ТРУДЫ ИНСТИТУТА МАТЕМАТИКИ И МЕХАНИКИ УрО РАН

Том 23 № 4 2017

УДК 517.977

ПРИМЕНЕНИЕ АЛЬТРУИСТИЧЕСКОГО И АГРЕССИВНОГО ТИПОВ ПОВЕДЕНИЯ В НЕАНТАГОНИСТИЧЕСКОЙ ПОЗИЦИОННОЙ ДИФФЕРЕНЦИАЛЬНОЙ ИГРЕ ДВУХ ЛИЦ НА ПЛОСКОСТИ

А. Ф. Клейменов[1]

Исследуется неантагонистическая позиционная дифференциальная игра двух лиц, в которой каждый из двух игроков помимо обычного, нормального (nor), типа поведения, ориентированного на максимизацию собственного функционала, может использовать другие типы поведения. В частности, это альтруистический (alt), агрессивный (agg) и парадоксальный (par) типы. Предполагается, что по ходу игры игроки могут осуществлять переключения своего поведения с одного типа на другой. В рассматриваемой игре каждый игрок одновременно с выбором позиционной стратегии выбирает также индикаторную функцию, определенную на всем отрезке игры и принимающую значения в множестве $\{nor, alt, agg, par\}$. Индикаторная функция игрока показывает динамику изменения типа поведения, которой придерживается этот игрок. Вводится понятие BT-решения такой игры. На BT-решениях использование игроками типов поведения, отличных от нормального, приводит к исходам, более предпочтительным для них, чем в игре только с нормальным типом поведения. На примере игры с динамикой простого движения на плоскости и фазовыми ограничениями иллюстрируется процедура построения BT-решений.

Ключевые слова: неантагонистическая позиционная дифференциальная игра, терминальные показатели качества, типы поведения игроков, альтруистический и агрессивный типы поведения, решения нэшевского типа.

A. F. Kleimenov. Application of the altruistic and aggressive types of behavior in a two-person non-zero-sum positional differential game on the plane.

A two-person non-zero-sum positional differential game is studied. In addition to the normal (nor) behavior aimed at maximizing the cost functional, each player can use other types of behavior, in particular, the altruistic (alt), aggressive (agg), and paradoxical (par) types. It is assumed that the players can switch between the types of behavior in the course of the game. Simultaneously with the choice of a positional strategy, each player also chooses an indicator function defined on the entire game interval and taking values in the set $\{nor, alt, agg, par\}$. The indicator function of each player shows the dynamics of behavior switchings of this player. The notion of BT-solution of the game is introduced. On BT-solutions, the application by a player of the behavior types different from normal produces a more favorable result for this player than a game with the normal behavior only. The procedure for constructing BT-solutions is exemplified by a planar game with simple motion dynamics and state constraints.

Keywords: non-zero-sum positional differential game, terminal cost functional, behavior types of the players, altruistic and aggressive types, Nash solution.

MSC: 20D10, 20D25

DOI: 10.21538/0134-4889-2017-23-4-181-191

Введение

В настоящей работе неантагонистическая позиционная дифференциальная игра двух лиц задается расширенным набором, состоящим из семи элементов. Первые три элемента описывают нормальную форму игры (см., например, [1; 2]), остальные доставляют дополнительную информацию, касающуюся, в частности, последовательности действий игроков, а также их возможностей (в одиночку или в составе коалиции) отклоняться от согласованного набора стратегий. В работе делается акцент на случае, когда каждый из двух игроков помимо обычного, *нормального* (nor), типа поведения, ориентированного на максимизацию собственного

[1] *25 лет назад* вышла первая в нашем журнале статья Анатолия Федоровича:

Клейменов А. Ф. Универсальное решение в неантагонистической позиционной игре с векторными критериями // Тр. Ин-та математики и механики УрО РАН. 1992. Т. 1. С. 97–105.

функционала, может использовать другие типы поведения, введенные в [3; 4], такие как *альтруистический* (alt), *агрессивный* (agg) и *парадоксальный* (par) типы. Допускается, что по ходу игры игроки могут осуществлять переключения своего поведения с одного типа на другой. Предполагается, что в игре каждый игрок одновременно с выбором позиционной стратегии выбирает также индикаторную функцию, определенную на всем отрезке игры и принимающую значения в множестве $\{nor, alt, agg, par\}$. Индикаторная функция игрока показывает динамику изменения типа поведения, которой придерживается этот игрок. Для каждой пары типов поведения игроков вводятся правила формирования управлений. Формализация позиционных стратегий в игре основана на формализации и результатах общей теории позиционных дифференциальных игр [5–7]. Вводится понятие BT-решения. Предложен пример игры с динамикой простого движения на плоскости и фазовым ограничением в двух вариантах. В первом варианте допускаем, что первый и второй игроки могут проявлять альтруизм по отношению к своему партнеру в течение некоторых промежутков времени. Во втором варианте помимо допущения об альтруизме игроков дополнительно допускаем, что каждый игрок может проявлять агрессию по отношению к другому игроку в течение некоторых промежутков времени, причем допускается случай взаимной агрессии. В обоих вариантах описаны множества BT-решений. Подробно приведены BT-решения, наилучшие для каждого из игроков. Настоящая статья является продолжением работ [8–10].

1. Расширенный набор элементов неантагонистической дифференциальной игры

Рассмотрим дифференциальную игру, заданную совокупностью следующих семи элементов 1°–7° [7]:

1°. Уравнения динамики, начальные условия, множество игроков — участников игры, ограничения на управления игроков.

2°. Формализация стратегий игроков и движений, порожденных стратегиями.

3°. Функционалы выигрыша игроков.

Отметим, что элементы 1°–3° описывают нормальную форму игры [1; 2]. Однако, как правило, для адекватного определения понятия решения игры требуется дополнительная информация об игре. В частности, эта информация может касаться возможностей отдельных игроков (или коалиций игроков) отклоняться от решения игры в те или иные моменты времени игры. Может также представить интерес характер реакции неотлонившихся игроков на состоявшееся отклонение. Далее, важна последовательность, в которой игроки осуществляют выбор своих стратегий.

Примем следующее определение.

О п р е д е л е н и е 1. Будем говорить, что имеет место отклонение какого-либо игрока (или коалиции игроков) от исходного набора стратегий, начиная с некоторого момента времени, если этот игрок (или коалиция игроков), начиная с этого момента, переключается на стратегию, отличную от стратегии в исходном наборе.

Перейдем теперь к описанию дополнительных элементов игры (элементы 4°–7°).

4°. Описание множества разрешенных коалиций отклонения (РКО) (т. е. коалиций, которым разрешается отклоняться) и для каждой РКО – множества разрешенных моментов отклонения (РМО).

5°. Описание реакции игроков, не вошедших в отклонившуюся коалицию, на отклонение этой коалиции.

Набор стратегий назовем *допустимым*, если ни одной РКО ни в один РМО невыгодно отклоняться от этого набора. Считаем, что отклонение игрока (или коалиции игроков) выгодно для него (или для нее), если отклонившийся игрок (или каждый член отклонившейся коалиции) получает в результате отклонения гарантированный выигрыш, больший, чем в игре без отклонений. Гарантированный выигрыш вычисляется с учетом оговоренной в 5° реакции неотклонившихся игроков. Очевидно, что если набор стратегий не является допустимым, то он не может быть выбран в качестве решения игры. Обозначим множество допустимых наборов стратегий через D.

Например, если в игре двух лиц положить, что множество РКО состоит из одноэлементных коалиций $\{1\},\{2\}$, а множество РМО является целым отрезком игры, то множество D совпадает с множеством равновесных по Нэшу решений [7]. Далее, если в игре двух лиц множество РКО состоит из единственной коалиции $\{i\}, i \in \{1,2\}$, то наборы стратегий из получающегося множества D, наилучшие для игрока $3-i$, представляют собою решения по Штакельбергу в игре с лидером — игроком $3-i$ [7].

Если же множество РКО — пустое множество, то на основе получающегося множества D можно строить различные типы кооперативных решений.

6°. Описание последовательности, в которой игроки осуществляют выбор своих стратегий.

7°. Описание дополнительных условий на допустимые наборы стратегий.

2. Неантагонистическая позиционная дифференциальная игра двух лиц. NE - и $P(NE)$-решения

В этом разделе зададим конкретные варианты элементов 1°–7°, определяющих неантагонистическую позиционную дифференциальную игру двух лиц (НПДИ).

1°. Динамика игры описывается уравнением

$$\dot{x} = f(t,x,u,v), \;\; t \in [t_0,\vartheta], \;\; x(t_0) = x_0, \tag{2.1}$$

где $x \in \mathbb{R}^n$, $u \in P \in comp(\mathbb{R}^p)$, $v \in Q \in comp(\mathbb{R}^q)$; ϑ — заданный момент окончания игры. Предполагается, что функция f непрерывна, липшицева по x, удовлетворяет условию подлинейного роста по x а также условию седловой точки в маленькой игре [5]. Игрок 1 ($P1$) и игрок 2 ($P2$) распоряжаются выбором управлений u и v соответственно.

2°. Обоим игрокам доступна информация о текущей позиции игры $\{t,x(t)\}$. Используемая здесь формализация позиционных стратегий и порождаемых ими движений аналогична формализации, введенной в [5;6], за исключением технических деталей [7].

Стратегия игрока 1 отождествляется с парой $U = \{u(t,x,\varepsilon),\beta_1(\varepsilon)\}$, где $u(\cdot)$ — произвольная функция позиции (t,x) и положительного параметра точности ε, принимающая значения в множестве P. Функция $\beta_1 : (0,\infty) \longrightarrow (0,\infty)$ непрерывная монотонная и удовлетворяет условию $\beta_1(\varepsilon) \longrightarrow 0$ при $\varepsilon \longrightarrow 0$. Для фиксированного ε величина $\beta_1(\varepsilon)$ является верхней границей шага разбиения отрезка $[t_0,\vartheta]$, которое игрок 1 применяет при формировании пошаговых движений.

Аналогично стратегия игрока 2 определяется как $V = \{v(t,x,\varepsilon),\beta_2(\varepsilon)\}$.

Рассматриваются два типа движений, порожденные парой (U,V) из начальной позиции (t_0,x_0): аппроксимационные (пошаговые) движения и идеальные (предельные) движения.

Аппроксимационное движение $x^{\varepsilon}_{\Delta}[t] = x\,[t,t_0,x_0,U,\varepsilon_1,\Delta_1,V,\varepsilon_2,\Delta_2]$ определяется для фиксированных значений параметров точности ε_1 и ε_2, для фиксированных разбиений отрезка $[t_0,\vartheta]$: $\Delta_1 = \{t_i^{(1)}\}$ и $\Delta_2 = \{t_k^{(2)}\}$, выбранных игроками и удовлетворяющих условиям $\delta(\Delta_i) \le \beta_i(\varepsilon_i),\ i=1,2$, где $\delta(\Delta_i){=}\max\limits_j(t^{(i)}_{j+1} - t^{(i)}_j)$.

Идеальное (предельное) движение $x(t) = x(t,t_0,x_0,U,V)$ есть равномерный предел последовательности аппроксимационных движений

$$\{x^{\varepsilon^s}_{\Delta^s}\,[t,t_0^s,x_0^s,U,\varepsilon_1^s,\Delta_1^s,V,\varepsilon_2^s,\Delta_2^s]\}$$

где $s \to \infty,\ \varepsilon_i^s \to 0,\ t_0^s \to t_0,\ x_0^s \to x_0,\ \ \delta(\Delta_i^s) \le \beta_i(\varepsilon_i^s),\ i=1,2.$

Законы управления $(U, \varepsilon_1, \Delta_1)$ и $(V, \varepsilon_2, \Delta_2)$ назовем *согласованными по параметру точности*, если $\varepsilon_1 = \varepsilon_2$. Согласованные законы управления порождают согласованные аппроксимационные движения, последовательности которых порождают согласованные предельные движения.

Множество предельных движений $X(t_0, x_0, U, V)$ есть компакт в пространстве $C[t_0, \vartheta]$.

3°. Функционал выигрыша игрока i берется в виде

$$I_i = \sigma_i(x(\vartheta)), \quad i = 1, 2, \tag{2.2}$$

где σ_i — непрерывная функция.

4°. Множество разрешенных коалиций отклонения состоит из двух одноэлементных коалиций $\{1\}$,$\{2\}$; для любой РКО множество разрешенных моментов отклонения есть отрезок $[t_0, \vartheta]$.

5°. Неотклонившийся игрок выбирает после отклонения какую угодно стратегию.

6°. Игроки выбирают свои стратегии одновременно.

7°. Дополнительные условия на допустимые наборы стратегий отсутствуют.

Заданное таким образом множество элементов 1°–7° определяет классическую неантагонистическую позиционную дифференциальную игру (НПДИ) [2; 7].

З а м е ч а н и е 1. Вместо описанных вариантов элементов 4° и 6° можно принять следующие:

4°. Множество РКО:$\{\{2\}\}$; для этой РКО множество РМО: $[t_0, \vartheta]$.

6°. Игрок 1 выбирает свою стратегию первым.

В этом случае в результате полученное множество элементов 1°–7° определит иерархическую игру Штакельберга с лидером — первым игроком [7].

О п р е д е л е н и е 2. Пара стратегий (U^N, V^N) образует нэшевское решение (NE-решение) игры НПДИ, если для любого движения $\overline{x}(\cdot) \in X(t_0, x_0, U^N, V^N)$, любого $\tau \in [t_0, \vartheta]$, любых стратегий U и V выполнены следующие неравенства:

$$max_{x(\cdot)} \sigma_1(x(\vartheta, \tau, \overline{x}(\tau), U, V^N)) \leq min_{x(\cdot)} \sigma_1(x(\vartheta, \tau, \overline{x}(\tau), U^N, V^N)), \tag{2.3}$$

$$max_{x(\cdot)} \sigma_2(x(\vartheta, \tau, \overline{x}(\tau), U^N, V]) \leq min_{x(\cdot)} \sigma_2(x(\vartheta, \tau, \overline{x}(\tau), U^N, V^N]), \tag{2.4}$$

причем операции min производятся по множеству согласованных движений, а операции max — по соответствующим множествам всех движений.

О п р е д е л е н и е 3. NE-решение (U^P, V^P), неулучшаемое по Парето относительно величин I_1, I_2 (2.2), называется $P(NE)$- решением НПДИ.

Введем вспомогательные антагонистические позиционные дифференциальные игры Γ_1 и Γ_2. Динамика обеих игр описывается уравнением (2.1). В игре Γ_i, $i = 1, 2$, игрок i максимизирует свой функционал I_i (2.2) , а игрок $3 - i$ противодействует ему. В [5; 6] показано, что обе игры Γ_1 и Γ_2 имеют универсальные седловые точки

$$u^{(i)}(t, x, \varepsilon), \; v^{(i)}(t, x, \varepsilon), \quad i = 1, 2, \tag{2.5}$$

и непрерывные функции цены $\gamma_i(t, x)$, $i = 1, 2$.

Свойство универсальности стратегий (2.5) означает, что они являются оптимальными не только для фиксированной начальной позиции (t_0, x_0), но и для любой позиции (t_*, x_*), рассматриваемой в качестве начальной.

Нетрудно видеть, что величина $\gamma_i(t, x)$ представляет собою гарантированный выигрыш игрока i в позиции (t, x) игры.

В [7] установлена структура NE- и $P(NE)$-решений, а именно показано, что все NE- и $P(NE)$-решения игры могут быть найдены в классе пар стратегий (U, V), порождающих единственное предельное движение (траекторию). Решающие стратегии, составляющие такую пару и порождающие траекторию $x^*(\cdot)$, имеют вид

$$U^0 = \{u^0(t, x, \varepsilon), \; \beta_1^0(\varepsilon)\}, \quad V^0 = \{v^0(t, x, \varepsilon), \; \beta_2^0(\varepsilon)\}, \tag{2.6}$$

$$u^0(t,x,\varepsilon)=\begin{cases}u^*(t,\varepsilon), & \text{если } \|x-x^*(t)\|<\varepsilon\varphi(t),\\ u^{(2)}(t,x,\varepsilon), & \text{если } \|x-x^*(t)\|\geq\varepsilon\varphi(t),\end{cases} \tag{2.7}$$

$$v^0(t,x,\varepsilon)=\begin{cases}v^*(t,\varepsilon), & \text{если } \|x-x^*(t)\|<\varepsilon\varphi(t),\\ v^{(1)}(t,x,\varepsilon), & \text{если } \|x-x^*(t)\|\geq\varepsilon\varphi(t)\end{cases}$$

для всех $t\in[t_0,\vartheta],\varepsilon>0$. В (2.6), (2.7) через $u^*(t,\varepsilon)$, $v^*(t,\varepsilon)$ обозначены семейства программных управлений, порождающих траекторию $x^*(t)$. Функция $\varphi(\cdot)$ вместе с функциями $\beta_1^0(\cdot)$ и $\beta_2^0(\cdot)$ выбираются таким образом, что ломаные Эйлера, порожденные парой (U^0,V^0) из начальной позиции (t_0,x_0), не выходят за пределы $\varepsilon\varphi(t)$-окрестности траектории $x^*(t)$. Функции $u^{(2)}(\cdot,\cdot,\cdot)$ и $v^{(1)}(\cdot,\cdot,\cdot)$ определены в (2.5).

Далее, для каждой NE- и $P(NE)$- траектории $x^*(t)$ имеет место следующее свойство.

С в о й с т в о A. Точка $t=\vartheta$ является точкой максимума функции гарантированного выигрыша игрока i, вычисленной вдоль этой траектории, т. е.

$$max_{t\in[t_0,\vartheta]}\gamma_i(t,x^*(t))=\gamma_i(\vartheta,x^*(\vartheta)),\quad i=1,2. \tag{2.8}$$

3. Новый вариант элемента 7°: типы поведения

Рассмотрим теперь новый вариант элемента 7°.

7°. Дополнительно предполагаем, что помимо обычного, *нормального* (nor), типа поведения, ориентированного на максимизацию собственных функционалов выигрыша (2.2), игроки могут использовать другие типы поведения, а именно *альтруистический*, *агрессивный* и *парадоксальный* типы ([3; 4]).

Эти три типа поведения могут быть формализованы следующим образом:

О п р е д е л е н и е 4. Скажем, что игрок 1 придерживается на отрезке $[t_*,t^*]$ альтруистического (alt) типа поведения по отношению к игроку 2, если на этом отрезке действия игрока 1 направлены на максимизацию функционала I_2 игрока 2.

О п р е д е л е н и е 5. Скажем, что игрок 1 придерживается на отрезке $[t_*,t^*]$ агрессивного (agg) типа поведения по отношению к игроку 2, если на этом отрезке действия игрока 1 направлены на минимизацию функционала I_2 игрока 2.

О п р е д е л е н и е 6. Скажем, что игрок 1 придерживается на отрезке $[t_*,t^*]$ парадоксального (par) типа поведения, если на этом отрезке действия игрока 1 направлены на минимизацию собственного функционала I_1.

Аналогично определяются альтруистический и агрессивный типы поведения игрока 2 по отношению к игроку 1, а также парадоксальный тип поведения игрока 2. Заметим, что агрессивный тип поведения игроков фактически используется в НПДИ в форме стратегий наказания, содержащихся в структуре решений игры (см., например, [7]).

Приведенные определения характеризуют экстремальные типы поведения игроков. В действительности же реальные индивидуумы ведут себя, как правило, частично нормально, частично альтруистично, частично агрессивно и частично парадоксально. Другими словами, смешанные типы поведения, по-видимому, больше согласуются с реальностью.

Если каждого игрока ограничить "чистыми" типами поведения, то в рассматриваемой игре двух лиц (2.1) с функционалами I_1 и I_2 (2.2) существует 16 возможных пар типов поведения:(nor,nor), (nor,alt), (nor,agg), (nor,par), (alt,nor), (alt,alt), (alt,agg), (alt,par), (agg,nor), (agg,alt), (agg,agg), (agg,par), (par,nor), (par,alt), (par,agg), (par,par). При этом в следующих четырех парах (nor,alt), (alt,nor), (agg,par) и (par,agg) интересы игроков совпадают и игроки решают командные задачи управления. В следующих четырех парах (nor,agg), (alt,par), (agg,nor) и (par,alt) игроки имеют противоположные интересы и, следовательно, разыгрываются антагонистические дифференциальные игры. Остальные 8 пар определяют неантагонистические дифференциальные игры.

Идея использования игроками возможности переключения по ходу игры своего поведения с одного типа на другой была применена для игры с кооперативной динамикой в работе [4] и для повторяющейся биматричной 2×2 игры в работе [8], что позволило получить новые решения в этих играх.

Распространение указанного подхода на неантагонистические позиционные дифференциальные игры приводит к новым постановкам задач. В частности, представляет интерес, как трансформируются выигрыши игроков, получаемые на нэшевских решениях. Актуальной становится задача минимизации времени "ненормального" поведения при условии достижения результата, более хорошего, чем при нормальном поведении игроков.

4. Формализация игры НПДИсТП. BT-решение игры

Таким образом, игроки могут по ходу игры переключаться с одного типа поведения на другой. Такую игру будем называть *неантагонистической позиционной дифференциальной игрой с типами поведения игроков* (НПДИсТП).

Далее будем полагать, что одновременно с выбором позиционной стратегии каждый игрок выбирает также свою индикаторную функцию, определенную на отрезке $t \in [t_0, \vartheta]$ и принимающую значение в множестве $\{nor, alt, agg, par\}$. Обозначим индикаторную функцию игрока i символом $\alpha_i : [t_0, \vartheta] \to \{nor, alt, agg, par\}$, $i = 1, 2$. Если индикаторная функция какого-то игрока принимает значение, скажем, alt на некотором отрезке времени, то этот игрок действует на этом отрезке как альтруист по отношению к своему партнеру по игре. Заметим также, что если индикаторные функции обоих игроков тождественно равны значению nor на всем промежутке игры, то имеем классическую НПДИ.

Таким образом, в рассматриваемой игре с различными типами поведения игроков игрок 1 управляет выбором пары *действий* {позиционная стратегия, индикаторная функция}: $(U,\ \alpha_1(\cdot))$, а игрок 2 управляет выбором пары действий $(V,\ \alpha_2(\cdot))$.

Как упоминалось в разд. 3, при выборе игроками различных типов поведения могут сложиться три вида задач принятия решений: задача командного управления, антагонистическая игра, неантагонистическая игра. Будем считать, что игроки в каждой из трех указанных ситуаций руководствуются следующим правилом.

П р а в и л о 1. Если на промежутке $(\tau_1, \tau_2) \subset [t_0, \vartheta]$ индикаторные функции игроков сгенерировали неантагонистическую игру, то на этом промежутке игроки выбирают одно из $P(NE)$- решений сложившейся игры. Если сложилась антагонистическая игра, то в качестве решения игроки реализуют седловую точку. Наконец, если сложилась задача командного управления, то игроки выбирают одну из пар управлений, обеспечиващих неубывание вдоль траектории функции цены γ_i, где i — номер игрока, функционал которого максимизируется.

Вообще говоря, один и тот же кусок траектории может быть отслежен несколькими парами типов поведения игроков, причем эти пары могут отличаться друг от друга временем использования *ненормальных* типов. Естественно ввести следующее правило.

П р а в и л о 2. При наличии нескольких пар типов поведения, отслеживающих некоторую часть траектории, игроки выбирают ту из них, которая минимизирует время использования ненормальных типов поведения.

Введем теперь определение понятия решения игры НПДИсТП. Отметим, что множество движений, порожденных парой действий $\{(U, \alpha_1(\cdot)), (V, \alpha_2(\cdot))\}$, совпадает с множеством движений, порожденных парой (U, V) в соответствующей НПДИ.

О п р е д е л е н и е 7. Пара $\{(U^0, \alpha_1^0(\cdot)), (V^0, \alpha_2^0(\cdot))\}$, согласованная с правилом 1, образует BT-решение игры НПДИсТП, если найдется порожденная парой траектория $x^{BT}(\cdot)$ и найдется $P(NE)$-решение в соответствующей игре НПДИ, порождающее траекторию $x^P(\cdot)$, такие, что выполнены неравенства

$$\sigma_i(x^{BT}(\vartheta)) \geq \sigma_i(x^P(\vartheta)), \ \ i = 1, 2, \tag{4.1}$$

причем по крайней мере одно из неравенств строгое.

О п р е д е л е н и е 8. BT-решение $\{(U^P, \alpha_1^P(\cdot)), (V^P, \alpha_2^P(\cdot))\}$, неулучшаемое по Парето относительно величин I_1, I_2 (2.2), назовем $P(BT)$-решением игры НПДИсТП.

З а д а ч а 1. Найти множество BT-решений.

З а д а ч а 2. Найти множество $P(BT)$-решений.

В общем случае задачи 1 и 2 решений не имеют. Однако вполне ожидаемо, что использование игроками в игре НПДИсТП типов поведения, отличных от нормального, может в ряде случаев привести к исходам, более предпочтительным для них, чем в соответствующей игре НПДИ только с нормальным типом поведения. Пример такого рода приводится в следующем разделе.

5. Пример

Пусть динамика (2.1) имеет вид

$$\dot{x} = u + v, \quad x, u, v \in \mathbb{R}^2, \quad \|u\| \leq 1, \quad \|v\| \leq 1, \quad 0 \leq t \leq \vartheta, \quad x(0) = x_0, \tag{5.1}$$

а функционалы выигрыша (2.2) суть

$$I_i = 20 - \|x(\vartheta) - a^{(i)}\|, \quad i = 1, 2, \tag{5.2}$$

т. е. игрок i стремится привести точку $x(\vartheta)$ как можно ближе к своей целевой точке $a^{(i)}$.

Зададим следующие значения параметров игры:

$$\vartheta = 5.0, \quad x_0 = (0, 0), \quad a^{(1)} = (-11, 9), \quad a^{(2)} = (11, 9). \tag{5.3}$$

В задаче имеется следующее фазовое ограничение в плоскости (x_1, x_2): траекториям системы (5.1) запрещается заходить во внутренность множества S , которое получается удалением из четырехугольника $Oabc$ двухзвенной ломаной Oeg (см. рисунок). Множество S состоит из двух частей S_1 и S_2, т. е. $S = S_1 \bigcup S_2$.

Координаты точек, задающих фазовое ограничение, следующие:

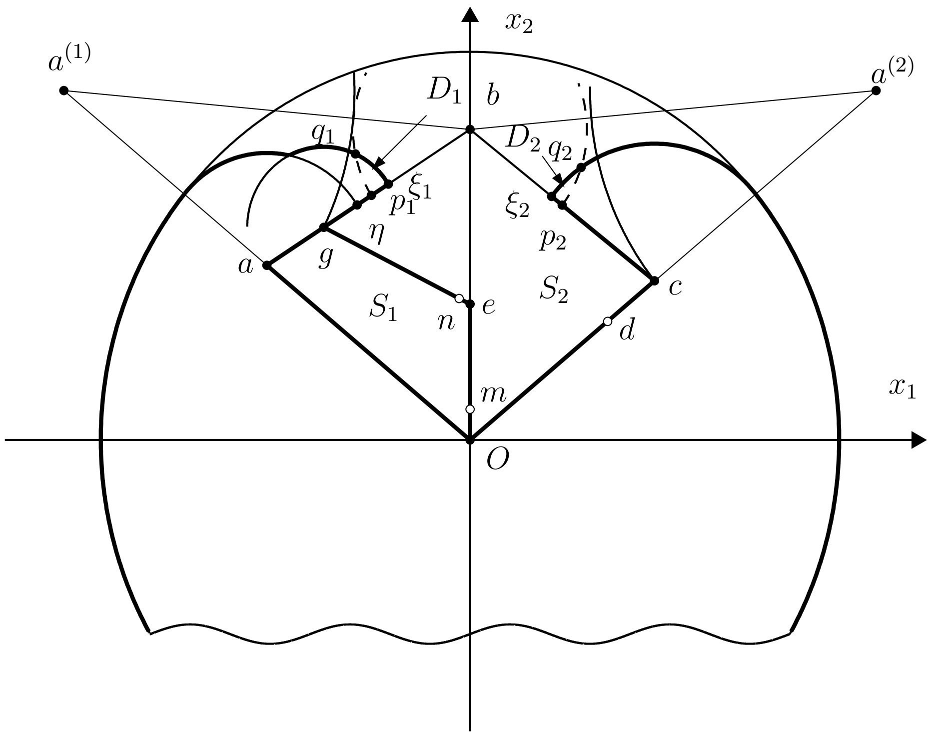
$$a = (-5.5, 4.5), \quad b = (0, 8), \quad c = (5, 4.1), \quad O = (0, 0), \quad e = (0, 3.5), \quad g = (-3.96, 5.48). \tag{5.4}$$

Можно проверить, что точка a лежит на отрезке $Oa^{(1)}$, а точка g — на отрезке ab. Отметим также, что точка c лежит чуть выше отрезка $Oa^{(2)}$.

Множество достижимости системы (5.1), построенное для момента $\vartheta = 5.0$, состоит из точек круга радиуса 10, расположенных не выше трехзвенника aOc и ограниченных также двумя дугами, соединяющими большую окружность со сторонами ab и bc четырехугольника (см. рисунок). При этом первая (составная) дуга состоит из дуги окружности с центром в точке a и радиусом $r_1 = |a\eta|$ и дуги окружности с центром в точке g и радиусом $r = |g\xi_1|$. Вторая дуга есть дуга окружности с центром в точке c и радиусом $r_2 = |c\xi_2|$.

Результаты приближенных вычислений: $r = 2.0726$, $r_1 = 2.8937$, $r_2 = 3.5335$, $\eta = (-3.0587, 6.0536)$, $xi_1 = (-2.2114, 6.5927)$, $\xi_2 = (2.2135, 6.2735)$. Кроме того, имеем (приближенно) $|Oa^{(1)}| = |Oa^{(2)}| = 14.2127$ и $|a^{(1)}b| = |a^{(2)}b| = 11.0454$.

На рисунке пунктирными линиями изображены дуги окружности L с центром в точке b и радиусом $r_3 = 3.1673 = 14.2127 - 11.0454$. Они пересекают стороны ab и bc в точках p_1 и p_2



Множество достижимости.

соответственно с координатами $p_1 = (-2.6721, 6.2995)$, $p_2 = (2.4974, 6.0520)$. По построению длины двузвенников $a^{(1)}bp_2$ и $a^{(2)}bp_1$ равны между собой и равны длине отрезков $Oa^{(1)}$ и $Oa^{(2)}$.

Функции цены $\gamma_1(t,x)$ и $\gamma_2(t,x)$, $0 \le t \le \vartheta, x \in \mathbb{R}^2 \backslash S$, (2.5) вспомогательных антагонистических игр Γ_1 и Γ_2 в данном примере будут

$$\gamma_i\,(t,x) = \begin{cases} 20 - \|x - a^{(i)}\|, & \text{если } xa^{(i)} \bigcap intS = \varnothing, \\ 20 - \rho_S\left(x, a^{(i)}\right) & \text{в противном случае.} \end{cases} \tag{5.5}$$

Здесь $i = 1, 2$, а через $\rho_S\left(x, a^{(i)}\right)$ обозначено наименьшее из двух расстояний от точки x до точки $a^{(i)}$, одно из которых вычисляется при обходе множества S по часовой стрелке, а другое — при обходе S против часовой стрелки.

Нетрудно проверить, что в игре НПДИ с нормальным типом поведения игроков траектория $x(t) \equiv 0$, $t \in [0, 5]$, (точка O) будет нэшевской траекторией. Далее, траектории, построенные вдоль ломаной Oeg, нэшевскими не являются, поскольку ни для одной из них не выполняется условие (2.8). Это подтверждается и тем, что окружность радиуса $|a^{(1)}g|$ с центром в точке $a^{(1)}$ не имеет общих точек с окружностью L (см. рисунок). Не являются нэшевскими и все траектории, огибающие множество S_2 справа, поскольку окружность радиуса $|a^{(2)}c|$ с центром в точке $a^{(2)}$ также не имеет общих точек с окружностью L. Очевидно, не являются нэшевскими и все траектории, огибающие множество S_1 слева. В итоге получается, что упомянутая траектория является единственной нэшевской траекторией, а, следовательно, и единственной $P(NE)$-траекторией; выигрыши игроков на ней определяются как $I_1 = I_2 = 5.7873$.

Перейдем теперь к игре НПДИсТП, в которой будем рассматривать два варианта.

Вариант *I*. Полагаем, что игроки 1 и 2 могут проявлять альтруизм по отношению к своему партнеру в течение некоторых промежутков времени.

Вариант *II*. Помимо допущения об альтруизме игроков дополнительно полагаем, что каждый игрок может проявлять агрессию по отношению к другому игроку в течение некоторых промежутков времени, причем допускается случай взаимной агрессии.

В множестве достижимости найдем все точки x, для которых выполнены неравенства

$$\sigma_i(x) \geq \sigma_i(O),\ i = 1, 2,\ \ \sigma_1(x) + \sigma_2(x) > \sigma_1(O) + \sigma_2(O) \tag{5.6}$$

Такие точки составят два множества D_1 и D_2 (см. рисунок). Множество D_1 ограничено отрезком $p_1\xi_1$, а также дугами p_1q_1 и $q_1\xi_1$ упомянутых выше окружностей. Множество D_2 ограничено отрезком $p_2\xi_2$, а также дугами ξ_2q_2 и q_2p_2 упомянутых окружностей. При этом на дуге p_1q_1 нестрогое неравенство (5.6) при $i = 2$ превращается в равенство, а на дуге p_2q_2 нестрогое неравенство (5.6) превращается в равенство при $i = 1$. В остальных точках множеств D_1 и D_2 нестрогие неравенства (5.6) при $i \in \{1, 2\}$ будут строгими.

В рамках варианта I построим BT-решение, приводящее в точку $\xi_2 \in D_2$.

Рассмотрим траекторию $Oc\xi_2$; выигрыши игроков на ней составляют $I_1 = 6.1474$, $I_2 = 10.8002$, т. е. каждый игрок выигрывает больше, чем на единственной $P(NE)$-траектории. Как следует из вышесказанного, траектория $Oc\xi_2$ не является нэшевской. Однако если удастся построить индикаторные функции-программы игроков, обеспечивающие движение по этой траектории, то тем самым будет построено BT-решение.

На стороне Oc найдем точку d, равноудаленную от точки $a^{(1)}$ как при обходе множества S по часовой стрелке, так и при обходе S против часовой стрелки. Получим $d = (3.7271, 3.0562)$. Далее, если двигаться по траектории $Oc\xi_2$ с максимальной скоростью при $t \in [0, 5]$, то время попадания в точку d будет $t = 2.4100$, а в точку c будет $t = 3.2330$. При этом можно убедиться, что при движении на промежутке $t \in [0, 2.4100]$ функция $\gamma_1(t, x)$ (5.5) монотонно убывает, а функция $\gamma_2(t, x)$ монотонно возрастает; при движении на промежутке $t \in [2.4100, 3.2330]$ обе функции $\gamma_1(t, x)$ и $\gamma_2(t, x)$ монотонно возрастают; наконец, при движении на оставшемся промежутке $t \in [3.2330, 5.0]$ функция $\gamma_1(t, x)$ монотонно возрастает, а функция $\gamma_2(t, x)$ убывает.

Проверяем, что на участке Od траектории пара (alt, nor), определяющая командную задачу управления, является единственной парой типов поведения, осуществляющей движение на участке в соответствии с правилом 1; это будет максимальный сдвиг в направлении точки $a^{(2)}$. На следующем участке dc будут уже четыре пары “кандидатов” (nor, nor), (alt, nor), (nor, alt) и (alt, alt), однако согласно правилу 2 три последних пары отбрасываются; оставшаяся пара определяет неантагонистическую игру, и движение на этом участке будет порождено $P(NE)$-решением игры. Наконец, на последнем участке $c\xi_2$ единственной парой типов поведения, осуществляющей движение на участке в соответствии с правилом 1, будет пара (nor, alt), определяющая командную задачу управления; движение представляет максимальный сдвиг в направлении точки ξ_2.

Таким образом, построены индикаторные функции-программы игроков

$$\alpha_1^{(1)}(t) = \{alt, t \in [0, 2.4100);\ \ nor, t \in [2.4100, 5]\}, \tag{5.7}$$

$$\alpha_2^{(1)}(t) = \{nor, t \in [0, 3.2330);\ \ alt, t \in [3.2330, 5]\}. \tag{5.8}$$

Обозначим через $(U^{(1)}, V^{(1)})$ пару стратегий игроков, порождающую предельное движение $Oc\xi_2$ при $t \in [0, 5]$ и согласованную с построенными индикаторными функциями. Дополнительно учтем, что ξ_2 — ближайшая к $a^{(1)}$ точка множества D_2, а также то обстоятельство, что в область D_1 не удается попасть, применяя только нормальный и альтруистический типы поведения. Тогда получаем следующее утверждение.

Утверждение 1. *В варианте I пара действий* $\{(U^{(1)}, \alpha_1^{(1)}(\cdot)), (V^{(1)}, \alpha_2^{(1)}(\cdot))\}$ (5.7), (5.8) *доставляет $P(BT)$-решение.*

Переходим к варианту II, в котором помимо предположения об альтруизме игроков дополнительно предполагается, что игроки могут использовать агрессивный тип поведения. Построим BT-решение, приводящее в точку $\xi_1 \in D_1$.

Найдем точку m, равноудаленную от точки $a^{(1)}$ как при обходе множества S_1 по часовой стрелке, так и при обходе S_1 против часовой стрелки. Найдем также точку n, равноудаленную от точки $a^{(2)}$ как при обходе множества S_2 по часовой стрелке, так и при обходе S_2 против часовой стрелки. Результаты вычислений: $m = (0, 0.7928)$, $n = (-0.2925, 3.6462)$.

Рассмотрим траекторию $Oeg\xi_1$; выигрыши игроков на ней составляют $I_1 = 10.8840, I_2 = 6.5359$, т. е. выигрыши обоих игроков на этой траектории больше, чем на единственной $P(NE)$-траектории. Как следует из вышесказанного, траектория $Oeg\xi_1$ не является нэшевской. Поэтому если удастся построить индикаторные функции-программы игроков, обеспечивающие движение по этой траектории, то тем самым будет построено BT-решение.

Прежде всего отметим, что при движении по траектории $Oeg\xi_1$ с максимальной скоростью при $t \in [0, 5]$ время попадания в точку m будет $t = 0.3964$, в точку n будет $t = 1.9135$, а в точку g будет $t = 3.9620$. Нетрудно проверить, что при таком движении по траектории $Oeg\xi_1$ на промежутке $t \in [0, 0.3964]$ обе функции $\gamma_1(t, x)$ и $\gamma_2(t, x)$ монотонно убывают; при движении на промежутке $t \in [0.3964, 1.9135]$ функция $\gamma_2(t, x)$ продолжает убывать, а функция $\gamma_1(t, x)$ возрастает; при движении на промежутке $t \in [1.9135, 3.9620]$ обе функции возрастают; наконец, на оставшемся промежутке $t \in [3.9620, 5]$ функция $\gamma_2(t, x)$ продолжает возрастать, а функция $\gamma_1(t, x)$ убывает.

Проверяем, что на участке Om траектории пара (agg, agg), определяющая неантагонистическую игру, является единственной парой типов поведения, осуществляющей движение на участке в соответствии с правилом 1; это будет движение, порожденное $P(NE)$-решением игры, наилучшим для обоих игроков. На следующем участке mn две пары осуществляют движение в соответствии с правилом 1, а именно (nor, alt) и (agg, alt), однако согласно правилу 2 остается только пара (nor, alt); она определяет командную задачу управления, в которой движение представляет максимальный сдвиг в направлении точки n. На участке ng будут уже четыре пары "кандидатов" (nor, nor), (alt, nor), (nor, alt) и (alt, alt), однако согласно правилу 2 три последние пары отбрасываются; оставшаяся пара определяет неантагонистическую игру и движение на этом участке будет порождено $P(NE)$-решением игры. Наконец, на последнем участке $g\xi_1$ единственной парой типов поведения будет пара (alt, nor), определяющая командную задачу управления; движение представляет максимальный сдвиг в направлении точки ξ_2.

Таким образом, построены индикаторные функции-программы игроков

$$\alpha_1^{(2)}(t) = \{agg, t \in [0, 0.3964); \quad nor, t \in [0.3964, 3.9620); \quad alt, t \in [3.9620, 5]\}, \tag{5.9}$$

$$\alpha_2^{(2)}(t) = \{agg, t \in [0, 0.3964); \quad alt, t \in [0.3964, 1.9135); \quad nor, t \in [1.9135, 5]\}. \tag{5.10}$$

Обозначим через $(U^{(2)}, V^{(2)})$ пару стратегий игроков, порождающую предельное движение $Oc\xi_1$ при $t \in [0, 5]$ и согласованную с построенными индикаторными функциями.

У т в е р ж д е н и е 2. *В варианте II пара действий* $\{(U^{(2)}, \alpha_1^{(2)}(\cdot)), (V^{(2)}, \alpha_2^{(2)}(\cdot))\}$ (5.9), (5.10) *доставляет* $P(BT)$*-решение.*

З а м е ч а н и е 2. Очевидно, что утверждение 1 справедливо и для варианта II.

Следуя схеме доказательств утверждений 1 и 2 (и учитывая также замечание 2), приходим к следующим утверждениям.

У т в е р ж д е н и е 3. *В варианте I множество D_2 состоит из тех и только тех точек, которые являются концами траекторий, порожденных BT-решениями игры.*

У т в е р ж д е н и е 4. *В варианте II множества D_1 и D_2 состоят из тех и только тех точек, которые являются концами траекторий, порожденных BT-решениями игры.*

СПИСОК ЛИТЕРАТУРЫ

1. **Оуэн Г.** Теория игр. М.: Мир, 1971. 230 с.
2. **Петросян Л.А., Зенкевич Н.А., Шевкопляс Е.В.** Теория игр. СПб.: БХВ-Петербург, 2012. 432 с.

3. **Клейменов А.Ф.** О решениях в неантагонистической позиционной дифференциальной игре // Прикл. математика и механика. 1997. Т. 61, вып. 5. С. 739–746.
4. **Kleimenov A.F., Kryazhimskii A.V.** Normal behavior, altruism and aggression in cooperative game dynamics. Laxenburg: IIASA, 1998. Interim Report IR-98-076. 47 p.
5. **Красовский Н.Н., Субботин А.И.** Позиционные дифференциальные игры. М.: Наука, 1974. 456 с.
6. **Красовский Н.Н.** Управление динамической системой. М.: Наука, 1985. 516 p.
7. **Клейменов А.Ф.** Неантагонистические позиционные дифференциальные игры. Екатеринбург: Наука, 1993. 185 с.
8. **Kleimenov A.F.** An approach to building dynamics for repeated bimatrix 2×2 games involving various behavior types // Dynamic and Control / ed. G. Leitman. London: Gordon and Breach, 1998. P. 195–204.
9. **Клейменов А.Ф.** Альтруистическое поведение в неантагонистической позиционной дифференциальной игре // Математическая теория игр и ее приложения. 2015. Т. 7, вып. 4. С. 40–55.
10. **Клейменов А.Ф.** Агрессивное поведение в неантагонистической позиционной дифференциальной игре // Математическая теория игр и ее приложения. 2016. Т. 8, вып. 4. С. 63–78.

Клейменов Анатолий Федорович
д-р физ.-мат. наук, ведущий науч. сотрудник
Институт математики и механики УрО РАН им. Н. Н. Красовского,
г. Екатеринбург
e-mail: kleimenov@imm.uran.ru

Поступила 10.09.2017

REFERENCES

1. Owen G. *Game theory.* NY: Saunders (W.B.) Co., Ltd., 1968, 228 p. ISBN: 0721670288 . Translated to Russian under the title *Teoriya igr*, Moscow, Mir Publ., 1971, 230 p.
2. Petrosyan L.A., Zenkevich N.A., Shevkoplyas E.V. *Teoriya igr* [Game theory]. Saint Petersburg, BKhV-Peterburg Publ., 2012, 432 p. ISBN: 978-5-9775-0484-3 .
3. Kleimenov A.F. On solutions in a nonantagonistic positional differential game. *J. Appl. Math. Mech.*, 1997, vol. 61, no. 5, pp. 717–723. doi: 10.1016/S0021-8928(97)00094-4 .
4. Kleimenov, A.F., Kryazhimskii A.V. *Normal behavior, altruism and aggression in cooperative game dynamics.* Interim Report IR-98-076, Laxenburg: IIASA, 1998, 47 p.
5. Krasovskii N.N., Subbotin A.I. *Game-theoretical control problems.* NY, Springer, 1988, 517 p. ISBN: 978-1-4612-8318-8. This book is substantially revised version of the monograph N.N. Krasovskii, A.I. Subbotin, *Pozitsionnye differentsial'nye igry*, Moscow, Nauka Publ., 1974. 456 p.
6. Krasovskii N.N. *Upravlenie dinamicheskoi sistemoi. Zadacha o minimume garantirovannogo rezul'tata.* [Control of a dynamic system. The problem of minimum guaranteed result]. Moscow, Nauka Publ., 1985, 516 p.
7. Kleimenov, A.F. *Nonantagonistic positional differential games* [Neantagonisticheskie pozitsionnye differentsial'nye igry]. Ekaterinburg, Nauka Publ., 1993, 185 p. ISBN: 5-7691-0353-1 .
8. Kleimenov A.F. An approach to building dynamics for repeated bimatrix 2×2 games involving various behavior types. In: *Dynamics and Control*, ed. G. Leitman, London, Gordon and Breach, 1999, pp. 195–204. ISBN: 90-5699-172-8 .
9. Kleimenov A.F. Altruistic behavior in a nonantagonistic positional differential game. *Autom. Remote Control*, 2017, vol. 78, no. 4, pp. 762–769. doi: 10.1134/S0005117917040178 .
10. Kleimenov A.F. Aggressive behavior in a non-antagonistic positional differential game. *Mat. Teor. Igr Pril.*, 2016, vol. 8, no. 4, pp. 63–78 (in Russian).

The paper was received by the Editorial Office on September 10, 2017.

Anatolii Fedorovich Kleimenov, Dr. Phys.-Math. Sci., Krasovskii Institute of Mathematics and Mechanics, Ural Branch of the Russian Academy of Sciences, Yekaterinburg, 620990 Russia,
e-mail: kleimenov@imm.uran.ru .