

*М.В. Губко
Д.А. Новиков
А.Г. Чхартишвили*

ЭЛЕМЕНТЫ ТЕОРИИ ИГР

Настоящий материал содержит описание основных понятий и моделей теории игр. В том числе кратко рассматриваются: некооперативные игры, кооперативные игры и иерархические игры. Для более полного ознакомления с проблематикой и результатами использования теоретико-игровых моделей в задачах управления организационными системами можно порекомендовать учебники и монографии [6, 9, 12, 16-18].

Некооперативные игры

Рассмотрим *игровую неопределенность*, отражающую совместное принятие решений несколькими агентами (при заданных управлениях со стороны центра), в рамках которой существенными являются предположения агента о множестве возможных значений *обстановки игры* (действий других агентов, выбираемых ими в рамках тех или иных неточно известных рассматриваемому агенту принципов их поведения).

Для описания коллективного поведения агентов недостаточно определить их предпочтения и правила индивидуального рационального выбора по отдельности. В случае, когда в системе имеется единственный агент, гипотеза его рационального (индивидуального) поведения предполагает, что агент ведет себя таким образом, чтобы выбором действия максимизировать значение своей целевой функции. В случае, когда агентов несколько, необходимо учитывать их взаимное влияние: в этом случае возникает *игра* – взаимодействие, в котором выигрыш каждого агента зависит как от его собственного действия, так и от действий других агентов. Если в силу гипотезы рационального поведения каждый из агентов стремится выбором действия максимизировать свою целевую функцию, то понятно, что в случае нескольких агентов индивидуально

рациональное действие каждого из них зависит от действий других агентов.

Рассмотрим теоретико-игровую модель *некооперативного взаимодействия* между n агентами, предполагая, что они принимают решения одновременно и независимо, не имея возможности договариваться о выбираемых действиях, перераспределять получаемую полезность (выигрыш) и т.д.

Каждый агент осуществляет выбор *действия* x_i , принадлежащего *допустимому множеству* X_i , $i \in N = \{1, 2, \dots, n\}$ – множеству агентов. Выбор действий агентами осуществляется однократно, одновременно и независимо.

Выигрыш i -го агента зависит от его собственного действия $x_i \in X_i$, от вектора действий

$$x_{-i} = (x_1, x_2, \dots, x_{i-1}, x_{i+1}, \dots, x_n) \in X_{-i} = \prod_{j \in N \setminus \{i\}} X_j$$

оппонентов $N \setminus \{i\}$ и от состояния природы¹ $\theta \in \Omega$, и описывается действительнозначной функцией выигрыша $f_i = f_i(\theta, x)$, где $x = (x_i, x_{-i}) = (x_1, x_2, \dots, x_n) \in X' = \prod_{j \in N} X_j$ – вектор действий всех

агентов. При фиксированном значении состояния природы совокупность $\Gamma = (N, \{X_i\}_{i \in N}, \{f_i(\cdot)\}_{i \in N})$ множества агентов, множеств их допустимых действий и целевых функций называется *игрой в нормальной форме*. *Решением игры (равновесием)* называется множество устойчивых в том или ином смысле векторов действий агентов [6].

В силу гипотезы рационального поведения каждый агент будет стремиться выбрать наилучшие для него (с точки зрения значения его целевой функции) действия при заданной обстановке. *Обстановкой* для него будет совокупность состояния природы $\theta \in \Omega$ и обстановки игры

$$x_{-i} = (x_1, x_2, \dots, x_{i-1}, x_{i+1}, \dots, x_n) \in X_{-i} = \prod_{j \in N \setminus \{i\}} X_j.$$

Следовательно, принцип принятия им решения о выбираемом действии (при фиксированных обстановке и состоянии природы)

¹ Состояние природы может быть, в том числе, вектором, компоненты которого отражают индивидуальные характеристики (типы) агентов.

можно записать следующим образом (BR обозначает *наилучший ответ – best response*):

$$BR_i(\theta, x_{-i}) = \operatorname{Arg} \max_{x_i \in X_i} f_i(\theta, x_i, x_{-i}), i \in N.$$

Рассмотрим возможные принципы принятия решений агентами, каждый из которых порождает соответствующую концепцию равновесия, то есть определяет, в каком смысле устойчивым должен быть прогнозируемый исход игры.

Равновесие в доминантных стратегиях. Если для некоторого агента при любом состоянии природы множество его наилучших ответов не зависит от обстановки, то оно составляет множество его доминантных стратегий (совокупность доминантных стратегий агентов называется *равновесием в доминантных стратегиях – РДС*) [6]. Если у каждого из агентов существует доминантная стратегия, то они могут принимать решения независимо, то есть выбирать действия, не имея никакой информации и не делая никаких предположений об обстановке. К сожалению, РДС существует далеко не во всех играх.

Для реализации агентами РДС, если последнее существует, достаточно знания каждым из них только своей целевой функции и допустимых множеств X' и Ω .

Гарантирующее равновесие. Той же информированностью должны обладать агенты для реализации *гарантирующего (максиминного) равновесия*, которое существует почти во всех играх:

$$x_i^* \in \operatorname{Arg} \max_{x_i \in X_i} \min_{x_{-i} \in X_{-i}} \min_{\theta \in \Omega} f_i(\theta, x_i, x_{-i}), i \in N.$$

Равновесие Нэша. Определим многозначное отображение

$$BR(\theta, x) = (BR_1(\theta, x_1); BR_2(\theta, x_2), \dots, BR_n(\theta, x_n)).$$

Равновесием Нэша [6] при состоянии природы θ (точнее – *параметрическим равновесием Нэша*) называется точка $x^*(\theta) \in X'$, удовлетворяющая следующему условию:

$$x^*(\theta) \in BR(\theta, x^*(\theta)).$$

Последнее вложение можно также записать в виде:

$$\forall i \in N, \forall y_i \in X_i \quad f_i(\theta, x^*(\theta)) \geq f_i(\theta, y_i, x_{-i}^*(\theta)).$$

Множество $E_N(\theta)$ всех точек вида $x^*(\theta)$ можно описать следующим образом:

$$E_N(\theta) = \{x \in X' \mid x_i \in BR_i(\theta, x_{-i}), i \in N\}.$$

Для реализации равновесия Нэша достаточно, чтобы рациональность агентов и все параметры игры, а также значение состояния природы были *общим знанием* [15], то есть каждый из агентов рационален, знает множество участников игры, целевые функции и допустимые множества всех агентов, а также знает значение состояния природы. Кроме того, он знает, что другие агенты знают это, а также то, что они знают, что он это знает и т.д. до бесконечности.

Субъективное равновесие. Рассмотренные виды равновесия являются частными случаями *субъективного равновесия*, которое определяется как вектор действий агентов, каждая компонента которого является наилучшим ответом соответствующего агента на ту обстановку игры, которая может реализоваться с его субъективной точки зрения. Рассмотрим возможные случаи.

Предположим, что i -ый агент рассчитывает на реализацию обстановки игры \hat{x}_{-i}^B ("B" обозначает beliefs; иногда используются термины «предположение», «догадка» – conjecture) и состояния природы $\hat{\theta}_i$, тогда он выберет

$$x_i^B \in BR_i(\hat{\theta}_i, \hat{x}_{-i}^B), i \in N.$$

Вектор x^B является *точечным субъективным равновесием*.

Отметим, что при таком определении «равновесия» не требуется обоснованности предположений агентов о действиях оппонентов, то есть, может оказаться, что $\exists i \in N: \hat{x}_{-i}^B \neq x_{-i}^B$. Обоснованное субъективное равновесие, то есть такое, что $\hat{x}_{-i}^B = x_{-i}^B$, $i \in N$, является равновесием Нэша (для этого, в частности, достаточно, чтобы все параметры игры были общим знанием, и чтобы каждый агент при построении \hat{x}_{-i}^B моделировал рациональное поведение оппонентов). В частном случае, если наилучший ответ каждого агента не зависит от предположений об обстановке, то субъективное равновесие является равновесием в доминантных стратегиях.

В более общем случае i -ый агент может рассчитывать на выбор оппонентами действий из множества $X_{-i}^B \subseteq X_i$ и реализацию

состояния природы из множества $\hat{\Omega}_i \subseteq \Omega$, $i \in N$. Тогда наилучшим ответом будет *гарантирующее субъективное равновесие*:

$$x_i(X_{-i}^B, \hat{\Omega}_i) \in \operatorname{Arg} \max_{x_i \in X_i} \min_{x_{-i} \in X_{-i}^B} \min_{\theta \in \Omega_i} f_i(\theta, x_i, x_{-i}), i \in N.$$

Если $X_{-i}^B = X_{-i}$, $\hat{\Omega}_i = \Omega$, $i \in N$, то $x_i(X_{-i}^B) = x_i^*$, $i \in N$, то есть гарантирующее субъективное равновесие является «классическим» гарантирующим равновесием. Разновидностью гарантирующего субъективного равновесия является Π -равновесие, подробно описанное в [1].

В еще более общем случае в качестве наилучшего ответа i -го агента можно рассматривать распределение вероятностей $p_i(x_i)$, где $p_i(\cdot) \in \Delta(X_i)$ – множеству всевозможных распределений на X_i , которое максимизирует ожидаемый выигрыш агента с учетом его представлений о распределении вероятностей $\mu_i(x_{-i}) \in \Delta(X_{-i})$ действий, выбираемых другими агентами, и распределении вероятностей $q_i(\theta) \in \Delta(\Omega)$ состояния природы (получим *Байесов принцип принятия решений*) [18]:

$$p_i(\mu_i(\cdot), q_i(\cdot), \cdot) \in \\ = \operatorname{Arg} \max_{p_i \in \Delta(X_i)} \int_{X', \Omega} f_i(\theta, x_i, x_{-i}) p_i(x_i) q_i(\theta) \mu_i(x_{-i}) d\theta dx, i \in N.$$

Таким образом, для реализации субъективного равновесия требуется минимальная информированность агентов – каждый из них должен знать свою целевую функцию $f_i(\cdot)$ и допустимые множества Ω и X' . Однако при такой информированности совокупность предположений агентов о состоянии природы и о поведении оппонентов могут быть *несогласованными*. Для достижения согласованности, то есть для того, чтобы предположения оправдывались, необходимы дополнительные предположения о взаимной информированности агентов. Наиболее сильным является предположение об общем знании, которое превращает субъективное точечное равновесие в равновесие Нэша, а совокупность Баийесовых принципов принятия решений – в равновесие Баийеса–Нэша.

Равновесие Баийеса–Нэша. Если в игре имеется неполная информация (см. [17, 18]), то Баийесова игра описывается следующим набором:

- множеством N агентов;

- множеством K' возможных типов агентов, где тип i -го агента $k_i \in K_i$, $i \in N$, вектор типов $k = (k_1, k_2, \dots, k_n) \in K' = \prod_{i \in N} K_i$;

- множеством $X' = \prod_{i \in N} X_i$ допустимых векторов действий агентов;

- набором функций полезности $u_i: K' \times X' \rightarrow \mathcal{R}^1$;

- представлениями $\mu_i(\cdot | k_i) \in \Delta(K_i)$, $i \in N$, агентов.

Равновесие Байеса-Нэша в игре с неполной информацией определяется как набор стратегий агентов вида $\sigma_i: K_i \rightarrow X_i$, $i \in N$, которые максимизируют соответствующие ожидаемые полезности

$$U_i(k_i, \sigma_i(\cdot), \sigma_{-i}(\cdot)) = \int_{k_{-i} \in \prod_{j \neq i} K_j} u_i(k, \sigma_i(k_i), \sigma_{-i}(k_{-i})) \mu_i(k_{-i} | k_i) dk_{-i}, \quad i \in N.$$

В Байесовых играх, как правило, предполагается, что представления $\{\mu_i(\cdot | \cdot)\}_{i \in N}$ являются общим знанием. Для этого, в частности, достаточно, чтобы они были *согласованы*, то есть выводились каждым из агентов по формуле Байеса из распределения $\mu(k) \in \Delta(K')$, которое является общим знанием.

Выше рассмотрены некоторые концепции решения некооперативных игр. Приведем основные понятия кооперативных игр, моделирующих взаимодействие агентов, которые имеют возможность образовывать коалиции, и в рамках этих коалиций договариваться о выбираемых действиях, перераспределять полезность и т.д. (отметим, что в настоящей работе рассматриваются, в основном, некооперативные модели – результаты исследования кооперативного взаимодействия участников организационных систем описаны в [5, 6, 16]).

П.1.2. Кооперативные игры

Кооперативная игра задается множеством игроков $N = \{1, \dots, n\}$ и характеристической функцией $v: 2^N \rightarrow \mathcal{R}$, ставящей в соответствие каждой коалиции игроков $S \subseteq N$ ее выигрыш.

Дележом игры (N, v) называется вектор $x = (x_1, \dots, x_n)$, для которого $\sum_{i \in N} x_i = v(N)$ (свойство эффективности), $x_i \geq v(\{i\})$,

$i \in N$ (свойство индивидуальной рациональности).

Решением кооперативной игры обычно считается множество дележей, которые реализуемы при рациональном поведении игроков. Различные концепции решения кооперативных игр отличаются предположениями о рациональном поведении игроков.

Говорят, что дележ x доминирует дележ y по коалиции S ($x \succ_S y$), если $\forall i \in S \quad x_i > y_i, \quad \sum_{i \in S} x_i \leq v(S)$. Если существует

такая коалиция S , что $x \succ_S y$, говорят, что дележ x доминирует дележ y . Множество недоминируемых дележей игры называется ее C -ядром.

Для заданного множества игроков N сбалансированным покрытием называется такое отображение δ_S множества собственных коалиций $2^N \setminus \{N\}$ в отрезок $[0, 1]$, что $\sum_{S: i \in S} \delta_S = 1$ для всех игро-

ков $i \in N$ (суммирование ведется по собственным коалициям, содержащим игрока i).

Необходимые и достаточные условия непустоты C -ядра даются теоремой О.Н. Бондаревой: C -ядро игры (N, v) не пусто тогда и только тогда, когда для любого сбалансированного покрытия δ_S

$$\sum_{S \subseteq N} \delta_S v(S) \leq v(N).$$

Игры с непустым C -ядром называются сбалансированными.

Кооперативная игра называется несущественной, если для произвольной коалиции $S \subseteq N \quad v(S) = \sum_{i \in S} v(\{i\})$, в противном

случае игра называется существенной. Несущественность игры означает нулевой эффект от кооперации игроков.

Игровая ситуация является сильным равновесием Нэша, если никакая коалиция не может выиграть, отклоняясь от равновесной ситуации. Множество сильных равновесий Нэша может оказаться пустым, однако если в некоторой игре с трансферабельной полезностью игроков имеется единственное сильное равновесие Нэша, то соответствующая кооперативная игра будет несущественной.

Концепция *решений в угрозах и контругрозах* основана на следующей идее. Пусть, например, в процессе игры трех лиц образовалась коалиционная структура $\{\{1, 2\}, \{3\}\}$, содержащая коалицию $T = \{1, 2\}$, в которую входят игроки с номерами 1 и 2. При распределении дохода коалиции $v(\{1, 2\})$ игроки 1 и 2 получают суммы x_1 и x_2 соответственно. Тогда, если игрок 1 недоволен таким распределением, он может сказать своему партнеру, что, если его доля дохода не будет увеличена, то он сформирует коалицию $S = \{1, 3\}$, где сможет рассчитывать на больший выигрыш. Если такая коалиция S может образоваться, то есть если игроку 3 выгодно сменить конфигурацию x на новую конфигурацию y , то такое заявление называется угрозой игрока 1 игроку 2. В свою очередь, игрок 2 может заявить игроку 1, что в случае подобных его действий он может предложить игроку 3 такую конфигурацию z коалиционной структуры $\{\{1\}, \{2, 3\}\}$, что игрок 3 получит больший доход, чем в конфигурации y , а сам игрок 2 получит не меньше, чем в исходной конфигурации x . Таким образом, игрок 2 выдвигает контругрозу, «защищающую» его долю x_2 .

Тогда распределение выигрыша коалиций некоторой коалиционной структуры между своими участниками является равновесием в угрозах и контругрозах, если на каждую угрозу произвольной коалиции K против любой другой коалиции L найдется контругроза коалиции L против коалиции K .

П.1.3. Иерархические игры

Если в рассматриваемых до сих пор моделях игровой неопределенности предполагалось, что игроки (агенты) выбирают свои стратегии одновременно и однократно (модели повторяющихся и дифференциальных игр в настоящей работе не рассматриваются – см. [5, 9, 13]), то в *иерархических играх* [2, 3, 4, 7, 8] существует фиксированный порядок ходов – первый ход делает центр, затем свои стратегии выбирают агенты. С этой точки зрения иерархические игры являются наиболее адекватным аппаратом описания задач управления организационными системами.

Для иерархических игр характерно использование максимального гарантированного результата (МГР) в качестве базовой кон-

цепции решения игры. При этом «пессимистичность» МГР (взятие минимума по множеству неопределенных параметров) компенсируется возможностью передачи информации между игроками, что, очевидно, снижает неопределенность при принятии решения.

Критерии эффективности (целевые функции) первого и второго игроков обозначим $w_1 = f_1(x_1, x_2)$ и $w_2 = f_2(x_1, x_2)$ соответственно. Выигрыши игроков зависят от их действий x_1 и x_2 из множеств действий X_1^0, X_2^0 .

Во всех моделях иерархических игр считается, что *первый игрок (центр)* имеет право первого хода. Его ход состоит в выборе *стратегии* \tilde{x}_1 . Понятие стратегии существенно отличается от понятия действия и тесно связано с информированностью первого игрока о поведении *второго игрока – агента*. Под стратегией игрока здесь и далее понимается правило его поведения, то есть правило выбора конкретного действия в зависимости от содержания и конкретного значения той информации, которую он получит в процессе игры. Выбирать же собственно действие центр может и после выбора действия агентом.

Самая простая стратегия центра состоит в выборе непосредственно действия x_1 (если поступления дополнительной информации о действии агента в процессе игры не ожидается), более сложная – в выборе функции $\tilde{x}_1(x_2)$ (если в процессе игры ожидается информация о действии агента). Также стратегия центра может состоять в сообщении агенту некоторой информации, например, информации о планах своего поведения в зависимости от выбора агентом действия. При этом агент должен быть уверен, что первый игрок может реализовать эту стратегию, то есть что первый игрок будет точно знать реализацию действия x_2 на момент выбора своего действия x_1 .

Например, если агент (выбирающий стратегию вторым) не ожидает информации о действии центра, то реализация права первого хода центра может состоять в сообщении центром агенту функции $\tilde{x}_1(x_2)$. Такое сообщение может рассматриваться, как обещание выбрать действие $x_1 = \tilde{x}_1(x_2)$ при выборе агентом действия x_2 . Тогда стратегия агента состоит в выборе действия в зависимости от сообщения центра, $x_2 = \tilde{x}_2(\tilde{x}_1(\cdot))$. Если при этом агент

доверяет сообщению центра, он должен выбрать действие x_2^* , реализующее

$$\max_{x_2 \in X_2^0} f_2(\tilde{x}_1(x_1), x_2).$$

Игра с описанным выше порядком функционирования называется для краткости игрой Γ_2 (примером такой игры служит, как раз, задача стимулирования в условиях информированности центра о действии агента – см. [11]) [2].

Если центр не ожидает информации о действии агента, и это известно агенту, то стратегия центра состоит, как уже было сказано, просто из выбора некоторого действия x_1^* . Стратегия агента состоит в выборе $x_2 = \tilde{x}_2(x_1^*)$ (он делает ход вторым, уже зная действие центра). Такая игра называется игрой Γ_1 (это, например, та же задача стимулирования, но уже в условиях отсутствия у центра информации о действии агента) [2].

Рассмотрим сначала игру Γ_1 . Пара действий (x_1^*, x_2^*) в игре Γ_1 называется *равновесием Штакельберга*, если

$$(1) \quad x_1^* \in \operatorname{Arg} \max_{x_1 \in X_1^0, x_2 \in R_2(x_1)} f_1(x_1, x_2),$$

$$(2) \quad x_2^* \in R_2(x_1^*) = \operatorname{Arg} \max_{x_2 \in X_2^0} f_2(x_1^*, x_2),$$

то есть $R_2(x_1^*)$ – функция наилучшего ответа агента на действие центра.

Равновесие в игре Γ_1 отличается от равновесия Штакельберга (1) тем, что при определении оптимальной стратегии первого игрока вычисляется минимум по множеству $R_2(x_1)$:

$$x_1^* \in \operatorname{Arg} \max_{x_1 \in X_1^0} \min_{x_2 \in R_2(x_1)} f_1(x_1, x_2).$$

В игре Γ_1 агент выбирает действие в условиях полной информированности, уже зная действие центра. Максимизация выигрыша выбором своего действия является здесь частным случаем применения принципа МГР. Равновесное по Штакельбергу действие центра также дает ему гарантированный результат, если центр уверен в том, что агент выбирает свое действие в соответствии с (2) и принципом благожелательности. Таким образом, равновесные

стратегии как центра, так и агента, являются для них и гарантирующими.

Однако ситуация, когда первый ход дает преимущество, все же более типична. Тогда, если порядок ходов определяется самими игроками, между ними возникает борьба за лидерство. Игре двух лиц в нормальной форме можно поставить в соответствие две игры Γ_1 (игры первого порядка), отличающиеся последовательностью ходов. Тогда борьба за лидерство (первый ход) определяется выгодностью перехода от исходной игры к какой-либо из иерархических игр первого порядка. Известно [6], что, если в игре двух лиц имеются хотя бы два различных оптимальных по Парето равновесия Нэша, то в этой игре имеет место борьба за первый ход.

Тем не менее, во многих случаях соответствующее игре Γ_1 поведение центра нельзя назвать эффективным (см. раздел 2.1 – если в задаче стимулирования центр будет первым выбирать действие (стимулирование агента, уровень зарплаты), а затем уже агент будет выбирать свое действие при заданном стимулировании, единственное равновесие Штакельберга будет состоять в том, что центр ничего не будет платить агенту, а агент, соответственно, не будет работать). Поэтому, когда центр наблюдает действие агента, он заинтересован сообщить агенту о своих планах по выбору действия в зависимости от действия агента, реализуя тем самым игру Γ_2 .

Далее приводится формулировка теоремы о максимальном гарантированном результате центра в игре типа Γ_2 . К этой игре сводятся многие модели управления, например, задача стимулирования в условиях полной информированности (см. вторую и третью главы). Определим необходимые для формулировки теоремы понятия.

Целевые функции игроков: $w_1 = f_1(x_1, x_2)$, $w_2 = f_2(x_1, x_2)$ непрерывны на компактных множествах $x_1 \in X_1^0$ $x_2 \in X_2^0$ допустимых действий.

Стратегия центра – $\tilde{x}_1 = \tilde{x}_1(x_2)$, то есть предполагается следующий порядок функционирования: игрок 1, обладая правом первого хода, сообщает игроку 2 план выбора своей стратегии в зависимости от выбранной игроком 2 стратегии x_2 . После этого второй игрок выбирает действие x_2 , максимизируя свою целевую

функцию с подставленной туда стратегией первого игрока, а затем первый игрок – действие $\tilde{x}_1(x_2)$.

Стратегия наказания $x_1'' = x_1''(x_2)$ определяется из условия

$$f_2(x_1''(x_2), x_2) = \min_{x_1 \in X_1^0} f_2(x_1, x_2).$$

Если стратегий наказания несколько, то будем называть *оптимальной стратегией наказания* ту из них, на которой достигается максимум выигрыша первого игрока.

Гарантированный результат второго игрока (при использовании первым игроком стратегии наказания) равен

$$L_2 = \max_{x_2 \in X_2^0} f_2(x_1''(x_2), x_2) = \max_{x_2 \in X_2^0} \min_{x_1 \in X_1^0} f_2(x_1, x_2).$$

Множество действий второго игрока, обеспечивающих ему максимальный выигрыш при использовании первым игроком стратегии наказания: $E_2 = \{x_2 \mid f_2(x_1''(x_2), x_2) = L_2\}$.

Множество достижимости $D = \{(x_1, x_2) : f_2(x_1, x_2) > L_2\}$ – это договорное множество рассматриваемой игры, то есть множество сочетаний стратегий первого и второго игроков, которые гарантировали бы второму результат, строго больший того, что тот может получить даже при наихудших для него действиях первого игрока (то есть при использовании первым игроком стратегии наказания).

Наилучший результат первого игрока на множестве достижимости $K = \begin{cases} \sup_{(x_1, x_2) \in D} f_1(x_1, x_2), & D \neq \emptyset \\ -\infty, & D = \emptyset \end{cases}$. Принадлежность ситуации множеству достижимости гарантирует реализуемость этого результата путем использования стратегии наказания.

Действие первого игрока, реализующее $K - \varepsilon$ при выборе вторым игроком рекомендуемого действия из D :

$$f_1(x_1^\varepsilon, x_2^\varepsilon) \geq K - \varepsilon, \quad (x_1^\varepsilon, x_2^\varepsilon) \in D \neq \emptyset.$$

$M = \inf_{x_2 \in E_2} \sup_{x_1 \in X_1^0} f_1(x_1, x_2)$ – гарантированный результат центра

при применении им стратегии наказания (так как стратегии второго игрока ограничены множеством E_2).

Стратегия $x_1^{\varepsilon e}(x_2)$ реализует (с точностью ε) наилучший ответ центра на действие x_2 агента (ε -доминантная стратегия), то есть

$$f_1(x_1^{ae}(x_2)) \geq \sup_{x_1 \in X_1^0} f_1(x_1, x_2) - \varepsilon.$$

Теорема Ю.Б. Гермейера [2]. В указанных условиях наибольший гарантированный результат центра равен $\max [K, M]$. При $K > M$ ε -оптимальная стратегия центра

$$\tilde{x}_1^\varepsilon(x_2) = \begin{cases} x_1^\varepsilon, & \text{при } x_2 = x_2^\varepsilon \\ x_1^*(x_2), & \text{при } x_2 \neq x_2^\varepsilon \end{cases}.$$

При $K \leq M$ оптимальная стратегия центра заключается в применении оптимальной стратегии наказания.

Каким же образом соотносятся выигрыши центра в играх Γ_1 и Γ_2 с одинаковыми функциями выигрыша? Существуют ли более рациональные для центра методы обмена информацией, дающие ему больший выигрыш? Ответ на эти вопросы дает рассмотрение информационных расширений игры, или *метаигр*.

Если центр не планирует самостоятельно получить информацию о действии агента, он может первым выбрать действие, реализуя игру Γ_1 . Однако ему можно порекомендовать и более сложное поведение. Центр может попросить агента сообщить ему свою стратегию $x_2 = \tilde{x}_2(x_1)$, которая основана на ожидаемой агентом информации о действии центра. Реализация права первого хода центром состоит в этом случае в сообщении агенту стратегии $\tilde{x}_1(\tilde{x}_2(x_1))$. Эту стратегию можно интерпретировать, как обещание центра выбрать действие $\tilde{x}_1(\tilde{x}_2(x_1))$ при условии, что агент обещает выбирать свое действие в соответствии с $\tilde{x}_2(x_1)$. Так образуется игра Γ_3 .

Если центр определяет порядок обмена информацией, он может выбирать, играть ему Γ_1 или Γ_3 . В обеих играх центр вынужден выбирать действие, не зная действия, выбранного агентом. Можно считать Γ_3 , в некотором роде, усложнением игры Γ_1 .

Аналогично тому, как, с помощью образования дополнительной «петли обратной связи», из Γ_1 была образована Γ_3 , можно усложнить и игру Γ_2 . Так образуется игра Γ_4 . В ней агент, ожидая от центра, как и в Γ_2 , информацию вида $\tilde{x}_1(x_2)$, формирует и сообщает центру свою стратегию $\tilde{x}_2(\tilde{x}_1)$. Центр, обладающий правом

первого хода, пользуется стратегиями $\tilde{\tilde{x}}_1(\tilde{\tilde{x}}_2)$, которые определяют, какую функцию $\tilde{x}_1(x_2)$ выберет центр в зависимости от сообщения агента $\tilde{\tilde{x}}_2$.

Таким же способом можно на основе Γ_3 построить игру Γ_5 , и так далее. В каждой из построенных четных игр Γ_{2m} , $m = 1, 2, \dots$, центр использует в качестве стратегий отображения множества стратегий агента в этой игре на множество стратегий центра в игре Γ_{2m-2} . Аналогично, стратегиями агента являются отображения множества стратегий центра в Γ_{2m} на множество стратегий агента в игре Γ_{2m-2} .

Такую рефлексию можно было бы наращивать бесконечно, переходя к все более сложным схемам обмена информацией, если бы рассмотрение этих игр увеличивало выигрыш центра (в интересах которого и проводится исследование всех метаигр). Однако имеет место следующий результат.

Теорема Н.С. Кукушкина [2]. Максимальный гарантированный результат центра в игре Γ_{2m} при $m > 1$ равен максимальному гарантированному результату центра в игре Γ_2 . В играх же Γ_{2m+1} при $m > 1$ максимальный гарантированный результат центра равен его максимальному гарантированному результату в игре Γ_3 .

Таким образом, при исследовании гарантированного результата центра можно ограничиться исследованием только игр Γ_1 , Γ_2 и Γ_3 . Кроме того, известно [2], что максимальный гарантированный результат центра в игре Γ_2 не меньше его гарантированного результата в игре Γ_3 , а тот, в свою очередь, не меньше гарантированного выигрыша в игре Γ_1 . Этот результат показывает, что Γ_2 является «идеальной» игрой для центра. Соответственно, если центр имеет возможность определять порядок и содержание обмена информацией, и, кроме того, при выборе своего действия знает действие, выбранное агентом, он должен играть Γ_2 . Если центр на момент выбора своего действия не знает действия агента – ему наиболее выгодна игра Γ_3 .

Игры и структуры. Выше мы рассмотрели основные понятия теории игр, перейдя от игр, в которых агенты выбирают свои действия одновременно (игра Γ_0 в нормальной форме или в форме характеристической функции) к иерархическим играм, в которых последовательность ходов фиксирована – первым делает ход

центр, а затем – агент. Можно усложнять модель и дальше, переходя к всё более сложным играм. Опишем общую картину (см. рисунок 1), которая позволяет увидеть логику перехода от более простых к более сложным задачам, чтобы более сложная задача могла быть декомпозирована на более простые.

Если имеется один субъект, принимающий решения (рисунок 1а), то он описывается с точки зрения гипотезы рационального поведения как стремящийся максимизировать свою целевую функцию. Далее можно усложнить модель и рассмотреть несколько субъектов на одном уровне (рисунок 1б), описав их взаимодействие игрой Γ_0 в нормальной форме. Если ввести иерархию, то для двух субъектов (рисунок 1в) их взаимодействие описывается игрой Γ_i , где $i = 1, 2$ или 3 .

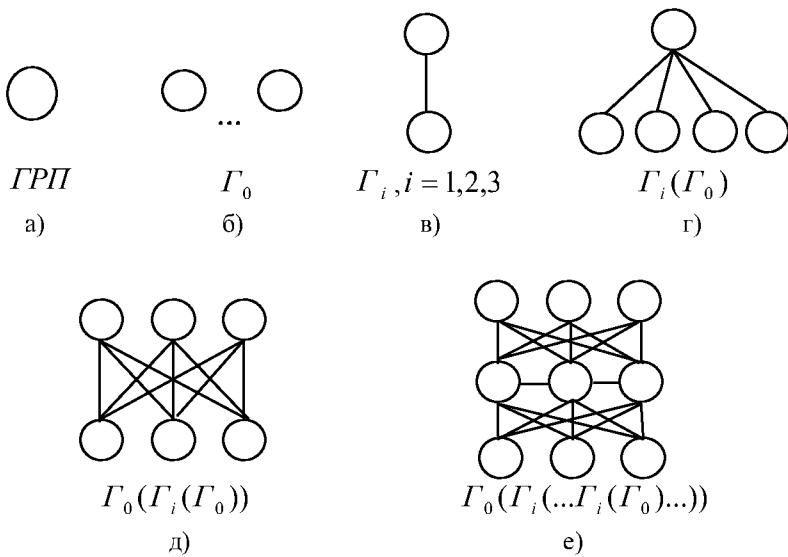


Рис. 1. Игры и структуры

Представим себе, что имеется структура «один начальник – несколько подчиненных» (рисунок 1г). Взаимодействие агентов,

находящихся на одном уровне, можно описывать игрой Γ_0 . Взаимодействие «начальник-подчиненный» описывается игрой Γ_i . Тогда условно такую структуру можно представить игрой Γ_i , определенной на игре Γ_0 , условно обозначив ее $\Gamma_i(\Gamma_0)$.

Далее, пусть есть несколько начальников (центров) и несколько подчиненных – агентов (рисунок 1д). На нижнем уровне агенты играют игру Γ_0 . Над ними центры играют иерархическую игру Γ_i , но центры, в свою очередь, разыгрывают на своем уровне игру Γ_0 . Итого, получили игру $\Gamma_i(\Gamma_i(\Gamma_0))$.

Можно взять более сложную структуру с более сложным взаимодействием (например, рисунок 1е). Это будет иерархическая игра между уровнями, и "обычная" игра на каждом из уровней: $\Gamma_0(\Gamma_i(\dots\Gamma_i(\Gamma_0)\dots))$.

Основная идея заключается в том, чтобы декомпозировать сложную структуру (игру) на набор более простых и воспользоваться результатами исследования последних. Оказывается, что между играми и структурами существует глубокая связь – момент принятия субъектом решений определяет его "место" в организационной иерархии (см. подробности в [10]).

Некооперативные игры

Рассмотрим игру, в которой участвуют агенты из множества $N = \{1, 2, \dots, n\}$. Если в ситуации присутствует неопределенный параметр $\theta \in \Omega$, то *структура информированности* I_i (как синоним будем употреблять термины *информационная структура* и *иерархия представлений*) i -го агента включает в себя следующие элементы. Во-первых, представление i -го агента о параметре θ – обозначим его θ_i , $\theta_i \in \Omega$. Во-вторых, представления i -го агента о представлениях других агентов о параметре θ – обозначим их θ_{ij} , $\theta_{ij} \in \Omega$, $j \in N$. В третьих, представления i -го агента о представлении j -го агента о представлении k -го агента – обозначим их θ_{ijk} , $\theta_{ijk} \in \Omega$, $j, k \in N$. И так далее.

Таким образом, структура информированности I_i i -го агента задается набором всевозможных значений вида $\theta_{ij_1 \dots j_l}$, где l пробегает множество целых неотрицательных чисел, $j_1, \dots, j_l \in N$, а $\theta_{ij_1 \dots j_l} \in \Omega$.

Аналогично задается структура информированности I игры в целом – набором значений $\theta_{i_1 \dots i_l}$, где l пробегает множество целых неотрицательных чисел, $i_1, \dots, i_l \in N$, а $\theta_{ij_1 \dots j_l} \in \Omega$. Подчеркнем, что структура информированности I «недоступна» наблюдению агентов, каждому из которых известна лишь некоторая ее часть (а именно – I_i). Таким образом, структура информированности – бесконечное n -дерево (то есть тип структуры постоянен и является n -деревом), вершинам которого соответствует конкретная информированность реальных и фантомных агентов.

Рефлексивной игрой Γ_i называется игра, описываемая следующим кортежем [15]:

$$\Gamma_i = \{N, (X_i)_{i \in N}, f_i(\cdot)_{i \in N}, \Omega, I\},$$

где N – множество реальных агентов, X_i – множество допустимых действий i -го агента, $f_i(\cdot): \Omega \times X' \rightarrow \mathbb{R}^1$ – его целевая функция, $i \in N$, Ω – множество возможных значений неопределенного параметра, I – структура информированности. Подчеркнем, что все элементы рефлексивной игры кроме структуры информированности являются общим знанием среди агентов, т.е.

- 1) эти элементы известны всем агентам;
- 2) всем агентам известно 1);
- 3) всем агентом известно 2) и т. д. до бесконечности.

Далее для формулировки некоторых определений и свойств нам понадобятся следующие обозначения:

Σ_+ – множество всевозможных конечных последовательностей индексов из N ;

Σ – объединение Σ_+ с пустой последовательностью;

$|\sigma|$ – количество индексов в последовательности σ (для пустой последовательности принимается равным нулю), которое выше было названо длиной последовательности индексов.

Если θ_i – представления i -го агента о неопределенном параметре, а θ_{ii} – представления i -го агента о собственном представлении, то естественно считать, что $\theta_{ii} = \theta_i$. Иными словами, i -й агент

правильно информирован о собственных представлениях, а также считает, что таковы и другие агенты и т. д. Формально это означает, что выполнена *аксиома автоинформированности*, которую далее будем предполагать выполненной:

$$\forall i \in N \forall \tau, \sigma \in \Sigma \theta_{\tau i \sigma} = \theta_{\tau i \sigma}.$$

Эта аксиома означает, в частности, что, зная θ_τ для всех $\tau \in \Sigma_+$, таких что $|\tau| = \gamma$, можно однозначно найти θ_τ для всех $\tau \in \Sigma_+$, таких что $|\tau| < \gamma$.

Наряду со структурами информированности I_i , $i \in N$, можно рассматривать структуры информированности I_{ij} (структуре информированности j -го агента в представлении i -го агента), I_{ijk} и т.д. Отождествляя структуру информированности с характеризуемым ею агентом, можно сказать, что, наряду с *n* реальными агентами (*i*-агентами, где $i \in N$) со структурами информированности I_i , в игре участвуют *фантомные агенты* (τ -агенты, где $\tau \in \Sigma_+$, $|\tau| \geq 2$) со структурами информированности $I_\tau = \{\theta_{\tau i \sigma}\}$, $\sigma \in \Sigma$, существующие в сознании реальных агентов.

Определим фундаментальное для дальнейших рассмотрений понятие тождественности структур информированности. Структуры информированности I_λ и I_μ ($\lambda, \mu \in \Sigma_+$) называются *тождественными*, если выполнены два условия:

1. $\theta_{\lambda \sigma} = \theta_{\mu \sigma}$ для любого $\sigma \in \Sigma$;
2. последние индексы в последовательностях λ и μ совпадают.

Будем обозначать тождественность структур информированности следующим образом: $I_\lambda = I_\mu$.

Понятие тождественности структур информированности позволяет определить их важное свойство – сложность. Заметим, что наряду со структурой I имеется счетное множество структур I_τ , $\tau \in \Sigma_+$, среди которых можно при помощи отношения тождественности выделить классы попарно нетождественных структур. Количество этих классов естественно считать *сложностью структуры информированности*.

Будем говорить, что структура информированности I имеет *конечную сложность* $v = v(I)$, если существует такой конечный набор попарно нетождественных структур $\{I_{\tau_1}, I_{\tau_2}, \dots, I_{\tau_v}\}$, $\tau_i \in \Sigma_+$, $i \in \{1, \dots, v\}$, что для любой структуры I_σ , $\sigma \in \Sigma_+$, най-

дется тождественная ей структура I_{τ_i} из этого набора. Если такого конечного набора не существует, будем говорить, что структура I имеет бесконечную сложность: $\nu(I) = \infty$.

Структуру информированности, имеющую конечную сложность, будем называть *конечной* (еще раз отметим, что при этом дерево структуры информированности все равно остается бесконечным). В противном случае структуру информированности будем называть *бесконечной*.

Ясно, что минимально возможная сложность структуры информированности в точности равна числу участвующих в игре реальных агентов (напомним, что по определению тождественности структур информированности они попарно различаются у реальных агентов).

Любой набор (конечный или счетный) попарно нетождественных структур I_τ , $\tau \in \Sigma_+$, такой, что любая структура I_σ , $\sigma \in \Sigma_+$, тождественна одной из них, назовем *базисом* структуры информированности I .

Если структура информированности I имеет конечную сложность, то можно определить максимальную длину последовательности индексов γ такую, что, зная все структуры I_τ , $\tau \in \Sigma_+$, $|\tau| = \gamma$, можно найти и все остальные структуры. Эта длина в определенном смысле характеризует ранг рефлексии, необходимый для описания структуры информированности.

Будем говорить, что структура информированности I , $\nu(I) < \infty$, имеет *конечную глубину* $\gamma = \gamma(I)$, если

1. для любой структуры I_σ , $\sigma \in \Sigma_+$, найдется тождественная ей структура I_τ , $\tau \in \Sigma_+$, $|\tau| \leq \gamma$;

2. для любого целого положительного числа ξ , $\xi < \gamma$, существует структура I_σ , $\sigma \in \Sigma_+$, не тождественная никакой из структур I_τ , $\tau \in \Sigma_+$, $|\tau| = \xi$.

Если $\nu(I) = \infty$, то и глубину будем считать бесконечной: $\gamma(I) = \infty$.

Понятия сложности и глубины структуры информированности игры можно рассматривать τ -субъективно. В частности, глубина структуры информированности игры с точки зрения τ -агента, $\tau \in \Sigma_+$, называется *рангом рефлексии* τ -агента.

Если задана структура I информированности игры, то тем самым задана и структура информированности каждого из агентов (как реальных, так и фантомных). Выбор τ -агентом своего действия x_τ в рамках гипотезы рационального поведения определяется его структурой информированности I_τ , поэтому, имея перед собой эту структуру, можно смоделировать его рассуждения и определить это его действие. Выбирая свое действие, агент моделирует действия других агентов (осуществляет рефлексию). Поэтому при определении исхода игры необходимо учитывать действия как реальных, так и фантомных агентов.

Набор действий $x_\tau^*, \tau \in \Sigma_+$, назовем *информационным равновесием* [15], если выполнены следующие условия:

1. структура информированности I имеет конечную сложность v ;
2. $\forall \lambda, \mu \in \Sigma \quad I_{\lambda i} = I_{\mu i} \Rightarrow x_{\lambda i}^* = x_{\mu i}^*$;
3. $\forall i \in N, \forall \sigma \in \Sigma$

$$x_{\sigma i}^* \in \operatorname{Arg} \max_{x_i \in X_i} f_i(\theta_{\sigma i}, x_{\sigma i 1}^*, \dots, x_{\sigma i, i-1}^*, x_i, x_{\sigma i, i+1}^*, \dots, x_{\sigma i, n}^*).$$

Первое условие в определении информационного равновесия означает, что в рефлексивной игре участвует конечное число реальных и фантомных агентов.

Второе условие отражает требование того, что одинаково информированные агенты выбирают одинаковые действия.

И, наконец, третье условие отражает рациональное поведение агентов – каждый из них стремится выбором собственного действия максимизировать свою целевую функцию, подставляя в нее действия других агентов, которые оказываются рациональными с точки зрения рассматриваемого агента в рамках имеющихся у него представлений о других агентах.

Удобным инструментом исследования информационного равновесия является *граф рефлексивной игры*, в котором вершины соответствуют реальным и фантомным агентам, и в каждую вершину-агента входят дуги (их число на единицу меньше числа реальных агентов), идущие из вершин-агентов, от действий которых в субъективном равновесии зависит выигрыш данного агента.

Одной из особенностей «классического» равновесия Нэша является его самоподдерживающийся характер – если игра повторяется несколько раз, и все игроки кроме i -го выбирают одни и те же

равновесные действия, то и i -му нет резона отклоняться от своего равновесного действия. Это обстоятельство очевидным образом связано с тем, что представления всех игроков о реальности адекватны – значение состояния природы является общим знанием.

В случае информационного равновесия ситуация, вообще говоря, может быть иной. Действительно, в результате однократного разыгрывания игры может оказаться, что какие-то из игроков (или даже все) наблюдают не тот результат, на который они рассчитывали. Это может быть связано как с неверным представлением о состоянии природы, так и с неадекватной информированностью о представлениях оппонентов. В любом случае, самоподдерживающийся характер равновесия нарушается – если игра повторяется, то действия игроков могут измениться.

Однако в некоторых случаях самоподдерживающийся характер равновесия может иметь место и при различных (и, вообще говоря, неверных) представлениях агентов. Говоря неформально, это происходит тогда, когда каждый агент (как реальный, так и фантомный) наблюдает тот результат игры, которого ожидает. Для формального изложения дополним кортеж, задающий рефлексивную игру, набором функций $w_i(\cdot): \Omega \times X' \rightarrow W_i$, $i \in N$, каждая из которых отображает вектор (θ, x) в элемент w_i некоторого множества W_i . Этот элемент w_i есть то, что i -й агент наблюдает в результате разыгрывания игры.

Функцию $w_i(\cdot)$ будем называть *функцией наблюдения* i -го агента [14]. Будем считать, что функции наблюдения являются общим знанием среди агентов.

Если $w_i(\theta, x) = (\theta, x)$, т. е. $W_i = \Omega \times X'$, то i -й агент наблюдает как состояние природы, так и действия всех агентов. Если, напротив, множество W_i состоит из одного элемента, то i -й агент ничего не наблюдает.

Пусть в рефлексивной игре существует информационное равновесие x_τ , $\tau \in \Sigma_+$ (напомним, что τ – произвольная непустая конечная последовательность индексов из N). Зафиксируем $i \in N$ и рассмотрим i -го агента. Он ожидает в результате игры пронаблюдать величину

$$w_i(\theta, x_1, \dots, x_{i-1}, x_i, x_{i+1}, \dots, x_m).$$

На самом же деле он наблюдает величину

$$w_i(\theta, x_1, \dots, x_{i-1}, x_i, x_{i+1}, \dots, x_n).$$

Поэтому требование стабильности для i -агента означает совпадение этих величин.

В общем случае, т. е. для ti -агента, $ti \in \Sigma_+$, условие стабильности определим следующим образом. Информационное равновесие x_{ti} , $ti \in \Sigma_+$, будем называть *стабильным* при заданной структуре информированности I , если для любого $ti \in \Sigma_+$ выполняется

$$w_i(\theta_{ti}, x_{ti1}, \dots, x_{ti,i-1}, x_{ti}, x_{ti,i+1}, \dots, x_{tin}) =$$

$$= w_i(\theta_t, x_{t1}, \dots, x_{t,i-1}, x_{ti}, x_{t,i+1}, \dots, x_{tn}).$$

Информационное равновесие, не являющееся стабильным, будем называть *нестабильным*.

Пусть набор действий x_{ti} , $ti \in \Sigma_+$, является стабильным информационным равновесием. Будем называть его *истинным* равновесием, если набор (x_1, \dots, x_n) является равновесием в условиях общего знания о состоянии природы θ (или о наборе (r_1, \dots, r_n) типов агентов). Из приведенного определения, в частности, следует, что в условиях общего знания любое информационное равновесие является истинным.

Стабильное информационное равновесие, не являющееся истинным, назовем *ложным*. Таким образом, ложное равновесие – это такое стабильное информационное равновесие, которое не является равновесием в случае одинаковой информированности агентов (в условиях общего знания).

Многочисленные примеры прикладных моделей информационного управления, использующие аппарат рефлексивных игр, можно найти в [14].

Литература

(работы, отмеченные звездочкой, можно найти в электронной библиотеке на сайте www.mtas.ru)

- 1 Бурков В.Н., Кондратьев В.В. Механизмы функционирования организационных систем. М.: Наука, 1981. – 384 с.
- 2 Гермейер Ю.Б. Игры с непротивоположными интересами. М.: Наука, 1976. – 327 с.
- 3 Горелик В.А., Горелов М.А., Кононенко А.Ф. Анализ конфликтных ситуаций в системах управления. М.: Радио и связь, 1991. – 288 с.

- 4 Горелик В.А., Кононенко А.Ф. Теоретико-игровые модели принятия решений в эколого-экономических системах. М.: Радио и связь, 1982. – 144 с.
- 5 *Губко М.В. Механизмы управления организационными системами с коалиционным взаимодействием участников. М.: ИПУ РАН, 2003. – 118 с.
- 6 Губко М.В., Новиков Д.А. Теория игр в управлении организационными системами. М.: Синтег, 2002. – 148 с.
- 7 Кононенко А.Ф., Халезов А.Д., Чумаков В.В. Принятие решений в условиях неопределенности. М.: ВЦ АН СССР, 1991. – 211 с.
- 8 Кукушкин Н.С., Морозов В.В. Теория неантагонистических игр. М.: МГУ, 1984.
- 9 Нейман Д., Моргенштерн О. Теория игр и экономическое поведение. М.: Наука, 1970.
- 10 *Новиков Д.А. Сетевые структуры и организационные системы. М.: ИПУ РАН, 2003. – 108 с.
- 11 *Новиков Д.А. Стимулирование в организационных системах. М.: Синтег, 2003. – 312 с.
- 12 *Новиков Д.А., Петраков С.Н. Курс теории активных систем. М.: Синтег, 1999. – 108 с.
- 13 *Новиков Д.А., Смирнов И.М., Шохина Т.Е. Механизмы управления динамическими активными системами. М.: ИПУ РАН, 2002. – 124 с.
- 14 *Новиков Д.А., Чхартишвили А.Г. Прикладные модели информационного управления. М.: ИПУ РАН, 2004. – 129 с.
- 15 Новиков Д.А., Чхартишвили А.Г. Рефлексивные игры. М.: Синтег, 2003. – 160 с.
- 16 Оуэн Г. Теория игр. М.: Мир, 1971.
- 17 Mas-Collie A., Whinston M.D., Green J.R. Microeconomic theory. N.Y.: Oxford Univ. Press, 1995. – 981 p.
- 18 Myerson R.B. Game theory: analysis of conflict. London: Harvard Univ. Press, 1991. – 568 p.