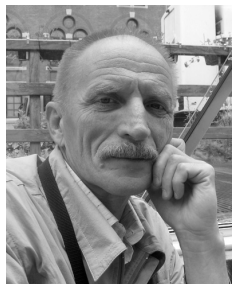


РЕФЛЕКСИВНЫЙ АНАЛИЗ БИМАТРИЧНЫХ ИГР КАК РЕШЕНИЕ ЗАДАЧИ ВЫБОРА В МОДЕЛЯХ ПРОТИВОДЕЙСТВИЯ

В. В. Карюкин, Ф. С. Чаусов



ВУНЦ «Военно-морская академия», профессор, доктор физико-математических наук



ВУНЦ «Военно-морская академия», профессор, доктор военных наук

1 ОСНОВНЫЕ ПОНЯТИЯ И ОПРЕДЕЛЕНИЯ

Одним из способов принятия решений в условиях неполной информированности является формирование игровой модели и определение ее как решения некоторой оптимизационной задачи. Далее мы будем рассматривать игры двух лиц, где в качестве решения используется понятие равновесия Нэша.

Введем необходимые понятия и определения. Рассмотрим двух игроков. Пусть в данной игре может выбрать по своему усмотрению один из возможных ходов (стратегий) A_i , $i \in \{1, n\} = N$, игрок B – m стратегий B_j , $j \in \{1, m\} = M$. Если игрок A выберет стратегию A_i , а игрок B – стратегию B_j , то первый игрок получает выигрыш A_{ij} , соответственно, выигрыш второго равен B_{ij} . Игры такого сорта называются биматричными. Выбор стратегии каждым из игроков производится одновременно и независимо друг от друга. В этом и заключается в данном случае неполная информированность.

Каждый игрок старается максимизировать свой выигрыш, хотя ему не известен будущий выбор противника. Поэтому в качестве решения игры используется решение некоторой оптимизационной задачи. Наиболее популярное в настоящий момент решение Неймана заключается в выборе стратегий A_i, B_j , так что [1, 2]:

$$A_{i_0} = \mathop{\text{Arg max}}_{i \in N} \min_{j \in M} A_{i,j}, \quad (1)$$

$$B_{j_0} = \mathop{\text{Arg min}}_{j \in M} \max_{i \in N} B_{i,j}.$$

Если решение данной задачи существует, то стратегии A_{i_0} , B_{j_0} реализуют равновесие Неймана и они называются максиминными (минимаксными). Отметим, что приведенная формулировка экстремальной задачи характерна именно для биматричных задач противоборства, в которой выигрыш одного игрока приводит к проигрышу другого.

Поскольку равновесие как Неймана, так и Нэша в чистых стратегиях существует не для всех игр, то вводится понятие смешанных стратегий.

Пусть на стратегиях игрока A задано распределение вероятностей x_i , $i \in N$ и пусть X – множество всевозможных распределений x_i . Аналогично для игрока B имеем распределение y_j , $j \in M$ и Y – множество таких распределений. Тогда выигрыш игроков описывается его математическим ожиданием:

$$\begin{aligned} F_A(x, y) &= x^t A y, & x \in R^n, & \quad y \in R^m \\ F_B(x, y) &= x^t B y, & x \in R^n, & \quad y \in R^m \end{aligned} \quad (2)$$

а максиминные смешанные стратегии есть решение задач:

$$\begin{aligned} x^* &= \mathop{\text{Arg max}}_{x \in X} \min_{y \in Y} F_A(x, y) \\ y^* &= \mathop{\text{Arg min}}_{y \in Y} \max_{x \in X} F_B(x, y). \end{aligned} \quad (3)$$

Равновесие Нэша определяется более сложным образом [2–6]. Определим его сразу для смешанных стратегий игроков. Пусть игрок произвел выбор, и решил воспользоваться своей какой-то конкретной смешанной стратегией y^* . Очевидно, что в этих условиях игрок должен использовать свою смешанную стратегию:

$$x'(y^*) \in \mathop{\text{Arg max}}_{x \in X} F_A(x, y^*), \quad (4)$$

которая дает ему максимальный выигрыш при данной стратегии игрока B .

Вектор $x'y^*$ называется наилучшим ответом игрока на выбор стратегии y^* игроком B . Символ $\in$ подчеркивает, что наилучший ответ может быть не единственным. В связи с этим вводится понятие множества наилучших ответов:

$$\text{Best}_-x(y^*) = \mathop{\text{Arg max}}_{x \in X} F_A(x, y^*). \quad (5)$$

Если $x' = x'(y)$ – наилучший ответ на y , то, очевидно, он должен удовлетворять неравенству:

$$F_A(x', y) \geq F_A(x, y). \quad (6)$$

Пусть $y' = y'(x)$ – наилучший ответ игрока B на выбор x , тогда:

$$F_B(x, y') \geq F_B(x, y). \quad (7)$$

Пара смешанных стратегий $x^* \in R^n$ и $y^* \in R^m$ реализует равновесие Нэша, если одновременно имеют место два неравенства [2]:

$$F_A(x^*, y^*) \geq F_A(x, y^*) \quad (8)$$

$$F_B(x^*, y^*) \geq F_B(x^*, y). \quad (8'')$$

Обозначим [6] $BR_X = \bigcup_{y \in Y} Best_x(y) \times \{y\}$ и $BR_Y = \bigcup_{x \in X} \{x\} \times Best_y(x)$.

Тогда множество равновесных по Нэшу точек (x^*, y^*) можно найти как результат пересечения BR_X и BR_Y . Отметим, что равновесие Нэша в смешанных стратегиях существует для всех конечных биматричных игр. Точек, равновесных по Нэшу, в игре может быть несколько.

В рассматриваемых далее играх элементы матриц интерпретируются как вероятности выигрыша соответствующего игрока. Поэтому считаем выполненным равенство $a_{i,j} + b_{i,j} = 1$.

2 ГЕОМЕТРИЧЕСКАЯ ИНТЕРПРЕТАЦИЯ РАВНОВЕСИЯ НЭША

Рассмотрим простейшую задачу поиска из класса задач противодействия. Пусть имеется три области поиска с разной степенью сложности обнаружения объекта. Игрок прячется, а игрок ищет. Игрок выбирает область, в которой предпочтительней прятаться, игрок область, в которой необходимо искать. После того, как выбор произведен, менять область поиска нельзя.

Если разведчик выбрал область, в которой находится объект поиска, то вероятности обнаружения объекта равны: 0,1 для первой области, 0,5 для области со средней сложностью поиска и 0,9 для последней области. Для первого игрока игра описывается матрицей A , и матрицей B — для второго:

$$A = \begin{pmatrix} 0,9 & 1 & 1 \\ 1 & 0,5 & 1 \\ 1 & 1 & 0,1 \end{pmatrix} \quad (9)$$

$$B = \begin{pmatrix} 0,1 & 0 & 0 \\ 0 & 0,5 & 0 \\ 0 & 0 & 0,9 \end{pmatrix}$$

Нетрудно найти максиминное решение игры для игрока. Оно является решением в чистых стратегиях и соответствует выбору области, в которой поиск затруднен в максимальной степени. Таким образом, если игрок все время выбирает данную область, то вероятность его успеха не может быть меньше 0,9. Это наилучший гарантированный результат. Аналогично, для игрока выбор той же области является минимаксным решением игры. Другими словами, вероятность его успеха не может превышать 0,1.

Найдем решение Нэша для данной игры. Для этого введем в рассмотрение целевые функции игроков в виде:

$$F_A(X, Y) = X' A Y \quad \text{где } X' = (x_1 \ x_2 \ 1 - x_1 - x_2)$$

$$F_B(X, Y) = X' B Y, \quad (10)$$

Градиент первой целевой функции по вектору X равен:

$$\frac{\partial}{\partial x_1} F_A(x_1, x_2, y_1, y_2) = -1,0 \cdot y_1 + 0,9 - 0,9 \cdot y_2$$

$$\frac{\partial}{\partial x_2} F_A(x_1, x_2, y_1, y_2) = -1,4 \cdot y_2 + 0,9 - 0,9 \cdot y_1 \quad (11)$$

Множество допустимых значений переменных $(y_1; y_2) \in \Delta$, где Δ – внутренние точки и граница треугольника с вершинами $(0; 0)$, $(0; 1)$, $(1; 0)$. Рассмотрим точки данного треугольника, для которых обе координаты вектора градиента положительны. Представим данное множество как объединение точек, для которых первая координата градиента больше второй, множества для которого вторая координата больше, и множества точек с равными координатами. Данные множества представлены на рисунке 1.

Множество наилучших ответов [3] $(x_1; x_2)$ игрока A на смешенные стратегии игрока из областей, приведенных на рисунке 1, приведены на рисунке 2.

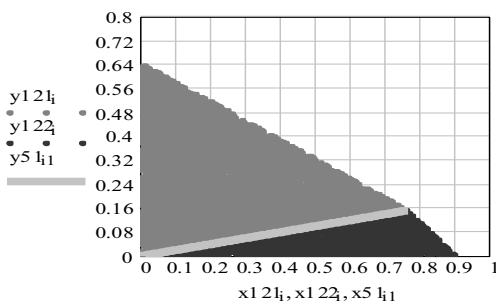


Рис. 1. Множество точек, для которых вектор градиента целевой функции объекта поиска имеет положительные координаты

Примечание: более темная область – множество точек, на которых обе координаты градиента положительны и вторая координата больше первой; более светлая область – множество точек, на которых обе координаты градиента положительны и вторая координата меньше первой; прямая – множество точек с равными положительными координатами градиента.

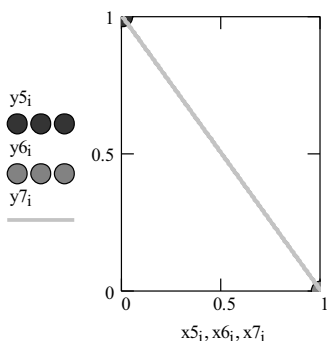


Рис. 2. Множество наилучших ответов игрока на смешанные стратегии игрока из рисунка 1

Примечание: Множество оптимальных субъективных стратегий первого игрока при условии, что второй игрок пользуется стратегиями из закрашенных областей рисунка 1. Здесь: вектор $x_1 = 1, x_2 = 0$ – оптимальный ответ на стратегию (y_1, y_2) из более светлого множества на рисунке 1; прямая – множество оптимальных стратегий $x_1 + x_2 = 1$ при выборе (y_1, y_2) на прямой $(y_2 = 0, 2y_1)$ на рисунке 1.

Множество точек $(y_1; y_2) \in \Delta$, для которых хотя бы одна координата отрицательна, изображено на рисунке 3.

Множество наилучших ответов на стратегии рисунка 3 приведено на рисунке 4.

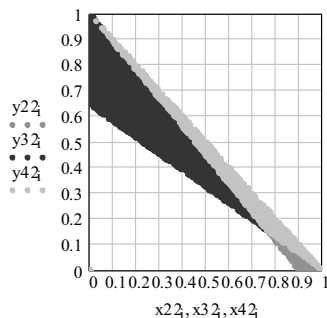


Рис. 3. Множество смешанных стратегий игрока, для которых хотя бы одна координата градиента целевой функции объекта поиска отрицательна.

Примечание: Здесь: самая темная область – первая координата положительна, вторая отрицательна; более светлая область – первая координата отрицательна, а вторая положительна; самая светлая область – обе координаты отрицательны.

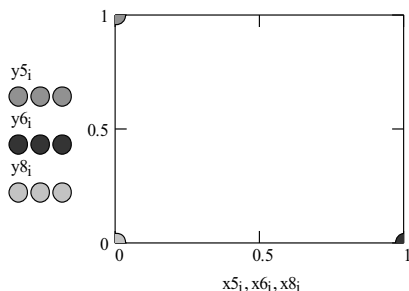


Рис. 4. Множество наилучших ответов на стратегии рисунка 3

Аналогичным образом можно найти множество оптимальных ответов второго игрока на смешанные стратегии первого. Составляя множество $BR_X \cap BR_Y$ (см. п. 1), находим равновесие Нэша в смешанных стратегиях: для объекта поиска

$$(x1^* x2^* x3^*) = (0,763 \ 0,152 \ 0,085),$$

и для ищущего

$$(y1^* y2^* y3^*) = (0,763 \ 0,152 \ 0,085). \quad (12)$$

Вероятности выигрыша (значения целевых функций) в точке Нэша равны $F_A(x^*, y^*) = 0,924$ и $F_B(x^*, y^*) = 0,076$. (13)

Как видно решение Нэша по сравнению с минимаксным решением увеличивает вероятность выигрыша объекта поиска и уменьшает вероятность выигрыша разведчика. Для данной игры это достаточно очевидный факт. При использовании минимаксного решения объект поиска всегда выбирает область с наиболее сложными условиями поиска, а при использовании решения Нэша он изредка (с вероятностью 0,237) меняет область, а разведчик ее не угадывает.

Однако, данный выигрыш есть математическое ожидание выигрыша, и использование смешанных стратегий, равновесных по Нэшу, оправдано при многократном повторении игры. При однократной игре увеличение математического ожидания выигрыша для объекта поиска связано с увеличением риска быть обнаруженным в случае необходимости реализации области, отличной от самой сложной для поиска.

Отметим еще одну особенность равновесия Нэша в задачах противодействия. Если один из игроков использует решение Нэша, то математическое ожидание выигрыша второго не зависит от выбранной им стратегии. Так, если первый игрок использует свои стратегии $(x1^* x2^* x3^*) = (0,763 \ 0,152 \ 0,085)$, то второй никак не может увеличить свой выигрыш.

Вообще говоря, свойство решения Нэша таково, что при отклонении вторым игроком от решения Нэша его выигрыш может лишь умень-

шиться. Отметим также, что если один из игроков использует свою оптимальную по Нэшу стратегию, то его выигрыш не зависит от того, какие стратегии использует второй игрок. Соответственно наоборот, если второй использует свои оптимальные смешанные стратегии $(y_1^*, y_2^*, y_3^*) = (0,763 \ 0,152 \ 0,085)$, то вне зависимости от стратегий первого он получает одно и то же значение функции ценности. В итоге для рассматриваемых нами задач противоборства правомерен вывод: если хотя бы один из игроков использует свои равновесные по Нэшу стратегии, то выигрыши обеих постоянны и равны оптимальным, вне зависимости от смены стратегий другим игроком.

3 ИГРЫ С УПРАВЛЯЮЩИМ ВОЗДЕЙСТВИЕМ

Для рассмотрения подобных игр понадобится другое представление игры. Игры, задаваемые матрицами, как это сделано в разделе 1, называются играми в нормальной форме. Представим игру из раздела 1 в развернутой форме, как это сделано на рисунке 5. Здесь пунктирными овалами выделены информационные множества (для игрока 1 — нижний овал, для игрока 2 — верхний овал), в которых они производят выбор. Овал означает, что, производя свой выбор, игроки не знают, какой выбор был сделан его противником. Приведенная развернутая форма игры точно соответствует нормальной форме, заданной матрицами. Цифры, написанные над терминальными вершинами графа, означают выигрыши первого и второго игрока соответственно. Рассмотрим влияние управляющего воздействия на примере рассмотренной ранее задачи разведки.

Для большей образности изложения будем теперь считать, что области поиска — это комнаты разной освещенности: темная с вероятностью обнаружения 0,1, комната средней освещенности с вероятностью 0,5, и светлая с вероятностью обнаружения 0,9.

Пусть правила игры остаются прежними, но введем в рассмотрение игрока 3, который повесил над комнатой средней освещенности красную лампочку.

Информационное наполнение воздействия игрока 3 может быть самым разнообразным. Например, обоим игрокам было сообщено, что средняя комната плохо (хорошо) соответствует игре в прятки. Формально введение красной лампочки можно рассматривать как указание на то, что все множество комнат разбивается на два подмножества: комнату средней освещенности и остальные. В силу изменившейся информационной обстановки игры, каждый из игроков сначала делает выбор между средней комнатой и остальными.

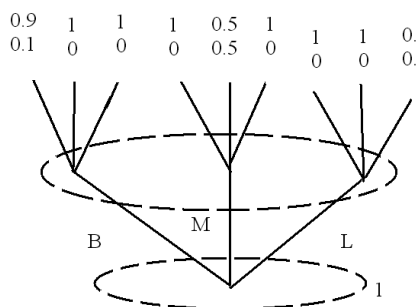


Рис. 5. Развернутая форма игры из пункта 2

Развернутая форма данной игры имеет вид, представленный на рисунке 6.

По сравнению с предыдущим деревом управляющее воздействие породило два новых информационных множества. Теперь множества 11 и 12, в которых производится выбор альтернатив для игрока, называются его агентами. Соответственно агенты 21 и 22 производят выбор в пользу игрока 2. Такая игра уже не описывается нормальной формой игры. Зельтман и Харшаньи для таких игр ввели форму названную ими стандартной формой. При наличии только по одному агенту для каждого из игроков стандартная форма совпадает с нормальной. Стандартная форма по существу задается двумя блочными матрицами, в которых каждый из блоков соответствует выбору игроком своих агентов.

Агент 11 выбирает, где прятаться в комнате с красной лампочкой или в остальных. Агент 21 производит тот же выбор для игрока 2. Агент 12

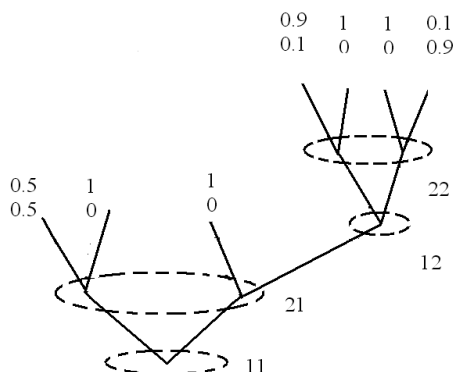


Рис. 6. Развернутая форма игры из пункта 1 с управляющим воздействием третьего игрока

выбирает из двух оставшихся комнат: самую темную или самую светлую. Агент 22 производит тот же выбор для игрока 2. Цифры над терминальными вершинами указывают выигрыши игроков в результате реализации соответствующего им пути в дереве игры. Верхняя цифра выигрыш игрока 1, нижняя – второго участника игры.

Рассмотрим для данной игры следующую ситуацию. Пусть оба игрока заметили появление красной лампочки. Это положение в точности соответствует предположению о рациональности поведения игроков. Поиск решения данной задачи несколько сложнее предыдущей, хотя, можно воспользоваться ее решением и пересчитать вероятности переходов в указанные множества новой задачи.

Оказывается, что полученные таким образом значения целевых функций в точности совпадают со значениями (13) и обладают свойствами оптимальности. То есть данная задача при объективном стремлении игроков к равновесию Нэша эквивалентна предыдущей по целевой функции.

Усложним постановку задачи условием, что третий игрок не только устанавливает красную лампочку над комнатой средней освещенности, но и выбирает какая из двух оставшихся комнат будет самой темной, а какая самой светлой. Здесь оба игрока знают, какая из комнат является комнатой средней освещенности из-за красной лампочки. Об остальных комнатах они знают, что одна из них является самой темной, а одна самой светлой, но какая именно – им неизвестно (см. рисунок 7).

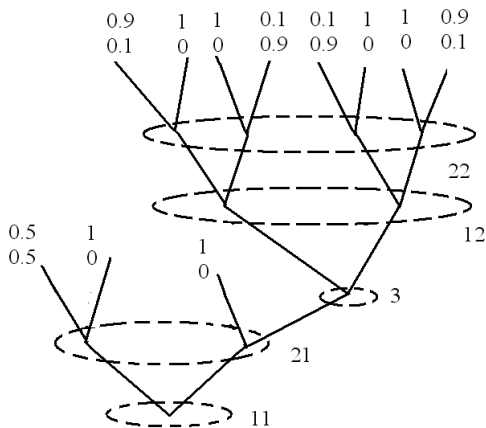


Рис. 7. Развернутая форма игры из пункта 1 с управляющим воздействием третьего игрока и отсутствием информации об освещенности двух комнат

Агент 11 выбирает, где прятаться в комнате с красной лампочкой или в остальных. Агент 21 производит тот же выбор для игрока 2. Агент 12 выбирает из двух оставшихся комнат: самую темную или самую светлую. Агент 22 производит тот же выбор для игрока 2. Игрок 3 выбирает, какую из двух комнат сделать самой светлой, а какую самой темной. Цифры над терминальными вершинами указывают выигрыши игроков в результате реализации соответствующего им пути в дереве игры. Верхняя цифра выигрыш игрока 1, нижняя – второго участника игры.

Нетрудно составить игры, в которых управляющее воздействие, подобное приведенному выше, делает один из игроков. Естественно рассматривать и задачи о наиболее выгодном управляющем воздействии.

Рассмотрим примеры управляющих воздействий объекта поиска. В начале игры двери во все комнаты закрыты и это известно обоим игрокам. Пусть ищущий обнаружил, что дверь в комнату средней освещенности оказалась слегка приоткрытой. Ясно, что решение, принимаемое ищущим в данной ситуации, зависит от качества и количества имеющейся у него информации.

Возможные варианты информации таковы:

- А Известно, что прячущийся – человек рассеянный, не придает особого значения участию в данной игре и не склонен к совершению ложных ходов. В этом случае можно считать, что вероятность выбора им средней комнаты равна единице и выбор ищущего очевиден.
- Б Пусть, наоборот, известно, что прячущийся – мастер на всякого рода хитрости. Ясно, что в этом случае ищущий должен сузить поиск до двух оставшихся комнат.
- В В более общей ситуации имеется набор логических утверждений за и против выбора комнаты средней освещенности в качестве объекта для поиска.

Можно считать, что мы имеем некоторую вспомогательную задачу в информационном пространстве утверждений (информационную игру), результатом решения которой является вероятность выбора прячущимся комнаты средней освещенности. Тогда оптимальная поведенческая стратегия ищущего есть решение игры, в которой ветвям агента 11 на рисунке 6 приписаны данные вероятности.

4 РЕФЛЕКСИВНЫЕ ИГРЫ

Рассмотрим процесс рефлексивного анализа игровой ситуации и построения рефлексивной игры на примере той же игры в прятки. Будем считать, что, если ищущий выбрал комнату, в которой находится прячу-

щийся, то он выиграл. Приведем фрагмент из файла, решающего данную задачу.

Задача [6]. Первый игрок прячется в одной из трех комнат разной освещенности, второй игрок выбирает комнату, в которой его следует искать. Выигрывает объект поиска, если разведчик выберет комнату, в которой отсутствует объект поиска. Если разведчик выбирает комнату, в которой находится объект поиска — он выигрывает.

Решение игры на основании рефлексивного анализа

Обозначим темную комнату — В, средней освещенности — М, и светлую комнату — L.

Рефлексивные стратегии игроков

Разведчик предпочитает искать там, где светлее. Объект поиска, естественно, предпочитает комнату, в которой его искать труднее.

Поэтому при **нулевом** ранге рефлексии выбор таков:

Объект поиска — самая темная (событие В), разведчик — самая светлая (событие L).

Первый ранг рефлексии: Разведчик догадывается, о выборе объекта поиска и, следовательно, его выбор — В. В свою очередь, объект поиска догадывается о приоритетах разведчика (событие L) и, следовательно, его выбор — $L = B \vee M$.

Второй ранг: разведчик — $L = B \vee M$, объект поиска — $B = L \vee M$.

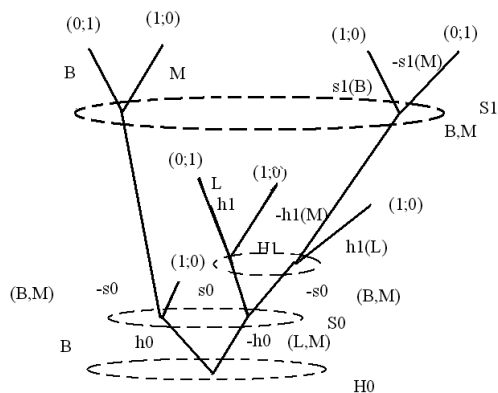
Третий ранг: разведчик — $B = L \vee M$, объект поиска — L.

Четвертый ранг: нетрудно убедиться в том, что он совпадает с нулевым.

Для построения дерева игры введем иерархию уровней рефлексии. Сначала реализуется нулевой уровень рефлексии. Если он не дает решения игры, то на дополнениях нулевого уровня реализуется первый уровень и т. д. В результате получаем дерево рефлексивной игры (дерево рефлексивной логики).

Агент H0 объекта поиска выбирает: прятаться в самой темной комнате (гипотеза h0) или в остальных (отрицание гипотезы h0). Агент S0 разведчика решает: прятаться в самой светлой комнате (гипотеза s0) или в остальных (отрицание гипотезы s0). Агент H1 решает для объекта поиска задачу первого ранга рефлексии: прятаться в самой светлой комнате (гипотеза h1) или в комнате средней освещенности (отрицание гипотезы h1). Агент S1 разведчика решает: прятаться в самой темной комнате (гипотеза s1) или в комнате средней освещенности (отрицание гипотезы s1). Цифры над терминальными вершинами указывают выигрыши игроков в результате реализации соответствующего им пути в дереве игры. Первая цифра выигрыш игрока 1, вторая — второго участника игры

Отметим важное обстоятельство полученных результатов: дерево рефлексивной логики было получено на основе анализа имеющейся ин-



формации о сложности поиска объекта. Решая данную информационную игру, получаем распределения вероятностей на множестве логических утверждений, связанных с игрой. Из данных распределений получаем распределение на множестве стратегий исходной игры, которые используются в качестве априорных распределений для нахождения оптимальной стратегии.

Вернемся к исходной рефлексивной игре с бинарной целевой функцией. Кроме данной игры, на рангах рефлексии можно сформулировать и игру в нормальной форме. При этом оказывается, что обе игры логически эквивалентны по значению целевых функций. Анализ игры в нормальной форме интересен и с точки зрения формальной логики, поскольку там возникает множество доминирующих и дублирующих стратегий.

Отметим, что как игры с частичной информированностью, так и построение моделей информированности о состоянии игры и о намерениях друг друга являются важными для решения задач поставленных в названии статьи и для общей теории игр.

В заключение авторы выражают уверенность, что многовариантный рефлексивный анализ [7] игровой ситуации, сопровождаемый строгой математической формализацией, позволит сделать теорию игр действенным инструментом в решении разнообразных прикладных задач [8].

ЛИТЕРАТУРА

- 1 Нейман Д., Моргенштерн О. Теория игр и экономическое поведение. М.: Наука. 1970.
- 2 Оуэн Г. Теория игр. М.: Мир, 1971.

- 3 *Харшаньи Д., Зельтен Р.* Общая теория выбора равновесия в играх. СПб.: СПбГУ, 2001.
- 4 *Вентцель Е. С.* Исследование операций. М.: Советское радио, 1972.
- 5 *Данилов В. И.* Лекции по теории игр. М.: Российская экономическая школа. 2002
- 6 *Новиков Д. А., Чхарташвили А. Г.* Рефлексивные игры. М.: СИНТЕГ, 2003.
- 7 *Левевер В. А.* Элементы логики рефлексивных игр. Проблемы инженерной психологии. Вып. 4. Ленинград 1966, с. 273–299.
- 8 *Чаусов Ф. С.* Рефлексивный анализ в сфере управления. Монография. СПб.: СПбВМИ, 2006.