Доклады Академии наук СССР
1974. Том 216, № 3

УДК 518.9 *МАТЕМАТИКА*

Ю. Б. ГЕРМЕЙЕР

СЛАБОУСТОЙЧИВЫЕ СОВМЕСТНЫЕ РЕШЕНИЯ В ПОВТОРЯЮЩИХСЯ ИГРАХ

(Представлено академиком А. А. Дородницыным 30 X 1973)

1. Наряду с принципом наилучшего гарантированного результата одним из основных методов принятия решений, рассматриваемых в теории игр ([1,3]), является принцип устойчивости решения, понимаемый как отсутствие выгоды в отклонении от этого решения. Классическим и наиболее используемым примером воплощения этого являются ситуации равновесия Нэша, которые обычно трактуются как способ принятия решений в бескоалиционных играх. Однако в таком качестве ситуации равновесия обладают определенными недостатками:

1) Необходимость достаточно точного знания каждым игроком интересов всех партнеров для нахождения ситуаций равновесия.

2) Обычная неединственность ситуаций равновесия при несовпадении их выгодности для различных игроков также не позволяет выбрать разумное бескоалиционное решение.

3) Возможность существования (во многих играх) решения, обеспечивающего сразу для всех игроков лучшие результаты, чем равновесные. Последнее обстоятельство хорошо видно в моделях иерархических систем ([4]). Еще один характерный пример можно задать с помощью критериев эффективности

$$W_i = x_i + \sum_{j\neq i}^{n} (1-x_j) \quad \text{при} \quad 0\leqslant x_i\leqslant 1; \quad 1\leqslant i\leqslant n; \tag{1}$$

здесь единственной ситуацией равновесия является $x_i=1$; $1\leqslant i\leqslant n$, обеспечивающая всем игрокам результат 1. Однако решение $x_i=0$, $1\leqslant i\leqslant n$, дает всем игрокам по $n-1$.

Первые два затруднения преодолеваются, если считать ситуацию равновесия совместным (коалиционным) решением всех игроков, вариантом неформального компромисса. При этом стремление к ситуациям равновесия в игре (W_i приведены к одной размерности)

$$W_i = f_i(x_1, \ldots, x_n), \quad x_i\in M_i, \quad i=1,\ldots, n, \tag{2}$$

может быть представлено как стремление к максимизации общего для всех игроков критерия ([4])

$$W_p = \min_{1\leqslant i\leqslant n} [f_i(x_1,\ldots, x_i,\ldots, x_n) - \max_{y_i\in M_i} f_i(x_1,\ldots, y_i,\ldots, x_n)]. \tag{3}$$

При таком подходе ситуация равновесия может трактоваться как слабо устойчивое решение, поскольку его нарушение одним из игроков не приносит последнему выгоды.

Заметим, что коалиционные решения в (2), дающие результаты, неулучшаемые одновременно для всех игроков, представимы в виде стремления к максимизации единого критерия

$$W_n = \min_{1\leqslant i\leqslant n} \rho_i [f_i(x_1,\ldots, x_n) - W_i^0], \tag{4}$$

где выбор $\bar\rho=(\rho_1,\ldots,\rho_n)$ и $\overline{W}^0=(W_1^0,\ldots,W_n^0)$ конкретизирует компромисс между игроками. Нетрудно убедиться, что компромиссные решения по (4), как правило, неустойчивы в одном повторении. Естественно поэтому стремиться так изменить понятие слабой устойчивости, чтобы хотя бы некоторые из компромиссных решений (4) оказались устойчивыми.

Такая возможность появляется при рассмотрении повторяющихся игр. При обсуждении ее полезно вспомнить, что для обеспечения существования ситуаций равновесия обычно используют смешанные стратегии, которые имеют реальный смысл только при достаточной повторяемости игры.

2. Будем предполагать, что результаты каждого повторения игры (2) и поведение игроков становятся известными к следующему повторению. Критерием эффективности i-го игрока в данном повторении игры (2) естественно считать сумму его результатов в m_i+1 повторении, начиная с данного. Совместное решение всех игроков в рассматриваемой реализации будем считать состоящим из двух компонент:

а) «нормальное» стационарное решение $\bar x^0=(x_1^0,\ldots,x_n^0)$, являющееся, вообще говоря, результатом неформального компромисса;

б) реализация $\min\limits_{j\neq i,\ x_j\in M_j} f_i(x_1,\ldots,\tilde x_i,\ldots,x_n)$ в m_i последующих повторениях, если в данном повторении i-й игрок нарушил решение $\bar x^0$ и в дальнейшем использовал стратегию $\tilde x_i=x_i(x_1,\ldots,x_{i-1},x_{i+1},\ldots,x_n)$. При этом предполагается, что в $m^0+1=\max\limits_{1\leqslant i\leqslant n} m_i+1$ повторениях нарушить коалиционное решение может лишь один игрок. Последнее предположение необязательно, если заранее известно, что все игроки осторожны и не договариваются между собой о совместном нарушении.

В указанных условиях i-му игроку нет смысла нарушать компромисс (решение x_i^0), если выполнено

$$\Phi_i(\bar x^0)=(m_i+1)f_i(\bar x^0)-\max_{y_i\in M_i} f_i(x_1^0,\ldots,y_i,\ldots,x_n^0)-$$
$$-m_i\max_{\tilde x_i}\min_{\substack{x_j\in M_j\\ j\neq i}} f_i(x_1,\ldots,\tilde x_i,\ldots,x_n)\geqslant 0. \tag{5}$$

Если (5) выполнено для всех i, то коалиционное решение в данном повторении можно называть слабо устойчивым. Стремление к «максимальной» слабой устойчивости на множестве P_0 допустимых $\bar x^0$ может быть выражено в виде стремления к максимизации по $\bar x^0\in P_0$

$$\min_{1\leqslant i\leqslant n}\gamma_i\Phi_i(\bar x^0),\quad \gamma_i>0,\quad 1\leqslant i\leqslant n. \tag{6}$$

При $m_i=0$ (6) превращаются в (3); таким образом, ситуации равновесия (если они существуют в чистых стратегиях) есть слабо устойчивые решения в играх, не рассчитанных на повторение. «Нормальные» решения, слабо устойчивые при данных m_i, очевидно остаются таковыми при больших m_i; однако при этом меняется вторая (угрожающая) часть полного решения и увеличивается уровень устойчивости, т. е. разность (5).

В качестве P_0 может выбираться любое непустое множество совместных решений, например, множество Парето, или множество решений, максимизирующих $W_{\bar\rho}$ (4) хоть при каких-то $\bar\rho>0$. Если m_i достаточно велики для данной игры (2), то слабо устойчивым становится любое решение, удовлетворяющее условиям

$$f_i(\bar x^0)>\max_{\tilde x_i}\min_{\substack{x_j\in M_j\\ j\neq i}} f_i(x_1,\ldots,\tilde x_i,\ldots,x_n)=L_i. \tag{7}$$

Таким образом, введение слабо устойчивых коалиционных решений обобщает ситуации равновесия, снимая все указанные их недостатки.

В примере (1) решение $\bar x=\bar 0$ является слабо устойчивым для $n\geqslant 3$ уже при $m_i=1$, причем разность (5) растет с ростом n. Заметим еще, что при

$n=2$ введенная слабая устойчивость является, конечно, достаточно исчерпывающим вариантом устойчивости.

3. Пусть $K_i(\bar{x}^0)$ — минимальное число наказывающих повторений, при котором еще выполнено (5), и

$$K(\bar{x}^0)=\max_{1\leqslant i\leqslant n} K_i(\bar{x}^0).$$

Если

$$K(\bar{x}^0)<\min_{1\leqslant i\leqslant n} m_i, \tag{8}$$

то наказания нарушителей могут быть сделаны все продолжительностью $K(\bar{x}^0)$ и предусматриваться повторением при возможном повторении нарушений. Такое построение наказаний снова обеспечивает устойчивость от одиночных нарушений, единообразно и не требует точного знания m_i при выполнении (8). Стремление к уменьшению $K(\bar{x}^0)$ увеличивает область возможных m_i, для которых обеспечивается слабая устойчивость, что весьма важно при неточной информации о реальных значениях m_i. Кроме того, минимизация $K(\bar{x}^0)$ может уменьшить продолжительность личных потерь игроков во время наказания. Все это приводит к целесообразности выбора $\bar{x}^0$ из P_0 на основе минимизации $K(\bar{x}^0)$.

Из (5) имеем приближенно

$$K(\bar{x}^0)+1\simeq\max_{1\leqslant i\leqslant n}\frac{\max\limits_{y_i\in M_i} f_i(x_1^0,\ldots,y_i,\ldots,x_n^0)-L_i}{f_i(\bar{x}^0)-L_i}. \tag{9}$$

Минимизация (9) по $\bar{\rho}$ при $\bar{x}^0$, максимизирующих (4), формально эквивалентна игре 2-х лиц с фиксированной последовательностью ходов ([4], [5]); численные методы решения этих задач (максиминов со связанными переменными) в настоящее время успешно разрабатываются ([6], [7]).

Поставленная задача легко решается для игр с неограниченными побочными платежами α_i

$$W_i=f_i(\bar{x})+\alpha_i,\quad \sum_{i=1}^{n}\alpha_i=0. \tag{10}$$

Коалиционное решение состоит здесь в выборе $\bar{x}=(x_1,\ldots,x_n)$ и вектора $\bar{\alpha}=(\alpha_1,\ldots,\alpha_n)$. Поскольку отказываться от получения $\alpha_i^0>0$ при коалиционном решении $(\bar{x}^0,\bar{\alpha}^0)$ невыгодно, то нарушение $\bar{\alpha}^0$ может состоять только в отказе отдать побочный платеж, если $\alpha_i^0<0$. Игрок, нарушающий $\bar{x}^0$, уже не получит $\alpha_i^0>0$, поскольку $\bar{\alpha}^0$ реализуется после выполнения $\bar{x}^0$; но он, конечно, не отдаст и сам побочный платеж, если $\alpha_i^0<0$. Наконец, если он хочет нарушить $\bar{\alpha}^0$, то ему целесообразно нарушить и x^0, максимизируя f_i по x_i. Итак, нарушитель будет максимизировать f_i по x_i и не участвовать в побочных платежах. Он по-прежнему наказывается минимизацией f_i по x_j, $j\neq i$. Максимизация (4), в котором f_i изменено на $f_i+\alpha_i$, в случае $W_i^0=L_i$ эквивалентна выбору $\bar{x}^0$, реализующего $\max\limits_{\bar{x}}\sum\limits_{i=1}^{n}f_i(\bar{x})$ и

$$\alpha_i^0+f_i(\bar{x}^0)=L_i+\sum_{j=1}^{n}(f_j(\bar{x}^0)-L_j)\Big/\rho_i\sum_{j=1}^{n}\frac{1}{\rho_j}.$$

Введем $b_i=\max\limits_{y_i\in M_i} f_i(x_1^0,\ldots,y_i,\ldots,x_n^0)-L_i$ и нормировку $\sum\limits_{i=1}^{n}(1/\rho_i)=1$. Тогда минимизация $K(\bar{x}^0,\bar{\alpha}^0)$ по $\bar{\rho}$ в предположении $b_i>0$, $1\leqslant i\leqslant n$, эквивалентна задаче

$$\min_{\bar{\rho}>0}\max_{1\leqslant i\leqslant n}\rho_i b_i=z,$$

в которой

$$\rho_i^0 = \left(\sum_{j=1}^n b_j\right) \Big/ b_i, \quad z = \sum_{i=1}^n b_i;$$

$$K(\bar{x}^0, \bar{\alpha}^0) = z \Big/ \sum_{i=1}^n (f_i(\bar{x}^0) - L_i);$$

$$f_i(\bar{x}^0) + \alpha_i^0 = L_i + \left(b_i \Big/ \sum_{j=1}^n b_j\right) \sum_{j=1}^n (f_j(\bar{x}^0) - L_j).$$

Аналогично находится и последовательность ε-оптимальных решений в случае, когда имеются $b_i \leqslant 0$.

4. Не представляет труда обобщение понятия устойчивости в повторяющихся играх на случай возможности одновременного нарушения совместного решения несколькими игроками. В дальнейших исследованиях по проблеме устойчивости должны занять свое место и вопросы обнаружения нарушителей и ограничения возможностей наказания за счет допустимости результатов наказывающих игроков в каждом повторении.

В заключение обратим внимание на иное использование повторений игры в интересах коалиции всех игроков. Если оставить параметры компромисса такими же, как в (4), то при r-кратном повторении игры для средних результатов вместо (4) имеем [8]

$$W_{\bar{\rho}}^r = \min_{1 \leqslant i \leqslant n} \rho_i \left(\frac{1}{r} \sum_{j=1}^r f_i(\bar{x}_j) - W_i^0\right). \tag{11}$$

Максимизация этого критерия должна теперь происходить по совокупности $(\bar{x}_1, \ldots, \bar{x}_r)$. Расширение множества стратегий при переходе от (4) к (11) приводит, вообще говоря, к увеличению средних результатов для всех игроков [9]. Нетрудно видеть, что максимум (11) не превысит максимума по возможным вероятностным мерам $\sigma(d\bar{x})$ выражения

$$\overline{W}_{\bar{\rho}} = \min_{1 \leqslant i \leqslant n} \rho_i \int f_i(\bar{x}) \sigma(d\bar{x}).$$

Как отмечено, например, в [4], использование (11) не может дать результатов больших, чем дают неограниченные побочные платежи в одном повторении.

Вычислительный центр
Академии наук СССР
Москва

Поступило
1 X 1973

ЦИТИРОВАННАЯ ЛИТЕРАТУРА

[1] *Н. Н. Воробьев*, УМН, т. 25, в. 2, 81 (1970). [2] *Р. Д. Льюс, Х. Райфа*, Игры и решения, ИЛ, 1961. [3] *Оуэн*, Теория игр, «Наука», 1971. [4] *Ю. Б. Гермейер*, Игры с непротивоположными интересами, М., 1972. [5] *Ю. Б. Гермейер*, ДАН, т. 198, № 5, 1001 (1971). [6] *В. А. Горелик*, Журн. вычислит. матем. и матем. физ., т. 12, 510, № 2 (1972). [7] *Д. А. Молодцов, В. В. Федоров*, Там же, т. 13, № 6 (1973). [8] *В. В. Морозов*, Там же, т. 11, № 3, 611 (1971). [9] *А. Ф. Кононенко, Н. С. Кукушкин*, ДАН, т. 209, № 6, 1274 (1973).