Computation of the Entropy of Fuzzy Hidden Markov Chain

R. Sujatha
Department of Mathematics
SSN College of Engineering
Chennai, India
sujathar@ssn.edu.in

T. M. Rajalaxmi
Department of Mathematics
SSN College of Engineering
Chennai, India
laxmi.raj18@gmail.com

Abstract
Modified Viterbi algorithm of FHMC [10] is the method for tracking the hidden states of a process from a sequence of given observation sequence. An important problem while tracking a process with an FHMC is estimating the uncertainty present in the solution. To overcome this kind of uncertainty we need to compute the entropy of a state sequence. The entropy of a possibilistic variable provides a measure of its uncertainty. In this correspondence we proposed an algorithm for computing the entropy of the most likelihood state sequence obtained from the modified Viterbi algorithm in addition to that this entropy is given in triangular fuzzy number on [0, 1].

Mathematics Subject Classification: 60J10, 03E72, 94A17

Keywords: Triangular fuzzy number, Possibility Space, Conditional possibility, Bayes’s possibility forumla, Fuzzy Markov Chain, Fuzzy Hidden Markov Chain

1 Introduction
Hidden Markov Model (HMM) [9] is a doubly stochastic process with an underlying stochastic process that is hidden which follows the Markov property, but
can only be observed through another set of stochastic processes that produce the sequence of observed symbols in such a way that the HMM constitutes of a initial probability, the transition probability and the output symbol observation probability. Real world problems for example prediction of web navigation using web log files are uncertain in nature. To this kind of imprecision data we cannot assign a certain probability values hence in this case classical HMM is less appropriate while fuzzy sets a theory in which everything has elasticity in addition to that the theory of uncertainty has been developed on the basis of fuzzy sets is the possibility theory. Possibility space is used to model the incomplete information in a flexible way. Operations involved in this space are minimum and maximum. Existence theorem on possibility space [4] says that possibilistic variable determines a normalized fuzzy set. Fuzzy variable is a possibilistic variable taking values in $\mathcal{R} = (-\infty, \infty)$ in this situation point possibility values are equal to the degree of membership values [4]. In this paper, we considered the situation that there are uncertainty in the possibility measure. But, type 2 fuzzy set is used to model the uncertainty in the membership values of the elements of the ordinary fuzzy set. Hence, in this paper, possibility values are converted into the type 2 fuzzy set whose grade of memberships are triangular fuzzy numbers on $[0,1]$ has been considered. A fuzzy Markov chain on the possibility space has the finite number of states and a possibilistic variables is state valued stochastic process whose possibility measure is $\sigma$ such that the chain satisfies the Markov property.

In this paper, we have introduced an algorithm to compute the entropy for the most likelihood state sequence obtained from modified Viterbi algorithm on possibility space. The entropy of a possibilistic variable provides a measure of its uncertainty. For a possibilistic variable $X$, the information conveyed by observing an outcome $x$ is $-\log_2 \sigma(X = x)$ bits. If the logarithm is base 2, the units of information are in bits. The entropy of the state sequence that explains an given observation sequence, given a model can be viewed as the minimum number of bits that on average will be needed to encode the state sequence (given the model and the observations) [5]. The higher this entropy the higher the uncertainty involved in tracking the hidden process with current model.

In Section 2 we have discussed the preliminaries and FHMC, algorithm for efficiently computing the entropy has explained in Section 3, in section 4 illustration has presented and finally concluded.

2 Preliminary Notes

Fuzzy sets, as its name implies, basically, a theory of graded concepts. Let $\Gamma$ be the universe of discourse whose generic element is denoted by $\omega$. 
Definition 2.1. A fuzzy set \([4]\) \(\tilde{A}\) defined on \(\Gamma\) is a mapping from \(\Gamma\) to the unit interval \([0, 1]\), \(\mu_{\tilde{A}}(\omega)\) is referred to as the membership function whose value at \(\omega\) signifies the grade of membership of \(\omega\) of the fuzzy set \(\tilde{A}\) and may vary from 0 to 1. \(\tilde{B}\) is said to be a type 2 fuzzy set \([4]\) if \(\forall \omega \in \Gamma, \mu_{\tilde{B}}(\omega)\) is a fuzzy set in itself defined on \([0, 1]\).

Definition 2.2. A normalized convex fuzzy set \(\tilde{A}\) on \(\Gamma\) whose membership function \(\mu_{\tilde{A}}\) is piecewise continuous is called the fuzzy number. \([4]\)

Definition 2.3. A triangular fuzzy number (TFN) \([4]\) \(\tilde{A} = (a_1, a_2, a_3)\) where \(a_1 < a_2 < a_3\) is a special type of fuzzy number and it satisfies,

1. the membership function \(\mu_{\tilde{A}}(x) = 1\) at \(x = a_2\);
2. the graph of \(y = \mu_{\tilde{A}}(x)\) on \([a_1, a_2]\) is a straight line from \((a_1, 0)\) to \((a_2, 1)\) and also on \([a_2, a_3]\) the graph is a straight line from \((a_2, 1)\) to \((a_3, 0)\);
3. \(\mu_{\tilde{A}}(x) = 0\) for \(x \leq a_1\) or \(x \geq a_3\).

The logarithm of the TFN \([2]\) has given by

\[
\log_2(a_1, a_2, a_3) = (\log_2 a_1, \log_2 a_2, \log_2 a_3)
\]

Definition 2.4. an \(\alpha -\) cut \([4]\) of a fuzzy set \(\tilde{A}\) denoted by \(\tilde{A}_\alpha\) is defined as

\[
\tilde{A}_\alpha = [a_1 + \alpha(a_2 - a_1), a_3 - \alpha(a_3 - a_2)] = [a^-, a^+],
\]

\(0 \leq \alpha \leq 1\). (1)

It is worth noting that contrary to what holds in set theory, \(\tilde{A} \cup \tilde{A}^c \neq \Gamma\) and \(\tilde{A} \cap \tilde{A}^c \neq \phi\) because it is not certain where \(\tilde{A}\) ends and \(\tilde{A}^c\) begins. This is the fundamental reason that places probability and fuzzy sets apart and mathematical apparatus of the theory of fuzzy sets provides a natural basis for the theory of possibility. This theory was coined by Zadeh \([11]\) in the late 1970s. Possibility theory is maxitive and not additive, i.e., the possibility of a disjunction of events is the maximum of the possibilities of each event

Definition 2.5. A possibility space \([4]\) is a triple \((\Gamma, \mathcal{F}, \sigma)\) where:

1. \(\mathcal{F}\) is a class of all subsets of \(\Gamma\), i.e., elements of \(\mathcal{F}\) represent the collection of events of interest in that experiment.

2. For every \(A \in \mathcal{F}\), the non-negative number \(\sigma(A)\) is the possibility that the event \(A\) occurs. The map \(A \rightarrow \sigma(A)\), called a possibility, if \(\sigma : \mathcal{F} \rightarrow [0, 1]\), with the following properties:

\(a) \ \sigma(\phi) = 0\) and \(\sigma(\Gamma) = 1\).
(b) For an arbitrary collection of sets \( A_i \in \mathcal{S} \), \( \sigma(\bigcup_{i \in I} A_i) = \sup_{i \in I} \sigma(A_i) \).

**Definition 2.6.** The conditional possibility [4] of \( A \in \mathcal{S} \) given \( B \) denoted by \( \sigma(A|B) \) is defined by,

\[
\sigma(A|B) = \begin{cases} 
1, & \text{if } \sigma(AB) = \sigma(B); \\
\sigma(AB), & \text{if } \sigma(AB) < \sigma(B). 
\end{cases}
\]

**Definition 2.7.** A possibilistic variable \( X \) [4] is a mapping from \( \Gamma \) to an arbitrary universe \( \mathbb{U} \) and the point possibility distribution function is given by

\[ g(x) = \sigma(X = x) \quad \forall x \in \mathbb{U}. \]

**Definition 2.8.** Possibilistic variable \( X \) is a fuzzy variable if \( X \) takes values in \( \mathbb{R} = (-\infty, \infty) \). If \( X \) is a fuzzy variable defined on possibility space we denote \( \mu_X(x) = \sigma(X = x) \quad \forall x \in \mathbb{R} \).

By the existence theorem of possibility space [4], possibilistic variable \( X \) determines a normalized fuzzy set defined on \( \mathcal{R} \).

**Definition 2.9.** Bayes Possibility formula [4] for the possibility space is defined for \( A_i \) is an arbitrary collection of sets such that \( \Gamma = \bigcup_i A_i \). Further \( B \in \mathcal{S} \). Then,

\[
\sigma(A_i|B) \quad \sigma(\bigcup_{i \in I} A_i) = \sup_{i \in I} \sigma(A_i). 
\]

**Definition 2.10.** Fuzzy Markov chain [8] on the possibility space has the finite number of states \( S = \{1, 2, \ldots, s\}; S \subset \mathcal{R} \) and a possibilistic variables \( X = \{X_n; n \in \mathbb{N}\} \) an \( S \) valued stochastic process on possibility space and whose possibility measure is \( \sigma \) such that the chain satisfies the Markov property, i.e., for all \( j \in S \) and for each time step \( n > 0 \) we have,

\[
\sigma(X_{n+1} = j|X_0, X_1, \ldots, X_n) = \sigma(X_{n+1} = j|X_n). \quad (2)
\]

**Definition 2.11.** The transition possibility \( \tilde{p}_{ij} \) of the system from state \( i \) to state \( j \) is defined as for each \( i, j \in S \)

\[
\tilde{p}_{ij} = \sigma(X_{n+1} = j|X_n = i). \quad (3)
\]

In the case if \( \tilde{p}_{ij} \) is independent of time then we can say that the chain is an homogeneous fuzzy Markov chain. Let \( \tilde{P} = (\tilde{p}_{ij}) \) is an \( s \times s \) matrix of transition possibilities.
Definition 2.12. Initial possibility vector of the system is denoted by $\tilde{p}^{(0)} = (\tilde{p}_1^{(0)}, \tilde{p}_2^{(0)}, \ldots, \tilde{p}_s^{(0)})$, where $\tilde{p}_i^{(0)} = \sigma(X_0 = i)$ is the possibility of being in the state $i$ initially.

To capture the vagueness involved in the system, TFN has been carried out to the elements of initial and transition possibility matrix of each state; Maximum TFN value among those can be done by comparison of TFN [6].

2.1 Fuzzy Hidden Markov Chain

Definition 2.13. Fuzzy Hidden Markov Chain FHMC on possibility space [10] is a bivariate discrete process $\{X_n, o_n\}_{n \geq 0}$, where $\{X_n\}$ is a fuzzy Markov chain on possibility space and $O = \{o_n\}$ is the sequence of observation such that the conditional possibility distribution of $o_n$ only depends on $X_n$.

Elements of Fuzzy Hidden Markov Chain $s$ the number of states in the chain, we have denoted the state at time step $n$ as $X_n$, $m$ the number of distinct observation symbols per state, i.e. the discrete output of the system; We have denoted the individual symbols as $V = \{v_1, v_2, \ldots, v_m\}$, the state transition possibility distribution matrix, denoted by $\tilde{P} = (\tilde{p}_{ij})$, the observation symbol possibility distribution in state $j$, denoted by $\tilde{B} = (\tilde{b}_j(k))$ where $\tilde{b}_j(k) = \tilde{\sigma}(v_k \mid X_n = j)$, the initial state possibility distribution $\tilde{p}^{(0)} = (\tilde{p}_i^{(0)})$, where $1 \leq i, j \leq s$ and $1 \leq k \leq m$.

TFN has been carried out to the elements of initial $\tilde{p}^{(0)}$, transition $\tilde{P}$ and observation possibilities $\tilde{B}$. A compact notation to indicate the complete parameter set of the model is denoted by $\tilde{\lambda} = (\tilde{P}, \tilde{B}, \tilde{p}^{(0)})$. Given an appropriate values of $s$, $m$, $\tilde{P}$, $\tilde{B}$ and $\tilde{p}^{(0)}$, the FHMC can be used as a generator to given an observation sequence $O = \{o_0, o_1, \ldots, o_{N-1}\}$ (where each observation $o_n$ is one of the symbol from $V$ at time step $n$, and $N − 1$ is the number of observations in the sequence).

Modified Viterbi Algorithm

Modified Viterbi algorithm [10] produces the most likely state sequence for the observations and the algorithm is as follows:

1. Initialization

$$\tilde{\gamma}_0(i) = \min \left[ \tilde{p}_i^{(0)}, \tilde{b}_i(o_0) \right], 1 \leq i \leq s,$$

$$\varphi_0(i) = 0.$$
2. Recursion
\[
\tilde{\gamma}_{n+1}(j) = \min \left\{ \max_{1 \leq i \leq s} [\min(\tilde{\gamma}_n(i), \tilde{p}_{ij})], \tilde{b}_j(o_{n+1}) \right\}, \\
\varphi_{n+1}(j) = \arg \max_{1 \leq i \leq s} [\min(\tilde{\gamma}_n(i), \tilde{p}_{ij})], 0 \leq n \leq N-2, 1 \leq j \leq s,
\]

3. Termination
\[
P^* = \max_{1 \leq i \leq s} [\tilde{\gamma}_{N-1}(i)], \\
X^*_N = \arg \max_{1 \leq i \leq s} [\tilde{\gamma}_{N-1}(i)].
\]

4. Path (state sequence) backtracking:
\[
X^*_n = \varphi_{n+1}(X^*_{n+1}), \quad n = N-2, N-3, \ldots, 0,
\]

The operation \(\arg \max\) has been done by [1]. Throughout this correspondence we will adopt the following notation. Subscripts will be used to identify a particular component in a sequence. Superscripts will be used to denote sequences of variables or symbols. For example, by \(X^n\) we will mean the sequence of \(n\) random variables \((X_1, X_2, \ldots, X_n)\).

3 Computation of the Entropy

In this section we presented an algorithm to compute the entropy \(\tilde{H}\) of the most likelihood state sequence obtained from the modified Viterbi algorithm on possibility space. Since we denoted the measure of entropy as triangular fuzzy number we used \(\tilde{\gamma}\).

Our proposed algorithm uses the following intermediate variables:

1. \(\tilde{H}_n(j) = \tilde{H}(X^{n-1}|X_n = j, O^n = o^n)\): entropy of all state sequences that lead to state \(j\) at time step \(n\), given the observations up to time step \(n\). For example, if there is just one possible path that leads to state \(j\) at time step \(n\), then \(\tilde{H}_n(j) = 0\).

2. \(\tilde{c}_n(j) = \sigma(X_n = j|O^n = o^n)\): possibility of being in state \(j\) at time step \(n\), given the observations up to time step \(n\).

The entropy \(\tilde{H}_n(j)\) can be computed recursively using the values from the previous step, \(\tilde{H}_{n-1}(i), 1 \leq i \leq s\). Let us assume that we are on state \(j\) at time step \(n\). Then, we can divide the path into two segments: the first contains the sequence of states up to time step \(n-2\) (let us call this possibilistic variable
as \( Y \), and the second contains just the state at time \( n - 1 \) (let us call this possibilistic variable as \( X \)) and since our aim is to find the maximum entropy associated with the state sequence we have,

\[
\tilde{H}(X, Y) = \max \left[ \tilde{H}(X), \tilde{H}(Y|X) \right]
\]  \hspace{1cm} (4)

The above expression has been done by the comparison of TFN [6]. In the equation (4), \( \tilde{H}(X) \) is the entropy associated with \( \sigma(X_{n-1} = i|X_n = j) \) (possibility that state \( i \) at time step \( n - 1 \) given state \( j \) at time step \( n \)) which we can compute easily using \( \tilde{c}_n(k), 1 \leq k \leq s \) and \( \tilde{c}_{n-1}(l), 1 \leq l \leq s \). \( \tilde{H}(Y|X) \) can be computed from \( \tilde{H}_{n-1}(k), 1 \leq k \leq s \) using

\[
\tilde{H}(Y|X) = \max_{x \in X} \left\{ \min_{i \in X} \left[ \sigma(X = x), \tilde{H}(Y|X = x) \right] \right\}
\]  \hspace{1cm} (5)

Let us analyze the recursion in detail. First, note that \( \tilde{c}_n(j) \) can be computed recursively

\[
\tilde{c}_n(j) = \frac{\max_{1 \leq i \leq s} \min_{1 \leq k \leq s} [\tilde{c}_{n-1}(i), \tilde{p}_{ij}, \tilde{b}_{ij}(o_n)]}{\max_{1 \leq k \leq s} \min_{1 \leq k \leq s} [\tilde{c}_{n-1}(i), \tilde{p}_{ik}, \tilde{b}_{ik}(o_n)]}
\]  \hspace{1cm} (6)

Division of the above expression has been carried out by the generalized division of TFN [7]. Next, for the recursion, we need the auxiliary possibilities

\[
\sigma(X_{n-1} = i|X_n = j, O^n = o^n)
\]

These can be computed using Baye’s possibility formula as follows:

\[
\sigma(X_{n-1} = i|X_n = j, O^n = o^n) = \begin{cases} 
1, & \min[\sigma(X_n = j|X_{n-1} = i), \sigma(X_{n-1} = i|O^{n-1} = o^{n-1})] = \\
\max_{1 \leq k \leq s} \min[\sigma(X_n = j|X_{n-1} = k), \sigma(X_{n-1} = k|O^{n-1} = o^{n-1})], & \text{if} \ \min[\sigma(X_n = j|X_{n-1} = i), \sigma(X_{n-1} = i|O^{n-1} = o^{n-1})] < \\
\min[\sigma(X_n = j|X_{n-1} = k), \sigma(X_{n-1} = k|O^{n-1} = o^{n-1})], & \text{if} \ \max_{1 \leq k \leq s} \min[\sigma(X_n = j|X_{n-1} = k), \sigma(X_{n-1} = k|O^{n-1} = o^{n-1})] < \\
\end{cases}
\]

i.e.,

\[
\begin{cases} 
1, & \text{if} \ \min[\tilde{p}_{ij}, \tilde{c}_{n-1}(i)] = \max_{1 \leq k \leq s} \min[\tilde{p}_{kj}, \tilde{c}_{n-1}(k)]; \\
\min[\tilde{p}_{ij}, \tilde{c}_{n-1}(i)], & \text{if} \ \min[\tilde{p}_{ij}, \tilde{c}_{n-1}(i)] < \max_{1 \leq k \leq s} \min[\tilde{p}_{kj}, \tilde{c}_{n-1}(k)].
\end{cases}
\]

The recursion on the intermediate entropies can be derived as follows:

\[
\tilde{H}_n(j) = \tilde{H}(X_{n-1}|X_n = j, O^n = o^n)
\]

\[
= \tilde{H}(X_{n-2}, X_{n-1}|X_n = j, O^n = o^n)
\]

\[
= \max \left[ \tilde{H}(X_{n-1}|X_n = j, O^n = o^n), \tilde{H}(X_{n-2}|X_{n-1}, X_n = j, O^n = o^n) \right]
\]
where \( \tilde{H}(X_{n-1}|X_n = j, O^n = o^n) \)

\[
= - \max_{1 \leq i \leq s} \left\{ \min \left[ \sigma(X_{n-1} = i|X_n = j, O^n = o^n), \log(\sigma(X_{n-1} = i|X_n = j, O^n = o^n)) \right] \right\}
\]

and \( \tilde{H}(X^{n-2}|X_{n-1}, X_n = j, O^n = o^n) \)

\[
= \max_{1 \leq i \leq s} \left\{ \min \left[ \sigma(X_{n-1} = i|X_n = j, O^n = o^n), \tilde{H}(X^{n-2}|X_{n-1} = i, X_n = j, O^n = o^n) \right] \right\}
\]

\[
= \max_{1 \leq i \leq s} \left\{ \min \left[ \sigma(X_{n-1} = i|X_n = j, O^n = o^n), \tilde{H}_{n-1}(i) \right] \right\}
\]

**Lemma 3.1.** The entropy of the state sequence up to time step \( n - 2 \) given the state at time step \( n - 1 \) and the observations up to \( n - 1 \) is conditionally independent on the state and observation at time step \( n \)

\[
\tilde{H}(X^{n-2}|X_{n-1} = i, X_n = j, O^n = o^n) = \tilde{H}_{n-1}(i)
\]

**Proof.** \( \tilde{H}(X^{n-2}|X_{n-1} = i, X_n = j, O^n = o^n) \)

\[
= \tilde{H}(X^{n-2}|X_{n-1} = i, O^{n-1} = o^{n-1}, X_n = j, O_n = o_n)
\]

\[
= \tilde{H}(X^{n-2}|X_{n-1} = i, O^{n-1} = o^{n-1})
\]

\[
= \tilde{H}_{n-1}(i)
\]

where the second step comes from the properties of FHMCs: states \( X_{n-k}, k \geq 2 \) and \( X_n \) are statistically independent given \( X_{n-1} \). The same applies to states \( X_{n-k}, k \geq 2 \) and observation \( O_n \) given \( X_{n-1} \). According to the basic properties of the entropy, \( \tilde{H}(X|Y = y) = \tilde{H}(X) \) if \( X \) and \( Y \) are independent. \( \Box \)

To finalize the algorithm, we need to compute \( \tilde{H}(X^N|O^N = o^N) \), which can be expanded using the basic properties of the entropy

\[
\tilde{H}(X^N|O^N = o^N)
\]

\[
= \max \left\{ \tilde{H}(X^{N-1}|X_N, O^N = o^N), \tilde{H}(X_N|O^N = o^N) \right\} \text{ by eq. (4)}
\]

\[
= \max \left\{ \left\{ \max_{1 \leq i \leq s} \left[ \min \left[ \sigma(X_N = i|O^N = o^N), \tilde{H}(X^{N-1}|X_N = i, O^N = o^N) \right] \right] \right\} \right\}
\]

\[
= \max \left\{ \left\{ \max_{1 \leq i \leq s} \left[ \min \left[ \sigma(X_N = i|O^N = o^N), \log(\sigma(X_N = i|O^N = o^N)) \right] \right] \right\} \right\}
\]

\[
= \max \left\{ \left\{ \max_{1 \leq i \leq s} \left[ \min \left( c_N(i), \tilde{H}_N(i) \right) \right] \right\} \right\}
\]

\[
\left\{ - \max_{1 \leq i \leq s} \left[ \min \left( c_N(i), \log(c_N(i)) \right) \right] \right\}
\]
Algorithm:

1. Initialization. For $1 \leq j \leq s$

\[
\tilde{H}_0(j) = (0, 0, 0) \\
\tilde{c}_0(j) = \min_{1 \leq i \leq s} \left[ \frac{\tilde{p}_i(0), \tilde{b}_j(o_0)}{\max_{1 \leq i \leq s} \left[ \tilde{p}_i(0), \tilde{b}_j(o_0) \right]} \right]
\]

2. Recursion. For $1 \leq j \leq s; 2 \leq n \leq N$

\[
\tilde{c}(j) = \max_{1 \leq i \leq s} \min_{1 \leq i \leq s} \left[ \sigma(X_{n-1} = i | X_n = j, O^n = o^n) \right]
\]

\[
\sigma(X_{n-1} = i | X_n = j, O^n = o^n) = \begin{cases} 
1, & \text{if } \min[\tilde{p}_{ij}, \tilde{c}_{n-1}(i)] = \max_{1 \leq k \leq s} \min_{1 \leq i \leq s} [\tilde{p}_{kj}, c_{n-1}(k)]; \\
\min[\tilde{p}_{ij}, \tilde{c}_{n-1}(i)], & \text{if } \min[\tilde{p}_{ij}, \tilde{c}_{n-1}(i)] < \max_{1 \leq k \leq s} \min_{1 \leq i \leq s} [\tilde{p}_{kj}, c_{n-1}(k)].
\end{cases}
\]

\[
\tilde{H}_n(j) = \max \left\{ \max_{1 \leq i \leq s} \left[ \frac{\sigma(X_{n-1} = i | X_n = j, O^n = o^n)}{\tilde{H}_{n-1}(i)} \right] \right\} - \max_{1 \leq i \leq s} \left\{ \min \left[ \sigma(X_{n-1} = i | X_n = j, O^n = o^n), \log(\sigma(X_{n-1} = i | X_n = j, O^n = o^n)) \right] \right\}
\]

3. Termination

\[
\tilde{H}(X^N | O^N = o^N) = \max \left\{ \max_{1 \leq i \leq s} \left[ \min \left( c_i(i), \tilde{H}(i) \right) \right] \right\} - \max_{1 \leq i \leq s} \left\{ \min \left( c_i(i), \log(c_i(i)) \right) \right\}
\]

4 Illustration

We have considered the FHMC discussed in the previous chapter. Let the state space $S$ be the following departments

\[
S = \{ EEE, MECH, BME \} = \{ E, M, B \},
\]

and each department’s attributes are namely About the Department, Faculty, News is the set of observations, i.e.,

\[
V = \{ \text{About the Department, Faculty, News} \} = \{ v_A, v_F, v_N \}.
\]
When we sketched out the data, we have experienced the sample observation sequence as

\[ O = o_0^F, o_1^A, o_2^A, o_3^F, o_4^N, o_5^A, o_6^N, o_7^F, o_8^A, o_9^N, o_{10}^N, o_{11}^N, o_{12}^N. \]

Converted TFN values of initial possibility vector, transition possibility matrix and observation possibility vector for each department are given below:

\[ \tilde{p}^{(0)} = \begin{bmatrix} 0.58, 0.65, 0.72 & 0.39, 0.48, 0.57 & 1.00, 1.00, 1.00 \end{bmatrix}, \]

<table>
<thead>
<tr>
<th></th>
<th>E</th>
<th>M</th>
<th>B</th>
</tr>
</thead>
<tbody>
<tr>
<td>E</td>
<td>(1.00, 1.00, 1.00)</td>
<td>(0.35, 0.49, 0.61)</td>
<td>(0.24, 0.35, 0.46)</td>
</tr>
<tr>
<td>M</td>
<td>(0.28, 0.39, 0.48)</td>
<td>(1.00, 1.00, 1.00)</td>
<td>(0.36, 0.5, 0.63)</td>
</tr>
<tr>
<td>B</td>
<td>(0.38, 0.52, 0.65)</td>
<td>(0.30, 0.45, 0.60)</td>
<td>(1.00, 1.00, 1.00)</td>
</tr>
</tbody>
</table>

Enhancing the modified Viterbi algorithm we obtained the state sequence which is the best to explain the given observation sequence is depicted below:

\[ X_{12}^* = B, X_{11}^* = \varphi_{12}(B) = B, \ldots, X_1^* = \varphi_2(E) = E, X_0^* = \varphi_1(E) = E. \]

Maximum likelihood path is depicted in the figure 1.

![Figure 1: Maximum likelihood path](image)

To verify that the above computed most likelihood path is free from uncertainty we need to calculate the entropy associated with this path. Using our proposed algorithm we have calculated the entropy of the path for each state at each time step. For calculation purpose we have the states EEE, MECH and BME as E, M and B respectively.

\[ \tilde{H}_0(E) = (0, 0, 0) \]

\[ \tilde{c}_0(E) = \min \left[ \tilde{p}_E^{(0)}, \tilde{b}_E(o_0 = F) \right] = (1.00, 1.00, 1.00) \]

\[ \tilde{c}_0(E) = \max \left\{ \min \left[ \tilde{p}_E^{(0)}, \tilde{b}_E(o_0 = F) \right], \min \left[ \tilde{p}_M^{(0)}, \tilde{b}_M(o_0 = F) \right], \min \left[ \tilde{p}_B^{(0)}, \tilde{b}_B(o_0 = F) \right] \right\} \]
The most likely state is obtained from modified Viterbi algorithm is given in bold. At time step 0 the state sequence as entropy value \( \tilde{H}_1(E) \) is depicted in Figure 2.

\[
\tilde{c}_1(E) = \frac{N}{D}
\]

where,

\[
N = \max \left\{ \min [\tilde{c}_0(E), \tilde{p}_{EE}, \tilde{b}_E(o_1 = A)], \min [\tilde{c}_0(M), \tilde{p}_{ME}, \tilde{b}_E(o_1 = A)], \min [\tilde{c}_0(B), \tilde{p}_{BE}, \tilde{b}_E(o_1 = A)] \right\}
\]

\[
D = \max \left\{ \max \left\{ \min [\tilde{c}_0(E), \tilde{p}_{EM}, \tilde{b}_M(o_1 = A)], \min [\tilde{c}_0(M), \tilde{p}_{MM}, \tilde{b}_M(o_1 = A)], \min [\tilde{c}_0(B), \tilde{p}_{BM}, \tilde{b}_M(o_1 = A)], \min [\tilde{c}_0(E), \tilde{p}_{EB}, \tilde{b}_B(o_1 = A)], \min [\tilde{c}_0(M), \tilde{p}_{MB}, \tilde{b}_B(o_1 = A)], \min [\tilde{c}_0(B), \tilde{p}_{BB}, \tilde{b}_B(o_1 = A)] \right\} \right\}
\]

\[
\tilde{H}_1(E) = \frac{(32, 46, .6)}{(55, 72, 88)} = (58, .63, .68)
\]

Similarly for the remaining states. The computed values are given in the table 1. One can notice in the table that the most likelihood state sequence obtained from modified Viterbi algorithm is given in bold. At time step 0 the most likelihood state is \( E \) which as entropy \( (0, 0, 0) \), but for the next time step the state sequence as entropy value \( (.79, .67, .56) \). While the system reaches the final state \( B \) the entropy associated with the path is \( (0, 0, 0) \) which shows that most likelihood path is free from uncertainty. Graph of the entropy measure associated with the most likelihood state sequence obtained from modified Viterbi algorithm is depicted in Figure 2.
<table>
<thead>
<tr>
<th>E</th>
<th>M</th>
<th>B</th>
<th>Entropy</th>
</tr>
</thead>
<tbody>
<tr>
<td>( H_0(E) = (0, 0, 0) )</td>
<td>( \bar{H}_0(M) = (0, 0, 0) )</td>
<td>( \bar{H}_0(B) = (0, 0, 0) )</td>
<td>(0, 0, 0)</td>
</tr>
<tr>
<td>( \bar{c}_0(E) = (1, 1, 1) )</td>
<td>( \bar{c}_0(M) = (1, 1, 1) )</td>
<td>( \bar{c}_0(B) = (1, 1, 1) )</td>
<td></td>
</tr>
<tr>
<td>( H_1(E) = (0, 0, 0) )</td>
<td>( \bar{H}_1(M) = (0, 0, 0) )</td>
<td>( \bar{H}_1(B) = (0, 0, 0) )</td>
<td>(.79, .67, .56)</td>
</tr>
<tr>
<td>( \bar{c}_1(E) = (.58, .63, .68) )</td>
<td>( \bar{c}_1(M) = (1, 1, 1) )</td>
<td>( \bar{c}_1(B) = (.25, .38, .45) )</td>
<td></td>
</tr>
<tr>
<td>( H_2(E) = (0, 0, 0) )</td>
<td>( \bar{H}_2(M) = (0, 0, 0) )</td>
<td>( \bar{H}_2(B) = (0, 0, 0) )</td>
<td>(1.6, 1.1, .74)</td>
</tr>
<tr>
<td>( \bar{c}_2(E) = (.32, .46, .6) )</td>
<td>( \bar{c}_2(M) = (1, 1, 1) )</td>
<td>( \bar{c}_2(B) = (.14, .27, .4) )</td>
<td></td>
</tr>
<tr>
<td>( H_3(E) = (0, 0, 0) )</td>
<td>( \bar{H}_3(M) = (0, 0, 0) )</td>
<td>( \bar{H}_3(B) = (0, 0, 0) )</td>
<td>(0, 0, 0)</td>
</tr>
<tr>
<td>( \bar{c}_3(E) = (.51, .97, 1) )</td>
<td>( \bar{c}_3(M) = (1, 1, 1) )</td>
<td>( \bar{c}_3(B) = (.47, .91, 1) )</td>
<td></td>
</tr>
<tr>
<td>( H_4(E) = (0, 0, 0) )</td>
<td>( \bar{H}_4(M) = (0, 0, 0) )</td>
<td>( \bar{H}_4(B) = (0, 0, 0) )</td>
<td>(2, 1.2, 0)</td>
</tr>
<tr>
<td>( \bar{c}_4(E) = (.25, .41, .95) )</td>
<td>( \bar{c}_4(M) = (.25, .43, 1) )</td>
<td>( \bar{c}_4(B) = (1, 1, 1) )</td>
<td></td>
</tr>
<tr>
<td>( H_5(E) = (0, 0, 0) )</td>
<td>( \bar{H}_5(M) = (0, 0, 0) )</td>
<td>( \bar{H}_5(B) = (0, 0, 0) )</td>
<td>(2.1, .76, 0)</td>
</tr>
<tr>
<td>( \bar{c}_5(E) = (1, 1, 1) )</td>
<td>( \bar{c}_5(M) = (1, 1, 1) )</td>
<td>( \bar{c}_5(B) = (.23, .59, 1) )</td>
<td></td>
</tr>
<tr>
<td>( H_6(E) = (0, 0, 0) )</td>
<td>( \bar{H}_6(M) = (0, 0, 0) )</td>
<td>( \bar{H}_6(B) = (0, 0, 0) )</td>
<td>(0, 0, 0)</td>
</tr>
<tr>
<td>( \bar{c}_6(E) = (.27, .64, 1) )</td>
<td>( \bar{c}_6(M) = (.25, .67, 1) )</td>
<td>( \bar{c}_6(B) = (1, 1, 1) )</td>
<td></td>
</tr>
<tr>
<td>( H_7(E) = (0, 0, 0) )</td>
<td>( \bar{H}_7(M) = (0, 0, 0) )</td>
<td>( \bar{H}_7(B) = (0, 0, 0) )</td>
<td>(.94, .6, .38)</td>
</tr>
<tr>
<td>( \bar{c}_7(E) = (1, 1, 1) )</td>
<td>( \bar{c}_7(M) = (.55, .72, .88) )</td>
<td>( \bar{c}_7(B) = (.52, .66, .77) )</td>
<td></td>
</tr>
<tr>
<td>( H_8(E) = (0, 0, 0) )</td>
<td>( \bar{H}_8(M) = (0, 0, 0) )</td>
<td>( \bar{H}_8(B) = (0, 0, 0) )</td>
<td>(1.7, 1.2, .84)</td>
</tr>
<tr>
<td>( \bar{c}_8(E) = (1, 1, 1) )</td>
<td>( \bar{c}_8(M) = (.32, .47, .63) )</td>
<td>( \bar{c}_8(B) = (.3, .43, .56) )</td>
<td></td>
</tr>
<tr>
<td>( H_9(E) = (0, 0, 0) )</td>
<td>( \bar{H}_9(M) = (0, 0, 0) )</td>
<td>( \bar{H}_9(B) = (0, 0, 0) )</td>
<td>(1.3, .86, .62)</td>
</tr>
<tr>
<td>( \bar{c}_9(E) = (.91, .93, .98) )</td>
<td>( \bar{c}_9(M) = (1, 1, 1) )</td>
<td>( \bar{c}_9(B) = (1, 1, 1) )</td>
<td></td>
</tr>
<tr>
<td>( H_{10}(E) = (0, 0, 0) )</td>
<td>( \bar{H}_{10}(M) = (0, 0, 0) )</td>
<td>( \bar{H}_{10}(B) = (0, 0, 0) )</td>
<td>(0, 0, 0)</td>
</tr>
<tr>
<td>( \bar{c}_{10}(E) = (.67, .69, .76) )</td>
<td>( \bar{c}_{10}(M) = (.62, .72, .86) )</td>
<td>( \bar{c}_{10}(B) = (1, 1, 1) )</td>
<td></td>
</tr>
<tr>
<td>( H_{11}(E) = (0, 0, 0) )</td>
<td>( \bar{H}_{11}(M) = (0, 0, 0) )</td>
<td>( \bar{H}_{11}(B) = (0, 0, 0) )</td>
<td>(0, 0, 0)</td>
</tr>
<tr>
<td>( \bar{c}_{11}(E) = (.27, .38, .5) )</td>
<td>( \bar{c}_{11}(M) = (.25, .4, .56) )</td>
<td>( \bar{c}_{11}(B) = (1, 1, 1) )</td>
<td></td>
</tr>
<tr>
<td>( H_{12}(E) = (0, 0, 0) )</td>
<td>( \bar{H}_{12}(M) = (0, 0, 0) )</td>
<td>( \bar{H}_{12}(B) = (0, 0, 0) )</td>
<td>(0, 0, 0)</td>
</tr>
<tr>
<td>( \bar{c}_{12}(E) = (.27, .38, .5) )</td>
<td>( \bar{c}_{12}(M) = (.25, .4, .56) )</td>
<td>( \bar{c}_{12}(B) = (1, 1, 1) )</td>
<td></td>
</tr>
</tbody>
</table>

Figure 2: Graph of the entropy values of most likelihood path.
5 Conclusion

The entropy of a possibilistic variable provides a measure of its uncertainty. Since possibility space itself is used to model the incomplete information in a flexible way which implies that the computation of the most likelihood state sequence obtained from modified Viterbi algorithm is itself gives the desired solution. But, still we need to verify is there any uncertainty associated with that state sequence. Hence we proposed the algorithm to compute the entropy associated with the state sequence and have showed with the illustration that the computed most likelihood path is free from uncertainty.

References


**Received: April, 2012**