

Е.Г. Драгалина-Черная
РАЦИОНАЛЬНОСТЬ ДЕЙСТВИЯ, РАЦИОНАЛЬНОСТЬ
ПРАВИЛА, РАЦИОНАЛЬНОСТЬ ЦЕЛИ:
РАССУЖДЕНИЕ КАК CASE-STUDY¹

В статье ставится задача систематизации многообразия форм рациональности в рассуждении. Предпринимается попытка экспликации нормативности логики на основе дихотомии 'рациональность действия / рациональность правила'. Стратегические ошибки в рассуждении объясняются с точки зрения рациональности цели.

The aim of this paper is to systematize the variety of rationality in reasoning. First, I'll try to adopt to logic the idea of rule-rationality which originates from behavioral economics. Second, I'll propose to shift the attention from rules to goals of reasoning and to the design of logical norms. What is gained, then, is a goal-rationality framework for modelling strategical mistakes in reasoning.

Ключевые слова: рациональность, действие, правило, цель, рассуждение, нормативность, логика.

Keywords: rationality, act, rule, goal, reasoning, normativity, logic.

1. Рациональность действия и нормативность логики

Со времен Аристотеля логику принято считать наукой о доказательном рассуждении. «Первая Аналитика» открывается следующим определением предмета исследования: «оно о доказательстве, и это дело доказывающей науки» (Аристотель, 1978: 119). Иммануил Кант обращает внимание на то, что общую логику следует рассматривать не как органон, а как канон правильного мышления, поскольку она способна сказать нам не о том, как действуют рассудок и разум, но лишь о том, как они должны действовать. «Под каноном, — пишет Кант, — я понимаю совокупность априорных

¹ В данной научной работе использованы результаты проекта «Рациональность в действии: интенции, интерпретации, интеракции», выполненного в рамках Программы фундаментальных исследований НИУ ВШЭ в 2015 году.

принципов правильного применения некоторых познавательных способностей вообще. Так, общая логика в своей аналитической части есть канон для рассудка и разума вообще, однако лишь со стороны формы, так как она отвлекается от всякого содержания» (Кант, 1994: 467).

Трактовка логики как нормативной теории правильных рассуждений остается мейнстримом среди современных подходов к демаркации дисциплинарных границ логики. Вместе с тем, психологические эксперименты последних десятилетий показали нетривиальность вычленения логической нормативности из всего многообразия факторов, влияющих на корректность и аргументативную убедительность обыденных рассуждений (см., например, Wason and Johnson-Laird, 1972; Newstead, Pollard, Evans and Allen, 1992; Evans, Newstead and Byrne, 1993; Fiddick, Cosmides and Tooby, 2000; Harman, 2002, Geurts, 2003; Pijnacker, Geurts, van Lambalgen, Buitelaar and Hagoort, 2010; Драгалина-Черная, 2014). Экспериментальные данные свидетельствуют о высокой частоте контекстных и культурных влияний, а также воздействии субъективных предпочтений и прошлого опыта на рассуждения агентов с ограниченными ресурсами о специализированной предметной области в определенной коммуникативной ситуации.

Так, в задаче на подавление (*Suppression Task*) Рут Бёрн (1989) испытуемым предлагалось сделать вывод из посылок (1) *Если ей надо написать эссе, она будет допоздна работать в библиотеке. Ей надо написать эссе.* Около 90% испытуемых приходили к заключению «*Она будет допоздна работать в библиотеке*». Однако при добавлении посылки (2) *Если библиотека будет открыта, она будет допоздна работать в библиотеке* лишь 60% процентов испытуемых делали такой вывод. Вместе с тем, при замене (2) посылкой (3) *Если у нее будет учебник, она будет допоздна работать в библиотеке* к выводу *Она будет допоздна работать в библиотеке* приходило практически столько же испытуемых, что и в случае (1).

Эффект «подавления» посылкой (2) правильного рассуждения по принципу *modus ponens*, выдерживающего «давление» посылки (3), нередко рассматривается как проявление нерациональности неискушенного в логике «человека с

улицы». Действительно, если правила логического вывода полагаются условием возможности рациональности как таковой, люди, статистически значимым образом нарушающие эти правила, должны быть признаны нерациональным. Но каким образом нерациональные существа могут адаптивно действовать, выживать и даже эволюционировать?

2. От действия к правилу действия: рациональность действия vs рациональность правила

Не только обыденные рассуждения, но и поведение реальных экономических агентов систематически нарушает вычислительную рациональность действия. Так, в знаменитой игре в ультиматум (*Ultimatum Bargaining Game*) (Güth, Schmittberger, Schwarze, 1982) двум игрокам выдают сто немецких марок и предлагают разделить их по взаимной договоренности. Игроки не встречались до игры и не встретятся после. Они общаются через компьютер, однако переговоры не ведут: первый игрок предлагает, другой соглашается или отвергает предложение, которое появляется на мониторе. Если второй игрок не принимает предложение первого игрока, никто ничего не получает. Стандартное предсказание для рационально действующих игроков, стремящихся к максимизации полезности: 99/1, то есть первому игроку рационально взять себе 99 марок, а второму — согласиться взять одну марку, ведь это лучше, чем ничего. Однако многочисленные эксперименты приводили к исходам между 50/50 и 65/35. Более того, второй игрок, которому предлагалось менее 30 процентов, как правило, отказывался, предпочитая 25 маркам «сладкую месть» нарушителю этических норм честной сделки.

Стремясь к синтезу принципов неоклассической и поведенческой экономик, лауреат Нобелевской премии Роберт Ауманн (2008) предложил объяснять нарушение рациональности действия (*act-rationality*) в игре в ультиматум особой рациональностью правила (*rule-rationality*). Идея рациональности правила восходит к различию, проведенному Джоном Ролзом между утилитаризмом действия и утилита-

ризмом правила, и состоит в том, что реальные экономические агенты стремятся к максимизации полезности не конкретного действия, но правила действия. Это правило не всегда принимается сознательно, зачастую следование ему является результатом эволюции и регулируется такими механизмами как «честь» и «альтруизм». Таким образом, конкретное действие, осуществляемое обычным человеком, не обязательно является рациональным с точки зрения максимизации «локальной полезности». Скорее люди следуют в своей повседневной практике правилам, которые приводят к оптимальным результатам «в целом», то есть в стандартных ситуациях, хотя в нетипичных случаях эти правила могут провоцировать ошибки и «локально нерациональное» поведение. Как отмечает Ауманн, люди «принимают правила или способы поведения, которые максимизируют меру тотальной, или нормальной, или ожидаемой полезности во всех тех ситуациях принятия решения, к которым это правило применимо; затем при принятии решения они выбирают действие, которое согласуется с принятым правилом» (Aumann, 2008: 2). К числу таких рациональных правил относятся стандарты справедливой сделки и наказания их нарушителей. Правило «Заботься о своей репутации – не позволяй себя унижать» эффективно работает в «повторяющихся играх» социального взаимодействия: тот, кто позволил себя унижить, рискует быть униженным вновь. Однако в игре в ультиматум оно не является оптимальным именно в силу нестандартности ситуации: отказ второго игрока взять «несправедливую» сумму денег не работает на его репутацию в силу анонимности и однократности его «альтруистического» выбора.

На первый взгляд, эффект подавления правильного рассуждения в задаче Бёрн представляется отличным подтверждением рациональности правила в обыденных рассуждениях с кондиционалами. Действительно, в таких рассуждениях испытуемые обычно исходят из предположения о замкнутости мира (*closed world assumption*) (Stenning and van Lambalgen, 2008). Согласно этому предположению, то, что явно не упомянуто как истинное, полагается ложным. В самом деле, если информация об исключениях отсутствует, то

предположить, что их вообще нет, как правило, рационально. Не принимать во внимание имплицитные исключения рационально, поскольку исключений в любой конкретной области обычно больше, чем правил, и их эксплицитная репрезентация является чрезвычайно ресурсоемкой, а зачастую практически нереализуемой задачей. В целом, рассуждение с предположением о замкнутости мира стратегически привлекательнее, чем дедукция, требующая с вычислительной точки зрения существенно большей затраты ресурсов. В типичных ситуациях это предположение работает эффективно и «в целом» максимизирует ожидаемую полезность. Исходя именно из этого предположения, люди зачастую оперируют правилом *modus ponens* в обыденных рассуждениях. *Modus ponens* с предположением о замкнутости мира перестает, однако, работать при наличии дополнительной посылки (3), поскольку факт, что библиотека закрыта, как раз и является тем исключением, которое ограничивает его применение.

Вместе с тем, теория логических ошибок, основанная на рациональности правила, вводит неочевидное допущение о принципиальной применимости такого типа рациональности к правилам логического вывода. Главный источник этой неочевидности в том, что правила вывода традиционно полагаются конститутивными правилами рациональности как таковой. Если регулятивное правило регулирует независимо от него существующие формы поведения, то конститутивные правила создают новые формы поведения. Подобно тому, как правила шахматной игры конституируют саму возможность этой игры, предположительно конститутивные правила вывода обеспечивают саму возможность рационального рассуждения. Как шахматная игра не существует вне правил шахматной игры, так и рациональное рассуждение не существует вне определяющих его правил вывода. Таким образом, если правило вывода конституирует возможность рационального рассуждения, попытка обосновать рациональность рассуждения апелляцией к этому правилу ведет к порочному кругу. Более того, конститутивность правил вывода в отношении рациональных рас-

суждений исключает приписывание различных мер полезности рассуждениям по различным правилам, а, следовательно, ранжирование таких рассуждений с точки зрения рациональности правила и, наконец, возможность логической ошибки как таковой.

«В знании, — утверждает Кант, — полностью согласующемся с законами рассудка, не бывает никакого заблуждения» (Кант, 1994: 215). Готлоб Фреге рассматривает рассуждения «логических чужаков», чьи «законы мышления» предположительно отличаются от наших, как род безумия, а не как ошибку или пример альтернативного мышления. «А что если, — задается он вопросом, — были бы обнаружены люди, чьи законы мышления категорически противоречат нашим и часто приводят поэтому к противоположным результатам даже на практике? Психологист в логике может лишь признать этот факт и сказать просто: те законы справедливы для них, эти законы справедливы для нас. Я должен сказать: мы имеем здесь неизвестную до настоящего времени разновидность безумия. Всякий, кто рассматривает законы логики как законы, которые предписывают способ, каковым должно думать, как законы истины, а не как естественные законы того, что человеческие существа полагают истиной, спросит, кто прав? Чьи законы полагания истинным соответствуют законам истины? Психологист в логике не может задать этот вопрос; в противном случае он признал бы законы истины, которые не были бы законами психологии» (Frege, 1964: 14). При этом Фреге отводит логике роль третейского судьи в разрешении конфликта мнений. «Если бы, — полагает он, — мы могли схватывать только то, что находится в нас, то конфликт мнений, основанный на взаимопонимании, был бы невозможен, поскольку отсутствовало бы общее основание, а никакая идея в психологическом смысле не способна обеспечить нам такого основания. Не было бы логики как признанного арбитра в конфликте мнений» (Frege, 1964: 17).

Таким образом, конфликт мнений может быть признан таковым лишь квалифицированным и объективным арбитром — логикой. «Логические чужаки», на которых не распространяется юрисдикция этого арбитра, не могут в

принципе вступить в рациональную полемику. Действительно, тот, кто нарушает правила футбола, удаляется с футбольного поля. Однако, уходя с поля, футболист все же подчиняется решению рефери и остается игроком в футбол. На каком основании удаляется с поля рациональности и подлежит дисквалификации «логический чужак», если он вообще не подпадает под юрисдикцию логики как идеального рефери игры по правилам рациональности? Кем назначен этот рефери, и кто уполномочил его совмещать судебские функции с законотворчеством на поле рациональности? Ответ на эти вопросы предполагает, на мой взгляд, переключение внимания с конститутивных правил логического вывода на его стратегическую цель, с априорной сущности логической нормы на ее дизайн.

3. От правила к дизайну правила: рациональность правила vs рациональность цели

Каждое действие, сообразное с целью, может быть представлено как игра. С теоретико-игровой точки зрения важно различать определяющие правила, которые специфицируют, как вообще играть в данную игру, и стратегические правила, следование которым позволяет максимизировать ожидаемую полезность, то есть играть в данную игру лучше (Хинтикка, 2000: 108). Определяющие правила логического вывода позволяют судить о том, является ли данное действие шагом логического вывода, но ничего не говорят об ожидаемой полезности этого шага. Поскольку полезность ассоциируется не с отдельным ходом, а со стратегией, мы можем говорить, что одно действие лучше или хуже другого, только если рассматриваем его как часть определенной стратегии. Искусство рассуждения состоит в стратегическом выборе и не может обсуждаться на уровне отдельных его шагов, иначе говоря, безотносительно к его стратегической цели. «Самостоятельная непсихологическая теория ошибок, — как подчеркивает Яаакко Хинтикка, — возможна лишь в той мере, в какой они рассматриваются не как нарушения логических “правил вывода”, а как стратегические ошибки. И из сказанного следует, что теорию стратегических ошибок в конечном счете невозможно построить как теорию ошибок вывода

и вообще как теорию того, какие конкретные шаги совершаются в логике или рассуждении» (Хинтиikka, 2000: 109).

Может ли, однако, логика нормировать стратегический выбор в рассуждении, исходя из деонтических полномочий рациональных агентов? «Главная роль институтов и смысл их существования, — отмечает Джон Сёрль, — не просто в ограничении действий людей, а скорее в создании новых видов отношений власти и влияния (power). В первую очередь человеческие институты создают новые возможности, поскольку учреждают особый тип полномочий, например права, обязанности, обязательства, санкции, разрешения, доверенности, требования и удостоверения. Все это я называю деонтическими полномочиями (deontic powers)» (Серл, 2007: 15). В институциональные отношения вступают не вещи или абстракции и даже не институты, но люди, наделенные деонтическими полномочиями: «то, что у меня есть ферзь на шахматной доске, не является следствием нахождения моей руки на определенном физическом объекте. Скорее это следствие того, что я обладаю некоторыми полномочиями производить действия относительно других фигур внутри формальной системы (то есть доски, хотя и не обязательно в физической форме). Точно так же наличие у меня в руках тысячи долларов не является следствием того, что у меня в руках стопки купюр, а скорее того, что я обладаю некоторыми деонтическими полномочиями. Теперь у меня есть право, то есть полномочие, делать покупки, которым я не обладал бы, не будь у меня денег» (Серл, 2007: 21). Какими же деонтическими полномочиями обладают рационально рассуждающий агент, с одной стороны, и дизайнер логической нормы, с другой? Может ли этот дизайнер руководствоваться в своем законотворчестве принципом толерантности Рудольфа Карнапа: «Наше дело не устанавливать запреты, но приходить к соглашению... В логике нет морали. Каждый вправе создавать свою собственную логику, то есть свою собственную форму языка, какую он пожелает» (Carnap, 1937: 51-52)?

Опасность конвенционализма в ответе на вопрос, что делает логику логикой, состоит в сближении логики с софи-

стикой, игнорирующей идеальную цель рассуждения. «Подобно этике, — замечает Фреге, — логика также может быть названа нормативной наукой. Как я должен мыслить, чтобы достичь цели, истины?» (Frege, 1964: 12). Солидаризируясь с Фреге в признании истины конечной целью логики, обязан ли, однако, дизайнер логической нормы взваливать на себя тяжкий груз его онтологии? «Мы должны признать, — настаивает Фреге, — что правила нашего мышления и нашего полагания истинным предписаны законами истины» (Frege, 1964: 12). Он убежден в том, что законы истины обладают вневременным объективным статусом в «третьем мире» вечных сущностей, вне всякой зависимости от того, осознает ли их кто-либо. Однако дизайнер логической нормы действует не в «третьем мире», а в институциональной реальности распределенных деонтических полномочий. Если конечной целью дизайнера холодильника является заморозка и безопасное хранение продуктов, то конечная цель дизайнера логической нормы состоит в достижении истины. «В целом истина значима для нас потому, — замечают Джордж Лакофф и Марк Джонсон, — что она способствует выживанию и позволяет нам выполнять свои функции в мире. Большинство накопленных нами истин — о наших телах, о людях, с которыми мы взаимодействуем, и о нашем ближайшем физическом и социальном окружении — используется в повседневной деятельности» (Лакофф и Джонсон, 2004: 188). Вместе с тем, собственные функции артефактов, ради осуществления которых они разрабатываются, не определяют однозначно их сущность (Драгалина-Черная, 2012). Хорошо известно, что практика коллективного использования артефактов зачастую обнаруживает их новые функции, повышающие их полезность в неожиданном направлении, не «запланированном» дизайнером, такие как, скажем, «мирные профессии» разработанных для военных целей девайсов или фармакологическое воздействие лекарственных препаратов на новые, не известные их создателям болезни. Логическая норма также зависит от коллективной *we*-интенциональности их создателей и пользователей. Нет внешних норм, в соответствии с которыми испы-

туемый в конкретном психологическом эксперименте обязан интерпретировать предъявленные ему послышки. Извлечение информации релевантно той цели, которую он ставит перед собой и которая может не совпадать с целью, подразумеваемой экспериментатором.

Согласно Чальзу-Сандерсу Пирсу, интерпретация представляет собой кооперативный диалог между Говорящим (*Utterer*), которого он также называет Утверждающим (*Assessor*) или Графистом (*Graphist, Graphist-mind*), и «творцом универсума», Интерпретатором (*Interpreter*), называемым также Оппонентом (*Opponent*) или Графеусом (*Grapheus, Grapheus-mind*). В неопубликованной работе 1905 года «Основание прагматицизма» («*The Basis of Pragmaticism*») Пирс утверждает, что «интерпретатор под руководством Графиста (*Graphist-mind*) обсуждает основание (*the rationale*) самой логики» (Peirce in Pietarinen, 2014: 125). При этом «Графист (*Graphist-mind*) и интерпретатор (*interpreter-mind*) должны обладать всеми характеристиками персонального интеллекта, наделенного дарами нравственного порядка» (Peirce in Pietarinen, 2014: 125). Таким образом, агенты, обсуждающие «основание самой логики», будучи теоретическими конструктами, обладают, по Пирсу, личностными чертами и моральными обязательствами.

Как отмечает Лукано Флориди, «мы должны по-новому взглянуть на природу норм: не только как на правила, устанавливающие отношения между агентами, но и как на конститутивных динамических агентов самих по себе» (Floridi, 2015:1012). Более того, следуя рекомендациям Пирса, полезно взглянуть на логические нормы как на включенных в диалог, наделенных деонтическими полномочиями и моральными обязательствами агентов. Нормативность логики означает, что тот, кто следует ее предписаниям, рассуждает формально для достижения конечной цели — истины. Согласно Пирсу, логика «не только предполагает правила, которым должно, но не необходимо следовать, но и оказывается анализом условий достижения чего-то, существенным ингредиентом чего является цель» (Пирс, 2001: 219). Он определяет логику как теорию взвешенного (*deliberate*) мышления, которое «контролируется с тем, чтобы сообразовать

его с какой-то целью или идеалом» (Пирс, 2001: 215). Формальность является условием «взвешенного» рассуждения, открытого критической оценке квалифицированного сообщества исследователей, а, следовательно, условием вменения ответственности рассуждающему агенту. «Все логические нормы, — отмечает Гордон Макфарлейн, — коренятся в схватывании мыслящим человеком выводов как имеющих определенную формальную структуру, ... за это схватывание он принимает на себя ответственность, критикуя или подвергаясь критике» (MacFarlane, 2004: 20). Макфарлейн подчеркивает также, что «мы предъявляем к логической правильности требование формальности, поскольку мы требуем от нее прозрачности, а прозрачности мы требуем в силу порождаемых ею аргументов и обязательств» (MacFarlane, 2004: 20).

Таким образом, стратегические ошибки в дизайне логических норм являются ошибками коллективной *we*-интенциональности, однако рассуждающий человек несет индивидуальную ответственность за следование этим нормам.

Литература

- Аристотель (1978) 'Первая Аналитика', *Аристотель. Сочинения в 4-х томах*, т. 2. М.: Мысль, 1978, с. 117-254.
- Драгалина-Черная, Е. Г. (2012) '«Добавленная реальность»: от интенций к институциям', *Онтологии артефактов: взаимодействие «естественных» и «искусственных» компонентов жизненного мира*. М.: Дело, 2012, с. 34-41.
- Драгалина-Черная, Е. Г. (2014) 'Поризматические сюрпризы в логике', *Рацио.ru*. 2014, № 13, с. 78-91.
- Кант, И. (1994) *Критика чистого разума*. М.: Мысль, 1994.
- Лакофф, Дж. и Джонсон, М. (2004) *Метафоры, которыми мы живем*. М.: Едиториал УРСС, 2004.
- Пирс, Ч.-С. (2001) *Принципы философии*. Том 2. Санкт-Петербург: Санкт-Петербургское философское общество, 2001.

- Серл, Дж. (2007) 'Что такое институт?', *Вопросы экономики*, 2007, №8, с. 5–27.
- Хинтиikka, Я. (2000) 'Действительно ли логика — ключ ко всякому хорошему рассуждению?', *Вопросы философии*, 2000, №11, с. 105-125.
- Aumann, R. J. (2008) *Rule-Rationality versus Act-Rationality* (Discussion Paper 497). Center for the study of rationality: The Hebrew University of Jerusalem, 2008.
- Byrne, R.M.J. (1989) 'Suppressing valid inferences with conditionals', *Cognition*, vol. 31, 1989, pp. 61–83.
- Carnap, R. (1937) *The Logical Syntax of Language*, New York, Harcourt, Brace, 1937.
- Evans, J., Newstead, S.L. and Byrne, R.M. (1993) *Human reasoning: the psychology of deduction*, Lawrence Erlbaum Associates, Hove, Sussex. 1993.
- Fiddick, L., Cosmides, L. and Tooby, J. (2000) 'No interpretation without representation: the role of domain-specific representations and inferences in the Wason selection task', *Cognition*, 2000, vol. 77, pp. 1–79.
- Floridi, L. (2015) 'Toleration and the Design of Norms', *Science and Engineering Ethics*, 2015, vol. 21, no. 5, pp. 1095-1123.
- Frege, G. (1964) *The Basic Laws of Arithmetic*, Berkeley: University of California Press, 1964.
- Geurts, B. (2003) 'Reasoning with quantifiers', *Cognition*, 2003, vol. 86, pp. 223–251.
- Güth, W., Schmittberger, R. and Schwarze B. (1982) 'An Experimental Analysis of Ultimatum Bargaining', *Journal of Economic Behavior and Organization*, 1982, vol. 3, pp. 367-388.
- Harman, G. (2002) 'Internal Critique: A Logic is not a Theory of Reasoning and a Theory of Reasoning is not a Logic', in D. M. Gabbay, R. H. Johnson, H. J. Ohlbach & J. Woods (Eds.), *Handbook of the Logic of Argument and Inference: The Turn Towards the Practical*. Amsterdam: Elsevier Science B.V., 2002, pp. 171–186.

- MacFarlane, J. (2004) *In What Sense (If Any) Is Logic Normative for Thought?* Draft of April 21, 2004.
- Newstead, S. E., Pollard, P., Evans, J. S. B. and Allen, J. L. (1992) 'The source of belief bias effects in syllogistic reasoning', *Cognition*, 1992, vol. 45, pp. 257–284.
- Pietarinen, A.-V.A. (2014) 'Logical and Linguistic Games from Peirce to Grice to Hintikka', *Theorema*, 2014, vol. 32, no. 2, pp. 121 – 136.
- Pijnacker, J., Geurts, B., van Lambalgen, M., Buitelaar J., and Hagoort, P. (2010) 'Reasoning with Exceptions: An Event-related Brain Potentials Study', *Journal of Cognitive Neuroscience*, 2010, vol. 23, no. 2, pp. 471–480.
- Stenning, K. and van Lambalgen, M. (2008) *Human Reasoning and Cognitive Science*, Cambridge: MIT Press., 2008.
- Wason, P. and Johnson-Laird, P.N. (1972) *The Psychology of Reasoning: Structure and Content*, Cambridge Mass., 1972.

Об авторе

Елена Григорьевна Драгалина-Черная – доктор философских наук, профессор, Национальный исследовательский университет «Высшая школа экономики», Школа философии, edragalina@gmail.com.

About author

Prof. Dr. *Elena Dragalina-Chernaya*, National Research University Higher School of Economics, School of Philosophy, edragalina@gmail.com.