Towards fuzzy linguistic Markov chains

Pablo J. Villacorta[1] **Jose L. Verdegay**[1] **David A. Pelta**[1]

[1]Models of Decision and Optimization (MODO) Research Group,
Dept. of Computer Science and AI, University of Granada (Spain)
Email: {pjvi, verdegay, dpelta}@decsai.ugr.es

Abstract

In this contribution we deal with the problem of doing computations with a Markov chain when the information about transition probabilities is expressed linguistically. This could be the case, for instance, if the process we are modeling is described by a human expert, for whom the use of linguistic labels is easier than being forced to give inexact numerical probabilities which, in turn, may yield an unstable chain. We address the uncertainty of linguistic judgments by introducing fuzzy probabilities, and carry on the calculation of the linguistic stationary distribution of the chain by resorting to an existing fuzzy approach with restricted matrix multiplication. Preliminary results are very promising and deserve further research.

Keywords: Fuzzy probabilities, linguistic probabilities, linguistic labels, Markov chain, stationary distribution.

1. Introduction

Assume a set of observations, $X_0, X_1, ..., X_r$ of the state of a system that evolves over time, taken at time instants $t = 0, 1, ..., r$. Such observations define a collection of random variables that take values in a given state space $S = \{1, ..., n\}$. The indexed sequence $\{X_t, t = 1, 2, ...\}$ is called a *stochastic process* [1] in discrete time. If, additionally, it has the so-called Markov property,

$$P[X_{t+1} = x_{t+1}|X_t = x_t, X_{t-1} = x_{t-1}, ..., X_0 = x_0]$$
$$= P[X_{t+1} = x_{t+1}|X_t = x_t]$$

then the process is said to be a *Markov chain.*

Markov chains are a well-known statistical model of a number of real physical, natural and social phenomenons. Examples include areas such as biology, medicine (disease expansion models[2]), economy, and also problems related to AI such as speech recognition [3]. See [4] for more examples.

A special property of Markov chains is time-homogeneity, which arises when $P[X_{t+h} = j|X_t = i]$ does not depend on concrete time instants $t + h$ and t but only on the time difference h between them. In that case, $P[X_{t+h} = j|X_t = i] = P[X_h = j|X_0 = i], \quad \forall t \geq 0$.

An n-state time-homogeneous Markov chain is described by its transition matrix $P = (p_{ij}), i, j = 1, ..., n$, where $p_{ij} = P[X_{t+1} = j|X_t = i]$. Matrix P collects the one-step transition probabilities. In general, transition probabilities after h steps are represented as $P^{(h)} = (p_{ij}^{(h)}) = P^h$. These matrices are stochastic, as every row adds to 1.

Matrix P contains all the information needed to study the behaviour of the chain. Since the process evolves along time, when this information is subject to uncertainty or small errors, the long-term behaviour can turn very different, thus it is important to have an exact description of the process we are modeling through the transition probabilities. Unfortunately, this is not always possible, due to insufficient or total lack of numerical data to construct the model. In those cases, a suitable solution consists on the introduction of fuzzy set set theory to cope with uncertainty in the transition probabilities, and use fuzzy methods to carry out the desired calculations.

Two main approaches have been presented in the literature on fuzzy Markov chains. The first one, probably the most extended, consists in relaxing the restriction of having the process described by a stochastic matrix, and using a fuzzy relation over $S \times S$ instead. A number of works have been published in this line, dealing with both theoretical aspects [5, 6] and successful applications, such as processor power [7], speech recognition [8], and multi-temporal image classification [9, 10].

The second approach, proposed by J. Buckley [11], considers the transition matrix as composed of fuzzy numbers, and uses restricted matrix multiplication to operate with them in a way that the constraint of being a well-formed probability distribution always holds. Some works devoted to this proposal are [11, 12, 13, 14, 15]. It is interesting to note that all the classic probability theory can be fuzzified this way. By doing so, the author addresses topics such as fuzzy random variables (discrete and continuous), fuzzy Markov chains, fuzzy queuing theory or fuzzy inventory control [11].

Due to the intuitive use of fuzzy numbers on the transition matrix, and their natural connection with linguistic variables to express uncertain information, we present here for the first time the concept of a fuzzy linguistic Markov chain, defined as a Markov chain with a linguistic transition matrix. It represents a step further in coping with uncertain or vague information in stochastic processes, which can be suitable in cases when no data about the sys-

tem are available, or these are vague and consist of mere linguistic descriptions given either by an observer or an expert who is able to express how the phenomenon works only in an approximate manner. We will show how linguistic labels and their underlying fuzzy numbers can be used to represent transition probabilities, and how they can be operated to calculate one of the most usual characteristics of a Markov chain, namely the (linguistic) stationary distribution. The work is a continuation towards the linguistic domain of an experimental study recently conducted by the authors [16], in which linguistic labels were mentioned just as a possible future research direction.

The remainder of this contribution is structured as follows. In section 2, the fuzzy Markov chains approach presented in [11] is reviewed. In section 3 we explain how linguistic probabilities can be used in conjunction with fuzzy Markov chains. Section 4 contains an illustrative example of our linguistic proposal. Finally, section 5 is devoted to conclusions and further work.

2. Fuzzy Markov chains

When one needs to represent an uncertain quantity, a suitable solution is using a fuzzy number [17]. As it is well known, a fuzzy number is a special kind of fuzzy set [18] defined over the real numbers, whose membership function fulfills some additional conditions. Before giving a formal definition, we introduce the concept of α-cut. Given a fuzzy set $\tilde{A}$ over a universe U, its α-cuts are defined as $\tilde{A}[\alpha] = \{x \in U : \mu_{\tilde{A}} \geq \alpha\} = [\tilde{A}_\alpha^L, \tilde{A}_\alpha^U], \forall \alpha \in [0,1]$.

Definition 1. A fuzzy set $\tilde{A}$ defined over $\mathbb{R}$ is a fuzzy number iff

- $\tilde{A}[\alpha]$ are nonempty convex sets $\forall \alpha \in [0,1]$.
- $\tilde{A}(\alpha)$ are compact sets $\forall \alpha \in [0,1]$.
- $\tilde{A}(\alpha) \subseteq \tilde{A}(\beta)$ for $\alpha > \beta$.
- $\tilde{A}(\alpha) = \bigcap_{\beta < \alpha} \tilde{A}(\beta)$ for $\alpha \in (0,1]$.

Following this notion, Buckley [11] suggests using a fuzzy number for those entries of the transition matrix that are uncertain. Crisp (certain) entries can be represented as singleton fuzzy numbers. Thus, the crisp transition matrix P is replaced by a fuzzy transition matrix $\tilde{P} = (\tilde{p}_{ij}), i, j = 1, ..., n$, with $\tilde{p}_{ij}$ are fuzzy numbers. We will later explain how to obtain the entries of $\tilde{P}$.

Although the method we will explain is valid for any kind of fuzzy number, in section 3 we will restrict our attention to Trapezoidal Fuzzy Numbers (TrFNs), represented as 4-tuples. The TrFN $\tilde{A} = (a, b, c, d)$, where $a, b, c, d \in \mathbb{R}, a \leq b \leq c \leq d$ has the following membership function:

$$\mu_{\tilde{A}}(x) = \begin{cases} (x-a)/(b-a) & a \leq x < b \\ 1 & b \leq x \leq c \\ (d-x)/(d-c) & c < x \leq d \\ 0 & \text{otherwise} \end{cases}$$

We will use the following operations with TrFNs. Let $\tilde{A} = (a_1, b_1, c_1, d_1)$ and $\tilde{B} = (a_2, b_2, c_2, d_2)$ be two TrFNs. Then:

$$\tilde{A} \oplus \tilde{B} = (a_1 + a_2, b_1 + b_2, c_1 + c_2, d_1 + d_2) \quad (1)$$

$$\text{Distance: } d(\tilde{A}, \tilde{B}) = (|a_1 - b_1| + 2|a_2 - b_2| + 2|a_3 - b_3| + |a_4 - b_4|)/6 \quad (2)$$

As mentioned in [11], the uncertainty is on the transition probabilities, but not in the fact that they must add to 1. To address this issue, Buckley proposes the notion of restricted matrix multiplication, that assures the probability restriction holds at the α-cut level when doing computations with the transition matrix, as follows. Let $\tilde{P}$ be the fuzzy transition matrix of an n-state Markov chain. Denote $\tilde{P}(\alpha) = (\tilde{p}_{ij}[\alpha])$ the matrix of intervals whose entries are the α-cuts of the fuzzy numbers for a given α. This interval matrix can be thought of as the (possibly infinite) set of all crisp matrices which could actually act as the transition matrix of our Markov chain, since we are uncertain about that fact. Therefore, every row i of $\tilde{P}(\alpha)$ represents a closed space of discrete probability distributions which row i may take. This space is defined for row i by the bounded n-dimensional hypercube $\tilde{p}_{i1}[\alpha] \times \tilde{p}_{i2}[\alpha] \times ... \times \tilde{p}_{in}[\alpha] = \times_{j=1}^n \tilde{p}_{ij}[\alpha]$.

Note not all n-tuples $(a_1, ..., a_n)$ belonging to the above hypercube are valid for row i of the transition matrix. As mentioned before, the restriction of being a well-formed probability distribution must hold all the time, thus it is also necessary that $a_1 + ... + a_n = 1$. Let $\Delta_n = \{(x_1, ..., x_n) : x_i \geq 0 \text{ and } \sum_{i=1}^n x_i = 1\}$. With this, Buckley defines the *domain* of row i as follows.

Definition 2. Let P be an $n \times n$ fuzzy transition matrix. The domain of row i for a given membership degree $\alpha \in [0,1]$ is the set

$$Dom_i(\alpha) = \left(\times_{j=1}^n \tilde{p}_{ij}[\alpha]\right) \bigcap \Delta_n =$$

$$\{(p_{i1}, ..., p_{in}) \in \mathbb{R}^n : p_{ij} \in [\tilde{p}_{ij_\alpha}^L, \tilde{p}_{ij_\alpha}^U] \wedge \sum_j p_{ij} = 1\}$$

The domain of the whole matrix for a given α is $Dom(\alpha) = \times_{i=1}^n Dom_i(\alpha)$. Note the elements of $Dom(\alpha)$ are matrices of dimensions $n \times n$.

2.1. Fuzzy stationary probabilities

Very often we are interested in the stationary distribution of a Markov chain, $\pi = (\pi_1, ..., \pi_n)$. Each probability π_i represents the amount of time (in percentage) that the chain spends in state i, which is a valuable information about the global behaviour of the system we are modeling. Analytically, the stationary distribution fulfills $\pi = \pi P$. Given the exact transition matrix P, several theorems deal with the existence of such distribution. An homogeneous Markov chain is said to be irreducible if all states form a single group so that every state is accessible

from every other state, be it in one or more than one step. Every finite, irreducible Markov chain with transition matrix P has a *stationary distribution* π such that $\pi = \pi P$ [1]. However, since the transition matrix is precisely what we are uncertain about, no theorems can be applied here, at least in their original formulation. Therefore, we assume the chain does pose a stationary distribution.

It is possible to compute the fuzzy stationary distribution of a fuzzy Markov chain, following the approach described previously. The key idea is that every fuzzy stationary probability π_i can be constructed from its α-cuts, and the lower and upper bounds of each α-cut can be computed as the minimum and maximum possible value for π_i when the crisp transition matrix lies in the feasible solution space defined by $Dom(\alpha)$ for each α. Formally,

$$
\begin{aligned}
\tilde{\pi}_i[\alpha] &= [\pi_{i\alpha}^L, \pi_{i\alpha}^U] \text{ where} \\
\pi_{i\alpha}^L &= \min\{w_i | w = wM : M \in Dom(\alpha)\} \ (3) \\
\pi_{i\alpha}^U &= \max\{w_i | w = wM : M \in Dom(\alpha)\} (4)
\end{aligned}
$$

It can be proved [11] that such intervals are closed, connected sets, thus $\tilde{\pi}_i$ can be constructed from them by the representation theorem [19].

Hence, the process to calculate $\pi_{i\alpha}^L$ and $\pi_{i\alpha}^U$ for a given α consists in finding, within the matrix space $Dom(\alpha)$, a matrix that minimizes the i-th component of its stationary distribution, and another matrix maximizing it, respectively. As suggested in [11], this search process should be accomplished using a heuristic constrained optimization technique since the objective function, which is the expression to obtain component w_i of the stationary vector w from transition matrix M, is a black-box in the general case of an n-dimensional matrix: no general analytical expression exists to express w_i as a function of the entries of M, as such expression would be impractical when n grows a bit, e.g. $n \geq 5$. Details on the optimization algorithm employed here are provided in the experiment section.

Once the α-cuts have been calculated, regression can be applied to the lower and to the upper bounds (separately) to obtain an analytical expression of their membership functions, so that a completely defined fuzzy number is finally constructed.

3. Linguistic probabilities

What has been explained in the previous section is valid for any kind of fuzzy number, no matter how they have been obtained. As mentioned in the introduction, most often we have empirical data from the phenomenon, and estimates of transition probabilities can be derived from them, either crisp or fuzzy. However, in case no data are available, we may have to elicit probabilities from a human expert. In those cases, it can be quite difficult for him to express his expertise using numerical crisp probabilities that, in addition, must meet the requirement of adding to 1. In order to solve this issue, here we
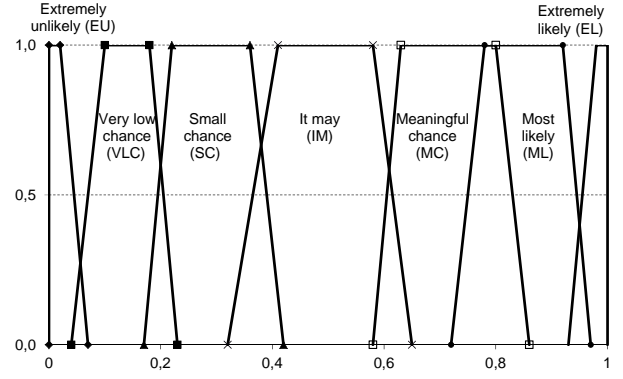

Figure 1: Membership functions of linguistic probability labels according to [21].

propose using natural language to express probabilities, thus modeling the transition probability as a linguistic variable[20] whose possible values are linguistic terms.

The concept of linguistic probability is not new. It was employed, for instance, by Bonissone [21]. As he points out, psychological studies cited in that work confirm that people are unwilling to give precise numerical estimates of probabilities, so it is reasonable to allow them to provide linguistic estimates of the likelihood of given statements that, in this case, are of the form *the system will move from state A to state B*. More recently, a proposal for a linguistic probability theory [22] explains how to extend the classical probability theory in a linguistic manner, with an example of a linguistic bayesian network. In the present contribution, we use linguistic probabilities as an additional layer that is placed on top of Buckley's fuzzy Markov chains approach explained in section 2.

We have defined 7 possible linguistic labels to evaluate the probability that the chain moves from one state to another. The underlying TrFNs are the mathematical structure that enables calculations with the labels. Their concrete values are those proposed in [21] after a psychological study (Fig. 1). Note some of the terms carry more uncertainty that others. We also consider crisp *Impossible* and *Sure*, whose TrFNs are respectively the singletons $(0, 0, 0, 0)$ and $(1, 1, 1, 1)$.

Because we are modelling a probability, the constraint of being a well-formed probability distribution must still hold. In the linguistic case, this is equivalent to the following condition [22]. Given a discrete linguistic probability distribution, $\tilde{\pi}_1, ..., \tilde{\pi}_n$, their sum must contain[1] the singleton fuzzy number 1_χ, which is defined by $\mu_{1_\chi}(x) = 1$ if $x = 1$, and 0 otherwise. This statement is valid regardless of the type of fuzzy numbers involved or how the sum has been defined, although in our experiments we employ Eq. (1). In a Markov chain, the use of linguistic probabilities leads to a linguistic

---

[1] $\tilde{A} \supseteq \tilde{B} \leftrightarrow \mu_{\tilde{A}}(x) \geq \mu_{\tilde{B}}(x) \forall x \in \mathbb{R}$

transition matrix. When expressing his judgments through linguistic terms, the expert should check the constraint explained above for every row:

$$\tilde{p}_{i1} \oplus \tilde{p}_{i2} \oplus ... \oplus \tilde{p}_{in} \supseteq 1_\chi, \forall i = 1, ..., n \qquad (5)$$

The reason is the following. The set $Dom_i(\alpha)$ is not empty when there exists at least one probability distribution whose elements fulfill all the interval (α-cuts) constraints of row i, and this happens $\forall \alpha \in [0, 1]$ if, and only if, the fuzzy numbers of row i, whose α-cuts constitute the box constraints, satisfy Eq. (5), as expressed in the following lemma.

Lemma 3. *Let* $\tilde{S} = \tilde{p}_{i1} \oplus ... \oplus \tilde{p}_{in}$. *Then,* $Dom_i(\alpha) \neq \emptyset \iff \tilde{S} \supseteq 1_\chi$, *i.e. iff Eq. (5) holds.*

Proof. a) If $\forall \alpha, Dom_i(\alpha)$ is not empty, then $\exists (p_{i1}, ..., p_{in}) \in \mathbb{R}^n : p_{ij} \in [\tilde{p}_{ij_\alpha}^L, \tilde{p}_{ij_\alpha}^U] \wedge \sum_j p_{ij} = 1$. Because $\tilde{p}_{ij_\alpha}^L \leq p_{ij} \leq \tilde{p}_{ij_\alpha}^U \ \forall j = 1, ..., n$, then $\sum_j \tilde{p}_{ij_\alpha}^L \leq \sum_j p_{ij} \leq \sum_j \tilde{p}_{ij_\alpha}^U$, i.e., $\sum_j \tilde{p}_{ij_\alpha}^L \leq 1 \leq \sum_j \tilde{p}_{ij_\alpha}^U$, which means that $1 \in [\sum_j \tilde{p}_{ij_\alpha}^L, \sum_j \tilde{p}_{ij_\alpha}^U]$, i.e. $1 \in [\tilde{S}_\alpha^L, \tilde{S}_\alpha^U]$, and this happens $\forall \alpha$, hence $\tilde{S} \supseteq 1_\chi$ and Eq. (5) is satisfied.

b) If Eq. (5) is satisfied, then $\forall \alpha, 1 \in [\tilde{S}_\alpha^L, \tilde{S}_\alpha^U]$ so $\sum_j \tilde{p}_{ij_\alpha}^L \leq 1 \leq \sum_j \tilde{p}_{ij_\alpha}^U$. It follows that, for the last element n, $\tilde{p}_{in_\alpha}^L \leq 1 - \sum_{j<n} \tilde{p}_{ij_\alpha}^L$. Moreover, $1 \leq \tilde{p}_{in_\alpha}^U$ so $1 - \sum_{j<n} \tilde{p}_{ij_\alpha}^L \leq \tilde{p}_{in_\alpha}^U$. Putting both together, $\left(1 - \sum_{j<n} \tilde{p}_{ij_\alpha}^L\right) \in [\tilde{p}_{in_\alpha}^L, \tilde{p}_{in_\alpha}^U]$. Thus we can easily obtain an element of $Dom_i(\alpha)$ whose components belong to the desired intervals, for instance $(\tilde{p}_{i1_\alpha}^L, \tilde{p}_{i2_\alpha}^L, ..., \tilde{p}_{i,(n-1)_\alpha}^L, 1 - \sum_{j<n} \tilde{p}_{ij_\alpha}^L) \in \mathbb{R}^n$. Since we have found an element of $Dom_i(\alpha)$, it means that $Dom_i(\alpha) \neq \emptyset$. $\qquad \square$

The above lemma shows the connection between two apparently different proposals on fuzzy probabilities, namely [11] and [22].

3.1. Output retranslation process

Since our system allows for a linguistic input, it is also expected that the system is able to provide a linguistic output. Therefore, after first computing the α-cuts of the stationary fuzzy probabilities and subsequently applying regression, it is necessary to assign a linguistic probability label to every fuzzy number obtained. This is known as *retranslation* and has been studied previously, for instance in [23]. The idea is to assign to every output TrFN the label of the *closest* TrFN of the reference set (Fig. 1).

Although other variants are possible, at this stage of research we use a simple measure of distance between TrFNs defined in Eq. (2). Note this metric can be replaced by any other that better fits the needs of the problem.

4. Example of application and results

We depart from the Markov chain of Fig. 2 whose transition matrix is known to us. It represents the
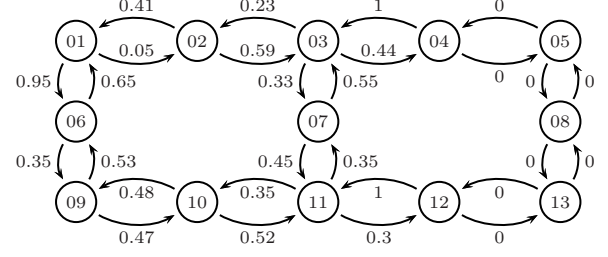

Figure 2: Markovian patrolling scheme of an autonomous robot in a map with 13 cells. Reproduced from [24]. Non accessible states 5, 8 and 13 were not considered for the transition matrix.

	1	2	3	4	5	6	7	8	9	10
1	-	VLC	-	-	ML	-	-	-	-	-
2	IM	-	IM	-	-	-	-	-	-	-
3	-	SC	-	SC	-	IM	-	-	-	-
4	-	-	EL	-	-	-	-	-	-	-
5	MC	-	-	-	-	-	SC	-	-	-
6	-	-	IM	-	-	-	-	-	IM	-
7	-	-	-	-	IM	-	-	IM	-	-
8	-	-	-	-	-	-	IM	-	IM	-
9	-	-	-	-	-	SC	-	IM	-	SC
10	-	-	-	-	-	-	-	-	EL	-

Table 1: Linguistic transition matrix estimated from a sequence of 200 observations.

randomized movement of an autonomous robot that is patrolling a bi-dimensional area divided in cells against an intruder who wants to attack some (unknown) location. The movement from one cell to any of the adjacent locations is randomized in order to be able to patrol a larger environment without letting any location always unvisited. In this way, an intruder who learns the robot's movement, even if he notices that the patrolling scheme is randomized and learns the robot's movement probabilities, does not have a guarantee that he will be able to successfully attack any given location, since there is always a non-zero probability of being caught. The movement probability to an adjacent location only depends on the current location, and not on the path followed by the robot to reach current location. Therefore, it is Markovian by definition, with each state of the chain representing one location within the grid. More details on this problem and the solution method that led to this Markov chain can be found in [24].

In a real instance of this problem, the transition probabilities shown in the figure would be unknown, so an expert that either has been watching the robot or knows (for some reason) its patrolling scheme in an approximate way would provide a linguistic transition matrix directly.

However, instead of asking an expert, in this preliminary study we have proceeded as follows. We generated a sequence of 200 observations of the evolution of the chain. Each observation is an integer number indicating the state at that time instant.
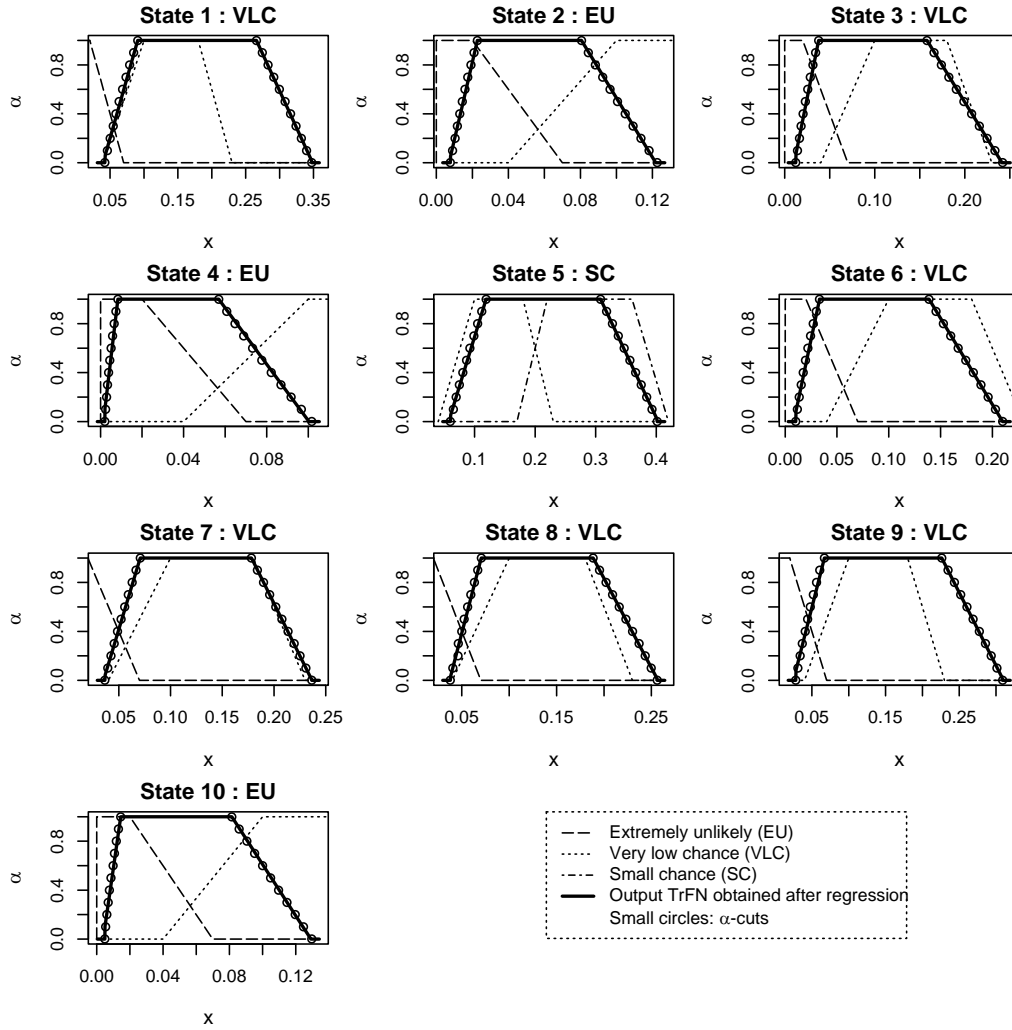
Figure 3: Fuzzy stationary probabilities (TrFNs) obtained and linguistic labels assigned (above each plot). For each state, we show the TrFN obtained, together with the closest labels of the reference set (Fig. 1).

Then, with this sequence, we estimated the transition probabilities as the proportion of times the chain moves from each state to another. Finally, we converted such numeric estimates into linguistic labels, by replacing each crisp probability estimate by the linguistic label to which it belongs most, according to Fig. 1. For instance, a crisp probability of 0.4 belongs simultaneously to fuzzy sets *Small chance* and *It may*, but the membership degree of the latter is higher, so it would be replaced by *It may (IM)*. The linguistic transition matrix obtained is shown in Table 1, where "-" stands for an impossible transition in the crisp sense. We made some adjustments to make sure Eq. (5) holds for every row, such as replacing in some cases the label for which a crisp estimate has the highest membership, by the immediately greater or immediate smaller label. Table 1 was obtained as a result of this conversion process. The TrFNs underlying the labels of the table are the input of our method.

We used the R programming language [25] for implementing the optimization problems of Eq. (3) and (4), and also for plotting the results. The opti-

mization algorithm employed was an R version [26] of Differential Evolution [27]. The computation of the α-cuts took approximately 10 minutes in a parallelized implementation that exploits all the cores over an Intel Core-i7 processor at 2.67 GHz with 6 GB RAM. We made use of the plotting facilities provided by the *FuzzyNumbers* package [28]. The results are displayed in Fig. 3.

A number of issues should be pointed out. The first one is that the membership function of the obtained TrFNs (thick line) is an almost perfect line at both sides of the core. This is just as expected if we take into account that the input were TrFNs with straight lines as well. However, in addition, this phenomenon confirms that the optimization process to compute the α-cuts is working fine, since the resulting lower bounds and upper bounds are visually perfectly aligned. If the expressions of the membership functions of input probabilities were more "exotic", then generalized TrFNs with other kind of membership functions at both sides (not lines but curves) would have been obtained because the α-cuts calculated would not be aligned. This does

not represent a problem since it is enough to define a proper distance measure that can cope with any kind of TrFN in order to assign a label to them.

It also proves that the general method outlined in [11] is effective in practice, provided the optimizer can deal properly with constrained optimization to make sure the solutions (i.e., matrices) evaluated belong to $Dom(\alpha)$ in each case.

Secondly, in this case only three different labels were required for all states. This depends heavily on the nature of the Markov chain being observed. Some Markov chains, as the one of this example, have very spread stationary probabilities, while other chains tend to be more concentrated around one state which is visited far more often than the rest. Therefore, this is not an issue caused by our method.

Note there are some difficult cases when assigning a linguistic to the TrFN, such as states 2 or 5. In those cases, a different distance measure might have yielded different results but, as mentioned in the previous section, our method proposal is independent of the distance measure employed and thus admits any function that fits the user's needs.

Finally, it should be recalled that, regardless the shape of the input probabilities, the output is linguistic and hence, easier to interpret than numerical values. In addition, since the transition matrix is uncertain, crisp stationary probabilities obtained with poor numerical punctual estimates of the transition probabilities may be misleading and lead to erroneous conclusions when comparing values of stationary probabilities that are very similar, or when we have too little information to do a reliable comparison. In those cases, our method will most likely assign the same linguistic label to those fuzzy stationary probabilities, which is probably more desirable and more robust.

5. Conclusions and further work

For the first time, we have developed a method to compute linguistic stationary probabilities of a Markov chain when the information about the transition probabilities is given in linguistic terms. We have addressed the uncertainty that is present in natural language judgments by using fuzzy numbers in the transition matrix. Furthermore, the fuzzy probabilities taken as reference set were obtained after a psychological study to better capture the uncertainty behind the linguistic evaluation of the likelihood of real-life events. We have employed an existing proposal to calculate the α-cuts of fuzzy stationary probabilities. Trapezoidal fuzzy numbers are built from the obtained α-cuts through a linear regression procedure to compute the membership function at both sides of the core. Finally, the output fuzzy numbers are assigned a linguistic term from the reference set.

Our proposal has been implemented in the R pro-

grammin language and has been successfully applied to a sample Markov chain. The obtained TrFNs preserve the shape of the input fuzzy numbers, and the output linguistic stationary probabilities are easier to interpret. In future studies, formal sensitivity analyses should be conducted to demonstrate the robustness of the method in problems where small variations in the transition probability matrix cause a big change in the stationary probabilities. Similar robustness studies have been conducted in [6] following the alternative fuzzy Markov chain approach that is based on fuzzy relations. Furthermore, other kind of measures from the Markov chain, such as first-passage time of a given state, could be investigated with the fuzzy linguistic methodology, as well as the application to real problems such as robotic patrolling.

Acknowledgments

References

[1] J. Medhi. *Stochastic Models in Queueing Theory, 2nd Ed.* Academic Press, Boston, 2002.

[2] P. K. Newton, J. Mason, K. Bethel, L. A. Bazhenova, J. Nieva, and P. Kuhn. A stochastic markov chain model to describe lung cancer growth and metastasis. *PLoS ONE*, 7(4):e34637, 04 2012.

[3] B. H. Juang and L. R. Rabiner. Hidden Markov Models for Speech Recognition. *Technometrics*, 33(3):251–272, 1991.

[4] W.K. Ching and M. K. Ng. *Markov Chains: Models, Algorithms and Applications.* Springer-Verlag, New York, 2006.

[5] M. Kurano, M. Yasuda, J. Nakagami, and Y. Yoshida. A limit theorem in some dynamic fuzzy systems. *Fuzzy Sets and Systems*, 51(1):83 – 88, 1992.

[6] K. E. Avrachenkov and E. Sanchez. Fuzzy Markov Chains and Decision-Making. *Fuzzy Optimization and Decision Making*, 1:143–159, 2002.

[7] R. Kruse, R. Buck-Emden, and R. Cordes. Processor Power Considerations - An Application of Fuzzy Markov Chains. *Fuzzy Sets and Systems*, 21(3):289–299, 1987.

[8] D. Tran and M. Wagner. Fuzzy hidden markov models for speech and speaker recognition. In *Proc. of NAFIPS*, pages 426–430, 1999.

[9] R. Q. Feitosa, G. A. O. P. Costa, G. L. A. Mota, and B. Feijó. Modeling Alternatives for Fuzzy Markov Chain-Based Classification

of Multitemporal Remote Sensing Data. *Pattern Recognition Letters*, (32):927 – 940, 2011.

[10] A. O. Alves, G. L. A. Mota, G. A. O. P. Costa, and R. Q. Feitosa. Estimation of Transition Possibilities for Fuzzy Markov Chains Applied to the Analysis of Multitemporal Image Sequences. In *Proc. of the 4th Int. Conf. on Geographic Object-Based Image Analysis (GEO-BIA)*, pages 367–371, 2012.

[11] J.J. Buckley. *Fuzzy Probabilities: New Approach and Applications, 2nd Edition*, volume 115 of *Studies in Fuzziness and Soft Computing*. Springer-Verlag, 2005.

[12] J. J. Buckley. Uncertain Probabilities III: the Continuous Case. *Soft Computing*, 8:200–206, 2004.

[13] J. J. Buckley and E. Eslami. Uncertain Probabilities I: the Discrete Case. *Soft Computing*, 7:500–505, 2003.

[14] J. J. Buckley and E. Eslami. Uncertain Probabilities II: the Continuous Case. *Soft Computing*, 8:193–199, 2004.

[15] J.J. Buckley and E. Eslami. Fuzzy Markov Chains: Uncertain Probabilities. *Mathware & Soft Computing*, 9(1), 2008.

[16] P.J. Villacorta and J.L. Verdegay. FuzzyStatProb: An R Package for the Estimation of Fuzzy Stationary Probabilities from a Sequence of Observations of an Unknown Markov Chain. *Submitted to the Journal of Statistical Software*, 2013.

[17] D. Dubois and H. Prade. Operations on fuzzy numbers. *International Journal of Systems Science*, 9(6):613–626, 1978.

[18] L.A. Zadeh. Fuzzy sets. *Information and Control*, 8(3):338 – 353, 1965.

[19] C.V. Negoita and D.A. Ralescu. *Applications of Fuzzy Sets to Systems Analysis*. John Wiley & Sons, 1975.

[20] L. A. Zadeh. The Concept of a Linguistic Variable and its Application to Approximate Reasoning - I. *Information Sciences*, 8(3):199 – 249, 1975.

[21] P. Bonissone and K. Decker. Selecting uncertainty calculi and granularity: An experiment in trading-off precision and complexity. In *Proc. of the Conf. Annual Conference on Uncertainty in Artificial Intelligence (UAI-85)*, pages 57 – 66, 1985.

[22] J. Halliwell and Q. Shen. Linguistic probabilities: Theory and application. *Soft Computing*, 13(2):169–183, 2008.

[23] R.R. Yager. On the retranslation process in Zadeh's paradigm of computing with words. *IEEE Transactions on Systems, Man, and Cybernetics - Part B: Cybernetics*, 34:1184–1195, 2004.

[24] F. Amigoni, N. Basilico, and N. Gatti. Finding the optimal strategies for robotic patrolling with adversaries in topologically-represented eenvironments. In *Proc. of the IEEE Conf. on Robotics and Automation*, pages 819–824, 2009.

[25] R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2012.

[26] K. M. Mullen, D. Ardia, D. L. Gil, D. Windover, and J. Cline. DEoptim: An R Package for Global Optimization by Differential Evolution. *Journal of Statistical Software*, 40(6):1–26, 2011.

[27] K.V. Price, R.M. Storn, and J.A. Lampinen. *Differential Evolution: A Practical Approach to Global Optimization*. Springer-Verlag - Natural Computing Series, 2005.

[28] M. Gagolewski. *FuzzyNumbers Package: Tools to Deal with Fuzzy Numbers in R*, 2012.