На правах рукописи

Семенищев Алексей Андреевич

ОПТИМАЛЬНОЕ УПРАВЛЕНИЕ В ПОВТОРЯЮЩЕЙСЯ БИМАТРИЧНОЙ ИГРЕ С КОНЕЧНОЙ ПАМЯТЬЮ И В ПОВЕДЕНЧЕСКОЙ МОДЕЛИ ФИРМЫ

05.13.18 – математическое моделирование,
численные методы и комплексы программ

АВТОРЕФЕРАТ

диссертации на соискание ученой степени
кандидата физико-математических наук

Екатеринбург – 2001

Работа выполнена в Уральском государственном университете им. А.М. Горького на кафедре теоретической механики

Научный руководитель: доктор физико-математических наук, профессор А. Ф. Клейменов

Официальные оппоненты: - доктор физико- математических наук, профессор А. Н. Красовский;
- член-корреспондент РАН, доктор физико- математических наук, профессор А. В. Кряжимский.

Ведущая организация: - Санкт-Петербургский государственный университет.

Защита состоится 19 сентября 2001 г. в 15^{00} часов на заседании диссертационного совета К 212.286.01 по присуждению ученой степени кандидата физико-математических наук в Уральском государственном университете им. А.М. Горького (620083, г. Екатеринбург, К-83, пр. Ленина, 51, комн. 248).

С диссертацией можно ознакомиться в библиотеке Уральского государственного университета.

Автореферат разослан "14" августа 2001 г.

Ученый секретарь
диссертационного совета
к. ф.-м. н., доцент В.Г. Пименов.

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Актуальность работы. Во многих технических, экономических и социальных процессах возникают задачи принятия решения, в которых важное место занимают игровые постановки. Основы теории статических игр были разработаны в 30 – 50-х гг. в трудах Дж. фон Неймана и О. Моргенштерна[1], Дж. Нэша[2], Х. Штакельберга[3] и других исследователей.

В конце 50-х – начале 60-х гг. в работах Р. Айзекса[4] были сформулированы первые задачи антагонистических динамических игр и были предложены методы их решения. В дальнейшем эти задачи исследовались многими отечественными и зарубежными учеными. Фундаментальный вклад в построение теории антагонистических дифференциальных игр принадлежит научным школам академиков Н. Н. Красовского и Л. С. Понтрягина. Этому направлению посвящены исследования Р. Айзекса[4], Н. Н. Красовского[5–7], А. Н. Красовского[7], А. В. Кряжимского[8], А. Б. Куржанского[9], Ю. С. Осипова[10], Л. А. Петросяна[11], Л. С. Понтрягина[12,13], А. И. Субботина[5,14], А. Г. Ченцова[14] и других авторов.

Большое внимание исследователей уделяется математическим моделям, формализуемым в рамках теории неантагонистических по-

---

[1] Нейман Дж., Моргенштерн О. *Теория игр и экономическое поведение.* Пер с англ. М.:Наука, 1970.

[2] Нэш Дж. *Бескоалиционные игры*// Матричные игры. М.:Физматгиз, 1961. С. 205-221.

[3] Von Stackelberg H. *The theory of the market economy.* London: Hodge, 1952.

[4] Айзекс Р. *Дифференциальные игры.* М.:Мир, 1967.

[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. 518 с.

[7] Krasovskii A. N., Krasovskii N. N. *Control under lack of information.* Birkhäuser, 1995. 322 p.

[8] Кряжимский А. В. *К теории позиционных дифференциальных игр сближения - уклонения*// Докл. АН СССР, 1978. Т.239. №4. С. 779-782.

[9] Куржанский А. Б. *Управление и наблюдение в условиях неопределенности.* М.:Наука, 1977.

[10] Осипов Ю. С. *Дифференциальная игра систем с последействием*// Докл. АН СССР, 1971. Т.196. № 4. С. 779-782.

[11] Петросян Л. А. *Одна игра преследования на плоскости*// Докл. АН АрмССР, 1965. Т.40. № 5.

[12] Понтрягин Л. С. *О линейных дифференциальных играх, I*// Докл. АН СССР,1967. Т.174. № 6.

[13] Понтрягин Л. С. *О линейных дифференциальных играх, II*// Докл. АН СССР,1967. Т.175. № 4.

[14] Субботин А. И., Ченцов А. Г. *Оптимизация гарантии в задачах управления.* М.:Наука, 1981.

зиционных дифференциальных игр. Такие модели возникают при описании динамических задач управления механическими, экономическими, биологическими системами, когда управление осуществляется разными участниками. При этом интересы участников не являются полностью противоположными, каждый из них оптимизирует собственный показатель качества и имеет свой собственный ресурс управления. В этом случае задача состоит в выработке управления, приемлемого для всех сторон, участвующих в игре. Основополагающие результаты в этом направлении были получены в работах Ю. Б. Гермейера[15], А. Ф. Кононенко[16], Л. А. Петросяна[17], Э. Мулена[18], В. И. Жуковского[19], А. Ф. Клейменова[20] и других исследователей.

Одна из основных проблем в неантагонистической игре состоит в выборе понятия решения, соответствующего содержанию задачи. В соответствии с различными принципами оптимальности выделяются равновесные но Нэшу решения[2], решения по Штакельбергу[3], кооперативное решение по Парето[1].

Важным частным случаем неантагонистических динамических игр являются повторяющиеся биматричные игры[21–23].

Среди них отметим бесконечно повторяющуюся игру, на каждом шаге которой разыгрывается известная биматричная игра «дилемма заключенного», предложенная А.У. Таккером. Активность, с которой изучается эта игра в последние годы, объясняется большим количеством ее интерпретаций с точки зрения социологии, психологии, биологии, экономики, философии, что отмечено в работах многих ис-

[15]Гермейер Ю. Б. *Игры с непротивоположными интересами.* М.:Наука, 1976.

[16]Кононенко А. Ф. *О равновесных позиционных стратегиях в неантагонистических дифференциальных играх*// Докл. АН СССР, 1976. Т.231. №2. С. 285-288.

[17]Петросян Л. А. *Устойчивость решений в дифференциальных играх со многими участниками*// Вест. ЛГУ, 1977. №19. С. 46-52.

[18]Мулен Э. *Теория игр с примерами из математической экономики.* М.:Мир, 1985.

[19]Жуковский В. И., Тынянский Н.Т. *Равновесные управления многокритериальных динамических систем:* М.:Изд-во МГУ, 1984.

[20]Клейменов А. Ф. *Неантагонистические позиционные дифференциальные игры:* Екатеринбург: Наука, 1993. 185 с.

[21]Льюс Р. Д., Райфа Х. *Игры и решения.* Пер. с англ. М.:ИЛ, 1961. 642 с.

[22]Axelrod R. *The evolution of cooperation.* New York: Basic Books, 1984.

[23]Nowak M., Sigmund K. *The Alternating Prisoner's Dilemma*// Journal of Theoretical Biology. 1994. Vol.168. P. 219-226.

следователей[21,22,24,25]. Основной вопрос, возникающий при анализе этой игры, следующий:какие условия необходимо создать игрокам, чтобы они проявили стремление к сотрудничеству[22].

Наиболее известным экспериментом, в ходе которого эмулировалось бесконечное повторение игры, стали два компьютерных турнира, описанных в монографии Р. Аксельрода[22]. Участниками турнира являлись программы, реализующие некоторые решающие правила в повторяющейся игре. Эффект бесконечного повторения достигался за счет того, что была задана вероятность продолжения игры, одна и та же для каждого шага. Турнир проводился по круговой схеме: каждая программа играла с каждой и, кроме того, со своей копией. В процессе принятия решения на программу не накладывалось никаких ограничений за исключением того, что ей не был известен алгоритм, по которому принимал решение партнер. Многие программы при выборе решения на текущем шаге использовали информацию о нескольких предыдущих шагах игры (например, так действовала стратегия Tit For Tat, ставшая победителем обоих турниров).

Вторым наиболее распространенным подходом к решению «дилеммы заключенного» является подход, основанный на рассмотрении динамического процесса как эволюционной игры[24–28].

Их идея заключается в том, что организуется партия повторяющихся игр, состоящая из довольно большого количества итераций, в некотором сообществе игроков – представителей различных популяций. Игроки одной популяции используют одну и ту же стратегию. Перед началом игры все игроки хорошо «перемешиваются», так что ни один из них не знает, с каким партнером он ведет игру. Более того, в некоторых экспериментах допускается, чтобы в ходе партии пары игроков разбивались и смешивались несколько раз.

По окончании партии подсчитываются накопленные выигрыши каждого игрока и на основе этих данных вычисляются рейтинги

[24]Maynard Smith J. *Evolution and the theory of games*. Cambrige Univ. Press, 1982. 222 p.

[25]Weibull J. W. *Evolutionary game theory*. Massachusetts Inst. of Techn. Press, 1995. 265 p.

[26]Friedman D. *Evolutionary games in economics//* Econometrica. Vol.59. № 3. P. 637-666.

[27]Hofbauer J., Sigmund K. *The theory of evolution and dynamic systems:* Cambridge Univ. Press, Cambridge, 1988.

[28]Kaniovski Yu. M., Kryazhimskii A. V., Young H. P. *Learning equilibria in games played by heterogeneous populations//* IIASA Interim Report IR-97-017.

каждой популяции. Количество игроков каждой популяции, допущенных к игре в следующей партии, берется пропорциональным ее рейтингу.

Следует также отметить подход к решению повторяющихся биматричных игр, заключающийся в смене типов поведения[29,30] игроков на протяжении повторяющейся игры.

Доказано[30,31], что выбор типа поведения по определенному алгоритму при некоторых условиях приводит к тому, что частота кооперации игроков в бесконечно повторяющейся игре стремится к единице. При этом существуют ситуации, в которых игроки должны проявлять альтруизм или агрессию по отношению к партнеру.

Данная работа примыкает по своей тематике к области упомянутых выше исследований.

Цель работы.
В первой главе диссертации изучается повторяющаяся биматричная игра типа «дилемма заключенного» с конечной памятью игроков в иерархической постановке. Целью исследования является разработка и обоснование алгоритма построения оптимальных стратегий лидера и ведомого.

Во второй главе изучается одна многокритериальная модель фирмы. Целью исследования является построение оптимального управления, приводящего систему на паретовское множество.

Методы исследования. В диссертации используются понятия и методы классической теории игр, теории графов, линейной алгебры, математического анализа, теории оптимального управления, теории дифференциальных уравнений.

Научная новизна. Основные результаты диссертации:

1. Разработан алгоритм построения оптимальных стратегий лидера и ведомого в бесконечно повторяющейся биматричной игре типа «дилемма заключенного» с конечной памятью игроков.

---

[29] Клейменов А. Ф. *О решениях в неантагонистической позиционной дифференциальной игре*// ПММ Т.61. Вып.5. 1997. С. 739-746.

[30] Kleimenov A. F., Kryazhimskii A. V. *Normal behavior, altruism and aggression in cooperative game dynamics*// IIASA Interim Report IR-98-076, 1998.

[31] Kleimenov A. F., Volegova E. I. *Problems of control by dynamics for repeated bimatrix 2×2 games with switching of local criteria for players*// Nonsmooth and discontinuous problems of control and optimization, Pergamon Press, 1999.

2. Построены оптимальные стратегии игроков в играх с одно- и двухшаговой памятью; для построения оптимальных стратегий в игре с двухшаговой памятью на основе указанного алгоритма разработана программа.

3. Доказано, что стратегия TIT FOR TAT оптимальна для лидера в игре с одношаговой памятью при произвольных параметрах игры, а также определены значения параметров, при которых эта стратегия является оптимальной в игре с двухшаговой памятью.

4. В одной многокритериальной модели фирмы указан алгоритм построения множества точек, оптимальных по Парето, при различных значениях параметров системы.

5. Доказано существование таких начальных состояний рассматриваемой модели, для которых использование нормальных типов поведения недостаточно для достижения паретовского множества.

6. В нескольких частных случаях построено оптимальное управление системой, приводящее ее на паретовское множество с минимальным временем использования типов поведения, не являющихся нормальными. Разработанные методы допускают компьютерную реализацию.

Результаты диссертационной работы являются новыми.

Теоретическая и практическая ценность. Результаты, полученные в первой главе диссертации, дополняют теорию повторяющихся биматричных игр и теорию эволюционных игр. Алгоритмы и методы, предложенные в ней, являются конструктивными и создают теоретическую основу для разработки численных методов.

Во второй главе диссертации получены результаты, применимые в рамках теории оптимального управления экономическими процессами.

Апробация работы. Основные результаты диссертации обсуждались и докладывались

- на международной конференции 11-th IFAC International Workshop «Control Applications of Optimization» (Санкт-Петербург, 2000 г.),

- на Воронежской весенней математической школе «Современные методы в теории краевых задач» (Воронеж, 1999 г.),

- на Всероссийской конференции «Общие проблемы управления и их приложения к математической экономике» (Тамбов, 2000 г.),

- на 30-ой молодежной конференции «Проблемы теоретической и прикладной математики» (Екатеринбург, 1999 г.),

- на научных семинарах Санкт-Петербургского государственного университета, кафедры теоретической механики Уральского государственного университета и отдела динамических систем Института математики и механики УрО РАН.

Структура и объем диссертации. Диссертация состоит из введения, двух глав и списка литературы. Объем диссертации 144 страниц. Библиография 61 наименование.

Публикации. Основные результаты диссертации опубликованы в 7 работах.

СОДЕРЖАНИЕ РАБОТЫ

Во **введении** дается общая характеристика работы, обзор литературы по исследуемой тематике, кратко изложено содержание работы. Также во введении обсуждаются различные постановки исследуемой задачи, полученные в их рамках результаты.

Первая глава состоит из пяти параграфов и посвящена изучению повторяющейся биматричной игры типа «дилемма заключенного».

Первый параграф носит вводный характер. В нем описывается статическая биматричная игра типа «дилемма заключенного», в которой каждый игрок имеет две стратегии, а именно кооперироваться с партнером (стратегия **C**) или отклоняться от сотрудничества (стратегия **D**). Матрица выигрышей первого игрока задается следующим образом:

$$A = \begin{pmatrix} p^+ & p^{--} \\ p^{++} & p^- \end{pmatrix} \tag{1}$$

при условии

$$p^{--} < p^- < p^+ < p^{++} \tag{2}$$

Матрица выигрышей второго игрока определяется следующим образом: $B = A^T$, задачей каждого игрока является максимизация выигрыша.

Формулируется основная проблема этой игры: при условии индивидуального рационализма каждый игрок получает выигрыш p^-, хотя при взаимном сотрудничестве игроки получают больший выигрыш, p^+.

Во втором параграфе исследуется бесконечно повторяющаяся дилемма заключенного. При бесконечном повторении игры на значения выигрышей накладывается дополнительное ограничение

$$\frac{p^{--} + p^{++}}{2} < p^+ \tag{3}$$

В рассматриваемой постановке игроки имеют конечную память; это означает, что для принятия решения на текущем шаге игры каждый игрок может использовать информацию о выигрышах первого игрока на m предыдущих шагах игры (очевидно, выигрыши второго игрока определяются однозначно).

Величины выигрышей нормируются таким образом, что $p^{--} = 0$, $p^{++} = 1$, отсюда следует, что игра задается по существу двумя параметрами, p^+ и p^-, принадлежащими области

$$P = \left\{(p^+, p^-) | \, 0 < p^- < p^+ < 1, p^+ > \frac{1}{2}\right\}$$

Множество всех состояний игры отождествляется со множеством вершин некоторого графа[32], называемого графом игры.

Стратегия игрока в бесконечно повторяющейся игре определяется как некоторая функция, заданная на множестве всех состояний игры и принимающая значения из множества $\{\mathbf{C}, \mathbf{D}\}$. Множество всех стратегий в игре с памятью m обозначается через $\mathcal{H}$.

При заданной стратегии $h^1 \in \mathcal{H}$ первого игрока и стратегии $h^2 \in \mathcal{H}$ второго игрока из каждой вершины графа игры, обозначаемого через $G(h^1, h^2)$, выходит единственная ориентированная дуга. Траектория игры определяется как путь[32] на графе $G(h^1, h^2)$, образованный в результате переходов между его вершинами в соответствии с заданными стратегиями игроков. В случае, если траектория содержит замкнутый участок, он называется циклом[32].

Во втором параграфе главы 1 доказывается важная лемма, описывающая динамику бесконечно повторяющейся игры.

Лемма 1 *При произвольных стратегиях игроков $h^1 \in \mathcal{H}$, $h^2 \in \mathcal{H}$ граф $G(h^1, h^2)$ содержит по крайней мере один цикл.*

[32] Оре О. *Теория графов.* Пер. с англ. М.:Наука, 1968. 352 с.

Выигрыш S_1 первого игрока и выигрыш S_2 второго игрока на цикле z определяются как среднее арифметическое их выигрышей на всех переходах, составляющих этот цикл:

$$S_1(z,p) = \frac{k_1 + k_2p^+ + k_3p^-}{k_1 + k_2 + k_3 + k_4}, \tag{4}$$

$$S_2(z,p) = \frac{k_2p^+ + k_3p^- + k_4}{k_1 + k_2 + k_3 + k_4} \tag{5}$$

Здесь через k_1 (или k_2, k_3, k_4) обозначено количество (кратность) выигрышей p^{++} (или p^+, p^-, p^{--}, соответственно) первого игрока на цикле z; вектор $p = (p^+, p^-)$ задает значения параметров игры.

Из леммы 1 следует, что при любых стратегиях игроков $h^1, h^2 \in \mathcal{H}$ игра носит циклический характер: начиная с некоторого шага траектория игры выходит на некоторый цикл z на графе игры и в дальнейшем остается на нем. При бесконечном повторении игры выигрышами, полученными в переходном режиме, можно пренебречь.

При заданных стратегиях игроков $h^1, h^2 \in \mathcal{H}$ на графе игры может существовать множество циклов $Z(h^1, h^2)$, являющееся неодноэлементным, причем выигрыши одного игрока на разных циклах могут отличаться.

Поэтому для определения выигрышей игроков при заданных стратегиях используется гарантированный подход:

$$f_1(h^1, h^2, p) = \min_{z \in Z(h^1,h^2)} S_1(z,p), \tag{6}$$

$$f_2(h^1, h^2, p) = \min_{z \in Z(h^1,h^2)} S_2(z,p) \tag{7}$$

Основным результатом второго параграфа является следующее утверждение.

Утверждение 1 *В симметричной постановке бесконечно повторяющейся дилеммы заключенного с конечной памятью игроков наибольшим гарантированным выигрышем каждого игрока является p^-.*

Из этого утверждения следует, что постановка игры с конечной памятью, в которой игроки равноправны, не решает основную проблему «дилеммы заключенного».

В третьем параграфе дается иерархическая постановка по Штакельбергу[3] повторяющейся «дилеммы заключенного» с конечной памятью с лидером – первым игроком. При этом считаются выполненными следующие условия:

$\mathbf{A}_1$. Лидер объявляет свою стратегию до начала игры и обязуется придерживаться ее.

$\mathbf{A}_2$. Ведомый рационален, то есть, зная стратегию лидера, он определяет свою стратегию так, чтобы максимизировать свой выигрыш.

При заданных параметров игры для каждой стратегии лидера $h^1 \in \mathcal{H}$ можно определить множество наилучших ответов ведомого,

$$\Phi(h^1, p) = \{h^2 \in \mathcal{H} |\, h^2 = \arg\max_{h^2 \in \mathcal{H}} f_2(h^1, h^2, p)\}, \tag{8}$$

которое, вообще говоря, может содержать более одной стратегии.

В этом случае гарантированный выигрыш лидера при использовании им стратегии $h^1 \in \mathcal{H}$ определяется по формуле

$$\mu(h^1, p) = \min_{h^2 \in \Phi(h^1, p)} f_1(h^1, h^2, p) \tag{9}$$

Таким образом, в иерархической постановке повторяющейся дилеммы заключенного с конечной памятью игроков возникает следующая задача.

Задача 1 *При заданных параметрах игры $p \in P$ найти стратегию лидера $\hat{h} = \hat{h}(p)$, максимизирующую его гарантированный выигрыш (9).*

Доказывается, что задача 1 разрешима в любой точке $p \in P$ за конечное число шагов методом полного перебора всех стратегий ведомого $h^2 \in \Phi(h^1, p)$ при каждой стратегии лидера $h^1 \in \mathcal{H}$.

Формулируется более общая задача.

Задача 2 *Построить разбиение области P на подобласти $\{P_i\}_{i=1}^n$, такие что для каждой подобласти P_i определена оптимальная стратегия лидера $\hat{h}_i$ (одна и та же для каждой точки $p \in P_i$).*

Разбиение, являющееся решением задачи 2, называется оптимальным.

В §3 главы 1 приведен алгоритм построения оптимального разбиения (АПОР), решающий эту задачу за конечное число шагов.

Работа алгоритма начинается выбором начального разбиения области P, при котором в каждой точке этой области лидер объявляет стратегию ALL D (опция D в каждой вершине графа игры). Оптимальным ответом ведомого при этом также является стратегия ALL D.

Дальнейшая работа алгоритма АПОР заключается в полном переборе всех стратегий лидера $h^1 \in \mathcal{H}$ и анализе рациональных откликов ведомого (8).

На каждом шаге этого процесса решается большое количество систем линейных неравенств. Эта работа выполняется с помощью специально разработанной программы. Существенно используется тот факт, что выигрыши игроков, задаваемые формулами (4) – (5), являются линейными функциями переменных p^+, p^-. В силу этого улучшение гарантированного выигрыша лидера происходит в областях, являющихся многоугольниками (возможно, невыпуклыми).

Поскольку количество стратегий обоих игроков конечно, то в результате их полного перебора получается оптимальное разбиение, которое является решением задачи 2.

Параграфы 4 и 5 первой главы посвящены изучению иерархической постановки повторяющейся «дилеммы заключенного» с памятью один и два шага, соответственно.

Анализ игры с одношаговой памятью в §4 проводится без применения вычислительной техники, так как количество возможных стратегий и циклов невелико – 16 и 24, соответственно. Основным результатом является определение оптимальной стратегии лидера. Во всех точках области P этой стратегией является TIT FOR TAT, описанная в работе Р.Аксельрода[22]. При этом выигрыши лидера и ведомого одинаковы и равны p^+.

При анализе игры с двухшаговой памятью в §5 предварительно анализируется проблема построения всех циклов с заданным набором кратностей. Для решения этой проблемы проводятся дополнительные теоретические построения, вводится способ представления цикла с помощью матрицы переходов.

Определение 1 *Пусть дан некоторый цикл z на графе игры. Матрица переходов этого цикла есть $T = \{t_{ij}\}_{i,j=1}^{4}$, где*

$$t_{ij} = \begin{cases} 1, & z \text{ содержит переходы с } i\text{-ой горизонтали на } j\text{-ую}, \\ 0, & \text{иначе.} \end{cases}$$

Показано, что на матрице переходов можно строить «отвлеченные» циклы. Каждому реальному циклу на графе игры соответствует единственный отвлеченный цикл, но каждому отвлеченному циклу можно поставить в соответствие несколько реальных циклов, по количеству различных перестановок кратностей этого цикла. Приведен алгоритм построения реального цикла по заданному отвлеченному, из которого, в частности, вытекает следующая лемма.

Лемма 2 *Если существует цикл z с набором кратностей $k_z = (k_1, k_2, k_3, k_4)$, то существует цикл $\tilde{z}$ с набором кратностей $k_{\tilde{z}} = (\tilde{k}_1, \tilde{k}_2, \tilde{k}_3, \tilde{k}_4)$, который является произвольной перестановкой набора кратностей k_z.*

Матрица переходов T, соответствующая циклу z с набором кратностей $k_z = (k_1, k_2, k_3, k_4)$, обладает следующим свойством: в ее i-ой строке (столбце) содержится в точности k_i единиц. Однако, не в каждой матрице, обладающей таким свойством, можно построить отвлеченный цикл, проходящий через все эти единицы, называемый полным циклом.

В §5 формулируются условия существования полного цикла, а также алгоритм его построения. Для этого используется следующее определение.

Определение 2 *Пусть дан набор чисел $k = (k_1, k_2, k_3, k_4)$, где $k_i \in \overline{0,4}$, $i = \overline{1,4}$. Назовем ассоциированным с ним набор чисел $n = (n_1, n_2, n_3, n_4)$, где n_i определяется как количество чисел в наборе k, не меньших i, $i = \overline{1,4}$.*

Основным результатом в исследовании вопроса существования цикла с заданным набором кратностей является следующий критерий.

Теорема 1 *Цикл z с набором кратностей $k = (k_1, k_2, k_3, k_4)$ на графе игры с двухшаговой памятью существует тогда и только тогда, когда выполнены условия*

$$\begin{aligned} n_1 &\geqslant k_1, \\ n_1 + n_2 &\geqslant k_1 + k_2, \\ n_1 + n_2 + n_3 &\geqslant k_1 + k_2 + k_3, \end{aligned} \tag{10}$$

где числа n_1, n_2, n_3 принадлежат набору $n = (n_1, n_2, n_3, n_4)$, ассоциированному с набором k.

Результатом применения алгоритма АПОР в игре с двухшаговой памятью является разбиение области $\boldsymbol{P}$ на шесть подобластей, в каждой из которых определены оптимальные стратегии лидера и ведомого. Существенное отличие от игры с одношаговой памятью заключается в том, что лишь в одной из этих подобластей лидер и ведомый имеют одинаковые выигрыши, равные p^{+}, а в остальных пяти подобластях выигрыш лидера больше выигрыша ведомого. В то же время, выигрыш ведомого в каждой точке области $\boldsymbol{P}$ превышает величину p^{-}.

Из этого следует вывод, что в иерархической постановке повторяющейся дилеммы заключенного увеличение памяти игроков с одного до двух шагов выгодно лидеру и не выгодно ведомому.

В §5 приводится некоторые количественные характеристики игры с двухшаговой памятью. В этой игре количество стратегий каждого игрока равно 65536, а количество циклов на графе игры равно 120538, что значительно больше аналогичных показателей игры с одношаговой памятью (16 и 24, соответственно). Все циклы были построены с помощью программы Memory2, разработанной при исследовании задачи. В их числе было найдено 20736 различных циклов, содержащих все 16 вершин графа игры с двухшаговой памятью.

Вторая глава посвящена изучению многокритериальной модели фирмы, описанной в работе Ф. П. де Вриза[33].

Идея поведенческой модели фирмы как противоположность моделям, рассматриваемым в неоклассической теории, была предложена в 60-х годах в работе Р. Сиерта и Дж. Марча[34].

Важным моментом, использованным в построении поведенческой модели фирмы, является введение желаемых уровней показателей качества отделов фирмы[35,36]. В поведенческой модели фирмы используется игровой подход к изучению ее динамики[37].

---

[33] De Vries F. P. *The Behavioral Firm and its Internal Game: Evolutionary Dynamics of Decision Making//* IIASA Interim Report IR-99-036, 1999.

[34] Cyert R., March J. *A Behavioral Theory of the Firm*. Prentice-Hall, Englewood Cliffs, 1963.

[35] Simon H. *A Behavioral Model of Rational Choice//* Quart. J. of Economics, 1955. Vol. 69. P. 99-118.

[36] Nelson R. R., Winter S. G. *An Evolutionary Theory of Economic Change*. The Belknap Press of Harvard Univ. Press, 1982.

[37] Bendor J., Mookherjee D., Ray D. *Aspirations, Adaptive Learning, and Cooperation in Repeated Games*. Mimeo, Indian Statistical Institute, 1994.

В первом параграфе приводится модель фирмы, состоящей из трех отделов: отдела производства, отдела продаж и отдела управления. Ее состояние описывается переменными (p, x), где $p \geqslant 0$ – стоимость единицы продукции, $x \geqslant 0$ – уровень слэка, или «ленивости» отдела производства.

Каждый отдел фирмы имеет свой показатель качества, зависящий от значений переменных p и x, а также от некоторых фиксированных положительных параметров системы $\theta, \alpha, \eta, \delta$.

Отдел производства (Product Department, PRD[33]) заинтересован в том, чтобы стоимость производства единицы продукции, вычисляемая по формуле

$$c(p,x) = \left(\frac{\theta}{\alpha}p^{\delta} + \eta\right)(1+x), \tag{11}$$

была как можно ближе к ее желаемому уровню, обозначаемому через $\tilde{c}$. Показателем качества PRD является величина

$$I_c(p,x) = |c(p,x) - \tilde{c}|. \tag{12}$$

В рассматриваемой модели считается, что объем продаж всегда соответствует спросу и определяется по формуле

$$w(p) = \frac{\alpha}{p^2}. \tag{13}$$

Отдел продаж (Sales Department, SLD[33]) заинтересован в том, чтобы объем продаж приближался к его желаемому уровню, обозначаемому через $\tilde{w}$. Таким образом, критерий качества SLD вычисляется по формуле

$$I_w(p) = |w(p) - \tilde{w}|. \tag{14}$$

Отдел менеджмента (Central Management Department, CMD[33]) заинтересован в том, чтобы прибыль фирмы, вычисляемая по формуле

$$\pi(p,x) = (p - c(p,x))w(p), \tag{15}$$

как можно меньше отличалась от ее желаемого уровня, обозначенного через $\tilde{\pi}$. Поэтому критерий качества CMD определяется формулой

$$I_\pi(p,x) = |\pi(p,x) - \tilde{\pi}|. \tag{16}$$

Задачей каждого отдела фирмы является минимизация соответствующего критерия, причем в рассматриваемой модели отделы менеджмента и производства управляют значениями переменных p и x, соответственно, а отдел продаж никак не может влиять на состояние фирмы.

В §1 второй главы вводятся в рассмотрение три важные кривые на плоскости (p,x), в точках которых показатели качества отделов фирмы принимают нулевые значения.

Решая уравнение $I_c(p,x)=0$, получим уравнение кривой

$$x=\tilde{\mu}(p)=\frac{\tilde{c}}{bp^2+\eta}-1. \tag{17}$$

Аналогично, из уравнения $I_\pi(p,x)=0$ найдем уравнение кривой

$$x=\tilde{\nu}(p)=\frac{p+\frac{\eta\tilde{\pi}}{\alpha b}}{bp^2+\eta}-\frac{\tilde{\pi}}{\alpha b}-1. \tag{18}$$

Наконец, показатель $I_w(p)$ достигает своего минимума на вертикальной прямой

$$p=\tilde{\varphi}=\sqrt{\frac{\alpha}{\tilde{w}}}. \tag{19}$$

Для постановки задачи управления фирмой необходимо заметить, что этот процесс по существу является кооперативным, и в нем должны быть учтены интересы всех участников, поэтому естественным образом возникает задача построения множества точек, оптимальных по Парето, относительно критериев $I_c(p,x), I_\pi(p,x), I_w(p)$. Это множество зависит от значений параметров $\tilde{c},\tilde{\pi},\tilde{w}$ и обозначается через $P(\tilde{c},\tilde{\pi},\tilde{w})$.

В [1] описано множество $P(\tilde{c},\tilde{\pi},\tilde{w})$ при произвольных $\tilde{c},\tilde{\pi},\tilde{w}$. В §2 второй главы диссертации рассматривается только случай, когда выполнено условие

$$\tilde{c}\tilde{\pi}<\frac{\alpha}{4} \tag{20}$$

В этом случае кривые $x=\tilde{\mu}(p)$ и $x=\tilde{\nu}(p)$ пересекаются в двух различных точках, левая из которых обозначена A_1, а правая – A_2. Область, ограниченная сверху кривой $x=\tilde{\nu}(p)$, а снизу кривой $x=\tilde{\mu}(p)$, обозначена через D. Через w_1 и w_2 обозначены параметры прямых семейства Φ_w, проходящих через точки A_1 и A_2, соответственно. В силу свойств линий семейства Φ_w имеет место неравенство $w_1>w_2$

Основным результатом второго параграфа являются два утверждения, описывающих структуру множества $P(\tilde{c}, \tilde{\pi}, \tilde{w})$ в зависимости от значения параметра $\tilde{w}$.

Утверждение 2 *Пусть выполнено условие $\tilde{w} > w_1$ ($\tilde{w} < w_2$), а параметры $\tilde{\pi}, \tilde{c}$ удовлетворяют условию (20). Обозначим через M точку пересечения прямой $p = \tilde{\varphi}$ с кривой $x = \tilde{\mu}(p)$, а через N – точку пересечения этой прямой с кривой $x = \tilde{\nu}(p)$.*

Тогда множество $P(\tilde{c}, \tilde{\pi}, \tilde{w})$ есть замыкание криволинейного треугольника A_1MN (соответственно, A_2MN).

Для исследования ситуации, когда выполнено условие $w_2 \leqslant \tilde{w} \leqslant w_1$, в §2 введена вспомогательная прямая γ, задаваемая параметрически, и следующие два множества.

Замыкание множества точек, расположенных в области D слева от прямой $p = \tilde{\varphi}$ и над кривой γ, обозначено D_L. Замыкание множества точек, расположенных в области D справа от прямой $p = \tilde{\varphi}$ и под кривой γ, обозначено D_R.

Формулируется следующее утверждение.

Утверждение 3 *При выполнении условия $w_2 \leqslant \tilde{w} \leqslant w_1$ оптимальными по Парето являются точки областей D_L, D_R и только они.*

В соответствии с принципами поведенческой модели фирмы, отделы менеджмента и производства не могут мгновенно изменить значения переменных p и x так, чтобы точка, задающая состояние системы, оказалась во множестве $P(\tilde{c}, \tilde{\pi}, \tilde{w})$. В связи с этим возникает задача приведения системы на множество $P(\tilde{c}, \tilde{\pi}, \tilde{w})$ с использованием управлений, имеющих достаточно простую экономическую интерпретацию.

Для определения класса допустимых управлений в параграфах 3 и 4 второй главы рассматриваются различные типы поведения[29,30] отделов менеджмента и производства.

Нормальным типом поведения отдела менеджмента называется такой способ изменения им переменной p, при котором значение критерия $I_\pi(p, x)$ монотонно убывает при постоянной величине переменной x.

Альтруистическим (*агрессивным*) поведением отдела менеджмента по отношению к отделу производства называется такое управление переменной p, при котором значение показателя $I_c(p, x)$ монотонно убывает (соответственно, возрастает) при постоянном значении переменной x.

Аналогично, ***альтруистическим*** (***агрессивным***) поведением отдела менеджмента по отношению к отделу продаж называется такое управление переменной p, при котором значение показателя $I_w(p)$ монотонно убывает (соответственно, возрастает) при постоянном значении переменной x.

Таким образом, определяется множество U допустимых типов поведения отдела менеджмента:

$$U = \{u^{nr}(p,x), u^{al}_{PRD}(p,x), u^{ag}_{PRD}(p,x), u^{al}_{SLD}(p,x), u^{ag}_{SLD}(p,x)\}$$

Аналогично вводятся типы поведения отдела производства. Управляя переменной x, этот отдел не может влиять на величину критерия $I_w(p)$, и поэтому имеет три различных типа поведения:

$$V = \{v^{nr}(p,x), v^{al}_{CMD}(p,x), v^{ag}_{CMD}(p,x)\}$$

Под управлением отдела менеджмента (производства) понимается выбор в каждой точке первого квадранта плоскости (p,x) некоторой функции $u(p,x) \in U$ (соответственно, $v(p,x) \in V$). При заданных управлениях движение системы описывается дифференциальными уравнениями

$$\begin{cases} \dot{p} &= u(p,x), \\ \dot{x} &= v(p,x). \end{cases} \tag{21}$$

В §5 второй главы ставится задача нахождения условий, при которых траектория движения системы (21) приходит на множество $P(\tilde{c},\tilde{\pi},\tilde{w})$ с использованием только нормальных типов поведения, то есть при выполнении в каждой точке первого квадранта плоскости (p,x) условия

$$u(p,x) = u^{nr}(p,x),\ v(p,x) = v^{nr}(p,x).$$

Доказано, что эти условия задают множество меры нуль в пространстве параметров $(\tilde{c},\tilde{\pi},\tilde{w})$, то есть почти всегда использования нормальных типов поведения недостаточно.

В связи с этим ставится задача приведения системы на множество $P(\tilde{c},\tilde{\pi},\tilde{w})$ с использованием типов поведения, не являющихся нормальными. При этом время использования таких типов поведения необходимо минимизировать.

Эта задача решена в §6 при выполнении условия $\tilde{\pi}\tilde{c} < \frac{\alpha}{4}$ и некоторых дополнительных ограничениях. Результаты исследования сформулированы в этом параграфе в виде ряда утверждений, в каждом из которых описывается оптимальное управление в некоторой области первого квадранта плоскости (p, x).

ПУБЛИКАЦИИ ПО ТЕМЕ ДИССЕРТАЦИИ

[1] Клейменов А.Ф., Семенищев А.А. *Построение множества оптимальных по Парето точек в одной многокритериальной задаче управления фирмой*/ ИММ УрО РАН. Екатеринбург, 2001. 87 с. Деп. в ВИНИТИ 27.03.2001, №752-В2001.

[2] Клейменов А.Ф., Семенищев А.А. *Построение решений в одной многокритериальной задаче управления фирмой*// Вестн. Тамбов. гос. ун-та. 2000. № 5. С. 458-459.

[3] Семенищев А.А. *Решение повторяющейся 2x2 биматричной игры с двухшаговой памятью в иерархической постановке*// Тез. докл. 30-ой регион. молодеж. конф. «Проблемы теорет. и прикл. математики» / ИММ УрО РАН. Екатеринбург, 1999. С. 68-69.

[4] Семенищев А.А. *Стратегии Штакельберга для повторяющейся 2*2 биматричной игры с двухшаговой памятью*// Тез. докл. Воронеж. весен. мат. шк. «Современные методы в теории краевых задач». Воронеж: Изд-во Воронеж. ун-та. 1999. С. 222.

[5] Семенищев А.А. *Решение повторяющейся Дилеммы Заключенного с конечной памятью игроков в иерархической постановке*/ Урал. ун-т. Екатеринбург, 2000. 63 с. Деп. в ВИНИТИ 24.10.00, №2706-В00.

[6] Kleimenov A.F., Semenishchev A.A. *Repeated Prisoner's Dilemma: Stackelberg Solution with Finite Memory*// Control Applications of Optimization: 11-th IFAC Workshop, July 3-6, 2000: Proc. Pergamon: Elsevier Sci., 2000. P. 567-574.

[7] Semenishchev A.A. *Solution of the Repeated Prisoner's Dilemma With Finite Memory In Hierarchical Statement*// Game Theory and Appl., 2001. Vol.6. P. 141-163.

Подписано в печать 07.08.2001. Формат 60×84 1/16.
Бумага типографская. Усл. печ. л. 1.
Тираж 70 экз. Заказ №141. Печать офсетная.
Екатеринбург, К-83, пр. Ленина, 51.
Типолаборатория УрГУ.