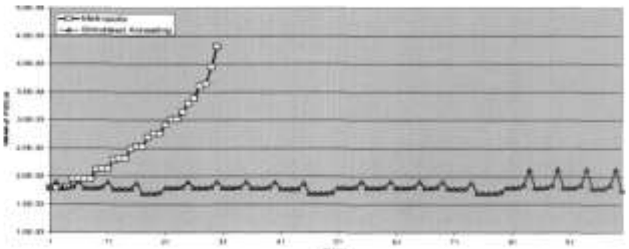
To implement Metropolis algorithm, for above casino problem, a point in the graph is characterized by a particular HMM model \, = (A , B,, Al)

i.e. (A , B t, H ,) and its value is defined by P (O | /i)for the gi› en observation sequence 0. Hence we can define a graph consisting of’ different point k k; kt and values at these points is defined by P(O | k) at any point i.

We start with as initial state \,, of \ of the system and make a small change to get to a new state \b Let ea and eb denote the values of P (O | \,), P (O | kb) respectively. If eb * ea then we update the current state to Otherwise we calculate de = (ea — eb) and update the current state to @ with probability p = *e' "* "'"” A random number between 0.5 and I is generated and if it is greater than p, we attempt the new state otherwise

leave it.





Metropolis algorithm graph for observation sequences O1,O2,O3,O4,and O5

6. CONCLUSION:

By exploiting certain structural properties of search space in some of the problems, it may be possible to have an efficient search. In this work, we have tried to explore different randomized search algorithms to reduce the search space of HMM learning problem. But often we are unable to find good structural properties in search spaces and reduce the search space. May of the above problems modeled as HMM learning problem and objective of HMM learning problem is to find a finite state automaton.

REFERENCES:

1. Guillermo L., C. Blum, E. Alba, "Evolutionary Computation" 24.

2.Lalit R. Bahal, Peter F. Brown, Peter V. de Souza, and Robert L. Mercer, "Maximum Mutual Information Estimation of Hidden Markov Model Parameters for Speech Recognition".

3.lie Yang Yangsheng Xu, and Chiou .$. Chen, "Hidden Marko›' Model Approach to Skill Learning and Its Application to Telerobotics", IEEE transactions on robotics and automation, vol. 10, no. 5, october 1994 62.

4.S. Kwong, C.W. Chau, K. Man, and K. Tang, "Optimisation of HMM topology and its model parameters by genetic al gorithms", presented at Pattern Recognition, 2001 , pp.509-522.

5. C. H. Papadimitriou, A.A. Schaffer and M. Yannakakis, "On the Complexity of Local Search", Proceedings of the twenty-second annual ACM symposium on Theory of computing Baltimore, Maryland, 1990, 438 - 445.

6.Vasiliki L. Kakali, Panaeiotis G. Sarigiannidis, Georgios I Papadimitriou , Andreas Pomportsis, "A novel HMM-based learning framework for impro›'ing dynamic w'ireless push system performance, Computers and Mathematics with Applications.

7. Kaleigh Smith, "HMM in Bioinformatics with application to gene finding in Human DNA", 308-761, Machine learning project (17- jan-2002).

8. F. Neumann and C. Witt, "Runtime analysis of a Simple Ant Colony Optimization Algorithms", In: Proc. of ISAAC 2006, LNCS 4288, Springer, Berlin, Germany, Pages 618- 627.

9. Ravindra Nath, and Renu Jain",Using Randomized Search Algorithms to Estimate HMM Learning Parameters" IEEE International Advanced computing Conference (IACC-2009).

10. Wegener, "Randomized Search Heuristics as an Alternative to Exact Optimization, Technical report", University of Dortmund. Dept of the Computer Science, February 2004.

11. Felipe Antonio Chegury Viana, "Giovanni lamin Kotinda, Domingos Alves Rade, Valder Steffen Jr.", Tuning dynamic vibration absorbers by using ant colony optimization", Computers and Structures 86 (2008) 1539—1549

12. J. Zhang and S. Li, "An Ant Colony Algorithm for Minimum Unsatisfiable Cores Extraction", School of Computer Science, National University Defence Technology, National Science Foundation of China (NSFC).

13. J. Kleinberg and E. Tardos, Algorithm Design , Cornell University Spring 2004.

14. L. R. Rabiner, "A Tutorial on Hidden Markov Models and Selected Applications in Speech Recobmition", proceedings of the IEEE, Vol. 77, No. 2, Feb 1989.

15. Y. Lifshits, S. Mozes, O. Weimann, M. Z. ukelson "Speeding Up HMM decoding and Training Exploiting sequence repetitions", Algorithimica (2009) 54. 379-399.

16. R. Durbin, R. Eddy and A. Krogh, Graeme Mite hison, "Biological Sequence Analysis", Cambridge University Press 2005.

17. Behrooz Nobakht, Dippel, Babak Loni," Stock Market Analysis and Prediction using Hidden Markov Models "

18. Daniel Hsu, Sham M. Kakade, Tong Zhang, " A Spectral Algorithm for Learning Hidden Markov Models".

19. Md. Rafiul Hassan and Baikunth Nath,"StockMarket Forecasnng Using Hidden Markov Model: A New Approach", Proceedings of the 2005 5th International Conference on Intelligent Systems Design and Applications (ISDA'05).

20. D. Sudholt, "On the Analysis of the Memetic Algorithm", In Maarten Keijzer Etal (Eds): Proceedings of the Genetic and Evolutionary Computation Conference (GECCO 2006), ACM Press, New York, NY, pages 493-500.