KYBERNETIKA — VOLUME *59* (2023), NUMBER 1, PAGES 160–178

AN EXTENDED VERSION OF AVERAGE MARKOV DECISION PROCESSES ON DISCRETE SPACES UNDER FUZZY ENVIRONMENT

Hugo Cruz-Suárez, Raúl Montes-de-Oca and R. Israel Ortega-Gutiérrez

The article presents an extension of the theory of standard Markov decision processes on discrete spaces and with the average cost as the objective function which permits to take into account a fuzzy average cost of a trapezoidal type. In this context, the fuzzy optimal control problem is considered with respect to two cases: the max-order of the fuzzy numbers and the average ranking order of the trapezoidal fuzzy numbers. Each of these cases extends the standard optimal control problem, and for each of them the optimal solution is related to a suitable standard optimal control problem, and it is obtained that (i) the optimal policy coincides with the optimal policy of this suitable standard control problem, and (ii) the fuzzy optimal value function is of a trapezoidal shape. Two models: a queueing system and a machine replacement problem are provided in order to examplify the theory given.

Keywords: Markov decision process, average criterion, trapezoidal fuzzy cost, max-order, average ranking

Classification: 90C40, 93C40

1. INTRODUCTION

This paper concerns discrete-time Markov decision processes (MDPs, in singular MDP) (see [9, 24], and [26]) on a denumerable state space under fuzzy preferences (see [6] and [29]). Specifically, a trapezoidal fuzzy cost-per-stage is taken into account and the other components of the Markov control model are considered as usual in the MDPs theory. The evolution of the system is roughly described as follows: in each stage the decision-maker observes the state of the system and chooses an admissible action. The sequence of actions is known as a policy. Here, the performance of a policy is measured by the fuzzy long-run expected average cost and the fuzzy optimal control problem consists in determining a policy that minimizes in an appropiate sense the performance criterion. In fact, the minimization is presented from two approaches: first by using the partial order on the α-cuts of fuzzy numbers and then by applying the average ranking order on the trapezoidal fuzzy numbers. It is relevant to comment that both approaches applied to the fuzzy optimal control problem extend the standard optimal control problem with respect to the average cost (see [1, 9, 24], and [26]).

DOI: 10.14736/kyb-2023-1-0160

In fuzzy theory literature, it has been noted that the decision making in real situations is taking place in an environment of uncertainty and imprecision (see, for instance, [12] and [28]), and in many cases the uncertainty is not characterized by a random processes. Then, a possibility to model this class of situations is to use the fuzzy theory proposed by L. Zadeh [29]. In this paper, under this approach an MDP with fuzzy cost is analyzed. Background around this topic is listed now and the presentation is given with respect to the performance criterion. In [3], a fuzzy total expected reward criterion is analyzed for an MDP with finite state space and with a trapezoidal fuzzy reward function. On the other hand, one of the most studied criteria in the literature is the discounted total expected reward/cost, see, for instance, [2, 14, 15, 16, 17] and [25]. In these works, the fuzzy approach is applied either in the reward/cost function ([2, 14, 15, 25]) or in the dynamic of the system ([14, 16, 17]), all of them under finite state and action spaces framework. In regards to the long-run expected average cost criterion, only the following two works were found: [10] and [13]. In [13], a Pareto optimal policy maximizing the average expected fuzzy reward under the max-order is found. And, in [10], using the concept of the random fuzzy variable in credibility theory they introduce a new model, called credibilistic MDPs, in order to treat the unknown transition matrices of classical MDPs with credibilistic information; it is important to mention that in both cases the expected average cost is analyzed under the assumption that the state and action spaces are finite.

The methodology applied in this article is described as follows. Based on a standard MDP a fuzzy MDP is induced, with the characteristic that all its components are the same as the standard MDP except for its cost function which is considered a fuzzy function. Under the assumption that the cost function is a fuzzy trapezoidal one, it is demonstrated that this fuzzy function, with a convenient theoretical framework, is a fuzzy random variable and a formula for its expectation is deduced. Based on this, it is shown that the fuzzy average cost is well defined, then the authors proceed to introduce the fuzzy optimal control problem subject to the average criterion and use the partial order on the α-cuts of fuzzy numbers (see [8]). And, under the assumption that the standard optimal control problem has an optimal solution it is demonstrated that the optimal policy of the fuzzy MDP coincides with the optimal policy of the standard MDP. Moreover, the fuzzy optimal value function is trapezoidal in shape, as expected. Later, a similar analysis is presented for the approach of average ranking order on the trapezoidal fuzzy numbers (see [21, 22] and [25]). In this approach, it is confirmed again that the policies of both optimal control problems, i. e., the standard problem and the fuzzy one coincide. Moreover, it is important to point out that the manuscript provides an adequate interpretation for the use of fuzzy costs in MDPs, see Remark 4.7.

Now, the organization of the paper is given. In Section 2, some basic definitions and properties related to fuzzy theory are presented. Section 3 addresses the topics of MDPs focused on the case of the long-run expected average cost. Subsequently, in Section 4 the fuzzy optimal control problem is introduced. Section 5 presents the analysis for fuzzy optimal control problem with respect to the average ranking. Section 6 illustrates the theory exposed in two examples. The first example is referring to a queueing model, and the second example shows a fuzzy machine replacement problem. To end the paper, in Section 7 some ideas for research in progress are provided.

Notation. In the article, the following standard mathematical symbols will be distinguished in the fuzzy context with an asterix symbol: " * " (or sometimes the symbol " ** ") . That is, in the fuzzy context, "$\leq$", "$+$" and "$\sum$", will be denoted by "$\leq^*$", "$+^*$" and "$\sum^*$", respectively. Similarly, in the fuzzy context, the expectation operator "E", the limit "lim" and the infimum "inf", will be denoted by "E^*", "lim^*" and "inf^*", respectively. Also, the notations: "$\leq^{**}$" and "inf^{**}" are used for a second comparison between trapezoidal fuzzy numbers referent to the ranking order. It is important to mention that the product of a real number λ and a fuzzy number Υ will be simply denoted as $\lambda\Upsilon$. Moreover, some special functions which appear as fuzzy quantities, say, the cost function, the optimal value function, and so on, will be distinguished with a "tilde"; for instance, the fuzzy cost function will be written as $\widetilde{C}$.

2. SOME FACTS ON FUZZY THEORY

The first part of this section presents some definitions and basic results about the fuzzy set theory (see [6, 27], and [29]).

Let Λ be a non-empty set. Then a *fuzzy set* Γ on Λ is defined in terms of the *membership function* Γ', which assigns to each element of Λ a real value from the interval $[0,1]$. The *α-cut* of Γ, denoted by Γ_α, is defined to be the set $\Gamma_\alpha := \{x \in \Lambda \mid \Gamma'(x) \geq \alpha\}$ $(0 < \alpha \leq 1)$ and Γ_0 is the *closure* of $\{x \in \Lambda \mid \Gamma'(x) > 0\}$ denoted by $cl\{x \in \Lambda \mid \Gamma'(x) > 0\}$.

Definition 2.1. A *fuzzy number* Γ is a fuzzy set defined on the set of real numbers $\mathbb{R}$ (i. e., taking $\Lambda = \mathbb{R}$ in the previous definition), which satisfies:

a) Γ' is normal, i. e., there exists $x_0 \in \mathbb{R}$ with $\Gamma'(x_0) = 1$;

b) Γ' is convex, i. e., Γ_α is convex for all $\alpha \in [0,1]$;

c) Γ' is upper-semicontinuous;

d) Γ_0 is compact.

The set of the fuzzy numbers will be denoted by $\mathfrak{F}(\mathbb{R})$.

Definition 2.2. A fuzzy number Γ is called a *trapezoidal fuzzy number* if its membership function has the following form:

$$\Gamma'(x) = \begin{cases} 0 & \text{if } x \leq l \\ \frac{x-l}{m-l} & \text{if } l < x \leq m \\ 1 & \text{if } m < x \leq n \\ \frac{p-x}{p-n} & \text{if } n < x \leq p \\ 0 & \text{if } p < x \ , \end{cases} \tag{1}$$

where l, m, n and p are real numbers, with $l < m \leq n < p$. A trapezoidal fuzzy number is simply denoted by (l, m, n, p).

Remark 2.3. a) The case in which $m = n$ in (1) will be named a *triangular fuzzy number* and it will be simply denoted as (l, m, p). And, considering the degenerated case in which $l = m = p$ in a triangular number, the *fuzzy representation* of the real number m is obtained with the membership function given by:

$$m'(x) = \begin{cases} 1 & \text{if } x = m \\ 0 & \text{if } x \neq m. \end{cases} \tag{2}$$

b) For a trapezoidal fuzzy number $\Gamma = (l, m, n, p)$ the corresponding α-cuts are given by $\Gamma_\alpha = [(m - l)\alpha + l, p - (p - n)\alpha]$, $\alpha \in [0, 1]$ (see [23]).

Definition 2.4. Let Γ and Υ be fuzzy numbers. If " $\star$ " denotes the addition or the scalar mutiplication, then it is defined as the fuzzy set $\Gamma \star \Upsilon$ with the membership function:

$$(\Gamma \star \Upsilon)'(u) = \sup_{u = x \star y} \min\{\Gamma'(x), \Upsilon'(y)\},$$

for all $u \in \mathbb{R}$.

As a consequence of Definition 2.4, it is possible to obtain the following result for trapezoidal fuzzy numbers (see [23]).

Lemma 2.5. If $H = (a_l, a_m, a_n, a_p)$ and $I = (b_l, b_m, b_n, b_p)$ are two trapezoidal fuzzy numbers and letting λ be a positive number, then it follows that

a) $\lambda H = (\lambda a_l, \lambda a_m, \lambda a_n, \lambda a_p)$, and

b) $H +^* I = (a_l + b_l, a_m + b_m, a_n + b_n, a_p + b_p)$. And, it also holds: if $\{(a_l^t, a_m^t, a_n^t, a_p^t) : 0 \leq t \leq N\}$ is a finite set of N trapezoidal fuzzy numbers, then

$$\sum_{t=0}^{N}{}^{*} (a_l^t, a_m^t, a_n^t, a_p^t) = \Big(\sum_{t=0}^{N} a_l^t, \sum_{t=0}^{N} a_m^t, \sum_{t=0}^{N} a_n^t, \sum_{t=0}^{N} a_p^t \Big).$$

Let $\mathbb{D}$ denote the set of all closed bounded intervals on the real line $\mathbb{R}$. For $\Psi = [a_l, a_u]$, $\Phi = [b_l, b_u] \in \mathbb{D}$ define

$$d(\Psi, \Phi) = \max(|a_l - b_l|, |a_u - b_u|).$$

It is possible to verify that d defines a metric on $\mathbb{D}$ and that $(\mathbb{D}, d)$ is a complete metric space (see [19]). Now, if $\widetilde{\eta} \in \mathfrak{F}(\mathbb{R})$, then $\widetilde{\eta}_\alpha$ is a compact set because its membership function is upper semicontinuous and has a compact support. Therefore, it is defined as $\hat{d} : \mathfrak{F}(\mathbb{R}) \times \mathfrak{F}(\mathbb{R}) \longrightarrow \mathbb{R}$ by

$$\hat{d}(\widetilde{\eta}, \widetilde{\mu}) = \sup_{\alpha \in [0,1]} d(\widetilde{\eta}_\alpha, \widetilde{\mu}_\alpha).$$

It is straightforward to see that $\hat{d}$ is a metric in $\mathfrak{F}(\mathbb{R})$ (see [19]).

Definition 2.6. A sequence $\{\widetilde{\eta}_t\}_{t=0}^{\infty}$ of fuzzy numbers is said to be *convergent* to the fuzzy number $\widetilde{\mu}$, written as $\lim^*_{t\longrightarrow\infty} \widetilde{\eta}_t = \widetilde{\mu}$ if and only if $\hat{d}(\widetilde{\eta}_t, \widetilde{\mu}) \longrightarrow 0$ as $t \longrightarrow \infty$.

Lemma 2.7. (Puri and Ralescu [19]) The metric space $(\mathfrak{F}(\mathbb{R}), \hat{d})$ is complete.

Now, for $\widetilde{\eta}, \widetilde{\mu} \in \mathfrak{F}(\mathbb{R})$, with α-cuts $\widetilde{\eta}_\alpha = [a_\alpha, b_\alpha]$ and $\widetilde{\mu}_\alpha = [c_\alpha, d_\alpha]$, $\alpha \in [0,1]$, respectively, define $\widetilde{\eta} \leqslant^* \widetilde{\mu}$ if and only if $a_\alpha \le c_\alpha$ and $b_\alpha \le d_\alpha$ for all $\alpha \in [0,1]$ (see [8]). It is not difficult to verify that the order "$\leq^*$ " is, in fact, a partial order on $\mathfrak{F}(\mathbb{R})$.

Remark 2.8. Take $w, z \in \mathbb{R}$, and let $\widetilde{w}$ and $\widetilde{z}$ be fuzzy numbers with membership functions given by $(\widetilde{w})'(x)=1$, $x = w$ and $(\widetilde{w})'(x) = 0$, $x \neq w$, and $(\widetilde{z})'(x)=1$, $x = z$ and $(\widetilde{z})'(x) = 0$, $x \neq z$. Then, it is easy to see that $\widetilde{w} \leqslant^* \widetilde{z}$ is equivalent to $w \le z$.

Moreover, the following *comparison* between the trapezoidal fuzzy numbers (see [21], [22], and [25]) is also introduced. Let $H = (a_l, a_m, a_n, a_p)$ be a trapezoidal fuzzy number. Then its *average ranking* $\Re(H)$ is defined as

$$\Re(H) = \frac{a_l + a_m + a_n + a_p}{4}.$$

Now, let $H = (a_l, a_m, a_n, a_p)$ and $I = (b_l, b_m, b_n, b_p)$ be trapezoidal fuzzy numbers. Hence, it is defined that $H \leqslant^{**} I$ if and only if

$$\Re(H) \le \Re(I). \tag{3}$$

Remark 2.9. In relation to the last comparison between trapezoidal fuzzy numbers, note that in the degenerate case for which in H: $a_l = a_m = a_n = a_p = a$ and in I: $b_l = b_m = b_n = b_p = b$, it results that $H \leqslant^{**} I$ if and only if $a \le b$.

Following [18] and [19], the next definitions on fuzzy random variables and their expectations are established. For this, $\mathfrak{C}(\mathbb{R})$ denotes the class of nonempty compact subsets of $\mathbb{R}$, and if $(\Omega_1, \mathcal{A}_1)$ and $(\Omega_2, \mathcal{A}_2)$ are measurable spaces, then $\mathcal{A}_1 \otimes \mathcal{A}_2$ denotes the corresponding product σ-algebra associated to the product space $\Omega_1 \times \Omega_2$.

Definition 2.10. Let $(\Omega, \mathcal{A})$ be a measurable space and $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ be the measurable space of the set of real numbers. A function $\widetilde{Y} : \Omega \longrightarrow \mathfrak{F}(\mathbb{R})$ is said to be a *fuzzy random variable* associated with $(\Omega, \mathcal{F})$, if the section $\widetilde{Y}_\alpha : \Omega \longrightarrow \mathfrak{C}(\mathbb{R})$ which is the α-cut function defined by $\widetilde{Y}_\alpha(\omega) = (\widetilde{Y}(\omega))_\alpha$ for all $\omega \in \mathit{\Omega}$ and $\alpha \in [0,1]$ satisfies that $Gr(\widetilde{Y}_\alpha) = \{(\omega, x) \in \Omega \times \mathbb{R} \mid x \in (\widetilde{Y}(\omega))_\alpha\} \in \mathcal{A} \otimes \mathcal{B}(\mathbb{R})$, for all $\alpha \in [0,1]$.

Definition 2.11. Given a probability space $(\Omega, \mathcal{A}, P)$, a fuzzy random variable $\widetilde{Y}$ associated to $(\Omega, \mathcal{A})$ is said to be an *integrably bounded fuzzy random variable* with respect to $(\Omega, \mathcal{A}, P)$ if there is a function $h : \Omega \longrightarrow \mathbb{R}$, $h \in L^1(\Omega, \mathcal{A}, P)$ such that $|x| \le h(\omega)$, for all $(\omega, x) \in \mathit{\Omega} \times \mathbb{R}$ with $x \in (\widetilde{Y}(\omega))_0 := \widetilde{Y}_0(\omega)$.

Definition 2.12. Given an integrably bounded fuzzy random variable $\widetilde{Y}$ associated to the probability space $(\Omega, \mathcal{A}, P)$, then the *fuzzy expected value* of $\widetilde{Y}$ in Aumann's sense is the unique fuzzy set of $\mathbb{R}$, $E^*[\widetilde{Y}]$ such that for each $\alpha \in [0,1]$:

$$\left(E^*[\widetilde{Y}]\right)_\alpha = \left\{\int_{\mathit{\Omega}} f(\omega)\, \mathrm{d}P(\omega) \mid f : \mathit{\Omega} \longrightarrow \mathbb{R}, f \in L^1(\Omega, \mathcal{A}, P), f(\omega) \in (\widetilde{Y}(\omega))_\alpha \;\; a.s.\ [P]\right\}.$$

3. STANDARD AVERAGE MARKOV DECISION PROCESSES

Markov Decision Models

Let $(X, A, \{A(i) \mid i \in X\}, Q, C)$ be the usual discrete-time Markov decision model (see [9, 24], and [26]), where the state space X is denumerable and the decision space A is finite. For each $i \in X$, $A(i) \subset A$, $A(i) \neq \varnothing$, is the subset of admissible actions at a state i. Let $\mathbb{K} := \{(i,a) \mid i \in X,\ a \in A(i)\}$. $Q = [p_{ij}(a)]$ is the controlled transition law on X given $\mathbb{K}$, for each $(i,a) \in \mathbb{K}$, $p_{ij}(a) \geq 0$ and $\sum_{j\in X} p_{ij}(a) = 1$. Finally, the cost per-stage C is a nonnegative function on $\mathbb{K}$.

Strategies

A *decision strategy* π is a (possibly randomized) rule for choosing actions, and at each time $t\,(t = 0, 1, \ldots)$ the decision prescribed by π may depend on the current state as well as on the history of the previous states and actions. The set of all strategies will be denoted by Π. Given an initial state $i \in X$ and $\pi \in \Pi$ there is a canonical space $(\Omega, \mathcal{A}, P_{i,\pi})$ with the corresponding state-action process $\{(x_t, a_t)\}$ (for details, see [9, 24], and [26]). $E_{i,\pi}$ denotes the expectation operator with respect to the probability measure $P_{i,\pi}$, and the stochastic process $\{x_t\}$ will be called Markov decision process. $\mathbb{F}$ denotes the set of functions $f : X \to A$ such that $f(i) \in A(i)$ for all $i \in X$. A strategy $\pi \in \Pi$ is *stationary* if there exists $f \in \mathbb{F}$ such that, under π, the action $f(x_t)$ is applied at each time t. The class of stationary strategies is naturally identified with $\mathbb{F}$.

Optimality Criterion

Given $\pi \in \Pi$ and initial state $x_0 = i \in X$, let

$$J_n(i,\pi) = \frac{1}{n} E_{i,\pi}\Big[\sum_{t=0}^{n-1} C(x_t, a_t)\Big],$$

for $n > 0$, and

$$J(i,\pi) = \lim_{n\to\infty} J_n(i,\pi) \tag{4}$$

be the *long-run expected average cost* when using the strategy π, given the initial state i.

Remark 3.1. See Proposition 8.1.1 in [20] and Proposition 6.1.1 in [26] for the existence of the limit in (4). Some authors consider in their works "limsup" in (4) instead of "lim", as it is taken into account in this paper (see [9] and [26]). Of course, if the limit exists in (4) it coincides with the corresponding limsup.

A strategy π_o is said to be *optimal* if $J(i,\pi_o) = J_o(i)$ for all $i \in X$, where

$$J_o(i) = \inf_{\pi\in\Pi} J(i,\pi), \tag{5}$$

$i \in X$. J_o defined in (5) is called the *optimal value function.*

Existence of Stationary Optimal Policies

It is important to observe that the conditions which ensure the existence of optimal policies requiere the existence of at least one policy $\pi \in \Pi$ such that $J(i,\pi) < \infty$, for all $i \in X$; otherwise the solution of the optimal decision problem would be trivial.

Lemma 3.2. Under certain Assumptions (see Remark 3.3 below) there exists a constant j_o and a function $h : X \longrightarrow \mathbb{R}$ such that

a) The pair (j_o, h) satisfies the following *Optimality Equation*:

$$j_o + h(i) = \inf_{a \in A(i)} \left[C(i,a) + \sum_j p_{ij}(a) h(j) \right], \tag{6}$$

for all $i \in X$.
Moreover,

b) There exists a stationary policy $f_o \in \mathbb{F}$ such that, for each $i \in X$, $f_o(i) \in A(i)$ minimizes the right-hand side of (6), and f_o is optimal, and $J(i, f_o) = j_o$, for all $i \in X$.

Remark 3.3. Lemma 3.2 holds under any of the following assumptions.

a) Assumptions given in Theorem 6.18, p. 146 in [24]. (These Assumptions support the validity of Example 2 below.)

b) Assumptions 4.2.1 and 5.5.1 given in Theorem 5.5.4, p. 97 in [9]. (In [9], average cost MDPs on Borel spaces are dealt with; this context obviously covers the case of discrete MDPs of this article.)

Remark 3.4. In Chapter 8 of [26] a detailed technique is developed which permits to obtain the optimal solution of the average cost MDPs on discrete spaces. This technique will be illustrated in Example 1 below.

4. FUZZY AVERAGE MARKOV DECISION PROCESSES

In this section the fuzzy optimal control problem is introduced. To this end, consider the following assumption.

Assumption 4.1. Let $(X, A, \{A(i) : i \in X\}, Q, C)$ be a decision model of a standard MDP for which it is supposed that there exists a stationary optimal policy f_o whose existence comes from Lemma 3.2. Let B, D, F, G, B_1, D_1, F_1, and G_1 be nonnegative real numbers such that $B < D \leq F < G$ and $B_1 < D_1 \leq F_1 < G_1$. It will be also supposed that

$$\begin{aligned} \widetilde{C}(i,a) &= (B,D,F,G)\, C(i,a) +^* (B_1, D_1, F_1, G_1) \\ &= (BC(i,a) + B_1, DC(i,a) + D_1, FC(i,a) + F_1, GC(i,a) + G_1), \end{aligned} \tag{7}$$

$i \in X$, $a \in A(i)$.

Now, the new Markov decision model $(X, A, \{A(i) : i \in X\}, Q, \widetilde{C})$ will be analyzed.

Lemma 4.2. Suppose that Assumption 4.1 holds. Take $i \in X$, $\pi \in \Pi$, and $t \geq 0$ such that $E_{i,\pi}[C(x_t, a_t)] < \infty$, and let $(\Omega, \mathcal{A}, P_{i,\pi})$ be the corresponding canonical space (see Section 3). Then, $\widetilde{C}(x_t, a_t)$ is a fuzzy random variable associated to $(\Omega, \mathcal{A}, P_{i,\pi})$, and

$$E^*_{i,\pi}[\widetilde{C}(x_t, a_t)] = E_{i,\pi}[C(x_t, a_t)](B, D, F, G) + (B_1, D_1, F_1, G_1). \tag{8}$$

P r o o f. Fix $i \in X$, $\pi \in \Pi$, $\alpha \in [0,1]$, and $t \geq 0$. Write $C_t := C(x_t, a_t)$, $C_t(\omega) := C(x_t(\omega), a_t(\omega))$, and $\widetilde{C}_t(\omega) := \widetilde{C}(x_t(\omega), a_t(\omega))$, $\omega \in \Omega$. Observe that as $\mathbb{K} \subseteq X \times A$ is denumerable (recall that X is denumerable and A is finite), it follows that C_t is a nonnegative discrete random variable with denumerable range denoted by $C_t[\Omega] = \{\lambda_1, \lambda_2, ...\}$ and consider the measurable sets $[C_t = \lambda_j] := \{\omega \in \Omega \mid C_t(\omega) = \lambda_j\}$, $j = 1, 2, ...$. Put $\Theta = (B, D, F, G)$ and $\Lambda = (B_1, D_1, F_1, G_1)$, with α-cuts $\Theta_\alpha = [\beta(\alpha), \gamma(\alpha)]$ and $\Lambda_\alpha = [\beta_1(\alpha), \gamma_1(\alpha)]$, respectively, where $\beta(\alpha) = D\alpha + (1-\alpha)B$, $\gamma(\alpha) = F\alpha + (1-\alpha)G$, $\beta_1(\alpha) = D_1\alpha + (1-\alpha)B_1$, and $\gamma_1(\alpha) = F_1\alpha + (1-\alpha)G_1$.
Consider the following multifunction given by

$$(\widetilde{C}_t)_\alpha(\omega) := (\widetilde{C}_t((\omega))_\alpha = [C_t(\omega)\beta(\alpha) + \beta_1(\alpha), C_t(\omega)\gamma(\alpha) + \gamma_1(\alpha)],$$

$\omega \in \Omega$.
Now, notice that

$$\begin{aligned} Gr((\widetilde{C}_t)_\alpha) &= \{(\omega, x) \in \Omega \times \mathbb{R} \mid x \in (\widetilde{C}_t((\omega))_\alpha\} \\ &= \{(\omega, x) \in \Omega \times \mathbb{R} \mid x \in [C_t(\omega)\beta(\alpha) + \beta_1(\alpha), C_t(\omega)\gamma(\alpha) + \gamma_1(\alpha)]\} \\ &= \bigcup_j ([C_t = \lambda_j] \times [\lambda_j\beta(\alpha) + \beta_1(\alpha), \lambda_j\gamma(\alpha) + \gamma_1(\alpha)]). \end{aligned}$$

Hence, $Gr((\widetilde{C}_t)_\alpha) \in \mathcal{A} \otimes \mathcal{B}(\mathbb{R})$. Since α is arbitrary, from Definition 2.10, it results that $\widetilde{C}_t$ is a fuzzy random variable. Next, note that, for each $\omega \in \Omega$,

$$(\widetilde{C}_t(\omega))_0 := (\widetilde{C}_t)_0(\omega) = [C_t(\omega)B + B_1, C_t(\omega)G + G_1].$$

Define $h : \Omega \longrightarrow \mathbb{R}$ given by

$$h(\omega) := C_t(\omega)G + G_1,$$

$\omega \in \Omega$. Then:

$$|x| \leq h(\omega),$$

$(\omega, x) \in \Omega \times \mathbb{R}$ with $x \in [C_t(\omega)B + B_1, C_t(\omega)G + G_1]$. Also, clearly $E_{i,\pi}[h] = GE_{i,\pi}[C_t] + G_1$ is finite. Therefore, from Definition 2.11, $\widetilde{C}_t$ is an integrably bounded fuzzy random variable with respect to $(\Omega, \mathcal{A}, P_{i,\pi})$. Now, from Definition 2.12, there is a unique fuzzy expected value $E^*_{i,\pi}[\widetilde{C}_t]$ and it is direct to verify from Definition 2.12 that,

$$(E^*_{i,\pi}[\widetilde{C}_t])_\alpha = [E_{i,\pi}[C_t]\beta(\alpha) + \beta_1(\alpha), E_{i,\pi}[C_t]\gamma(\alpha) + \gamma_1(\alpha)]$$

which is the α-cut of the trapezoidal number given for

$$E_{i,\pi}[C(x_t, a_t)](B, D, F, G) +^* (B_1, D_1, F_1, G_1),$$

that is,

$$E^*_{i,\pi}[\widetilde{C}_t] = E_{i,\pi}[C_t](B, D, F, G) +^* (B_1, D_1, F_1, G_1),$$

or equivalently, (8) holds. □

Let $(X, A, \{A(i) : i \in X\}, Q, \widetilde{C})$ be a fixed decision model which satisfies Assumption 4.1. For each policy $\pi \in \Pi$ and state $i \in X$, it is defined that

$$\widetilde{J}_n(i,\pi) := \frac{1}{n} \sum_{t=0}^{n-1}{}^{*} E^{*}_{i,\pi}[\widetilde{C}(x_t, a_t)],$$

where n is a positive integer.

Lemma 4.3. Suppose that Assumption 4.1 holds. Then, for each $i \in X$ and $\pi \in \Pi$ such that $J(i,\pi) < \infty$, $\{\widetilde{J}_n(i,\pi)\}_{n=0}^{+\infty}$ converges and

$$\begin{aligned}
\widetilde{J}(i,\pi) &:= lim^{*}_{n\to\infty} \widetilde{J}_n(i,\pi) = lim^{*}_{n\to\infty} \frac{1}{n} \sum_{t=0}^{n-1}{}^{*} E^{*}_{i,\pi}[\widetilde{C}(x_t, a_t)] \\
&= (BJ(i,\pi), DJ(i,\pi), FJ(i,\pi), GJ(i,\pi)) +^{*} (B_1, D_1, F_1, G_1) \\
&= (BJ(i,\pi) + B_1, DJ(i,\pi) + D_1, FJ(i,\pi) + F_1, GJ(i,\pi) + G_1)\,.
\end{aligned}$$

Proof. Fix $\pi \in \Pi$, $i \in X$ such that $J(i,\pi) < \infty$ and $\alpha \in [0,1]$. Observe that $J(i,\pi) < \infty$ implies that $E_{i,\pi}[C(x_t, a_t)] < \infty$, for all $t \geq 0$. Then, from Lemmas 2.5 and 4.2, it follows that $\widetilde{J}_n(i,\pi)$, $n \geq 1$ is well-defined and it is given by

$$\widetilde{J}_n(i,\pi) = (BJ_n(i,\pi) + B_1, DJ_n(i,\pi) + D_1, FJ_n(i,\pi) + F_1, GJ_n(i,\pi) + G_1),$$

and its α-cut is given by

$$\begin{aligned}
(\widetilde{J}_n(i,\pi))_\alpha &= [(DJ_n(i,\pi) + D_1)\alpha + (1-\alpha)(BJ_n(i,\pi) + B_1), (FJ_n(i,\pi) + F_1)\alpha \\
&+ (1-\alpha)(GJ_n(i,\pi) + G_1].
\end{aligned}$$

Write

$$\Gamma = (BJ(i,\pi) + B_1, DJ(i,\pi) + D_1, FJ(i,\pi) + F_1,\ GJ(i,\pi) + G_1),$$

with α-cut given by

$$\begin{aligned}
\Gamma_\alpha &= [(DJ(i,\pi) + D_1)\alpha + (1-\alpha)(BJ(i,\pi) + B_1), (FJ(i,\pi) + F_1)\alpha \\
&+ (1-\alpha)(GJ(i,\pi) + G_1)].
\end{aligned}$$

Now, it is direct to prove that

$$d(\Gamma_\alpha, (\widetilde{J}_n(i,\pi))_\alpha) = |J(i,\pi) - J_n(i,\pi)|\, max\{\alpha D + (1-\alpha)B, \alpha F + (1-\alpha)G\}.$$

Then, as α is arbitrary and $B < D \leq F < G$

$$\begin{aligned}
\hat{d}(\Gamma, \widetilde{J}_n(i,\pi)) &= sup_{\alpha\in[0,1]} d(\Gamma_\alpha, (\widetilde{J}_n(i,\pi))_\alpha) \\
&= |J(i,\pi) - J_n(i,\pi)|\, sup_{\alpha\in[0,1]}\{\alpha F + (1-\alpha)G\} \\
&\leq |J(i,\pi) - J_n(i,\pi)|\, G.
\end{aligned} \tag{9}$$

Hence, when $n \longrightarrow \infty$ in (9), it follows that

$$\lim_{n\longrightarrow\infty} \hat{d}(\Gamma, \widetilde{J}_n(i,\pi)) = 0.$$

Consequently, since π and i are arbitrary, the result follows. $\square$

For each policy $\pi \in \Pi$ and state $i \in X$ such that $J(i,\pi) < \infty$,

$$\widetilde{J}(i,\pi) = lim^*_{n\to\infty} \frac{1}{n} \sum_{t=0}^{n-1}{}^{*} E^*_{i,\pi}[\widetilde{C}(x_t,a_t)]$$

will be named the *fuzzy average cost*, when the policy π is applied, given the initial state i.

Fuzzy Optimal Control Problem (FOCP)

Therefore, the *fuzzy optimal problem* is as follows: determine $\pi_o \in \Pi$ (if it exists) such that:

$$\widetilde{J}(i,\pi_o) \leq^* \widetilde{J}(i,\pi) \tag{10}$$

for all $i \in X$, and $\pi \in \Pi$. In this case, it is possible to write

$$\widetilde{J}(i,\pi_o) = \inf_{\pi\in\Pi}{}^{*} \widetilde{J}(i,\pi), \tag{11}$$

$i \in X$, and it is said that π_o is *optimal.* Moreover, the function $\widetilde{J}_o(i) = \widetilde{J}(i,\pi_o)$, $i \in X$, will be called the *optimal fuzzy cost function.*

Remark 4.4. It is easy to see that in the (degenerate) case when in the decision model $\widetilde{C}(i,a)$ is a fuzzy quantity with a membership function given by:

$$(\widetilde{C}(i,a))'(x) = \begin{cases} 1 & \text{if } x = C(i,a) \\ 0 & \text{if } x \neq C(i,a), \end{cases}$$

for all $i \in X$ and $a \in A(i)$, then the fuzzy optimal control problem described in (10) and (11) is reduced to the optimal control problem described in (5) (See also Remark 2.8.)

Theorem 4.5. Suppose that Assumption 4.1 holds. Then, with respect to the max-order, the following statements hold.

a) The optimal policy of the fuzzy decision problem is f_o whose existence is guaranteed in Assumption 4.1.

b) The optimal fuzzy cost function is given by the fuzzy constant

$$\widetilde{J}_o = (B,D,F,G)\, j_o +^* (B_1,D_1,F_1,G_1)\,,$$

$i \in X$.

Proof. Fix $\pi \in \Pi$, $i \in X$ such that $J(i,\pi) < \infty$ and $\alpha \in [0,1]$. Now, since $J(i,f_o) \leq J(i,\pi)$, it results that

$$(DJ(i,f_o)+D_1)\alpha+(1-\alpha)(BJ(i,f_o)+B_1) \leq (DJ(i,\pi)+D_1)\alpha+(1-\alpha)(BJ(i,\pi)+B_1),$$

and

$$(FJ(i,f_o)+F_1)\alpha+(1-\alpha)(GJ(i,f_o)+G_1) \leq (FJ(i,\pi)+F_1)\alpha+(1-\alpha)(GJ(i,\pi)+G_1).$$

Since

$$[(DJ(i,f_o)+D_1)\alpha+(1-\alpha)(BJ(i,f_o)+B_1),(FJ(i,f_o)+F_1)\alpha+(1-\alpha)(GJ(i,f_o)+G_1)],$$

is the α-cut of $\widetilde{J}(i,f_o)$ and as π, i and α are arbitrary, it results that

$$\widetilde{J}(i,f_o) \leq^* \widetilde{J}(i,\pi),$$

for all π and i. Hence, f_o is optimal with respect to the max-order. Therefore, Theorem 4.5 follows from Lemmas 3.2 and 4.3.

□

Corollary 4.6. Let ε be a positive number. Suppose that Assumption 4.1 holds and that $\widetilde{C}(i,a)$ is specifically given by

$$\begin{aligned}\widetilde{C}(i,a) &= (1/2,1,1,2)\,C(i,a) +^* (\varepsilon,2\varepsilon,3\varepsilon,4\varepsilon)\\ &= \left(\frac{1}{2}C(i,a)+\varepsilon, C(i,a)+2\varepsilon, C(i,a)+3\varepsilon, 2C(i,a)+4\varepsilon\right),\end{aligned} \tag{12}$$

for all $i \in X$ and $a \in A(i)$. Then, with respect to the max-order, the optimal policy of the fuzzy decision problem is f_o whose existence is guaranteed in Assumption 4.1, and the optimal fuzzy cost function is given by the fuzzy constant

$$\widetilde{J}_o = \left(\frac{1}{2}j_o+\varepsilon, j_o+2\varepsilon, j_o+3\varepsilon, j_o+4\varepsilon\right). \tag{13}$$

Proof. Follows directly from Theorem 4.5. □

Remark 4.7. Note that (12) models the fact that "the cost is approximately in the interval $[C(i,a)+2\varepsilon, C(i,a)+3\varepsilon]$" rather than paying $C(i,a)$ as it happens in the standard model. Moreover, (13) models the fact that "the optimal cost is approximately in the interval $[j_o+2\varepsilon, j_o+3\varepsilon]$" instead of receiving j_o as in the standard MDP.

5. COMPARISON WITH THE AVERAGE RANKING

In this section the ranking comparison between trapezoidal fuzzy numbers provided in (3) will be used.

Fuzzy Optimal Control Problem with Respect to the Average Ranking (FOCPAR)

Consider the formulation of the fuzzy MDPs and Assumption 4.1 given in Section 4. So the *fuzzy optimal problem with respect to the average ranking* is as follows: determine $\pi_o^r \in \Pi$ (if it exists) such that:

$$\widetilde{J}(i, \pi_o^r) \leq^{**} \widetilde{J}(i, \pi) \tag{14}$$

for all $i \in X$, and $\pi \in \Pi$. In this case, it is possible to write

$$\widetilde{J}(i, \pi_o^r) = \inf_{\pi \in \Pi}{}^{**} \tilde{J}(i, \pi), \tag{15}$$

$i \in X$, and it is said that π_o^r is *optimal.* Moreover, the function $\widetilde{J^r}_o(i) = \widetilde{J}(i, \pi_o^r)$, $i \in X$, will be called the *optimal fuzzy cost function with respect to the average ranking.*

Remark 5.1. If in (7) (see Assumption 4.1) the degenerate case is considered in which $B = D = F = G = 1$ and $B_1 = D_1 = F_1 = G_1 = 0$, it is obtained that

$$\widetilde{C}(i, a) = C(i, a)\widetilde{1}$$

for all $i \in X$ and $a \in A(i)$, and in this situation it is easy to obtain that

$$\widetilde{J}(i, \pi) = J(i, \pi)\widetilde{1}$$

and that

$$\Re(\widetilde{J}(i, \pi)) = J(i, \pi)$$

for all $\pi \in \Pi$ and $i \in X$. It yields that the optimal control problem described in (14) and (15) is reduced to the standard optimal control problem given in (5).

Theorem 5.2. Suppose that Assumption 4.1 holds. Then, with respect to the average ranking, the following statements hold.

a) The optimal policy of the fuzzy decision problem is f_o whose existence is guaranteed in Assumption 4.1.

b) The optimal fuzzy cost function is given by the fuzzy constant

$$\widetilde{J^r}_o = (B, D, F, G)\, j_o +^* (B_1, D_1, F_1, G_1)\,,$$

$i \in X$.

P r o o f. Fix $\pi \in \Pi$, $i \in X$ such that $J(i, \pi) < \infty$. From (3) and Lemma 4.3, it is obtained that

$$\Re(\widetilde{J}(i, \pi)) = J(i, \pi)\Re((B, D, E, F)) + \Re((B_1, D_1, E_1, F_1)).$$

Now, since $J(i, \pi) \geq J(i, f_o) \geq 0$, it follows that

$$\Re(\widetilde{J}(i, \pi)) \geq J(i, f_o)\Re((B, D, E, F)) + \Re((B_1, D_1, E_1, F_1)),$$

i. e.,

$$\Re(\widetilde{J}(i, \pi)) \geq \Re((\widetilde{J}(i, f_o)).$$

Since π and i are arbitrary, it results that f_o is optimal with respect to the average ranking. Therefore, Theorem 5.2 follows from Lemmas 3.2 and 4.3. □

The next corollary is a direct consequence of Theorem 5.2.

Corollary 5.3. Let ε be a positive number. Suppose that Assumption 4.1 holds and that $\widetilde{C}(i,a)$ is specifically given by

$$\begin{aligned}\widetilde{C}(i,a) &= (1/2,1,1,2)\,C(i,a) +^{*} (\varepsilon, 2\varepsilon, 3\varepsilon, 4\varepsilon)\\ &= \left(\frac{1}{2}C(i,a)+\varepsilon, C(i,a)+2\varepsilon, C(i,a)+3\varepsilon, 2C(i,a)+4\varepsilon\right),\end{aligned}$$

for all $i \in X$ and $a \in A(i)$. Then, with respect to the average ranking, the optimal policy of the fuzzy decision problem is f_o whose existence is guaranteed in Assumption 4.1, and the optimal fuzzy cost function is given by the fuzzy constant

$$\widetilde{J^r}_o = \left(\frac{1}{2}j_o+\varepsilon, j_o+2\varepsilon, j_o+3\varepsilon, j_o+4\varepsilon\right).$$

6. EXAMPLES

Example 1: A queueing model

This example presents a version of a fuzzy queue. The standard case of this example has been addressed in different references, see for instance, [5] and [26]. This model considers a single-server with service rate control. There is a probability $0 < p < 1$ which denotes a single customer arriving, then $1-p$ represents the probability of no arrival. The queuing system can hold at most two customers, one in service and one waiting for a service. If a customer arrives to an empty system, then it may enter service immediately. Moreover, the system can be operating with one of the following two service rates: $0 < a_1 < a_2 < 1$. Under this framework, the stochastic process of interest corresponds to the number of customers in the queueing system. This model can be identified as a Markov decision model as follows: $X = \{0,1,2\}$, $A = \{a_1,a_2\}$ and with the following transition law:

$$\begin{aligned}p_{00}(a) &= 1-p+ap,\\ p_{01}(a) &= (1-a)p,\\ p_{10}(a) &= 1(1-p),\\ p_{11}(a) &= ap+(1-a)(1-p),\\ p_{10}(a) &= (1-a)p,\\ p_{21}(a) &= 1(1-p),\\ p_{22}(a) &= 1-a(1-p),\end{aligned}$$

with $a \in A$, in the other cases the transition law is zero. Moreover, it is assumed that there exists a trapezoidal fuzzy holding cost, defined by

$$\widetilde{L}(i,a) := Hi(1/2,1,1,2),$$

$(i,a) \in \mathbb{K}$, where H is a positive number, furthermore, a service cost $M(a)$ is considered for $a = a_1, a_2$, then

$$\widetilde{M}(a) = M(a)(1/2,1,1,2).$$

Therefore, the fuzzy cost function is given by

$$\widetilde{C}(i,a) = (Hi + M(a))(1/2, 1, 1, 2) + (\varepsilon, 2\varepsilon, 3\varepsilon, 4\varepsilon), \tag{16}$$

$(i, a) \in \mathbb{K}$ and $\varepsilon > 0$.

In particular, a numerical experiment is considered. To this end, assume that $H = 2$, $p = 0.6$, $a_1 = 0.3$, $a_2 = 0.5$, $M(a_1) = 3$ and $M(a_2) = 4$. To solve this numerical case, the value iteration algorithm (VIA) is considered, which is exposed in [26]:

1. Set $n = 0$, $\varepsilon > 0$, $u_0 := 0$ and let $x \in X$ be a distinguished state.
2. Set $w_n(i) := \min_a\{C(i,a) + E_i[u_n(x_1)]\}$.
3. If $n = 0$, set $\delta = 1$. If $n \geq 1$, then set $\delta = |w_n(x) - w_{n-1}(x)|$. If $\delta < \epsilon$, go to step 6.
4. Set $u_{n+1} = w_n(i) - w_n(x)$.
5. Go to step 2, and replace n by $n + 1$.
6. Print $w_n(x)$ and a stationary policy realizing $\min_a\{C(i,a) + E_i[u_n(x_1)]\}$.

Then, consider $x = 0$ as a distinguishable state, this implies that $u_n(0) = 0$ for all $n \geq 1$. Consequently, the following equations are valid:

$$\begin{aligned} w_n(0) &= 3 + \min\{0.42u_n(1), 1 + 0.3u_n(1)\}, \\ w_n(1) &= 5 + \min\{0.46u_n(1) + 0.42u_n(2), 1 + 0.5u_n(1) + 0.3u_n(2)\}, \\ w_n(2) &= 7 + \min\{0.12u_n(1) + 0.88u_n(2), 1 + 0.2u_n(1) + 0.8u_n(2)\} \end{aligned}$$

Now, an implementation of VIA in software R returns the summarized results in Table 1. In Table 1, f_1 is defined as $f_1(i) = a_1$ for all $i \in X$ and f_2 is given by

$$f_2(i) = \begin{cases} a_1 & \text{if} \quad i \in \{0, 2\} \\ a_2 & \text{if} \quad i = 1. \end{cases}$$

In summary, it is concluded from Table 1 that the average cost optimal policy for the crisp model is f_2 and the long-run expected average optimal cost is $j_o = 6.26$.

Lemma 6.1. With respect to both problems, the FOCP and the FOCPAR, the optimal policy of the fuzzy service rate control queues problem is

$$f_o(i) = \begin{cases} a_1 & \text{if} \quad i \in \{0, 2\} \\ a_2 & \text{if} \quad i = 1 \end{cases}$$

$i \in X$, and the corresponding optimal fuzzy cost function is given by

$$\widetilde{J}_o(i) = \widetilde{J^r}_o(i) = (\frac{1}{2}j_o + \varepsilon, j_o + 2\varepsilon, j_o + 3\varepsilon, j_o + 4\varepsilon),$$

$i \in X$ with $j_o = 6.26$, and $\varepsilon > 0$.

n	u_n	w_n	strategy	n	u_n	w_n	strategy
	0.00	3.00			7.52	6.16	
0	0.00	5.00	f_1	10	13.29	13.75	f_2
		7.00				19.6	
	2.00	3.84			7.64	6.2	
1	4.00	7.6	f_1	12	13.55	13.88	f_2
		10.76				19.84	
	3.76	4.57			7.70	6.24	
2	6.92	9.63	f_1	14	13.69	13.96	f_2
		13.54				19.98	
	5.05	5.12			7.74	6.25	
3	8.96	11.09	f_1	16	13.77	14.00	f_2
		15.49				20.05	
	5.96	5.5			7.75	6.26	
4	10.36	12.09	f_2	17	13.79	14.02	f_2
		16.84				20.07	
	6.58	5.76			7.75	6.26	
5	11.33	12.69	f_2	18	13.81	14.03	f_2
		17.76				20.09	

Tab. 1. VIA.

Proof. This lemma is a direct consequence of Corollaries 4.6 and 5.3. □

Example 2: A machine replacement problem

Suppose that there is a machine which is in any one of the states $0, 1, 2, \ldots$. Suppose that at the beginning of each day, the state of the machine is noted and a decision upon whether or not to replace the machine is made. If the decision to replace is made, then the fact that the machine is instantaneously replaced by a new machine whose state is 0 is assumed.

There is a non-negative cost associated with each state and available action in that state. The cost is associated with the state-action pair. The cost of replacing the machine will be denoted by $R > 0$, and furthermore, a maintenance cost $C(i)$ is incurred each day that the machine is supposed to be in state i.

Let p_{ij} represent the probability that a machine in state i at the beginning of one day will be in state j at the beginning of the next day. The action space have two actions in which action 1 is the replacement action and action 2 is the nonreplacement action.

The one-stage costs and transition probabilities are given by

$$\begin{cases} C(i,1) = R + C(0), \\ C(i,2) = C(i), & i \geq 0 \end{cases}$$

$$\begin{cases} p_{ij}(1) = p_{0j}, \\ p_{ij}(2) = p_{ij}, & i \geq 0. \end{cases}$$

Furthermore, the following assumptions on the costs and transition probabilities are imposed.

Assumption 6.2. (i) $\{C(i), i \geq 0\}$ is a bounded, increasing sequence

(ii) $\sum_{j=k}^{\infty} p_{ij}$ is an increasing function of i, for each $k \geq 0$.

Hence, (i) asserts that the maintenance cost is an increasing function of the state; and (ii) asserts an increasing failure rate, that is, the increasing failure rate assumption states that machines in better states are more likely to remain in better states than machines in worse states.

Now, let ε be a fixed positive number and suppose that

$$\widetilde{C}(i,a) = \left(\frac{1}{2}C(i,a) + \varepsilon, C(i,a) + 2\varepsilon, C(i,a) + 3\varepsilon, 2C(i,a) + 4\varepsilon\right),$$

$(i,a) \in \mathbb{K}$.

The optimal solution of this example is given in terms of the next auxiliar optimal control problem widely studied and known in the literature of MDPs as the *total discounted cost problem* (see [9, 24], and [26]). To define this problem, consider $0 < \beta < 1$. Hence, given $\pi \in \Pi$ and initial state $x_0 = i \in X$, let

$$V_\beta(i,\pi) = E_{i,\pi}\Big[\sum_{t=0}^{\infty} \beta^t C(x_t, a_t)\Big],$$

be the *expected total discounted cost* when using the strategy π, given the initial state i. And

$$V_\beta(i) = \inf_{\pi \in \Pi} V_\beta(i,\pi), \tag{17}$$

$i \in X$. V_β defined in (17) is called the *discounted optimal value function.*

Lemma 6.3. Under Assumptions 6.2, with respect to both problems, the FOCP and the FOCPAR, the optimal fuzzy cost function is given, for $\varepsilon > 0$, by

$$\widetilde{J_o} = \widetilde{J^r}_o = \left(\frac{1}{2}j_o + \varepsilon, j_o + 2\varepsilon, j_o + 3\varepsilon, j_o + 4\varepsilon\right),$$

where

$$j_o + h(i) = \min \left\{ R + C(0) + \sum_{j=0}^{\infty} p_{0j} h(j); C(i) + \sum_{j=0}^{\infty} p_{ij} h(j) \right\}, \quad i \geq 0$$

with a constant j_o and an increasing function h which are given by

$$j_o = \lim_{\beta \longrightarrow 1} (1-\beta) V_\beta(0),$$

and for a suitable sequence $\{\beta_n\} \longrightarrow 1$,

$$h(i) = \lim_{n \longrightarrow \infty} (V_{\beta_n}(i) - V_{\beta_n}(0)),$$

$i \in X$. Moreover, there exists an integer i^*, $i^* \leq \infty$ given by

$$i^* = \max \left\{ i : C(i) + \sum_{j=0}^{\infty} p_{ij} h(j) \leq R + C(0) + \sum_{j=0}^{\infty} p_{0j} h(j) \right\},$$

where the optimal policy f_o is characterized by the fact that it is replaced in all states greater than i^*.

Proof. It is a direct consequence of the examples given in [24] pp. 129 and 147, and Corollaries 4.6 and 5.3. □

7. RESEARCH IN PROGRESS

The authors are developing a theory of MDPs with different objective functions like the total discounted cost, the total cost, or the average cost, incorporating as the main assumptions the concept of generalized trapezoidal costs with both the average ranking (see [7] and [11]) or with other ranking criteria (see [4]). In fact, the concept of generalized trapezoidal fuzzy number extends Definition 2.2.

(Received July 8, 2022)

REFERENCES

[1] A. Arapostathis, V. S. Borkar, E. Fernández-Gaucherand, M. K. Gosh, and S. I. Marcus: Discrete-time controlled Markov processes with average cost criterion: a survey. SIAM J. Control Optim. *32* (1993), 2, 282–344. DOI:10.1137/0331018

[2] K. Carrero-Vera, H. Cruz-Suárez, and R. Montes-de-Oca: Discounted Markov decision processes with fuzzy rewards induced by non-fuzzy systems. In: Proc. 10th International Conference on Operations Research and Enterprise Systems ICORES 2021, pp. 49–59.

[3] K. Carrero-Vera, H. Cruz-Suárez, and R. Montes-de-Oca: Markov decision proceses on finite spaces with fuzzy total reward. Kybernetika *58* (2022), 2, 180–199. DOI:10.14736/kyb-2022-2-0180

[4] S. J. Chen and S. M. Chen: Fuzzy risk analysis based on the ranking of generalized trapezoidal fuzzy numbers. Appl. Intell. *26* (2007), 1, 1–11. DOI:10.1007/s10489-006-0003-5

[5] Y. L. Chung and Z. N. Tsai: A quantized water-filling packet scheduling scheme for downlink transmissions in LTE-advanced systems with carrier aggregation. In: SoftCOM 2010, 18th International Conference on Software, Telecommunications and Computer Networks IEEE (2010), pp. 275–279.

[6] P. Diamond and P. Kloeden: Metric Spaces of Fuzzy Sets: Theory and Applications. World Scientific, Singapore 1994.

[7] A. Ebrahimnejad: A simplified new approach for solving fuzzy transportation problems with generalized trapezoidal fuzzy numbers. Appl. Soft Comput. *19* (2014), 171–176. DOI:10.1016/j.asoc.2014.01.041

[8] N. Furukawa: Parametric orders on fuzzy numbers and their roles in fuzzy optimization problems. Optimization *40* (1997), 171–192. DOI:10.1080/02331939708844307

[9] O. Hernández-Lerma and J. B. Lasserre: Discrete-Time Markov Control Processes: Basic Optimality Criteria. Springer-Verlag, New York, 1996.

[10] M. Kageyama: Credibilistic Markov decision processes: the average case. J. Comput. Appl. Math. *224* (2009), 1, 140–145. DOI:10.1016/j.cam.2008.04.035

[11] A. Kaur and A. Kumar: A new approach for solving fuzzy transportation problems using generalized trapezoidal fuzzy numbers. Appl. Soft Comput. *12* (2012), 3, 1201–1213. DOI:10.1016/j.asoc.2011.10.014

[12] E. Konstantin, E. Avrachenkov, and E. Sanchez: Fuzzy Markov chains and decision-making. Fuzzy Optim. Decis. Making *1* (2002), 12, 143–159.

[13] M. Kurano, M. Yasuda, J. Nakagami, and Y. Yoshida: A fuzzy treatment of uncertain Markov decision processes: average case. In: Proc. ASSM2000 International Conference on Applied Stochastic System Modeling, Kyoto 2000, pp. 148–157.

[14] M. Kurano, M. Yasuda, J. Nakagami, and Y. Yoshida: Markov-type fuzzy decision processes with a discounted reward on a closed interval. Eur. J. Oper. Res. *92* (1996), 3, 649–662. DOI:10.1016/0377-2217(95)00140-9

[15] M. Kurano, M. Yasuda, J. Nakagami, and Y. Yoshida: Markov decision processes with fuzzy rewards. J. Nonlinear Convex Anal. *4* (1996), 1, 105–116.

[16] M. Kurano, M. Yasuda, J. Nakagami, and Y. Yoshida: Perceptive evaluation for the optimal discounted reward in Markov decision processes. In: International Conference on Modeling Decisions for Artificial Intelligence, Springer 2005, pp. 283–293.

[17] M. Kurano, M. Yasuda, J. Nakagami, and Y. Yoshida: A fuzzy approach to Markov decision processes with uncertain transition probabilities. Fuzzy Sets and Systems *157* (2006), 19, 2674–2682. DOI:10.1016/j.fss.2004.10.023

[18] M. López-Díaz and D. A. Ralescu: Tools for fuzzy random variables: embeddings and measurabilities. Comput. Statist. Data Anal. *51* (2006), 109–114. DOI:10.1016/j.csda.2006.04.017

[19] M. L. Puri and D. A. Ralescu: Fuzzy random variable. J. Math. Anal. Appl. *114* (1986), 402–422. DOI:10.1016/0022-247x(86)90093-4

[20] M. Puterman: Markov Decision Processes: Discrete Stochastic Dynamic Programming. Wiley, 1994.

[21] D. Rani, and T. R. Gulati : A new approach to solve unbalanced transportation problems in imprecise environment. J. Transp. Secur. *7* (2014), 3, 277–287. DOI:10.1007/s12198-014-0143-5

[22] D. Rani, T. R. Gulati, and A. Kumar : A method for unbalanced transportation problems in fuzzy environment. Sadhana *39* (2014),3, 573–581. DOI:10.1007/s12046-014-0243-8

[23] S. Rezvani and M. Molani: Representation of trapezoidal fuzzy numbers with shape function. Ann. Fuzzy Math. Inform. *8* (2014), 89–112.

[24] S. Ross: Applied Probability Models with Optimization Applications. Holden Day, 1996.

[25] A. Semmouri, M. Jourhmane, and Z. Belhallaj: Discounted Markov decision processes with fuzzy costs. Ann. Oper. Res. *295* (2020), 769–786. DOI:10.1007/s10479-020-03783-6

[26] L. Sennott: Stochastic Dynamic Programming and Control of Queueing Systems. Systems. Wiley, New York 1999.

[27] A. Syropoulos and T. Grammenos: A Modern Introduction to Fuzzy Mathematics. Wiley, New Jersey 2020.

[28] J. Wang, X. Ma, Z. Xu, and J. Zhan: Three-way multi-attribute decision making under hesitant fuzzy environments. Inform. Sci. *552* (2021), 328–351. DOI:10.1016/j.ins.2020.12.005

[29] L. Zadeh: Fuzzy sets. Inform. Control *8* (1965), 338–353. DOI:10.1016/S0019-9958(65)90241-X

Hugo Cruz-Suárez, Facultad de Ciencias Físico Matemáticas, Benemérita Universidad Autónoma de Puebla, Av. San Claudio y Río Verde, Col. San Manuel, CU, Puebla, Pue. 72570. México.
e-mail: hcs@fcfm.buap.mx

Raúl Montes-de-Oca, Departamento de Matemáticas, Universidad Autónoma Metropolitana-Iztapalapa, Av. Ferrocarril San Rafael Atlixco 186, Col. Leyes de Reforma 1^a *Sección, Alcaldía Iztapalapa, C. P. 09310, CDMX. México.*
e-mail: momr@xanum.uam.mx

R. Israel Ortega-Gutiérrez, Facultad de Ciencias Físico Matemáticas, Benemérita Universidad Autónoma de Puebla, Av. San Claudio y Río Verde, Col. San Manuel, CU, Puebla, Pue. 72570. México.
e-mail: rei.ortegag@correo.buap.mx