УДК 681.5

С.И. Доценко

ИГРОВЫЕ СИТУАЦИИ В МОДИФИЦИРОВАННОЙ ЗАДАЧЕ ОПТИМАЛЬНОГО ВЫБОРА

Для задачи оптимального выбора наилучшего или второго по качеству объекта рассмотрена игровая ситуация, в которой участвуют два игрока, осуществляющие свой выбор на двух различных множествах объектов. Данная игровая ситуация рассмотрена в двух модификациях в зависимости от информации, доступной игрокам. В обеих ситуациях найдены оптимальные стратегии игроков, обеспечивающие равновесие по Нэшу. Также рассмотрена альтернативная игровая ситуация, в которой два игрока осуществляют просмотр на одном множестве объектов и целью каждого из них является выбор объекта, лучшего чем у противника.

Введение

Задача выбора наилучшего объекта на сегодняшний день является одной из классических задач исследования операций, а именно стохастической оптимизации. Изначально она была придумана Мартином Гарднером в середине 20-го века как головоломка. В последствии оказалось, что данная задача и ее множественные модификации могут служить иллюстративными примерами в теории оптимальной остановки марковских процессов и теории игр. В основе анализа многошаговых игр лежит так называемая концепция сложного рационального поведения игроков, когда каждый игрок исходит из таких предпосылок: «Я веду себя оптимальным образом на каждом шаге. Я знаю, что противник ведет себя оптимальным образом. Я знаю, что противник знает, что я веду себя оптимальным образом и т.д.».

В [1] была рассмотрена задача выбора наилучшего объекта, сформулированная следующим образом. Пусть некто в случайном порядке знакомится с n объектами и хочет выбрать среди них наилучший. При этом после ознакомления с очередным объектом нужно либо остановить на нем свой выбор, либо отвергнуть его, возвращаться к ранее просмотренным объектам нельзя. Объекты являются упорядоченными определенным образом по качеству, т.е. качества любых двух объектов сравнимы между собой. «Ознакомление в случайном порядке» означает, что изначально все $n!$ перестановок, задающих порядок просмотра объектов, равновероятны.

Объект, наилучший среди всех n, в дальнейшем будем называть *наилучшим*, а объект, лучший среди k просмотренных, – *максимальным*. Очевидно, что в ходе просмотра следует анализировать целесообразность остановки выбора на некотором объекте, только если он является максимальным. При этом оказывается, что первый объект является максимальным и индексы максимальных объектов образуют цепь Маркова с переходными вероятностями $p(k, j) = \frac{k}{j(j-1)},\ j > k.$ Независимо от того, был ли k-й элемент максимальным или нет, вероятность того, что среди элементов с индексами $k+1,...,n$ минимальный индекс максимального элемента будет j, равна $\frac{k}{j(j-1)}$, $j>k$ и с вероятностью $\frac{k}{n}$ в последовательности $k+1,...,n$ не встретится ни одного максимального элемента.

Доказано, что для того, чтобы выбрать наилучший объект из n, нужно придерживаться такой стратегии: вначале пропустить все элементы с индексами $1,...,k^*-1$ и затем остановить свой выбор на первом максимальном элементе, индекс которого не меньше k^*, где k^* определяется из двойного неравенства $\frac{1}{k^*}+...+\frac{1}{n-1}\le 1<\frac{1}{k^*-1}+...+\frac{1}{n-1}$. Оказывается, что при $n\to\infty$, $\frac{k}{n}\to\frac{1}{e}$, а вероятность выбора наилучшего объекта при соблюдении описанной стратегии стремится к 1 / e. Множество индексов элементов, на которых необходимо остановиться, если они окажутся максимальными, будем обозначать $\tilde{A}^*=\{k^*,...,n\}$.

В [2] было рассмотрено обобщение задачи из [1]. Пусть в результате просмотра требуется найти не обязательно наилучший объект, а такой, ранг которого среди n не превышает r (т.е. объект является r-м по качеству среди всех). Обозначим ранг k-го по счету объекта среди m просмотренных через $R(a_k\mid m)$. Состояние просмотра будем описывать парой (k, i), где k — номер просматриваемого объекта, i — его ранг среди k просмотренных, т.е. $R(a_k\mid k)=i$.

В [1] было показано, что если изначально все перестановки $a_1,...,a_n$ равновероятны, то независимо от взаимного расположения $a_1,...,a_{k-1}$ ранг a_k в последовательности $a_1,...,a_k$ с равными вероятностями $1/k$ принимает одно из возможных значений $1,...,k$ (назовем это свойство (*)). С использованием этого свойства в [2] было показано, что опорное множество состояний $\tilde{A}$, на которых следует остановить просмотр, обладает такими признаками: $(k,i)\in\tilde{A}\Rightarrow(k+1,i)\in\tilde{A}$, $(k,i)\in\tilde{A}\Rightarrow(k,i-1)\in\tilde{A}$.

В случае если ценность выбранного объекта зависит от ранга и функция ценности является невозрастающей (т.е. пусть за выбор i-го по качеству объекта $i=\overline{1,r}$ выплачивается выигрыш α_i, причем $\alpha_1\ge\alpha_2\ge...\ge\alpha_r$, то указанные свойства опорного множества сохраняются. Поэтому существует набор целых чисел $s_1\le s_2\le...\le s_r\le n$, такой, что $\tilde{A}=\{(s_1,1),...,(n,1)\}\cup\{(s_2,2),...,(n,2)\}\cup...\cup\{(s_r,r),...,(s_r,n)\}$, а оптимальная стратегия, максимизирующая ожидаемую величину выигрыша, имеет следующий вид:

1) пропустить все объекты от 1 до s_1-1;

2) остановиться на первом максимальном объекте, индекс которого не превышает s_2-1, а если такового не найдется, то перейти к п.3;

3) остановиться на первом объекте, индекс которого не превышает s_3-1, а ранг не превышает 2, и т.д.;

r+1) начиная с элемента s_r, остановиться на элементе, ранг которого не превышает *r*.

Конкретные значения $s_1,..., s_r$ зависят от *n* и от $\alpha_1,..,\alpha_r$.

Рассмотрим частный случай $r=2, \alpha_1=1, \alpha_2=\alpha \in [0;1]$. Найдем предельные значения s_1/n и s_2/n при $n \to \infty$ в зависимости от параметра α. Для этого предварительно найдем некоторые вероятности, вычисление которых базируется на свойстве (*) и правилах суммы и произведения вероятностей:

0) вероятность того, что в ряду $k+1, k+2,...$ ближайший максимальный элемент будет иметь индекс *j*,

$$P(k,j) = P\big(R(a_{k+1} \mid k+1)>1, R\big(a_{k+2} \mid k+2\big)>1,..., R\big(a_{j-1} \mid j-1\big)>1, R\big(a_j \mid j\big)=1\big)=$$

$$= P\big(R(a_{k+1} \mid k+1)>1\big)\times ... \times P\big(R\big(a_{j-1} \mid j-1\big)>1\big)\times P\big(R\big(a_j \mid j\big)=1\big)=\frac{k}{j(j-1)};$$

1) вероятность того, что элемент окажется лучшим из *n* при условии, что он является лучшим из *k*,

$$P_1(k,n) = P\big(R(a_k \mid n)=1 \mid R(a_k \mid k)=1\big)=$$

$$= P\big(R(a_{k+1} \mid k+1)>1\big)\times ... \times P\big(R(a_n \mid n)>1\big)=\frac{k}{n};$$

2) вероятность того, что элемент окажется вторым из *n* при условии, что он является лучшим среди *k*,

$$P_2(k,n) = P\big(R(a_k \mid n)=2 \mid R(a_k \mid k)=1\big)=$$

$$= \sum_{j=k+1}^{n} P\big(R(a_{k+1} \mid k+1)>1\big)\times ... \times = P\big(R\big(a_{j-1} \mid j-1\big)>1\big)\times$$

$$\times P\big(R\big(a_j \mid j\big)=1\big)\times P\big(R\big(a_{j+1} \mid j+1\big)>1\big)\times ... \times P\big(R(a_n \mid n)>1\big)=\frac{k(n-k)}{n(n-1)};$$

3) вероятность того, что элемент окажется вторым из *n* при условии, что он является вторым среди *k*,

$$P_3(k,n) = P\big(R(a_k \mid n)=2 \mid R(a_k \mid k)=2\big)=$$

$$= P\big(R(a_{k+1} \mid k+1)>2\big)\times ... \times = P\big(R(a_n \mid n)>2\big) = \frac{k(k-1)}{n(n-1)};$$

4) вероятность того, что при дальнейшем просмотре элементов $k+1,..., n$ среди элементов $k+1,..., j-1$ не будет встречаться максимальных или вторых по качеству, а *j*-й элемент окажется лучшим среди просмотренных, равна

$$P_{41}(k,j)=P(R(a_{k+1}\mid k+1)>2,\,R(a_{k+2}\mid k+2)>2,\dots,R(a_{j-1}\mid j-1)>$$
$$>2,\,R(a_j\mid j)=1=P_{42}(k,j)=P(R(a_{k+1}\mid k+1)>2,\,R\big(a_{k+2}\mid k+2\big)>$$
$$>2,\dots,R(a_{j-1}\mid j-1)>2,\,R(a_j\mid j)\;=2)=P(R\big(a_{k+2}\mid k+2\big)>2)\times$$
$$\times\dots\times P(R(a_{k+1}\mid k+1)>2)\times P(R(a_{j-1}\mid j-1)>2)\times P(R(a_j\mid j)=1)=$$
$$=P(R(a_{k+1}\mid k+1)>2)\times P(R(a_{k+2}\mid k+2)>2)\times\dots\times P(R(a_{j-1}\mid j-1)>$$
$$>2)\times P(R\big(a_j\mid j\big)=2)=\frac{k-1}{k+1}\times\frac{k}{k+2}\times\dots\times\frac{j-3}{j-1}\times\frac{1}{j}=\frac{k(k-1)}{j(j-1)(j-2)}.$$

Пусть $f((k,i))$ — средний выигрыш при остановке на элементе $(k,\ i)$, тогда

$$f((k,1))=P_1(k,n)\times 1+P_2(k,n)\times\alpha=\frac{k}{n}+\alpha\frac{k(n-k)}{n(n-1)},$$
$$f((k,2))=\alpha P_3(k,n)=\alpha\frac{k(k-1)}{n(n-1)}.$$

Поскольку по определению $(s_2-1{,}2)\notin\tilde{A}$, а $(s_2,2),\dots(n,2)\in\tilde{A}$, $(s_2,1),\dots(n,1)\in\tilde{A}$, то

$$f(s_2-1{,}2)\le\sum_{j=s2}^{n}P_{41}\big(s_2-1,j\big)f(j,1)+\sum_{j=s2}^{n}P_{42}\big(s_2-1,j\big)f(j,2).$$

Обозначим $\lim\limits_{n\to\infty}\frac{s_2}{n}=\gamma$, тогда, после подстановок, элементарных преобразований и перехода к пределу в неравенстве имеем:

$$\alpha\le\int_{\gamma}^{1}\frac{1}{t^2}dt+\alpha\int_{\gamma}^{1}\frac{1-t}{t^2}dt+\alpha\int_{\gamma}^{1}\frac{1}{t}dt=\frac{1}{\gamma}-1+\alpha\left(\frac{1}{\gamma}-1\right)\qquad(1)$$

С другой стороны,

$$f(s_2,2)=P_3(s_2,n)\times\alpha\ge\sum_{j=s2+1}^{n}P_{41}\big(s_2,j\big)f(j,1)+\sum_{j=s2+1}^{n}P_{42}\big(s_2,\ j\big)f(j,2),$$

отсюда, после аналогичных преобразований, получим неравенство (1) с противоположным знаком, отсюда

$$\gamma=\frac{\alpha+1}{2\alpha+1}.\qquad(2)$$

Исследуем асимптотическое поведение величины $\beta=\frac{s_1}{n}$ при $n\to\infty$.

Поскольку $(s_1-1{,}1)\notin\tilde{A}$, а $(s_1,1),\dots(n,1)\in\tilde{A}$, $(s_2,2),\dots(n,2)\in\tilde{A}$, то

$$f(s_1-1{,}1)\le\sum_{j=s1}^{s2-1}P(s_1-1,j)\times f(j,1)+$$
$$+\frac{s_1-1}{s_2-1}\sum_{j=s2}^{n}\big(P_{41}(s_2-1,j)\times f(j,1)+P_{42}\big(s_2-1,\ j\big)\times f(j,2)\big);$$

$$f(s_1,1) \ge \sum_{j=s1+1}^{s2-1} P(s_1, j) \times f(j,1) +$$

$$+\frac{s_1}{s_2-1} \sum_{j=s2}^{n} \left(P_{41}(s_2-1, j) \times f(j,1) + P_{42}(s_2-1, j) \times f(j,2)\right).$$

Переход к пределу в данных неравенствах приводит к равенству $1+\alpha-\alpha\beta=(1+\alpha)\ln\left(\frac{\gamma}{\beta}\right)+\alpha\beta$, отсюда β — корень уравнения

$$\frac{2\alpha}{\alpha+1}\beta-\ln(\beta)=1-\ln\left(\frac{\alpha+1}{2\alpha+1}\right). \tag{3}$$

При этом вероятности того, что выбранный элемент является первым/вторым сходятся к таким величинам:

$$\pi_1 = \sum_{j=s1}^{s2-1} P(s_1-1, j)P_1(j,n) + P_1(s_1-1, s_2-1) \times \sum_{j=s2}^{n} P_{41}(s_2-1, j) \times P_1(j,n),$$

$$\pi_1 \overset{n\to\infty}{\to} \beta\ln\left(\frac{\gamma}{\beta}\right)+\beta(1-\gamma)=\beta\left(\frac{\alpha}{2\alpha+1}+\ln\left(\frac{\alpha+1}{2\alpha+1}\right)-\ln(\beta)\right),$$

$$\pi_2 = \sum_{j=s1}^{s2-1} P(s_1-1, j)P_2(j,n) + P(s_1-1, s_2-1)\times$$

$$\times\left(\sum_{j=s2}^{n}\left(P_{41}(s_2-1, j)\times P_2(j,n)+P_{42}(s_2-1, j)\times P_3(j,n)\right)\right),$$

$$\pi_2 \overset{n\to\infty}{\to} \beta\ln\frac{\gamma}{\beta}-\beta(\gamma-\beta)+\beta(1-\gamma)=\beta\left(\ln\left(\frac{\alpha+1}{2\alpha+1}\right)-\ln(\beta)+\beta-\frac{1}{2\alpha+1}\right),$$

а средняя величина выигрыша с учетом (3) равна
$v=\pi_1+\alpha\pi_2=\beta(1+\alpha-\alpha\beta)$.

Обозначим предельные значения

$$\begin{aligned} g_1(\alpha,\beta)&=\beta\left(\frac{\alpha}{2\alpha+1}+\ln\left(\frac{\alpha+1}{2\alpha+1}\right)-\ln(\beta)\right),\\ g_2(\alpha,\beta)&=\beta\left(\ln\left(\frac{\alpha+1}{2\alpha+1}\right)-\ln(\beta)+\beta-\frac{1}{2\alpha+1}\right).\end{aligned} \tag{4}$$

Назовем стратегией 1–2 следующую последовательность действий. Пусть задано $0\le\alpha\le1$ — отношение важности второго по качеству элемента к важности наилучшего. Найдем β и γ по формулам (3) и (2). В ходе просмотра следует придерживаться такой стратегии:

1) пропустить все элементы от 1 до $\beta n-1$;

2) при просмотре элементов с номерами от βn до $\gamma n-1$ остановиться на первом максимальном элементе, если таковой найдется, иначе перейти к п. 3;

3) при просмотре элементов с номерами от γn до n остановиться на первом элементе, текущий ранг которого не превышает 2.

Первая игровая ситуация — борьба за выбор лучшего объекта при выборе из разных множеств

Пусть в выборе принимают участие два игрока, которые осуществляют свой выбор на двух разных множествах элементов. Для каждого из игроков возможны три исхода просмотра:

1) выбран наилучший (1-й) элемент;

2) выбран второй по качеству (2-й) элемент;

3) не выбран ни один из элементов из п. 1, 2.

Считается, что в данной нумерации чем меньше номер, тем лучше исход. Игрок, исход выбора которого оказался хуже, платит единичную сумму другому игроку. Если исходы игроков одинаковы, то фиксируется ничья и платежей не производится.

Рассмотрим два варианта игры.

1. Несимметричная игра

Пусть исход просмотра 1-го игрока становится известным 2-му игроку, до того как он начнет просмотр.

Пусть исход 1-го игрока — 1. Тогда 2-й игрок должен выбрать стратегию, максимизирующую вероятность выбора 1-го по качеству элемента, т.е. положить $\alpha=1$. Тогда $\beta=\frac{1}{e}$, $\gamma=1$, т.е. второй игрок должен придерживаться стратегии классической задачи, что обеспечит выбор 1-го по качеству элемента с вероятностью $\frac{1}{e}$, и, таким образом, средний выигрыш 2-го игрока составит $+\frac{1}{e}\times 0+\left(1-\frac{1}{e}\right)\times(-1)\approx -0{,}6321$.

Пусть исход 1-го игрока — 2. Тогда если 2-й игрок придерживается стратегии, обеспечивающей ему выбор 1-го/2-го элемента с вероятностями p_1 и p_2 соответственно, то его средний выигрыш равен

$$p_1\times 1+p_2\times 0+(1-p_1-p_2)\times(-1)=2p_1+p_2-1=2\left(p_1+\frac{1}{2}p_2\right)-1,$$

т.е. коэффициенты важности выбора 1-го и 2-го элементов относятся как 2 к 1, значит он должен придерживаться стратегии 1–2, положив $\alpha=1/2$, тогда $\beta\approx 0{,}34794$, $\gamma=0{,}75$, $\max\left(p_1+\frac{1}{2}p_2\right)\approx 0{,}46138$ и, таким образом, средний выигрыш 2-го игрока равен $2\times 0{,}46138-1\approx -0{,}0772$.

Пусть исход 1-го игрока — 3. Тогда ожидаемый выигрыш 2-го игрока равен $(p_1+p_2)\times 1+(1-p_1-p_2)\times 0=p_1+p_2$. Положив $\alpha=1$, находим, что $\max(p_1+p_2)$ достигается при $\beta\approx 0{,}3470$, $\gamma=2/3$ и равно 0,5736.

Проанализируем поведение 1-го игрока. Если он придерживается стратегии, обеспечивающей ему выбор 1-го/2-го элемента с вероятностями p_1 и p_2 соответственно, то его средний выигрыш равен

$$0{,}6321p_1+0{,}0772p_2-0{,}5736(1-p_1-p_2)=1{,}2057p_1+0{,}6508p_2-0{,}5736=$$
$$=1{,}2057(p_1+0{,}5398p_2)-0{,}5736.$$

Положив $\alpha \approx 0{,}5398$, находим $\beta \approx 0{,}3476$, $\gamma \approx 0{,}7404$, $\max(p_1 + 0{,}5398 p_2) \approx 0{,}4710$ и, таким образом, выигрыш 1-го игрока составляет $1{,}2057 \times 0{,}4710 - 0{,}5736 \approx -0{,}0057$.

То, что выигрыш 1-го игрока отрицательный, закономерно, поскольку игра почти симметрична за исключением того, что 2-й игрок знает исход выбора 1-го и адаптирует свою стратегию в соответствии с этим исходом.

Таким образом, оптимальные стратегии игроков такие: 1-й игрок применяет стратегию 1–2 с $\beta \approx 0{,}3476, \gamma \approx 0{,}7404$. Узнав об исходе выбора 1-го игрока, 2-й игрок применяет стратегию 1–2. Параметры стратегии приведены в табл. 1.

Таблица 1

Исход	β	γ
1	1 / e	1
2	0,3479	3 / 4
3	0,3470	2 / 3

2. Симметричная игра

Пусть игроки осуществляют просмотр независимо друг от друга, т.е. исход просмотра каждого из игроков держится в тайне от другого до момента сопоставления исходов. Найдем пару равновесных стратегий методом нащупывания по Курно.

Пусть 1-й игрок применил некоторую стратегию, при которой исходы 1 и 2 реализуются с вероятностями q_1, q_2 соответственно, а второй игрок противопоставил стратегию, при которой эти исходы реализуются с вероятностями r_1, r_2. Тогда выигрыш второго игрока составляет

$$(1+q_2)r_1 + (1-q_1)r_2 - (q_1+q_2) = (1+q_2)\left[r_1 + \left(\frac{1-q_1}{1+q_2}\right)r_2\right] - (q_1+q_2).$$

Таким образом, наилучшим ответом второго игрока будет применение стратегии 1–2 при $\alpha = \dfrac{1-q_1}{1+q_2}$.

Положим $\alpha_1 = 0{,}5$. Пусть β_i — корень уравнения (3), соответствующий α_i.

Обозначим $q_1^{(i)} = g_1(\alpha_i, \beta_i)$, $q_2^{(i)} = g_2(\alpha_i, \beta_i)$, согласно (4). Положим $\alpha_{i+1} = \dfrac{1-q_1^{(i)}}{1+q_2^{(i)}}$.

При итеративном применении процедуры нащупывания по Курно получаются такие предельные значения:

$$\alpha \approx 0{,}5318, \beta \approx 0{,}3467, \gamma \approx 0{,}7423, p_1 \approx 0{,}3533, p_2 \approx 0{,}2161.$$

Заметим, что найденные значения α, β, γ близки к соответствующим значениям для несимметричной игры.

Данная стратегия дает нулевой средний выигрыш против такой же стратегии и положительный выигрыш против любой другой стратегии.

Вторая игровая ситуация — борьба за выбор лучшего объекта при выборе из одного множества

Пусть теперь игроки осуществляют просмотр на одном множестве элементов. При этом каждый объект просматривается обоими игроками одновременно при следующих условиях. Как и в [1], каждый из игроков имеет возможность выбора — останавливаться на текущем объекте или продолжать просмотр. И точно также отсутствует возможность возврата к ранее просмотренным объектам. Если на одном из объектов желает остановиться только один из игроков, то он останавливается на нем и выбывает из игры, другой игрок продолжает просмотр. Если же на некотором объекте желают остановиться оба игрока, то право остановки разыгрывается между ними жеребьевкой:

1-й/2-й игрок выигрывает жребий с вероятностью $\alpha/1-\alpha$ соответственно. Выигравший жребий останавливается и выбывает из игры, проигравший продолжает просмотр.

В [5] была рассмотрена задача, в которой целью каждого из игроков является выбор наилучшего объекта с максимальной вероятностью. Далее рассмотрим задачу, аналогичную по постановке, рассмотренной в [5]. Пусть теперь целью каждого из игроков является выбор объекта, лучшего, чем тот, который выберет противник (в этом случае игрок выигрывает). Оказывается, что такое, на первый взгляд, несущественное изменение постановки задачи приводит к радикальному изменению стратегий игроков.

В [1] было показано, что если изначально все перестановки элементов $a_1,\dots,a_n$ равновероятны, то независимо от взаимного расположения элементов $a_1,\dots,a_{k-1}$ ранг элемента a_k в последовательности $a_1,\dots,a_k$ с равными вероятностями $1/k$ принимает одно из возможных значений $1,\dots,k$ (где ранг 1 означает, что a_k — наилучший, k — наихудший среди $a_1,\dots,a_k$). Тогда, согласно формуле произведения вероятностей, вероятность того, что объект является i-м по качеству среди n при условии, что он является i-м по качеству среди k (или, другими словами, вероятность того, что в последовательности $a_{k+1},\dots,a_n$ не встретится ни одного объекта лучше a_k), равна

$$p(k,i)=\frac{k+1-i}{k+1}\times\frac{k+2-i}{k+2}\times\dots\times\frac{n-i}{n}=\frac{(n-i)!k!}{(k-i)!n!}=\frac{k(k-1)\dots(k-i+1)}{n(n-1)\dots(n-i+1)}. \qquad (5)$$

Рассмотрим частный случай $\alpha=1$. Тогда второй игрок сможет остановиться на очередном элементе, если этого не захотел сделать первый игрок.

Текущее состояние игры характеризуется парой (k,i), где k — номер просматриваемого элемента, i — его ранг среди k просмотренных.

Если один из игроков останавливается на некотором элементе, то об этом становится известно другому игроку и его дальнейшая стратегия становится тривиальной — остановиться на первом элементе лучше a_k, если таковой встретится. При этом с вероятностью $p(k,i)$ такой элемент не встретится и выигрывает игрок, сделавший остановку и с дополнительной вероятностью — встретится и выигрывает его противник, продолживший просмотр. Поэтому игру можно переформулировать таким образом: если остановку сделал один из игроков, то он получает выигрыш $f(k,i)=2p(k,i)-1$ (эта величина может быть и отрицательной), его противник получает выигрыш $-f(k,i)$ и на этом игра заканчивается. Если остановку не сделал ни один из игроков, то игра продолжается и с появлением элемента a_{k+1} с равными вероятностями $\frac{1}{k+1}$ система переходит в одно из состояний $(k+1, j)$, $j=\overline{1,k+1}$.

При $k=n$ согласно формуле (5) $p(n,i)=1$, следовательно, $f(n,i)=1$ при всех $i=\overline{1,n}$. По сути это означает, что если до n-го шага оба игрока не остановились на каких-либо предыдущих элементах, то на n-м шаге первый игрок выбирает a_n (который может иметь какой угодно ранг) и выигрывает у второго игрока, не выбравшего никакого элемента.

Найдем оптимальные стратегии игроков. Стратегии будем описывать множествами $\widetilde{A}_1, \widetilde{A}_2$, где $\widetilde{A}_1, \widetilde{A}_2$ — подмножества $\{(k,i) \mid k=\overline{1,n},\ i=\overline{1,n}\}$. При этом $(k,i)\in \widetilde{A}_j$, $j=1;2$ означает, что j-й игрок намерен остановиться на элементе (k,i), если будет такая возможность.

Назовем ценой игры $v(k,i)$ средний выигрыш первого игрока при условии, что игра находится в состоянии (k,i) и в дальнейшем игроки придерживаются своих оптимальных стратегий. Обозначим $V(k)=\frac{1}{k}\sum_{i=1}^{k} v(k,i)$ — средний выигрыш на шаге k (усредненный по всем состояниям $v(k,i)$). Чтобы найти оптимальные стратегии игроков, проанализируем игру методом динамического программирования, двигаясь по множеству состояний в обратном порядке, начиная с последнего шага.

Очевидно, что $v(n,i)=1$, $i=1,n$ и, таким образом, $V(n)=1$.

Пусть определено $V(k+1)$, тогда

$$v(k,i)=\max(f(k,i), \min(-f(k,i), V(k+1))). \tag{6}$$

Действительно, пусть игра находится в состоянии (k,i). Тогда первый игрок имеет альтернативу – останавливаться или нет. Если первый игрок остановился, то его выигрыш составил $f(k,i)$, если нет, то право выбора, останавливаться или нет, перешло ко второму игроку, т.е. он может остановиться (и тогда выигрыш первого игрока составит $-f(k,i)$ или не остановиться (тогда игра продолжается и ожидаемый выигрыш первого игрока равен $V(k+1)$). Поскольку второй игрок стремится минимизировать выигрыш первого, то он выбирает минимум из этих двух величин. Поскольку первый игрок понимает это и стремится максимизировать свой выигрыш, то он выберет максимум из $f(k,i)$ и $\min(-f(k,i), V(k+1))$. Покажем, что $V(k)>0$ для всех k. Действительно, $V(n)=1>0$. Если $V(k+1)>0$, то согласно (2) все $v(k,i)\geq 0$ и есть такие, которые строго больше нуля, поэтому $V(k)>0$.

Будем полагать, что $(k,i)\in\tilde{A}_1$, если $f(k,i)>\min(-f(k,i),V(k+1))$. Поскольку $V(k+1)>0$, то последнее условие эквивалентно $f(k,i)>0$. Далее, $(k,i)\in\tilde{A}_2$, если $-f(k,i)\le V(k+1)\Rightarrow f(k,i)\ge -V(k+1)$, поэтому $\tilde{A}_2\supset\tilde{A}_1$.

Рассмотрим случай $\frac{1}{2}<\alpha<1$. Тогда каждый из игроков имеет альтернативу — останавливаться или нет, так что в состоянии $(k,\ i)$ выигрыш первого игрока описывается матричной игрой (табл. 2).

Пусть $f(k,i)>0$. Тогда $f(k,i)>(2\alpha-1)f(k,i)>-f(k,i)$ и $-f(k,i)<V(k+1)$, таким образом, игра разрешима по доминированию – обоим игрокам выгодно останавливаться, и, таким образом, $(k,i)\in\tilde{A}_1,(k,i)\in\tilde{A}_2,V(k,i)=f(k,i)$.

Таблица 2

2-й игрок / 1-й игрок	останавливаться	не останавливаться
останавливаться	$(2\alpha-1)f(k,i)$	$f(k,i)$
не останавливаться	$-f(k,i)$	$V(k+1)$

Пусть $f(k,i)\le 0$. Тогда $f(k,i)\le(2\alpha-1)f(k,i)\le -f(k,i)$ и $f(k,i)<V(k+1)$. В этом случае стратегия первого игрока «останавливаться» доминируется стратегией «не останавливаться», значит $(k,i)\notin\tilde{A}_1$. Если при этом $-f(k,i)\le V(k+1)$, то оптимальная стратегия второго игрока – останавливаться, и, таким образом, $(k,i)\in\tilde{A}_2$ и $v(k,i)=-f(k,i)$, а если $-f(k,i)>V(k+1)$, то не останавливаться и $(k,i)\notin\Gamma_2$ и $v(k,i)=V(k+1)$.

Объединяя, имеем:

$$\tilde{A}_1=\{(k,i)\mid f(k,i)>0\},\quad \tilde{A}=\{(k,i)\mid f(k,i)\ge -V(k+1)\},$$

$$v(k,i)=\begin{cases}(2\alpha-1)f(k,i),\ f(k,i)>0,\\ \min(-f(k,i),V(k+1)),\ f(k,i)\le 0;\end{cases}\quad V(k)=\frac{1}{k}\sum_{i=1}^{k}v(k,i).$$

Таким образом, средний выигрыш первого игрока в самом начале, при условии, что оба игрока придерживаются своих оптимальных стратегий, равен $U=V(1)$, а вероятность выигрыша равна $p=\frac{V(1)+1}{2}$.

Определим начальные условия. Если оба игрока не сделали своего выбора до n-го (т.е. самого последнего) шага игры, то выигрывает игрок, которому достался последний элемент, независимо от его ранга, поэтому оба игрока должны на него претендовать. Следовательно, при $k=n$ для всех $i=1,n$ справедливы соотношения: $f(n,i)=1,\ v(n,\ i)=2\alpha-1,\ (n,i)\in\tilde{A}_1,(n,\ i)\in\tilde{A}_2$.

Рассмотрим случай $\alpha=\frac{1}{2}$. Тогда данная антагонистическая игра становится симметричной, и цена такой игры равна нулю. Заметим, что стратегия $\tilde{A}=\{(k,i)\mid f(k,i)>0\}$ обеспечивает каждому игроку неотрицательный выигрыш против любой стратегии противника, и, разумеется, нулевой выигрыш против такой же стратегии.

Для некоторых значений n и α были сделаны расчеты в пакете Mathcad. В табл. 3 приведены стратегии игроков при $n = 20$, $\alpha = \frac{1}{2}; \frac{3}{4}; 1$, в табл. 4 — значения U при различных значениях n и α.

Таблица 3

$\tilde{A}_1; \tilde{A}_2, \alpha = 1/2$	
k	i
11–14	1
15–16	1–2
17	1–3
18	1–5
19	1–9

$\tilde{A}_2,\ \alpha = 3/4$	
k	i
9–12	1
13–14	2
15–16	3
17	5
18	7
19	15

$\tilde{A}_2,\ \alpha = 1$	
k	i
8–11	1
12–14	1–2
15	3
16	4
17	6
18	9
19	1–19

Таблица 4

n	α				
	1 / 2	5 / 8	3 / 4	7 / 8	1
10	0	0,087	0,154	0,206	0,238
100	0	0,077	0,136	0,183	0,222
1000	0	0,073	0,131	0,178	0,218

Выводы. При всех $\alpha \in \left[\frac{1}{2}; 1\right]$ оптимальная стратегия первого игрока остается неизменной: $\tilde{A}_1 = \{(k, i) \mid f(k, i) > 0\}$, а стратегия второго игрока зависит от α и может быть найдена по приведенным выше формулам. Можно показать, что $\tilde{A}_2$ сужается с ростом α от $\frac{1}{2}$ до 1, т.е. $\tilde{A}_2(\alpha) \supset \tilde{A}_2(\beta)$ при $\frac{1}{2} \le \alpha < \beta \le 1$.

Рассмотренные задачи являются лишь одной из многочисленных возможных игровых ситуаций, возникающих при рассмотрении задачи оптимального выбора. К сожалению, приходится прийти к выводу, что нет общей методики решения задач данного типа — для каждой задачи требуется индивидуальный подход. Неизменной остается лишь использование марковского свойства игровых ситуаций и концепции сложного рационального поведения игроков.

1. *Дынкин Е.Б., Юшкевич А.А.* Теоремы и задачи о процессах Маркова. — Москва: Наука, 1976. — С. 91–102.
2. *Гусейн-Заде С. М.* Задача выбора и оптимальное правило остановки последовательности независимых испытаний // Теория вероятности и ее применение. — 1966. — 11, № 3. — С. 534–537.
3. *Мазалов В.В.* Математическая теория игр и приложения. — СПб.: Лань, 2010. — 446 с.
4. *Мазалов В.В.* Игровые моменты остановки. — Новосибирск: Наука, 1987. — 191 с.
5. *Доценко С.И.* Задача выбора наилучшего объекта как игра двух лиц // Кибернетика и Вычислительная техника. — 2011. — Вып. 164, — С. 43–53.

Киевский национальный
университет имени Тараса Шевченко

Получено 11.06.2012