

# Application of Smooth Ergodic Hidden Markov Model in Text to Speech Systems

Armin Ghayoori, Faramarz Hendessi, and Asrar Sheikh

**Abstract**—In developing a text-to-speech system, it is well known that the accuracy of information extracted from a text is crucial to produce high quality synthesized speech. In this paper, a new scheme for converting text into its equivalent phonetic spelling is introduced and developed. This method is applicable to many applications in text to speech converting systems and has many advantages over other methods. The proposed method can also complement the other methods with a purpose of improving their performance. The proposed method is a probabilistic model and is based on Smooth Ergodic Hidden Markov Model. This model can be considered as an extension to HMM. The proposed method is applied to Persian language and its accuracy in converting text to speech phonetics is evaluated using simulations.

**Keywords**—Hidden Markov Models, text, synthesis.

## I. INTRODUCTION

THE text to speech synthesis (TTS) systems consist of the text analyzing process and converting the observation into speech waveform. In other words, association between text and speech waveform must be defined. In recent years, text to speech synthesis technologies for different languages are growing rapidly [1] [2].

During the quarter of century, probabilistic models became the mainstay in defining the association between words and sounds and are extensively used in modeling the processes of both speech perception and generation.

These models work in terms of phonetic units, which are representative and intuitive. Taken together, the set of phonetic units spans the range of sounds used to produce phonetic spellings of every word, as is found in any good dictionary.

The text to speech system model may be thought of consisting two black boxes, which associate words with sounds. The first box takes in raw text and produces strings of phonetic units, which represent the phonetic spelling of each word in the text. The second box generates speech waveform based on the outputs of the first box. The association between

each word and its phonetic spelling is most often made with artificial neural networks (ANNs) [3], [4], [5]. The linkage between sub-word units and complete word phonetic spelling can be made with the use of a stochastic finite state automaton (SFSA) that defines a distribution over the possible phonetic spellings of a word.

In the ANN framework, a restricted database exists, which is used as a training set. In this structure, by giving examples of words and their pronunciation to the network, the system will be able to find association between the words and their phonetic spelling. It is therefore possible to generalize and obtain outputs even when the ANN is presented with previously unseen words. However, in some complex languages such as Persian, ANN may cause some illegal errors, which result in degradation of word intelligibility [6], [7]. The contributions of this paper are in the application of a new probabilistic modeling framework, called Smooth Ergodic Hidden Markov Model, to the problem of text to speech systems, and in showing how it can be used to address the problem with the HMM.

Smooth Ergodic Hidden Markov Model (SEHMM) provides an ideal framework in which to formulate probabilistic model that are simultaneously expressive, precise, and compact. Because of the dynamic nature of any language, the ergodicity of the model helps a lot in adapting to new words.

In this paper, we first briefly review the theory of Hidden Markov Models and then extend it to idea of Smooth Ergodic Hidden Markov Model (SEHMM). The developed model is then applied to the Persian text to speech system.

## II. HIDDEN AND SMOOTH HIDDEN MARKOV MODEL

Hidden Markov Models (HMMs) belong to a powerful class of modeling technique that represents discrete state processes [8],[9]. The main idea behind the Hidden Markov Model is that an observation sequence  $\mathbf{O}$  generated by a system is represented as one of finite number of states. At each time step, the system makes a transition from its current state to another one, and according to state specific probability distribution emits an observation symbol. A Hidden Markov Model is defined by the number of states,  $N$ ; the number of different observation symbols,  $M$ ; the state transition probabilities,  $A = \{a_{ij}\}$ ; the state observation symbol probability distributions,  $B = \{b_j(k)\}$ ; and the initial state distribution,  $\pi = \{\pi_i\}$ . The appropriate values of  $M, N, A, B$

Manuscript received December 18, 2004.

Armin Ghayoori and Faramarz Hendessi are with the Department of Electrical and Computer Engineering, Isfahan University of Technology, Isfahan, Iran.

Asrar U. H. Sheikh is with the Department of Electrical Engineering, King Fahd University of Petroleum & Minerals, KFUPM Box 5038, Dhahran 31261, Saudi Arabia (phone: +966 3 860 1182; fax: +966 3 860 1183; e-mail: asrarhaq@kfupm.edu.sa).

The authors acknowledge the support of the King Fahd University of Petroleum & Minerals and the Isfahan University of Technology in carrying out this research.

and  $\pi$  specify an HMM which can be used as generation of observation symbols.

$$O = (o_1, o_2, \dots, o_T)$$

Every HMM can be compactly denoted as  $\lambda = (A, B, \pi)$  to indicate the complete set of model parameter.

In an Ergodic HMM, every state of the model can be reached (in a single step) from every other state Fig. 1. This model seems to be the most practical for any language, because it assumes that every letter is reachable from every other. This assumption is very close to reality, and enables the model to adapt itself to any new words that are added to the language.

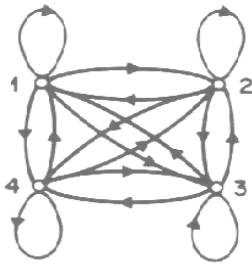


Fig. 1 A 4 States Ergodic Markov Model

In general, the following needs to be considered when applying HMMs [8]:

1. Given the model parameters  $\lambda$ , what is the probability of observing a particular sequence  $O = (o_1, o_2, \dots, o_T)$ ?
2. Given the model parameters  $\lambda$  and the observation sequence, what is the state sequence of system?
3. How to adjust the model parameters to maximize  $P(O | \lambda)$ ?

In the case of the text to speech systems, the second requirement is not relevant because the state sequence is determined by the text under process. The first requirement reduces to a simple form of finding the probability of observation sequence, given the model parameters and state sequence. For example, consider the state sequence  $Q = (q_1, q_2, \dots, q_T)$  and the model parameters  $\lambda = (A, B, \pi)$ . The probability of the corresponding observation sequence  $O = (o_1, o_2, \dots, o_T)$  is computed as follows:

$$P(O | Q, \lambda) = P(o_1, o_2, \dots, o_T | q_1, q_2, \dots, q_T, \lambda) \quad (1)$$

When a statistical independency is considered between observations, (1) reduces to:

$$P(O | Q, \lambda) = \prod_{t=1}^T P(o_t | q_t, \lambda) = b_{q_1}(o_1) b_{q_2}(o_2) \dots b_{q_T}(o_T)$$

These conditional probabilities may be stored in CPTs (Conditional Probability Tables).

The third issue is the most difficult. It does not have a known analytical solution. However, in the context of text to speech systems, this problem may be solved by counting the

number of occurrences over the training set. In this case, a training set consisting of words with their corresponding phonetic spellings is prepared. Each word in this training set is assigned a repetition coefficient, which is the indicator of the occurrence probability of the word in the language for which the system is being designed. Since there are dependencies between the corresponding phonetic units of each letter of the word and preceding and future letters, the model discussed is not properly matched to model the language. So to consider these dependencies, a new model, which is called Smooth Ergodic HMM, will be developed.

SEHMM is an Ergodic Hidden Markov Model in which the observation is conditioned on the current state as well as the adjacent events (states and/or observations). We consider two Classes of SEHMM models. In the class I model, dependency is assumed to be on the adjacent states, on which the observation is conditioned. Different number of states results in different kinds of SEHMMs. These are called  $m$ - $n$  SEHMMs,  $m$  and  $n$  are left and right adjacent states respectively. A 1-1 SEHMM is depicted in Fig. 2.

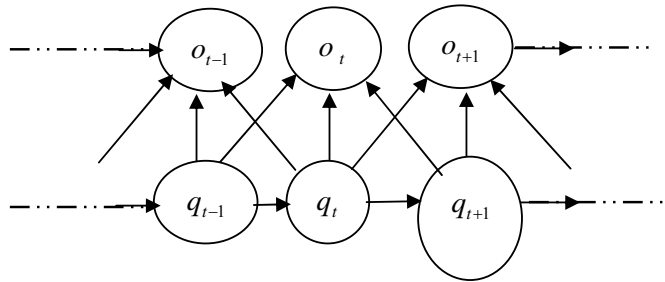


Fig. 2 A 1-1 SEHMM

The observation probabilities of this set of  $m$ - $n$  SEHMMs have the following form:

$$B = \{b_j(k)\}$$

$$b_{S_{i_0}, \dots, S_{i+n}}(k) = P[v_k \text{ at}$$

$$t | q_t = S_{i_0}, q_{t-m} = S_{i_0}, \dots, q_{t+n} = S_{i+n}],$$

$$1 \leq i_{t-m}, \dots, i_0, \dots, i_{t+n} \leq N$$

$$1 \leq k \leq M$$

For calculation of observing probabilities for a particular sequence of observation  $O = (o_1, o_2, o_3, \dots, o_T)$  for  $m$ - $n$  SEHMM, we have:

**Theorem 1:** The observation probability of the observation sequence  $O = (o_1, o_2, o_3, \dots, o_T)$  in the class I SEHMM is:

$$Q'_t = (q_{t-m}, \dots, q_t, \dots, q_{t+n})$$

$$P(O | Q, \lambda) = \prod_{t=1}^T P(o_t | Q'_t, \lambda) \quad (2)$$

where  $Q'_t$  is the set of states, which the observation of system

at time  $t$  is assumed to be dependant on,  $o_t$  is the system observation at time  $t$  and  $\lambda$  is the model parameters.

**Proof:**  $P(O | Q, \lambda) = P(o_1, o_2, \dots, o_T | Q, \lambda)$

Since the probability of appearing  $o_1, o_2, \dots, o_T$  conditioned on  $Q, \lambda$  are independent so we have

$$P(O | Q, \lambda) = P(o_1 | Q, \lambda)P(o_2 | Q, \lambda) \dots P(o_T | Q, \lambda)$$

By the assumption,  $o_1$  is only dependant on  $Q'_1$ ,  $o_2$  is only dependant on  $Q'_2, \dots$  and  $o_T$  is only dependant on  $Q'_T$ , then we have

$$P(O | Q, \lambda) = P(o_1 | Q, \lambda)P(o_2 | Q, \lambda) \dots P(o_T | Q, \lambda) = P(o_1 | Q'_1, \lambda)P(o_2 | Q'_2, \lambda) \dots P(o_T | Q'_T, \lambda) =$$

$$\prod_{t=1}^T P(o_t | Q'_t, \lambda)$$

and the theorem is proved.

In equation (2), it is apparent that the observations are supposed to be independent. Thus we have:

$$P(O | Q, \lambda) = b_{Q'_1}(o_1)b_{Q'_2}(o_2) \dots b_{Q'_T}(o_T)$$

$$b_{Q'_t}(o_t) = P(o_t | Q'_t, \lambda)$$

In the class II SEHMM, the observation probabilities are conditioned on previous observations of the system in addition to the current, left and right states of the system. We use a general name for this class of SEHMM as  $m$ - $n$ - $l$  SEHMM, where  $m$  and  $n$  denote the number of adjacent states and  $l$  being the number of previous observations of system on which the current observation is conditioned. A 1-1-1 SEHMM is depicted in Fig. 3.

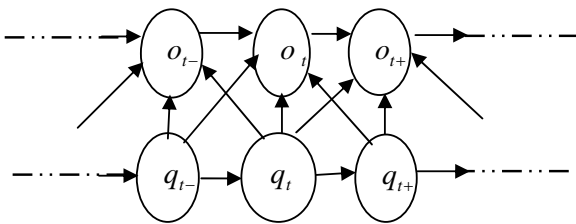


Fig. 3 A 1-1-1 SEHMM

The observation probabilities of this set of  $m$ - $n$ - $l$  SEHMMs have the following form:

$$B = \{b_j(k)\}$$

$$b_{S_{i_{t-m}, \dots, S_{i_{t+n}}, o_{t-1}, \dots, o_{t-l}}}(k) = P[v_k \text{ at}$$

$$t | q_t = S_{i_0}, q_{t-m} = S_{i_{-m}, \dots, q_{t+n} = S_{i_{t+n}}, o_{t-1}, \dots, o_{t-l}],$$

$$1 \leq i_{t-m}, \dots, i_0, \dots, i_{t+n} \leq N$$

$$1 \leq k \leq M$$

The probability that the observation sequence  $O = (o_1, o_2, o_3, \dots, o_T)$  at output of the system in the  $m$ - $n$ - $l$  SEHMM is as follows:

**Theorem 2:** The observation probability of the observation sequence  $O = (o_1, o_2, o_3, \dots, o_T)$  in the class II SEHMM is:

$$Q'_t = (q_{t-m}, \dots, q_t, \dots, q_{t+n})$$

$$O'_t = (o_{t-l}, \dots, o_{t-1})$$

$$P(O | Q, \lambda) = \prod_{t=1}^T P(o_t | O'_t, Q'_t, \lambda) \quad (3)$$

where  $Q'_t$  is the set of adjacent states, which appeared in conditional probability of the observation and  $O'_t$  is the previous observations of the system, which are effective on the current observation of the system.

**Proof:** We use mathematical induction to prove this theorem. Starting with  $T=1$

$$\text{For } T = 1: Q'_1 = (q_1, \dots, q_{1+n})$$

$$O'_1 = \text{Non}$$

$$P(O | Q, \lambda) = p(o_1 | Q_1, \lambda) = p(o_1 | O_1, Q_1, \lambda)$$

$$= \prod_{t=1}^1 P(o_t | O'_t, Q'_t, \lambda)$$

Now assume that the assertion has been proved for particular value of  $T = T'$

For  $T = T'$

$$Q'_t = (q_{t-m}, \dots, q_t, \dots, q_{t+n})$$

$$O'_t = (o_{t-l}, \dots, o_{t-1})$$

$$P(O | Q, \lambda) = \prod_{t=1}^{T'} P(o_t | O'_t, Q'_t, \lambda)$$

Now using this, we shall deduce the corresponding result for  $T = T' + 1$

For  $T = T' + 1$

$$Q'_t = (q_{t-m}, \dots, q_t, \dots, q_{t+n})$$

$$O'_{t+1} = (o_{t-l}, \dots, o_{t-1})$$

$$P(O | Q, \lambda) = P(o_1, \dots, o_{T'+1} | Q, \lambda)$$

$$= P(o_{T'+1} | Q, \lambda)P(o_1, \dots, o_{T'} | Q, \lambda, o_{T'+1})$$

Since  $(o_1, \dots, o_{T'})$  is independent of  $o_{T'+1}$  then

$$\begin{aligned}
 P(O | Q, \lambda) &= P(o_{T'+1} | Q, \lambda) P(o_1, \dots, o_{T'} | Q, \lambda) = \\
 &= P(o_{T'+1} | Q, \lambda) \prod_{t=1}^{T'} P(o_t | O'_t, Q'_t, \lambda) \\
 &= \prod_{t=1}^{T'+1} P(o_t | O'_t, Q'_t, \lambda)
 \end{aligned}$$

And the theorem is proved.

The training phase of different classes of SEHMM is based on choosing  $\lambda = (A, B, \pi)$  such that  $P(O | \lambda)$  is maximized. This criterion is achieved using an iterative procedure such as the EM (expectation-modification) method. In the case of discrete CPTs, the crux of the EM algorithm is extremely simple and obtains simply by counting. As in the first class of SEHMM the observation probabilities according to EM obey the following formula:

$$\begin{aligned}
 b_{S_{i-k}, \dots, S_{i+m}}(k) &= \frac{\text{Expected number of times in state sequence } S_{i-k}, \dots, S_{i+m} \text{ and observing symbol, } v_k}{\text{Expected number of times in state, } S_{i-k}, \dots, S_{i+m}} \\
 &= \frac{\sum_{w \in (O_t = v_k)} P_w}{\sum_w P_w}
 \end{aligned}$$

where  $w$  is a word, which has state sequence  $S_{i-k}, \dots, S_{i+m}$  and  $P_w$  is the generality of  $w$ . These generality factor can be assumed as the probability of the occurrences of the word in the text. Similar procedure can be applied to the second class of the SEHMM as follows:

$$\begin{aligned}
 b_{S_{i-k}, \dots, S_{i+m}, O_{t-1}, \dots, O_{t-l}}(k) &= \frac{\text{Expected number of times in state sequence, } S_{i-k}, \dots, S_{i+m} \text{ and observing sequence } O_{t-1}, \dots, O_{t-l}, O_t}{\text{Expected number of times in state } S_{i-k}, \dots, S_{i+m} \text{ and sequence } O_{t-1}, \dots, O_{t-l}} \\
 &= \frac{\sum_{w \in (O_{t-1}, \dots, O_{t-l}, O_t = v_k)} P_w}{\sum_w P_w}
 \end{aligned}$$

where  $w$  is a word, which has state sequence  $S_{i-k}, \dots, S_{i+m}$  and  $P_w$  is the generality of  $w$ .

Any of the estimation algorithms discussed earlier converges to the standard ML criterion that is it uses a training sequence of observations  $O$  to derive a set of model parameters  $\lambda$ , yielding  $\lambda_{ML} = \arg(\max_{\lambda} p(O | \lambda))$ . In the remaining sections of the paper, we concentrate on simulation results on the performance of the two classes of SEHMMs.

### III. APPLICATION OF SEHMM TO PERSIAN TTS

Persian is a complex language. Extraction of the correct phonetic spelling of a Persian word is a difficult process. The major complexity arises because of two kinds of vowels - short and long. Short vowels always appear as diacritic marks. Persian language has 3 short vowels and 3 long vowels. These are listed in Table I.

Short vowels are omitted in conventional Persian text and it is difficult, even for a Persian native speaker, to properly vowelize an unknown word. Long vowels are appearing along with consonants in a word. But the alphabets which are representative of long vowels (“ا”, “و” and “ی”) also appear as consonants in some other different words. For example, consider the words “رَوَد” (/rævæd/) and “بُود” (/bu:d/). In the first word, “و” is a consonant but in the second word it appears as a long vowel. So to interpret the use of long vowels is not an easy task. In this paper the task of phonetic spelling extraction is accomplished by using the proposed probabilistic model. Since 32 alphabets exist in Persian, the number of different states to be considered in our proposed model for TTS purpose is chosen to be 33 (Space is considered to be a separate state). In the next section, different kinds of SEHMM are examined to adjust model’s parameters.

TABLE I  
VOWEL DESCRIPTION OF PERSIAN LANGUAGE

Vowel	Type	Phonetic Description	Example
َ	Short Vowel	/æ/	بَار (/bær/)
ُ	Short Vowel	/u/	بُر (/bur/)
ِ	Short Vowel	/e/	بِر (/ber/)
و	Long Vowel	/u:/	بُور (/bu:r/)
ی	Long Vowel	/i:/	بیر (/bi:r/)
ا	Long Vowel	/a:/	بَار (/ba:r/)

### IV. SIMULATION RESULTS

A set of 10,000 samples is chosen for training and testing the proposed probabilistic model. Performance evaluations of different SEHMMs are accomplished in the form of diagrams. The performance criteria are based on the output error of the models and their complexity. Our goal is to achieve a model with low complexity and least probability of error.

Fig. 4 examines the error of Class I SEHMM by adapting the parameters of the model. It is evident from the results that by increasing the number of model parameters the errors decrease significantly in different Class I models. However, the performance difference between 1-2 SEHMM and 1-3 SEHMM is not significant. By increasing the number of the adjacent states in  $m-n$  SEHMM, the complexity of the model increases logarithmically. This increase is depicted in Fig. 5. It seems that the 1-2 SEHMM gives the best results and has the lowest complexity. The higher order SEHMM decreases the complexity of the system significantly, but does not result in significant improvement in the performance.

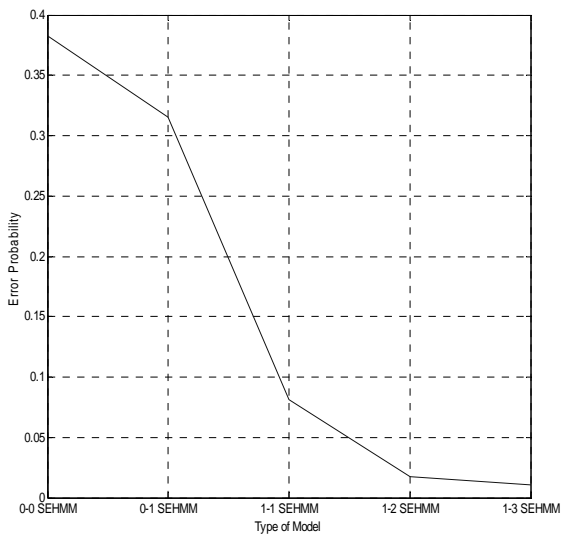


Fig. 4 Performance Evaluation of different types of m-n SEHMMs

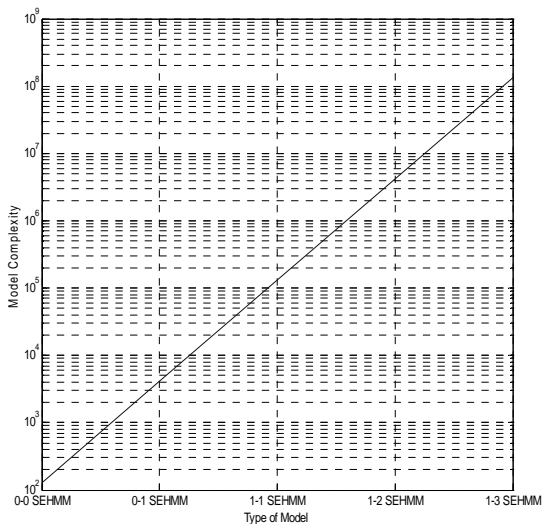


Fig. 5 Complexity Evaluation of different types of class m-n SEHMMs

There are some limitations in the Persian vowelization process and every sequence of observations cannot appear in the output of the system. This implies the idea of conditioning the current observation of the system on the previous observations as well as adjacent states. The idea is implemented in the Class II of the SEHMM. Fig. 6 illustrates the error of the different kinds of Class II SEHMM. It is seen that the error increases rapidly with an increase in the order of the SEHMM. The complexity of system is depicted in Fig. 7. It is apparent that the complexity increases linearly in the logarithmic scale as the number of adjacent states increases.

The performance difference between 1-1-1 SEHMM and 1-2-1 SEHMM illustrate that in the higher order SEHMM, there would be no major improvements in system performances.

But it is clear from Fig. 7 that the complexity is the linear function of number of adjacent states.

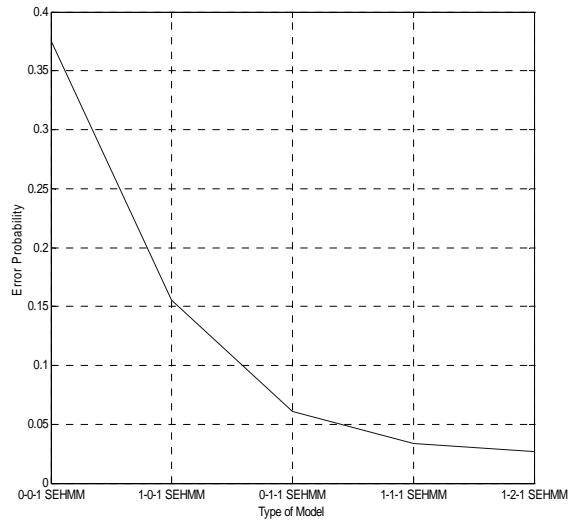


Fig. 6 Evaluation of the performance of different orders class m-n-1 SEHMMs

The other interesting points in the *m-n-l* SEHMMs pertain to the performance difference between 1-0-1 SEHMM and 0-1-1 SEHMM. As evident in Figures 6 and 7, there is major difference between the performances though the complexity of these two systems is the same. It is for the reason that in the 1-0-1 SEHMM, there is no information about the future of the system and the decision is made based on current and previous states and past observation. In the 0-1-1 SEHMM, the previous observation and future state carry information about the past and future of the system.

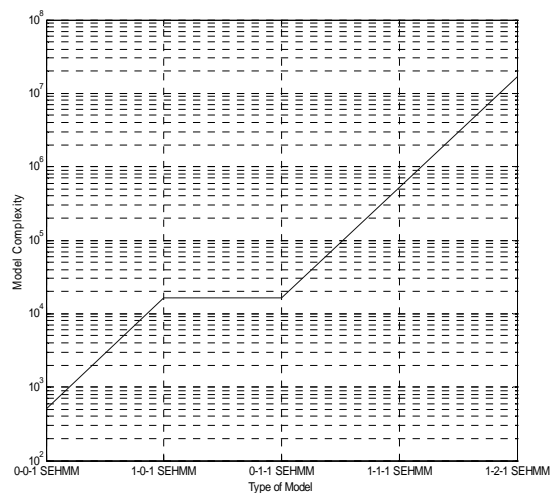


Fig. 7 Evaluation of the Complexity of different orders m-n-1 SEHMMs

The performance comparison of the two classes of SEHMM versus system complexity is illustrated in Figure 8. It is clear that in the Class II of SEHMM, higher performance can be achieved with lower complexity.

## V. PREPROCESSING

Before the words are processed by the SEHMM, some preprocessing is useful. Because of the complexity of the phonetic spelling extraction process, applying some Persian language rules can greatly increase the extraction accuracy. For example, if certain pairs of letters appear in a particular position in a word, there is no ambiguity about the phonetic spelling, so it can be determined using a rule database. This database specifies the positions and their corresponding phonetic spelling. These rules can also be incorporated into the neural network, but this increases the amount of information and rules that the neural network should learn. Thus, using these preprocessing rules, the resources required to achieve a desired accuracy are reduced.

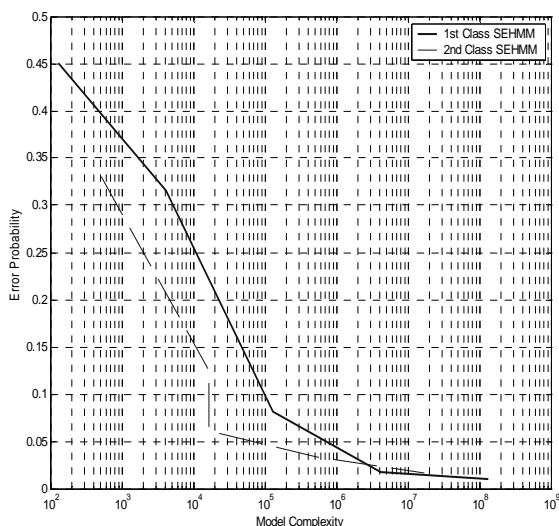


Fig. 8 The error probability of the 1<sup>st</sup> and 2<sup>nd</sup> class of SEHMMs versus system complexity

Some of the words in Arabic are widely used in Persian. Arabic is a rule based language in which the phonetic spelling extraction process obeys known rules. Thus Arabic rules are also stored in the database to easily obtain the phonetic spelling of these words.

There are some exceptions to the Persian rules, which mean there are words for which the rules cannot be applied. These words and their corresponding phonetic spelling are stored in a dictionary.

Preprocessing is also used to distinguish the suffix and/or prefix of each word in the text. There are many suffixes and prefixes in the Persian language, and they should be separated from the original word before the

word enters the neural network. These suffixes and prefixes result in many different words. Thus by learning just a main word, the neural network can correctly extract the phonetic spelling of its variations.

When a suffix and/or prefix is identified, it is separated from the word as the last stage of preprocessing. The word is now ready for being processed by SEHMM. The flow chart of this modified system is shown in Fig 9. For our training set, using preprocessing increases the performance of the 0-1-1 SEHMM from 94% to 98%. The effect of this preprocessing is dependent on text under process and the number of words in the text on which the preprocessing can be applied.

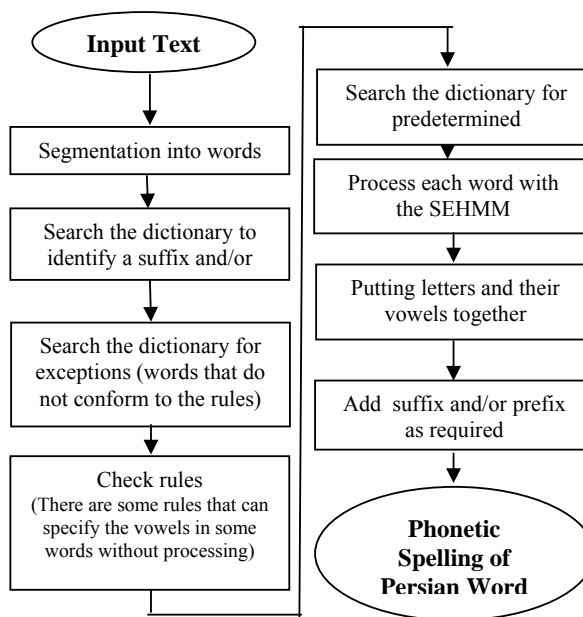


Fig. 9 A flowchart of the phonetic spelling extraction process

## VI. CONCLUSION

Probabilistic model demonstrates very high capability in modeling natural phenomena. In this paper a new model called Smooth Ergodic Hidden Markov Model is proposed for text to speech systems. This model is divided into two major classes, which are called  $m$ - $n$  SEHMM and  $m$ - $n$ - $l$  SEHMM. The system performances are shown for different values of  $m$ ,  $n$  and  $l$ . It is shown that increasing  $m$ ,  $n$  and  $l$  values improves the performance of the system rapidly for the  $m$  and  $n$ . However, considering more adjacent states/observations does not improve the performance noticeably, but the addition of the system order increases the complexity significantly. After comparing different versions of Class I and Class II of SEHMM, it seems that the 0-1-1 SEHMM is the best choice for Persian TTS systems, though for achieving better performance 1-1-1 SEHMM also can be used. In comparison

to the traditional method of mapping words into their phonetic spelling, in which the neural networks is used, the performance of 80-85% could be achieved [10]. But by using the new probabilistic model the system performance increased to 98%. This performance improvement is due to the fact that the SEHMM can be trained over the entire language.

## REFERENCES

- [1] R. Sproat, J. Hu, H. Chen, "Emu: An e-mail preprocessor for text-to-speech," Proc. IEEE Workshop on Multimedia Signal Proc., pp. 239-244, Dec. 1998.
- [2] C.-H. Wu and J.-H. Chen, "Speech activated telephony e-mail reader (SATER) based on speaker verification and text-to-speech conversion," IEEE Trans. Consumer Electronics, vol. 43, no. 3, pp. 707-716, Aug. 1997.
- [3] Sejnowski, T.J. and C.R. Rosenberg, "NETTalk: A Parallel network that learns to read aloud", Electrical Engineering and Computer Science, Technical Report JHU/EECS-86/01, Johns Hopkins University, Baltimore, 1986.
- [4] Sejnowski, T.J. and C.R. Rosenberg, "Parallel networks that learn to pronounce English text", Complex Systems, vol.1, 145-168, 1987.
- [5] *Neural Networks in Text-to-Speech Systems for the Greek Language*, 10<sup>th</sup> IEEE Mediterranean Electro-technical Conference, MELECON, pp. 574-577, May 2000.
- [6] F. Hendessi, A. Ghayoori, T. A. Gulliver, "A new text-to-speech system for Persian using a neural network and a SEHMM", Accepted for publication in the ACM Trans. Asian Lang. Proc., p.24, 2004.
- [7] F. Hendessi, A. Ghayoori, "Text-to-phoneme Conversion using Smooth Ergodic Hidden Markov Model", Proceedings of the 12<sup>th</sup> Iranian Conference on Electrical Engineering, May 2004.
- [8] L.R. Rabiner, "A tutorial on hidden Markov models and selected applications in speech recognition," Proc. IEEE, Vol. 77, No. 2, pp. 257-286, Feb. 1989.
- [9] Baum, T. Petrie, G. Soules & N. Weiss, "A maximization technique occurring in the statistical analysis of probabilistic functions of markov chains", Annuals of Mathematical Statistics pp. 41.164-171, 1970.
- [10] F. Hendessi, A. Ghayoori, "Text to Phoneme Conversion in Persian using Neural Networks", Proceedings of 9<sup>th</sup> annual of Iran computer conference, 2004.

**Armin Ghayoori** received the M. Sc. degree in electrical engineering from the Isfahan University of Technology, Isfahan, Iran, 2004. Since June 2004, he has been with the electrical and computer engineering department research center, Isfahan University of Technology, Isfahan, Iran. His general interests lie in the areas of signal processing and communications.

**Faramarz Hendessi** received his B. Sc. degree in electrical engineering from the Balochestan University, Zhaidan, Iran in 1986 and M. Sc. degree from the Isfahan University of Technology in electrical engineering in 1988. He received his Ph.D. degree in Systems and Computer Engineering from Carleton University, Ottawa, Canada in 1993. At present he is an Assistant Professor in the Electrical Engineering Department, Isfahan University of Technology and the Head IT Center and Mobile Research Group. His research interests lay in interference modeling in cellular systems.

**Asrar U. H. Sheikh** graduated from the University of Engineering and Technology, Lahore, Pakistan with first class honors and received his M.Sc. and Ph.D. degrees from the University of Birmingham, England, in 1966 and 1969 respectively. After completing teaching assignments in several countries, he returned to Birmingham as a Research Fellow in 1975. He worked at Carleton University from 1981 to 1997, first as an Associate Professor and later as a Professor and Associate Chairman for Graduate Studies. He was the Founder Director of PCS Research Laboratory at Carleton University. Before taking position of Bugshan/Bell Lab Chair in Telecommunications at King Fahd University of Petroleum and Minerals in April 2000, he was a Professor and Associate Head of the Department of Electronic and Information Engineering at Hong Kong Polytechnic University. He is also the founder director of Wireless Information Systems Research (WISR) Centre. At KFUPM he has established Telecommunications Research Laboratory.

Professor Sheikh is the author of a recently published book, *Wireless Communications – Theory & Techniques* (Orwell, Mass. USA, Kluwer Academic Publishers 2004). He has published over 230 papers in international journals and conference proceedings. He also authored or co-authored 30 technical reports. Dr. Sheikh is a co-recipient of Paul Adorian Premium from IERE (London) for his work on impulsive noise characterization. He was awarded teaching achievement awards in 1984 and 1986, and Research Achievement Award in 1994, all by Carleton University. He has organized and chaired many technical sessions at several international conferences. Dr. Sheikh is actively involved in several international conferences mainly as a member of Technical Program Committees. He Chaired the Technical Program of VTC'98. He is an editor of IEEE Transaction on Wireless Communications, a Technical Associate Editor of IEEE Communication Magazine. He is on the Editorial Board of *Wireless Personal Communications*, and *Wireless Communications and Mobile Computing*. He was a co-guest editor of the Special Issue of WPC on Interference. Dr. Sheikh is also on the reviewer panels of many IEEE and IEE Transactions and Journals. Dr. Sheikh has been consultant to many private companies and government agencies. His current interests are in signal processing in communications, mitigation of interference, spread spectrum and 3G and beyond systems. His other interests include helping developing countries in education and research. He had assignments under UNDP's sustained Development Program. He is a Fellow of the IEEE and a Fellow of the IEE. Dr. Sheikh is also listed in *Marquis Who's Whos in the world* and *Who's Who in Science and Engineering*.