УДК 519.83

С.И. ДОЦЕНКО, А.В. МАРИНИЧ

ИГРЫ В ПОДСКАЗКУ, ВЫМОГАТЕЛЬСТВО И УГАДЫВАНИЕ В ЗАДАЧЕ ОПТИМАЛЬНОГО ВЫБОРА

Аннотация. Рассмотрена игровая задача оптимального выбора, в которой один из игроков стремится получить часть выигрыша, причитающегося другому игроку за нахождение наилучшего элемента. Механизмами воздействия одного игрока на другого является предложение более благоприятных условий поиска либо, наоборот, угроза создания менее благоприятных условий поиска наилучшего элемента. Найдены оптимальные стратегии игроков, образующие равновесие по Нэшу, и исследовано асимптотическое поведение найденных стратегий для случая, когда количество просматриваемых объектов стремится к бесконечности.

Ключевые слова: задача оптимального выбора, остановка цепи Маркова, матричная игра, равновесие по Нэшу, смешанная стратегия, пороговая стратегия.

ВВЕДЕНИЕ

В работе [1] была рассмотрена задача выбора наилучшего объекта, сформулированная следующим образом. Пусть в случайном порядке среди n объектов необходимо выбрать 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 упорядочены определенным образом по качеству, т.е. качества любых двух объектов сравнимы между собой. Ознакомление в случайном порядке предполагает, что изначально все $n!$ перестановок, задающих порядок просмотра объектов, равновероятны.

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

Доказано, что для выбора наилучшего объекта из n нужно придерживаться такой стратегии: вначале пропустить все элементы с индексами $1, \ldots, k_n - 1$ и затем остановить свой выбор на первом максимальном элементе, индекс которого не меньший k_n, при этом k_n определяется из двойного неравенства

$$\frac{1}{k_n} + \ldots + \frac{1}{n-1} \le 1 < \frac{1}{k_n - 1} + \ldots + \frac{1}{n-1}. \qquad (1)$$

Оказывается, что $\dfrac{k_n}{n} \to \dfrac{1}{e}$ при $n \to \infty$, а вероятность выбора наилучшего объекта при соблюдении описанной стратегии стремится к e^{-1}.

Пусть первый игрок просматривает каждый элемент до того, как его просмотрит второй игрок. За нахождение наилучшего элемента второй игрок получает единичный выигрыш. Первый игрок не может получить плату за обнаружение наилучшего элемента, но он может помочь сделать это второму, дав подсказку и получив за это от него определенную компенсацию. Определим правила подсказки и платы за нее. Перед началом игры первый игрок на свое усмотрение устанавливает плату за подсказку. Установленная плата не может изменяться на протяжении всей игры. При появлении очередного максимального элемента первый игрок имеет возможность просмотреть все элементы и предложить второму игроку купить за установленную ранее цену информацию о том, является ли данный максимальный элемент наилучшим. Такую возможность первый игрок имеет один раз, после использования которой он выбывает из игры.

Определим ценность подсказки, как разность вероятностей нахождения наилучшего элемента при использовании подсказки и без использования подсказки.

Исходя из концепции сложного рационального поведения, второй игрок купит подсказку, если ценность подсказки будет не меньшей, чем запрашиваемая цена. Найдем оптимальные стратегии игроков в данной игре.

Пусть второй игрок просматривает k-й элемент, который является максимальным. Рассмотрим отдельно случаи, когда $k < k_n$ и когда $k \geq k_n$.

Пусть $k < k_n$. Тогда без использования подсказки второму игроку следует пропустить все элементы до $k_n - 1$ и остановиться на первом максимальном элементе, начиная с k_n. Таким образом, вероятность выигрыша составит

$$\sum_{j=k_n}^{n} \frac{k_n - 1}{j(j-1)} \frac{j}{n} = \frac{k_n - 1}{n} \sum_{j=k_n-1}^{n-1} \frac{1}{j}.$$

При использовании подсказки вероятность выигрыша составит $\frac{k}{n} + \left(1 - \frac{k}{n}\right) \frac{k_n - 1}{n} \sum_{j=k_n-1}^{n-1} \frac{1}{j}$ и, таким образом, ценность подсказки определится как

$$f_1(n,k) := \frac{k}{n} \left(1 - \frac{k_n - 1}{n} \sum_{j=k_n-1}^{n-1} \frac{1}{j}\right). \tag{2}$$

Пусть $k \geq k_n$. Тогда без использования подсказки второму игроку следует немедленно остановиться на k-м элементе и вероятность выигрыша составит k / n. В случае использования подсказки с вероятностью k / n данный максимальный элемент окажется наилучшим и на нем следует остановиться, а с дополнительной вероятностью данный элемент не будет наилучшим и остановиться следует на следующем максимальном элементе. Величина выигрыша составит

$$\frac{k}{n} \cdot 1 + \left(1 - \frac{k}{n}\right) \sum_{j=k+1}^{n} \frac{k}{j(j-1)} \cdot \frac{j}{n} = \frac{k}{n} + \left(1 - \frac{k}{n}\right) \frac{k}{n} \sum_{j=k}^{n-1} \frac{1}{j}.$$

Таким образом, ценность подсказки составляет

$$f_2(n,k) := \left(1 - \frac{k}{n}\right) \frac{k}{n} \sum_{j=k}^{n-1} \frac{1}{j}. \tag{3}$$

Заметим, что при фиксированном n последовательность $(f_1(n,k))_{k \geq 1}$ не убывает на множестве $\{1, 2, \ldots, k_n\}$, а последовательность $(f_2(n,k))_{k \geq 1}$ не возрастает на множестве $\{k_n, \ldots, n\}$. Поэтому величины

$$\underline{k}_n(\alpha) := \sup\{k < k_n : f_1(n,k) \le \alpha\} \tag{4}$$

и

$$\bar{k}_n(\alpha) := \inf\{k \ge k_n : f_2(n,k) \ge \alpha\}, \tag{5}$$

в силу неравенств $\sup\{f_1(n,k) : k < k_n\} \le \frac{k_n}{n}\left(1-\frac{k_n}{n}\right) \le \inf\{f_2(n,k) : k \ge k_n\}$ корректно определены, по крайней мере, при $\alpha \le \frac{k_n}{n}\left(1-\frac{k_n}{n}\right)$, и выполняется неравенство $\underline{k}_n(\alpha)+1 \le \bar{k}_n(\alpha)$. Величина $\frac{k_n}{n}\left(1-\frac{k_n}{n}\right)$ представляет собой верхнее значение для цены подсказки, т.е. при цене, выше данной, подсказка не будет куплена первым игроком ни при каких обстоятельствах.

При данном $\alpha \le \frac{k_n}{n}\left(1-\frac{k_n}{n}\right)$ подсказка будет продана, если в множестве индексов $\{\underline{k}_n(\alpha)+1, \ldots, \bar{k}_n(\alpha)\}$ окажется максимальный элемент. Вероятность этого события равна $p_n(\alpha) = \sum_{j=\underline{k}_n(\alpha)+1}^{\bar{k}_n(\alpha)} \frac{\underline{k}_n(\alpha)}{j(j-1)} = \underline{k}_n(\alpha)\left(\frac{1}{\underline{k}_n(\alpha)} - \frac{1}{\bar{k}_n(\alpha)}\right) = 1 - \frac{\underline{k}_n(\alpha)}{\bar{k}_n(\alpha)}$, а средний выигрыш первого игрока составляет $g_{1,n}(\alpha) = \alpha \cdot p_n(\alpha)$. Пусть $M_n := \sup\left\{g_{1,n}(\alpha) : \alpha \in \left[0, \frac{k_n}{n}(1-\frac{k_n}{n})\right]\right\}$; тогда при фиксированном n оптимальной ценой подсказки для первого игрока будет значение $\alpha_0 := = \inf\left\{\alpha \in \left[0, \frac{k_n}{n}\left(1-\frac{k_n}{n}\right)\right] : g_{1,n}(\alpha) = M_n\right\}$. Заметим, что указанные величины корректно определены в силу кусочной непрерывности функции $\alpha \to g_{1,n}(\alpha)$ на компакте $\left[0, \frac{k_n}{n}\left(1-\frac{k_n}{n}\right)\right]$.

Поскольку найти значения α_0 и M_n для любого фиксированного n в явном виде не представляется возможным, исследуем их асимптотические свойства для $n \to \infty$. Пусть $k,\ n \to \infty$ и $k = tn + o(n)$ для некоторого $t \in [0, 1]$; тогда предельные значения для (2) и (3) имеют соответственно вид

$$f_1(t) := \lim_{n,k\to\infty} f_1(n,k) = (1-e^{-1})t \tag{6}$$

и

$$f_2(t) := \lim_{n,k\to\infty} f_2(n,k) = -(1-t)t\ln t. \tag{7}$$

Предельное верхнее значение для цены подсказки составляет $\lim_{n\to\infty} \frac{k_n}{n}\left(1-\frac{k_n}{n}\right) = = e^{-1}(1-e^{-1}) \approx 0{,}233$, а функции $\lim_{n\to\infty} \underline{k}_n(\alpha) =: \underline{X}(\alpha)$ и $\lim_{n\to\infty} \bar{k}_n(\alpha) =: \bar{X}(\alpha)$, где $\underline{X}(\alpha) = \frac{\alpha}{1-e^{-1}}$ и $\bar{X}(\alpha)$ — корень уравнения $-(1-t)t \cdot \ln(t) = \alpha$ из промежутка $[e^{-1}, 1]$, определены на отрезке $[0, e^{-1}(1-e^{-1})]$. При этом средний выигрыш первого игрока составит $g_1(\alpha)=\alpha\left(1-\dfrac{\underline{X}(\alpha)}{\overline{X}(\alpha)}\right)$. В силу соотношения $-(1-\overline{X}(\alpha))\overline{X}(\alpha)\ln\overline{X}(\alpha)=\alpha$ функцию $\alpha\to g_1(\alpha)$ можно представить в виде

$$g_1(\alpha)=-(1-\overline{X}(\alpha))\overline{X}(\alpha)\ln\overline{X}(\alpha)-(1-e^{-1})(1-\overline{X}(\alpha))^2\,\overline{X}(\alpha)\ln^2\overline{X}(\alpha).$$

Заметим, что функция $t\to-(1-t)t\ln t-(1-e^{-1})(1-t)^2 t\ln^2 t$ унимодальная на интервале $[e^{-1},1]$ (численное решение показывает, что максимум достигается в точке $t_0\approx 0{,}566$), а функция $\alpha\to\overline{X}(\alpha)$ монотонно убывает на интервале $[0,e^{-1}(1-e^{-1})]$. Следовательно, функция $\alpha\to g_1(\alpha)$ унимодальная и достигает максимума в точке $\alpha_0\in[0,e^{-1}(1-e^{-1})]$ такой, что $\overline{X}(\alpha_0)=t_0$. Следовательно, $\alpha_0=-t_0(1-t_0)\ln t_0\approx 0{,}14$ и $g_1(\alpha_0)\approx 0{,}085$.

Средний суммарный выигрыш первого и второго игроков составит

$$g_{1,2}(\alpha_0)=\int\limits_{\underline{X}(\alpha_0)}^{1/e}\frac{\underline{X}(\alpha_0)}{t^2}(f_1(t)+t)dt+$$

$$+\int\limits_{1/e}^{\overline{X}(\alpha_0)}\frac{\underline{X}(\alpha_0)}{t^2}(f_2(t)+t)dt+\int\limits_{\overline{X}(\alpha_0)}^{1}\frac{\underline{X}(\alpha_0)}{t^2}t\,dt\approx 0{,}480\,.$$

Вычитая из этой величины среднюю выплату первому игроку за подсказку, получаем

$$g_2(\alpha_0)=g_{1,2}(\alpha_0)-g_1(\alpha_0)\approx 0.480-0.085\approx 0{,}395.$$

Отметим, что средний выигрыш в классической задаче оптимального выбора (т.е. без использования подсказки) составляет $e^{-1}\approx 0{,}368$.

Найдем цену анархии в данной игре. Так, в работе [3] она определяется, как отношение суммарного выигрыша в случае корпоративного поведения к суммарному выигрышу в ситуации равновесия по Нэшу (если равновесий по Нэшу несколько, то берется минимальный из суммарных выигрышей). Пусть оба игрока действуют сообща так, чтобы максимизировать суммарный выигрыш. Тогда данная игровая ситуация сводится к задаче оптимального выбора с возможностью двух остановок. Такая задача рассмотрена в [5]. Было найдено, что первую остановку следует делать на первом максимальном элементе после элемента $e^{-3/2}n$, а вторую — после элемента $e^{-1}n$. При этом вероятность нахождения наилучшего элемента равна $e^{-3/2}+(1-e^{-3/2})e^{-1}\approx 0{,}509$. Таким образом, цена анархии составляет $0{,}509/0{,}480\approx 1{,}06$.

ИГРА В ВЫМОГАТЕЛЬСТВО

Аналогично рассмотренной игре в подсказку в игре в вымогательство второй игрок может получить единичный выигрыш из внешнего источника, найдя наилучший элемент, а первый игрок рассчитывает получить некоторую сумму от второго игрока. Однако в отличие от предыдущей игры первый игрок (назовем его вымогателем) не оказывает помощи второму игроку в поиске наилучшего элемента, а наоборот, угрожает созданием помех в этом поиске. Чтобы эта угроза не была реализована, второй игрок должен выплатить некоторую компенсацию.

До начала игры вымогатель назначает фиксированную плату, которую он требует от первого игрока; впоследствии она не может изменяться. Пусть на всем протяжении просмотра вымогатель может один раз заблокировать максимальный элемент. Тогда у второго игрока возникает альтернатива: выплатить вымогателю требуемую установленную сумму и остановиться на заблокированном элементе либо ничего не платить и продолжить просмотр. В любом случае после блокирования полномочия вымогателя исчерпываются.

Найдем оптимальные стратегии вымогателя и просматривающего. Очевидно, что вымогатель не должен закрывать элементы с индексами, меньшими k_n. Иначе его полномочия на этом исчерпаются, а просматривающий ничего не выплатит и будет придерживаться стратегии «остановиться на первом максимальном элементе, начиная с k_n». Его ожидаемый выигрыш при этом составит (как в классической задаче) $1/e+o(1)$. Вымогатель должен заблокировать первый максимальный элемент с индексом, большим k_n, иначе на этом элементе остановится просматривающий (он имеет такие намерения согласно оптимальной стратегии классической задачи). Таким образом, план действий вымогателя найден, осталось найти величину платы (обозначим ее α).

Пусть вымогатель заблокировал некоторый элемент $k \ge k_n$. Тогда если просматривающий потребует разблокировать этот элемент и остановится на нем, то его выигрыш составит $\varphi_{1,n}(k,\alpha):=\frac{k}{n}-\alpha$, а если он продолжит просмотр, то выигрыш составит $\varphi_{2,n}(k):=\frac{k}{n}\sum_{j=k}^{n-1}\frac{1}{j}$. Поскольку второй игрок максимизирует собственный выигрыш, то условием платы вымогателю суммы α является условие $\varphi_{1,n}(k,\alpha)\ge\varphi_{2,n}(k)$; отсюда

$$\alpha \le \frac{k}{n}\left(1-\sum_{j=k}^{n-1}\frac{1}{j}\right). \tag{8}$$

Поскольку правая часть (8) при фиксированном n монотонно возрастает по k (при $k\ge k_n$) от $\frac{k_n}{n}\left(1-\sum_{j=k_n}^{n-1}\frac{1}{j}\right)$ до единицы, то при любом фиксированном $\alpha\in\left[\frac{k_n}{n}\left(1-\sum_{j=k_n}^{n-1}\frac{1}{j}\right);1\right]$ неравенство (8) будет справедливо при $k\in\{t_n(\alpha),\ldots,n\}$, где $t_n(\alpha):=\inf\left(k\ge k_n : \alpha\le\frac{k}{n}\left(1-\sum_{j=k}^{n-1}\frac{1}{j}\right)\right)$. Кроме того, $t_n(\alpha)$ монотонно растет с ростом α и, таким образом, множество $\{t_n(\alpha),\ldots,n\}$ сужается.

Предположим, что просматривающий не знает о существовании вымогателя до того, как тот начнет действовать. Тогда он заплатит выкуп, если первый максимальный элемент с индексом, не меньшим k_n, появится в множестве $\{t_n(\alpha),\ldots,n\}$. Это событие (обозначим его вероятность $p_n(\alpha)$) является пересечением следующих двух событий:

1) в множестве объектов с номерами $\{k_n,\ldots,t_n(\alpha)-1\}$ не появилось ни одного максимального (вероятность такого события составляет $\frac{k_n-1}{t_n(\alpha)-1}$);

2) в множестве объектов с номерами $\{t_n(\alpha),\ldots,n\}$ появился хотя бы один максимальный элемент (вероятность этого события составляет $1-\frac{t_n(\alpha)-1}{n}$).

Так как эти события являются независимыми, то

$$p_n(\alpha)=\frac{k_n-1}{t_n(\alpha)-1}\left(1-\frac{t_n(\alpha)-1}{n}\right)=\frac{k_n-1}{t_n(\alpha)-1}-\frac{k_n-1}{n}; \qquad (9)$$

таким образом, ожидаемый выигрыш вымогателя составляет

$$g_{1,n}(\alpha)=\alpha\cdot p_n(\alpha)=\alpha(k_n-1)\left(\frac{1}{t_n(\alpha)-1}-\frac{1}{n}\right). \qquad (10)$$

При этом средний выигрыш просматривающего игрока составляет

$$g_{2,n}(\alpha)=\sum_{i=k_n}^{t_n(\alpha)-1}\frac{k_n-1}{i(i-1)}\frac{i}{n}\sum_{j=i}^{n-1}\frac{1}{j}+\sum_{i=t_n(\alpha)}^{n}\frac{k_n-1}{i(i-1)}\left(\frac{i}{n}-\alpha\right)=$$

$$=\frac{k_n-1}{n}\sum_{i=k_n-1}^{t_n(\alpha)-2}\frac{1}{i}\sum_{j=i}^{n-1}\frac{1}{j}+\sum_{i=t_n(\alpha)}^{n}\frac{k_n-1}{i(i-1)}\left(\frac{i}{n}-\alpha\right)=$$

$$=\frac{k_n-1}{n}\sum_{i=k_n-1}^{t_n(\alpha)-2}\frac{1}{i}\sum_{j=i}^{n-1}\frac{1}{j}+\frac{k_n-1}{n}\sum_{i=t_n(\alpha)-1}^{n-1}\frac{1}{i}-\alpha(k_n-1)\left(\frac{1}{t_n(\alpha)-1}-\frac{1}{n}\right).$$

Сумма первых двух слагаемых в этой формуле — это величина выигрыша обоих игроков:

$$g_{1,2,n}(\alpha)=\frac{k_n-1}{n}\sum_{i=k_n-1}^{t_n(\alpha)-2}\frac{1}{i}\sum_{j=i}^{n-1}\frac{1}{j}+\frac{k_n-1}{n}\sum_{i=t_n(\alpha)-1}^{n-1}\frac{1}{i}. \qquad (11)$$

Исследуем теперь асимптотические свойства найденных величин, когда количество просматриваемых объектов стремится к бесконечности. Пусть, как и ранее, $k,\ n\to\infty$ и $k=tn+o(n)$ для некоторого $t\in[0,1]$. Тогда

$$\varphi_1(t,\alpha):=\lim_{n,k\to\infty}\varphi_{1,n}(k,\alpha)=t-\alpha,\quad \varphi_2(t,\alpha):=\lim_{n,k\to\infty}\varphi(k,t)=-t\ln t.$$

Далее, $\lim\limits_{n\to\infty}\frac{t_n(\alpha)}{n}=:t(\alpha),\ \alpha\in[0,1]$, где $t(\alpha)$ — корень уравнения $t(1+\ln t)=\alpha$ из интервала $[e^{-1},1]$. (Функцию $t(\alpha)$ можно представить в терминах функции Ламберта $W(\cdot)$ таким образом: $t(\alpha)=e^{W(\alpha e)-1}$, где W — функция, обратная к $x\to xe^x$ на полуоси $x\ge e^{-1}$.) Кроме того, $p(\alpha):=\lim\limits_{n\to\infty}p_n(\alpha)=\frac{1-t(\alpha)}{et(\alpha)}$; $g_1(\alpha):=$ $:=\lim\limits_{n\to\infty}g_{1,n}(\alpha)=\alpha\frac{1-t(\alpha)}{et(\alpha)}$;

$$g_{1,2,n}(\alpha):=\lim_{n\to\infty}g_{1,2,n}(\alpha)=\frac{1}{e}\int_{1/e}^{t(\alpha)}\frac{1}{u}\int_{u}^{1}\frac{1}{v}\,dvdu+\frac{1}{e}\int_{t(\alpha)}^{1}\frac{1}{u}\,du-\frac{\alpha}{e}\left(\frac{1-t(\alpha)}{t(\alpha)}\right)=$$

$$=\frac{1}{2e}(1-\ln^2 t(\alpha))-\frac{1}{e}\ln t(\alpha)-\frac{\alpha(1-t(\alpha))}{et(\alpha)};$$

$$g_{1,2,n}(\alpha):=\lim_{n\to\infty}g_{1,2,n}(\alpha)=\frac{1}{2e}(1-\ln^2 t(\alpha))-\frac{1}{e}\ln t(\alpha).$$

Рассмотрим также вспомогательную функцию $d(\alpha) = \frac{1}{e} - g_2(\alpha)$, описывающую ущерб, причиняемый просматривающему вымогателем.

Используя соотношение $t(\alpha)(1 + \ln t(\alpha)) = \alpha$, получаем $g_1(\alpha) = \frac{1}{e}(1 - t(\alpha))(1 + \ln t(\alpha))$. Функция $x \to (1 - x)(1 + \ln x)$ вогнута на промежутке $[e^{-1}, 1]$ и принимает единственное максимальное значение, приблизительно равное 0,073, в точке $x_0 \approx 0{,}644$. Функция $\alpha \to t(\alpha)$ возрастает и принимает значения на интервале $[e^{-1}, 1]$, поэтому функция $g_1(\alpha)$ унимодальная и принимает максимальное значение, также приблизительно равное 0,073, в точке $\alpha_0 := x_0(1 + \ln x_0) \approx 0{,}361$. Аналогичным образом можно констатировать, что функции $g_2(\alpha)$ и $g_{1,2,n}(\alpha)$ монотонно убывают, а функция $d_2(\alpha)$ соответственно возрастает, при этом $g_2(\alpha_0) \approx 0{,}237$, $g_{1,2,n}(\alpha_0) \approx 0{,}310$, $d_2(\alpha_0) \approx 0{,}131$, $g_{1,2,n}(\alpha) = 0{,}310$, $d_2(\alpha) = 0{,}131$. Для наглядности в табл. 1 представлена табуляция рассмотренных величин по α в диапазоне от 0 до единицы с шагом 0,1.

Таблица 1

Функция от α	Значения аргумента α в диапазоне										
	0	0,1	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,9	1
$t(\alpha)$	0,368	0,458	0,535	0,604	0,669	0,730	0,788	0,844	0,897	0,949	1
$g_1(\alpha)$	0	0,044	0,064	0,072	0,073	0,068	0,059	0,048	0,034	0,018	0
$g_2(\alpha)$	0,368	0,316	0,278	0,250	0,229	0,213	0,202	0,193	0,188	0,185	0,184
$g_{1,2,n}(\alpha)$	0,368	0,359	0,342	0,323	0,302	0,282	0,261	0,241	0,222	0,203	0,184
$d_2(\alpha)$	0	0,052	0,09	0,118	0,139	0,154	0,166	0,174	0,18	0,183	0,184

ИГРА В УГАДЫВАНИЕ

Пусть два игрока одновременно просматривают элементы и в процессе просмотра первый игрок должен второму игроку задать вопрос: является ли текущий элемент наилучшим? Если второй игрок неправильно отвечает на этот вопрос, то он выплачивает единичный выигрыш первому игроку, в противном случае он ничего не платит. Момент формулирования данного вопроса принадлежит первому игроку. Если первый игрок не задаст свой вопрос на протяжении всего просмотра элементов, то он ничего не получает.

Вначале найдем оптимальную стратегию второго игрока — она оказывается достаточно простой. Если вопрос задается об элементе, который не является максимальным, то следует отрицательный ответ, поскольку этот элемент заведомо не является наилучшим. Если k-й по счету элемент является максимальным, то он является наилучшим с вероятностью k/n. Таким образом, чтобы максимизировать вероятность угадывания, второму игроку следует давать отрицательный ответ, если $k < n/2$, и положительный ответ, если $k > (n/2)$. Если k — четное число и $k = n/2$, то независимо от ответа вероятность угадывания будет равна $1/2$.

Рассмотрим задачу остановки марковского процесса на множестве максимальных элементов. Сформулированная ниже лемма использовалась при анализе игры в угадывание.

Лемма 1. Пусть имеется n элементов, расположенных в случайном порядке, причем все $n!$ перестановок равновероятны. Пусть проводится просмотр элементов и просматривающему разрешено остановиться и получить выигрыш $f(i) = \min\left\{\frac{i}{n}, 1 - \frac{i}{n}\right\}$, только если i-й элемент является максимальным среди уже просмотренных. Тогда оптимальная стратегия, максимизирующая ожидаемый выигрыш, имеет пороговый вид — пропустить все элементы от первого до $(l_n - 1)$-го и остановиться на первом максимальном элементе с номером $j \ge l_n$, где

$$l_n := \inf\left\{k \in N : H_{k-1} \ge 2H_{[n/2]-1} - H_{n-1} + \frac{n}{[n/2]} - 2\right\}, \tag{12}$$

$H_n := \sum_{k=1}^{n} \frac{1}{k}$ — n-е гармоническое число.

Доказательство. Согласно результатам работы [1, §5, гл. III] оптимальная стратегия имеет такой вид: остановиться на первом максимальном элементе, индекс которого лежит в опорном множестве $\Gamma_n := \{1 \le k \le n : f(k) = v(k)\}$, где функция v определяется рекурсивно равенством

$$v(k) := \max\left\{f(k),\ k\sum_{j=k+1}^{n} \frac{1}{j(j-1)} v(j)\right\}$$

и является эксцессивной мажорантой функции выигрыша $f(\cdot)$. Проверкой убедимся, что для функции $f(i) = \min\left\{\frac{i}{n},\ 1 - \frac{i}{n}\right\}$ опорное множество имеет вид $\Gamma_n := \{l_n, l_n + 1, \ldots, n\}$. Так как функция $f(\cdot)$ убывает на множестве $i > [n/2]$, то $v(k) = f(k)$ при $k > [n/2]$. Пусть $k \le [n/2]$, тогда

$$v(k) = \max\left\{\frac{k}{n},\ k\sum_{j=k+1}^{[n/2]} \frac{1}{j(j-1)} v(j) + k\sum_{j=[n/2]+1}^{n} \frac{f(j)}{j(j-1)}\right\} =$$

$$= \max\left\{\frac{k}{n},\ k\sum_{j=k+1}^{[n/2]} \frac{1}{j(j-1)} v(j) + k\sum_{j=[n/2]+1}^{n} \frac{1 - j/n}{j(j-1)}\right\} =$$

$$= k\max\left\{\frac{1}{n},\ \sum_{j=k+1}^{[n/2]} \frac{1}{j(j-1)} v(j) + \left(\frac{1}{[n/2]} - \frac{1}{n}\right) - \frac{1}{n}(H_{n-1} - H_{[n/2]-1})\right\}.$$

Так как последовательность $\left(\sum_{j=k+1}^{[n/2]} \frac{1}{j(j-1)} v(j)\right)_{k \le [n/2]}$ убывает по k, то $v(k) = f(k)$ для тех и только тех $k \le [n/2]$, для которых выполняется неравенство $\sum_{j=k+1}^{[n/2]} \frac{1}{j(j-1)} f(j) \le \frac{1}{n}(H_{n-1} - H_{[n/2]-1}) + \left(\frac{1}{[n/2]} - \frac{1}{n}\right) - \frac{1}{n}$. Подставим выражение для $f(j)$; тогда $v(k) = f(k)$ для тех и только тех $k \le [n/2]$, для которых $H_{k-1} \ge 2H_{[n/2]-1} - H_{n-1} - \frac{n}{[n/2]} + 2$.

Доказательство завершено.

Далее будем находить оптимальную стратегию первого игрока. При оптимальном поведении второго игрока первому игроку следует задавать вопрос только о максимальных элементах, причем его функция выигрыша в зависимости от индекса элемента имеет вид

$$f(k) = \min\left(\frac{k}{n},\ 1 - \frac{k}{n}\right) = \begin{cases} \frac{k}{n}, & k \le \frac{n}{2}, \\ 1 - \frac{k}{n}, & k > \frac{n}{2}. \end{cases}$$

Из леммы следует, что оптимальная стратегия имеет пороговый вид — пропустить все элементы от первого до $(l_n - 1)$-го и остановиться на первом максимальном элементе с номером $j \geq l_n$, где l_n определяется условием (12). Как и ранее, рассмотрим предельный случай, когда $k = nt + o(n)$ при $n \to \infty$. Предельной функцией для $f(k)$ будет функция $\varphi(t) = \max(t, 1-t)$. Из соотношения (12) и асимптотического соотношения $\sum_{k=1}^{n} \frac{1}{k} = \log n + \gamma + o(1)$, $n \to \infty$, имеем $\lim_{n\to\infty} \frac{l_n}{n} = \frac{1}{4}$.

Таким образом, асимптотически оптимальная стратегия первого игрока — пропустить четвертую часть элементов и задать вопрос второму игроку относительно первого максимального элемента, который следует после этого. Отметим, что при этом средний выигрыш первого игрока также будет равен $1/4$ части.

ЗАКЛЮЧЕНИЕ

Рассмотренные три типа игры: игра в подсказку, вымогательство и угадывание, а также изложенная в работе [8] игра в цензуру в задаче оптимального выбора, абсолютно различны по целям игроков и вытекающим из них оптимальным стратегиям. Однако их объединяет следующее.

— Все рассмотренные игры представляются играми двух игроков, которые являются надстройкой над задачей оптимального выбора.

— Все рассмотренные игры являются партизанскими (игроки имеют совершенно различные цели и наборы допустимых стратегий).

— Во всех случаях существует единственное равновесие по Нешу, и асимптотическое поведение точек равновесия найдены в явном виде.

Исходя из этого можно провести стохастическое моделирование и сравнить результаты моделирования с полученными в статье теоретическими значениями.

СПИСОК ЛИТЕРАТУРЫ

1. Д ы н к и н Е. Б., Ю ш к е в и ч А. А. Теоремы и задачи о процессах Маркова. —М.: Наука. — 232 с.
2. Г у с е й н - З а д е С. М. Задача выбора и оптимальное правило остановки последовательности независимых испытаний // ТВП. — 1966. — **11**, № 3. — С. 534–537.
3. М а з а л о в В. В. Математическая теория игр и приложения. — СПБ.: Лань, 2010. — 446 с.
4. М а з а л о в В. В. Игровые моменты остановки. — Новосибирск: Наука, 1987. — 191 с.
5. Д о ц е н к о С. Задача выбора наилучшего объекта как игра двух лиц // Кибернетика и вычислительная техника. — 2011. — Вып. 164. — С. 43–53.
6. Д о ц е н к о С. Игровые ситуации в модифицированной задаче оптимального выбора // Там же. — 2012. — Вып. 168. — С. 3–8.
7. M e l i s s a d e C a r v a l h o, L u c a s M. C h a v e s, R i c a r d o M a r t i n s. Variations of secretary problem via game theory and linear programming // J. Comput. Sci. — 2008. — **7**, N 3. — P. 78–82.
8. Д о ц е н к о С., М а р и н и ч А. Игра в цензуру в задаче оптимального выбора // Кибернетика и системный анализ. — 2013. — № 5. — С. 124–131.
9. C h i k r i i A. Conflict-Controlled Processes. — Kluwer Academic Publishers, April 1997. — 424 p.

Поступила 19.04.2013