УДК 519.8

О методе «неполных наказаний» в неантагонистических позиционных дифференциальных играх двух лиц

© А. Н. Джафаров[1]

Аннотация. В работе получены почти совпадающие необходимые и достаточные условия существования в неантагонистической позиционной дифференциальной игре равновесных по Нэшу ситуаций. Доказательство основано на построении равновесных позиционных стратегий с помощью стратегий неполного наказания.

Ключевые слова: дифференциальная игра, ситуация равновесия по Нэшу, стратегия неполного наказания, множество Нэша.

1. Введение

Известно, что анализ горизонтальных взаимодействий элементов ДИС (Динамических иерархических систем) и синтез механизмов этих взаимодействий может быть сведена к решению бескоалиционной позиционной дифференциальной игры многих лиц. При этом в качестве решения указанной игры можно принять ситуацию равновесия по Нэщу, которая достаточно адекватна принятому условию «равноправности» находящихся на одном уровне иерархии элементов. В качестве конкретного примера возможного приложения развиваемой теории можно привести теоретико-игровую модель развития экономики с учетом финансовых взаимоотношений, учитывающую взаимодействия равноправных элементов нижнего уровня в двухуровневой ИС [1].

Дифференциальная игра многих лиц называется позиционной, если выбор управляющего вектора каждого игрока осуществляется по принципу обратной связи, т.е., если при вычислении управляющих воздействий в нужный для игроков момент времени игроки опираются лишь на сведения о реализовавшейся в этот момент времени позиции игры. Основу теории антагонистических позиционных дифференциальных игр составляют работы, выполненные в Свердловске. В ВЦ АН СССР развита теория неантагонистических позиционных дифференциальных игр, которая опирается на результаты этих работ. Этот подход позволяет описать практически все множество решений неантагонистической позиционной дифференциальной игры двух лиц. При этом предложена схема исследования неантагонистических позиционных дифференциальных игр двух лиц, заключающаяся в анализе множества взаимовыгодных траекторий и применении (угрозы применения) стратегий наказания в случае отклонения одного из игроков от намеченной траектории. Такой подход, при распространении на дифференциальные игры 3-х и более игроков встречается ряд трудностей, принципиальная из которых заключается в неопределенности, связанной с необходимостью применения (угрозы применения) стратегий наказания. Необходимо отметить, что под множеством решений указанных игр подразумевается множество равновесных (ε-равновесных) по Нэшу наборов позиционных стратегий. Такой подход к решению бескоалиционных позиционных дифференциальных игр многих лиц является одним из распространенных в последнее время подходов. Таким образом вислучае $N > 2$, где N-число игроков, наблюдая отклонение текущей позиции от намеченной, игроки

[1]Преподаватель Сумгаитского филиала Азербайджанского Института Учителей, г.Сумгаит; ali19562001@mail.ru.

не могут определить, по чьей вине произошло это отклонение. Но даже, если кто-то из игроков в состоянии определить отклонившегося игрока (по определению равновесия по Нэшу, в каждый момент времени может отклониться один и только один игрок), то в силу бескоалиционной игры, правила которой запрещают любые взаимоотношения игроков как до начала, так и в процессе игры вплоть до ее окончания, такая информированность об отклонившемся игроке, как правило, недостаточна для эффективного применения стратегий наказания.

В то же время [3-11] из самого наличия целого семейства равновесных (по Нэшу) позиционных стратегий следует, что точки равновесия не могут являться, строго говоря, бескоалиционным (несогласованным) решением позиционной дифференциальной игры нескольких лиц. Стало быть, бескоалиционность позиционной дифференциальной игры многих лиц выражается в изолированных действиях игроков в процессе игры вплоть до ее окончания. До начала же игры игрокам разрешается, строго говоря, необходимо, договариваться о выборе определенной траектории из целого семейства таких траекторий.

2. Постановка задачи

Итак, рассмотрим следующую игру:

$$\dot{x} = f\left(t, x, u, v\right), \quad t \in [t_0, T] \tag{2.1}$$

$$x\left(t_0\right) = x_0 \tag{2.2}$$

$$u \in P, \quad v \in Q, \tag{2.3}$$

$$J_1\left(u, v\right) = g_1\left(x\left(T\right)\right), \tag{2.4}$$

$$J_2\left(u, v\right) = g_2\left(x\left(T\right)\right), \tag{2.5}$$

где $x - n$-мерной вектор состояния, t-время, t_0 и T-фиксированные моменты времени, характеризующие моменты начала и окончания игры, $f\left(t, x, u, v\right)$-заданная функция, отражающая динамические свойства системы (2.1)-(2.5). u, v-векторы управляющих воздействий (управления) игроков, стесненные ограничениями (2.3), где P, Q-некоторые непустые компакты в евклидовых пространствах R^n. Интересы всех игроков полностью выражаются в их стремлении к максимизации своих функций выигрыша (2.4) и (2.5). Предполагается, что игроки имеют абсолютно точную информацию о параметрах системы (2.1)-(2.5) до начала и в течение всей игры вплоть до ее окончания, а также о реализуемых позициях во время игры, т.е. каждому игроку известна текущая позиция системы. В каждый момент времени выбор управляющего вектора осуществляется на основе знания позиции $\left(t, x\left(t\right)\right)$, т.е. по принципу обратной связи. Иначе говоря, при вычислении управляющих воздействий $u\left(t_i\right)$ и $v(t_j)$, $i, j = 1, 2, ...$ в нужный для игроков момент времени $t_i, t_j \in [t_0, T]$ на основе так или иначе выбранных законов управления игроки опираются лишь на сведения о реализовавшейся в момент времени $t = t_j$ позиции игры $\left(t_j, x\left(t_j\right)\right)$. Кроме того, предполагается, что игроки действуют изолированно друг от друга в течений всей игры вплоть до ее окончания, в чем и выражается бескоалиционность дифференциальной игры (2.1)-(2.5). Функции g_j предполагаются непрерывными, а вектор-функция $f\left(t, x, u, v\right)$ удовлетворяет стандартным требованиям, обеспечивающим существование, единственность и продолжимость решения задачи Коши (2.1), (2.2) вплоть до момента окончания игры - T.

К классу допустимых стратегий игроков отнесем множества программных U, V и позиционных U, ϑ стратегий, элементы которых является измеримыми по Борелю по всем своим аргументам функциями $u_t = u(t), v_t = v(t),\ \ t \in [t_0, T]$ и $u_{xt} = u(t, x(t))$, $v_{xt} = v(t, x(t)),\ \ t \in [t_0, T],\ \ x \in R^n$ соответственно удовлетворяющими ограничениям (2.3).

Следуя [1-10] через $X^P[t_0, x_0, u_{xt}, V]$ и $X^P[t_0, x_0, U, v_{xt}]$ соответственно обозначим множество всех движений, порожденных позиционной стратегией u_{xt} при любых реализациях управлений второго игрока и позиционной стратегией vxt при любых реализациях управлений первого игрока. Введем функционалы $l_1 : [t_0, T] \times R^n \times \vartheta \to R$ и $l_2 : [t_0, T] \times R^n \times U \to R$ следующим образом: $\forall ((\tau, x_\tau) \in [t_0, T] \times R^n \& u_{xt} \in U \& v_{xt} \in \vartheta)$ положим $l_1(\tau, x_\tau, v_{xt}) = \max\limits_{X[\tau, x_\tau, U, v_{xt}]} g_1(x(T))$, $l_2(\tau, x_\tau, u_{xt}) = \max\limits_{X[\tau, x_\tau, u_{xt}, V]} g_2(x(T))$.

Для любых начальных позиций $(\tau, x_\tau) \in [t_0, T] \times R^n$ очевидны равенства

$$\inf_{u_{xt} \in U} l_2(\tau, x_\tau, u_{xt}) = L_2(\tau, x_\tau), \quad \inf_{v_{xt} \in \vartheta} l_1(\tau, x_\tau, v_{xt}) = L_1(\tau, x_\tau), \tag{2.6}$$

$$l_2\left(\tau, x_\tau, u_{xt}^{\Pi}\right) = L_2(\tau, x_\tau), \quad l_1\left(\tau, x_\tau, v_{xt}^{\Pi}\right) = L_1(\tau, x_\tau), \tag{2.7}$$

где $L_i(\cdot),\ \ i = 1, 2$ минимаксные выигрыши игроков в соответствующих вспомогательных антагонистических играх (см. [1-10]), а u_{xt}^{Π} и v_{xt}^{Π} стратегии полного наказания, которые, по определению, при любых $(\tau, x_\tau) \in [t_0, T] \times R^n$ удовлетворяют равенствам $L_1(\tau, x_\tau) = \max\limits_{X\left[\tau, x_\tau, U, v_{xt}^{\Pi}\right]} g_1(x(T))$, $L_2(\tau, x_\tau) = \max\limits_{X\left[\tau, x_\tau, u_{xt}^{\Pi}, V\right]} g_2(x(T))$.

Через $X^P[t_0, x_0, U, V]$ обозначим множество решений задачи Коши (2.1), (2.2), порожденных всеми подмножества этого множества

$$D(u_{xt}, v_{xt}) = \Big\{ x_t \in X^p[t_0, x_0, U, V] / g_1(x(t)) > l_1(t, x(t), v_{xt}),$$

$$g_2(x(T)) > l_2(t, x(t), u_{xt})\ \ \forall t \in [t_0, T) \Big\},$$

$$\widehat{D}(u_{xt}, v_{xt}) = \Big\{ x_t \in X^p[t_0, x_0, U, V] / g_1(x(t)) \geq l_1(t, x(t), v_{xt}),$$

$$g_2(x(T)) \geq l_2(t, x(t), u_{xt})\ \ \forall t \in [t_0, T] \Big\}.$$

Введем еще множества

$$\Omega = \{(u_{xt}, v_{xt}) \in U \times \vartheta \,| D(u_{xt}, v_{xt}) \neq \text{Ш}\}, \hat{\Omega} = \left\{ (u_{xt}, v_{xt}) \in U \times \vartheta \,\middle|\, \hat{D}(u_{xt}, v_{xt}) \neq \emptyset \right\}.$$

Содержательный смысл этих множеств очевиден. Например, множество Ω состоит из всех пар позиционных стратегий таких, что существует траектория x_t, порожденная программными стратегиями, для которой в каждый момент времени $t \in [t_0, T)$ значения функционалов $l_i(t, x(t), \cdot)$ (соответствующих фиксированной паре из Ω) на этой траектории соответственно меньше значений этих же функционалов в момент окончания игры T.

О п р е д е л е н и е 2.1. *Набор позиционных стратегий* (u_{xt}^0, v_{xt}^0) *образует в игре* (2.1)-(2.5) *ситуацию равновесия по Нэщу, если:*

a) *набор* (u_{xt}^0, v_{xt}^0) *порождает единственное решение* x_t^0 *задачи Коши* (2.1), (2.2), *которое является и единственным движением, т.е.* $\forall\ t \in [t_0, T]$ *имеет место равенство*

$$\left\{ x^0(\tau),\ \ \tau \in [t, T] \right\} = X\left[t, x^0(t),\ u_{xt}^0, v_{xt}^0\right];$$

b) $\forall\ t \in [t_0, T]$ *выполняются равенства*

$$\max_{X[t, x^0(t), U, v_{xt}^0]} g_1(x(T)) = g_1\left(x^0(T)\right),$$

$$\max_{X[t,x^0(t),u^0_{xt},V]} g_2(x(T)) = g_2\left(x^0(T)\right)$$

О п р е д е л е н и е 2.2. *В игре* (2.1)-(2.4) *существует ситуация* ε*-равновесия по Нэщу, если* $\forall\ \varepsilon > 0$ *найдется набор позиционных стратегий* ($u^\varepsilon_{xt}, v^\varepsilon_{xt}$) *такой, что:*

a) существует единственное решение x^ε_t *задачи Коши* (2.1), (2.2), *которое является и единственным движением, т.е. при всех* $t \in [t_0, T]$ *выполняется равенство*

$$\{x^\varepsilon(\tau),\quad \tau \in [t, T]\} = X\left[t, x^\varepsilon(t), u^\varepsilon_{xt}, v^\varepsilon_{xt}\right];$$

b) $\forall t \in [t_0, T]$ *справедливы неравенства*

$$\max_{X[t,x(t),u^\varepsilon_{xt},V]} g_2(x(T)) \le g_2(x^\varepsilon(T)) + \varepsilon$$

$$\max_{X[t,x^\varepsilon(t),U,v^\varepsilon_{xt}]} g_1(x(T)) \le g_1(x^\varepsilon(T)) + \varepsilon$$

Поскольку изложение полученных результатов проводится в форме описания множества равновесных по Нэщу решений игры (2.1)-(2.5), кажется целесообразным введение самого термина «множество Нэша». Так, множество всех равновесных в смысле определения 1 решений игры (2.1)-(2.5) назовем множеством Нэша, множество всех ε-равновесных в смысле определения 2 решений-ε-множеством Нэша.

3. Результаты

Т е о р е м а 3.1. *Пусть выполняется соотношение* $\Omega \ne \emptyset$.

Тогда в игре (2.1)-(2.5) *существует ситуация равновесия по Нэшу на классе позиционных стратегий.*

Доказательство следует из равенств (2.6)-(2.7) и аналогичной теоремы из [2,3].

Т е о р е м а 3.2. *Пусть в игре* (2.1)-(2.5) *существует ситуация равновесия по Нэшу на классе позиционных стратегий. Тогда имеет место неравенство* $\hat{\Omega} \ne \emptyset$.

Доказательство непосредственно следует из определения.

Т е о р е м а 3.3. *Неравенства* $\hat{\Omega} \ne \emptyset$ *необходимо и достаточно для существования в игре* (2.1)-(2.5) *ситуации* ε*- равновесия по Нэшу в позиционных стратегиях.*

Обозначив через U^n множество всех пар стратегий полного наказания, введем следующее

О п р е д е л е н и е 3.1. *Элемент непустого множества* $\hat{\Omega}\backslash U^n$ *назовем парой стратегий неполного наказания.*

Стратегии полного и неполного наказания объединим под общим названием стратегий наказания.

С л е д с т в и е 3.1. *В игре* (2.1)-(2.5) *существует ситуация* ε*- равновесия тогда и только тогда, когда существует пара стратегий наказания.*

Таким образом, понятия «пара стратегий наказания» и «ситуация ε-равновесия» в известном смысле эквивалентны. Более того, любая пара равновесных стратегий суть пара стратегий наказания. Обратное, вообще говоря, не верно (см. Пример 1.1 из [6]).

В заключение следует отметить вытекающий из данной работы вывод о том, что оказывается, предложенная в ВЦ АН СССР теория неантагонистических позиционных дифференциальных игр обходится и без основ теории антагонистических позиционных дифференциальных игр.

Список литературы

1. Горелик В.А., Кононенко А.Ф. Теоретико-игровые модели принятия решений в эколого-экономических системах. – М.: Радио и связь, 1982, – 144 с.

2. Джафаров А.Н. О формализации и редукции коалиционных игр с объединением ресурсов // Докл. АН Азерб. ССР. – 1985. – Т. 41. – С. 7-9.

3. Джафаров А.Н., Кононенко А.Ф. Об одном методе редукции дифференциальных игр, I // Изв. АН Азерб. ССР, Сер. физ.-техн. и мат. наук. – 1985. – Т. 41, №1. – С. 103-106.

4. Джафаров А.Н., Кононенко А.Ф. Об одном методе редукции дифференциальных игр, I // Изв. АН Азерб. ССР, Сер. физ.-техн. и мат. наук. – 1985. – Т. 41, №2. – С. 100-103.

5. Джафаров А.Н. Об одном семействе достаточных условий существования равновесных позиционных стратегий в дифференциальных играх нескольких игроков // Тр. конф. молод. Уч. АН Азерб. ССР, Баку. – 1987. – С. 19-20.

6. Джафаров А.Н. Структура и свойства множества Нэша дифференциальных игр N лиц / Автореф. дисс. на соиск. уч. ст. кан. физ.-мат. наук, – М.: М. ВЦ АН СССР, 1988.

7. Джафаров А.Н. О стратегиях «Неполного наказания» в неантагонистических дифференциальных играх двух лиц // Кибернетика. – 1989. – №1 – С. 134-137.

8. Джафаров А.Н. Некоторые условия существования равновесных по Нэшу позиционных стратегий в бескоалиционных линейно-квадратичных играх многих лиц // Изв. ВУЗ. Математика. – 1991. – №4 – С. 80-82.

9. Джафаров А.Н. Об одном классе бескоалиционных дифференциальных игр трёх лиц // Автоматика и Телемеханика. – 1992. – №1 – С. 24-29.

10. Djafarov A. N. On a reduction of non-antagonistic positional differentional games to a pair of parametric extremal problems. Multiple criteria and game problems under uncertainty. – The 4th International Workshop (8-14 September, 1996), Moscow. – 1996. – p. 40.

11. Джафаров А.Н. Редукция неантагонистических позиционных дифференциальных игр к параметрическим задачам оптимизации. Спектральная теория и ее приложения / Тезисы Международной конференции, посвященной 80-летнему юбилею академика Ф. Г. Максудова. Baku. – 2010. – С. 158-159.

About the method of «incomplete punishments» in nonantagonistic item differential games of two persons.

© A. N. Djafarov[2]

Abstract. In the work almost conterminous necessary and sufficient living conditions in nonantagonistic item differential game of equilibrium situations on Nash are obtained. The proof is based on construction of equilibrium item strategy by means of strategy of incomplete punishment.

Key Words: differential game, a balance situation on Nash, strategy of incomplete punishment, Nash's set.

[2]The teacher of Sumgait branch of the Azerbaijan Institute of Teachers, Sumgait; ali19562001@mail.ru.