IOSR Journal of Mathematics (IOSR-JM)
e-ISSN: 2278-5728, p-ISSN: 2319-765X. Volume 11, Issue 3 Ver. I (May - Jun. 2015), PP 32-41
www.iosrjournals.org

Semi - Probabilistic Automata

P Suguna
Government Arts College (Autonomous), Bharathiar University, Coimbatore, India

Abstract: *Observing the environment and recoganizing patterns for the purpose of decision making are fundamental to any scientific enquiry. Pattern recognition is a scientific discipline so much so that it enables perception in machines and also it has applications in diverse technology areas. Among the scientific community, statistical pattern recognition has received considerable attention in recent years. The statistical pattern recognition challenges are mostly approached by Hidden Markov Models (HMMs). A Hidden Markov Model (HMM) is a probabilistic mathematical discrete structure with the state emission probabilities apart from consisting the components of a probabilistic finite state automaton (PFA). Over the years, researches have been carried out to study the relations between HMM and PFA. Probabilistic finite state automata are mathematical models constructed to generate distributions over a set of strings. The computation of the probability of generating a string as a total, and a string with given prefix or suffix have important applications in the field of parsing. In this attempt, the Semi - Probabilistic Finite State Automata (Semi-PA), the most general class of Probabilistic Automata is discussed in detail.*
AMS Classification: *68Q10 and 68Q45.*
Key words: *Semi-PA, extended transition function and transition probability.*

I. Introduction

Over the years, many researchers have attempted various formal as well as stochastic models to learn, infer and identify or approximate the behavioural pattern of a system in a scientific way. Probabilistic finite state automata (PFA) are a fundamental structural model built to deal with the problem of probabalizing a structured space by adding probabilities to structure and also they are used to implement other finite state models [1].

This paper presents an overview of probabilities of strings generated by Semi – probabilistic automata (Semi-PA), the most general class of probabilistic finite state automata. The probabilistic languages and probability distributions generated by Semi-PA, in particular probabilities of strings generated by Semi-PA are discussed through extended transition functions. The challenges of parsing strings of any arbitrary length (possibly infinite) generated by semi-PA are also dealt with. The equivalence of two semi-probabilistic finite automata, the equivalence of automata with single initial state to multiple initial states and also single initial state to a designated final state are proved.

1.1 Probabilistic Languages

A nonempty set Σ of symbols is called alphabet and a finite sequence of symbols over Σ is called a word or a string. Usually words are denoted by the letters u, v, w,.... The number of symbols appearing in a word u is called its length and is denoted as $|u|$. ε is used to denote a word of zero length. If u and v are any two words, then the new word uv called concatenation of u and v is formed by adjoining the symbols in v to those symbols in u. For each word u, $\varepsilon u = u\varepsilon = u$. In the word $uv = w$, u is called prefix of w with respect to the suffix v.

For any $n \in N$, the symbol Σ^n denote the set of words of length n over the alphabet Σ. Σ^* is the set of all words of finite length including the empty word ε. Σ^∞ is the set of all words including the words of infinite length. For any word u, $u\Sigma^\infty$ denote the set of infinite words with prefix u. $\Sigma^{\leq n}$ denote the set of words of length less than or equal to n. In general, a subset L of Σ^* is defined as a language over Σ.

II. Semi - Probabilistic Automata

This section presents an overview of probabilities of strings generated by Semi – Probabilistic automata (Semi-PA).The proofs of certain theorems of Dupont [3] have not been cited in the literature. The interesting aspect of this paper is to provide few new results supported by examples.

Definition 2.1: A Semi-Probabilistic automaton (Semi-PA) is a mathematical model consists of the following components.

i) $\sum$ - a finite set of alphabets

ii) $Q = \{q_1, q_2, ... q_n\}$ - a finite set of states

iii) ϕ: $Q \times \sum \times Q \to [0,1]$ - a mapping defining the transition probability from one state to another

iv) $i : Q \to [0,1]$ - a mapping defining the initial probability for each state and

v) $\tau : Q \to [0,1]$ - a mapping defining the final probability for each state

with the conditions

$\sum_{q \in Q} i(q) = 1$, and $\forall q \in Q$, $\tau(q) + \sum_{a \in \Sigma} \sum_{q' \in Q} \phi(q, a, q') = 1$

In general, a Semi-PA is represented as $A = (\sum, Q, \phi, i, \tau)$.

Typically, a Semi-Probabilistic automaton is represented as directed graphs with labels on each edge. The initial and final probabilities of each state are given as an ordered pair, and each edge carries an alphabet and its transition probability. The following is an example of a Semi-PA.

Example 2.2: Let $A = (\sum, Q, \phi, i, \tau)$ where $Q = \{ q_0, q_1, q_2, q_3, q_4 \}$; $\sum = \{H, T\}$, ϕ is defined as $\phi(q_0, H, q_1) = 1/2$; $\phi(q_0, T, q_1) = 0$; $\phi(q_1, T, q_3) = 3/4$; $\phi(q_1, H, q_2) = 0$; $\phi(q_3, H, q_4) = 7/8$; $\phi(q_3, T, q_2) = 0$; $\phi(q_4, H, q_4) = 1/4$; $\phi(q_4, T, q_4) = 1/4$ and $\phi(q', a, q'') = 0$ for all other choices of $q', q'' \in Q$ and $a \in \sum$.

$i(q_0) = 1$; $i(q_i) = 0$ for $i = 1,2,3,4$. $\tau(q_0) = 1/2$; $\tau(q_1) = 1/4$; $\tau(q_2) = 1$; $\tau(q_3) = 1/8$; $\tau(q_4) = 1/2$.

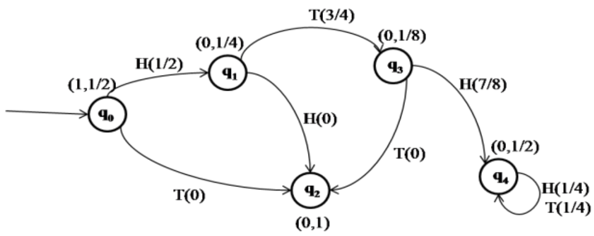

Fig. 2.1

The state for which $i(q) \neq 0$ is called initial state and the state for which $\tau(q) \neq 0$ is called final state. It is not necessary that start states and final states always be different. Also for some $q \in Q$, $i(q)$ and $\tau(q)$ both may be equal to zero. A Semi-PA may also have more than one start state and more than one final state.

The computation of probability of substrings generated by a given Semi-PA is discussed in detail in the following pages.

Definition 2.3: Let $A = (\sum, Q, \phi, i, \tau)$ be a Semi-PA. A function $\hat{\phi}$ defined from $Q \times \Sigma^* \times Q$ to [0, 1] such that

i) $\hat{\phi}(q, \varepsilon, q') = \begin{cases} 1 & \text{if } q = q' \\ 0 & \text{otherwise} \end{cases}$

ii) For $w \in \sum^*$ and $w = ua$ where a is the last symbol of w,

$\hat{\phi}(q, w, q') = \sum_{q'' \in Q} \hat{\phi}(q, u, q'') \, \phi(q'', a, q')$ is called extended transition function of A.

$\hat{\phi}(q, w, q')$ denotes the probability of reaching state q' from state q while generating the word w and $\phi(q, a, q')$ denotes the transition probability of the symbol $a \in \sum$ while reaching q' from q.

Definition 2.4: Let $A = (\sum, Q, \phi, i, \tau)$ be a Semi-PA. Then the probability of generating a word w by A is defined as $P_A(w) = \sum_{q, q' \in Q} i(q) \, \hat{\phi}(q, w, q') \, \tau(q')$.

Similarly the probability of generating a word with prefix w by A is defined as $\overline{P}_A(w) = \sum_{q,q' \in Q} i(q)$ $\hat{\phi}(q, w, q')$.

In particular $P_A(\varepsilon) = \sum_{q \in Q} i(q)\tau(q)$ and $\overline{P}_A(\varepsilon) = \sum_{q \in Q} i(q) = 1$.

Theorem 2.5: Let $A = (\Sigma, Q, \phi, i, \tau)$ be a Semi-PA. Then $\overline{P}_A(w) = P_A(w) + \overline{P}_A(w\Sigma)$.

Proof: $\overline{P}_A(w) = \sum_{q,q' \in Q} i(q)\ \hat{\phi}(q, w, q')$

$$P_A(w) = \sum_{q,q' \in Q} i(q)\ \hat{\phi}(q, w, q')\ \tau(q')$$

$$\overline{P}_A(w\Sigma) = \sum_{a \in \Sigma} \overline{P}_A(wa)$$

$$\overline{P}_A(wa) = \sum_{q,q'' \in Q} i(q)\ \hat{\phi}(q, wa, q'') = \sum_{q,q'' \in Q} i(q) [\sum_{q' \in Q} \hat{\phi}(q, w, q')\ \phi(q', a, q'')]$$

$$= \sum_{q,q',q'' \in Q} i(q)\ \hat{\phi}(q, w, q')\ \phi(q', a, q'')$$

Therefore $\overline{P}_A(w\Sigma) = \sum_{a \in \Sigma} \overline{P}_A(wa) = \sum_{a \in \Sigma} \sum_{q,q',q'' \in Q} i(q)\ \hat{\phi}(q, w, q')\ \phi(q', a, q'')$

$$P_A(w) + \overline{P}_A(w\Sigma) = \sum_{q,q' \in Q} i(q)\ \hat{\phi}(q, w, q')\ \tau(q') + \sum_{a \in \Sigma} \sum_{q,q',q'' \in Q} i(q)\ \hat{\phi}(q, w, q')\ \phi(q', a, q'')$$

$$= \sum_{q,q' \in Q} i(q)\ \hat{\phi}(q, w, q') [\ \tau(q') + \sum_{a \in \Sigma} \sum_{q'' \in Q} \phi(q', a, q'')]$$

$$= \sum_{q,q' \in Q} i(q)\ \hat{\phi}(q, w, q'),$$

$$= \overline{P}_A(w)$$

Theorem 2.6: If $A = (\Sigma, Q, \phi, i, \tau)$, is a Semi-PA then

(i) $\overline{P}_A(\Sigma^0) = 1$

(ii) $\overline{P}_A(\Sigma^k) = P_A(\Sigma^k) + \overline{P}_A(\Sigma^{k+1})$

(iii) $P_A(\Sigma^{\le n}) + \overline{P}_A(\Sigma^{n+1}) = 1$

Proof: Proof of (i) It is easy to see that $\overline{P}_A(\Sigma) = \overline{P}_A(\Sigma^1) = \sum_{a \in \Sigma} \overline{P}_A(a)$

and $\overline{P}_A(\Sigma^0) = \overline{P}_A(\varepsilon) = \sum_{q \in Q} i(q)\ \hat{\phi}(q, \varepsilon, q') = \sum_{q \in Q} i(q) = 1$

Proof of (ii) By definition, $\overline{P}_A(\Sigma^k) = \sum_{w \in \Sigma^k} \overline{P}_A(w) = \sum_{w \in \Sigma^k} \sum_{q,q' \in Q} i(q)\ \hat{\phi}(q, w, q')$

and $P_A(\Sigma^k) = \sum_{w \in \Sigma^k} P_A(w) = \sum_{w \in \Sigma^k} \sum_{q,q' \in Q} i(q)\ \hat{\phi}(q, w, q')\ \tau(q')$

Also, $\overline{P}_A(\Sigma^{k+1}) = \overline{P}_A(\Sigma^k\ \Sigma) = \sum_{wa \in \Sigma^{k+1}} \overline{P}_A(wa)$

$$= \sum_{wa \in \Sigma^{k+1}} \sum_{q,q'' \in Q} i(q)\ \hat{\phi}(q, wa, q'')$$

$$= \sum_{wa \in \Sigma^{k+1}} \sum_{q,q',q'' \in Q} i(q)\ \hat{\phi}(q, w, q')\ \phi(q', a, q'')$$

Therefore

$$P_A(\Sigma^k)+\overline{P}_A(\Sigma^{k+1}) = \sum_{w\in\Sigma^k}\sum_{q,q'\in Q} i(q)\,\hat{\phi}(q,w,q')\,\tau(q') + \sum_{wa\in\Sigma^{k+1}}\sum_{q,q',q''\in Q} i(q)\,\hat{\phi}(q,w,q')\,\phi(q',a,q'')$$

$$= \sum_{w\in\Sigma^k}\sum_{q,q'\in Q} i(q)\,\hat{\phi}(q,w,q')\,[\,\tau(q')+\sum_{a\in\Sigma}\sum_{q''\in Q}\phi(q',a,q'')\,]$$

$$= \sum_{w\in\Sigma^k}\sum_{q,q'\in Q} i(q)\,\hat{\phi}(q,w,q'),$$

$$= \overline{P}_A(\Sigma^k)$$

Proof of (iii) For any integer k, $\overline{P}_A(\Sigma^k)=P_A(\Sigma^k)+\overline{P}_A(\Sigma^{k+1})$

Suppose if k varies from 0 to n then,

$\overline{P}_A(\Sigma^0)=P_A(\Sigma^0)+\overline{P}_A(\Sigma^1)$; $\overline{P}_A(\Sigma^1)=P_A(\Sigma^1)+\overline{P}_A(\Sigma^2)$; …………………..

$\overline{P}_A(\Sigma^{n-1})=P_A(\Sigma^{n-1})+\overline{P}_A(\Sigma^n)$; $\overline{P}_A(\Sigma^n)=P_A(\Sigma^n)+\overline{P}_A(\Sigma^{n+1})$

Adding these equalities, it can be found that

$\overline{P}_A(\Sigma^0)+\overline{P}_A(\Sigma^1)+\ldots\ldots\ldots\ldots+\overline{P}_A(\Sigma^{n-1})+\overline{P}_A(\Sigma^n)$

$= P_A(\Sigma^0)+\overline{P}_A(\Sigma^1)+P_A(\Sigma^1)+\overline{P}_A(\Sigma^2)+\ldots\ldots\ldots\ldots+P_A(\Sigma^{n-1})+\overline{P}_A(\Sigma^n)+P_A(\Sigma^n)+\overline{P}_A(\Sigma^{n+1})$

Therefore $\overline{P}_A(\Sigma^0)=P_A(\Sigma^{\le n})+\overline{P}_A(\Sigma^{n+1})$

As $\overline{P}_A(\Sigma^0)$ = 1, it is easy to see that $P_A(\Sigma^{\le n})+\overline{P}_A(\Sigma^{n+1})=1$

There are varieties of finite-state models have been discussed in the literature to generate probability distributions on the strings over an alphabet. Many of such models aim at predicting the next symbol in the string, thereby describing probability distributions over each $\Sigma^n, n>0$. The main objective of these attempts is to know the amount of information one needs from the prefix to compute the next state probability. In the following theorems, we discuss the probability of strings over an alphabet with a given prefix.

Lemma 2.7: Let $A=(\Sigma,Q,\phi,i,\tau)$ be a Semi-PA. Then for any $u\in\Sigma^*$, $\overline{P}_A(u\Sigma)=\overline{P}_A(u)$.

Theorem 2.8: Let $A=(\Sigma,Q,\phi,i,\tau)$ be a Semi-PA. Then for any $u\in\Sigma^*$ and for any integer k, $\overline{P}_A(u\Sigma^k)=\overline{P}_A(u)$.

Definition 2.9: Let $A=(\Sigma,Q,\phi,i,\tau)$ be a Semi-PA with a single initial state $q\in Q$. Then $i(q)$ = 1 and $i(q')=0$ for $q'\in Q-q$.

It follows from the above definition that if $A=(\Sigma,Q,\phi,i,\tau)$ is a Semi-PA such that $i(q)$ = 1 for some $q\in Q$, then for any word $w\in\Sigma^*$,

$$\overline{P}_A(w)=\sum_{q'\in Q}\hat{\phi}(q,w,q') \text{ and } P_A(w)=\sum_{q'\in Q}\hat{\phi}(q,w,q')\tau(q')$$

In a deterministic Semi-PA $\forall a\in\Sigma$ there exist at most one path such that $\phi(q,a,q')>0$, whose underlying graph is one which has only one start state. For a deterministic Semi-PA it is possible to define a transition function ϕ on $Q\times\Sigma\times Q$ to [0,1] in a natural way as $\phi(q,a,q')$ such that $\phi(q,a,q')>0$.

Definition 2.10: Let $A=(\Sigma,Q,\phi,i,\tau)$ be a Semi-PA. A is called deterministic Semi-PA if

(i) there exist a $q\in Q$ such that $i(q)$ = 1

(ii) $\forall a\in\Sigma$ there is at most one pair (q,q') of states such that $\phi(q,a,q')>0$

The transition function ϕ is defined on $Q\times\Sigma\times Q\to[0,1]$, such that $i(q_0)=1, q_0\in Q$ and for some $q'\in Q$, $\phi(q_0,a,Q)=\phi(q_0,a,q')$.

Result 2.11: Let $A = (\Sigma, Q, \phi, i, \tau)$ be a deterministic Semi-PA then for any word $w \in \Sigma^*$, $P_A(w) = \hat{\phi}(q, w, q')\ \tau(q')$ and $\overline{P}_A(w) = \hat{\phi}(q, w, q')$.

Theorem 2.12: A Semi-PA $A = (\Sigma, Q, \phi, i, \tau)$ defines a semi-distribution over Σ^*.

Theorem 2.13: Let $A = (\Sigma, Q, \phi, i, \tau)$ be a Semi-PA. Also let $w = aub$ where $a, b \in \Sigma$ and $u \in \Sigma^*$. Then $\hat{\phi}(q, aw', q') = \hat{\phi}(q, w''b, q')$, where $w' = ub$ and $w'' = au$.

Definition 2.14: Let $A = (\Sigma, Q, \phi, i, \tau)$ be a Semi-PA. The function ϕ is defined on $Q \times \Sigma \times Q$ to [0,1], then

(i) $\phi(q, a, Q) = \sum_{q' \in Q} \phi(q, a, q')$ is the transition probability of the symbol $a \in \Sigma$

(ii) $\phi(q, \Sigma, Q) = \sum_{a \in \Sigma} \sum_{q' \in Q} \phi(q, a, q')$ is the transition probability of generating all symbols $a \in \Sigma$ while reaching all possible states $q' \in Q$ from q.

Result 2.15: Let $A = (\Sigma, Q, \phi, i, \tau)$ be a Semi-PA, then $\phi(q, \Sigma, Q) \le 1$.

Definition 2.16: Let $A = (\Sigma, Q, \phi, i, \tau)$ be a Semi-PA. The function defined on $Q \times \Sigma^* \times Q$ to [0, 1]such that $\hat{\phi}(q, \Sigma^n, Q) = \sum_{w \in \Sigma^n} \sum_{q' \in Q} \hat{\phi}(q, w, q')$, $q \in Q$ is called the probability of generating a word w of length n, by the Semi-PA A, while reaching the state q' from q.

Result 2.17: Let $A = (\Sigma, Q, \phi, i, \tau)$ be a Semi-PA. For any integer 'n' and any states $q, q'' \in Q$ $\hat{\phi}(q, \Sigma^n, Q) = \sum_{q'' \in Q} \phi(q, \Sigma, q'') \hat{\phi}(q'', \Sigma^{n-1}, Q)$

Definition 2.18: Two Semi-PA are equivalent if they define the same semi-distribution.

Theorem 2.19: Any Semi-PA is equivalent to a Semi-PA with a single initial state.

Theorem 2.20: If A and A' are equivalent, then $\overline{P}_A(w) = \overline{P}_{A'}(w)$.

The following example highlights the theorem 2.19.

Example 2.21: Consider the following Semi-PA $A = (\Sigma, Q, \phi, i, \tau)$

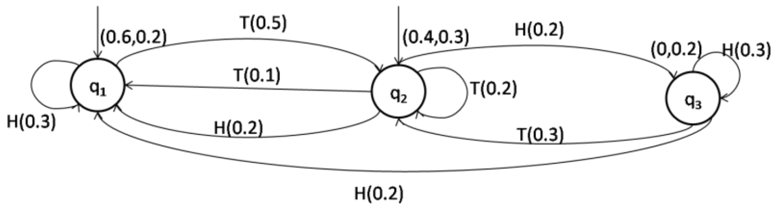

Fig. 2.2

where $Q = \{q_1, q_2, q_3\}$; $\Sigma = \{H, T\}$ and ϕ is defined such that

$\phi(q_1, H, q_1) = 0.3$; $\phi(q_1, T, q_2) = 0.5$; $\phi(q_2, T, q_2) = 0.2$; $\phi(q_2, H, q_1) = 0.2$; $\phi(q_2, T, q_1) = 0.1$; $\phi(q_2, H, q_3) = 0.2$; $\phi(q_3, H, q_1) = 0.2$; $\phi(q_2, H, q_1) = 0.2$; $\phi(q_3, H, q_3) = 0.3$ and $\phi(q', a, q'') = 0$ for all other choices of $q', q'' \in Q$ and $a \in \Sigma$. In this example both q_1 and q_2 are two initial states with $i(q_1) = 0.6$ and $i(q_2) = 0.4$.

Also q_1, q_2 and q_3 are final states with $\tau(q_1) = 0.2$ $\tau(q_2) = 0.3$, $\tau(q_3) = 0.2$.

The equivalent Semi-PA $A' = (\Sigma, Q', \phi', i', \tau')$ is constructed as given in the following diagram Fig 2.3.

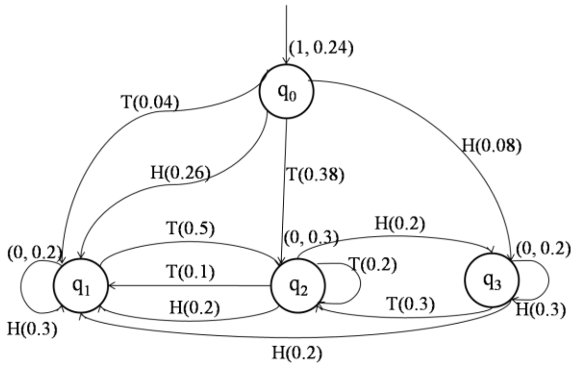

Fig. 2.3

The number of transitions from the initial state q_0 of A' to a state $q' \in Q$ is decided based on the transitions from a start state of A to q'.

The transition function ϕ' of equivalent Semi-PA A' is defined as follows:

$\phi'(q_0, H, q_1) = \sum_{q \in Q} i(q)\phi(q, H, q_1)$ =0.26; Similarly $\phi'(q_0, T, q_1) = \sum_{q \in Q} i(q)\phi(q, T, q_1)$ = 0.04

Also $\phi'(q_0, T, q_2) = \sum_{q \in Q} i(q)\phi(q, T, q_2)$ = 0.38; $\phi'(q_0, H, q_3) = \sum_{q \in Q} i(q)\phi(q, H, q_3)$ =0.08

Thus $\tau'(q_0) = \sum_{q' \in Q} i(q')\tau(q')$ = 0.24

The three conditions to prove the Theorem 2.19 can be verified for a newly constructed Semi-PA A'. They are as follows.

(i) $\tau'(q_0) + \sum_{a \in \Sigma} \sum_{q' \in \Sigma} \phi'(q_0, a, q')$

$= \tau'(q_0) + \phi'(q_0, T, q_1) + \phi'(q_0, H, q_1) + \phi'(q_0, T, q_2) + \phi'(q_0, H, q_3)$

= 0.24 + 0.04 + 0.26 + 0.38 + 0.08 =1.

(ii) $P_A(\varepsilon)$ = $\sum_{q \in Q} i(q)\tau(q)$ = 0.24= $\tau'(q_0)$ = $P_{A'}(\varepsilon)$

(iii)To show $P_A(w) = P_{A'}(w)$

Let $w = HTHT$ where $a = H$ and $u = THT$,

Then, $P_A(au)$ = $\sum_{q, q' \in Q} i(q)\hat{\phi}(q, au, q')\tau(q')$ = $\sum_{q, q' \in Q} i(q)\hat{\phi}(q, HTHT, q')\tau(q')$ = 7.392×10^{-3}

$P_{A'}(au)$ = $\sum_{q' \in Q} i(q_0)\hat{\phi}(q_0, au, q')\tau(q')$ = 7.392×10^{-3}

Therefore $P_A(au)$ = $P_{A'}(au)$

Hence the Semi-PA $A' = (\Sigma, Q', \phi', i', \tau')$ is equivalent to the Semi-PA $A = (\Sigma, Q, \phi, i, \tau)$.

Note that the converse of the above theorem 2.19 is also true. Before claiming that an equivalent Semi-PA with multiple initial states can be constructed from a given Semi-PA with single initial state, the method of calculating initial probabilities in such cases is provided.

Consider the Semi-PA $A' = (\Sigma, Q', \phi', i', \tau')$ with $i(q_0)$ =1 and $i(q_1)$ = $i(q_2)$ = $i(q_3)$ = 0.

Consider the following Semi-PA represented in Fig 2.4.

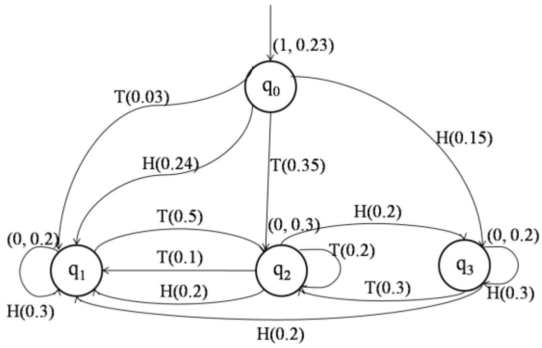

Fig.2.4

It is known that $\phi'(q,a,q') = \sum_{q''\in Q} i(q'')\phi(q'',a,q')$

$$\phi'(q_0,H,q_1) = \sum_{q\in Q} i(q)\phi(q,H,q_1) = i(q_1)\phi(q_1,H,q_1) + i(q_2)\phi(q_2,H,q_1) + i(q_3)\phi(q_3,H,q_1)$$

(2.1)

$$\phi'(q_0,T,q_1) = \sum_{q\in Q} i(q)\phi(q,T,q_1) = i(q_2)\phi(q_2,T,q_1)$$

(2.2)

$$\phi'(q_0,T,q_2) = \sum_{q\in Q} i(q)\phi(q,T,q_2) = i(q_1)\phi(q_1,T,q_2) + i(q_2)\phi(q_2,T,q_2) + i(q_3)\phi(q_3,T,q_2)$$

(2.3)

$$\phi'(q_0,H,q_3) = \sum_{q\in Q} i(q)\phi(q,H,q_3) = i(q_2)\phi(q_2,H,q_3) + i(q_3)\phi(q_3,H,q_3)$$

(2.4)

From equation (2.2), it is observed that 0.03 = $i(q_2)$ (0.01) $\therefore$ $i(q_2)$ = 0.3

Also from (2.4), we have 0.15 = (0.3× 0.2) + $i(q_3)$ (0.3)$\therefore$ $i(q_3)$ = 0.3

From (2.1), we have 0.24 = $i(q_1)$ (0.3) + (0.3× 0.2) + (0.3× 0.2)$\therefore$ $i(q_1)$ = 0.4

Finally we get $\sum_{q\in Q} i(q) = i(q_1) + i(q_2) + i(q_3)$ = 0.4 + 0.3 + 0.3 = 1.

Hence the equivalent Semi-PA $A = (\Sigma, Q, \phi, i, \tau)$ with multiple initial states for the above Semi-PA $A' = (\Sigma, Q', \phi', i', \tau')$ with a single initial state is as follows:

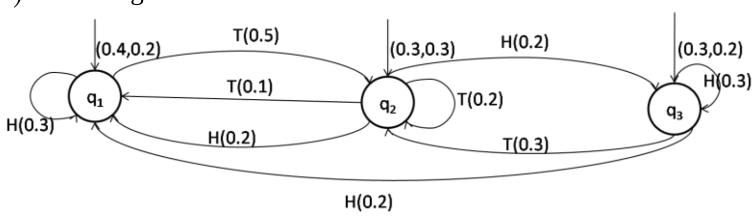

Fig. 2.5

Theorem 2.22: A Semi-PA with a single initial state is equivalent to any Semi-PA with multiple initial states.

Proof: Let $A' = (\Sigma, Q', \phi', q_0, \tau')$ be a Semi-PA with single initial state q_0.

Let $A = (\Sigma, Q, \phi, i, \tau)$ be defined as follows.

i. $\forall a \in \Sigma$ $\phi(q,a,q') = \phi'(q,a,q')$ if $q, q' \in Q$

ii. $\forall q' \in Q$ $\sum_{q\in Q} i(q)\phi(q,a,q') = \phi'(q_0,a,q')$, $a \in \Sigma$

iii. $\tau(q) = \tau'(q)$ if $q \in Q$

To show A is a Semi-PA, we have to prove

(i) $\tau(q)+\sum_{a\in\Sigma}\sum_{q'\in Q}\phi(q,a,q')$ = 1

(ii) $\sum_{q\in Q} i(q)=1$

Proof of (i)

Consider $\tau(q)+\sum_{a\in\Sigma}\sum_{q'\in Q}\phi(q,a,q')$ = $\tau(q')+\sum_{a\in\Sigma}\sum_{q'\in Q}\phi'(q,a,q')$ = 1

Proof of (ii)

Using the equation $\phi'(q,a,q')$ = $\sum_{q''\in Q} i(q'')\phi(q'',a,q')$

As the transition probability $\phi'(q,a,q')$ and $\phi(q'',a,q')$ are known, we arrive to simultaneous equations satisfying the condition; number of equations greater than or equal to the number of unknowns. By solving these equations $\sum_{q\in Q} i(q)=1$ can be easily proved.

Therefore A is a Semi-PA.

Next to prove A' and A are equivalent Semi-PA's, we have to show

(iii) $P_{A'}(w)=P_A(w)$, where $w=au$ and $a\in\Sigma$

(iv) $P_{A'}(\varepsilon)=P_A(\varepsilon)$

Proof of (iii), Now $P_{A'}(w)=P_{A'}(au)$

$$= \sum_{q'\in Q}\hat{\phi}'(q_0,au,q')\tau'(q') \quad = \sum_{q'\in Q} i'(q_0)\hat{\phi}'(q_0,au,q')\tau'(q')$$

$$= \sum_{q'',q'\in Q} i'(q_0)\phi'(q_0,a,q'')\hat{\phi}'(q'',u,q')\tau'(q')$$

$$= \sum_{q,q'',q'\in Q} i(q)\phi(q,a,q'')\hat{\phi}(q'',u,q')\tau(q')$$

$$= P_A(au) \quad = P_A(w)$$

Proof of (iv), Consider $P_{A'}(\varepsilon)$ = $i'(q_0)\tau'(q_0)$ = $\sum_{q\in Q} i(q)\tau(q)$ = $P_A(\varepsilon)$

Hence A' and A defines the same semi-distribution.

A Semi-PA A'' can be thought of having a designated initial state and a designated final state, with a special end-of-word symbol for reaching the final state from every state. In general it is represented as $A''=(\Sigma'',Q'',\phi'',i'',\tau'')$. As in any Semi-PA the outgoing edges of each state generate symbols drawn from Σ'', with each edge a probability value being associated such that for every state the probabilities of all outgoing edges sum up to one.

Theorem 2.23: A Semi-PA with single initial state is equivalent to a Semi-PA with single initial state and designated final state provided one considers a special end-of-word symbol for reaching final state from every state.

Proof: Let $A'=(\Sigma,Q',\phi',i',\tau')$ be a Semi-PA with single initial state q_0.

Let $A''=(\Sigma'',Q'',\phi'',i'',\tau'')$ with single initial state q_0 and single final state q_f be defined as follows:

i. $Q''=Q'\cup\{q_f\}$

ii. $\Sigma''=\Sigma\cup\xi$, where ξ is the special final symbol such that every state reaches the final state generating ξ.

iii. $\forall a\in\Sigma''$, $\phi''(q,a,q')$ = $\left\{\begin{array}{ll} \phi'(q,a,q'), & \text{if } q,q'\in Q \\ i'(q)\phi'(q,a,q') & \text{if } q=q_0 \\ 0 & \text{if } q=q_f \end{array}\right\}$

iv. $\phi''(q,\xi,q_f) = \tau'(q)$

v. $i''(q) = \begin{cases} 1 & \text{if } q = q_0 \\ 0 & \text{otherwise} \end{cases}$

vi. $\tau''(q) = \begin{cases} 1 & \text{if } q = q_f \\ 0 & \text{otherwise} \end{cases}$

To show A'' is a Semi-PA with single initial state q_0 and a designated final state q_f, the following conditions should hold. $\tau''(q)+\sum_{a\in\Sigma''}\sum_{q'\in Q''}\phi''(q,a,q') = 1$ as $Q'' = Q' \cup \{q_f\}$ and $\Sigma'' = \Sigma \cup \xi$.

Consider $\tau''(q)+\sum_{a\in\Sigma''}\sum_{q'\in Q''}\phi''(q,a,q')$

case (i) if $q = q_f$ then $\tau''(q_f)+\sum_{a\in\Sigma''}\sum_{q'\in Q''}\phi''(q_f,a,q') = 1$

case (ii) if $q \neq q_f$ then $\tau''(q)+\sum_{a\in\Sigma''}\sum_{q'\in Q''}\phi''(q,a,q') = 0 + \sum_{a\in\Sigma\cup\xi}\sum_{q'\in Q''}\phi''(q,a,q')$

$$= \sum_{a\in\Sigma}\sum_{q'\in Q'}\phi''(q,a,q')+\phi''(q',\xi,q_f) = \sum_{a\in\Sigma}\sum_{q'\in Q}\phi'(q,a,q')+\tau'(q) = 1.$$

case (iii) if $q = q_0$ then $\tau''(q)+\sum_{a\in\Sigma''}\sum_{q'\in Q''}\phi''(q,a,q') = \tau''(q_0)+\sum_{a\in\Sigma''}\sum_{q'\in Q''}\phi''(q_0,a,q')$

$$= \tau''(q_0)+\sum_{a\in\Sigma''}\sum_{q'\in Q''}i'(q_0)\phi'(q_0,a,q') = 0 + \sum_{a\in\Sigma\cup\xi}\sum_{q'\in Q'\cup\{q_f\}}i'(q_0)\phi'(q_0,a,q')$$

$$= \sum_{a\in\Sigma}\sum_{q'\in Q'}i'(q_0)\phi'(q_0,a,q')+\phi'(q_0,\xi,q_f) = \sum_{a\in\Sigma}\sum_{q'\in Q'}\phi'(q_0,a,q')+\tau'(q_0)$$

$= 1.$

Hence A'' represents a Semi-PA with single initial state q_0 and a designated final state q_f .

To prove A' and A'' are equivalent, we need to prove

(i) $P_{A'}(\varepsilon) = P_{A''}(\varepsilon)$, where ε is an empty word.

(ii) $P_{A'}(au) = P_{A''}(au)$, for any word u and any symbol a .

Proof of (i) Consider $P_{A'}(\varepsilon) = i'(q_0)\tau'(q_0) = 1\times \phi''(q_0,\xi,q_f) = P_{A''}(\varepsilon)$

Proof of (ii)Next, let $w\in\Sigma^*$ such that $w = au$ where $a\in\Sigma$ and $u\in\Sigma^*$.

Then $P_{A'}(w) = P_{A'}(au) = \sum_{q,q'\in Q'} i'(q_0)\hat{\phi}'(q_0,au,q')\tau'(q')$

$$= \sum_{q',q''\in Q'} i'(q_0)\phi'(q_0,a,q'')\hat{\phi}'(q'',u,q')\tau'(q')$$

$$= \sum_{q',q''\in Q'} \phi'(q_0,a,q'')\hat{\phi}'(q'',u,q')\tau'(q')$$

$$= \sum_{q',q''\in Q'\cup\{q_f\}} \phi''(q_0,a,q'')\hat{\phi}''(q'',u,q')\phi''(q',\xi,q_f)$$

$$= \sum_{q',q''\in Q''} i''(q_0)\phi''(q_0,a,q'')\hat{\phi}''(q'',u,q')\phi''(q',\xi,q_f)$$

$= P_{A''}(au)$

$= P_{A''}(w)$

Hence A' and A'' defines the same semi-distribution and therefore they are equivalent Semi-PA.

Example 2.24: Consider a Semi-PA A' with single initial state.

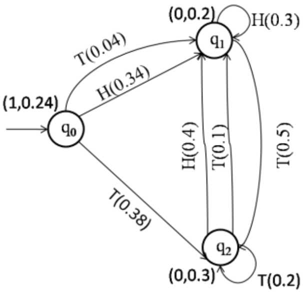

Fig. 2.6

The following is an equivalent Semi-PA A'' with single initial state and designated final state.

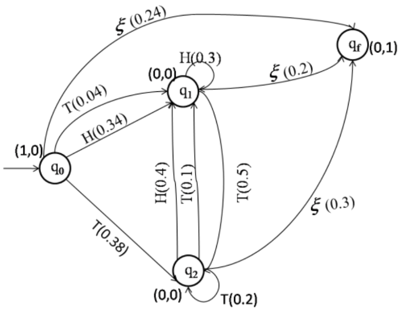

Fig.2.7

III. Conclusion

There are various types of finite state machines proposed in the literature to generate or model probability distribution on the stings over an alphabet. Some models assume probabilities on states, others on transitions. Semi-PA is one such finite state machine assuming probabilities on its transitions, and based not on an acyclic automaton, so that it enables us to define distributions over the set of strings of infinite length.

A number of results centered on the issues related to parsing strings generated by Semi-Probabilistic Automata in terms of extended transition function are provided. One can study the relationship between these models and probabilistic automata. Also, the important tasks of parsing the strings of any arbitrary length (possibly infinite) generated by Semi-Probabilistic Automata are to be explored. This work can also be extended to other probabilistic models like hidden markov models, finite state tree automata.

References

[1]. Paz. (1971). Introduction to Probabilistic Automata, New York: Academic Press.

[2]. L. Rabiner, A tutorial on hidden markov models and selected applications in speech recognition, Proc. IEEE. vol. 77, (1989) pp. 257-286.

[3]. P. Dupont, F. Denis, and Y. Esposito, Links between probabilistic automata and hidden markov models: probability distributions, learning models and induction algorithms, Pattern Recognition, vol. 38, (2004) pp. 1349- 1371.

[4]. R. Durbin, S. R. Eddy, A. Krogh and G. Mitchison. (1998). Biological sequence analysis: probabilistic models of proteins and nucleic acids, Cambridge, UK: Cambridge University Press.

[5]. J.E. Hopcroft and J.D. Ullman. (1969). Formal Languages and their relation to Automata, Addison-Wesley Publishing Company.

[6]. E. Vidal, F. Thollard, C. de la Higuera F. Casacuberta, and R.C. Carrasco, Probabilistic finite state automata – part I, IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 27, No.7, (2005) pp. 1013-1025.

[7]. Wetherell C.S, Probabilistic languages: A Review and Some Open Questions, Computing Surveys, vol.12, No.4, (1980) pp. 361 - 379.

[8]. Yasubumi Sakakibara, Grammatical inference in bioinformatics, IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 27, No.7, July (2005) pp.1051-1062.