

ИЗБОР, ПРЕДПОЧИТАНИЕ, РАЦИОНАЛНОСТ

Росен Люцканов

CHOICE, PREFERENCE, RATIONALITY

Rosen Lutskanov

Abstract: The paper sketches a theoretical framework that has the required expressive power to deal with the phenomenon of preference reversal. This phenomenon is studied in three different contexts: rational choice theory, epistemology, and ethics. A common underlying theme is identified: the so-called higher-order interactions stemming from the epistemic value of the available alternatives. In order to capture them I employ the concept of a relational system. It is shown that this mathematical structure, along with several crucial concepts (regularity, stratifiedness, coherence), is enough to capture the essential features of preference reversal. Building on this insight, the paper proposes several new lines of research, based on the common structure exemplified by these seemingly different domains.

Keywords: rational choice theory, preference reversal, epistemic menus, relational systems.

Преди много години един млад студент се записва в лекционен курс, изнасян от един вече утвърден учен. Ретроспективно погледнато, срещата между тях не изглежда случайна, тъй като студентът е изключително любознателен, а утвърденият учен – една от най-ярките звезди на интелектуалния небосклон по онова време. Студентът се казва Кенет Ероу, а ученият – Алфред Тарски. Една от последиците от срещата им е появата на нова област на научни изследвания: формално-аксиоматичната теория на рационалния и социалния избор¹.

От много години философите, психолозите и икономистите се интересуват от възможността да прочетат мислите на хората, вглеждайки се в тяхното поведение. За целта те се фокусират върху изборите, които хората правят в различен тип ситуации. Когато избираме една алтернатива при положение, че е достъпна и друга алтернатива казваме, че проявяваме предпочитание към първата за сметка на втората. Когато тази релация на предпочитание представлява строга линейна наредба (релация, която е ирефлексивна, асиметрична, транзитивна и свързана) казваме, че тя рационализира нашите избори. Всяка система от избори, която може да бъде рационализирана по този начин има редица желани свойства: винаги може да бъде избрана точно една алтернатива и изборите в рамките на различни възможни „менюта“ от достъпни алтернативи неизбежно се съгласуват едни с други.

¹ Историята е предадена подробно в (Arrow 2014, 145).

Следвайки разработените от Тарски методи в областта на теорията на релациите, Ероу успява да покаже, че определени изисквания към системата от избори, които се осъществяват при наличието на различни „менюта“ от достъпни алтернативи гарантират съществуването на релация на предпочитание с желаните свойства. По-късно са открити различни аксиоми, които могат да служат за тази цел и да бъдат оправдани интуитивно. За да бъдат проучени тези аксиоматични системи са използвани различни математически техники².

Години по-късно, друг любознателен младеж се запознава с теорията на рационалния избор, чиито основи поставят изследванията на Ероу. Неговото име е Амартия Сен. В една от статиите си той демонстрира, че при определени условия релацията на предпочитание се преобръща: дори да предпочетем една алтернатива пред друга алтернатива в рамките на едно меню е напълно възможно и същевременно оправдано да предпочетем втората пред първата в рамките на друго меню. Това хвърля сянка на съмнение върху приравняването на рационалност и максимизиращо поведение, съответно се превръща в един от основните нерешени проблеми в теорията на рационалния избор. Имайки предвид това обстоятелство, ще се спрем накратко на един мисловен експеримент, който показва къде е проблемът.

Примерът, който привежда Сен е крайно елементарен³. Той ни приканва да си представим човек, към когото далечен познат отправя покана за чай. Менюто, от което той трябва да избере има два елемента: $x =$ „да приема поканата и да пия чай“ и $y =$ „да си остана вкъщи“. Според Сен при меню от вида $\{x, y\}$ няма нищо необичайно в това да изберем x , $C(\{x, y\}) = x$, иначе казано, да проявим предпочитание към x за сметка на y . Проблемът е там, че ако отправената покана бъде модифицирана, „да пием чай или да смъркаме кокаин“, което означава, че към менюто се прибавя трети елемент z , то същият този човек, без нищо в неговите разбирания за света и ценностната му ориентация да се е променило, би имал пълното право да преобърне предпочитанията си и да си остане вкъщи: $C(\{x, y, z\}) = y$.

Какви поуки да изведем от този пример? Този въпрос няма очевиден отговор. Според самия Сен примерът показва, че менюто, от което избираме, има „епистемична стойност“: то ни казва какъв е характерът на алтернативите, между които избираме. Как обаче да отчетем това обстоятелство? Настоящата статия предлага отговор на този въпрос. Неформално погледнато, той се свежда до това, че реконструкцията на система от проявени предпочитания изисква да изследваме не само релациите между достъпни алтернативи (елементите на менюто), а също и релациите между алтернативи и предпочитания между алтернативи (които задават тяхната епистемична стойност). Иначе казано, използваното понятие за релация трябва да бъде разширено по подходящ начин, за да отчете двойствената роля на алтернативите: на неща, между които избираме и на неща, които са определящи за начина, по който избираме.

Основен елемент на предложената формализация е понятието за релационна система. Това понятие също се появява в контекста на логиката и показва, че икономистите не могат да се справят добре с анализа на феномените, от които се интересуват, възползвайки се от еднократна доза логически „допинг“. Дефиницията за релационна система е рекурсивна: $S_0 = \{x, y, z, \dots\}$; $S_1 = S_0 \cup S_0 \times S_0$, ..., $S_k = S_{k-1} \cup S_{k-1} \times S_{k-1}$, ... Под *релационна система* \mathcal{R} с базис S_0 от ред k ще разбираме произволно подмножество на S_k . Както се вижда, стандартното понятие за релация отговаря на релационните системи от първи ред, така че в случая наистина имаме естествено обобщение на това понятие⁴.

За щастие, релационната система, която ще ни интересува тук е крайно елементарна: тя е построена върху множество от три алтернативи и е от втори ред. В нея ще имаме (а) елементи от вида x, y, z (алтернативи), (б) двойки от вида $(x, y), (y, z), (x, z)$ (релация на предпочитание между алтернативи, където (а, б) означава, че алтернативата а е предпочитана пред алтернативата б при сравнение по двойки), (в) двойка от вида $(z, (x, y))$ (мета-релация, изразяваща епистемичната

² За авторитетно въведение в темата, вж. (Bossert, Suzumura 2010), най-вече гл. 1 и 2.

³ Вж. (Sen 1993, 130–131).

⁴ Макар да има и други еквивалентни формулировки, тук се придържам към (Gabbay, Schlechta 2009, 417).

стойност на z по отношение на наредбата на двойката, съставена от x и y). Дори такава крайно елементарна информационна система изисква въвеждането на определени понятия, за да бъде анализирана по-подробно.

Първо ще дефинираме рекурсивно понятието за *основа* на даден елемент $M \in \mathcal{R}$, от вида $M = (A, B)$ като положим, че $g(M) = g(A) \cup g(B)$. Иначе казано, основата на даден елемент на релационната система е съставена от елементите на нейния базис, които участват в неговата конструкция съобразно с горната дефиниция. Използвайки това понятие ще въведем изискването за *регулярност*: за всяко $M \in \mathcal{R}$, от вида (A, B) ще изискваме $g(A) \cap g(B) = \emptyset$, т.е. основата на A и основата на B да нямат общи елементи. Това блокира възможността да съществуват конструкции от вида $(x, (x, y))$, в които x носи информация за себе си, преобръщайки отношението между нея самата и друга алтернатива.

По-нататък ще въведем още едно понятие: за даден елемент $M \in \mathcal{R}$ ще казваме, че неговият *ранг* $r(M)$ е минималната стойност на k , за която $M \in S_k$. При това положение очевидно алтернативите ще имат ранг 0, релацията на предпочитание ще има ранг 1, а мета-релацията, задаваща епистемична стойност – ранг 2. Ще казваме, че релационната система е *йерархична*, когато за всяко $M \in \mathcal{R}$ от вида (A, B) е изпълнено следното: $r(A) < r(B)$. Нашият пример е точно такъв: релацията на предпочитание между x и y (която е от ранг 1) е преобръната от алтернативата z (която е от ранг 0). Засега ще отменим въпроса дали си струва да бъдат разглеждани случаи, в които релационната система не е йерархична, макар да не съм склонен да изключа тази възможност.

Накрая ще добавим още едно изискване, което изглежда интуитивно защитимо: за дадена релационна система \mathcal{R} от ранг k под *ограничение* на \mathcal{R} до ранг m , $L^m(\mathcal{R})$, където $0 < m \leq k$ ще имаме предвид релационна система, която е конструирана по следния начин: (1) полагаме $L^m(\mathcal{R}) = \mathcal{R}$ и $i = m$; (2) ако $i > 1$, то за всеки елемент, от ранг i , който има вида $(A, (B, C))$ елиминираме $(A, (B, C))$ и (B, C) и добавяме (C, B) към $L^m(\mathcal{R})$. След това полагаме $i = i - 1$ и повтаряме стъпка (2). Процедурата по конструиране завършва щом получим $i = 1$. В нашия пример $L^1(\mathcal{R})$ ще включва наредените двойки (x, y) , (y, z) , (x, z) , докато $L^2(\mathcal{R})$ ще включва наредените двойки (y, x) , (x, z) , (y, z) поради наличието на $(z, (x, y))$. Ще казваме, че релационната система \mathcal{R} е *кохерентна*, когато всяко ограничение на \mathcal{R} съдържа линейна наредба на базиса на \mathcal{R} . В общия случай от дефиницията следва, че за всяко ограничение на релационната система ще имаме линейна наредба, която го рационализира и че при прехода от един ранг към следващия се допускат определен тип пермутации на елементите на наредбата в резултат от отчитането на релации от по-високо ниво.

След като въведохме понятието за релационна система и дефинициите за регулярност, йерархичност и кохерентност разполагаме с всичко необходимо, за да опишем въображаемия пример на Сен. Релационната система, която му съотнасяме има три елемента от ранг 0 ($x =$ „да пием чай“, $y =$ „да си остана вкъщи“, $z =$ „да смъркаме кокаин“); три елемента от ранг 1 ($(x, y) =$ „при избор между това да пием чай и да си остана вкъщи бих избрал да пием чай“, $(y, z) =$ „при избор между това да си остана вкъщи и да смъркаме кокаин бих избрал да си остана вкъщи“, (x, z) при избор между това да пием чай и да смъркаме кокаин бих избрал да пием чай); и един елемент от ранг 2 ($(z, (x, y)) =$ „присъствието на алтернативата да смъркаме кокаин води до преобръщане на предпочитанието ми да приема поканата, вместо да си остана вкъщи“). Можем да отбележим, че изискването за регулярност е спазено (автоматично за елементите от ранг 0 и 1, но и за елемента от ранг 2, тъй като z се различава от x и y). Изискването за йерархичност също е спазено, тъй като $r(\{z\}) < r(\{x, y\})$. Накрая, изискването за кохерентност също е спазено, тъй като и двете ограничения на релационната система задават линейна наредба.

Преди да продължим нататък ще си позволя да опиша накратко различните релационни системи върху „меню“ съставено от три елемента, които удовлетворяват формулираните по-горе изисквания. Както ще видим, това е лесно, тъй като броят им не е голям, но същевременно е полезно с оглед на възможните приложения на мини-теорията, която скицирам тук. На първо място, за система, в която елементите от ранг 0 са три (x, y, z) ще имаме система от първи ред, включваща елементи от вида (x, y) , (y, z) и (x, z) , за които изискването за кохерентност е спазено (разбира се, тъй като три елемента могат да бъдат наредени линейно по 6 различни начина, ще имаме 6

различни, но взаимно изоморфни системи от този тип; това уточнение важи и за случаите, които предстои да разгледаме). Това отговаря на тъждествената пермутация, която запазва поредността на елементите на базиса.

На второ място, ще имаме система от втори ред, която отговаря на описания от Сен случай: към разгледаната по-горе конфигурация ще имаме елемент от ранг 2 от вида $(z, (x, y))$, както и такава, която отговаря на случая $(x, (y, z))$. Третият комбинаторно възможен случай не може да се реализира при направените допускания, тъй като $(y, (x, z))$ задава циклична наредба на алтернативите в ограничението от втори ред на разглежданата релационна система, съответно изискването за кохерентност се нарушава. Съответно, тук ще имаме две различни линейни наредби за елементите от ранг 2, които отговарят на двете транспозиции на линейната наредба, съотнесена на ограничението от първи ред.

На трето място, ще разгледаме възможните системи от втори ред, които съдържат два елемента от ранг 2. Тук отново възможните случаи са два (с точност до изоморфизъм): първо, ще имаме системата с елементи от ранг 2 $(x, (y, z))$ и $(y, (x, z))$, при която последната алтернатива в наредбата (z) става първа и второ ще имаме системата $(y, (x, z))$ и $(z, (x, y))$, при която първата алтернатива (x) в наредбата става последна. Тези случаи отговарят на цикличните пермутации за множество с три елемента. Третият комбинаторно възможен случай, който има вида $(x, (y, z))$ и $(z, (y, z))$ не е допустим, тъй като задава циклична наредба на ограничението на системата от втори ред.

Накрая, съществува точно един възможен случай в който система с базис от три елемента има три елемента от ранг 2. Това (отново с точност до изоморфизъм) е системата с елементи $(x, (y, z))$, $(y, (x, z))$ и $(z, (x, y))$. При нея се наблюдава пълно преобръщане на релацията, съставена от елементи от ранг 1: нито една от наредените двойки не се запазва; това е случаят на огледалната пермутация на първоначалната наредба. За мен това е изроден случай в смисъл, че при него изборът в рамките на нито едно двуелементно множество от елементи на базиса не осигурява информация за тяхната „реална“ наредба, тъй като винаги третият елемент има епистемична стойност, която преобръща отношението между тях.

Това подсказва следната възможност: да говорим за нива на рационализация, съответно за нива на рационалност. Стандартният подход на Ероу отговаря на пълната рационализация, която се асоциира с тъждествената пермутация, запазваща изцяло изходната наредба. Последният разгледан случай отговаря на нулева рационализация, която се асоциира с огледална пермутация, преобръщаща напълно изходната наредба. Между тези две крайни точки попадат останалите два случая, отговарящи на другите пермутации.

Общото правило във всички случаи е следното: колкото по-малко се променя първоначалната наредба, толкова по-„рационален“ е изборът поради самото естество на ситуацията, в която той се осъществява. Струва ми се, че това е важна модификация на стандартния подход: рационалността на дадена система от избори не зависи единствено от субекта, който избира, съответно от неговите предпочитания, но и от ситуацията, в която се прави избор, тъй като алтернативите, от които избираме може да предоставят информация за това, което избираме, която да повлияе на проявените ни предпочитания.

До този момент скицирах концептуалния апарат, който ни позволява да изразим формално епистемичната стойност на алтернативите и процесът на преобръщане на предпочитанията. Струва ми се, че идеите, които бяха формулирани до този момент имат значение, което не се ограничава в рамките на икономиката, в частност на теорията на рационалния избор. Разглеждайки понятието за предпочитание по-общо и по-абстрактно, можем да открием преобръщане на предпочитанията и в други контексти. Тук ще разгледам съвсем накратко два примера: един почерпен от епистемологията и втори – от етиката, акцентирайки върху общите им характеристики.

За начало предлагам да разгледаме един казус, който на пръв поглед няма нищо общо с този на Сен, но на втори поглед според мен е идентичен на неговия. Имам предвид ... демонът, който Декарт използва, за да се освободи от някои от най-дълбоко вложените в ума му, възприети

сякаш веднъж завинаги мнения⁵. Отклонявайки се леко от оригиналния пример и от ролята, която е възложена на демона от Декарт в него, нека изберем произволна перцептивна ситуация: (струва ми се, че) виждам пред себе си червена ябълка. При това положение можем да разгледаме следното „меню“: $x =$ „имам сетивно усещане, каквото е налице, когато пред мен има червена ябълка и знам, че пред мен има червена ябълка“ и $y =$ „имам сетивно усещане, каквото е налице, когато пред мен има червена ябълка и не знам, че пред мен има червена ябълка“. Имайки предвид вероятността да стрешим при подобен тип перцептивна ситуация можем да въведем релация на епистемично предпочитание: (x, y) , която изразява обстоятелството, че в рамките на „менюто“ съставено от тези две алтернативи имам основание да предпочета x пред y , т.е. да приема, че имам знание.

Следвайки примера на Сен, въвеждаме трета, патологична алтернатива, която има епистемична стойност и чиято роля е аналогична на тази на кокаина в неговия пример: $z =$ „възможно е да съществува зъл демон, който е решен да ме заблуждава и той създава в мен измамно усещане, каквото е налице, когато пред мен има червена ябълка“. В нейно присъствие релацията на епистемично предпочитание се преобръща, иначе казано, налице е мета-релация от вида $(z, (x, y))$. В допълнение на това можем да отбележим, че отношението между x и y от една страна и z от друга страна е абсолютно същото като това в примера на Сен: (x, z) и (y, z) , тъй като основанията в подкрепа на x и y са по-силни от тези в подкрепа на z (по дефиниция не съществува разсъждение или сетивно възприятие, което обосновава убеждението, че съществува зъл демон, който ме мами, тъй като злият демон не би допуснал наличието на улики, които обосновават достоверното убеждение в неговото съществуване). Преди да изоставим приложната епистемологична демонология бих отбелязал, че разгледаните по-горе конфигурации ни осигуряват указания, следвайки които можем да конструираме все по-странни и по-странни сценарии, разколебаващи увереността ни в нашите мнения (задача, която няма да бъде решена тук)⁶.

В заключение бих направил още една препратка – този път в полето на етиката. Макар преобръщането на предпочитанията първоначално да е установено във връзка с инструменталните предпочитания, които изследва икономиката, няма да се наложи да търсим дълго, за да го установим и в полето на моралните предпочитания. Да вземем например най-известният и широко обсъждан мисловен експеримент в етиката: казусът с трамвая на Филипа Фут, заедно с една от добре познатите му разновидности, в която за да бъдат спасени петимата върху релсите шестият човек (наречен просто „дебелият мъж“) трябва да бъде блъснат на пътя на трамвая. В рамките на психологията на моралните съждения е установено отдавна, че е относително лесно да създадем предпочитание към активната намеса ($x =$ „да дръпнеш ръчката, която пренасочва трамвая по друг път“) за сметка на пасивното проследяване на развитието на ситуацията ($y =$ „да оставиш трамвая да продължи по пътя си“). От друга страна, известно е също, че когато въведем трета алтернатива ($z =$ „да блъснеш дебелия мъж върху релсите“) не просто ненамесата изглежда за предпочитане в рамките на менюто $\{y, z\}$, а се променя и първоначалното предпочитание в рамките на менюто $\{x, y\}$. Иначе казано, в мисловния експеримент на Фут получаваме същата конфигурация както в случая с кокаина на Сен и в мисловния експеримент с демона на Декарт⁷.

⁵ Макар че това едва ли е необходимо, ще припомня, че фигурата на „злия дух, изключително мощен и хитър, вложил всичките си усилия да ме заблуждава“ се появява в края на първото от Декартовите „Размишления върху първата философия“, които тук цитирам по новия превод на Мориарти (Descartes 2008, 16).

⁶ Друг важен въпрос, на който тук няма да имам възможност да обърна подобаващо внимание е това каква е природата на убежденията и релацията на епистемично предпочитание между тях. Казаното тук е напълно съвместимо с възможността убежденията да се третират като първични (primal beliefs), което пък означава, че нареждането и пренареждането им се осъществяват от предрефлексивни механизми на ниско ниво. За повече по темата, вж. (Angelova-Elchinova 2024, 126).

⁷ За оригиналната версия на мисловния експеримент, вж. (Foot 1967). За модифицираната версия, вж. (Singer 2005, 340). Относно психологическите механизми, които водят до промяна в модифицираната версия, вж. (Nado et al. 2020, 705).

Макар че казаното очевидно е само първа стъпка към разработването на обща формална теория на преобръщането на предпочитанията в различни контексти, струва ми се, че дори на този етап в изложението има резултат, който си струва да бъде отбелязан: макар и това да не е очевидно, налице са повтарящи се теми, които свързват част от най-трудните нерешени проблеми в икономиката, епистемологията и етиката. Откриването на връзката между тях очевидно е, ако не друго, то поне първа стъпка към търсеното решение. В добавка си струва да отбележим, че всеки от тези инак лишени от връзка теоретични контексти осигурява указания относно това накъде да поемем. Дали те могат да бъдат съчетани в единна теория е въпрос, чийто отговор ще потърся другаде.

ЛИТЕРАТУРА/REFERENCES

- Angelova-Elchinova, M. (2024).** Norms of belief and non-propositional primal beliefs, *Balkan Journal of Philosophy*, 16(2), 117–130. doi: 10.5840/bjp202416215.
- Arrow, L. (2014).** The origins of the impossibility theorem. In: J. Maskin & A. Sen (Eds.) *The Arrow Impossibility Theorem* (pp. 143–148). New York: Columbia University Press.
- Bossert, W., & Suzumura, K. (2010).** *Consistency, Choice, and Rationality*. Cambridge, Mass.: Harvard University Press.
- Descartes, R. (2008).** *Meditations on First Philosophy*, translated with an introduction and notes by Michael Moriarty. Oxford: Oxford University Press.
- Foot, Ph. (1967).** The problem of abortion and the doctrine of the double effect, *Oxford Review* 5, 1–7.
- Gabbay, D., & Schlechta, K. (2009).** Reactive preferential structures and nonmonotonic consequence. *Review of Symbolic Logic*, 2(2), 414–450. doi:10.1017/S1755020309090248.
- Nado, J. et al. (2020).** Moral Judgment. In: S. Robins et al. (Eds.) *Routledge Companion to Philosophy of Psychology*, 2nd ed. (pp. 697–707). New York: Routledge.
- Sen, A. (2002).** Internal consistency of choice. In: *Rationality and Freedom* (pp. 121–157). Cambridge, Mass.: Harvard University Press.
- Singer, P. (2005).** Ethics and intuitions, *Journal of Ethics* 9, 331–352. doi: 10.1007/s10892-005-3508-y.