В.И. Данилов[1]
ЦЭМИ РАН, Москва

За пределами классической рациональности: двухэтапная рационализация

Многочисленные данные свидетельствуют о том, что поведение (выбор) экономических агентов часто не удовлетворяет строгим требованиям рациональности. Одно из возможных объяснений состоит в том, что выбор производится в два (или больше) последовательных этапа. Сначала из исходного меню выделяется некоторое его подмножество, а затем из него выбирается наилучший (в соответствии с некоторым линейным порядком) элемент. Результирующий выбор уже не является рациональным в классическом смысле. Тем не менее в ряде случаев он обладает некоторыми характерными свойствами, позволяющими отличать его от произвольного нерационального поведения. В настоящей статье дается обзор современных работ по двухэтапному выбору. В ряде случаев удается дать исчерпывающее описание свойств двухэтапного выбора в зависимости от свойств рациональности первого этапа.

Ключевые слова: *функции выбора; условия наследования, отбрасывания, согласия; фильтр; предпочтение; аксиома выявленного предпочтения; функции Плотта.*

Классификация JEL: D71.

Введение

Классическая рациональность поведения (в данном случае – выбора одной альтернативы среди нескольких) понимается как выбор, производимый в соответствии с линейным порядком (или функцией полезности). Предполагается, что индивид, принимающий решение, обладает встроенным линейным порядком и из каждого предъявления (меню) выбирает наилучшую альтернативу (т.е. альтернативу с наивысшей полезностью). Так как его внутренняя полезность, или предпочтение, – явление, прямо не наблюдаемое, были найдены простые условия, накладываемые на его выбор (например, выявленное предпочтение или независимость от посторонних альтернатив; подробнее об этом в разд. 2), которые эквивалентны рациональному поведению.

Экспериментально было показано, что люди[2] очень часто ведут себя нерационально, проявляя нетранзитивности и другие отклонения от совершенной рациональности. Когда эти отклонения не слишком велики, говорят об ограниченной рациональности. Экспериментально были выявлены причины такой ограниченной рациональности (главная из них – сложность (и связанная с этим дороговизна) анализа всех альтернатив из данного меню) и эвристические принципы преодоления сложности анализа, например, с помощью предварительного сокращения списка выбираемых альтернатив. Заметим, что подобные процедуры часто применяются при голосовании (когда на первом этапе отбираются два наиболее популярных кандидата), или на музы-

[1] Автор благодарит Ф.Л. Зака и рецензента за полезные замечания.

[2] И не только люди. В работе (Manzini, Mariotti, 2007) приводится ссылка на эксперименты, указывающие, что серые сойки ведут себя даже менее рационально, чем люди.

Журнал НЭА, №2 (26), 2015, с. 12–35

кальных конкурсах, или спортивных состязаниях (когда сначала отбираются финалисты, а затем уже среди них – победитель).

До недавнего времени такой подход к объяснению нерациональности оставался эвристическим. И только в последние 7–10 лет поиски рационального объяснения нерационального, на первый взгляд, выбора вышли на уровень строгих утверждений. Практически сразу появилось несколько статей, где обсуждались различные механизмы и даны их аксиоматические характеризации (например, в статье (Cherepanov et al., 2013) приведен список источников, который содержит около 70 наименований). Чаще всего речь идет о последовательной рационализации выбора. Имеется в виду, что (окончательный) выбор производится на протяжении нескольких последовательных этапов. Далее мы обсудим несколько результатов о характеризации, ограничиваясь случаем двухэтапного (или двухступенчатого) выбора. Предполагается, что выбор на обоих этапах происходит в некотором смысле рационально. Так как рациональность многозначного выбора, происходящего на первом этапе, может пониматься по-разному, мы обсуждаем встречающиеся возможности в отдельных разделах.

Указанный подход напоминает рассмотренный в (Зак, 2014а, 2014б) обзор работ по теории игр, посвященных моделированию принятия решений в условиях неполной рациональности участников. Этот подход тоже был мотивирован желанием средствами математического моделирования объяснить поведение людей, не подчиняющееся правилам рационального выбора (первым признаком которого часто служит нарушение знаменитой слабой аксиомы выявленного предпочтения – WARP). Там тоже принятие решений осуществляется в два этапа: сначала выбирается меню, а затем – альтернатива из этого меню.

Однако имеется и принципиальная разница. Там индивид (лицо, принимающее решения, ЛПР) на первом этапе выбирает меню, выбирает свободно, из соображений привлекательности этого меню для последующих целей. Поэтому основной объект изучения – отношения предпочтения между различными меню, гиперотношения[3], в терминологии А. Малишевского[4]. У нас же выбор (точнее – сужение) меню происходит независимо от воли ЛПР, т.е. это не результат его свободного выбора, хотя принимаемое решение может быть связано с бессознательными процессами его психики.

С математической точки зрения речь идет о композиции двух функций выбора. Начиная с классической работы (Айзерман, Малишевский, 1981) (см. также книгу (Айзерман, Алескеров, 1990)), известно, что рациональность при композиции сильно портится (если вообще остается). Интуитивно понятно, что рациональность композиции получается только при достаточно сильной взаимной увязанности функций выбора, действующих на первом и втором этапах.

[3] Отношения между меню будем называть гиперотношениями, чтобы не путать с отношениями предпочтения между альтернативами.

[4] В этой связи можно указать на работу (Danilov et al., 2014), где прослеживаются тесные связи между гиперотношениями и функциями выбора.

Журнал НЭА, №2 (26), 2015, с. 12–35

Исследователи из школы М.А. Айзермана много сделали в этом направлении. Не претендуя на полноту списка, укажем работы (Aizerman, Malishevski, 1986; Aleskerov, Cinar, 2008; Айзерман, Алескеров, 1990).

В данном обзоре будет исследован случай, когда нет никакой особой связи между выборами на первом и втором этапах. Дело в том, что на этих этапах, как правило, действуют разные механизмы (см. разд. 9). Тем не менее в ряде случаев удается дать характеризацию двухэтапной рациональности, причем в довольно простых терминах. Это частично отвечает на вопрос о том, до какой степени ослабляется рациональность при двухэтапном выборе и насколько широкий класс функций выбора при этом получается.

После изложения кратких сведений из теории классической рациональности мы приводим три главные аксиомы рациональности многозначного выбора: наследования, отбрасывания и согласия. Затем подробно рассматриваем случаи, когда фильтр, отбирающий финалистов, удовлетворяет тем или иным комбинациям основных аксиом. В последнем разделе обсуждаются мотивировки фильтрации на первом этапе выбора.

1. Классический рациональный выбор

Функцией выбора (или ФВ)[5] называется отображение $f: 2^X \to 2^X$ такое, что $f(A) \subseteq A$ для любого $A \subseteq X$. Здесь X – множество всех альтернатив, предполагаемое для простоты конечным. Непустые подмножества X называются *меню. Точной* называется ФВ с одноэлементным исходом, т.е. когда $f(A)$ состоит из единственного элемента для любого меню A. В этом случае вместо $f(A) = \{a\}$ мы запишем $f(A) = a$. В дальнейшем окончательный выбор f считается точным, а промежуточный или предварительный выбор (фильтр) – не обязательно.

Классические работы по данной теме начинали с исследования точного выбора. Здесь не было сомнений, какой выбор считать рациональным. Конечно выбор, направляемый линейным порядком. Пусть «<» – линейный порядок на множестве X, т.е. транзитивное, полное иррефлексивное бинарное отношение. Мы говорим, что выбор *направляется* линейным порядком <, если из каждого меню A выбирается его наибольший (относительно <) элемент. Это описание выбора со стороны механизма, или внутренней мотивировки. С внешней стороны (т.е. для наблюдателя) описание процесса выбора тоже хорошо известно и называется WARP, *слабой аксиомой выявленного предпочтения* Самуэльсона:

WARP: *если* $x = f(S)$, $y \in S$ (*и отлично от* x) *и* $x \in T$, *то* $f(T) \neq y$.

В 1950 г. Нэш сформулировал другое условие рациональности (точного) выбора, которое назвал *независимостью от посторонних альтернатив*:

IIA: если $f(A) \in B \subseteq A$, то $f(B) = f(A)$.

Следующая теорема является классической в теории выбора.

[5] Функция выбора – неудачное название, в данном случае следовало бы говорить о соответствии.

Журнал НЭА, №2 (26), 2015, с. 12–35

Теорема. *Пусть f – точная ФВ (на конечном X). Эквивалентны следующие три утверждения:*

1) *f удовлетворяет WARP;*
2) *f удовлетворяет IIA;*
3) *f вполне рациональна, т.е. выбирает наибольший элемент относительно некоторого линейного порядка.*

Доказательство. $1) \Rightarrow 2)$. Пусть $x = f(A) \in B \subseteq A$. Тогда для любого y из B, отличного от x, $y \neq f(B)$ (в силу WARP). Значит, $f(B)$ может быть равно только x.

$2) \Rightarrow 3)$. Пусть f удовлетворяет IIA. Тогда легко построить линейный порядок, или ранжирование <, с помощью следующей простой операции пилинга. Старшим элементом объявляется $x_1 = f(X)$, следующим – $x_2 = f(X - \{x_1\})$ и т.д. Почему выбор, направляемый этим порядком <, совпадает с f? Проверим, например, что если меню B содержит x_1, то $f(B) = x_1$ (остальное доказывается аналогично). Это следует из IIA, примененного к $A = X$.

$3) \Rightarrow 1)$. Очевидно, так как $y < x = f(S)$, а $f(T) \geq x$. ■

Таким образом, мы видим, что вопросов с пониманием и характеризацией рациональности однозначного выбора нет. Другое дело – для многозначного выбора. Тут рациональность можно понимать по-разному. Можно идти от механизма выбора. Например, считать выбор сильно рациональным, если он максимизирует слабый порядок; просто рациональным, если максимизирует предпорядок; и слабо рациональным, если выбирает недоминируемые элементы относительно некоторого бинарного отношения (вдобавок иррефлексивного или асимметричного). А можно идти от аксиом, ослабляя, например, аксиому IIA (точнее, приспосабливая ее к многозначности). Делать это можно разными способами[6]. Укажем два самых естественных, обозначая многозначную ФВ как F. Самое бесхитростное обобщение называется аксиомой отбрасывания.

Аксиома отбрасывания (О). *Пусть $F(A) \subseteq B \subseteq A$, тогда $F(B) = F(A)$.*

Немного иное прочтение аксиомы IIA приводит к аксиоме наследования.

Аксиома наследования (Н). *Пусть x выбирается из A и принадлежит меньшему множеству $B \subseteq A$, тогда x выбирается* и в B. *То есть $F(A) \cap B \subseteq F(B)$, когда $B \subseteq A$.*

В каком-то смысле это самая важная аксиома в теории рациональности; без нее трудно что-либо делать. Имеется несколько эквивалентных формулировок аксиомы **Н**, из которых отметим симметричную формулировку: $F(A \cup B) \subseteq F(A) \cup F(B)$.

Наконец, упомянем так называемую аксиому согласия.

Аксиома согласия (С). $F(A) \cap F(B) \subseteq F(A \cup B)$.

Формально она напоминает симметричную формулировку **Н**, поэтому интересно исследовать, а что будет, если потребовать и то,

---

[6] См. (Manzini, Mariotti, 2007): «В то время как есть один способ быть рациональным, имеется много способов быть нерациональным».

Журнал НЭА, №2 (26), 2015, с. 12–35

и другое. Ответ известен (Айзерман, Алескеров, 1990) – теорема Сена – это в точности слаборациональные ФВ, т.е. ФВ, рационализуемые бинарными отношениями.

Аксиому согласия трудно рассматривать как самостоятельное требование рациональности. Обычно она встречается в комбинации с условием наследования и обеспечивает бинарность выбора. Приведем один простой результат. Пусть f – точная ФВ. *Базисным отношением*, ассоциированным с f, называется отношение $<_f$, которое определяется так: $x <_f y$, если $f(x, y) = y$. Очевидно, что это отношение – полное и асимметричное. В остальном оно может быть совершенно произвольным и иметь циклы. Если все же оно ациклично, то вместе с аксиомой согласия ацикличность $<_f$ гарантирует рациональность ФВ f. Этот результат дает основание для простейшей таксономии нарушения рациональности. По мнению (Manzini, Mariotti, 2007, 2012), циклы в попарном выборе и нарушение условия согласия **C** отражают различные поведенческие аспекты.

2. Двухэтапный выбор

Начнем с того, что нарушение рациональности в классическом смысле еще не свидетельствует о полной невменяемости индивида, принимающего решение. Оно может иметь вполне разумное объяснение, основанное на обстоятельствах, скрытых от внешнего наблюдателя. Поясним это примером, взятым из пионерской работы (Manzini, Mariotti, 2007) и проясняющим суть двухэтапного выбора.

Пример. Представим, что у ЛПР есть три дороги, чтобы добраться на работу: короткая, средняя и длинная. Многочисленные наблюдения над выбором дороги (предполагается, что иногда некоторые дороги закрываются на ремонт и становятся недоступны) опытным путем была выявлена следующая закономерность. Когда доступны все три дороги, ЛПР выбирает среднюю. Из короткой и средней он выбирает короткую, из средней и длинной – среднюю, но вот из короткой и длинной – длинную. Налицо цикл в бинарных (попарных) сравнениях, и тем самым – нарушение рациональности!

Однако ЛПР вполне разумно объясняет свое поведение. Дело в том, что движение по короткой дороге сильно затруднено пробками по сравнению с длинной дорогой. Пробки встречаются и на средней дороге, но сравнить их с пробками на короткой и длинной дорогах затруднительно – бывает по-разному. Таким образом, доступность длинной дороги сразу выключает из рассмотрения короткую. Мы видим в этом объяснении скрытую рациональность, хотя и ограниченную. Полная рациональность требовала бы привлечения среднего времени, затрачиваемого на дорогу. Но это в данном примере невозможно, причем не из-за неразумности ЛПР, а по объективным причинам – невозможности бесспорного сравнения продолжительности преодоления пробок на среднем пути.

Журнал НЭА, №2 (26), 2015, с. 12–35

Этот пример демонстрирует главную особенность рассматриваемого далее подхода. Выбор происходит в два этапа. На первом этапе из исходного меню A по тем или иным причинам отбирается сокращенный список $F(A)$. В нашем примере такой причиной были пробки на дорогах; в других случаях причины могут быть менее понятными (внешняя привлекательность, броскость альтернатив, сложность перебора всех потенциально доступных альтернатив, незнание некоторых возможностей и т.п.). Мы не будем рассматривать такие детали (немного об этом будет сказано в разд. 9) и сразу перейдем к формальной модели.

Предполагается, что окончательный (точный) выбор f происходит в два этапа. Сначала альтернативы из исходного меню A проходят отборочное сито, или *фильтр* F, в результате чего формируется укороченный список (тут используются разные термины типа *рассматриваемого* (consideration) или *выступающего* (salient) множества, фильтра *внимания* (attention) и т.п.; терминология еще не устоялась) $F(A) \subseteq A$. На втором этапе, пользуясь уже окончательной функцией выбора g, мы получаем исход $g(F(A)) = f(A)$.

Интересуясь рациональностью, естественно считать, что обе ФВ – F и g – в каком-то смысле рациональны. При этом F реально многозначная, так как в противном случае второй этап выбора является фиктивным. Доводочная же функция g должна быть точной (однозначной). Но так как она реально действует не на всех меню (а только на меню вида $F(A)$), мы не можем применить теорему о рациональности. Поэтому мы просто постулируем, что выбор на втором этапе направляется некоторым бинарным отношением $<$. Это означает, что отбирается альтернатива, которая доминирует над всеми другими.

Определение. Точная ФВ f *порождается* (или *объясняется*) парой $(F,<)$ (где F – вспомогательная ФВ, или *фильтр*, а $<$ – асимметричное бинарное отношение (*предпочтение*)), если $f(A) \in F(A)$ для любого меню A и для любого $a \in F(A) - \{f(A)\}$ выполняется $a < f(A)$.

Мы говорим также, что f *рационализируется с помощью фильтра* F и что f *сильно рационализируется*, если предпочтение $<$ является линейным порядком.

Что касается фильтра F, то тут большой простор для понимания рациональности, о чем мы уже говорили в предыдущем разделе. Возможны несколько комбинаций условий, например:

1) F удовлетворяет наследованию **Н**;

2) F удовлетворяет отбрасыванию **О**;

3) F порождается некоторым бинарным отношением, т.е. типа **НС**. (Следует сделать одно терминологическое уточнение: ФВ F *порождается* бинарным отношением $\ll$ и обозначается $F = \mathrm{Max}(\ll)$, если $\mathrm{Max}(\ll)(A) = \{a \in A,$ не существует $x \in A$: $a \ll x\}$.)

4) F удовлетворяет **НО**, т.е. является функцией Плотта (см. разд. 7);

Журнал НЭА, №2 (26), 2015, с. 12–35

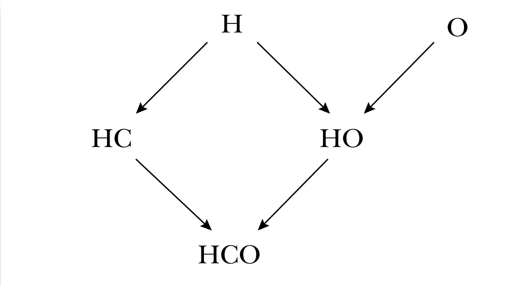

Рис. 1

5) F удовлетворяет **HCO**, т.е. порождается некоторым транзитивным (и асимметричным) бинарным отношением. Этот случай наиболее близок к классической рациональности (рис. 1).

Руководствуясь такой схемой, приведем соответствующие теоремы характеризации. Ситуация 3 была исследована в работе (Manzini, Mariotti, 2007); ситуация 2 – в (Masatlioglu et al., 2012); ситуация 1 – в (Cherepanov et al., 2013; Lleras et al., 2012); ситуации 4–5 – в (Lleras et al., 2012). В основном мы следуем упомянутым работам, хотя временами используем унифицированную терминологию, приводим новые результаты, формулировки и/или доказательства. Если теорема снабжена ссылкой на источник, это указывает на формулировку, данную в этом источнике.

3. Случай трех альтернатив

Прежде чем переходить к общим результатам о характеризации рационализуемости с помощью тех или иных фильтров, полезно разобраться с частным случаем, когда имеется всего три альтернативы.

Будем обозначать наши альтернативы символами 1, 2, 3. Имеется четыре существенно разных точных ФВ на множестве из трех альтернатив. Альтернативу, выбираемую из 123, обозначим через 3; альтернативу, выбираемую из 12, обозначим через 2. После этого остаются 4 возможности (выбрать из 13 и 23). Расположим их в порядке убывания рациональности (с указанием соответствующего базисного отношения; $x \to y$ обозначает $x >_f y$).

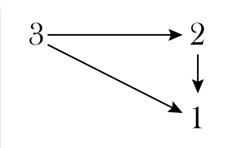

Рис. 2

Случай 1. Классически рациональный выбор, когда $f(13)=3$ и $f(23)=3$ (рис. 2). Базисное отношение $<_f$ совпадает с естественным порядком 1<2<3. Фильтр можно считать тождественным, $F(A)=A$.

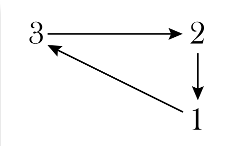

Рис. 3

Случай 2. $f(13)=1$, а $f(23)=3$ (рис. 3). Здесь мы видим цикл в базисном отношении $>_f$: $1 >_f 3 >_f 2 >_f 1$. Фактически это – пример, рассмотренный в начале раздела. Данная ФВ рационализуется с помощью бинарного отношения $2 \gg 1$; в качестве предпочтения можно взять, например, линейный порядок 2<3<1. При таком объяснении фильтр удовлетворяет условиям **H**, **C** и **O**.

Журнал НЭА, №2 (26), 2015, с. 12–35

Однако эта же функция f может иметь и другое объяснение. Фильтр порождается (уже нетранзитивным) бинарным отношением $1 \ll 2 \ll 3$ и предпочтением 2>1>3. Такая рационализация представляется менее рациональной, поскольку фильтр удовлетворяет лишь условиям **H** и **C**. Иначе говоря, рационал на первом этапе не является транзитивным бинарным отношением. Этот пример показывает, что могут существовать довольно отличающиеся двухэтапные рационализации одной и той же точной ФВ. В частности, довольно сильно могут отличаться предпочтения. В рассмотренных выше двух предпочтения общим является только соотношение 1>3; и как будет показано в разд. 4 (лемма $1'$), это не случайно.

Случай 3. $f(13)=3$, $f(23)=2$ (рис. 4).

Данная ФВ рационализируется плоттовским фильтром (т.е. удовлетворяющим **H** и **O**, см. разд. 7). А именно фильтром F, который тождествен всюду, кроме 123, и $F(123)=13$ (это единственная нерациональная функция Плотта). В качестве предпочтения < надо взять линейный порядок 1<3<2 (совпадающий с базисным отношением).

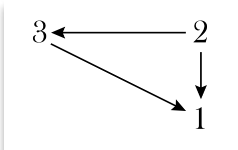

Рис. 4

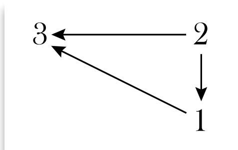

Рис. 5

Случай 4. $f(13)=1$, $f(23)=2$ (рис. 5). Базисное отношение 3<1<2 является линейным порядком.

Случай 4a. Данная ФВ f рационализируется фильтром, удовлетворяющим **H**. Например, можно взять фильтр F, который тождествен на всех меню, кроме $X=123$; $F(123)=3$. Очевидно, этот фильтр наследственный. Предпочтение 3<1<2.

Случай 4b. Эта же ФВ рационализируется фильтром, удовлетворяющим **O**. Например, можно взять фильтр G, тождественный на 123 и совпадающий с f на остальных меню. В качестве предпочтения можно взять порядок 1<2<3.

Случай 4c. Фильтр F удовлетворяет **H**, но не **O**. Фильтр G удовлетворяет **O**, но не **H**. Можно показать (см. разд. 7), что не существует фильтра, удовлетворяющего одновременно **H** и **O**, который бы рационализировал f. Так что и в этом случае можно двумя совершенно разными способами рационализировать (объяснять) эту ФВ.

Отметим, что все четыре точные ФВ на множестве из трех альтернатив могут быть рационализированы с помощью наследственного фильтра. В случае четырех и более альтернатив это уже не так.

4. Случай наследственного фильтра

Каким условиям должна удовлетворять ФВ f, чтобы рационализироваться с помощью наследственного фильтра? Авторы статьи (Cherepanov et al., 2013) (следуя более раннему предложению (Manzini,

Журнал НЭА, №2 (26), 2015, с. 12–35

Mariotti, 2007)) формулируют такое условие, называя его WWARP, т.е. слабым WARP. Мы же предпочитаем назвать его условием *слабой наследственности*, **WH**. Потом мы увидим, что это условие также и достаточно для рационализации с помощью наследственного фильтра.

Предположим, что мы имеем ситуацию $x, y \in S \subseteq T$, как в WARP. И допустим, что $x = f(T)$. Аксиома WARP говорит, что $y \neq f(S)$ (и даже что $f(S) = x$). Ослабление WARP утверждает то же, но только при дополнительном условии, что $f(x, y) = x$, или если все же $f(S) = y$, то $f(x, y) = y$.

Аксиома слабой наследственности (**WH**). *Пусть* $S \subseteq T$ *и* $x = f(T) \in S$. *Если* $y = f(S)$ *отлично от* x, *то* $f(x, y) = y$.

В таком виде аксиома **WH** выглядит как ослабление **H**. Наследование утверждало бы, что $y = f(S)$ совпадает с $x = f(T)$. Ослабление показывает, что $f(x, y) = y$. Интуитивно y лучше x (о чем говорит выбор из x, y и S). Но в большем меню T альтернатива y не выбирается по той причине, что не проходит фильтр $F(T)$.

Лемма 1. *Если ФВ* f *рационализируется с помощью наследственного фильтра* F, *то* f *удовлетворяет аксиоме* **WH**.

Д о к а з а т е л ь с т в о. В самом деле, пусть $S \subseteq T$ и $x = f(T)$ лежит в S. В частности, $x \in F(T)$ и, ввиду наследственности F, альтернатива x принадлежит $F(S)$. Альтернатива y, выбираемая из S, также принадлежит $F(S)$, и тем более $F(\{x, y\})$. Предположим, вопреки **WH**, что $f(x, y) = x$. Так как y лежит в $F(\{x, y\})$, то $y < x$. Но тогда y доминируется альтернативой x, лежащей в $F(S)$, и не может выбираться в S. Возникает противоречие. ■

Свойству слабой наследственности удобнее придать несколько иную форму. Для этого стоит ввести важное бинарное отношение $\prec^{\text{H}}$ (Н указывает на связь с наследованием): для различных альтернатив x и y

$$x \prec^{\text{H}} y \text{, если } \exists S \subseteq T:\ x = f(T) \in S \text{ и } y = f(S). \tag{1}$$

Интуитивный смысл тут такой. Мы полагаем альтернативу x хуже, чем y на основании того, что y выбирается из меньшего меню S, хотя альтернатива x тоже доступна. Выбор x из большего множества T объясняется тем, что в T альтернатива y не проходит фильтр.

Лемму 1 можно теперь переформулировать следующим образом.

Лемма 1'. *Пусть ФВ* f *порождается парой* $(F, <)$ *с наследственным фильтром* F. *Тогда* $\prec^{\text{H}}$ *является подотношением отношения* $<$.

Д о к а з а т е л ь с т в о. Пусть $x \prec^{\text{H}} y$, т.е. для некоторых S и T выполнено $x \in S \subseteq T$, $x = f(T)$, $y = f(S)$. Повторяя рассуждение леммы 1, мы заключаем, что x, как и y, принадлежит $F(S)$. А так как из

Журнал НЭА, № 2 (26), 2015, с. 12–35

S выбирается y, то неверно, что $y < x$. В силу той же наследственности x и y принадлежат $F(x, y)$. Если не выполняется условие, что $x < y$, то $f(x, y)$ должна была бы одновременно быть и x, и y, что противоречит однозначности f. Значит, $x < y$. ■

В частности, отношение $\prec^{\text{H}}$ асимметрично как подотношение асимметричного отношения предпочтения $<$.

В качестве иллюстрации обратимся к ФВ, рассмотренной в случае 2 разд. 3. Легко убедиться, что для такой ФВ отношение $\prec^{\text{H}}$ состоит из единственной пары $3 \prec^{\text{H}} 1$. И как видно из обсуждения этого примера в разд. 3, большего о предпочтениях ничего сказать нельзя.

Теорема 1 (Cherepanov et al., 2013). *Пусть f – точная ФВ. Эквивалентны следующие утверждения:*

1) f *рационализируется с помощью наследственного фильтра;*

2) f *удовлетворяет условию слабой наследственности* **WH**;

3) *бинарное отношение* $\prec^{H}$ *асимметрично.*

Доказательство. Импликации $1) \Rightarrow 2) \Rightarrow 3)$ были установлены выше. Поэтому остается доказать, что $3 \Rightarrow 1$. Для этого в (Cherepanov et al., 2013) был явно построен соответствующий фильтр $F = F^{*}$. А именно для меню A они полагают

$$F^{*}(A) = \{f(B), \text{ где } f(B) \in A, \ A \subseteq B\}. \quad (2)$$

Легко показать, что построенный фильтр F^{*} наследственный. В самом деле, если некоторый элемент x выбирается из A и принадлежит меньшему множеству $A' \subseteq A$, то $x = f(B)$, $A \subseteq B$, и тем более $A' \subseteq B$, так что $x \in F^{*}(A')$.

Что касается предпочтения $<$, то оно берется равным $\prec^{\text{H}}$, благо последнее бинарное отношение асимметрично. Утверждается, что пара $(F^{*}, <)$ порождает f.

Пусть $x = f(A)$. Тогда очевидно $x \in F^{*}(A)$, и все будет доказано, если мы покажем, что x доминирует любую другую альтернативу из $F^{*}(A)$. Пусть $y \in F^{*}(A)$ и отлично от x. Это значит, что $y = f(B)$ для $A \subseteq B$. Что и означает, что $y \prec^{H} x$, т.е. $y < x$. ■

Замечание 1. Конструкция наследственной ФВ F^{*} по f носит общий характер. Если f – произвольная (не обязательно точная) ФВ, то по ней можно построить наследственное расширение (оболочку) f^{*} по формуле

$$f^{*}(A) = \bigcup_{B \supseteq A} (A \cap f(B)) = A \cap \Big(\bigcup_{B \supseteq A} f(B)\Big).$$

Это минимальная наследственная ФВ, содержащая f.

Замечание 2. Характеризация в терминах отношения $\prec^{\text{H}}$ хороша тем, что позволяет без труда дать характеризацию сильной рационализации с помощью наследственного фильтра. *Необходимым*

Журнал НЭА, №2 (26), 2015, с. 12–35

и достаточным условием сильной рационализации (с наследственным фильтром) является требование ацикличности бинарного отношения $\prec^{\text{Н}}$.

В одну сторону (ацикличность $\prec^{\text{Н}}$) это вытекает из того, что в силу леммы 1′ $\prec^{\text{Н}}$ является подотношением линейного порядка <. Проверим в другую сторону. В силу теоремы Шпильрайна ацикличное отношение $\prec^{\text{Н}}$ продолжается до некоторого линейного порядка, который мы обозначим как < и примем за предпочтение, действуюшее на втором этапе. Дальнейшее доказательство проводится, как в доказательстве теоремы 1.

Замечание 3. Условию ацикличности отношения $\prec^{\text{Н}}$ можно придать немного иной вид. Известно, что ацикличность отношения $\prec$ эквивалентна тому, что в любом меню S есть недоминируемая альтернатива s^*. Поэтому ацикличность $\prec^{\text{Н}}$ можно записать в виде: для любого меню S существует альтернатива $s^* \in S$, которая не доминируется (в смысле $\prec^{\text{Н}}$) никакой другой альтернативой из S. То есть если $s \in S$ и отлична от s^*, то неверно $s^* \prec^{\text{Н}} s$. Соотношение $s^* \prec^{\text{Н}} s$ расписывается как существование $s^* \in A \subseteq B$ таких, что $s = f(A)$, $s^* = f(B)$. Поэтому недоминирование s^* переписывается как утверждение: как только мы имеем $s^* \in A \subseteq B$, $s^* = f(B)$ и $f(A) \in S$, то $f(A) = s^*$. Окончательно условие ацикличности $\prec^{\text{Н}}$ (или сильной рационализуемости ФВ f с помощью наследственного фильтра) принимает вид: *для любого меню S существует альтернатива $s^* \in S$, которая обладает следующим свойством: если $s^* \in A$, $f(A) \in S$ и $f(B) = s^*$ для некоторого $B \supseteq A$, то $f(A) = s^*$*. Именно в таком виде это условие было сформулировано в работе (Lleras et al., 2012).

5. Случай бинарного фильтра

Теперь мы обсудим случай, исследованный в работе (Manzini, Mariotti, 2007), когда фильтр F удовлетворяет аксиомам **H** и **C**. Согласно теореме Сена это эквивалентно существованию некоторого бинарного отношения $\ll$ такого, что $F = \max(\ll)$. Это дает основания называть такой фильтр *бинарным.*

То, что фильтр F генерируется бинарным отношением, накладывает дополнительное ограничение на ФВ f. А именно f (как и F) удовлетворяет аксиоме согласия **C**. Имеет место следующая лемма.

Лемма 2. *Если точная ФВ f рационализируется с помощью бинарного фильтра, то она удовлетворяет аксиоме* **C**.

Д о к а з а т е л ь с т в о. Пусть $x = f(S) = f(T)$; мы должны показать, что $x = f(S \cup T)$. Начнем с того, что x принадлежит $F(S)$ и $F(T)$ (так как выбирается в S и T). Фильтр F удовлетворяет **C**, поэтому x принадлежит $F(S \cup T)$. Но как уже отмечалось при обсуждении свойства **H**, $F(S \cup T) \subseteq F(S) \cup F(T)$. Поэтому если x доминируется каким-то

Журнал НЭА, №2 (26), 2015, с. 12–35

элементом z, лежащим в $F(S \cup T)$, то он доминируется элементом, лежащим в $F(S)$ или в $F(T)$. Но таких элементов нет. ■

Эта лемма показывает, что не всякий выбор из трех альтернатив может быть оправдан с помощью какого-либо бинарного фильтра. Это согласуется с наблюдениями из разд. 3.

Итак, если точная ФВ f рационализируется с помощью бинарного фильтра, то она обладает двумя характерными свойствами: **C** и **WH** (или асимметрией отношения $\prec^{H}$). В (Manzini, Mariotti, 2007) было показано, что верно и обратное утверждение.

Теорема 2 (Manzini, Mariotti, 2007). *Пусть f – точная ФВ. Эквивалентны следующие утверждения:*

1) f *удовлетворяет* **WH** *и* **C***;*

2) f *рационализируется с помощью бинарного фильтра;*

3) f *удовлетворяет аксиоме* **C***, а отношение* $\prec^{H}$ *асимметрично.*

Доказательство. Почти все импликации уже проверены, и остается только показать, что $3 \Rightarrow 2$. А это следует из того, что наследственный фильтр F^*, построенный по формуле (2), удовлетворяет аксиоме **C**. Напомним, что $F^*(A) = \{f(B),\ \text{где}\ f(B) \in A,\ A \subseteq B\}$.

Предположим, что x принадлежит и $F^*(A)$, и $F^*(B)$. Это значит, что $x = f(A')$ для $A' \supseteq A$, а также $x = f(B')$ для $B' \supseteq B$. Так как ФВ f удовлетворяет аксиоме **C**, $x = f(A' \cup B')$. А так как $(A \cup B) \subseteq (A' \cup B')$, то $x \in F^*(A \cup B)$. ■

Замечание 4. Фильтр, удовлетворяющий условиям **H** и **C**, генерируется некоторым бинарным отношением $\ll$. Полезно выписать это отношение более явно, тем более что в (Manzini, Mariotti, 2007) теорема была доказана именно с помощью такого отношения. Существует естественный способ определить отношение $\ll$. А именно для различных x и y мы полагаем $y \ll x$, если $F(x, y) = \{x\}$. Вспоминая определение F^*, можно переписать это так: для любого меню T, содержащего x и y, $f(T) \neq y$. Отсюда мы приходим к формуле для $\ll$, предложенной в (Manzini, Mariotti, 2007):

$x \ll y$, если для любого меню S, содержащего y, $f(S)$ отличен от x.

Легко убедиться, что данное отношение $\ll$ порождает фильтр F^*. Действительно, альтернатива a из A принадлежит $\mathrm{Max}(\ll)(A)$, если она не доминируется никакой другой альтернативой b из A. То есть для любой альтернативы b из $A - \{a\}$ существует такое меню $B(b)$, содержащее b, что $f(B(b)) = a$. Пусть теперь B – объединение всех $B(b), b \in A - a$. Тогда очевидно, что B содержит A. В силу аксиомы согласия $a = f(B)$. Но это означает, что $a \in F^*(A)$.

Противоположное включение $F^*(A) \subseteq \mathrm{Max}(\ll)(A)$ уже не требует аксиомы согласия.

Журнал НЭА, №2 (26), 2015, с. 12–35

Замечание 5. Как и в предыдущем разделе (см. замечание 3), сильная рационализуемость получается, если мы заменим в теореме 2 условие асимметрии отношения $\prec^{\text{H}}$ на более сильное условие ацикличности.

Замечание 6. Бинарное отношение $\ll$, построенное в замечании 1, объясняет выбор f. Однако, как показывает пример 2 из разд. 3, могут быть и меньшие бинарные отношения, объясняющие тот же выбор (см. также разд. 8).

6. Фильтры с условием отбрасывания

Теперь, следуя работе (Masatlioglu et al, 2012), рассмотрим другой фундаментальный случай, когда фильтр F удовлетворяет аксиоме отбрасывания **O**.

Пусть элемент y принадлежит S, но не $F(S)$. В силу аксиомы **O** для фильтра F мы получаем $F(S-y)=F(S)$ и поэтому $f(S-y)=f(S)$. Если удаление из S некоторого y (отличного от $f(S)$) меняет выбор ($f(S-y)\neq f(S)$), то y принадлежит $F(S)$ и поэтому хуже, чем $f(S)$. Это дает основание ввести следующее отношение $\prec^{\text{O}}$: для разных x и y полагаем:

$$y \prec^{\text{O}} x, \text{ если } x=f(S) \text{ для некоторого меню S и } f(S-y)\neq f(S). \quad (3)$$

Данное рассуждение можно оформить как лемму.

Лемма 3. *Если точная ФВ f рационализируется фильтром, удовлетворяющим аксиоме* **O**, *то отношение $\prec^{\text{O}}$ является подотношением предпочтения $<$.*

В частности, отношение $\prec^{\text{O}}$ асимметрично. Верно и обратное утверждение.

Теорема 3. *Точная ФВ f тогда и только тогда рационализируется фильтром, удовлетворяющим условию* **O**, *когда отношение $\prec^{\text{O}}$ асимметрично.*

Доказательство. Определим фильтр $F=F^{\text{O}}$ формулой, предложенной в (Masatlioglu et al., 2012):

$$\text{для } A\subseteq X\text{: } F^{O}(A)=\{f(A)\}\cup\{a\in A, a\prec^{O} f(A)\}. \quad (4)$$

Мы утверждаем, что этот фильтр удовлетворяет аксиоме **O**. Пусть s принадлежит S, но не $F^{\text{O}}(S)$; нужно проверить, что $F^{\text{O}}(S-s)=F^{\text{O}}(S)$. Так как $s\notin F^{\text{O}}(S)$, то соотношение $s\prec^{\text{O}} f(S)$ не имеет места. По определению $\prec^{\text{O}}$ это означает, что для любого T, для которого $f(T)=f(S)$, выполняется равенство $f(T-s)=f(S)$. В частности, для $T=S$ получаем $f(S-s)=f(S)$. Отсюда $F^{\text{O}}(S-s)=F^{\text{O}}(S)$.

Теперь положим $<=\prec^{\text{O}}$. По определению фильтра F^{O}, элемент $f(A)$ доминирует все другие альтернативы из $F^{\text{O}}(A)$.

Следствие. Точная ФВ f сильно рационализируется фильтром со свойством **O** тогда и только тогда, когда отношение $\prec^{\text{O}}$ ациклично.

Журнал НЭА, №2 (26), 2015, с. 12–35

Ацикличность $\prec^{\mathrm{O}}$ следует из леммы 3. Обратно: пусть отношение $\prec^{\mathrm{O}}$ ациклично, и $<$ – произвольный линейный порядок, продолжающий $\prec^{\mathrm{O}}$. Очевидно, что $f(A)$ – максимальный (уже относительно $<$) элемент в $F^{\mathrm{O}}(A)$.

Фактически это следствие было установлено в (Masatlioglu et al., 2012). Единственное отличие состоит в том, что там ацикличность $\prec^{\mathrm{O}}$ формулируется в терминах недоминируемости: в любом меню S существует альтернатива s^*, которая не доминируется (с помощью $\prec^{\mathrm{O}}$) никакой другой альтернативой из S. Расшифровывая недоминируемость, мы приходим к условию (названному в (Masatlioglu et al., 2012) как WARP(LA)):
в любом меню S существует альтернатива $s^* \in S$, обладающая свойством: $f(A) = s^*$ как только $f(A) \in S$ и $f(A - s^*) \neq f(A)$.

Теперь мы можем сформулировать результат, установленный в (Masatlioglu et al., 2012).

Теорема 3′ (Masatlioglu et al., 2012). *Пусть f – точная ФВ. Эквивалентны условия:*

1) *f сильно рационализируется с помощью фильтра, удовлетворяющего условию отбрасывания* **O**;

2) *бинарное отношение $\prec^{O}$ ациклично;*

3) *f удовлетворяет условию* WARP(LA).

7. Плоттовский фильтр

В этом разделе мы обсудим двухэтапную рационализацию в том случае, когда фильтр F является функцией выбора Плотта, т.е. удовлетворяет одновременно двум фундаментальным требованиям наследования **H** и отбрасывания **O**. В сильной постановке эти вопросы были решены в работе (Lleras et al., 2012), однако мы начнем со слабой рационализации, когда предпочтение удовлетворяет лишь условию асимметрии.

Ранее (см. (1), (3)) мы ввели два бинарных отношения $\prec^{\mathrm{H}}$ и $\prec^{\mathrm{O}}$, связанных с точной ФВ f. Рассмотрим теперь объединение этих отношений, $\prec^{\mathrm{HO}} = \prec^{\mathrm{H}} \cup \prec^{\mathrm{O}}$. Иначе говоря, $x \prec^{\mathrm{HO}} y$ тогда и только тогда, когда $x \prec^{\mathrm{H}} y$ или $x \prec^{\mathrm{O}} y$.

Предложение 1. *Если точная ФВ рационализируема с помощью плоттовского фильтра, то отношение $\prec^{\mathrm{H}} \cup \prec^{\mathrm{O}}$ асимметрично.*

Это следствие лемм 1′ и 3. Из него вытекает сделанное в разд. 3 (случай 4с) утверждение. В самом деле, если вычислить отношение $\prec^{\mathrm{H}}$ для этой ФВ, то $3 \prec^{\mathrm{H}} 1$ и $3 \prec^{\mathrm{H}} 2$. Если же вычислить $\prec^{\mathrm{O}}$, то получим противоположные соотношения. Разумеется, это противоречит асимметрии $\prec^{\mathrm{H}} \cup \prec^{\mathrm{O}}$.

Журнал НЭА, №2 (26), 2015, с. 12–35

Гипотеза. Если отношение $\prec^{\mathrm{H}} \cup \prec^{\mathrm{O}}$ асимметрично, то ФВ f рационализируется плоттовским фильтром.

Ясно, что в качестве предпочтения надо брать отношение $\prec^{\mathrm{H}} \cup \prec^{\mathrm{O}}$. Трудность состоит в построении плоттовского фильтра F. Дело в том, что в случае **H** использовалась конструкция (2), а в случае **O** – (4). Тем не менее в работе (Lleras et al., 2012) утверждается, что гипотеза верна в варианте сильной рационализуемости. А именно имеет место следующая теорема.

Теорема 4 (Lleras et al., 2012). *Точная ФВ f сильно рационализируется с помощью плоттовского фильтра тогда и только тогда, когда бинарное отношение $\prec^{\mathrm{H}} \cup \prec^{\mathrm{O}}$ ациклично.*

Мы не приводим доказательство этой теоремы ввиду его сложности.

Следствие. Любое из следующих условий достаточно для сильной рационализации точной ФВ f плоттовским фильтром:
- отношение $\prec^{\mathrm{H}}$ ациклично, а отношение $\prec^{\mathrm{O}}$ является подотношением транзитивного замыкания $\prec^{\mathrm{H}}$;
- отношение $\prec^{\mathrm{O}}$ ациклично, а отношение $\prec^{\mathrm{H}}$ является подотношением транзитивного замыкания $\prec^{\mathrm{O}}$.

Замечание 7. На самом деле в (Lleras et al., 2012) вместо отношения $\prec^{\mathrm{H}} \cup \prec^{\mathrm{O}}$ было использовано отношение $\prec$. А именно для различных x и y авторы полагают $x \prec y$, если существуют S и $T: x \in S \subseteq T$, $y = f(S)$ и $f(T) \neq f(T-x)$. В (Lleras et al., 2012) теорема 4 выглядит так: f *рационализируется плоттовким фильтром тогда и только тогда, когда отношение $\prec$ ациклично.*

Эквивалентость такой формулировки с приведенной выше следует из того, что отношение $\prec^{\mathrm{H}} \cup \prec^{\mathrm{O}}$ и отношение $\prec$ имеют одно и то же транзитивное замыкание. Это видно из следующих наблюдений:

1) если $x \prec y$, то $x \prec^{\mathrm{O}} f(T)$ и $f(T) \prec^{\mathrm{H}} f(S) = y$;
2) если $x \prec^{\mathrm{H}} y$ (и значит $x = f(T) \in S \subseteq T$ и $y = f(S)$), то $f(T-x) \neq f(T) = x$ и поэтому $x \prec y$;
3) если $x \prec^{\mathrm{O}} y$ (и значит $y = f(A)$ и $f(A-x)$ отлично от $f(A) = y$), мы можем взять $T = A$ и $S = \{y\}$ и (так как $f(T) \neq f(T-x)$) получить $x \prec y$.

8. Случай паретовского фильтра

Рассмотрим теперь тот случай, когда фильтр удовлетворяет всем трем базисным условиям **H**, **O** и **C**. Как известно, это эквивалентно тому, что фильтр порождается транзитивным иррефлексивным бинарным отношением. Последнее можно представить как пересечение нескольких линейных порядков. И недоминируемые альтерна-

Журнал НЭА, №2 (26), 2015, с. 12–35

тивы – это в точности паретовские альтернативы, откуда и название – паретовский фильтр. Кроме того, мы будем интересоваться сильной рационализуемостью.

Зафиксируем линейный порядок <, и пусть f – точная ФВ. Для дальнейшего изложения удобно ввести бинарное отношение *выявленного доминирования* $\ll_f$. А именно:

$$x \ll_f y, \text{ если } x > y, \ f(x,y) = y$$

(можно сказать, что пара (x,y) *аномальная*, потому что x предпочтительнее, чем y, но не выбирается из $\{x,y\}$. Конечно, это объясняется тем, что x не проходит фильтр). Очевидно, что отношение $\ll_f$ асимметрично.

Теорема 5. *Пусть $<$ – линейный порядок, а f – точная ФВ. Эквивалентны утверждения:*

4) *f порождается парой $(F,<)$, где F – паретовский фильтр;*
5) *f порождается парой $(F,<)$, где F – плоттовский фильтр, а f удовлетворяет* **C***;*
6) *построенноевышеотношение $\ll_f$ транзитивно, а $f = \max(\mathrm{Max}(\ll_f))$;*
7) *f порождается парой $(F,<)$, где $F = \mathrm{Max}(\ll)$, а $\ll$ – транзитивное отношение.*

Доказательство. $1) \Rightarrow 2)$ в силу леммы 2. Центральной является импликация $2) \Rightarrow 3)$. Транзитивность $\ll_f$ следует из того, что ФВ F плоттовская.

Лемма А. *Если фильтр F плоттовский, то отношение $\ll_f$ транзитивно.*

Доказательство леммы А. Пусть $x \gg_f y$ и $y \gg_f z$; нужно показать, что $x \gg_f z$. Рассмотрим $F(xyz)$. Мы утверждаем, что $y \notin F(xyz)$. Предположим противное, т.е. $y \in F(xyz)$. В силу наследственности F значение y принадлежит $F(xy)$, и так как $y > x$, то $f(xy) = y$, вопреки $x \gg_f y$.

По тем же причинам $z \notin F(xyz)$. Значит, $F(xyz) = \{x\}$. А так как фильтр F удовлетворяет **O**, то, отбрасывая y, мы получаем $F(xz) = \{x\}$, откуда $f(xz) = x$. Так как из транзитивности < мы имеем $z > x$, мы получаем $x \gg_f z$. ■

Покажем, что $f(A)$ – наибольший элемент в $\mathrm{Max}(\ll_f)(A)$. Для этого нужно проверить, что $f(A)$ принадлежит $\mathrm{Max}(\ll_f)(A)$ (т.е. не доминируется ничем в A) и что любая большая альтернатива доминируется чем-то из A. Это устанавливается в следующих двух леммах.

Лемма В. *Альтернативы из $F(A)$ не доминируются никакими элементами из A.*

Доказательство. В самом деле, допустим, что $a \in F(A)$, $x \in A$ и $x \gg_f a$. В силу наследственности фильтра F альтернатива a принадлежит $F(\{x,a\})$. А так как $x < a$, то $f(x,a) = a$, вопреки тому, что $f(x,a) = x$.

Журнал НЭА, №2 (26), 2015, с. 12–35

Лемма С. *Пусть* f *как в п. 2 теоремы 5. Если для* $a \in A$ *выполнено* $a > f(A)$, *то* $a \ll_f x$ *для некоторого элемента* x *из* $F(A)$.

Доказательство. Пусть $a > f(A)$. Положим $A' = F(A) \cup a$. Тогда

$$F(A) \subseteq A' \subseteq A,$$

и в силу **O** для F мы имеем $F(A') = F(A)$, откуда $f(A') = f(A)$. Заменяя A на A', мы можем считать, что $A = F(A) \cup \{a\}$. Пусть теперь x – произвольный элемент из $A - a = F(A)$. Если соотношение $x \gg_f a$ не имеет места, то (так как $x \leq f(A) < a$) выполнено $f(x,a) = a$. И так для любого x из $A - a$. В силу условия **С** для функции f мы заключаем, что $f(A) = a$, вопреки предположению $a > f(A)$. Итак, мы доказали, что $2) \Rightarrow 3)$.

Импликация $3) \Rightarrow 4)$ тривиальна. Импликация $4) \Rightarrow 1)$ следует из упомянутого в начале раздела классического факта. ■

Объединяя теоремы 4 и 5, получаем следствие.

Следствие (Lleras et al., 2012). Точная ФВ f сильно рационализируется с помощью паретовского фильтра тогда и только тогда, когда отношение $\prec^{\mathrm{H}} \cup \prec^{\mathrm{O}}$ ациклично, а f удовлетворяет **С**.

9. Механизмы фильтрации

До этого момента мы в основном занимались математической стороной двухэтапной рационализации. Вопросы формирования фильтров оставались без обсуждения. Однако этой теме авторы обозреваемых статей уделяли много внимания, так как те или иные условия на фильтры связаны с их генезисом. Здесь мы вкратце обсудим эти вопросы.

1. Рационализация. Чтобы оправдать свой интерес к наследственным фильтрам, авторы (Cherepanov et al., 2013) ссылаются на психоаналитика Джонса, который писал в 1908 г., что люди (как рациональные создания) чувствуют необходимость иметь объяснение своему поведению. И что они бессознательно подыскивают оправдания своим поступкам. Психологи отмечают ту легкость, с которой люди придумывают самые неправдоподобные объяснения. Неспособность подыскать рациональное объяснение оказывает влияние на поведение, ограничивая выбор. В (Cherepanov et al., 2013) приводится не совсем понятный пример. Управляющий может и имеет побуждения совершить некое мошенничество. Однако если он не находит убедительного рационального объяснения, которое легитимизирует это мошенничество, он отказывается от него.

Пока это звучит убедительно. Но следующий шаг, который делается в (Cherepanov et al., 2013), не выглядит достаточно обоснованным. Авторы сводят рационалы к бинарным отношениям и счи-

Журнал НЭА, №2 (26), 2015, с. 12–35

тают, что (в данном меню) альтернатива оправдана, если она доминирует все другие альтернативы из этого меню с помощью некоторого рационала. Нам представляется неправильной подмена оправдания некоторой альтернативы как логического рассуждения (типа доказательства теоремы) проверкой ее недоминируемости. Поэтому ссылки на Джонса выглядят притянутыми.

Легко убедиться, что указанным выше способом получаются наследственные фильтры и только они. Этот факт хорошо известен в литературе по теории выбора (см., например, (Айзерман, Алескеров, 1990)). Тем не менее авторы (Cherepanov et al., 2013) (хотя и небезупречно) доказывают его.

Теория бинарных фильтров (разд. 5), развиваемая в (Manzini, Mariotti, 2007), оказывается частным случаем, когда для выбора финалистов используется один рационал. Теория плоттовских фильтров также оказывается частным случаем этой схемы, когда рационалы являются не просто бинарными отношениями, но линейными порядками (теорема Айзермана–Малишевского, см. (Айзерман, Алескеров, 1990)).

2. Рассмотрение. В работе (Lleras et al., 2012) обоснование наследственности основано на соображении, что обширность и/или сложность меню приводит к тому, что многие альтернативы выпадают из рассмотрения. Авторы этой работы называют наследственные фильтры фильтрами консидерации. Аксиому наследственности они оправдывают тем, что если некоторая альтернатива выпала из рассмотрения в меньшем множестве, она выпадает из рассмотрения и в большем. Это условие они понимают как наиболее базисное условие согласованности выбора финалистов. В процессе исследования авторы (Cherepanov et al., 2013) добавляют к этому базисному условию дополнительные, примерно как это делали мы ранее. Стоит отметить, что в формальном изучении двухэтапного рационального выбора авторы (Lleras et al., 2012) продвинулись дальше всех.

3. Категоризация. С иных позиций к оправданию наследственности подошли в работах (Manzini, Mariotti, 2010, 2012). Авторы исходили из того, что категоризация является центральным когнитивным механизмом. Под категоризацией понимается стремление разложить все по полочкам, отнести каждую альтернативу к той или иной категории, к тому или иному подмножеству сходных в чем-то объектов. Сталкиваясь со сложной проблемой (например, выбора ресторана), человек не сравнивает достоинства каждого ресторана в отдельности, а сначала сравнивает грубые объекты (категории) (например, тип кухни – итальянская, мексиканская, китайская и т.д.). И лишь сделав

Журнал НЭА, №2 (26), 2015, с. 12–35

выбор в пользу, например, итальянской кухни, человек решает, в какой из итальянских ресторанов он пойдет. Впрочем, категоризация тех же ресторанов может происходить и на другой основе.

Более формально, категория – это просто непустое подмножество альтернатив, меню. Считается, что на множестве меню имеется некоторое бинарное отношение $\triangleleft$. Если $R \triangleright S$, то категория R исключает (затеняет, заслоняет) альтернативы из S. Более формально, фильтр $F_{\triangleleft}$ образуется по правилу: альтернатива $x \in S$ не попадает в фильтр, если она принадлежит категории $R \subseteq S$, которая доминируется другой категорией R' ($R \triangleleft R'$), также лежащей внутри S.

Согласно работе (Айзерман, Алескеров, 1990) такой механизм формирует любые наследственные фильтры. Поэтому (вслед за (Manzini, Mariotti, 2012)) мы получаем характеризацию такого категоризационного выбора с помощью аксиомы **WH**. Впрочем, авторы (Manzini, Mariotti, 2012) считают это совпадение «довольно удивительным».

Нам же представляется более удивительным замечание этих авторов (сделанное в (Manzini, Mariotti, 2010)), что их теория категоризации может быть отделена от теории рационализации (Cherepanov et al., 2013). А именно они формулируют утверждение в (Manzini, Mariotti, 2010, Theorem 2), которое звучит так: *точная ФВ f сильно рационализируется с помощью категоризации* (читай – с помощью наследственного фильтра) *тогда и только тогда, когда она удовлетворяет условию* WWARP, *а базисное отношение $<_f$ ациклично.* Этому заявлению противоречит пример из разд. 2, где базисное отношение как раз циклическое. Авторы явно что-то напутали в формулировке.

Нам кажется, что идею категоризации более естественно формализовать следующим образом. Зафиксируем множество $\mathcal{C} \subseteq 2^X$ категорий, а также некоторую вспомогательную ФВ $\mathcal{F}$ на множестве $\mathcal{C}$. Эти данные позволяют построить фильтр F по следующему правилу. Пусть дано меню $A \subseteq X$. Сначала мы формируем релевантное множество категорий, в которые могут попасть элементы из A. А именно

$$\mathcal{C}(A) = \{C \in \mathcal{C}, C \cap A \neq \varnothing\}.$$

Например, $\mathcal{C}(a) = \{C \in \mathcal{C}, a \in C\}$. И тогда $\mathcal{C}(A) = \cup_{a \in A} \mathcal{C}(a)$.

Далее, пользуясь ФВ $\mathcal{F}$, мы из множества $\mathcal{C}(A)$ отбираем побеждающие категории, т.е. образуем $\mathcal{F}(\mathcal{C}(A))$. На последнем шаге мы дезагрегируем победителей, полагая:

$$F(A) = A \cap (\cup_{C \in \mathcal{F}(\mathcal{C}(A))} C) = \cup_{C \in \mathcal{F}(\mathcal{C}(A))} (A \cap C).$$

Например, вспомогательная ФВ $\mathcal{F}$ может генерироваться линейным порядком $\prec$ на $\mathcal{C}$. Тогда $F(A)$ состоит из элементов A, попавших в наиболее привлекательную категорию среди категорий, задевающих A.

Журнал НЭА, №2 (26), 2015, с. 12–35

Легко показать (мы оставляем это читателю), что имеет место следующая лемма.

Лемма. *Если ФВ $\mathcal{F}$ обладает свойством* **H** *или* **O**, *то тем же свойством обладает фильтр F.*

4. Фильтр внимания. Авторы статьи (Masatlioglu et al., 2012) называют фильтр, удовлетворяющий аксиоме отбрасывания, *фильтром внимания* (attention filter). Они исходят из того, что фильтр проходят альтернативы, чем-то привлекшие внимание ЛПР. Иначе говоря, альтернативы не попадают в фильтр, если они не привлекли внимания, попросту говоря, остаются незамеченными, невидимыми, неосознанными. Аксиома отбрасывания в этом случае выглядит очень естественно: если кто-то уберет невидимую альтернативу, это пройдет незамеченным для ЛПР.

5. Эксперименты. В работе (Manzini, Mariotti, 2010) приведены результаты некоторых лабораторных экспериментов о функциях выбора реальных испытуемых на множествах из трех–четырех альтернатив. Вот краткие выводы, сделанные на их основе:

а) большинство испытуемых нарушают аксиому WARP, т.е. их выбор не вполне рационален;

б) эти нарушения чаще связаны с эффектами меню, чем с цикличностью бинарного выбора (разд. 1);

в) выбор большинства испытуемых удовлетворял аксиоме WWARP. Что, может быть, не очень удивительно, учитывая, что выбор среди трех альтернатив всегда рационализируется с помощью наследственного фильтра (разд. 3).

6. Пороговый фильтр. В работе (Manzini et al., 2013) был рассмотрен еще один интересный способ образования фильтра финалистов F. Он основан на идее порога, приведенной в (Айзерман, Адескеров, 1990). В формировании фильтра участвуют две дополнительные функции: одна h задана на X, а другая, θ, задана на множестве меню. Альтернатива $x \in S$ включается в $F(S)$, если $h(x) \geq (S)$.

Например, если θ монотонно зависит от меню, получающаяся ФВ удовлетворяет условию наследования.

Авторы (Manzini et al., 2013) приводят характеризацию (неточных в общем случае) ФВ, рационализируемых такими пороговыми фильтрами. Попутно они отмечают, что любая точная ФВ допускает рационализацию пороговым фильтром. Получается слишком широкая модель, если не делать специальных предположений на пороговую функцию θ.

7. Случайный выбор. В работе (Manzini, Mariotti, 2014) авторы исследовали еще один интересный способ фильтрации альтернатив, на этот раз основанный на случайности. Имеется в виду, что у каждой

Журнал НЭА, №2 (26), 2015, с. 12–35

альтернативы a есть некоторая (заданная изначально) вероятность $\gamma(a)$ быть рассмотренной (т.е. включенной в множество $F(A)$). Выбор из последнего осуществляется уже с помощью фиксированного линейного порядка <. Иными словами, $F(A)$ – это уже случайное подмножество в A; соответственно, окончательный выбор $f(A)$ – случайный элемент A.

Математически случайный выбор f задается набором неотрицательных чисел (вероятнослей выбора a из A) $p(a,A)$, где $a \in A \subseteq X$, удовлетворяющим неравенствам $\sum_{a\in A} p(a,A) \le 1$ (допускается, что не будет выбрано ни одного элемента).

Конкретно, в модели Манзини–Мариотти (известной как *правило случайного рассмотрения*) случайность задается отображением $\gamma : X \to (0,1]$ ($\gamma(x)$ понимается как вероятность рассмотрения извлечения альтернативы). При фиксированном правиле γ и линейном порядке <, вероятности $p(a,A)$ определяются формулой

$$p(a,A) = \gamma(a) \prod_{b\in A, a<b} (1-\gamma(b)).$$

Удивительно, что и в этом случае существует довольно простая характеризация такого случайного выбора. Более того, как управляющий линейный порядок <, так и вероятности наблюдения альтернатив однозначно восстанавливаются по f. Опишем кратко, как это делается.

Прежде всего, если $b < a$, то удаление a из меню повышает шансы b быть выбранной. Действительно,

$$p(b,A) = p(b,A-a)(1-\gamma(A)) < p(b,A-a).$$

Это позволяет однозначно восстановить отношение <. После этого число $\gamma(a)$ восстанавливается как $p(a,L(a))$, где $L(a) = \{x, x \le a\}$ – нижний контур a.

На самом деле столь же легко описать те правила случайного выбора, которые порождаются механизмом случайного рассмотрения. Оказывается, что для этого должны выполняться три свойства:

1) *монотонность по меню*: вероятность выбора в большем множестве меньше, чем в меньшем;

2) *независимость от исключения альтернатив*: отношение $p(a,A) / p(a,A-b)$ не зависит от A, если только меню A содержит a и b. Положим $a \prec b$, если $p(a,A) < p(a,A-b)$ и $p(b,A) = p(b,A-a)$ для некоторого (на самом деле, любого) меню A, содержащего a и b.

3) *ацикличность отношения* $\prec$. Легко убедиться, что отношение $\prec$ полное и, значит, является линейным порядком. Поэтому его можно взять как предпочтение <. Про вероятности извлечения мы уже говорили.

Журнал НЭА, №2 (26), 2015, с. 12–35

Авторы (Manzini, Mariotti, 2014) сравнивают свою концепцию случайного выбора с теорией Льюса и теорией случайной полезности, но это уже другая тема.

ЛИТЕРАТУРА

Айзерман М.А., Алескеров Ф.Т. (1990). Выбор вариантов; основы теории. М.: Наука.

Айзерман М.А., Малишевский А.В. (1981). Некоторые аспекты общей теории выбора лучших вариантов // *Автоматика и телемеханика.* № 2. С. 65–83.

Зак Ф.Л. (2014а). Психологические игры в теории выбора. I. Искушение, перфекционизм, самообман // *Журнал Новой экономической ассоциации.* № 1. С. 12–32.

Зак Ф.Л. (2014б). Психологические игры в теории выбора. II. Стыд, сожаление, эгоизм и альтруизм // *Журнал Новой экономической ассоциации.* № 2. С. 12–40.

Aizerman M., Malishevski A. (1986). Conditions for Universal Reduction of a Two-Stage Extremization Problem to One-Stage Problem // *Journal of Mathematicsl Analysis and Applications.* Vol. 118. P. 361–388.

Aleskerov F., Cinar Y. (2008). q-Pareto-scalar Two-Stage Extremization Model and Its Reducibility to One-Stage Model // *Theory and Decision.*Vol. 65. P. 325–338.

Cherepanov V., Feddersen T., Sandroni A. (2013). Rationalization // *Theoretical Economics.* Vol. 8. P. 775–800.

Danilov V., Koshevoy G., Savaglio E. (2014). Hyperrelations, Choice Functions, and Orderings of Opportunity Sets. Forthcomming in Social choice and welfare'.

Lleras J.S., Masatlioglu Y., Nakajima D., Ozbay E. (2012). When more is less: Limited consideration. Unpublished paper.

Manzini P., Mariotti M. (2007). Sequentially Rationalizable Choice // *American Economic Review.* Vol. 97. P. 1824–1839.

Manzini P., Mariotti M. (2010). Revealed Preferences and Boundedly Rational Choice Procedures: an Experiment. Unpublished Paper.

Manzini P., Mariotti M. (2012). Categorize then Choose: Boundedly Rational Choice and Welfare // *Journal of the European Economic Association.* Vol. 10. P. 1141–1165.

Manzini P., Mariotti M. (2014). Stochastic Choice and Consideration Sets // *Econometrica.* Vol. 82. P. 1153–1176.

Manzini P., Mariotti M., Tyson C.J. (2013). Two-stage Threshold Representations // *Theoretical Economics.* Vol. 6. P. 875–882.

Masatlioglu Y., Nakajima D., Ozbay E. (2012). Revealed Attention // *American Economic Review.* Vol. 102. P. 2183–2205.

REFERENCES (with English translation or transliteration)

Aizerman M., Malishevski A. (1986). Conditions for Universal Reduction of a Two-Stage Extremization Problem to One-Stage Problem. *Journal of Mathematicsl Analysis and Applications* 118, 361–388.

Журнал НЭА, №2 (26), 2015, с. 12–35

Aizerman M., Malishevski A. (1986). Conditions for Universal Reduction of a Two-Stage Extremization Problem to One-Stage Problem. *Journal of Mathematicsl Analysis and Applications* 118, 361–388.

Aleskerov F., Cinar Y. (2008). q-Pareto-scalar Two-Stage Extremization Model and Its Reducibility to One-Stage Model. *Theory and Decision* 65, 325–338.

Ayzerman M.A., Aleskerov F.T. (1990). Choice of Variants (Foundations of the Theory). Moscow: Nauka (in Russian).

Ayzerman M.A., Malishevskiy A.V. (1981). General Theory of Best Variants Choice: Some Aspects. *Avtomatika i telemekhanika* 2, 65–83 (in Russian).

Cherepanov V., Feddersen T., Sandroni A. (2013). Rationalization. *Theoretical Economics* 8, 775–800.

Danilov V., Koshevoy G., Savaglio E. (2014). Hyperrelations, Choice Functions, and Orderings of Opportunity Sets. Forthcomming in Social Choice and Welfare'.

Lleras J.S., Masatlioglu Y., Nakajima D., Ozbay E. (2012). When More is Less: Limited Consideration. Unpublished Paper.

Manzini P., Mariotti M. (2007). Sequentially Rationalizable Choice. *American Economic Review* 97, 1824–1839.

Manzini P., Mariotti M. (2010). Revealed Preferences and Boundedly Rational Choice Procedures: an Experiment. Unpublished Paper.

Manzini P., Mariotti M. (2012). Categorize then Choose: Boundedly Rational Choice and Welfare. *Journal of the European Economic Association* 10, 1141–1165.

Manzini P., Mariotti M. (2014). Stochastic Choice and Consideration Sets. *Econometrica* 82, 1153–1176.

Manzini P., Mariotti M., Tyson C.J. (2013). Two-Stage Threshold Representations. *Theoretical Economics* 6, 875–882.

Masatlioglu Y., Nakajima D., Ozbay E. (2012). Revealed Attention. *American Economic Review* 102, 2183–2205.

Zak F.L. (2014a). Psychological Games in the Theory of Choice. I. Temptation, Perfectionism, Self-deception. *Zhurnal Novoy ekonomicheskoy assotsiatsii* 1, 12–32 (in Russian).

Zak F.L. (2014b). Psychological Games in the Theory of Choice. II. Psychological Games in the Theory of Choice: Shame, Regret, Egoism and Altruism. *Zhurnal Novoy ekonomicheskoy assotsiatsii* 2, 12–40 (in Russian).

Поступила в редакцию 26 ноября 2014 года

Журнал НЭА,
№2 (26), 2015,
с. 12–35

V.I. Danilov
The Central Economic Mathematical Institute RAS, Moscow, Russia

Beyond Classical Rationality: Two-Stage Rationalization

Numerous facts indicate that behaviour (choice) of economic agents does not satisfy strong rationality criterions. One of possible explanation is in that the choice goes in two (or more) sequential stages. In the first stage, one chooses a subset F(A) from an initial menu A. Then the best element (with respert to some linear order) is chosen from F(A). The resulting choice is not rational in the classical sense. However it possesses some specific properties which distinguish it from arbitrary non-rational behaviour. In the paper we give a survey of recent works on two-stage choice. Axiomatic characterizations of such a choice are given depending on rationality conditions of the first step choice.

Keywords: *choice function; axioms of heredity, outcast, and concordance; filter; preferences; axiom of revealed preference; Plott function.*

JEL Classification: D71.