

УДК 621.391.1

© 2003 г. П.В. Голубцов, В.А. Любецкий

## СТОХАСТИЧЕСКИЕ ДИНАМИЧЕСКИЕ ИГРЫ С ИНФОРМАЦИЕЙ РАЗЛИЧНОГО ТИПА

Рассматриваются динамические игры с дискретным временем и стохастической средой в ситуациях с различными информационными структурами игры (и равновесием по Нэшу). Оказывается, что информационная структура игры, т.е. характер и объем доступной игрокам информации, а особенно асимметрия в поступающей к ним информации, оказывают неожиданное влияние на результаты игры – средние выигрыши игроков. Также предлагаются новые алгоритмы для нахождения равновесия по Нэшу в таких играх. Приводятся подробные и систематизированные в виде графических зависимостей результаты компьютерного счета.

### § 1. Введение

В работе рассматривается некоторый класс стохастических динамических (с дискретным временем) игр с явным учетом информации, доступной в этих играх двум (или нескольким) игрокам. Рассматриваются различные варианты информационной структуры игры (иными словами, игры-поведения) с целью изучить *роль информации, доступной игрокам, при построении ими оптимальной стратегии игры* (иными словами, с целью изучить эффекты *контекстного использования информации в процессе поведения* игроков). Таким образом, речь идет о семантической стороне информации (о *семантической информации*).

Два независимых игрока осуществляют управление общей динамической системой с дискретным временем, в которой имеются случайные возмущающие факторы. Управляющие решения принимаются на каждом временном шаге в условиях большей или меньшей информации об этих факторах. Каждый игрок выбирает свою стратегию управления так, чтобы максимизировать свой суммарный пошаговый выигрыш за достаточно протяженный отрезок времени при условии, что его соперник придерживается фиксированной стратегии. Таким образом, ход игры описывается динамическим равновесием по Нэшу.

Упомянутые факторы задаются случайным марковским параметром (который может быть многомерным).

Рассматриваются игры с информационными структурами [1], при которых игроки обладают различными уровнями знаний о реализациях случайного параметра: например, обладают текущей информацией, информацией с задержкой [2, 3] или информацией, полученной из неточных измерений случайного параметра; информация может быть асимметричной, когда, например, один игрок имеет текущую информацию, а другой получает ее с запаздыванием или неточно; также рассматриваются кооперативные игры с различными типами информационных структур и т.д.

Для каждого варианта игры мы предлагаем алгоритм с целью ее сведения к задаче многократного нахождения равновесия по Нэшу (с непрерывными и дискретными параметрами игры) в  $n$ -мерном пространстве. А в пункте 5.3 предлагаем алгоритм,

который сводит рассматриваемую там игру к многоэтапному поиску точки равновесия по Нэшу уже в функциональном пространстве, что в свою очередь приводит к пошаговой процедуре оптимизации (на каждом этапе).

В приложении и в пунктах 4.1 и 5.3 доказываются простые утверждения, цель которых скорее в том, чтобы пояснить используемые понятия. Вместо серьезных доказательств мы провели обширный счет по нашим алгоритмам, который показал их быструю сходимость к естественным результатам для большого числа исходных данных.

Затем описываются результаты компьютерного моделирования игр, в которых доступная игрокам информация оказывает нетривиальное влияние на выбор ими оптимальной стратегии игры (на поведение игроков). Авторы думают, что наблюдение таких эффектов о влиянии доступной игрокам информации на ход игры (т.е. на результат использования информации в контексте игры или иной “деятельности организма”) проясняет само понятие “семантической информации”.

В качестве простого примера динамической игры и для удобной интерпретации результатов в работе рассматривается хорошо известная задача морского рыболовства [4, 5].

Авторам кажутся содержательными и в какой-то мере неожиданными следующие наблюдения о влиянии информации на поведение игроков.

При игре с симметричной текущей информацией (“полная информация”) снижение себестоимости лова приводит к снижению доходов игроков (рис. 2, область около точки 0,3 на сплошном графике). При игре с неполной информацией наблюдается аналогичная ситуация, но существует область, в которой эта неполная информация приводит к более высоким их доходам по сравнению с полной информацией. При кооперативном поведении игроков (в частности, их симметричном доступе к информации) получаются наивно ожидаемые результаты игры.

При повышении точности доступной игрокам информации во всех вариантах их доступа к ней (т.е. при увеличении точности измерений случайного параметра), начиная с некоторого значения точности, доходы обоих игроков снижаются. При кооперативном поведении игроков результаты игры снова получаются наивно ожидаемые (см. рис. 3 и 4).

На рис. 5 видно, что при разных значениях природной изменчивости результаты игры зависят от информационной структуры игры; например, иногда минимальная информация оказывается более выгодной для игроков, чем полная.

На рис. 6 видно, что дополнительная информация выгодна игрокам только в том случае, если уровень их сотрудничества достаточно высокий.

Отмеченные особенности, в сущности, контекстного (игрового) влияния информации не кажутся авторам очевидными из каких-либо общих соображений и представляются им интересными для возможной формализации (не колмогоровского или шенноновского) понятия информации – *семантической информации*.

В заключение скажем об интерпретации рассматриваемых здесь задач, которую авторы имели в виду при их разработке. На интуитивном уровне имеются представления о своеобразном противоборстве и сотрудничестве двух сил, о которых говорят словами: организм (или: клетка, популяция) и среда, индивидуум и общество, человек и судьба, и т.д. Клетка (скажем, бактерии) противостоит среде и сотрудничает с ней на основе генома – своеобразной записи большого числа инструкций по биохимическому производству и по транспорту из окружающей среды специфических химических веществ (и еще инструкций по запуску упомянутых выше инструкций, заключенных в регуляторных областях генома, и т.д.). В геноме перемешаны и в некотором смысле неотличимы друг от друга такие два типа инструкций. Сложные комплексы инструкций одновременно включаются и выключаются в зависимости от текущей ситуации (в клетке и вокруг нее). Цель со стороны клетки состоит в

сохранении ряда ее параметров; цель со стороны среды – в сохранении ряда других параметров в интересах уже популяции (ограничительных для клетки и частично антагонистических ее интересам). Важно, что обе стороны – клетка и среда – в каждый момент их действий (принятия решений) обладают неполной, частично общей и частично раздельной (иногда противоречивой) информацией. Вопрос о характере их сотрудничества и, в частности, (оптимальной) стратегии обмена между ними информацией имеет важное эволюционное значение.

Например, среда выделяет сигнал о количестве других клеток (того же и/или иных видов) в окрестности данной клетки (феромон), в результате клетка может перестать делиться, изменить вирулентность, продуцировать антибиотик (чтобы бороться с конкурентами в среде), участвовать в образовании биопленки и т.п.

Другой пример такого рода сложных “игровых” взаимодействий – это механизм споруляции, когда клетки популяции образуют сложную форму (головку и ножку): клетки, вошедшие в головку, превращаются в споры и улетают благодаря тому, что находятся на высоте, а остальные погибают. Можно упомянуть симбиотические бактерии, живущие в светящемся органе каракатицы и взаимодействующие с ней; морские водоросли, колонизированные бактериями, которые выделяют аналоги бактериальных феромонов, чтобы разрушить популяцию бактерий.

Здесь возможны нетривиальные стратегии взаимодействия, относящиеся к различным его уровням: отдельная клетка и ее окружение, внутри клона (среди потомков одной клетки), внутри сообщества клонов и т.д. Результат взаимодействия бывает неожиданным: например, внутри популяции-клона одна успешная клетка может обмануть все другие, размножиться, уничтожив сообщество, и в результате клон погибает. Взаимодействие между клонами может быть уже симбиотическим или антагонистическим. Предпочтение может отдаваться судьбе отдельной особи или общему благу.

Кажется, что обычный математический аппарат плохо описывает такого рода системы (про которые к настоящему времени известно как раз много конкретных естественнонаучных сведений). Можно представить себе пространство состояний “клетка (популяция)–среда”, в котором происходят переходы типа  $R \rightarrow R_+$  и по текущему состоянию  $R$  вычисляется его оценка  $v$ . Та или иная область в этом пространстве доступна в зависимости от полученной пары  $S = \langle S^\alpha, S^\beta \rangle$ , где  $S^\alpha$  – результат реакции клетки на  $R^\alpha$ , а  $S^\beta$  – результат деятельности среды в присутствии текущего  $R^\beta$ , при этом  $R^\alpha$  и  $R^\beta$  – результаты расщепления текущего состояния  $R$  клетки и среды. Авторы предполагают обсудить интерпретации этого типа для системы клетка–среда в следующей публикации.

Структура работы такова: §§ 2 и 3 содержат описание стохастической игры двух лиц с дискретным временем и общую постановку задачи, в §§ 4 и 5 содержатся уточнения рассматриваемых задач и описания алгоритмов для построения стратегий, образующих равновесие по Нэшу. Эти алгоритмы состоят в сведении исходной задачи к процедуре динамического программирования, обобщающей классическую процедуру теории оптимального управления (см., например, [6]). В § 6 рассматривается метод введения в игру частичной или полной кооперации игроков. В § 7 приводятся результаты компьютерного моделирования. Наконец, в § 8 представлены общие сведения об играх, используемые в основной части статьи.

## § 2. Описание динамической игры

В данном параграфе мы приведем довольно общее описание неантагонистической динамической игры двух лиц с дискретным временем

$$R \xrightarrow{\rho(R, \nu, p^\alpha, p^\beta)} R_+.$$

Здесь  $R$  и  $R_+$  – соответственно, текущее и следующее состояние системы;  $\nu$  – случайный параметр, зависящий от шага (времени);  $p^\alpha$  и  $p^\beta$  – решения (действия) игроков. Мы считаем, что состояние системы на следующем шаге  $R_+$  и текущие выигрыши  $v^\alpha$  и  $v^\beta$  игроков зависят от текущего состояния  $R$ , от решений  $p^\alpha$ ,  $p^\beta$  и реализации случайного параметра  $\nu$  на данном шаге игры, т.е.

$$R_+ = \rho(R, \nu, p^\alpha, p^\beta), \\ v^\alpha = v^\alpha(R, \nu, p^\alpha, p^\beta), \quad v^\beta = v^\beta(R, \nu, p^\alpha, p^\beta).$$

Сами решения игроков определяются текущим состоянием системы  $R$  и, возможно, некоторой дополнительной информацией  $\xi^\alpha$  и  $\xi^\beta$  о случайном возмущении  $\nu$ :

$$p^\alpha = P^\alpha(R, \xi^\alpha), \quad p^\beta = P^\beta(R, \xi^\beta),$$

где  $P^\alpha$  и  $P^\beta$  – стратегии управления на данном шаге, зависящие от информации, имеющейся у игрока  $\alpha$  или  $\beta$ .

В дальнейшем мы будем полагать, что случайные элементы  $\nu_t$ , где  $t$  – время, независимы и одинаково распределены или, в более общем случае, образуют цепь Маркова с известным переходным распределением. В частности, если  $\nu_t$  принимает конечное число  $n$  различных значений, то его стохастическое поведение полностью определяется соответствующей стохастической матрицей размера  $n \times n$ .

### § 3. Игра с бесконечным горизонтом

Вкратце опишем постановку задачи выбора оптимальных стратегий (управления) для неантагонистической динамической игры. Точные постановки и технические детали сведения этой задачи выбора оптимальных стратегий к соответствующим процедурам динамического программирования рассмотрены в §§ 4 и 5.

Ради простоты рассмотрим в этом параграфе игру с бесконечным горизонтом. Обычно в играх с бесконечным горизонтом считается, что ситуация стационарна и, в частности, отображения  $\rho$ ,  $v^\alpha$  и  $v^\beta$  не зависят от времени  $t$ . Мы также принимаем это в текущем параграфе.

Задача оптимального управления (в которых для двух игроков) состоит в построении оптимальных стационарных стратегий управления  $P^\alpha$  и  $P^\beta$  таких, что средние суммарные пошаговые выигрыши (разные слагаемые могут иметь коэффициенты, отражающие, например, то, что значимость “дохода” убывает для игрока с годами)

$$U^\alpha(R_0, P^\alpha, P^\beta) = E \sum_{t=0}^{\infty} \gamma_\alpha^t v^\alpha(R_t, p_t^\alpha, \nu_t) = E \sum_{t=0}^{\infty} \gamma_\alpha^t v^\alpha(R_t, P^\alpha(R_t, \xi_t^\alpha), \nu_t)$$

и аналогичное выражение для игрока  $\beta$  были бы максимальными. Здесь математическое ожидание берется по случайным элементам  $\nu_t$  и (в случаях информации, получаемой из измерений) по случайным измерениям  $\xi_t^\alpha$  и  $\xi_t^\beta$ .

Таким образом, мы получаем игру двух лиц, состоящую в нахождении равновесия по Нэшу

$$\begin{cases} \max_{P^\alpha} U^\alpha(R, P^\alpha, P^\beta), \\ \max_{P^\beta} U^\beta(R, P^\alpha, P^\beta) \end{cases}$$

для всех возможных значений  $R$ . Утверждение “пара стратегий  $\hat{P}^\alpha, \hat{P}^\beta$  является равновесной по Нэшу” означает по определению, что для всех возможных значений  $R$

и всех отображений  $P^\alpha, P^\beta$  выполняется

$$\begin{cases} U^\alpha(R, \hat{P}^\alpha, \hat{P}^\beta) \geq U^\alpha(R, P^\alpha, \hat{P}^\beta), \\ U^\beta(R, \hat{P}^\alpha, \hat{P}^\beta) \geq U^\beta(R, \hat{P}^\alpha, P^\beta). \end{cases}$$

Оптимальные стратегии  $\hat{P}^\alpha, \hat{P}^\beta$  (а отсюда и отвечающие им оптимальные средние выигрыши  $\hat{U}^\alpha, \hat{U}^\beta$ ) обычно строят переходом к пределу  $T \rightarrow \infty$  в соответствующих играх с конечным горизонтом  $T$ . Алгоритмы решения этой задачи (и даже в нестационарном случае) для игры с конечным горизонтом  $T$  излагаются в §§ 4 и 5.

#### § 4. Игра с конечным горизонтом: игроки имеют результаты точных наблюдений

Рассмотрим игру с конечным горизонтом, которая по определению заканчивается в момент времени  $T$ . Поскольку различные моменты времени неравнозначны, приходится рассматривать нестационарную ситуацию. Мы разрешаем функциям  $v_\tau^\alpha$  и  $v_\tau^\beta$  отличаться в различные моменты времени  $\tau$ . Обозначим управляющую стратегию игрока  $\alpha$  в момент  $\tau$  через  $P_\tau^\alpha$ , а последовательность отображений  $P_\tau^\alpha$ , начиная с момента  $t$  до момента  $T$ , через

$$P_t^\alpha = \langle P_t^\alpha, P_{t+1}^\alpha, \dots, P_T^\alpha \rangle = \langle P_t^\alpha, P_{t+1}^\alpha \rangle.$$

Аналогично для игрока  $\beta$ .

В этом параграфе предполагается, что случайные элементы  $\nu_t$  образуют цепь Маркова (с конечным или бесконечным числом состояний) с заданной переходной вероятностью.

Каждый игрок в момент времени  $t$  знает текущее состояние системы  $R_t$  и имеет некоторую информацию о случайном элементе  $\nu_t$  (или более раннем  $\nu_{t-1} \dots$ ). Информация о  $\nu_t$  составляет, как минимум, переходное распределение для цепи Маркова.

В этом параграфе рассматриваются два простейших случая: в момент  $t$  игрок знает реализацию случайного элемента  $\nu_{t-1}$ , а также все предыдущие значения  $\nu_{t-2}, \dots$  (т.е. знает эти значения для всех предыдущих моментов времени); или, кроме того, он знает и само текущее значение  $\nu_t$ .

В следующем параграфе рассматривается случай, когда игрокам доступна лишь неточная информация о  $\nu_t$ , получаемая в результате некоторого измерения. Структура знаний игроков может быть симметричной, когда они оба располагают одинаковой информацией, или асимметричной, когда, например, один игрок имеет текущую информацию о  $\nu_t$ , а другой знает лишь предыдущую реализацию  $\nu_{t-1}$ .

**4.1. Игроки имеют текущую информацию.** Положим, что оба игрока обладают одинаковой текущей информацией о случайном элементе  $\nu_t$  в момент времени  $t$ .

Тогда средний выигрыш игрока  $\alpha$  в момент  $t$  равен

$$V_t^\alpha(R_t, \nu_t, P_t^\alpha, P_t^\beta) = E_{(\nu_{t+1}, \nu_{t+2}, \dots, \nu_T | \nu_t)} \sum_{\tau=t}^T \gamma_\alpha^{\tau-t} v_\tau^\alpha(R_\tau, \nu_\tau, p_\tau^\alpha, p_\tau^\beta) \quad (1)$$

(аналогичное выражение для игрока  $\beta$ ), где

$$p_\tau^\alpha = P_\tau^\alpha(R_\tau, \nu_\tau), \quad p_\tau^\beta = P_\tau^\beta(R_\tau, \nu_\tau).$$

Пара  $\langle \hat{P}_t^\alpha, \hat{P}_t^\beta \rangle$  является точкой равновесия по Нэшу для пары  $\langle V_t^\alpha, V_t^\beta \rangle$ , если для всех возможных значений  $R_t$  и  $\nu_t$  выполняется

$$\begin{cases} V_t^\alpha(R_t, \nu_t, \hat{P}_t^\alpha, \hat{P}_t^\beta) = \max_{P_t^\alpha} V_t^\alpha(R_t, \nu_t, P_t^\alpha, \hat{P}_t^\beta), \\ V_t^\beta(R_t, \nu_t, \hat{P}_t^\alpha, \hat{P}_t^\beta) = \max_{P_t^\beta} V_t^\beta(R_t, \nu_t, \hat{P}_t^\alpha, P_t^\beta). \end{cases} \quad (2)$$

Будем обозначать соответствующие выигрыши

$$\begin{cases} \hat{V}_t^\alpha(R_t, \nu_t) = V_t^\alpha(R_t, \nu_t, \hat{P}_t^\alpha, \hat{P}_t^\beta), \\ \hat{V}_t^\beta(R_t, \nu_t) = V_t^\beta(R_t, \nu_t, \hat{P}_t^\alpha, \hat{P}_t^\beta). \end{cases}$$

Поскольку случайные элементы  $\nu_t$  образуют цепь Маркова, последующее значение полностью определяется текущим значением и переходным распределением (т.е. распределением  $\nu_{t+1}$  при фиксированном  $\nu_t$ ), математическое ожидание  $E_{(\nu_{t+1}, \nu_{t+2}, \dots, \nu_T | \nu_t)}$  может быть записано как последовательность условных математических ожиданий:

$$E_{(\nu_{t+1}, \nu_{t+2}, \dots, \nu_T | \nu_t)} = E_{(\nu_{t+1} | \nu_t)} E_{(\nu_{t+2} | \nu_{t+1})} \dots E_{(\nu_T | \nu_{T-1})}.$$

Отсюда следует, что выигрыш  $V_t^\alpha$  может быть выражен через текущий выигрыш  $v_t^\alpha$  и  $V_{t+1}^\alpha$ :

$$\begin{aligned} V_t^\alpha(R_t, \nu_t, P_t^\alpha, P_t^\beta) = \\ = v_t^\alpha(R_t, \nu_t, p_t^\alpha, p_t^\beta) + \gamma_\alpha E_{(\nu_{t+1} | \nu_t)} V_{t+1}^\alpha(\rho(R_t, \nu_t, p_t^\alpha, p_t^\beta), \nu_{t+1}, P_{t+1}^\alpha, P_{t+1}^\beta). \end{aligned} \quad (3)$$

Заметим, что выражение (3) можно использовать как альтернативное (рекурсивное) определение выигрыша.

Выражение (3) позволяет свести задачу (2) посредством процедуры динамического программирования к последовательности принципиально более простых задач. Предположим, что  $\langle \hat{P}_{t+1}^\alpha, \hat{P}_{t+1}^\beta \rangle$  – пара стратегий, равновесных по Нэшу, начиная с момента  $t+1$ , а  $\langle \hat{V}_{t+1}^\alpha, \hat{V}_{t+1}^\beta \rangle$  – соответствующие выигрыши.

Авторы предлагают по аналогии с динамическим программированием ввести функцию

$$\tilde{V}_t^\alpha(R_t, \nu_t, P_t^\alpha, P_t^\beta) = V_t^\alpha(R_t, \nu_t, \langle P_t^\alpha, \hat{P}_{t+1}^\alpha \rangle, \langle P_t^\beta, \hat{P}_{t+1}^\beta \rangle),$$

описывающую выигрыш игрока  $\alpha$ , который соответствует произвольным стратегиям  $P_t^\alpha$  и  $P_t^\beta$  в момент  $t$  и оптимальным “хвостовым” последовательностям  $\hat{P}_t^\alpha$  и  $\hat{P}_t^\beta$ , а также аналогичную функцию для игрока  $\beta$ .

Тогда оптимальные стратегии  $\langle \hat{P}_t^\alpha, \hat{P}_t^\beta \rangle$  в момент  $t$  могут быть получены путем нахождения точек равновесия по Нэшу при всех возможных значениях  $R_t$  и  $\nu_t$  для функций

$$\begin{cases} \tilde{V}_t^\alpha(R_t, \nu_t, p_t^\alpha, p_t^\beta) = v_t^\alpha(R_t, \nu_t, p_t^\alpha, p_t^\beta) + \gamma_\alpha E_{(\nu_{t+1} | \nu_t)} \hat{V}_{t+1}^\alpha(\rho(R_t, p_t^\alpha, p_t^\beta, \nu_t), \nu_{t+1}), \\ \tilde{V}_t^\beta(R_t, \nu_t, p_t^\alpha, p_t^\beta) = v_t^\beta(R_t, \nu_t, p_t^\alpha, p_t^\beta) + \gamma_\beta E_{(\nu_{t+1} | \nu_t)} \hat{V}_{t+1}^\beta(\rho(R_t, p_t^\alpha, p_t^\beta, \nu_t), \nu_{t+1}) \end{cases} \quad (4)$$

относительно переменных  $\langle p_t^\alpha, p_t^\beta \rangle$ . А именно,  $\hat{P}_t^\alpha(R_t, \nu_t) = \hat{p}_t^\alpha$  и  $\hat{P}_t^\beta(R_t, \nu_t) = \hat{p}_t^\beta$ , где пара  $\langle \hat{p}_t^\alpha, \hat{p}_t^\beta \rangle$  – точка равновесия для этих функций при фиксированных  $R_t$  и  $\nu_t$ . Таким образом, равновесные по Нэшу стратегии  $\langle \hat{V}_t^\alpha, \hat{V}_t^\beta \rangle$  могут быть построены рекурсивно:

$$\hat{P}_t^\alpha = \langle \hat{P}_t^\alpha, \hat{P}_{t+1}^\alpha \rangle, \quad \hat{P}_t^\beta = \langle \hat{P}_t^\beta, \hat{P}_{t+1}^\beta \rangle. \quad (5)$$

Чтобы сформулировать это утверждение более точно, примем, что  $R \in \mathcal{R}$  – пространству состояний системы,  $\nu \in \mathcal{N}$  – пространству значений случайного параметра,  $\mathcal{D}^\alpha$  и  $\mathcal{D}^\beta$  – пространства решений  $p^\alpha$  и  $p^\beta$  игроков  $\alpha$  и  $\beta$  соответственно. Будем считать, что  $\mathcal{R}, \mathcal{N}, \mathcal{D}^\alpha, \mathcal{D}^\beta$  – измеримые пространства, а  $v_t^\alpha, v_t^\beta : \mathcal{R} \times \mathcal{N} \times \mathcal{D}^\alpha \times \mathcal{D}^\beta \rightarrow \mathbb{R}$  и  $\rho : \mathcal{R} \times \mathcal{N} \times \mathcal{D}^\alpha \times \mathcal{D}^\beta \rightarrow \mathcal{R}$  – измеримые отображения.

На каждом шаге  $t$  стратегия  $P_t^\alpha$  игрока  $\alpha$  является измеримым отображением из  $\mathcal{R} \times \mathcal{N}$  в  $\mathcal{D}^\alpha$ . Обозначим через  $\tilde{\mathcal{D}}^\alpha$  пространство измеримых отображений вида  $\mathcal{R} \times \mathcal{N} \rightarrow \mathcal{D}^\alpha$ . Аналогично, пусть  $\tilde{\mathcal{D}}^\beta$  – пространство измеримых отображений вида  $\mathcal{R} \times \mathcal{N} \rightarrow \mathcal{D}^\beta$ .

Тогда игру с конечным горизонтом  $T$  можно рассматривать как игру двух лиц

$$G_{t,R,\nu} = \langle \tilde{\mathcal{D}}_t^\alpha, \tilde{\mathcal{D}}_t^\beta, V_t^\alpha, V_t^\beta \rangle_{R,\nu}, \quad R \in \mathcal{R}, \quad \nu \in \mathcal{N}, \quad t = 1, \dots, T,$$

с пространствами стратегий

$$\tilde{\mathcal{D}}_t^\alpha = \prod_{\tau=t}^T \tilde{\mathcal{D}}^\alpha, \quad \tilde{\mathcal{D}}_t^\beta = \prod_{\tau=t}^T \tilde{\mathcal{D}}^\beta,$$

и функциями выигрыша  $V_t^\alpha$  и  $V_t^\beta$ , определяемыми согласно (1).

Точнее, в произвольный момент времени  $t$  речь идет о семействе игр  $\{G_{t,R,\nu} | R \in \mathcal{R}, \nu \in \mathcal{N}\}$ , параметризованном значениями  $R$  и  $\nu$ . Пара стратегий  $\langle \hat{P}_t^\alpha, \hat{P}_t^\beta \rangle$  оптимальна для такого семейства игр (образует точку равновесия по Нэшу), если для произвольных стратегий  $P_t^\alpha$  и  $P_t^\beta$  и любых  $R \in \mathcal{R}$  и  $\nu \in \mathcal{N}$  выполняются неравенства

$$V_t^\alpha(R_t, \nu_t, \hat{P}_t^\alpha, \hat{P}_t^\beta) \geq V_t^\alpha(R_t, \nu_t, P_t^\alpha, \hat{P}_t^\beta), \\ V_t^\beta(R_t, \nu_t, \hat{P}_t^\alpha, \hat{P}_t^\beta) \geq V_t^\beta(R_t, \nu_t, \hat{P}_t^\alpha, P_t^\beta).$$

Рассмотрим теперь семейство игр, определяемых функциями выигрышей  $\tilde{V}_t^\alpha$  и  $\tilde{V}_t^\beta$ ,

$$\tilde{G}_{t,R,\nu} = \langle \mathcal{D}^\alpha, \mathcal{D}^\beta, \tilde{V}_t^\alpha, \tilde{V}_t^\beta \rangle_{R,\nu}, \quad R \in \mathcal{R}, \quad \nu \in \mathcal{N}, \quad t = 1, \dots, T.$$

Следующее предложение позволяет свести проблему построения оптимальной стратегии для игры  $G_{t,R,\nu}$  к многократному построению оптимальных решений для принципиально более простых игр  $\tilde{G}_{t,R,\nu}$ .

**Предложение 4.1.** Пусть для всех  $t$ , начиная с  $t = T$  и до  $t = 1$ , и произвольных  $R \in \mathcal{R}$  и  $\nu \in \mathcal{N}$  игра  $\tilde{G}_{t,R,\nu}$  имеет точку равновесия по Нэшу  $\langle \hat{P}_{R,\nu}^\alpha, \hat{P}_{R,\nu}^\beta \rangle$ ,

а отображения

$$\hat{P}_t^\alpha(R, \nu) = \hat{p}_{R, \nu}^\alpha, \quad \hat{P}_t^\beta(R, \nu) = \hat{p}_{R, \nu}^\beta, \quad \forall R \in \mathcal{R}, \quad \forall \nu \in \mathcal{N}, \quad (6)$$

измеримы по  $R$  и  $\nu$ . Тогда для любого  $t$  существует пара стратегий  $\langle \hat{P}_t^\alpha, \hat{P}_t^\beta \rangle$ , образующая точку равновесия по Нэшу для семейства игр  $\{G_{t, R, \nu} | R \in \mathcal{R}, \nu \in \mathcal{N}\}$ .

Оптимальные стратегии  $\hat{P}_t^\alpha$  и  $\hat{P}_t^\beta$  определяются рекуррентным образом соотношениями (5) и (6) через оптимальные стратегии  $\hat{P}_{t+1}^\alpha, \hat{P}_{t+1}^\beta$  в следующий момент времени и точки равновесия  $\langle \hat{p}_{R, \nu}^\alpha, \hat{p}_{R, \nu}^\beta \rangle$  для игр  $\tilde{G}_{t, R, \nu}$ .

Функции выигрыша  $\tilde{V}_t^\alpha$  и  $\tilde{V}_t^\beta$  для игры  $\tilde{G}_{t, R, \nu}$  выражаются через функции выигрыша для игр в последующий момент времени  $\tilde{G}_{t+1, R, \nu}$  посредством выражений (4) и равенств

$$\begin{aligned} \hat{V}_t^\alpha(R, \nu) &= \tilde{V}_t^\alpha(R, \nu, \hat{p}_{R, \nu}^\alpha, \hat{p}_{R, \nu}^\beta), \\ \hat{V}_t^\beta(R, \nu) &= \tilde{V}_t^\beta(R, \nu, \hat{p}_{R, \nu}^\alpha, \hat{p}_{R, \nu}^\beta). \end{aligned}$$

Доказательство. Достаточно убедиться в справедливости следующего “индуктивного” утверждения для любого  $t = 1, \dots, T$ . Если существует точка равновесия по Нэшу  $\langle \hat{P}_{t+1}^\alpha, \hat{P}_{t+1}^\beta \rangle$  для семейства игр  $\{G_{t+1, R, \nu} | R \in \mathcal{R}, \nu \in \mathcal{N}\}$  и для произвольных  $R \in \mathcal{R}, \nu \in \mathcal{N}$  существует точка равновесия по Нэшу  $\langle \hat{p}_{R, \nu}^\alpha, \hat{p}_{R, \nu}^\beta \rangle$  для игры  $\tilde{G}_{t, R, \nu}$ , то  $\langle \hat{P}_t^\alpha, \hat{P}_t^\beta \rangle$  образует точку равновесия по Нэшу для семейства игр  $\{G_{t, R, \nu} | R \in \mathcal{R}, \nu \in \mathcal{N}\}$ , где  $\hat{P}_t^\alpha$  и  $\hat{P}_t^\beta$  определяются формулами (5) и (6).

Итак, пусть для некоторого момента  $t$  пара  $\langle \hat{P}_{t+1}^\alpha, \hat{P}_{t+1}^\beta \rangle$  определяет оптимальные стратегии для семейства игр  $\{G_{t+1, R, \nu} | R \in \mathcal{R}, \nu \in \mathcal{N}\}$ . Пусть  $P_t^\alpha$  – произвольная стратегия игрока  $\alpha$ . Рассмотрим функцию выигрыша  $V_t^\alpha(R_t, \nu_t, P_t^\alpha, \hat{P}_t^\beta)$ . Используя равенство (3), получаем:

$$\begin{aligned} V_t^\alpha(R, \nu, P_t^\alpha, \hat{P}_t^\beta) &= \\ &= v_t^\alpha(R, \nu, p_{R, \nu}^\alpha, \hat{p}_{R, \nu}^\beta) + \gamma_\alpha E_{(\nu_+ | \nu)} V_{t+1}^\alpha(\rho(R, \nu, p_{R, \nu}^\alpha, \hat{p}_{R, \nu}^\beta), \nu_+, P_{t+1}^\alpha, \hat{P}_{t+1}^\beta) \leq \\ &\leq v_t^\alpha(R, \nu, p_{R, \nu}^\alpha, \hat{p}_{R, \nu}^\beta) + \gamma_\alpha E_{(\nu_+ | \nu)} V_{t+1}^\alpha(\rho(R, \nu, p_{R, \nu}^\alpha, \hat{p}_{R, \nu}^\beta), \nu_+, \hat{P}_{t+1}^\alpha, \hat{P}_{t+1}^\beta) = \\ &= v_t^\alpha(R, \nu, p_{R, \nu}^\alpha, \hat{p}_{R, \nu}^\beta) + \gamma_\alpha E_{(\nu_+ | \nu)} \hat{V}_{t+1}^\alpha(\rho(R, \nu, p_{R, \nu}^\alpha, \hat{p}_{R, \nu}^\beta), \nu_+) = \\ &= \tilde{V}_t^\alpha(R, \nu, p_{R, \nu}^\alpha, \hat{p}_{R, \nu}^\beta) \leq \tilde{V}_t^\alpha(R, \nu, \hat{p}_{R, \nu}^\alpha, \hat{p}_{R, \nu}^\beta) = V_t^\alpha(R, \nu, \hat{P}_t^\alpha, \hat{P}_t^\beta). \end{aligned}$$

Здесь  $p_{R, \nu}^\alpha = P_t^\alpha(R, \nu)$ .

Аналогично,  $V_t^\beta(R, \nu, \hat{P}_t^\alpha, P_t^\beta) \leq V_t^\beta(R, \nu, \hat{P}_t^\alpha, \hat{P}_t^\beta)$  для любой стратегии  $P_t^\beta$  игрока  $\beta$ .

Поэтому пара  $\langle \hat{P}_t^\alpha, \hat{P}_t^\beta \rangle$  образует точку равновесия по Нэшу в игре  $G_{t, R, \nu}$  для всех  $R \in \mathcal{R}, \nu \in \mathcal{N}$ . ▲



Заметим, что точки равновесия  $\langle \hat{p}_{R,\nu}^\alpha, \hat{p}_{R,\nu}^\beta \rangle$  для игр  $\tilde{G}_{t,R,\nu}$  можно строить путем итерационной процедуры, рассмотренной в предложении 8.1.

**4.2. Игроки имеют информацию с задержкой.** Этот случай похож на случай с текущей информацией. Отличие состоит в том, что стратегии и выигрыши в момент времени  $t$  зависят не от  $\nu_t$ , а от  $\nu_{t-1}$ , т.е.

$$p_t^\alpha = P_t^\alpha(R_t, \nu_{t-1}), \quad p_t^\beta = P_t^\beta(R_t, \nu_{t-1}),$$

и

$$\begin{aligned} V_t^\alpha(R_t, \nu_{t-1}, P_t^\alpha, P_t^\beta) &= E_{(\nu_t, \nu_{t+1}, \dots, \nu_T | \nu_{t-1})} \sum_{\tau=t}^T \gamma_\alpha^{\tau-t} v_\tau^\alpha(R_\tau, \nu_\tau, p_\tau^\alpha, p_\tau^\beta) = \\ &= E_{(\nu_t | \nu_{t-1})} E_{(\nu_{t+1}, \dots, \nu_T | \nu_t)} \sum_{\tau=t}^T \gamma_\alpha^{\tau-t} v_\tau^\alpha(R_\tau, \nu_\tau, p_\tau^\alpha, p_\tau^\beta) = \\ &= E_{(\nu_t | \nu_{t-1})} \dot{V}^{\alpha t}(R_t, \nu_t, P_t^\alpha, P_t^\beta), \end{aligned}$$

где  $\dot{V}^{\alpha t}$  обозначает функцию выигрыша для задачи с “текущей информацией”.

Таким образом, игра, в которой оба игрока получают информацию с задержкой, приводит к процедуре динамического программирования для следующей задачи отыскания точки равновесия по Нэшу на каждом шаге:

$$\begin{cases} \tilde{V}_t^\alpha(R_t, \nu_{t-1}, p_t^\alpha, p_t^\beta) = E_{(\nu_t | \nu_{t-1})} \left[ v_t^\alpha(R_t, \nu_t, p_t^\alpha, p_t^\beta) + \gamma_\alpha \hat{V}_{t+1}^\alpha(\rho(R_t, p_t^\alpha, p_t^\beta, \nu_t), \nu_t) \right], \\ \tilde{V}_t^\beta(R_t, \nu_{t-1}, p_t^\alpha, p_t^\beta) = E_{(\nu_t | \nu_{t-1})} \left[ v_t^\beta(R_t, \nu_t, p_t^\alpha, p_t^\beta) + \gamma_\beta \hat{V}_{t+1}^\beta(\rho(R_t, p_t^\alpha, p_t^\beta, \nu_t), \nu_t) \right], \end{cases}$$

где  $\hat{V}_t^\alpha(R_t, \nu_{t-1})$  и  $\hat{V}_t^\beta(R_t, \nu_{t-1})$  – значения этих функций в точке равновесия. Оптимальные управляющие стратегии  $\hat{P}_t^\alpha$  и  $\hat{P}_t^\beta$  определяются как  $\hat{P}_t^\alpha(R_t, \nu_{t-1}) = \hat{p}_t^\alpha$  и  $\hat{P}_t^\beta(R_t, \nu_{t-1}) = \hat{p}_t^\beta$ , где пара  $\langle \hat{p}_t^\alpha, \hat{p}_t^\beta \rangle$  обеспечивает равновесие по Нэшу для этих функций при заданных  $R_t$  и  $\nu_{t-1}$ .

**4.3. Игроки имеют асимметричную информацию: текущая против задержанной.** Предположим теперь, что первый игрок располагает текущей информацией о  $\nu$ , в то время, как второй игрок получает информацию с запаздыванием. Таким образом, стратегия первого игрока зависит от  $\nu_t$  и  $\nu_{t-1}$ , а стратегия второго игрока зависит только от  $\nu_{t-1}$ , т.е.

$$p_t^\alpha = P_t^\alpha(R_t, \nu_{t-1}, \nu_t), \quad p_t^\beta = P_t^\beta(R_t, \nu_{t-1}).$$

В этом случае на каждом шаге динамического программирования получаем следующую (асимметричную) задачу на равновесие по Нэшу:

$$\begin{cases} \tilde{V}_t^\alpha(R_t, \nu_{t-1}, \nu_t, p_t^\alpha, p_t^\beta) = v_t^\alpha(R_t, \nu_t, p_t^\alpha, p_t^\beta) + \gamma_\alpha E_{(\nu_{t+1} | \nu_t)} \hat{V}_{t+1}^\alpha(\rho(R_t, p_t^\alpha, p_t^\beta, \nu_t), \nu_{t+1}), \\ \tilde{V}_t^\beta(R_t, \nu_{t-1}, p_t^\alpha, p_t^\beta) = E_{(\nu_t | \nu_{t-1})} \left[ v_t^\beta(R_t, \nu_t, p_t^\alpha, p_t^\beta) + \gamma_\beta \hat{V}_{t+1}^\beta(\rho(R_t, p_t^\alpha, p_t^\beta, \nu_t), \nu_t) \right]. \end{cases}$$

Теперь первый игрок использует знание  $\nu_{t-1}$  для того, чтобы определить управление второго игрока в момент времени  $t$ . Однако, хотя второй игрок может построить стратегию первого игрока, он не может узнать его реальное управление (поскольку он не знает  $\nu_t$ ). Вместо этого ему доступно лишь (условное) вероятностное распределение для  $\nu_t$ .

Строго говоря, решением задачи поиска равновесия по Нэшу является пара отображений  $\langle \hat{P}_t^\alpha(R_t, \nu_{t-1}, \nu_t), \hat{P}_t^\beta(R_t, \nu_{t-1}) \rangle$ , которая среди всех пар отображений вида  $\langle P_t^\alpha(R_t, \nu_{t-1}, \nu_t), P_t^\beta(R_t, \nu_{t-1}) \rangle$  обеспечивает равновесие по Нэшу для следующей пары функций и всех значений  $R_t, \nu_{t-1}$  и  $\nu_t$ :

$$\begin{cases} \tilde{V}_t^\alpha(R_t, \nu_{t-1}, \nu_t, P_t^\alpha, P_t^\beta) = v_t^\alpha(R_t, \nu_t, P_t^\alpha(R_t, \nu_{t-1}, \nu_t), P_t^\beta(R_t, \nu_{t-1})) + \\ + \gamma_\alpha E_{(\nu_{t+1}|\nu_t)} \hat{V}_{t+1}^\alpha(\rho(R_t, \nu_t, P_t^\alpha(R_t, \nu_{t-1}, \nu_t), P_t^\beta(R_t, \nu_{t-1})), \nu_{t+1}), \\ \tilde{V}_t^\beta(R_t, \nu_{t-1}, P_t^\alpha, P_t^\beta) = E_{(\nu_t|\nu_{t-1})} \left[ v_t^\beta(R_t, \nu_t, P_t^\alpha(R_t, \nu_{t-1}, \nu_t), P_t^\beta(R_t, \nu_{t-1})) + \right. \\ \left. + \gamma_\beta \hat{V}_{t+1}^\beta(\rho(R_t, \nu_t, P_t^\alpha(R_t, \nu_{t-1}, \nu_t), P_t^\beta(R_t, \nu_{t-1})), \nu_t) \right]. \end{cases}$$

Заметим, что каждый игрок может “уточнять” свою стратегию “поточечно”, т.е. для всех возможных значений аргументов своей стратегии, при фиксированной стратегии другого игрока. Это может быть использовано для итерационного вычисления оптимальных стратегий. А именно, для некоторой  $\alpha$ -стратегии мы можем найти оптимальную “ответную”  $\beta$ -стратегию. Затем фиксируем  $\beta$ -стратегию и находим соответствующую “ответную”  $\alpha$ -стратегию, и т.д.

## § 5. Игра с конечным горизонтом: игроки имеют результаты неточных наблюдений

Предположение § 4 о том, что игроки получают точное значение реализации случайного элемента  $\nu_t$  является идеализацией. На практике значение случайного параметра определяется не точно, с некоторой погрешностью. В этом параграфе мы введем понятие измерения (с погрешностью) случайного параметра. В частности, это позволит формализовать представление о различной степени информированности игроков.

Неточное наблюдение параметра  $\nu_t$  может быть описано переходной вероятностью из пространства состояний параметра  $\nu_t$  в пространство наблюдений  $\xi_t$ . Если оба эти пространства конечны, скажем, числа состояний для  $\nu_t$  и  $\xi_t$  равны  $n$  и  $m$  соответственно, то такое измерение полностью определяется стохастической матрицей размерности  $m \times n$ , в которой  $i$ -й столбец задает распределение наблюдения  $\xi_t$  при условии, что  $\nu_t$  принимает  $i$ -е значение.

Итак, положим, что информация, которой обладает игрок в момент времени  $t$  состоит из текущего состояния системы  $R_t$  и результата неточного измерения  $\xi_t$  текущего значения параметра  $\nu_t$ . Игроки могут получать результаты различных измерений  $\xi_t^\alpha$  и  $\xi_t^\beta$ , или это могут быть результаты одного и того же измерения. Теперь стратегии игроков  $\alpha$  и  $\beta$  в момент времени  $t$  зависят от  $R_t$  и  $\xi_t^\alpha$  или  $\xi_t^\beta$  соответственно, т.е.

$$p_t^\alpha = P_t^\alpha(R_t, \xi_t^\alpha), \quad p_t^\beta = P_t^\beta(R_t, \xi_t^\beta).$$

Если игроки располагают одинаковой информацией, т.е.  $\xi_t^\alpha = \xi_t^\beta$ , то обозначим ее  $\xi_t$ . Кроме того, будем предполагать, что случайные элементы  $\nu_t$  независимы и одинаково распределены.

**5.1. Игроки имеют симметричную информацию: максимизация среднего выигрыша.** Пусть стратегии обоих игроков зависят от общего результата измерения  $\xi_t$ :

$$p_t^\alpha = P_t^\alpha(R_t, \xi_t), \quad p_t^\beta = P_t^\beta(R_t, \xi_t),$$

и каждый стремится максимизировать свой средний выигрыш  $U^\alpha$  и  $U^\beta$ , где

$$\begin{aligned} U_t^\alpha(R_t, P_t^\alpha, P_t^\beta) &= \\ &= E_{(\nu_t, \xi_t, \nu_{t+1}, \xi_{t+1}, \dots, \nu_T, \xi_T)} \sum_{\tau=t}^T \gamma_\alpha^{\tau-t} v_\tau^\alpha(R_\tau, \nu_\tau, P_\tau^\alpha(R_\tau, \xi_\tau), P_\tau^\beta(R_\tau, \xi_\tau)). \end{aligned} \quad (7)$$

Как и в рассмотренных выше случаях, можно выразить  $U_t^\alpha$  через текущий выигрыш  $v_t^\alpha$  и средний выигрыш для следующего шага  $U_{t+1}^\alpha$ . В результате получим следующее рекурсивное выражение:

$$\begin{aligned} U_t^\alpha(R_t, P_t^\alpha, P_t^\beta) &= E_{\nu_t} E_{(\xi_t|\nu_t)} E_{\nu_{t+1}} E_{(\xi_{t+1}|\nu_{t+1})} \dots E_{\nu_T} E_{(\xi_T|\nu_T)} \sum_{\tau=t}^T \gamma_\alpha^{\tau-t} v_\tau^\alpha(R_\tau, \nu_\tau, p_\tau^\alpha, p_\tau^\beta) = \\ &= E_{\nu_t} E_{(\xi_t|\nu_t)} \left[ v_t^\alpha(R_t, \nu_t, p_t^\alpha, p_t^\beta) + \right. \\ &\quad \left. + \gamma_\alpha E_{(\xi_{t+1}, \nu_{t+1})} \dots E_{(\xi_T, \nu_T)} \sum_{\tau=t+1}^T \gamma_\alpha^{\tau-(t+1)} v_\tau^\alpha(R_\tau, \nu_\tau, p_\tau^\alpha, p_\tau^\beta) \right] = \\ &= E_{\nu_t} E_{(\xi_t|\nu_t)} \left[ v_t^\alpha(R_t, \nu_t, P_t^\alpha(R_t, \xi_t), P_t^\beta(R_t, \xi_t)) + \right. \\ &\quad \left. + \gamma_\alpha U_{t+1}^\alpha(R_t, \nu_t, P_t^\alpha(R_t, \xi_t), P_t^\beta(R_t, \xi_t)), P_{t+1}^\alpha, P_{t+1}^\beta) \right]. \end{aligned}$$

Пусть, как и раньше,  $\hat{P}_t^\alpha, \hat{P}_t^\beta$  обозначают оптимальные последовательности стратегий, начиная с момента  $t$ , а  $\hat{U}_t^\alpha(R_t), \hat{U}_t^\beta(R_t)$  – соответствующие средние выигрыши, т.е.

$$\hat{U}_t^\alpha(R_t) = U_t^\alpha(R_t, \hat{P}_t^\alpha, \hat{P}_t^\beta), \quad \hat{U}_t^\beta(R_t) = U_t^\beta(R_t, \hat{P}_t^\alpha, \hat{P}_t^\beta),$$

и  $\tilde{U}_t^\alpha(R_t, P_t^\alpha, P_t^\beta), \tilde{U}_t^\beta(R_t, P_t^\alpha, P_t^\beta)$  – средний дисконтированный выигрыш для оптимальных “хвостовых” стратегий, т.е.

$$\begin{aligned} \tilde{U}_t^\alpha(R_t, P_t^\alpha, P_t^\beta) &= U_t^\alpha(R_t, \langle P_t^\alpha, \hat{P}_{t+1}^\alpha \rangle, \langle P_t^\beta, \hat{P}_{t+1}^\beta \rangle) = \\ &= E_{\nu_t} E_{(\xi_t|\nu_t)} \left[ v_t^\alpha(R_t, \nu_t, P_t^\alpha(R_t, \xi_t), P_t^\beta(R_t, \xi_t)) + \right. \\ &\quad \left. + \gamma_\alpha \hat{U}_{t+1}^\alpha(R_t, \nu_t, P_t^\alpha(R_t, \xi_t), P_t^\beta(R_t, \xi_t)) \right]. \end{aligned}$$

Таким образом, стратегии  $\hat{P}_t^\alpha$  и  $\hat{P}_t^\beta$  в момент  $t$  обеспечивают равновесие по Нэшу для пары функций

$$\begin{cases} \tilde{U}_t^\alpha(R_t, P_t^\alpha, P_t^\beta), \\ \tilde{U}_t^\beta(R_t, P_t^\alpha, P_t^\beta). \end{cases} \quad (8)$$

Эти оптимальные стратегии можно построить путем итераций, поточно относительно  $R$  (но не относительно  $\xi$ ), посредством следующей итерационной процедуры. Для некоторого заданного состояния  $R$  и некоторой фиксированной стратегии  $P_{(1)}^\alpha$  игрока  $\alpha$  найдем стратегию оптимального отклика  $P_{(1)}^\beta$  игрока  $\beta$ . Затем найдем оп-

тимальный отклик  $P_{(2)}^\alpha$  игрока  $\alpha$  на  $P_{(1)}^\beta$ , и т.д. Итак,

$$P_{(i)}^\beta = \arg \max_{P^\beta} \tilde{U}_t^\beta (R, P_{(i)}^\alpha, P^\beta),$$

$$P_{(i+1)}^\alpha = \arg \max_{P^\alpha} \tilde{U}_t^\alpha (R, P^\alpha, P_{(i)}^\beta).$$

На каждом шаге итераций оптимизация проводится по элементу функционального пространства, а именно, по отображению  $P_{(i)}^\alpha(R, \xi)$ , которое фактически можно рассматривать как функцию только одной переменной  $\xi$  (поскольку мы можем фиксировать  $R$ , но вынуждены проводить усреднение по всем значениям переменной  $\xi$ ).

Авторы предлагают следующую поточечную конструкцию для получения стратегий, обеспечивающих равновесие по Нэшу в двух функциональных пространствах. Иными словами, авторы предлагают способ перехода от задачи построения равновесия по Нэшу в функциональном пространстве к такой же задаче в конечномерном арифметическом пространстве. Определим

$$W_t^\alpha(R, \nu, \xi, P^\alpha, P^\beta) =$$

$$= v_t^\alpha(R, \nu, P^\alpha(R, \xi), P^\beta(R, \xi)) + \gamma_\alpha \hat{U}_{t+1}^\alpha(\rho_t(R, \nu, P^\alpha(R, \xi), P^\beta(R, \xi)))$$

и аналогичную функцию  $W_t^\beta(R, \nu, \xi, P^\alpha, P^\beta)$  для игрока  $\beta$ . Тогда функция  $\tilde{U}_t^\alpha$  записывается в виде

$$\tilde{U}_t^\alpha(R, P^\alpha, P^\beta) = E_\nu E_{(\nu|\xi)} W_t^\alpha(R, \nu, \xi, P^\alpha, P^\beta) = E_\xi E_{(\nu|\xi)} W_t^\alpha(R, \nu, \xi, P^\alpha, P^\beta),$$

где  $E_{(\nu|\xi)}$  – условное распределение случайного элемента  $\nu$  для заданного  $\xi$ . Здесь использовалось, что математическое ожидание относительно совместного распределения  $\nu$  и  $\xi$  представимо в виде

$$E_{(\nu, \xi)} = E_\nu E_{(\xi|\nu)} = E_\xi E_{(\nu|\xi)}.$$

Предположим, что для некоторого фиксированного  $\xi$  стратегии  $\hat{P}^\alpha$  и  $\hat{P}^\beta$  обеспечивают равновесие по Нэшу для функций

$$\begin{cases} \tilde{V}^\alpha(R, \xi, P^\alpha, P^\beta) = E_{(\nu|\xi)} W_t^\alpha(R, \nu, \xi, P^\alpha, P^\beta), \\ \tilde{V}^\beta(R, \xi, P^\alpha, P^\beta) = E_{(\nu|\xi)} W_t^\beta(R, \nu, \xi, P^\alpha, P^\beta). \end{cases} \quad (9)$$

Это означает, что для всех  $R$  и  $\xi$  и произвольных стратегий  $P^\alpha$  и  $P^\beta$  выполняется

$$\begin{cases} \tilde{V}^\alpha(R, \xi, \hat{P}^\alpha, \hat{P}^\beta) \geq \tilde{V}^\alpha(R, \xi, P^\alpha, \hat{P}^\beta), \\ \tilde{V}^\beta(R, \xi, \hat{P}^\alpha, \hat{P}^\beta) \geq \tilde{V}^\beta(R, \xi, \hat{P}^\alpha, P^\beta). \end{cases}$$

Отсюда получаем

$$\begin{cases} \tilde{U}^\alpha(R, \hat{P}^\alpha, \hat{P}^\beta) = E_\xi \tilde{V}^\alpha(R, \xi, \hat{P}^\alpha, \hat{P}^\beta) \geq E_\xi \tilde{V}^\alpha(R, \xi, P^\alpha, \hat{P}^\beta) = \tilde{U}^\alpha(R, P^\alpha, \hat{P}^\beta), \\ \tilde{U}^\beta(R, \hat{P}^\alpha, \hat{P}^\beta) = E_\xi \tilde{V}^\beta(R, \xi, \hat{P}^\alpha, \hat{P}^\beta) \geq E_\xi \tilde{V}^\beta(R, \xi, \hat{P}^\alpha, P^\beta) = \tilde{U}^\beta(R, \hat{P}^\alpha, P^\beta). \end{cases}$$

Таким образом, исходная задача построения стратегий, обеспечивающих равновесие по Нэшу для пары функций (8) сводится к аналогичной задаче для функций (9), а для этой второй задачи стратегии, обеспечивающие равновесие по Нэшу, вычисляются “поточечно”.

Можно прийти к такой же задаче поиска равновесия по Нэшу, изначально рассматривая игру с *условными* выигрышами

$$V_t^\alpha(R_t, \xi_t, P_t^\alpha, P_t^\beta) = E_{(\nu_t, \nu_{t+1}, \xi_{t+1}, \dots | \xi_t)} \sum_{\tau=t}^T \gamma_\alpha^{\tau-t} v_\tau^\alpha(R_\tau, \nu_\tau, p_\tau^\alpha, p_\tau^\beta) \quad (10)$$

и аналогичное выражение для  $\beta$ . Это будет проделано далее.

**5.2. Игроки имеют симметричную информацию: максимизация условного выигрыша.** В предыдущем пункте оптимизационная задача решалась как задача максимизации *средних* выигрышей, где усреднение бралось по всем возможным реализациям случайного параметра  $\nu_t$  и всем возможным реализациям результата измерения  $\xi_t$ .

Возможно, более естественной является постановка, в которой максимизируются не *средние*, а *условные* (по отношению к результатам измерений) выигрыши согласно выражению (10).

Переписывая  $V_t^\alpha$  в рекурсивной форме, получим

$$\begin{aligned} V_t^\alpha(R_t, \xi_t, P_t^\alpha, P_t^\beta) &= E_{(\nu_t | \xi_t)} E_{(\nu_{t+1}, \xi_{t+1}, \dots)} \sum_{\tau=t}^T \gamma_\alpha^{\tau-t} v_\tau^\alpha(R_\tau, \nu_\tau, p_\tau^\alpha, p_\tau^\beta) = \\ &= E_{(\nu_t | \xi_t)} \left[ v_t^\alpha(R_t, \nu_t, p_t^\alpha, p_t^\beta) + \gamma_\alpha E_{(\nu_{t+1}, \xi_{t+1})} \sum_{\tau=t+1}^T \gamma_\alpha^{\tau-(t+1)} v_\tau^\alpha(R_\tau, \nu_\tau, p_\tau^\alpha, p_\tau^\beta) \right] = \\ &= E_{(\nu_t | \xi_t)} \left[ v_t^\alpha(R_t, \nu_t, p_t^\alpha, p_t^\beta) + \gamma_\alpha E_{(\xi_{t+1})} V_{t+1}^\alpha(\rho_t(R_t, \nu_t, p_t^\alpha, p_t^\beta), \xi_{t+1}, P_{t+1}^\alpha, P_{t+1}^\beta) \right]. \end{aligned}$$

Здесь математическое ожидание  $E_{(\nu_{t+1}, \xi_{t+1})}$  относительно совместного распределения пары  $(\nu_{t+1}, \xi_{t+1})$  представимо как последовательность операций математического ожидания, т.е.

$$E_{(\nu_{t+1}, \xi_{t+1})} = E_{(\xi_{t+1})} E_{(\nu_{t+1} | \xi_{t+1})}.$$

Это приводит к процедуре динамического программирования, при которой на каждом шаге строится точка  $\langle p_t^\alpha, p_t^\beta \rangle$ , обеспечивающая равновесие по Нэшу для следующих функций:

$$\begin{cases} \tilde{V}_t^\alpha(R_t, \xi_t, p_t^\alpha, p_t^\beta) = E_{(\nu_t | \xi_t)} \left[ v_t^\alpha(R_t, \nu_t, p_t^\alpha, p_t^\beta) + \gamma_\alpha E_{(\xi_{t+1})} \hat{V}_{t+1}^\alpha(\rho_t(R_t, \nu_t, p_t^\alpha, p_t^\beta), \xi_{t+1}) \right], \\ \tilde{V}_t^\beta(R_t, \xi_t, p_t^\alpha, p_t^\beta) = E_{(\nu_t | \xi_t)} \left[ v_t^\beta(R_t, \nu_t, p_t^\alpha, p_t^\beta) + \gamma_\beta E_{(\xi_{t+1})} \hat{V}_{t+1}^\beta(\rho_t(R_t, \nu_t, p_t^\alpha, p_t^\beta), \xi_{t+1}) \right]. \end{cases}$$

Оптимальные стратегии  $\langle \hat{P}_t^\alpha, \hat{P}_t^\beta \rangle$  теперь находятся “поточно” как  $\hat{P}_t^\alpha(R_t, \xi_t) = \hat{p}_t^\alpha$  и  $\hat{P}_t^\beta(R_t, \xi_t) = \hat{p}_t^\beta$ , где  $\langle \hat{p}_t^\alpha, \hat{p}_t^\beta \rangle$  – точка равновесия по Нэшу для рассмотренных выше функций.

Точнее, вычисление оптимальных стратегий  $\langle \hat{P}_t^\alpha, \hat{P}_t^\beta \rangle$  на каждом шаге может выполняться двумя различными способами:

(а) Для фиксированных  $R$  и  $\xi$  находим точку равновесия по Нэшу  $\langle \hat{p}^\alpha, \hat{p}^\beta \rangle$  для функций  $\tilde{V}_t^\alpha(R, \xi, p^\alpha, p^\beta)$  и  $\tilde{V}_t^\beta(R, \xi, p^\alpha, p^\beta)$  и определяем  $\hat{P}_t^\alpha(R, \xi) = \hat{p}^\alpha$  и  $\hat{P}_t^\beta(R, \xi) = \hat{p}^\beta$ . В данном случае задача сводится к поточечному нахождению равновесия по Нэшу для всех возможных значений  $R$  и  $\xi$ .

Отметим, что точка  $\langle \hat{p}^\alpha, \hat{p}^\beta \rangle$ , в которой достигается равновесие, также может быть получена итерационно: для фиксированного начального приближения  $p_{(1)}^\alpha$  найдем оптимальное ответное управление  $p_{(1)}^\beta$ , т.е.

$$p_{(1)}^\beta = \arg \max_{p^\beta} \tilde{V}_t^\beta \left( R, \xi, p_{(1)}^\alpha, p^\beta \right),$$

затем найдем оптимальный ответ  $p_{(2)}^\alpha$  для  $p_{(1)}^\beta$ , и т.д. Согласно предложения 8.1 последовательность  $\langle p_{(i)}^\alpha, p_{(i)}^\beta \rangle$ , если она сходится, то сходится к точке равновесия по Нэшу  $\langle \hat{p}^\alpha, \hat{p}^\beta \rangle$ .

(b) Для фиксированной стратегии  $P_{(1)}^\alpha$  найдем оптимальную ответную стратегию  $P_{(1)}^\beta$ , т.е. такую стратегию, что для всех  $R, \xi$  отображение  $P_{(1)}^\beta$  доставляет максимум функционалу  $\tilde{V}_t^\beta(R, \xi, P_{(1)}^\alpha(R, \xi), P^\beta(R, \xi))$  относительно  $P^\beta$ . Затем, аналогичным образом находим  $P_{(2)}^\alpha$ , и т.д. Легко видеть, что на каждом шаге оптимальная ответная стратегия может строиться поточечно (для каждой комбинации  $R, \xi$ ) при любой (!) стратегии соперника.

**5.3. Игроки имеют различную информацию.** В более общей ситуации игроки получают информацию из *различных* измерений  $\xi_t^\alpha \in \mathcal{X}^\alpha$  и  $\xi_t^\beta \in \mathcal{X}^\beta$  случайного параметра  $\nu_t \in \mathcal{N}$ . В этом случае стратегия каждого игрока зависит от всей доступной ему информации, т.е.  $p_t^\alpha = P_t^\alpha(R, \xi^\alpha)$ ,  $p_t^\beta = P_t^\beta(R, \xi^\beta)$ , где  $P_t^\alpha$  и  $P_t^\beta$  – измеримые отображения:

$$P_t^\alpha : \mathcal{R} \times \mathcal{X}^\alpha \rightarrow \mathcal{D}^\alpha, \quad P_t^\beta : \mathcal{R} \times \mathcal{X}^\beta \rightarrow \mathcal{D}^\beta.$$

Обозначим пространства измеримых отображений из  $\mathcal{R} \times \mathcal{X}^\alpha$  в  $\mathcal{D}^\alpha$  и из  $\mathcal{R} \times \mathcal{X}^\beta$  в  $\mathcal{D}^\beta$  как  $\tilde{\mathcal{D}}^\alpha$  и  $\tilde{\mathcal{D}}^\beta$  соответственно. Полные стратегии игроков  $P_t^\alpha$  и  $P_t^\beta$  являются элементами соответственно пространств

$$\tilde{\mathcal{D}}_t^\alpha = \prod_{\tau=t}^T \tilde{\mathcal{D}}^\alpha, \quad \tilde{\mathcal{D}}_t^\beta = \prod_{\tau=t}^T \tilde{\mathcal{D}}^\beta.$$

Рассмотрим игру  $G_{t,R} = \langle \tilde{\mathcal{D}}_t^\alpha, \tilde{\mathcal{D}}_t^\beta, U_t^\alpha, U_t^\beta \rangle_R$ ,  $R \in \mathcal{R}$ ,  $t = 1, \dots, T$ , в которой каждый игрок максимизирует свой средний суммарный пошаговый выигрыш:

$$U_t^\alpha \left( R_t, P_t^\alpha, P_t^\beta \right) = E_{(\nu_t, \xi_t^\alpha, \xi_t^\beta, \dots)} \sum_{\tau=t}^T \gamma_\alpha^{\tau-t} v_\tau^\alpha \left( R_\tau, \nu_\tau, P_\tau^\alpha(R_\tau, \xi_\tau^\alpha), P_\tau^\beta(R_\tau, \xi_\tau^\beta) \right)$$

и аналогичное выражение для  $U_t^\beta(R_t, P_t^\alpha, P_t^\beta)$ .

Поскольку  $U_t^\alpha$  может быть выражено через  $U_{t+1}^\alpha$ :

$$\begin{aligned} U_t^\alpha \left( R_t, P_t^\alpha, P_t^\beta \right) &= E_{(\nu_t, \xi_t^\alpha, \xi_t^\beta, \dots)} \left[ v_t^\alpha \left( R_t, \nu_t, P_t^\alpha(R_t, \xi_t^\alpha), P_t^\beta(R_t, \xi_t^\beta) \right) + \right. \\ &+ \gamma_\alpha \sum_{\tau=t+1}^T \gamma_\alpha^{\tau-t} v_\tau^\alpha \left( R_\tau, \nu_\tau, P_\tau^\alpha(R_\tau, \xi_\tau^\alpha), P_\tau^\beta(R_\tau, \xi_\tau^\beta) \right) \left. \right] = \\ &= E_{(\nu_t, \xi_t^\alpha, \xi_t^\beta)} \left[ v_t^\alpha \left( R_t, \nu_t, P_t^\alpha(R_t, \xi_t^\alpha), P_t^\beta(R_t, \xi_t^\beta) \right) + \right. \\ &+ \gamma_\alpha U_{t+1}^\alpha \left( \rho_t \left( R_t, \nu_t, P_t^\alpha(R_t, \xi_t^\alpha), P_t^\beta(R_t, \xi_t^\beta) \right), \xi_{t+1}^\alpha, P_{t+1}^\alpha, P_{t+1}^\beta \right) \left. \right], \end{aligned}$$

то проблема построения пары оптимальных стратегий  $\langle \hat{P}_t^\alpha, \hat{P}_t^\beta \rangle$  также решается посредством процедуры динамического программирования, в которой для каждого момента  $t$  ищется пара оптимальных стратегий  $\langle \hat{P}_t^\alpha, \hat{P}_t^\beta \rangle$  в некоторой более простой игре при условии, что уже найдено решение  $\langle \hat{P}_{t+1}^\alpha, \hat{P}_{t+1}^\beta \rangle$  для игры  $G_{t+1,R}$ .

А именно, пусть

$$\hat{U}_t^\alpha(R_t) = U_t^\alpha(R_t, P_t^\alpha, P_t^\beta), \quad \hat{U}_t^\beta(R_t) = U_t^\beta(R_t, P_t^\alpha, P_t^\beta), \quad R \in \mathcal{R}, \quad t = 1, \dots, T$$

– оптимальные выигрыши игроков. Обозначим через

$$\begin{aligned} \tilde{U}_t^\alpha(R_t, P_t^\alpha, P_t^\beta) &= E_{(\nu_t, \xi_t^\alpha, \xi_t^\beta)} \left[ v_t^\alpha(R_t, \nu_t, P_t^\alpha(R_t, \xi_t^\alpha), P_t^\beta(R_t, \xi_t^\beta)) + \right. \\ &\quad \left. + \gamma_\alpha \hat{U}_{t+1}^\alpha(\rho_t(R_t, \nu_t, P_t^\alpha(R_t, \xi_t^\alpha), P_t^\beta(R_t, \xi_t^\beta)), \xi_{t+1}^\alpha) \right], \\ \tilde{U}_t^\beta(R_t, P_t^\alpha, P_t^\beta) &= E_{(\nu_t, \xi_t^\alpha, \xi_t^\beta)} \left[ v_t^\beta(R_t, \nu_t, P_t^\alpha(R_t, \xi_t^\alpha), P_t^\beta(R_t, \xi_t^\beta)) + \right. \\ &\quad \left. + \gamma_\beta \hat{U}_{t+1}^\beta(\rho_t(R_t, \nu_t, P_t^\alpha(R_t, \xi_t^\alpha), P_t^\beta(R_t, \xi_t^\beta)), \xi_{t+1}^\beta) \right] \end{aligned} \quad (11)$$

выигрыши игроков при условии, что оба игрока используют произвольные стратегии  $P_t^\alpha$  и  $P_t^\beta$  в момент времени  $t$  и оптимальные стратегии  $\hat{P}_{t+1}^\alpha$  и  $\hat{P}_{t+1}^\beta$  в последующие моменты времени.

Рассмотрим теперь игру, возникающую на шаге  $t$  (пусть шаги от  $\tau = T$  до  $\tau = t+1$  уже пройдены) при каждом фиксированном  $R \in \mathcal{R}$ :

$$\hat{G}_{t,R} = \langle \hat{\mathcal{D}}^\alpha, \hat{\mathcal{D}}^\beta, \tilde{U}_t^\alpha, \tilde{U}_t^\beta \rangle_R, \quad R \in \mathcal{R}, \quad t = 1, \dots, T, \quad (12)$$

где  $\hat{\mathcal{D}}^\alpha$  и  $\hat{\mathcal{D}}^\beta$  – пространства измеримых отображений  $\mathcal{X}^\alpha \rightarrow \mathcal{D}^\alpha$  и  $\mathcal{X}^\beta \rightarrow \mathcal{D}^\beta$  соответственно.

**Предложение 5.1.** Пусть для всех  $t$ , начиная от  $t = T$  и до  $t = 1$ , для произвольных  $R \in \mathcal{R}$  и  $\nu \in \mathcal{N}$  игра  $\hat{G}_{t,R}$  имеет точку равновесия по Нэшу  $\langle \hat{p}_R^\alpha, \hat{p}_R^\beta \rangle$  (где  $\hat{p}_R^\alpha \in \hat{\mathcal{D}}^\alpha$ ,  $\hat{p}_R^\beta \in \hat{\mathcal{D}}^\beta$ ), а отображения

$$\hat{P}_t^\alpha(R, \xi^\alpha) = \hat{p}_R^\alpha(\xi^\alpha), \quad \hat{P}_t^\beta(R, \xi^\beta) = \hat{p}_R^\beta(\xi^\beta), \quad \forall R \in \mathcal{R}, \quad \forall \xi^\alpha \in \mathcal{X}^\alpha, \quad \forall \xi^\beta \in \mathcal{X}^\beta, \quad (13)$$

измеримы по  $R$ ,  $\xi^\alpha$  и  $\xi^\beta$ . Тогда для любого  $t$  существует пара стратегий  $\langle \hat{P}_t^\alpha, \hat{P}_t^\beta \rangle$ , образующая точку равновесия по Нэшу для семейства игр  $\{G_{t,R,\nu} | R \in \mathcal{R}\}$ .

Оптимальные стратегии  $\hat{P}_t^\alpha$  и  $\hat{P}_t^\beta$  определяются рекуррентным образом через оптимальные стратегии  $\hat{P}_{t+1}^\alpha$ ,  $\hat{P}_{t+1}^\beta$  в последующий момент времени и точки равновесия  $\langle \hat{p}_R^\alpha, \hat{p}_R^\beta \rangle$  для игр  $\tilde{G}_{t,R}$  соотношениями (5) и (13).

Функции выигрыша  $\tilde{U}_t^\alpha$  и  $\tilde{U}_t^\beta$  для игры  $\hat{G}_{t,R}$  выражаются через функции выигрыша для игр в последующий момент времени  $\tilde{G}_{t+1,R}$  посредством выражений (11) и равенств

$$\hat{U}_t^\alpha(R) = \tilde{U}_t^\alpha(R, \hat{p}_R^\alpha, \hat{p}_R^\beta), \quad \hat{U}_t^\beta(R) = \tilde{U}_t^\beta(R, \hat{p}_R^\alpha, \hat{p}_R^\beta), \quad R \in \mathcal{R}.$$

**Доказательство.** Проводится по той же схеме, что и доказательство предложения 4.1. ▲

В отличие от предложения 4.1, где задача построения оптимальных стратегий свелась к проблеме отыскания точек равновесия в играх простейшего типа (в которых пространства стратегий игроков совпадают с пространствами их решений  $\mathcal{D}^\alpha$  и  $\mathcal{D}^\beta$ ), и на каждом шаге можно было поточно (по  $R$  и  $\nu$ ) строить оптимальные стратегии игроков  $\hat{P}_t^\alpha$  и  $\hat{P}_t^\beta$ , предложение 5.1 требует на каждом шаге рассмотрения игр  $\hat{G}_{t,R}$  (в которых пространства стратегий  $\hat{\mathcal{D}}^\alpha$  и  $\hat{\mathcal{D}}^\beta$  представляют собой пространства отображений) и построения отображений  $\hat{p}_R^\alpha(\xi^\alpha)$  и  $\hat{p}_R^\beta(\xi^\beta)$  (поточно лишь по  $R$ ). Однако само построение этих отображений может также быть выполнено поточно (по  $\xi^\alpha$  и  $\xi^\beta$ ), если воспользоваться итерационной процедурой предложения 8.3. Действительно, при фиксированных  $R$  и  $t$  игра (12) является игрой типа (14).

Комбинация результатов предложений 5.1 и 4.1 приводит к следующей процедуре построения оптимальных стратегий.

Рассмотрим условный относительно  $\xi_t^\alpha$  суммарный пошаговый выигрыш игрока  $\alpha$ :

$$V_t^\alpha(R_t, \xi_t^\alpha, P_t^\alpha, P_t^\beta) = E_{(\nu_t, \xi_t^\beta, \dots, \xi_t^\beta)} \sum_{\tau=t}^T \gamma_\alpha^{\tau-t} v_\tau^\alpha(R_\tau, \nu_\tau, P_\tau^\alpha(R_\tau, \xi_\tau^\alpha), P_\tau^\beta(R_\tau, \xi_\tau^\beta))$$

и аналогичное выражение для  $V_t^\beta(R_t, \xi_t^\beta, P_t^\alpha, P_t^\beta)$ .

Очевидно,  $U_t^\alpha$  связано с  $V_t^\alpha$  следующим образом:

$$U_t^\alpha(R_t, P_t^\alpha, P_t^\beta) = E_{\xi_t^\alpha} V_t^\alpha(R_t, \xi_t^\alpha, P_t^\alpha, P_t^\beta),$$

где  $E_{\xi_t^\alpha}$  – математическое ожидание по полному распределению случайного элемента  $\xi_t^\alpha$ .

Также  $V_t^\alpha$  выражается через  $V_{t+1}^\alpha$ :

$$V_t^\alpha(R_t, \xi_t^\alpha, P_t^\alpha, P_t^\beta) = E_{(\nu_t, \xi_t^\beta | \xi_t^\alpha)} \left[ v_t^\alpha(R_t, \nu_t, P_t^\alpha(R_t, \xi_t^\alpha), P_t^\beta(R_t, \xi_t^\beta)) + \gamma_\alpha E_{\xi_{t+1}^\alpha} V_{t+1}^\alpha(\rho_t(R_t, \nu_t, P_t^\alpha(R_t, \xi_t^\alpha), P_t^\beta(R_t, \xi_t^\beta)), \xi_{t+1}^\alpha, P_{t+1}^\alpha, P_{t+1}^\beta) \right].$$

Совместно с предложениями 5.1 и 4.1 это приводит к следующей процедуре динамического программирования, в которой на каждом шаге строится пара стратегий  $\langle \hat{P}_t^\alpha, \hat{P}_t^\beta \rangle$ , обеспечивающих оптимальные отклики друг для друга относительно функций:

$$\begin{cases} \tilde{V}_t^\alpha(R_t, \xi_t^\alpha, p_t^\alpha, p_t^\beta) = E_{(\nu_t, \xi_t^\beta | \xi_t^\alpha)} \left[ v_t^\alpha(R_t, \nu_t, p_t^\alpha, p_t^\beta(R_t, \xi_t^\beta)) + \gamma_\alpha E_{\xi_{t+1}^\alpha} \tilde{V}_{t+1}^\alpha(\rho_t(R_t, \nu_t, p_t^\alpha, p_t^\beta(R_t, \xi_t^\beta)), \xi_{t+1}^\alpha) \right], \\ \tilde{V}_t^\beta(R_t, \xi_t^\beta, p_t^\alpha, p_t^\beta) = E_{(\nu_t, \xi_t^\alpha | \xi_t^\beta)} \left[ v_t^\beta(R_t, \nu_t, p_t^\alpha(R_t, \xi_t^\alpha), p_t^\beta) + \gamma_\beta E_{\xi_{t+1}^\beta} \tilde{V}_{t+1}^\beta(\rho_t(R_t, \nu_t, p_t^\alpha(R_t, \xi_t^\alpha), p_t^\beta), \xi_{t+1}^\beta) \right]. \end{cases}$$



А именно, оптимальная пара  $\langle \hat{P}_t^\alpha, \hat{P}_t^\beta \rangle$  должна удовлетворять условиям

$$\begin{cases} \tilde{V}_t^\alpha(R_t, \xi_t^\alpha, p_t^\alpha, \hat{P}_t^\beta) \leq \tilde{V}_t^\alpha(R_t, \xi_t^\alpha, \hat{P}_t^\alpha(R_t, \xi_t^\alpha), \hat{P}_t^\beta), \\ \tilde{V}_t^\beta(R_t, \xi_t^\beta, \hat{P}_t^\alpha, p_t^\beta) \leq \tilde{V}_t^\beta(R_t, \xi_t^\beta, \hat{P}_t^\beta, \hat{P}_t^\alpha(R_t, \xi_t^\beta)). \end{cases}$$

В данном случае мы не можем использовать алгоритм (а) из предыдущего пункта для построения оптимальных стратегий, поскольку теперь игроки (даже зная стратегии  $P_t^\alpha$  и  $P_t^\beta$  друг друга) не могут предсказывать реальные управления  $p_t^\alpha$  и  $p_t^\beta$  друг у друга (поскольку они определяются используемыми игроками результатами разных измерений  $\xi_t^\alpha$  и  $\xi_t^\beta$ ). Однако мы можем использовать итерационную конструкцию (b) предыдущего пункта.

## § 6. Кооперативное поведение игроков

В рассмотренной модели несложно определить ситуации кооперации или противоборства игроков. Для этого модифицируем выигрыши  $V^\alpha$  и  $V^\beta$  так, чтобы отразить “заботу” одного игрока о другом.

А именно, игрок  $\alpha$  “учитывает” интерес игрока  $\beta$ , если он оптимизирует линейную комбинацию их выигрышей

$$V^\alpha = c_{\alpha\alpha}V^\alpha + c_{\alpha\beta}V^\beta$$

вместо только своего выигрыша  $V^\alpha$ . Аналогично, пусть  $\beta$  максимизирует линейную комбинацию

$$V^\beta = c_{\beta\alpha}V^\alpha + c_{\beta\beta}V^\beta.$$

Фактически, эта игра описывается (в случае  $c_{\alpha\alpha} + c_{\beta\alpha} = 1$  и  $c_{\alpha\beta} + c_{\beta\beta} = 1$ ) игрой с побочными платежами, когда игрок знает, что он получит определенную часть дохода второго игрока.

Если  $c_{\alpha\beta} > 0$ , то игрок  $\alpha$  старается в некоторой мере повысить доход игрока  $\beta$ , а если  $c_{\alpha\beta} < 0$ , то  $\alpha$  заботится об интересах  $\beta$  больше, чем о своих. Напротив, условие  $c_{\beta\alpha} < 0$  означает, что  $\alpha$  пытается “навредить”  $\beta$ , возможно, с целью “исключения” его из игры.

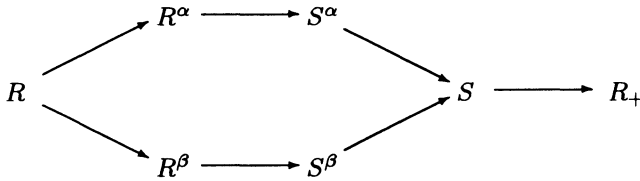
Интересный частный случай получается, если  $c_{\alpha\alpha} = c_{\alpha\beta} = c_{\beta\alpha} = c_{\beta\beta} = \frac{1}{2}$ . Фактически, он описывает ситуацию с одним игроком (монополистом).

Если коэффициенты при слагаемых в суммарном выигрыше равны, т.е.  $\gamma_\alpha = \gamma_\beta = \gamma$ , то введение кооперативных коэффициентов в постановку задачи почти не отражается на алгоритме динамического программирования для ее решения. Действительно, выражения типа  $v_t^\alpha(R_t, \nu_t, p_t^\alpha, p_t^\beta)$  просто заменяются на  $c_{\alpha\alpha}v_t^\alpha(R_t, \nu_t, p_t^\alpha, p_t^\beta) + c_{\alpha\beta}v_t^\beta(R_t, \nu_t, p_t^\alpha, p_t^\beta)$ . Например, в случае “текущей информации” мы приходим к следующей задаче на равновесие по Нэшу на каждом шаге динамического программирования:

$$\begin{cases} \tilde{V}_t^\alpha(R_t, \nu_t, p_t^\alpha, p_t^\beta) = c_{\alpha\alpha}v_t^\alpha(R_t, \nu_t, p_t^\alpha, p_t^\beta) + c_{\alpha\beta}v_t^\beta(R_t, \nu_t, p_t^\alpha, p_t^\beta) + \\ + \gamma E_{(\nu_{t+1}|\nu_t)} \hat{V}_{t+1}^\alpha(\rho(R_t, p_t^\alpha, p_t^\beta, \nu_t), \nu_{t+1}), \\ \tilde{V}_t^\beta(R_t, \nu_t, p_t^\alpha, p_t^\beta) = c_{\beta\alpha}v_t^\alpha(R_t, \nu_t, p_t^\alpha, p_t^\beta) + c_{\beta\beta}v_t^\beta(R_t, \nu_t, p_t^\alpha, p_t^\beta) + \\ + \gamma E_{(\nu_{t+1}|\nu_t)} \hat{V}_{t+1}^\beta(\rho(R_t, p_t^\alpha, p_t^\beta, \nu_t), \nu_{t+1}). \end{cases}$$

## § 7. Компьютерное моделирование

**7.1. Лов рыбы в расщепленном потоке.** В модели лова рыбы в расщепленном потоке мы полагаем, что каждый игрок (рыболовный флот) ведет лов в отдельном потоке и случайный коэффициент распределения рыбы по потокам неизвестен или известен игрокам с некоторой точностью. Такая игра иллюстрируется диаграммой:



Здесь  $R$  определяет полный объем рыбы доступной для лова или, иными словами, “ресурс”, а  $R^\alpha$  и  $R^\beta$  – частичные объемы доступные игрокам  $\alpha$  и  $\beta$  соответственно. Таким образом,

$$R^\alpha = \theta^\alpha R, \quad R^\beta = \theta^\beta R,$$

где

$$\theta^\alpha = \theta, \quad \theta^\beta = 1 - \theta,$$

и  $\theta$  является случайным распределяющим фактором. Остатки рыбы в потоках после лова обозначаются  $S^\alpha$  и  $S^\beta$  соответственно. Они определяются как

$$S^\alpha = \sigma^\alpha(R^\alpha, p^\alpha), \quad S^\beta = \sigma^\beta(R^\beta, p^\beta).$$

Здесь  $p^\alpha$  и  $p^\beta$  – управляющие решения игроков для данного сезона, которые определяют, какая часть доступной рыбы вылавливается. Определим решение игрока как относительную часть оставляемого им ресурса, т.е.

$$S^\alpha = p^\alpha R^\alpha, \quad S^\beta = p^\beta R^\beta.$$

Наконец, остатки отдельных потоков объединяются в полный остаток данного года

$$S = S^\alpha + S^\beta,$$

который определит объем  $R_+$  рыбы в следующем году посредством функции “размножения”

$$R_+ = F(S).$$

Управляющие стратегии игроков (стратегии лова) зависят, как от информации о структуре игры в целом, так и от специфической информации, которой игрок обладает на момент принятия решения. Всегда будем считать, что оба игрока знают текущий полный объем ресурса  $R$ , а также, у каждого есть некоторая информация  $\xi^\alpha$  и  $\xi^\beta$  о текущем случайном возмущении  $\theta$ . Таким образом,

$$p^\alpha = P^\alpha(R, \xi^\alpha), \quad p^\beta = P^\beta(R, \xi^\beta),$$

где  $P^\alpha$  и  $P^\beta$  – решающие стратегии игроков.

Степень осведомленности игроков о случайном распределяющем факторе (парамetre)  $\theta$  может быть различной. Во всех случаях будем считать, что игроки знают, по меньшей мере, стохастические свойства случайного параметра. Кроме того, игро-

ки могут обладать дополнительной информацией: например, иметь полную текущую информацию (значение этого параметра в данном сезоне), или могут получать результаты неточных измерений текущего значения параметра; или, наконец, могут не иметь более никакой дополнительной информации. Вместе с тем, структура знаний может быть асимметричной, т.е. игроки информированы в различной степени.

В каждом сезоне игрок получает годовой доход (выигрыш за год)  $v_{spl}^\alpha$  или  $v_{spl}^\beta$ , который является заданной функцией от ресурса в его потоке и от его решения, т.е.

$$v_{spl}^\alpha = v_{spl}^\alpha(R^\alpha, p^\alpha), \quad v_{spl}^\beta = v_{spl}^\beta(R^\beta, p^\beta).$$

Полный доход игрока в этой динамической игре определяется как сумма (с весовыми коэффициентами) его годовых доходов на протяжении всего времени игры.

Легко видеть, что лов рыбы в разделенных потоках можно рассматривать как частный случай динамической игры, определенной выше. Действительно, вспоминая обозначения § 2, положим  $\nu = \theta$  и функцию  $\rho$  определим как

$$\rho(R, \theta, p^\alpha, p^\beta) = F((\theta p^\alpha + (1 - \theta)p^\beta)R),$$

а функции  $v^\alpha$  и  $v^\beta$  выразим через соответствующие функции  $v_{spl}^\alpha$  и  $v_{spl}^\beta$  следующим образом:

$$v^\alpha(R, \theta, p^\alpha, p^\beta) = v_{spl}^\alpha(\theta R, p^\alpha), \quad v^\beta(R, \theta, p^\alpha, p^\beta) = v_{spl}^\beta((1 - \theta)R, p^\beta).$$

В примерах рассмотренных далее считается, что случайные величины  $\theta_t$  независимы и одинаково распределены.

**7.2. Детали компьютерной модели.** Итак, в процессе моделирования мы используем алгоритмы, описанные в §§ 4 и 5, а также выбираем целый ряд числовых и функциональных параметров: функцию годового дохода, описанную в 7.2.1; функцию роста среды, описанную в 7.2.2; а также другие уже числовые параметры, приведенные в 7.2.3. Особое внимание уделяется вариантам информационной структуры игры, изучение которой является нашей целью (они описываются в 7.2.4).

Все эти параметры и варианты изменялись нами с некоторым “шагом” в довольно широких пределах, так что всего было обчислено около тысячи различных комбинаций. Важно, что эффекты, о которых мы пишем в 7.3, наблюдались при всех этих комбинациях с большей или меньшей выраженностью. Конечно, мы приводим результаты счета только для одной комбинации параметров, которая типична, как с точки зрения наблюдаемых нами эффектов, так и с точки зрения содержания модели, о которой говорится в пункте 7.1 (т.е. числовые значения, которые указаны далее, считаются естественными и с точки зрения биологии процесса).

Как обычно в ситуации компьютерного эксперимента, мы не можем утверждать, что во всех случаях нашего счета алгоритм сходил к искомой точке равновесия. Это связано, как с отсутствием простого и удобного критерия сходимости алгоритма, так и с возможными эффектами компьютерной реализации любого алгоритма. Нами просчитано около 30 тысяч вариантов игр, в 96% случаев наблюдалась сходимость в следующем “практическом” смысле.

Пусть  $x_t = \langle P_t^\alpha, P_t^\beta \rangle$ , а  $\delta_t$  – обычная невязка итерационной процедуры. Если на некоторой итерации эта невязка становилась меньше, чем фиксированный порог  $\delta_0$ , и оставалась таковой не менее, чем  $d_0$  последующих итераций, то счет прекращался, и для исходных данных алгоритм считался сходящимся. Если такое условие не выполнялось более, чем  $D_0$  итераций, то счет прекращался, и алгоритм считался расходящимся.

Заметим, что даже для одних данных для каждой итерации приходилось вычислять оптимальную ответную стратегию как функцию двух переменных: непрерывной  $R$  (использовалось задание на сетке из 21 или более узлов и линейная интерполяция между узлами сетки) и дискретной  $\theta$  или  $\xi$  (принимавшей несколько, как правило два значения), что требует значительного объема вычислений.

Теоретический анализ вопросов о единственности пары стратегий, образующей равновесие по Нэшу, и об устойчивости нашего алгоритма к выбору начальной пары стратегий требует отдельного глубокого исследования.

Счет, который мы провели для этой статьи, был основан на выборе в качестве начальной пары стратегий (соответствующей последнему моменту времени  $T$ ) естественных по смыслу игры стратегий, заключающихся в отсутствии лова. Для остальных моментов времени  $t$  в качестве начального приближения выбиралась пара оптимальных стратегий, полученная на предыдущем шаге игры (т.е. для следующего момента времени  $t+1$ ). На всех шагах алгоритма игрок выбирал свою стратегию как реакцию на выбранную другим игроком стратегию, которая строилась глобальной оптимизацией на соответствующей сетке.

Для проверки устойчивости алгоритма к выбору начальной пары стратегий (и единственности точки равновесия по Нэшу) применялся следующий обычный прием. Начальные пары стратегий игроков выбирались случайным образом в пространстве всех таких пар. Во всех рассмотренных нами случаях алгоритм сходил к той же самой точке равновесия Нэша, т.е. результат не зависел от выбора начальной пары стратегий.

**7.2.1. Функция мгновенного выигрыша игрока.** Решения игроков состоят в фиксировании размера той части ресурса, которая не будет ими выловлена, т.е.

$$S^\alpha = p^\alpha R^\alpha, \quad S^\beta = p^\beta R^\beta.$$

Тогда уловы игроков соответственно равны

$$H^\alpha = (1 - p^\alpha)R^\alpha, \quad H^\beta = (1 - p^\beta)R^\beta.$$

В нашей модели учитывается зависимость стоимости лова от объема доступного игроку ресурса, а именно, удельная стоимость лова игрока обратно пропорциональна объему ресурса в его потоке. Таким образом, полная стоимость лова игрока получается как интеграл от объема доступного ему ресурса в пределах от оставшегося объема  $S = pR$  до начального объема ресурса  $R$ :

$$\text{cost} = \int_S^R \frac{c}{x} dx = -c \log(p).$$

Тогда его годовой доход равен

$$v(R, p) = H - \text{cost} = H + c \log(p).$$

**7.2.2. Функция роста среды.** Здесь используется функция роста среды, которая учитывает возможность полного уничтожения соответствующего ресурса, когда его объем упадет ниже некоторого критического уровня. Это означает, что должно выполняться  $F(S) < S$  при достаточно малых  $S$ .

Мы рассматривали довольно широкий класс функций  $F(S)$  роста среды, который состоял из полиномов третьей степени, проходящих через точку  $(0, 0)$ :

$$F(S) = a_1 S + a_2 S^2 + a_3 S^3.$$

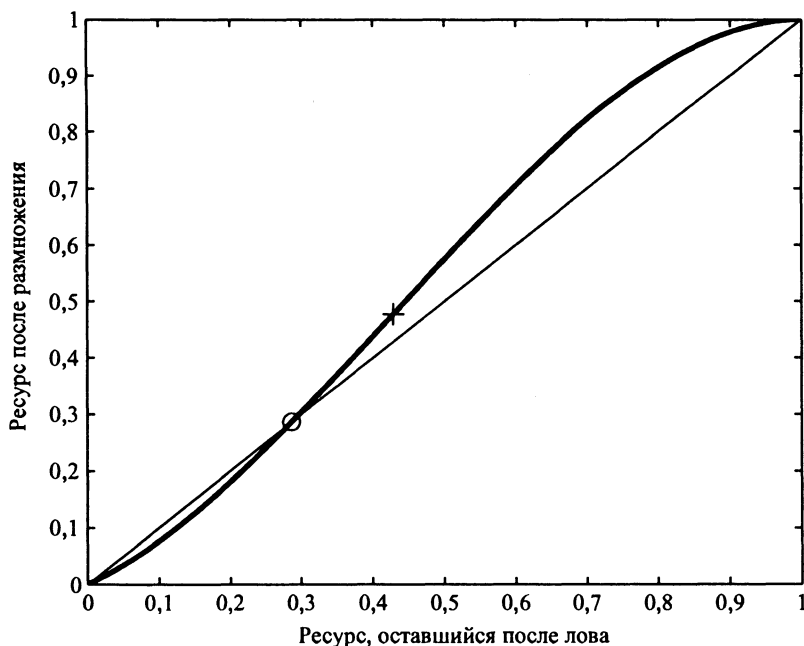


Рис. 1. Функция роста среды. Окружностью отмечен критический уровень ресурса

Поскольку  $a_1$  равен производной этой функции в нуле, то при значениях  $0 < a_1 < 1$  возникает упомянутая возможность полного уничтожения ресурса. Далее приводятся результаты счета для естественного случая  $a_1 = 0,6$ . Функция  $F(S)$  должна где-то (на положительной полуоси) принимать максимальное значение. Также представлены результаты счета для случая, когда этот максимум достигается в точке  $(1, 1)$ . Итак, конкретный вид этой функции дается здесь выражением

$$F(S) = 0,6S + 1,8S^2 - 1,6S^3,$$

ее график приведен на рис. 1.

**7.2.3. Параметры компьютерного моделирования.** Результаты моделирования, представленные в работе, относятся к игре из пункта 7.1 с отдельными потоками. Параметры принимают следующие значения:

- функция выигрыша (7.2.1) включает  $c = 0,2$ ;
- функция роста (7.2.2) имеет вид, показанный на рис. 1;
- состояния случайного параметра  $\theta$  равны:  $\theta_1 = 0,1$  и  $\theta_2 = 0,9$ , в различные моменты времени его состояния независимы, одинаково распределены и наступают с равными вероятностями  $P(\theta = \theta_1) = P(\theta = \theta_2) = 0,5$ ;
- дисконтный коэффициент равен  $\gamma_\alpha = \gamma_\beta = 0,9$ ;
- коэффициенты кооперации из § 6 равны  $c_{\alpha\alpha} = 1$ ,  $c_{\alpha\beta} = 0$ ,  $c_{\beta\alpha} = 0$ ,  $c_{\beta\beta} = 1$  (для чистого соперничества) и  $c_{\alpha\alpha} = 0,5$ ,  $c_{\alpha\beta} = 0,5$ ,  $c_{\beta\alpha} = 0,5$ ,  $c_{\beta\beta} = 0,5$  (для полной кооперации).

Чтобы плавно менять “степень полноты” информации о случайном параметре  $\theta$ , используем неточные измерения (см. [5]) и меняем параметр “точности измерения”  $\pi$ ,

который определяет матрицу измерения

$$M = \begin{pmatrix} P(\xi = \xi_1 | \theta = \theta_1) & P(\xi = \xi_1 | \theta = \theta_2) \\ P(\xi = \xi_2 | \theta = \theta_1) & P(\xi = \xi_2 | \theta = \theta_2) \end{pmatrix},$$

т.е. матрицу переходного распределения наблюдаемой величины  $\xi$  для различных значений  $\theta$ , следующим образом:

$$M = \begin{pmatrix} (1 + \pi)/2 & (1 - \pi)/2 \\ (1 - \pi)/2 & (1 + \pi)/2 \end{pmatrix}.$$

Таким образом, если  $\pi = 1$  (максимальная точность), то  $M = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ . Это отвечает “тождественному” измерению. Если  $\pi = 0$  (минимальная точность), то  $M = \begin{pmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{pmatrix}$  – это означает “независимость” результатов измерения от состояния  $\theta$ .

Выбирался достаточно длинный промежуток усреднения, примерно в 2000 лет.

**7.2.4. Основные типы информационной структуры игры.** Нами были выделены и рассмотрены в компьютерном эксперименте, в частности, следующие типы информационной структуры игры (в скобках приводится название и обозначение для каждого из этих типов).

1. Игроки в равной мере имеют в каждый текущий момент  $t$  времени (и во все предыдущие моменты) реализацию случайного параметра  $\theta$  (“текущая информация”, Cur).

2. Игроки в равной мере знают только само вероятностное распределение случайного дискретного параметра  $\theta$  (“минимальная информация”, Min).

3. Игроки имеют асимметричную информацию: первый имеет текущую информацию по пункту 1, а второй – лишь минимальную по пункту 2 (“асимметричная информация”, Cur-Min).

4. Игроки имеют в равной мере неполную информацию: они получают реализацию случайной величины  $\xi$  измерения случайного параметра  $\theta$  (“неполная информация, полученная из измерений”, Meas).

5. Игроки имеют асимметричную информацию: первый получает неполную информацию по пункту 4, а второй – лишь минимальную по пункту 2 (“асимметричная неполная информация”, Meas-Min).

Приводятся также результаты моделирования кооперативного поведения игроков для симметричных информационных структур 1, 2 и 4 (так как при кооперативном поведении представляется естественным, что игроки полностью обмениваются всей доступной им информацией).

### 7.3. Результаты компьютерного моделирования.

**7.3.1. Влияние стоимости лова.** Исследуем влияние на результат игры удельной стоимости лова  $c$  (см. 7.2.1) для различных информационных структур.

Представленные на рис. 2 графики показывают средние доходы в зависимости от удельной стоимости лова для следующих пяти типов игр: кооперативная ловля с текущей информацией, кооперативная ловля с минимальной информацией, некооперативная ловля с текущей информацией, некооперативная ловля с минимальной информацией, и, наконец, некооперативная ловля с асимметричной информацией (текущая против минимальной).

Наиболее интересным в случаях, когда игроки соперничают, является резкое сокращение среднего дохода при снижении удельной стоимости лова. Это связано с

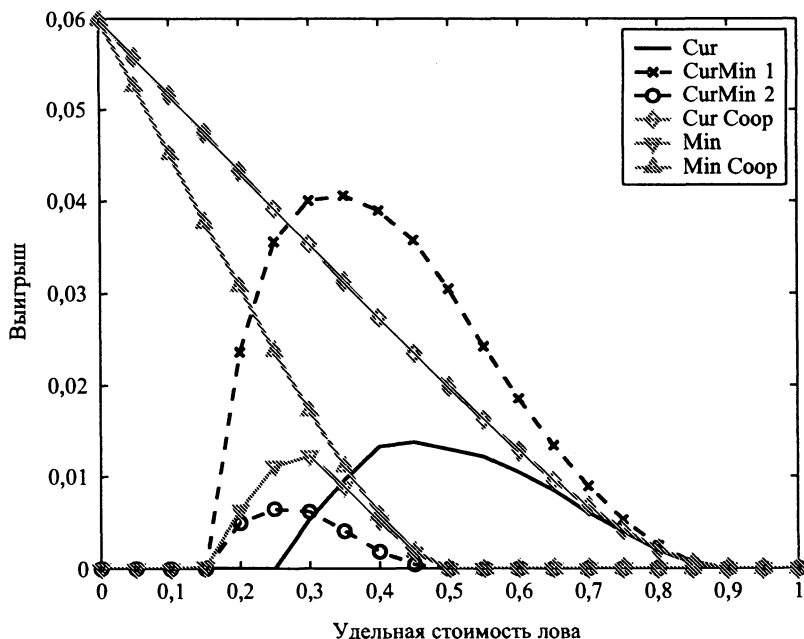


Рис. 2. Влияние удельной стоимости лова для различных информационных структур: "Cur" – оба игрока имеют одну и ту же текущую информацию (включая все предшествующие моменты времени); "Cur-Min" – игроки имеют асимметричную информацию: игрок 1 обладает текущей информацией, а игрок 2 минимальной; "Cur-Coop" – кооперативное поведение игроков и совместная текущая информация; "Min" – оба игрока имеют в равной мере минимальную информацию; "Min-Coop" – кооперативное поведение игроков и совместная минимальная информация

тем, что при низкой удельной стоимости лова конкуренция носит весьма острый характер, и это приводит к резкому снижению объема ресурса. Данный эффект не наблюдается, когда игроки сотрудничают и, таким образом, оказываются в состоянии поддерживать объем рыбы на достаточно высоком уровне.

Кроме того, из рис. 2 видно, что информационное преимущество в случае информационно-асимметричной игры чрезвычайно выгодно для первого игрока ("Cur-Min 1"). Более того, ему не выгодно делиться информацией с соперником и, таким образом, перейти к игре с полной симметричной информацией ("Cur"). Что касается второго игрока ("Cur-Min 2") в асимметричном случае, то он как правило предпочел бы получить дополнительную информацию, но не для низкой стоимости лова ( $< 0,3$ ), где односторонний недостаток информации более выгоден (даже для него), чем полная симметричная информация ("Cur"). Заметим также, что при достаточно низких значениях удельной стоимости лова ( $< 0,35$ ), симметричная минимальная информация ("Min") становится предпочтительнее полной информации ("Cur"). Это, по-видимому, связано с тем, что при недостатке информации возрастает риск случайного уничтожения ресурса и, как следствие этого, поведение игроков становится более "осторожным".

Как уже отмечалось, соперничество становится особенно разрушительным при низкой стоимости лова, когда можно легко уничтожить ресурс. Однако, если игроки сотрудничают, их доход становится заметно выше, особенно когда стоимость лова

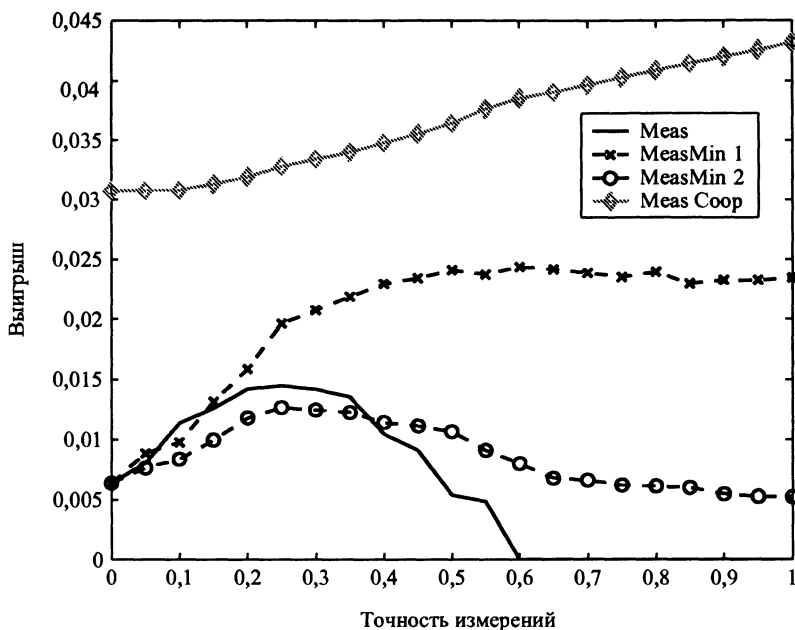


Рис. 3. Влияние точности измерений: “Meas” – оба игрока имеют одинаковую неполную информацию; “Meas-Min” – первый игрок получает неполную информацию, а второй лишь минимальную; “Meas Coop” – кооперация при наличии одинаковой неполной информации

мала. При этом они в состоянии поддерживать объем возобновляемого ресурса на довольно высоком уровне.

**7.3.2. Игроки получают информацию из измерений.** В данном наборе численных экспериментов (рис. 3) информация о текущем значении  $\theta$  получается из неточных измерений. При этом точность измерений является переменным параметром, изменяющимся от 0 (отсутствие информации) до 1 (полная информация).

Заметим, что когда игроки сотрудничают (“Meas Coop”), ситуация вполне естественная: чем точнее информация, тем выше доход. Суммарный доход в такой кооперативной игре всегда выше, чем суммарные доходы в некооперативных играх.

Как видно из рис. 3, в некооперативных играх с симметричной (“Meas”) и асимметричной (“Meas-Min”) информационными структурами, повышение точности информации до определенного уровня выгодно обоим игрокам, даже игроку, которому не доступны результаты измерений (“Meas-Min 2” в случае асимметричной информации). Однако дальнейшее повышение точности информации приводит к значительному снижению доходов, поскольку стратегии игроков становятся более агрессивными.

Кроме того, при низкой точности измерений ситуация, когда второй игрок также имеет доступ к результатам измерений (“Meas”), более выгодна для него, чем одностороннее отсутствие у него этой информации (“Meas-Min 2”). Однако, начиная с определенного уровня точности измерений ( $> 0,38$ ), такая дополнительная информация становится невыгодной для него.

**7.3.3. Влияние информационной асимметрии при наличии асимметрии в природных условиях.** В данном эксперименте природные условия несколько благоприятствуют второму игроку (рис. 4). А именно,  $\theta$  принимает значения 0,1 и 0,8 с рав-



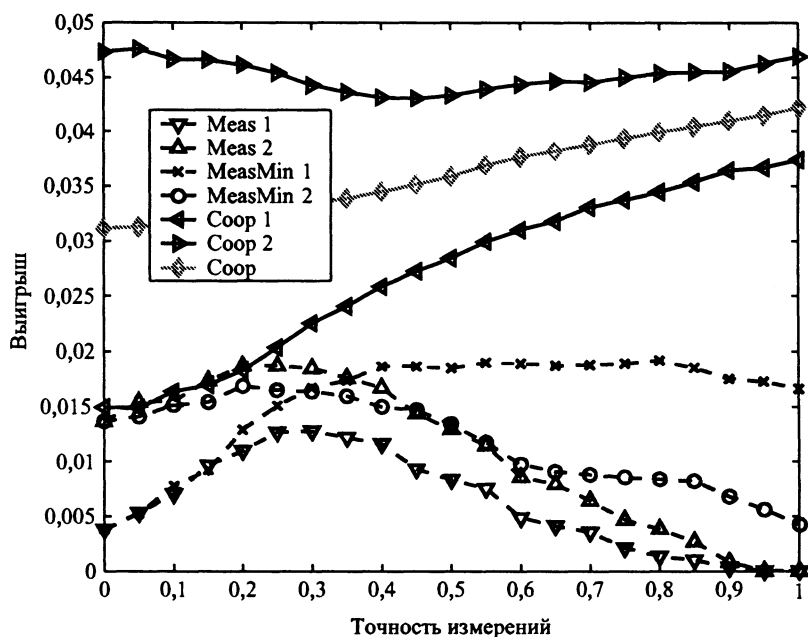


Рис. 4. Влияние точности измерений при асимметричных природных условиях: “Meas” – оба игрока получают одинаковую неполную информацию; “Meas-Min” – первый игрок получает неполную информацию, а второй лишь минимальную; “Coop” – кооперация при наличии одинаковой неполной информации (кроме реальных доходов игроков “Coop 1” и “Coop 2” представлен их усредненный доход “Coop”)

ными вероятностями (так что ресурс первого игрока составляет 0,1 или 0,8 от всего ресурса, в то время как второй игрок получает соответственно 0,9 или 0,2 от всего ресурса). Таким образом, в среднем ресурс у первого игрока ниже, чем у второго.

Как и следует ожидать, если игроки обладают одинаковой информацией, доход второго игрока (“Meas 2”) всегда будет выше, чем первого игрока (“Meas 1”). Однако, когда первый игрок (“Meas-Min 1”) обладает серьезным информационным преимуществом и точность  $> 0,3$ , то это перевесит природное преимущество второго игрока (“Meas-Min 2”).

С другой стороны, сумма доходов игроков будет максимальной при условии кооперации, когда игроки обмениваются информацией и максимизируют их суммарный доход. Однако в этом случае вследствие природной асимметрии их непосредственные доходы от уловов не будут равны (“Coop 1” and “Coop 2”). Поэтому такого рода сотрудничество можно рассматривать как “проявление доброй воли”. Но игроки могут договориться о перераспределении их суммарного дохода, что приведет к “побочным платежам” одного игрока другому. Случай равного распределения доходов также изображен на рисунке рис. 4 (“Coop”).

**7.3.4. Влияние самой природной изменчивости.** Мы меняем степень изменчивости коэффициента  $\theta$  (рис. 5). Точнее,  $\theta$  принимает одно из двух значений:  $\theta_1$  или  $\theta_2 = 1 - \theta_1$ , где  $\theta_1$  может быть любым числом от 0 до 0,5. Так при  $\theta_1 = 0,5$  изменчивость отсутствует ( $\theta = 0,5$  всегда), в то время, как при  $\theta_1 = 0$  изменчивость максимальна ( $\theta$  принимает значения 0 или 1).

Возрастание дохода в кооперативной игре с полной информацией (“Cur Coop”), а также при возрастании изменчивости  $\theta$  (и убывании  $\theta_1$ ) вполне естественно. Дей-

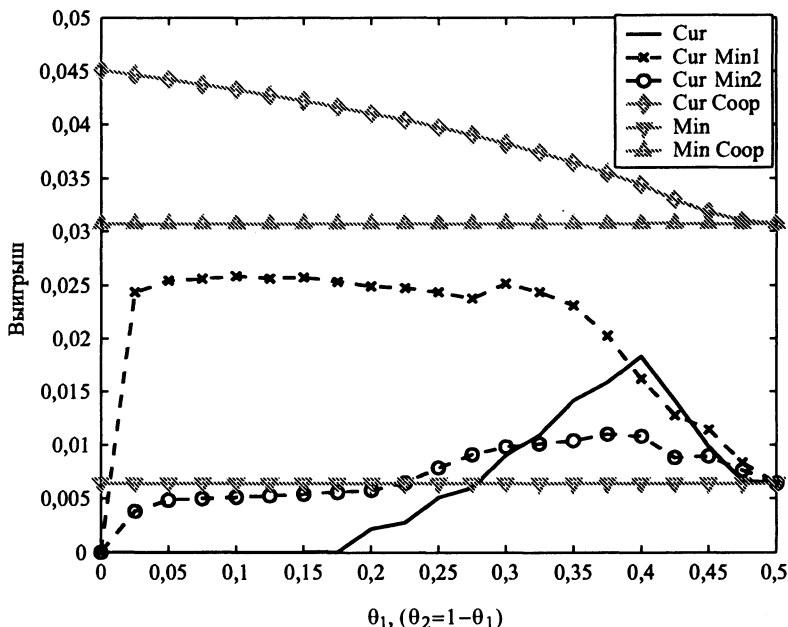


Рис. 5. Влияние степени природной изменчивости: “Cur” – текущая информация; “Cur-Min” – асимметричная информация, когда игрок 1 обладает текущей информацией, а игрок 2 – минимальной; “Cur-Coop” – кооперация с текущей информацией; “Min” – минимальная информация; “Min-Coop” – кооперация с минимальной информацией

ствительно, при высокой вариабельности  $\theta$  фактически весь ресурс направляется в один из двух потоков, и это приводит к снижению средней удельной стоимости лова (см. 7.2.1). Поскольку эта кооперативная игра полностью симметрична, средние годовые доходы игроков одинаковы.

Похоже, что возрастание дохода в некооперативной игре при уменьшении  $\theta_1$  от 0,5 до 0,4 может иметь такое же объяснение. Однако при малых  $\theta_1$  роль соперничества (особенно при полной информации “Cur”) становится доминирующей. Действительно, если весь ресурс попадает в один поток, соответствующий флот может практически полностью выловить его при сравнительно низких затратах.

**7.3.5. Переход игроков от соперничества к кооперации.** В § 6 был описан простой способ включения “кооперации” в модель игры. На рис. 6 представлена зависимость среднего выигрыша от “степени сотрудничества” для случаев минимальной и полной информации.

Нулевая степень кооперации соответствует чисто соперническому поведению, когда каждый игрок максимизирует только свой доход, в то время, как степень кооперации 1 означает, что оба игрока максимизируют полный доход. При промежуточных степенях кооперации каждый игрок максимизирует выпуклую комбинацию своего дохода и дохода соперника.

Точнее, коэффициенты кооперации определяются через степень кооперации, обозначаемую  $s$ , следующим образом:  $c_{\alpha\alpha} = 1 - c/2$ ,  $c_{\alpha\beta} = c/2$ ,  $c_{\beta\alpha} = c/2$ ,  $c_{\beta\beta} = 1 - c/2$ . Таким образом, когда  $s$  изменяется от 0 до 1, коэффициенты кооперации изменяются от “отсутствия кооперации” до “полной кооперации”.

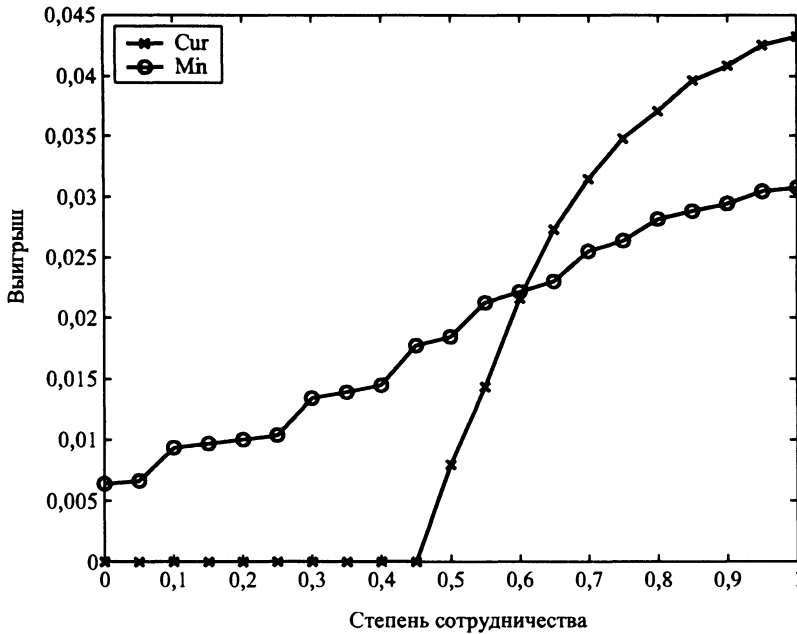


Рис. 6. Влияние степени сотрудничества при текущей и минимальной информации: “Cur” – текущая информация, “Min” – минимальная информация

Из рис. 6 видно, что дополнительная информация выгодна для игроков только в том случае, если уровень кооперации достаточно высок. При низком уровне кооперации, и особенно при соперничестве, дополнительная информация приводит к критическому сокращению ресурса и нулевому доходу.

### § 8. Приложение: общие сведения об играх

Напомним, что неантогонистическая игра двух лиц определяется совокупностью  $g = \langle D^\alpha, D^\beta, v^\alpha, v^\beta \rangle$ , где  $D^\alpha$  и  $D^\beta$  – пространства ходов (решений) игроков  $\alpha$  и  $\beta$ , а  $v^\alpha$  и  $v^\beta$  – их функции выигрыша соответственно. Точнее,  $v^\alpha, v^\beta : D^\alpha \times D^\beta \rightarrow \mathbf{R}$ ,  $v^\alpha(p^\alpha, p^\beta)$  и  $v^\beta(p^\alpha, p^\beta)$  – выигрыши игроков  $\alpha$  и  $\beta$ , если их решения равны  $p^\alpha$  и  $p^\beta$  соответственно.

Классическая задача построения оптимальных решений в такой игре состоит в нахождении точки равновесия по Нэшу, а именно, такой пары решений  $\langle \hat{p}^\alpha, \hat{p}^\beta \rangle$ , для которой отклонение каждого игрока от его оптимальной стратегии в одностороннем порядке приведет только к его потерям, т.е.

$$\begin{aligned} \forall p^\alpha \in D^\alpha \quad v^\alpha(p^\alpha, \hat{p}^\beta) &\leq v^\alpha(\hat{p}^\alpha, \hat{p}^\beta), \\ \forall p^\beta \in D^\beta \quad v^\alpha(\hat{p}^\alpha, p^\beta) &\leq v^\alpha(\hat{p}^\alpha, \hat{p}^\beta). \end{aligned}$$

Обозначим  $\pi^\alpha : D^\beta \rightarrow P(D^\alpha)$  многозначное отображение<sup>1</sup>, определяющее множество оптимальных откликов игрока  $\alpha$  на решение игрока  $\beta$ . Точнее, пусть

$$\pi^\alpha(p^\beta) = \arg \max_{p^\alpha} (p^\alpha, p^\beta).$$

<sup>1</sup> Здесь и далее  $P(\mathcal{X})$  обозначает множество всех подмножеств множества  $\mathcal{X}$ .

Аналогично определим  $\pi^\beta : \mathcal{D}^\alpha \rightarrow \mathcal{P}(\mathcal{D}^\beta)$ :

$$\pi^\beta(p^\alpha) = \arg \max_{p^\beta} (p^\alpha, p^\beta).$$

Очевидно, пара  $\langle \hat{p}^\alpha, \hat{p}^\beta \rangle$  образует точку равновесия по Нэшу, если

$$\hat{p}^\alpha \in \pi^\alpha(\hat{p}^\beta), \quad \hat{p}^\beta \in \pi^\beta(\hat{p}^\alpha).$$

Нередко точка равновесия по Нэшу  $\langle \hat{p}^\alpha, \hat{p}^\beta \rangle$  может быть построена путем следующей итерационной процедуры.

Пусть  $p_0^\alpha$  – некоторое произвольное решение игрока  $\alpha$  (начальное приближение). Пусть  $p_0^\beta$  – оптимальный отклик игрока  $\beta$  на решение  $p_0^\alpha$  ( $p_0^\beta \in \pi^\beta(p_0^\alpha)$ ). Аналогично, пусть  $p_1^\alpha$  – оптимальный отклик на  $p_0^\beta$  ( $p_1^\alpha \in \pi^\alpha(p_0^\beta)$ ), и т.д.:

$$p_n^\alpha \in \pi^\alpha(p_{n-1}^\beta), \quad p_n^\beta \in \pi^\beta(p_n^\alpha).$$

**Предложение 8.1.** Пусть  $\mathcal{D}^\alpha$  и  $\mathcal{D}^\beta$  – метрические пространства,  $\pi^\alpha$  и  $\pi^\beta$  – всюду определенные однозначные непрерывные отображения. Если существует хотя бы один из пределов

$$\lim_{n \rightarrow \infty} p_n^\alpha = \lim_{n \rightarrow \infty} (\pi^\alpha \circ \pi^\beta)^n(p_0^\alpha) = \hat{p}^\alpha$$

или

$$\lim_{n \rightarrow \infty} p_n^\beta = \lim_{n \rightarrow \infty} ((\pi^\beta \circ \pi^\alpha)^n \circ \pi^\beta)(p_0^\alpha) = \hat{p}^\beta,$$

то существует и второй, и пара  $\langle \hat{p}^\alpha, \hat{p}^\beta \rangle$  образует точку равновесия по Нэшу для игры  $g = \langle \mathcal{D}^\alpha, \mathcal{D}^\beta, v^\alpha, v^\beta \rangle$ .

**Доказательство.** Пусть существует  $\lim_{n \rightarrow \infty} p_n^\alpha = \hat{p}^\alpha$ . Докажем, что второй предел  $\lim_{n \rightarrow \infty} p_n^\beta$  также существует и равен  $\hat{p}^\beta$ . Действительно,

$$\hat{p}^\beta = \lim_{n \rightarrow \infty} p_n^\beta = \lim_{n \rightarrow \infty} \pi^\beta(p_n^\alpha) = \pi^\beta \left( \lim_{n \rightarrow \infty} p_n^\alpha \right) = \pi^\beta(\hat{p}^\alpha).$$

Аналогично, если существует  $\lim_{n \rightarrow \infty} p_n^\beta = \hat{p}^\beta$ , то существует и  $\lim_{n \rightarrow \infty} p_n^\alpha$ , и

$$\hat{p}^\alpha = \lim_{n \rightarrow \infty} p_n^\alpha = \pi^\alpha(\hat{p}^\beta).$$

Таким образом, пара  $\langle \hat{p}^\alpha, \hat{p}^\beta \rangle$  удовлетворяет условиям  $\hat{p}^\alpha \in \pi^\alpha(\hat{p}^\beta)$  и  $\hat{p}^\beta \in \pi^\beta(\hat{p}^\alpha)$ , т.е. образует точку равновесия по Нэшу.  $\blacktriangle$

Пусть теперь результат игры зависит не только от решений игроков, но и от некоторого случайного параметра  $\nu$  (имеющего известное распределение) из пространства  $\mathcal{N}$ . А именно, пусть  $v^\alpha, v^\beta : \mathcal{D}^\alpha \times \mathcal{D}^\beta \times \mathcal{N} \rightarrow \mathbb{R}$ . Поскольку параметр  $\nu$  игрокам не известен, то естественно ставить задачу о построении оптимальных решений как задачу оптимизации средних выигрышей

$$V^\alpha(p^\alpha, p^\beta) = E_\nu v^\alpha(p^\alpha, p^\beta, \nu)$$

и

$$V^\beta(p^\alpha, p^\beta) = E_\nu v^\beta(p^\alpha, p^\beta, \nu).$$

Мы приходим к игре  $\langle D^\alpha, D^\beta, V^\alpha, V^\beta \rangle$ , которая отличается от исходной игры  $\langle D^\alpha, D^\beta, v^\alpha, v^\beta \rangle$  только функциями выигрыша ( $V^\alpha, V^\beta$  вместо  $v^\alpha, v^\beta$ ).

Ситуация несколько усложняется, если игрокам доступна некоторая информация о случайном параметре  $\nu$ . А именно, пусть игрокам  $\alpha$  и  $\beta$  доступны результаты неточных измерений параметра  $\nu$ , т.е. реализации случайных элементов  $\xi^\alpha \in \mathcal{X}^\alpha$ ,  $\xi^\beta \in \mathcal{X}^\beta$ . Взаимосвязь между  $\nu$ ,  $\xi^\alpha$  и  $\xi^\beta$  описывается заданным совместным распределением на пространстве  $\mathcal{N} \times \mathcal{X}^\alpha \times \mathcal{X}^\beta$ .

Пусть, например, игроку  $\alpha$  доступно наблюдение случайного элемента  $\xi^\alpha$ . Тогда его решение  $p^\alpha$  может зависеть от  $\xi^\alpha$ , т.е. определяется как  $p^\alpha = P^\alpha(\xi^\alpha)$ , где  $P^\alpha$  – некоторая решающая стратегия игрока  $\alpha$  – измеримое отображение из  $\mathcal{X}^\alpha$  в  $D^\alpha$ . Аналогично, игрок  $\beta$  может принимать свои решения на основании наблюдений  $\xi^\beta$ , т.е.  $p^\beta = P^\beta(\xi^\beta)$ .

При использовании стратегий  $P^\alpha$  и  $P^\beta$  средние выигрыши игроков определяются как

$$\begin{aligned} V^\alpha(P^\alpha, P^\beta) &= E_{(\nu, \xi^\alpha, \xi^\beta)} v^\alpha(P^\alpha(\xi^\alpha), P^\beta(\xi^\beta), \nu), \\ V^\beta(P^\alpha, P^\beta) &= E_{(\nu, \xi^\alpha, \xi^\beta)} v^\beta(P^\alpha(\xi^\alpha), P^\beta(\xi^\beta), \nu). \end{aligned}$$

Таким образом, в ситуации с неточными наблюдениями случайного параметра мы приходим к новой игре

$$G = \langle \tilde{D}^\alpha, \tilde{D}^\beta, V^\alpha, V^\beta \rangle.$$

Здесь  $\tilde{D}^\alpha$  и  $\tilde{D}^\beta$  – функциональные пространства, а именно, пространства измеримых отображений вида  $\mathcal{X}^\alpha \rightarrow D^\alpha$  и  $\mathcal{X}^\beta \rightarrow D^\beta$  соответственно, а  $V^\alpha$  и  $V^\beta$  – функционалы, определенные на пространстве  $\tilde{D}^\alpha \times \tilde{D}^\beta$ .

Проблема построения точки равновесия по Нэшу для такой игры принципиально более сложная, чем для исходной игры  $g$ . Действительно, в исходной игре стратегии, как правило, являются элементами конечномерных линейных пространств (или даже конечных множеств), в то время как в игре с неточными измерениями стратегии представляют из себя элементы функциональных пространств.

Однако в некоторых случаях (например, когда оба игрока получают одинаковую информацию) оказывается возможным свести построение оптимальных стратегий в игре  $G = \langle \tilde{D}^\alpha, \tilde{D}^\beta, V^\alpha, V^\beta \rangle$  к построению стратегий в играх типа  $\langle D^\alpha, D^\beta, v^\alpha, v^\beta \rangle$ .

Итак, пусть оба игрока получают результаты одного и того же наблюдения, т.е.  $\xi^\alpha = \xi^\beta = \xi \in \mathcal{X} = \mathcal{X}^\alpha = \mathcal{X}^\beta$ .

Для каждого фиксированного  $\xi \in \mathcal{X}$  рассмотрим функции условного выигрыша

$$\begin{aligned} V_\xi^\alpha(p^\alpha, p^\beta) &= E_{(\nu|\xi)} v^\alpha(p^\alpha, p^\beta, \nu), \\ V_\xi^\beta(p^\alpha, p^\beta) &= E_{(\nu|\xi)} v^\beta(p^\alpha, p^\beta, \nu), \end{aligned}$$

где  $E_{(\nu|\xi)}$  – операция условного математического ожидания по  $\nu$  относительно  $\xi$ .

Для  $\xi \in \mathcal{X}$  рассмотрим “условную” игру

$$G_\xi = \langle D^\alpha, D^\beta, V_\xi^\alpha, V_\xi^\beta \rangle,$$

отличающуюся от игры  $G$  лишь заменой распределения случайного элемента  $\nu$  на условное распределение  $\nu$  при фиксированном  $\xi$ .

**Предложение 8.2 (принцип Байеса).** Пусть для любого  $\xi \in \mathcal{X}$  существует точка равновесия по Нэшу  $\langle \hat{p}_\xi^\alpha, \hat{p}_\xi^\beta \rangle$  для игры  $G_\xi = \langle D^\alpha, D^\beta, V_\xi^\alpha, V_\xi^\beta \rangle$ , и пусть

отображения  $\hat{P}^\alpha$  и  $\hat{P}^\beta$ , определенные как  $\hat{P}^\alpha(\xi) = \hat{p}_\xi^\alpha$  и  $\hat{P}^\beta(\xi) = \hat{p}_\xi^\beta$ , измеримы. Тогда пара  $\langle \hat{P}_\xi^\alpha, \hat{P}_\xi^\beta \rangle$  образует точку равновесия по Нэшу в игре  $G = \langle \tilde{D}^\alpha, \tilde{D}^\beta, V^\alpha, V^\beta \rangle$ .

Доказательство. Пусть  $P^\alpha : \mathcal{X} \rightarrow \mathcal{D}^\alpha$  и  $P^\beta : \mathcal{X} \rightarrow \mathcal{D}^\beta$  – произвольные измеримые отображения. Поскольку операция математического ожидания  $E_{(\nu, \xi)}$  может быть записана как композиция  $E_{(\nu, \xi)} = E_\xi E_{(\nu | \xi)}$  математических ожиданий по маргинальному распределению  $\xi$  и условному распределению  $\xi$  относительно  $\nu$ , то

$$V^\alpha(P^\alpha, P^\beta) = E_\xi E_{(\nu | \xi)} v^\alpha(P^\alpha(\xi), P^\beta(\xi), \nu),$$

$$V^\beta(P^\alpha, P^\beta) = E_\xi E_{(\nu | \xi)} v^\beta(P^\alpha(\xi), P^\beta(\xi), \nu).$$

Тогда

$$V^\alpha(P^\alpha, P^\beta) = E_\xi V_\xi^\alpha(P^\alpha(\xi), P^\beta(\xi)),$$

$$V^\beta(P^\alpha, P^\beta) = E_\xi V_\xi^\beta(P^\alpha(\xi), P^\beta(\xi)).$$

Рассмотрим теперь  $V^\alpha(P^\alpha, \hat{P}^\beta)$ . Согласно определению,  $V_\xi^\alpha(P^\alpha(\xi), \hat{p}_\xi^\beta) \leq V_\xi^\alpha(\hat{p}_\xi^\alpha, \hat{p}_\xi^\beta)$  для всех  $\xi \in \mathcal{X}$ . Отсюда получаем

$$V^\alpha(P^\alpha, \hat{P}^\beta) = E_\xi V_\xi^\alpha(P^\alpha(\xi), \hat{P}^\beta(\xi)) \leq E_\xi V_\xi^\alpha(\hat{P}^\alpha(\xi), \hat{P}^\beta(\xi)) = V^\alpha(\hat{P}^\alpha, \hat{P}^\beta).$$

Таким образом,  $V^\alpha(P^\alpha, \hat{P}^\beta) \leq V^\alpha(\hat{P}^\alpha, \hat{P}^\beta)$ . Аналогично доказывается, что  $V^\beta(\hat{P}^\alpha, P^\beta) \leq V^\beta(\hat{P}^\alpha, \hat{P}^\beta)$ , т.е.  $\langle \hat{P}^\alpha, \hat{P}^\beta \rangle$  – точка равновесия по Нэшу в игре  $G$ .  $\blacktriangle$

Таким образом, в игре с одинаковой информацией  $G$  пара оптимальных стратегий  $\langle \hat{P}^\alpha, \hat{P}^\beta \rangle$  может быть построена “поточечно”, путем рассмотрения игр  $G_\xi$  для всех возможных значений  $\xi \in \mathcal{X}$ . При этом оптимальные решения для условной игры могут строиться как это было описано для игры  $g$  в предложении 8.1.

Наконец, для игры  $G$  в общем случае построение оптимальных стратегий также может выполняться “поточечно” с использованием более сложной итерационной процедуры.

Итак, рассмотрим общую игру

$$G = \langle \tilde{D}^\alpha, \tilde{D}^\beta, V^\alpha, V^\beta \rangle \quad (14)$$

для случая различных наблюдений  $\xi^\alpha \in \mathcal{X}^\alpha$  и  $\xi^\beta \in \mathcal{X}^\beta$  игроков  $\alpha$  и  $\beta$  соответственно.

Для некоторого фиксированного значения  $\xi^\alpha \in \mathcal{X}^\alpha$  рассмотрим условный выигрыш

$$V_{\xi^\alpha}^\alpha(p^\alpha, P^\beta) = E_{(\nu, \xi^\beta | \xi^\alpha)} v^\alpha(p^\alpha, P^\beta(\xi^\beta), \nu).$$

Аналогично определим  $V_\xi^\alpha$ :

$$V_{\xi^\beta}^\beta(P^\alpha, p^\beta) = E_{(\nu, \xi^\alpha | \xi^\beta)} v^\beta(P^\alpha(\xi^\alpha), p^\beta, \nu).$$

Пусть  $\pi_{\xi^\alpha}^\alpha : \tilde{D}^\beta \rightarrow \mathcal{P}(\mathcal{D}^\alpha)$  – многозначное отображение, определяющее множество оптимальных реакций  $p^\alpha$  игрока  $\alpha$  на стратегию  $P^\beta$  игрока  $\beta$  при условии наблю-

дения  $\xi^\alpha$ . А именно, пусть

$$\pi_{\xi^\alpha}^\alpha(P^\beta) = \arg \max_{p^\alpha} V_{\xi^\alpha}^\alpha(p^\alpha, P^\beta).$$

Аналогично определим  $\pi_{\xi^\beta}^\beta : \tilde{D}^\alpha \rightarrow \mathcal{P}(\mathcal{D}^\beta)$  для всех  $\xi^\beta \in \mathcal{X}^\beta$ .

Рассмотрим следующую итерационную процедуру. Выберем произвольную стратегию игрока  $\alpha$  – измеримое отображение  $P_0^\alpha : \mathcal{X}^\alpha \rightarrow \mathcal{D}^\alpha$ . Построим оптимальную реакцию игрока  $\beta$  следующим образом. Для любого  $\xi^\beta \in \mathcal{X}^\beta$  выберем  $P_0^\beta(\xi^\beta) \in \pi_{\xi^\beta}^\beta(P_0^\alpha)$ . Аналогично определим  $P_1^\alpha : \mathcal{X}^\alpha \rightarrow \mathcal{D}^\alpha$  так, чтобы  $P_1^\alpha(\xi^\alpha) \in \pi_{\xi^\alpha}^\alpha(P_0^\beta)$  для всех  $\xi^\alpha \in \mathcal{X}^\alpha$ , и т.д.:

$$\begin{aligned} P_n^\alpha(\xi^\alpha) &\in \pi_{\xi^\alpha}^\alpha(P_{n-1}^\beta) \quad \forall \xi^\alpha \in \mathcal{X}^\alpha, \\ P_n^\beta(\xi^\beta) &\in \pi_{\xi^\beta}^\beta(P_n^\alpha) \quad \forall \xi^\beta \in \mathcal{X}^\beta. \end{aligned}$$

**Предложение 8.3.** Пусть  $\mathcal{D}^\alpha$  и  $\mathcal{D}^\beta$  – метрические пространства,  $\pi_{\xi^\alpha}^\alpha$  и  $\pi_{\xi^\beta}^\beta$  – отображения

$$\begin{aligned} \pi_{\xi^\alpha}^\alpha : \tilde{D}^\beta &\rightarrow \mathcal{P}(\mathcal{D}^\alpha) \quad \pi_{\xi^\alpha}^\alpha(P^\beta) = \arg \max_{p^\alpha} V_{\xi^\alpha}^\alpha(p^\alpha, P^\beta) \quad \forall \xi^\alpha \in \mathcal{X}^\alpha, \\ \pi_{\xi^\beta}^\beta : \tilde{D}^\alpha &\rightarrow \mathcal{P}(\mathcal{D}^\beta) \quad \pi_{\xi^\beta}^\beta(P^\alpha) = \arg \max_{p^\beta} V_{\xi^\beta}^\beta(p^\beta, P^\alpha) \quad \forall \xi^\beta \in \mathcal{X}^\beta, \end{aligned}$$

всюду определенные однозначные равномерно непрерывные относительно  $\xi^\alpha \in \mathcal{X}^\alpha$  и  $\xi^\beta \in \mathcal{X}^\beta$  соответственно, такие, что для всех фиксированных измеримых отображений  $P^\alpha \in \tilde{D}^\alpha$  и  $P^\beta \in \tilde{D}^\beta$  отображения  $\pi_{\xi^\alpha}^\alpha(P^\beta)$  и  $\pi_{\xi^\beta}^\beta(P^\alpha)$  измеримы по  $\xi^\alpha$  и  $\xi^\beta$  соответственно. Если существует один из пределов  $\lim_{n \rightarrow \infty} P_n^\alpha = \hat{P}^\alpha$  или  $\lim_{n \rightarrow \infty} P_n^\beta = \hat{P}^\beta$ , то существует и другой, и  $\langle \hat{P}^\alpha, \hat{P}^\beta \rangle$  образуют точку равновесия по Нэшу для игры  $G$ .

**Доказательство.** Рассмотрим отображение  $\hat{\pi}^\alpha$ , определяемое как  $\hat{\pi}^\alpha(P^\beta)(\xi^\alpha) = \pi_{\xi^\alpha}^\alpha(P^\beta)$ . По условию для любого измеримого  $P^\beta \in \tilde{D}^\beta$  отображение  $\hat{\pi}^\alpha(P^\beta) : \mathcal{X}^\alpha \rightarrow \mathcal{D}^\alpha$  также является измеримым, т.е.  $\hat{\pi}^\alpha : \tilde{D}^\beta \rightarrow \tilde{D}^\alpha$ . Аналогично определим  $\hat{\pi}^\beta : \tilde{D}^\alpha \rightarrow \tilde{D}^\beta$ :  $\hat{\pi}^\beta(P^\alpha)(\xi^\beta) = \pi_{\xi^\beta}^\beta(P^\alpha)$ .

Пространства  $\tilde{D}^\alpha$  и  $\tilde{D}^\beta$  являются метрическими с равномерной метрикой, определяемой метрикой пространств  $\mathcal{D}^\alpha$  и  $\mathcal{D}^\beta$  соответственно. Тогда из условия равномерно непрерывности отображений  $\pi_{\xi^\alpha}^\alpha$  и  $\pi_{\xi^\beta}^\beta$  сразу следует, что отображения  $\hat{\pi}^\alpha$  и  $\hat{\pi}^\beta$  непрерывны.

Действительно, равномерно непрерывность функций  $\pi_{\xi^\alpha}^\alpha$ ,  $\xi^\alpha \in \mathcal{X}^\alpha$ , в точке  $P^\beta \in \tilde{D}^\beta$  означает, что

$$\forall \varepsilon > 0 \exists \delta \forall P \in \tilde{D}^\beta \left[ d_{\tilde{D}^\beta}(P, P^\beta) < \delta \Rightarrow d_{\mathcal{D}^\alpha}(\pi^\alpha(P), \pi^\alpha(P^\beta)) < \varepsilon \right].$$

Это эквивалентно, поскольку  $d_{\tilde{D}^\alpha}(P_1, P_2) = \sup_{\xi^\alpha} d_{\mathcal{D}^\alpha}(P_1, P_2)$ , непрерывности отображения  $\hat{\pi}_{\xi^\alpha}^\alpha$  в точке  $P^\beta \in \tilde{D}^\beta$ :

$$\forall \varepsilon > 0 \exists \delta \forall P \in \tilde{D}^\beta \left[ d_{\tilde{D}^\beta}(P, P^\beta) < \delta \Rightarrow d_{\tilde{D}^\alpha}(\hat{\pi}^\alpha(P), \hat{\pi}^\alpha(P^\beta)) < \varepsilon \right].$$

Таким образом, выполнены условия предложения 8.1 (с очевидной заменой  $\mathcal{D}^\alpha \rightarrow \tilde{D}^\alpha$ ,  $\mathcal{D}^\beta \rightarrow \tilde{D}^\beta$ ,  $\pi^\alpha \rightarrow \hat{\pi}^\alpha$ ,  $\pi^\beta \rightarrow \hat{\pi}^\beta$ ,  $v^\alpha \rightarrow V^\alpha$ ,  $v^\beta \rightarrow V^\beta$ ).  $\blacktriangle$

## СПИСОК ЛИТЕРАТУРЫ

1. *Geanakoplos J.* Common Knowledge // Handbook of Game Theory with Economic Applications. Vol. II, Chap. 40. Amsterdam: Elsevier, 1984.
2. *McKelvey R.* The Split-Stream Harvesting Game. Part I: Mathematical Analysis // NCAR Technical Note. National Center for Atmospheric Research. Boulder, Colorado, 2001. P. 1–42.
3. *McKelvey R., Gripe G.* The Split-Stream Harvesting Game. Part II: Numerical and Simulation Studies // NCAR Technical Note. National Center for Atmospheric Research. Boulder, Colorado, 2001. P. 43–57.
4. *Clark C.W.* Restricted Access to Common-Property Fishery Resources: a Game-Theoretic Analysis // Dynamic Optimization and Mathematical Economics. New York: Plenum, 1980. P. 117–132.
5. *Levhari D., Mirman L.J.* The Great Fish War: an Example Using a Dynamic Cournot-Nash Solution // Bell J. Economics. 1980. V. 11. P. 322–344.
6. *Hernandes-Lerma O., Lasserre J.B.* Discrete-time Markov Control Processes. Basic Optimality Criteria. New York: Springer, 1996.

*Голубцов Петр Викторович*  
Московский государственный университет  
им. М.В. Ломоносова  
p\_v\_g@mail.ru  
*Любецкий Василий Александрович*  
Институт проблем передачи информации РАН  
lyubetsk@iitp.ru

Поступила в редакцию  
29.04.2002