Non-cooperative n-person semi-Markov game

By
Kensuke TANAKA* and Hisafumi HOMMA*

(Received October 31, 1978)

1. Introduction

This paper is a continuation of our paper [8] in which we gave a concept of a two-person zero-sum semi-Markov game. Here, we consider a non-cooperative n-person semi-Markov game which is an extension of a two-person semi-Markov game. In the game, all players observe the present state of the system and then choose actions independently according to the full knowledge of the history of the system up to the present state. As a result of the actions and the duration time of the present state, each player gains a reward respectively and the system moves to a new state which is governed by the known conditional distribution. Then, we consider the optimization problem to maximize the limit of expected reward of each player gained during the first m transitions divided by the expected length of the first m transitions as the game proceeds to the infinite future. And, we show that the game has an equilibrium point and all players have the equilibrium stationary strategies under this criterion and some conditions.

This paper consists of four sections. In Section 2, we give the formulation of the problem treated by us in this paper. In Section 3, we show that such a game has an equilibrium point and all players have the equilibium stationary strategies. In Section 4, a sufficient condition to ensure a important assumption is given.

2. The formulation of the problem

In this paper, we define "non-cooperative n-person semi-Markov game" by a set of $(2n+3)$ objects:

$$(S, A^{(1)}, A^{(2)}, \cdots, A^{(n)}, q, F, r^{(1)}, r^{(2)}, \cdots r^{(n)}).$$

Here, S is a non-empty Borel subset of a Polish space, the set of statets of a system; each $A^{(i)}$ is a non-empty Borel subset of a Polish space, the set of actions available to player i, $i=1, 2, \cdots, n$; q is a distribution which governs the law of jump of the system, it associates Borel measurably with each

* Department of Mathematics, Niigata University, Niigata 950–21, Japan.

$$(s, a^{(1)}, a^{(2)}, \cdots, a^{(n)}) \in S \times \prod_{k=1}^{n} A^{(k)}$$

a probability measure $q(\cdot|s, a^{(1)}, a^{(2)}, \cdots, a^{(n)})$ on the Borel measurable space $(S, B(S))$, where $B(S)$ is the σ-field generated by the metric on S; $F(\cdot|s, a^{(1)}, a^{(2)}, \cdots, a^{(n)}, s')$ is a distribution of time until the transition from s to s' occurs, given that the next state is s'; $r^{(i)}$, a reward function of player i, is a bounded Borel measurable on

$$S \times \prod_{k=1}^{n} A^{(k)} \times R^+,$$

where R^+ is a non-negative real line.

At successive random times, all players observe the state of the system and classify it into one of the possible states $s \in S$. Then, according to the full knowledge of the history of the system as it has evolved to the present state s, each player i chooses independently an action $a^{(i)} \in A^{(i)}$, $i=1, 2, \cdots, n$, without collaboration with any of the others. As a consequence of the actions chosen by the players and the duration time of s, each player i gains a reward $r^{(i)}(s, a^{(1)}, a^{(2)}, \cdots, a^{(n)}, t)$ unit of money and the system jumps to a new state s' according to the distribution $q(\cdot|s, a^{(1)}, a^{(2)}, \cdots, a^{(n)})$ after some duration of state s. Then the whole process is repeated from the state s'. In this paper, our optimization problem is to maximize the limit of expected reward of each player gained during the first m transitions divided by the expected length of the first m transitions, respectively, as the game proceeds to the infinite future.

A strategy $\pi^{(i)}$ for each player i is a sequence of $\pi_1^{(i)}, \pi_2^{(i)}, \cdots$, where $\pi_m^{(i)}$ specifies the mth action to be chosen by player i by associating Borel measurably with each history $h_m=(s_1, a_1^{(1)}, a_1^{(2)}, \cdots, a_1^{(n)}, t_1, s_2, a_2^{(1)}, \cdots, a_2^{(n)}, t_2, \cdots, s_{m-1}, a_{m-1}^{(1)}, \cdots, t_{m-1}, s_m)$ of the system a probability measure $\pi_m^{(i)}(\cdot|h_m)$ on $(A^{(i)}, B(A^{(i)}))$, where s_m, $a_m^{(i)}$, $i=1, 2, \cdots, n$, and t_m are the mth state, the mth action chosen by player i, $i=1, 2, \cdots, n$, and the mth duration time, respectively. A strategy $\pi^{(i)}$ is, particularly, said to be stationary if there is a Borel measurable mapping $\mu^{(i)}$ from S into $P(A^{(i)})$, where $P(A^{(i)})$ is the set of all probability measures on $(A^{(i)}, B(A^{(i)}))$, such that $\pi_m^{(i)}=\mu^{(i)}$ for all m and in this case, $\pi^{(i)}$ is denoted by $\mu^{(i)}$. Each $\Pi^{(i)}$, $i=1, 2, \cdots, n$, denotes the class of all strategies for player i, respectively.

In order to ensure that the transitions do not take place too quickly, we need to assume the following:

Assumption 1. *There exists $\delta>0$, $\varepsilon>0$ such that for all $s \in S$ and $a^{(i)} \in A^{(i)}$, $i=1, 2, \cdots, n$,*

$$\int_S F(\delta|s, a^{(1)}, a^{(2)}, \cdots, a^{(n)}, s')dq(s'|s, a^{(1)}, a^{(2)}, \cdots, a^{(n)}) < 1-\varepsilon.$$

When the system starts in a state $s_1 \in S$ and a set of strategies $(\pi^{(1)}, \pi^{(2)}, \cdots, \pi^{(n)})$, $\pi^{(i)} \in \Pi^{(i)}$, $i=1, 2, \cdots, n$, is used, the total expected gain for each player i up to the mth transitions is defined to be

$$\phi^{(i)}(m, s_1, \pi^{(2)}, \cdots, \pi^{(n)})=$$
$$=\mathrm{E}_{\pi^{(1)}, \pi^{(2)}, \cdots, \pi^{(n)}}\Big[\sum_{j=1}^{m} r^{(i)}(s_j, a_j^{(1)}, \cdots, a_j^{(n)}, t_j)\Big]$$

and the expected average gain for each player i is defined to be

$$\phi^{(i)}(s_1, \pi^{(1)}, \pi^{(2)}, \cdots, \pi^{(n)})=$$
$$=\overline{\lim_{m\to\infty}}\frac{\phi^{(i)}(m, s_1, \pi^{(1)}, \pi^{(2)}, \cdots, \pi^{(n)})}{\mathrm{E}_{\pi^{(1)}, \pi^{(2)}, \cdots, \pi^{(n)}}\Big[\sum_{j=1}^{m} t_j \,|\, s_1\Big]}$$

Then, a set of strategies $(\pi^{(1)}, \pi^{(2)}, \cdots, \pi^{(n)})$ is called an equilibrium point if, for all i and $s\in S$,

$$\phi^{(i)}(s, \pi^{(1)}, \pi^{(2)}, \cdots, \pi^{(n)})=$$
$$=\sup_{\sigma^{(i)}\in\Pi^{(i)}}\phi^{(i)}(s, \pi^{(1)}, \cdots, \sigma^{(i)}, \cdots, \pi^{(n)})$$

and each $\pi^{(i)}$, $i=1, 2, \cdots, n$, is called an equilibrium strategy for player i, respectively.

3. Existence of equilibrium stationary strategies in the n-person semi-Markov game

In this section, we show the existence of equilibrium stationary strategies in the game. Firstly, we assume the following notations:

$$\overline{a}=(a^{(1)}, a^{(2)}, \cdots, a^{(n)})\in\prod_{k=1}^{n} A^{(k)},$$

$$\overline{\mu}=(\mu^{(1)}, \mu^{(2)}, \cdots, \mu^{(n)})\in\Big(\prod_{k=1}^{n} P(A^{(k)})\Big)^S,$$

$$\widehat{\mu}^{(i)}=(\mu^{(1)}, \mu^{(2)}, \cdots, \mu^{(i-1)}, \mu^{(i+1)}, \cdots, \mu^{(n)})\in\Big(\prod_{\substack{k=1\\k\neq i}}^{n} P(A^{(k)})\Big)^S$$

and

$$(\overline{\mu}\,;\sigma^{(i)})=(\mu^{(1)}, \cdots, \mu^{(i-1)}, \sigma^{(i)}, \mu^{(i+1)}, \cdots, \mu^{(n)})\quad\text{for}\quad \sigma^{(i)}\in(P(A^{(i)}))^S,$$

where

$$\Big(\prod_{k=1}^{n} P(A^{(k)})\Big)^S$$

denotes a set of all mappings from S into

$$\prod_{k=1}^{n} P(A^{(k)}).$$

Secondly, throughout the paper, we assume the following:

ASSUMPTIOH 2. *(i) S and each $A^{(k)}$, $k=1, 2, \cdots, n$, are compact metric space, (ii) whenever $s_m \longrightarrow s_0$ and $a_m^{(i)} \longrightarrow a_0^{(i)}$, $i=1, 2, \cdots, n$, as $m \longrightarrow \infty$, $q(\cdot \mid s_m, \overline{a}_m)$ converges to weakly to $q(\cdot \mid s_0, \overline{a}_0)$.*

ASSUMPTION 3. (i)

$$\int_0^\infty t dF(t \mid s, \overline{a}, s') = \tau(s, \overline{a}, s')$$

is a continuous function on

$$S \times \prod_{k=1}^{n} A^{(k)} \times S,$$

(ii) *For each i,*

$$\int_0^\infty r^{(i)}(s, \overline{a}, t) dF(t \mid s, \overline{a}, s') = r^{(i)}(s, \overline{a}, s')$$

is a continuous function on

$$S \times \prod_{k=1}^{n} A^{(k)} \times S.$$

From these assumptions we can prove the following lemma.

LEMMA 3. 1 *$\overline{\tau}(s, \overline{a})$ and each $\overline{r^{(i)}}(s, \overline{a})$, $i=1, 2, \cdots, n$, are continuous functions on*

$$S \times \prod_{k=1}^{n} A^{(k)},$$

where

$$\overline{\tau}(s, \overline{a}) = \int_S \tau(s, \overline{a}, s') dq(s' \mid s, \overline{a})$$

and

$$\overline{r^{(i)}}(s, \overline{a}) = \int_S r^{(i)}(s, \overline{a}, s') dq(s' \mid s, \overline{a}).$$

This proof is shown in our paper [8].

Then, from Assumption 2 (i), (ii), Assumption 3 (i) and Lemma 3. 1, we can prove that $\overline{\tau}(s, \overline{a})$ is a bounded continuous function on

$$S \times \prod_{k=1}^{n} A^{(k)}.$$

Similarly, each $\overline{r^{(i)}}(s, \overline{a})$, $i=1, 2, \cdots, n$, is a bounded continuous function on

$$S \times \prod_{k=1}^{n} A^{(k)}.$$

Since by Assumption 2 (i) each $P(A^{(k)})$ endowed with weak topology is a compact metric space,

$$\prod_{k=1}^{n} P(A^{(k)})$$

is a compact metric space. Throughout the paper, we assume that each $P(A^{(k)})$, $k=1, 2, \cdots, n$, is endowed with weak topology.

LEMMA 3.2 *Let* $u(s, \bar{a})$ *be a continuous, real-valued function on*

$$S \times \prod_{k=1}^{n} A^{(k)}.$$

Then, under Assumption 2 (i),

$$u(s, \bar{\mu}) = \int_{A^{(1)}} \cdots\cdots \int_{A^{(n)}} u(s, \bar{a}) d(\bar{\mu})$$

is continuous on

$$S \times \prod_{k=1}^{n} P(A^{(k)}),$$

where

$$d(\bar{\mu}) = \prod_{k=1}^{n} d\mu^{(k)}(a^{(k)}).$$

LEMMA 3.3 *Let u be a bounded, continuous function on* $X \times Y$, *where X is a Borel subset of a Polish space and Y is a compact metric space. Then,* $u(x) = \max_{y \in Y} u(x, y)$ *is continuous.*

LEMMA 3.4 *Let u be a bounded, continuous function on* $X \times Y$, *where X is a Borel subset of a Polish space and Y is a compact metric space. Then, there exists a Borel measurable mapping f from X into Y such that* $u(x, f(x)) = \max_{y \in Y} u(x, y)$, *for* $x \in X$.

The proofs of these lemmas are given in [2].

ASSUMPTION 4. *For each i, there exist a Borel measurable mapping* $\bar{\mu}_* \in (\prod_{k=1}^{n} P(A^{(k)}))^S$, *a continuous function* $u^{(i)}(s, \widehat{\mu_*^{(i)}})$ *and constant* $d^{(i)}(\widehat{\mu_*^{(i)}})$ *on S such that, for each s,*

$$u^{(i)}(s, \widehat{\mu_*^{(i)}}) = \max_{\sigma^{(i)} \in P(A^{(i)})} \Big\{ \bar{r}^{(i)}(s, \bar{\mu}_*; \sigma^{(i)}) + \\ + \int_S u^{(i)}(s', \widehat{\mu_*^{(i)}}) dq(s' | s, \bar{\mu}_*; \sigma^{(i)}) - d^{(i)}(\widehat{\mu_*^{(i)}}) \, \bar{\tau}(s, \bar{\mu}*; \sigma^{(i)}) \Big\} \tag{3.1}$$

and

$$u^{(i)}(s, \widehat{\mu_*^{(i)}}) = \bar{r}^{(i)}(s, \bar{\mu}_*) + \int_S u^{(i)}(s', \widehat{\mu_*^{(i)}}) dq(s' | s, \bar{\mu}_*) - \\ - d^{(i)}(\widehat{\mu_*^{(i)}}) \bar{\tau}(s, \bar{\mu}_*), \tag{3.2}$$

where, for each $\overline{\mu}\in\prod_{k=1}^{n} P(A^{(k)})$ *and* $E\in B(S)$,

$$\overline{r}^{(i)}(s,\overline{\mu})=\int_{A^{(1)}}\cdots\cdots\int_{A^{(n)}}\overline{r}^{(i)}(s,\overline{a})d(\overline{\mu}),$$

$$\overline{\tau}(s,\overline{\mu})=\int_{A^{(1)}}\cdots\cdots\int_{A^{(n)}}\overline{\tau}(s,\overline{a})d(\overline{\mu})$$

and

$$q(E|s,\overline{\mu})=\int_{A^{(1)}}\cdots\cdots\int_{A^{(n)}}q(E|s,\overline{a})d(\overline{\mu}).$$

Then, we can prove the following theorem.

THEOREM 3. 1 *The game has an equilibrium point and all players have the equilibrium stationary strategies.*

PROOF. For a set of the stationary strategies $\widehat{\mu_*}^{(i)}$ and any strategy $\sigma^{(i)}$ for player i, we have

$$\mathrm{E}_{\overline{\mu}_*;\,\sigma^{(i)}}\Big[\sum_{j=2}^{m+1}\Big\{u^{(i)}(s_j,\widehat{\mu_*}^{(i)})-\mathrm{E}_{\overline{\mu}_*;\,\sigma^{(i)}}[u^{(i)}(s_j,\widehat{\mu_*}^{(i)})\,|\,h_{j-1}]\Big\}\Big]=0. \tag{3. 3}$$

But, from (3. 1), it holds that, for each j,

$$\begin{aligned}
&\mathrm{E}_{\overline{\mu}_*;\,\sigma^{(i)}}[u^{(i)}(s_j,\widehat{\mu_*}^{(i)})\,|\,h_{j-1}]=\\
&=\int_S u^{(i)}(s',\widehat{\mu_*}^{(i)})dq(s'|s_{j-1},\overline{\mu}_*;\lambda_{j-1}^{(i)})\\
&=\Big\{\overline{r}^{(i)}(s_{j-1},\overline{\mu}_*;\lambda_{j-1}^{(i)})-d^{(i)}(\widehat{\mu_*}^{(i)})\,\overline{\tau}(s_{j-1},\overline{\mu}_*;\lambda_{j-1}^{(i)})+\\
&+\int_S u^{(i)}(s',\widehat{\mu_*}^{(i)})dq(s'|s_{j-1},\overline{\mu}_*;\lambda_{j-1}^{(i)})\Big\}-\Big\{\overline{r}^{(i)}(s_{j-1},\\
&\overline{\mu}_*;\lambda_{j-1}^{(i)})-d^{(i)}(\widehat{\mu_*}^{(i)})\,\overline{\tau}(s_{j-1},\overline{\mu}_*;\lambda_{j-1}^{(i)})\Big\}\\
&\leqq u^{(i)}(s_{j-1},\widehat{\mu_*}^{(i)})-\Big\{\overline{r}^{(i)}(s_{j-1},\overline{\mu}_*;\lambda_{j-1}^{(i)})-\\
&-d^{(i)}(\widehat{\mu_*}^{(i)})\,\overline{\tau}(s_{j-1},\overline{\mu}_*;\lambda_{j-1}^{(i)})\Big\},
\end{aligned} \tag{3. 4}$$

where $\lambda_{j-1}^{(i)}$ denotes a probability measure on $(A^{(i)}, B(A^{(i)}))$ determined by $\sigma_{j-1}^{(i)}(\cdot|h_{j-1})$. Hence, from (3. 3) and (3. 4), we have

$$\begin{aligned}
0\geqq\mathrm{E}_{\overline{\mu}_*;\,\sigma^{(i)}}&\Big[\sum_{j=2}^{m+1}\Big\{u^{(i)}(s_j,\widehat{\mu_*}^{(i)})-u^{(i)}(s_{j-1},\widehat{\mu_*}^{(i)})+\\
&+\overline{r}^{(i)}(s_{j-1},\overline{\mu}_*;\lambda_{j-1}^{(i)})-d^{(i)}(\widehat{\mu_*}^{(i)})\,\overline{\tau}(s_{j-1},\overline{\mu}_*;\lambda_{j-1}^{(i)}\Big\}\Big].
\end{aligned} \tag{3. 5}$$

From (3. 5), we get

$$d^{(i)}(\widehat{\mu_*^{(i)}}) \geqq \frac{\mathrm{E}_{\bar{\mu}_*;\,\sigma^{(i)}}\left[\sum_{j=1}^{m} r^{(i)}(s_j, a_j^{(1)}, a_j^{(2)}, \cdots, a_j^{(n)}, t_j)\right]}{\mathrm{E}_{\bar{\mu}_*;\,\sigma^{(i)}}\left[\sum_{j=1}^{m} t_j \mid s_1\right]} +$$

$$+\frac{\mathrm{E}_{\bar{\mu}_*;\,\sigma^{(i)}}\left[u^{(i)}(s_{m+1}, \widehat{\mu_*^{(i)}}) - u^{(i)}(s_1, \widehat{\mu_*^{(i)}})\right]}{\mathrm{E}_{\bar{\mu}_*;\,\sigma^{(i)}}\left[\sum_{j=1}^{m} t_j \mid s_1\right]}. \tag{3. 6}$$

By Assumption 1, it is easy to see that

$$\mathrm{E}_{\bar{\mu}_*;\,\sigma^{(i)}}\left[\sum_{j=1}^{m} t_j \mid s_1\right] \geqq m\varepsilon\delta \longrightarrow \infty \text{ as } m \longrightarrow \infty \tag{3. 7}$$

Since $u^{(i)}(s, \widehat{\mu_*^{(i)}})$ is bounded on S, from (3. 6) and (3. 7), we get, for any strategy $\sigma^{(i)}$ for player i,

$$d^{(i)}(\widehat{\mu_*^{(i)}}) \geqq \varlimsup_{m\to\infty} \frac{\mathrm{E}_{\bar{\mu}_*;\,\sigma^{(i)}}\left[\sum_{j=1}^{m} r^{(i)}(s_j, a_j^{(1)}, a_j^{(2)}, \cdots, a_j^{(n)}, t_j)\right]}{\mathrm{E}_{\bar{\mu}_*;\,\sigma^{(i)}}\left[\sum_{j=1}^{m} t_j \mid s_1\right]}. \tag{3. 8}$$

Thus, from (3. 8), it holds that for any $s_1 \in S$

$$d^{(i)}(\widehat{\mu_*^{(i)}}) \geqq \sup_{\sigma^{(i)} \in \Pi^{(i)}} \phi^{(i)}(s_1, \bar{\mu}_*; \sigma^{(i)}). \tag{3. 9}$$

Similary, from Assumption 4, it holds that for any $s_1 \in S$

$$d^{(i)}(\widehat{\mu_*^{(i)}}) = \phi^{(i)}(s_1, \bar{\mu}_*). \tag{3. 10}$$

From (3. 9) and (3. 10), a set of stationary strategies $\bar{\mu}^* = (\mu_*^{(1)}, \mu_*^{(2)}, \cdots, \mu_*^{(n)})$ is an equilibrium point and each ith element of $\bar{\mu}_*$ is an equilibrium stationary strategy for player i. Thus, the theorem is proved.

4. Sufficient condition of Assumption 4

In this section, we give the sufficient condition to ensure Assumption 4.

Let $C(S)$ denote the family of all bounded, continuous functions on S. For $u \in C(S)$ we define $\|u\| = \max_{s\in S} |u(s)|$. Then, $(C(S), d)$ is a complete metric space, where $d(v, u) = \|v-u\|$ for each $v, u \in C(S)$.

First we define an operator $T_\alpha^{(i)}: C(S) \longrightarrow C(S)$ as follows: for $\alpha > 0$

$$(T_\alpha^{(i)}u)(s) = \max_{\sigma^{(i)} \in P(A^{(i)})} \Big[\int_{A^{(1)}} \cdots\cdots \int_{A^{(n)}} e^{-\alpha\bar{\tau}(s, \bar{a})} \Big\{\overline{r}^{(i)}(s, \bar{a}) +$$

$$+\int_S u(s')dq(s'|s,\bar{a})\Big\}d(\bar{\mu};\sigma^{(i)})\Big], \tag{4.1}$$

where $d(\bar{\mu};\sigma^{(i)})=\prod_{j=1}^{i-1}d\mu^{(j)}(a^{(j)})d\sigma^{(i)}(a^{(i)})\prod_{j=i+1}^{n}d\mu^{(j)}(a^{(j)})$.

LEMMA 4.1 *The operator $T_\alpha^{(i)}$ is a contraction mapping on $C(S)$ for any $\alpha>0$ and $\bar{\mu}$.*

PROOF. It is easily proved that, for any $u, v\in C(S)$,

$$\|T_\alpha^{(i)}u-T_\alpha^{(i)}v\|$$
$$\leqq \max_{\sigma(i)\in P(A^{(i)})}\Big[\int_{A^{(1)}}\cdots\cdots\int_{A^{(n)}} e^{-\alpha\bar{\tau}(s,\bar{a})}\|u-v\|d(\mu;\bar{\sigma}^{(i)})\Big] \tag{4.2}$$

and, by Assumption 1,

$$\bar{\tau}(s,\bar{a})\geqq\delta\varepsilon. \tag{4.3}$$

From (4.2) and (4.3), we have

$$\|T_\alpha^{(i)}u-T_\alpha^{(i)}v\|\leqq e^{-\alpha\delta\varepsilon}\|u-v\|.$$

The lemma is proved.

Since $C(S)$ is a complete metric space, $T_\alpha^{(i)}$ has a unique fixed point in $C(S)$, by virtue of the Banach fixed point theorem. Let $u_\alpha^{(i)}(\widehat{\mu^{(i)}})$ be the unique fixed point of $T_\alpha^{(i)}$. Then it holds that, for each $s\in S$ and $\bar{\mu}\in(\prod_{k=1}^{n}P(A^{(k)}))^S$,

$$u_\alpha^{(i)}(s,\widehat{\mu^{(i)}})=\max_{\sigma(i)\in P(A^{(i)})}\Big[\int_{A^{(1)}}\cdots\cdots\int_{A^{(n)}} e^{-\alpha\bar{\tau}(s,\bar{a})}\Big\{r^{(i)}(s,\bar{a})$$
$$+\int_S u_\alpha^{(i)}(s',\widehat{\mu^{(i)}})\,dq(s'|s,\bar{a})\Big\}d(\bar{\mu};\sigma^{(i)})\Big], \tag{4.4}$$

where $u_\alpha^{(i)}(s,\widehat{\mu^{(i)}})$ is a value of $u_\alpha^{(i)}(\widehat{\mu^{(i)}})$ at $s\in S$.

Moreover, we can prove that the fixed point $u_\alpha^{(i)}(\widehat{\mu^{(i)}})$ is continuous in $\bar{\mu}$.

LEMMA 4.2 *If $\bar{\mu}_l(s)\in\prod_{k=1}^{n}P(A^{(k)})$ for all l and $\bar{\mu}_l(s)\Rightarrow\bar{\mu}_0(s)\in\prod_{k=1}^{n}P(A^{(k)})$ as $l\longrightarrow\infty$, it holds that, for each i,*

$$\|u_\alpha^{(i)}(\widehat{\mu_l^{(i)}})-u_\alpha^{(i)}(\widehat{\mu_0^{(i)}})\|\longrightarrow 0 \quad \text{as } l\longrightarrow\infty,$$

where the notation $\Rightarrow$ denotes weak convergence and $u_\alpha^{(i)}(\widehat{\mu^{(i)}})$ is a fixed point of $T_\alpha^{(i)}$ for $\widehat{\mu}$.

PROOF. Since $u_\alpha^{(i)}(\widehat{\mu^{(i)}})$ is the fixed point of $T_\alpha^{(i)}$ for $\bar{\mu}$, we have, for every i and s,

$$|u_\alpha^{(i)}(s,\widehat{\mu_l^{(i)}})-u_\alpha^{(i)}(s,\widehat{\mu_0^{(i)}})|$$
$$\leqq\max_{\sigma(i)\in P(A^{(i)})}\Big|\int_{A^{(1)}}\cdots\cdots\int_{A^{(n)}}e^{-\alpha\bar{\tau}(s,\bar{a})}\bar{r}^{(i)}(s,\bar{a})d(\bar{\mu}_l;\sigma^{(i)})-$$
$$-\int_{A^{(1)}}\cdots\cdots\int_{A^{(n)}}e^{-\alpha\bar{\tau}(s,\bar{a})}\bar{r}^{(i)}(s,\bar{a})d(\bar{\mu}_0;\sigma^{(i)})\Big|+$$

$$+\max_{\sigma^{(i)}\ni P(A^{(i)})}\int_{A^{(1)}}\cdots\cdots\int_{A^{(n)}} e^{-\sigma\bar{\tau}(s,\bar{a})}\int_S |u_{\alpha^{(i)}}(s',\widehat{\mu_l}^{(i)})-$$

$$-u_{\alpha^{(i)}}(s',\widehat{\mu_0}^{(i)})|dq(s'|s,\bar{a})d(\overline{\mu_l};\sigma^{(i)})+$$

$$+\max_{\sigma^{(i)}\in P(A^{(i)})}\left|\int_{A^{(1)}}\cdots\cdots\int_{A^{(n)}} e^{-\alpha\bar{\tau}(s,\bar{a})}\int_S u_{\alpha^{(i)}}(s',\widehat{\mu_0}^{(i)})dq(s'|s,\bar{a})d(\overline{\mu_l};\sigma^{(i)})\right.$$

$$\left.-\int_{A^{(1)}}\cdots\cdots\int_{A^{(n)}} e^{-\alpha\bar{\tau}(s,\bar{a})}\int_S u_{\alpha^{(i)}}(s'|s,\widehat{\mu_0}^{(i)})dq(s'|s,\bar{a})d(\overline{\mu_0};\sigma^{(i)})\right| \quad (4.5)$$

From Assumption 2, Lemma 3. 1 and Lemma 3. 2, the first term and the third term in right-hand side of (4. 5) become zero uniformly in $\sigma^{(i)}$ as $l\longrightarrow\infty$. Hence, by using the definition of norm and taking the superior limit as $l\longrightarrow\infty$, we obtain

$$\overline{\lim_{l\to\infty}}\ \|u_{\alpha^{(i)}}(\widehat{\mu_l}^{(i)})-u_{\alpha^{(i)}}(\widehat{\mu_0}^{(i)})\|$$

$$\leqq e^{-\alpha\delta\varepsilon}\ \overline{\lim_{l\to\infty}}\ \|u_{\alpha^{(i)}}(\widehat{\mu_l}^{(i)})-u_{\alpha^{(i)}}(\widehat{\mu_0}^{(i)})\|. \quad (4.6)$$

From (4. 6) it holds that, for each i,

$$\|u_{\alpha^{(i)}}(\widehat{\mu_l}^{(i)})-u_{\alpha^{(i)}}(\widehat{\mu_0}^{(i)})\|\longrightarrow 0 \text{ as } l\longrightarrow\infty,$$

because $0<e^{-\alpha\delta\varepsilon}<1$. Thus, the lemma is proved.

Next, for the fixed point $u_{\alpha^{(i)}}(s,\widehat{\mu^{(i)}})$ of $T_{\alpha^{(i)}}$ by Lemma 3. 1, Lemma 3. 2, Lemma 3. 3 and Lemma 3. 4, we can introduce the following notations: for each i and s,

$$K_{\alpha^{(i)}}(s,\bar{\mu};\sigma^{(i)})=\int_{A^{(1)}}\cdots\cdots\int_{A^{(n)}} e^{-\alpha\tau(s,\bar{a})}\Big\{\bar{r}^{(i)}(s,\bar{a})+$$

$$+\int_S u_{\alpha^{(i)}}(s',\widehat{\mu^{(i)}})dq(s'|s,\bar{a})\Big\}d(\bar{\mu};\sigma^{(i)})$$

$$\text{for}\quad \sigma^{(i)}(s)\in P(A^{(i)})$$

and

$$G_{\alpha^{(i)}}(\widehat{\mu^{(i)}})=\Big\{\lambda^{(i)};\ K_{\alpha^{(i)}}(s,\bar{\mu};\lambda^{(i)})=\max_{\sigma^{(i)}\in P(A^{(i)})}K_{\alpha^{(i)}}(s,\bar{\mu};\sigma^{(i)})\Big\}\ .$$

Then,

$$\prod_{k=1}^{n}P(A^{(k)})$$

is a compact convex set of locally convex space and $G_{\alpha^{(i)}}(\widehat{\mu^{(i)}})$ is a non-empty, closed and convex subset. So we define a mapping G_α:

$$\Big(\prod_{k=1}^{n}P(A^{(k)})\Big)^S\longrightarrow\Big(\prod_{k=1}^{n}P(A^{(k)})\Big)^S$$

as follows: for each $\overline{\mu}$

$$G_\alpha(\overline{\mu})=\left\{(\lambda^{(1)},\lambda^{(2)},\cdots,\lambda^{(n)});\ \lambda^{(i)}\in G^{(i)}(\overline{\mu}^{(i)})\ \textit{for all } i\right\}.$$

Hence, in order to apply the fixed point theorem in [1], the following lemma is important.

LEMMA 4.3 *The mapping G_α is upper semi-continuous.*

PROOF. It will be sufficient to show that, if $\overline{\lambda}_l\Rightarrow\overline{\lambda}_0$, $\overline{\mu}_l\Rightarrow\overline{\mu}_0$ as $l\longrightarrow\infty$ and $\overline{\lambda}_l\in G_\alpha(\overline{\mu}_l)$ for all l, then $\overline{\lambda}_0\in G_\alpha(\overline{\mu}_0)$. In fact, from Lemma 4. 1, we have for each s and $\overline{\lambda}_l=(\lambda_l^{(1)},\lambda_l^{(2)},\cdots\lambda_l^{(n)})\in G_\alpha(\overline{\mu}_l)$

$$u_\alpha^{(i)}(s,\widehat{\mu_l}^{(i)})=\max_{\sigma^{(i)}\in P(A^{(i)})}K_\alpha^{(i)}(s,\overline{\mu}_l;\sigma^{(i)})$$

$$=K_\alpha^{(i)}(s,\overline{\mu}_l;\lambda_l^{(i)}) \tag{4.7}$$

and, for any $\sigma^{(i)}(s)\in P(A^{(i)})$,

$$u_\alpha^{(i)}(s,\widehat{\mu_l}^{(i)})\geqq K_\alpha^{(i)}(s,\overline{\mu}_l;\sigma^{(i)}). \tag{4.8}$$

So passing to the limit and using Lemma 4. 2, (4. 7) and (4. 8) can be written as

$$u_\alpha^{(i)}(s,\widehat{\mu_0}^{(i)})=K_\alpha^{(i)}(s,\overline{\mu}_0;\lambda_0^{(i)})$$

and for any $\sigma^{(i)}(s)\in P(A^{(i)})$

$$u_\alpha^{(i)}(s,\widehat{\mu_0}^{(i)})\geqq K_\alpha^{(i)}(s,\overline{\mu}_0;\sigma^{(i)}),$$

respectively. Thus, the lemma is proved.

Then, we can conclude, from Fan's theorem in [1] and Lemma 3. 4, that there exists a Borel measurable mapping $\overline{\mu}_*=(\mu_*^{(1)},\mu_*^{(2)},\cdots,\mu_*^{(n)})$ from S into

$$\prod_{k=1}^{n}P(A^{(k)})$$

such that $\overline{\mu}_*\in G_\alpha(\overline{\mu}_*)$, that is, for all i and s,

$$u_\alpha^{(i)}(s,\widehat{\mu_*}^{(i)})=K_\alpha^{(i)}(s,\overline{\mu}_*),$$

and, for any $\sigma^{(i)}(s)\in P(A^{(i)})$,

$$u_\alpha^{(i)}(s,\widehat{\mu_*}^{(i)})\geqq K_\alpha^{(i)}(s,\overline{\mu}_*;\sigma^{(i)}).$$

Now, fix some state s_0 and let

$$f_\alpha^{(i)}(s,\widehat{\mu_*}^{(i)})=u_\alpha^{(i)}(s,\widehat{\mu_*}^{(i)})-u_\alpha^{(i)}(s_0,\widehat{\mu_*}^{(i)}). \tag{4.9}$$

From (4. 9), we get

$$u_\alpha^{(i)}(s_0,\widehat{\mu_*}^{(i)})+f_\alpha^{(i)}(s,\widehat{\mu_*}^{(i)})=$$

$$=\max_{\sigma^{(i)}\in P(A^{(i)})}\Big[\int_{A^{(1)}}\cdots\cdots\int_{A^{(n)}}e^{-\alpha\,\bar{\tau}(s,\bar{a})}\Big\{\bar{r}^{(i)}(s,\bar{a})+$$

$$+\int_S f_\alpha{}^{(i)}(s',\widehat{\mu_*}^{(i)})dq(s'\,|\,s,\bar{a})\Big\}\;d(\bar{\mu}_*;\sigma^{(i)})+$$

$$+u_\alpha{}^{(i)}(s_0,\widehat{\mu_*}^{(i)})\int_{A^{(1)}}\cdots\cdots\int_{A^{(n)}}e^{-\alpha\,\bar{\tau}(s,\bar{a})}d(\bar{\mu}_*;\sigma^{(i)})\Big]\,. \tag{4.10}$$

But

$$\int_{A^{(1)}}\cdots\cdots\int_{A^{(n)}}e^{-\alpha\,\bar{\tau}(s,\bar{a})}d(\bar{\mu}_*;\sigma^{(i)})=1-\alpha\,\bar{\tau}(s,\bar{\mu}_*;\sigma^{(i)})+$$

$$+\int_{A^{(1)}}\cdots\cdots\int_{A^{(n)}}R(\alpha)d(\bar{\mu}^*;\sigma^{(i)}). \tag{4.11}$$

Hence, from (4. 10) and (4. 11), we get

$$f_\alpha{}^{(i)}(s,\widehat{\mu_*}^{(i)})=\max_{\sigma^{(i)}\in P(A^{(i)})}\Big[\int_{A^{(1)}}\cdots\cdots\int_{A^{(n)}}e^{-\alpha\,\bar{\tau}(s,\bar{a})}\Big\{\bar{r}^{(i)}(s,\bar{a})+$$

$$+\int_S f_\alpha{}^{(i)}(s',\widehat{\mu_*}^{(i)})dq((s'\,|\,s,\bar{a})\Big\}\;d(\bar{\mu}_*;\sigma^{(i)})-$$

$$-\alpha\,u_\alpha{}^{(i)}(s_0,\widehat{\mu_*}^{(i)})\;\bar{\tau}(s,\bar{\mu}_*;\sigma^{(i)})+u_\alpha{}^{(i)}(s_0,\widehat{\mu_*}^{(i)})\,\Sigma\Big] \tag{4.12}$$

where

$$\Sigma=\int_{A^{(1)}}\cdots\cdots\int_{A^{(n)}}R(\alpha)d(\bar{\mu}_*;\sigma^{(i)})\,.$$

Similarly, we get

$$f_\alpha{}^{(i)}(s,\widehat{\mu_*}^{(i)})=\int_{A^{(1)}}\cdots\cdots\int_{A^{(n)}}e^{-\alpha\,\bar{\tau}(s,\bar{a})}\Big\{\bar{r}^{(i)}(s,\bar{a})+$$

$$+\int_S f_\alpha{}^{(i)}(s';\widehat{\mu_*}^{(i)})dq(s'\,|\,s,\bar{a})\Big\}\;d(\bar{\mu}_*)-$$

$$-\alpha\,u_\alpha{}^{(i)}(s,\widehat{\mu_*}^{(i)})\widehat{\tau}(s,\bar{\mu}_*)+u_\alpha{}^{(i)}(s,\bar{\mu}_*{}^{(i)})\Sigma^*. \tag{4.13}$$

where

$$\Sigma^*=\int_{A^{(1)}}\cdots\cdots\int_{A^{(n)}}R(\alpha)d(\bar{\mu}_*).$$

Then, we can prove the following theorem.

THEDREM 4. 1 *If* $\{f_\alpha{}^{(i)}(s,\widehat{\mu_*}^{(i)}),\ 0<\alpha<c\}$ *is a uniformly bounded, equi-continuous family of functions on S for some* $0<c<\infty$, *then, Assumption 4 holds.*

PROOF. By Ascoli-Arzela's theorem there exist a sequence $\alpha_v\longrightarrow 0$ and a continuous function $u^{(i)}(s,\widehat{\mu_*}^{(i)})$ such that $f_{\alpha_v}{}^{(i)}(s,\widehat{\mu_*}^{(i)})$ converges uniformly to $u^{(i)}(s,\widehat{\mu_*}^{(i)})$ on S.

Now we show that $\{\alpha u_\alpha^{(i)}(s_0, \widehat{\mu_*}^{(i)}), 0<\alpha<c\}$ is bounded. For the Borel measurable mapping $\overline{\mu}_*$, it holds that

$$u_\alpha^{(i)}(s, \widehat{\mu_*}^{(i)})=\int_{A^{(1)}}\cdots\cdots\int_{A^{(n)}} e^{-\alpha\,\overline{\tau}(s,\overline{a})}\Big\{\overline{r}^{(i)}(s,\overline{a})+$$

$$+\int_S u_\alpha^{(i)}(s', \widehat{\mu_*}^{(i)})dq(s'\,|\,s,\overline{a})\Big\}d(\overline{\mu}_*). \tag{4.14}$$

From (4. 14), it is easy to see that for each $s\in S$,

$$u_\alpha^{(i)}(s, \widehat{\mu_*}^{(i)})=\mathrm{E}_{\overline{\mu}_*}\left[\sum_{m=1}^{\infty} e^{-\alpha\sum_{j=1}^{m}\overline{\tau}(s_j,\overline{a}_j)}\,\overline{r}^{(i)}(s_m,\overline{a}_m)\right], \tag{4.15}$$

where

$$\overline{a}_j=(a_j^{(1)}, a_j^{(2)}, \cdots, a_j^{(n)})\in\prod_{k=1}^{n} A^{(k)}.$$

Then, since $\overline{r}^{(i)}$ is bounded on

$$S\times\prod_{k=1}^{n} A^{(k)}$$

and $|\overline{\tau}|\geqq\varepsilon\delta$, from (4. 15) $|\alpha u_\alpha^{(i)}(s, \widehat{\mu_*}^{(i)})|$ is bounded. Hence we require that $\alpha_v u_{\alpha_v}^{(i)}(s_0, \widehat{\mu_*}^{(i)})$ converges to $d^{(i)}(\widehat{\mu_*}^{(i)})$ as $\alpha_v\longrightarrow 0$ and we can show that $u_{\alpha_v}^{(i)}(s_0, \widehat{\mu_*}^{(i)})\sum$ converges uniformly to zero in $(\overline{\mu}_*; \sigma^{(i)})$ as $\alpha_v\longrightarrow 0$. Thus, from (4. 12), we get (3. 1). Similarly, from (4. 13), we can obtain (3. 2). Thus, the thorem is proved.

References

[1] FAN, K.. *Fixed-point and minimax theorems in locally convex topological linear space*, Proc. Nat Acad. Sci. USA., 88 (1952), 121–126.

[2] MAITRA, A. and PARTHASARATHY, T., *On stochastic games*, J. Opti. Theory and Appl., 5 (1970), 289–300.

[3] NASH, J. F., *Equilibrium points in n-person games*, Proc. Nat. Acad. Sci. USA. 36 (1950), 48–49.

[4] PARTHASARATHY, T., *Discounted, positive and non-cooperative stochastic games*, Inter. J. Game Theory, 2 (1973), 25–37.

[5] ROSS, M., *Average cost semi-Markov decision processes*, J. Appli. Prob., 7 (1970), 649–656.

[6] TAKAHASHI, M., *Equilibrium points of stochastic non-cooperative n-person games*, J. Sci. Hiroshima Univ., 28 (1964), 95–99.

[7] TANAKA, K., IWASE, S. and WAKUTA, K., *On Markov games with the expected average reward criterion*, Sci. Rep. Niigata Univ., 13 (1976), 31–41.

[8] TANAKA, K. and WAKUTA, K., *On semi-Markov games*, Sci. Rep. Niigata Univ., 13 (1976), 55–64.

[9] TANAKA, K. and WAKUTA, K., *On continuous time Markov games with the expected average reward criterion*, Sci. Rep. Niigata Univ., 14 (1977), 15–24.

[10] TANAKA, K. and WAKUTA, K., *On continuous time Markov games with countable state space*, J. OR. Japan, 21 (1978), 17–28.

[11] TANAKA, K. and HOMMA, H., *Continuous time non-cooperative n-person Markov games*, Bull. Math. Statist., 18 (1978), 93–105.