Scientific
Research

Discrete-Time Hybrid Decision Processes: The Discounted Case

Buheeerdun Yang[1], Pingjun Hou[1], Masayuki Kageyama[2]
[1]Graduate School of Science, Chiba University, Chiba, Japan
[2]Nagoya City University, Nagoya, Japan
Email: kageyama@sda.nagoya-cu.ac.jp

ABSTRACT

This paper is a sequel to Kageyama *et al.* [1], in which a Markov-type hybrid process has been constructed and the corresponding discounted total reward has been characterized by the recursive equation. The objective of this paper is to formulate a hybrid decision process and to give the existence and characterization of optimal policies.

Keywords: Hybrid Decision Process; Discounted Reward Criteria; Optimal Equation; Chance Space; Fixed Point Theorem

1. Introduction

The credibility theory, developed by Liu [2], is useful in dealing with uncertainty in human thinking. In the real world, we often encounter the complex problem with human thinking, which could not be treated only by probability theory. To deal with such complex problem, Li and Liu [3] have introduced a more flexible uncertain theory, called chance theory, which is a hybrid of probability and credibility. Also, recently, Kageyama *et al.* [1] have given a method of constructing a Markov-type hybrid process from stochastic kernel and credibilistic kernel. Imagining the much wider applications of hybrid processes in the near future, it is meaningful to consider the case where the behavior of hybrid processes given in Kageyama *et al.* [1] may be influenced by a suitable choice of decisions or actions. The objective of this paper is to formulate a hybrid decision process, referring a modeling of stochastic control system known as a Markov decision process (cf. [4,5]), and to give the existence and characterization of optimal policies.

In the remainder of this section, we shall establish the notation that will be used throughout the paper and recall the chance measure and hybrid variables whose expected values are defined. For any non-empty set X, a function $g : X \mapsto [0,0.5]$ is said to satisfy condition $\mathscr{K}$ with X if

$$\sup_{x \in X} g(x) = 0.5, \qquad (1.1)$$

and

$$\sup_{x \neq x^*, x \in X} g(x) = 0.5 \quad \text{if } g(x^*) = 0.5. \qquad (1.2)$$

The set of such functions will be denoted by $\mathscr{K}(X)$.

A Borel set is a Borel subset of a metric space. Let X and Y be arbitrary Borel sets. We denote by $\mathcal{B}(X)$ and $\mathscr{P}(X)$ the sets of the Borel subsets and the power set of X respectively. Let $\mathcal{P}(X)$ be the set of probability measures on X. The subset $\Lambda \subset X \times Y$ is called an event if $\Lambda(x) \in \mathcal{B}(Y)$ for all $x \in X$, where

$$\Lambda(x) = \{ y \in Y \mid (x,y) \in \Lambda \}. \qquad (1.3)$$

The set of all events will be denoted by $\mathscr{P}(X) \times \mathcal{B}(Y)$.

For any $(g,p) \in \mathscr{K}(X) \times \mathcal{P}(Y)$, a function δ on $\mathscr{P}(X) \times \mathcal{B}(Y)$ is defined by, for each $\Lambda \in \mathscr{P}(X) \times \mathcal{B}(Y)$,

$$\delta(\Lambda \mid g,p) = \sup_{x \in X} g(x) \wedge p(\Lambda(x)). \qquad (1.4)$$

where $a \wedge b = \min\{a,b\}$ for any real numbers a,b.

For any $(g,p) \in \mathscr{K}(X) \times \mathcal{P}(Y)$ the chance measure $\mathrm{Ch}(\cdot \mid g,p)$ on $X \times Y$ is given (cf. [3,6]) as follows: for any $\Lambda \in \mathscr{P}(X) \times \mathcal{B}(Y)$,

$$\mathrm{Ch}(\Lambda \mid g,p) = \begin{cases} \delta(\Lambda \mid g,p) & \text{if } \delta(\Lambda \mid g,p) < 0.5, \\ 1 - \delta(\bar{\Lambda} \mid g,p) & \text{if } \delta(\Lambda \mid g,p) = 0.5, \end{cases} \qquad (1.5)$$

where $\bar{\Lambda}$ is a complement of a set Λ. The triplet $\left\{ X \times Y, \mathscr{P}(X) \times \boldsymbol{B}(Y), \mathrm{Ch}(\cdot | g, p) \right\}$ is a chance space constructed from $(g, p) \in \mathscr{K}(X) \times P(Y)$.

The function $\tilde{r} : X \times Y \mapsto \mathbb{R}^+ := [0, \infty)$ is called a hybrid variable if the set

$\left\{ (x, y) \in X \times Y \middle| \tilde{r}(x, y) \le a \right\} \in \mathscr{P}(X) \times \mathcal{B}(Y)$ for any

$a \in \mathbb{R}^+$. We denote by $\mathscr{K}(Y | X)$ the set of all functions $\boldsymbol{q}(y | x)$ on $X \times Y$ such that $\boldsymbol{q}(\cdot | x) \in \mathscr{K}(Y)$ for each $x \in X$. The function $\boldsymbol{q} \in \mathscr{K}(Y | X)$ is called a credibilistic kernel ([7]). Let us denote by $\mathcal{P}(Y | X)$ the set of all stochastic kernel on Y given X ([8]).

In Section 2, We define a hybrid decision process by using the credibilistic kernel and stochastic kernel, which is analyzed to show the existence and functional characterization of optimal policies in Section 3. An example is given in Section 4.

The expected value $E(\tilde{r} | g, p)$ of the hybrid variable $\tilde{r} : X \times Y \to \mathbb{R}^+$ is defined by the Choquet integral:

$$E(\tilde{r} | g, p) = \int_0^\infty \mathrm{Ch}(\{\tilde{r} \ge t\} | g, p) \mathrm{d}t.$$

2. Hybrid Decision Processes

The state and parameter of some dynamic systems will be denoted respectively by points in a Borel set X and a finite set $\Theta := \{\theta_1, \theta_2, \cdots, \theta_l,\}$. Let $A := \{a_1, a_2, \cdots, a_k\}$ be a finite action space.

Let $X := \mathscr{K}(\Theta) \times \mathcal{P}(S)$, where

$$\mathscr{K}(\Theta) := \Big\{ g = \big(g(\theta_1), g(\theta_2), \cdots, g(\theta_l) \big) \Big| $$
$$g \text{ satisfies the condition } \mathscr{K} \Big\}.$$

The discrete-time Markov-type hybrid decision process with the state space X is a five-tuple $(X, A, \boldsymbol{q}, \boldsymbol{p}, \tilde{r})$, where $\boldsymbol{q}(\theta' | \theta, a) \in \mathscr{K}(\Theta | \Theta \times A)$ is a credibilistic kernel, $\boldsymbol{p}(y | x, a) \in P(Y | X \times A)$ a stochastic kernel and $\tilde{r} : \Theta \times S \times A \to \mathcal{R}^+ := [0, \infty)$ a reward function.

For any initial state $(g_0, p_0) \in X$, if the action $a \in A$ is chosen, then two things happen.

1) The state (g_0, p_0) of the system moves to the new state $(g_1, p_1) \in X$ as the following state equation:

$$g_1(\theta') = \max_{\theta \in \Theta} g_0(\theta) \wedge \boldsymbol{q}(\theta' | \theta, a) =: T_q(g_0 | a)(\theta')$$
$$\text{for any } \theta' \in \Theta; \tag{2.1}$$

$$p_1(B) = \int_X \boldsymbol{p}(B | x, a) p_0(\mathrm{d}x) =: T_p(p_0 | a)(B)$$
$$\text{for any } B \in \mathcal{B}(X). \tag{2.2}$$

2) The expected value $r(g_0, p_0, a)$ of reward function $\tilde{r}$ occurs, where

$$r(g_1, p_1, a) := \int \mathrm{Ch}\left(\left\{ (\theta, x) \middle| \tilde{r}(\theta, x, a) \ge t \right\} \middle| g_0, p_0 \right) \mathrm{d}t. \tag{2.3}$$

Once the state has been translated into the new state, a new action is chosen and the processes is repeated.

The metric on the state space X will be defined by, for any state $(g, p), (g', p') \in X$,

$$\rho\big((g, p), (g', p') \big) := \max_{\theta \in \Theta} |g(\theta) - g'(\theta)| + \|p - p'\|_{TV}, \tag{2.4}$$

where $\|p - p'\|_{TV}$ is the total variation metric which is defined by

$$\|p - p'\|_{TV} := 2 \sup_{\Lambda \in \mathcal{B}(X)} |p(B) - p'(B)|.$$

We denote by $\mathcal{B}(X)$ the Borel subsets of X generated by the ρ-metric topology.

The stationary policy $f : X \to A$ is measurable if, for any action $a \in A$, $B_a \in \mathcal{B}(X)$, where

$$B_a := \left\{ (g, p) \in X \middle| f(g, p) = a \right\} \in \mathcal{B}(X). \tag{2.5}$$

We denote by $\mathbb{F}$ the set of all measurable stationary policies.

For any initial state $g_0 = g, p_0 = p$ with $(g, p) \in X$, under the stationary policy $f \in \mathbb{F}$, we define the total discounted reward function $(0 < \beta < 1)$ by

$$\varphi_f(g, p) = \sum_{t=0}^\infty \beta^t r(g_t, p_t, f(g_t, p_t)) \tag{2.6}$$

where

$$\begin{cases} g_0 = g, p_0 = p, \\ g_t = T_q(g_{t-1} | f(g_{t-1}, p_{t-1})), \\ p_t = T_p(p_{t-1} | f(g_{t-1}, p_{t-1})), t \ge 1. \end{cases} \tag{2.7}$$

The value function $\varphi(g, p)$ on X is given as

$$\varphi(g, p) := \sup_{f \in \mathbb{F}} \varphi_f(g, p). \tag{2.8}$$

The policy $f^* \in \mathbb{F}$ is called optimal if, for any $(g, p) \in X$,

$$\varphi(g, p) = \varphi_{f^*}(g, p). \tag{2.9}$$

3. Analysis

In this section, we will utilize the method of dynamic programming (cf. [4,5,9]) to drive the discounted optimality equation, from which the existence of optimal policies is shown. As first, we show the measurability of the total discounted reward function and the value function.

Lemma 3.1. *For any stationary policy $f \in \mathbb{F}$, the function $r(g_t, p_t, f(g_t, p_t))$ is a measurable function of the initial state $(g_0, p_0) = (g, p) \in X$.*

Proof. For the case of $t = 1$, it suffices to prove that

$$\mathbb{C}_d := \left\{ (g, p) \middle| r(g_1, p_1, f(g_1, p_1)) \le d \right\} \in \mathcal{B}(X)$$

for any $d \in \mathbb{R}^+$,

where g_1 and p_1 are given by (2.1) and (2.2). From (1.4) and (1.5), for any $a \in A$,

$\mathrm{Ch}\left(\left\{(\theta, x) \in \Theta \times S \middle| \tilde{r}(\theta, x, a) \geq t\right\} \middle| g, p\right)$ is measurable in $(g, p) \in X$, so that, by (1.5), $r(g, p, a)$ is measurable in $(g, p) \in X$.

The set $\mathbb{C}_d$ is rewritten by

$$\mathbb{C}_d = \bigcup_{a \in A}\left\{(g, p)\middle| r(g_1, p_1, a) \leq d\right\}$$
$$\bigcap\left(T_q(\cdot | a), T_p(\cdot | a)\right)^{-1}\left(\mathcal{B}_a\right),$$

where $\mathcal{B}_a$ is given in (2.5). Since A is finite, together with (2.1), (2.2) and (2.5), $\mathbb{C}_d \in \mathcal{B}(X)$ follows.

For the case of $t = 0$ and $t \geq 2$, the measurability of $r(g_t, p_t, f(g_t, p_t))$ can be proved, by induction, similarly to the above. This completes the proof. □

Theorem 3.1. *For any stationary policy* $f \in \mathbb{F}$, *the discounted total reward function* $\varphi_f(g, p)$ *is measurable on the state space* X.

Proof. From Lemma 3.1 and the definition of $\varphi_f(g, p)$, the assertion holds obviously. □

Let $\bar{X}$ denote the class of all bounded measurable functions on X. For $\varphi, \varphi' \in \bar{X}$, we define the metric $\bar{\rho}$ on $\bar{X}$ by

$$\bar{\rho}(\varphi, \varphi') := \sup_{(g, p) \in X}\left|\varphi(g, p) - \varphi'(g, p)\right|. \qquad (3.1)$$

Then, it is clear that the space $(\bar{X}, \bar{\rho})$ is complete.

For any policy $f \in \mathbb{F}$, $(g, p) \in X$ and $h \in \bar{X}$, we define the operator U_f on $\bar{X}$ as follows:

$$U_f h(g, p) = r(g, p, f(g, p))$$
$$+ \beta h\left(T_q\left(g | f(g, p)\right), T_p\left(p | f(g, p)\right)\right). \qquad (3.2)$$

Lemma 3.2. *The operator* U_f *is a contraction on the space* $\bar{X}$.

Proof. For any state $(g, p) \in X$, we have

$$\left|U_f h(g, p) - U_f h'(g, p)\right|$$
$$= \beta\left|h\left(T_q\left(g | f(g, p)\right), T_p\left(p | f(g, p)\right)\right)\right.$$
$$\left. - h'\left(T_q\left(g | f(g, p)\right), T_p\left(p | f(g, p)\right)\right)\right|$$
$$\leq \beta \sup_{(g, p) \in X}\left|h(g, p) - h'(g, p)\right| = \beta\bar{\rho}(h, h').$$

Thus, we have $\bar{\rho}\left(U_f h, U_f h'\right) \leq \beta\bar{\rho}(h, h')$. This completes the proof. □

Theorem 3.2. *The discounted total value* φ_f *is a unique fixed point of the operator* U_f, *i.e.*

$$\varphi_f = U_f \varphi_f. \qquad (3.3)$$

Proof. As r is a non-negative and bounded function, there exists a $M > 0$ such that $0 \leq r \leq M$. So, we have

that $0 \leq \varphi_f \leq \dfrac{M}{1 - \beta}$ for all $(g, p) \in X$, which shows $\varphi_f \in \bar{X}$ from Theorem 3.1. Since

$$\varphi_f(g, p)$$
$$= r\left(g_0, p_0, f(g_0, p_0)\right) + \sum_{t=1}^{\infty}\beta^t r\left(g_t, p_t, f(g_t, p_t)\right)$$
$$= r\left(g_0, p_0, f(g_0, p_0)\right) + \beta\sum_{t=1}^{\infty}\beta^{t-1} r\left(g_t, p_t, f(g_t, p_t)\right)$$
$$= r\left(g_0, p_0, f(g_0, p_0)\right) + \beta\varphi_f(g_1, p_1),$$

we have $\varphi_f = U_f \varphi_f$.

From the Banach fixed point theorem, the conclusion of the theorem follows. □

In order to describe the optimality equation with respect to the value function $\varphi(g, p)$, we define the operator U on $\bar{X}$ as following:

$$Uh(g, p) = \max_{a \in A}\left\{r(g, p, a) + \alpha h\left(T_q\left(g | a\right), T_p\left(p | a\right)\right)\right\},$$
$$(g, p) \in X, \ h \in \bar{X}.$$

$$(3.4)$$

Then, we have the following.

Lemma 3.3. *The operator* U *is a contraction with modulus* α *on the space* $\bar{X}$.

Proof. From the definitions of $\bar{X}$ and U, it is obviously that U is the mapping from $\bar{X}$ to $\bar{X}$. For any $(g, p) \in X$ and any $h, h' \in \bar{X}$,

$$\left|Uh(g, p) - Uh'(g, p)\right|$$
$$\leq \alpha \max_{a \in A}\left|h\left(T_q\left(g | a\right), T_p\left(p | a\right)\right) - h'\left(T_q\left(g | a\right), T_p\left(p | a\right)\right)\right|$$
$$\leq \alpha \sup_{(g, p) \in X}\left|h(g, p) - h'(g, p)\right|$$
$$= \alpha\bar{\rho}(h, h').$$

Thus, we have $\bar{\rho}\left(Uh, Uh'\right) \leq \alpha\bar{\rho}(h, h')$, which completes the proof. □

Lemma 3.4. *The value function* φ *is a bounded and measurable function. i.e.,* $\varphi \in \bar{X}$.

Proof. For any $f \in \mathbb{F}$, we observe that $\varphi_f \in \bar{X}$ from Theorem 3.1. By Lemma 3.3, it holds that $U^n\varphi_f \to \varphi$ as $n \to \infty$ (cf. [4,5]). Since $U^n\varphi_f$ is a measurable, φ is a measurable. Noting that $0 \leq \varphi \leq M/(1 - \beta)$, we have $\varphi \in \bar{X}$, as described. □

Here, we can state the main result which shows the existence of optimal policies.

Theorem 3.3. *It holds that*

1) The value function φ is the unique fixed point of the operator U.

2) An optimal stationary policy $f^* \in \mathbb{F}$ exists.

3) The policy $f \in \mathbb{F}$ such that $U\varphi = U_f\varphi$ is the optimal policy.

Proof. For 1), we have, for any $(g, p) \in X$,

$$\varphi_f(g,p) = \sum_{t=0}^{\infty} \beta^t r(g_t, p_t, f(g,p))$$

$$= r(g_0, p_0, f(g_0, p_0)) + \sum_{t=1}^{\infty} \beta^t r(g_t, p_t, f(g_t, p_t))$$

$$\le r(g_0, p_0, f(g_0, p_0)) + \beta \varphi_f(g_1, p_1)$$

$$\le r(g_0, p_0, f(g_0, p_0)) + \beta \overline{\varphi}(g, p)$$

$$\le \max_{a \in A} \left[r(g_0, p_0, a) + \beta \overline{\varphi}(g_1, p_1) \right]$$

$$= U\overline{\varphi}(g, p).$$

Therefore, $\sup_{f \in \mathbb{F}} \varphi_f(g,p) = \overline{\varphi}(g,p) \le U\overline{\varphi}$.

Let us prove that $\overline{\varphi} \ge U\overline{\varphi}$. Because A is finite, there exists $f^* \in \mathbb{F}$ which satisfies

$$U\overline{\varphi}(g,p) = r(g, p, f^*(g,p)) + \beta \overline{\varphi}(g_1, p_1), \ (g,p) \in X.$$

From $\overline{\varphi} \le U\overline{\varphi}$, we get

$$U\overline{\varphi}(g,p) \le r(g, p, f^*(g,p)) + \beta U\overline{\varphi}(g_1, p_1),$$
$$(g,p) \in X.$$

Repeating the above inequality, we have

$$U\overline{\varphi}(g,p)$$
$$\le \sum_{t=0}^{n} \beta^t r(g_t, p_t, f^*(g_t, p_t)) + \beta^{n+1} \overline{\varphi}(g_{n+1}, p_{n+1}).$$

As $\beta^{n+1} \overline{\varphi}(g_{n+1}, p_{n+1}) \to 0$ as $n \to \infty$, we have that $U\overline{\varphi} \le \varphi_{f^*} \le \overline{\varphi}$. Thus, $\overline{\varphi} = U\overline{\varphi}$ follows.

For 2) and 3), we have been already shown in the proof of 1). □

4. The Controlled Floating Exchange Rate System

In this section, we will give an example for application of our model.

For simplicity, let us denote by (μ, τ) a normal distribution with the mean μ $(0 < \mu < \infty)$ and variance τ $(\tau > 0)$. Let S be $(-\infty, \infty)$, X_t be the amount of wealth at time t $(t \ge 0)$, whose controlled system equation will be defined by:

$$X_{t+1} = X_t + a_t Z_t, \tag{4.1}$$

where Z_t is a sequence of i.i.d. income random variables with the normal distribution $(1,1)$ and a_t is an action at time t, selected from the action space $A := \{a_0, a_1, a_2, \cdots, a_l\}$ with $a_0 = 0 < a_1 < a_2 < \cdots < a_{l-1} < a_l = 1$.

Let $\Theta = \{0.6, 0.7, 0.8, 0.9, 1.0\}$ be the possible exchange rate. Then, the real income will be defined by

$$\tilde{r}(\theta, x, a) = \theta d(x, a) \quad x \in S, a \in A, \ \theta \in \Theta, \tag{4.2}$$

where, $d(x, a)$ is a reward when the action $a \in A$ is

taken to the state x and the exchange rate $\theta \in \Theta$.

We assume that X_0 is distributed with the normal distribution (μ, τ). Then, X_t is distributed with the normal distribution $\left(\mu + \sum_{k=1}^{t} a_k, \tau + \sum_{k=1}^{t} a_k^2\right)$. So, we can restrict the state space of the corresponding hybrid decision process to the following:

$$X = \left\{ (g, (\mu, \tau)) \middle| g \in \mathcal{K}(\Theta), -\infty < \mu < \infty, \tau > 0 \right\}, \tag{4.3}$$

where $(\mu, 0)$ denotes one point distribution with $\mathbf{1}_{\{\mu\}}(x) = 1$ if $x = \mu$, $\mathbf{1}_{\{\mu\}}(x) = 0$ if otherwise.

The credibilistic kernel is given as follows

$$q = \left(q(\theta_j | \theta_i) \right) = \begin{bmatrix} 0.5 & 0.5 & 0 & 0 & 0 \\ 0 & 0.5 & 0.5 & 0 & 0 \\ 0 & 0 & 0.5 & 0.5 & 0 \\ 0 & 0 & 0 & 0.5 & 0.5 \\ 0.5 & 0 & 0 & 0 & 0.5 \end{bmatrix}. \tag{4.4}$$

Here, q is assumed to be independent of $a \in A$. The corresponding operate T_q has a fixed point $g^* = (0.5, 0.5, 0.5, 0.5, 0.5)$, *i.e.*, $g^* = T_q g^*$ (cf. [7]).

So, putting $\varphi_0(\mu, \tau) = \varphi(g^*, (\mu, \tau))$, we have the following optimal equation by Theorem 3.3,

$$\varphi_0(\mu, \tau) = \max_{a \in A} \left\{ r(g^*, (\mu, \tau), a) + \alpha \varphi_0(\mu + a, \tau + a^2) \right\}$$
$$\text{for all} -\infty < \mu < \infty, \tau \ge 0,$$
$$\tag{4.5}$$

where, $r(g, (\mu, \tau), a)$ is defined by (2.3).

It is shown in [7] that for any $g \in \mathcal{K}(\Theta)$, $T_q^{(n)}(g) \to g^*$ $(n \to \infty)$ and that there exist $n_0 \ge 1$ such that $T_q^{(n_0)}(g) = g^*$ for all $g \in \mathcal{K}(\Theta)$. Letting $G_0 = \{g^*\}$, define the sequence $\{G_k\}$ by $G_k = T_q^{-1}(G_{k-1}) = \{g \in \mathcal{K}(\Theta) | T_q(g) \in G_{k-1}\}$ $(k \ge 1)$.

Then, it clearly holds that $\bigcup_{k=0}^{n_0} G_k = \mathcal{K}(\Theta)$. Putting $\varphi_k(g, \mu, \tau) = \varphi(g, (\mu, \tau))$ for $g \in G_k$, from Theorem 3.3 we have that

$$\varphi_1(g, \mu, \tau) = \max_{a \in A} \left\{ r(g, (\mu, \tau), a) + \alpha \varphi_0(\mu + a, \tau + a^2) \right\}$$
$$\text{for } g \in G_1,$$
$$\tag{4.6}$$

and

$$\varphi_k(g, \mu, \tau) = \max_{a \in A} \left\{ r(g, (\mu, \tau), a) + \alpha \varphi_{k-1}(\mu + a, \tau + a^2) \right\}$$
$$\text{for } g \in G_k \ (k \ge 2).$$
$$\tag{4.7}$$

Noting that if $g \in G_k$ then $T_q \in G_{k-1}$ and that $\mathcal{K}(\Theta) = \bigcup_{k=0}^{n_0} G_k$, the value function $\varphi(g, (\mu, \tau))$ can be obtained recursively by (4.5)-(4.7).

REFERENCES

[1] M. Kageyama, B. Yang and P. Hou, "Discrete-Time Hybrid Processes and Discounted Total Expected Values," *Fuzzy Optimization Decision Making*, Vol. 10, No. 4, 2011, pp. 341-355. http://dx.doi.org/10.1007/s10700-011-9109-2

[2] B. Liu, "Uncertain Theory," 2nd Edition, Springer-Verlag, Berlin, 2007.

[3] X. Li and B. Liu, "Chance Measure for Hybrid Events with Fuzziness and Randomness," *Soft Computing*, Vol. 13, No. 2, 2009, pp. 105-115. http://dx.doi.org/10.1007/s00500-008-0308-x

[4] O. Hernandez-Lerma and J. B. Lasserre, "Discrete-Time Markov Control Processes, Basic Optimality Criteria," Springer, New York, 1996. http://dx.doi.org/10.1007/978-1-4612-0729-0

[5] M. Puterman, "Markov Decision Processes," John Wiley and Sons, Inc., 1994. http://dx.doi.org/10.1002/9780470316887

[6] B. Liu, "Uncertain Theory," 3rd Edition, UTLAB, 2009.

[7] M. Kageyama and K. Iwamura, "Discrete Time Credibilistic Processes: Construction and Convergences," *Information Sciences*, Vol. 179, No. 24, 2009, pp. 4277-4283.

[8] J. L. Doob, "Stochastic Processes," John Wiley, New York, 1953.

[9] R. E. Bellman, "Dynamic Programming," Princeton University Press, Princeton, 1957.