

Научная статья

УДК 165

DOI 10.25205/2541-7517-2021-19-3-33-43

Рациональные основания, контрфактические высказывания и «невозможные миры» в философском обосновании мысленных экспериментов

Валентин Никонович Карпович

Институт философии и права
Сибирского отделения Российской академии наук
Новосибирск, Россия
kkvvn@gmail.com, <https://orcid.org/0000-0002-7897-7685>

Аннотация

Дэвид Льюис в своей теории законов отвергает подлинность невозможных миров на том основании, что противоречие, заключенное в рамках его модификатора «в (мире) w», равносильно противоречию во всей теории, что представляется неприемлемым. Вместе с тем, все чаще в философии рассматриваются невозможные, контрфактические миры и конструируются мысленные эксперименты с ними. Возникает необходимость расширения теории возможных миров на подлинные, конкретные, но невозможные миры. Один из способов сделать это заключается в отказе от принятого Льюисом классического отрицания на том основании, что оно ведет к проблемам полноты и совместимости миров. Такой подход совместим с экстенциональной метафизикой Льюиса, хотя и ведет к некоторым потерям полноты описания в семантике.

Ключевые слова

реализм, диалетизм, противоречие, возможный мир, экстенциональная семантика, контрфактические высказывания, классическая логика, философия, метафизика

Для цитирования

Карпович В. Н. Рациональные основания, контрфактические высказывания и «невозможные миры» в философском обосновании мысленных экспериментов // Сибирский философский журнал. 2021. Т. 19, № 3. С. 33–43. DOI 10.25205/2541-7517-2021-19-3-33-43

© Карпович В. Н., 2021

Rational Reasons, Counterfactual Statements and “Impossible Worlds” in the Philosophical Justifications of Thought Experiments

Valentin N. Karpovich

Institute of Philosophy and Law
of the Siberian Branch of the Russian Academy of Sciences
Novosibirsk, Russian Federation
kkvvn@gmail.com, <https://orcid.org/0000-0002-7897-7685>

Abstract

In his theory of natural laws David Lewis rejects the authenticity of impossible worlds on the grounds that the contradiction contained within his modifier “in (the world) w ” is tantamount to a contradiction in the whole theory, which seems unacceptable. At the same time, in philosophical discourse very often researchers use counterfactual situations and thought experiments with impossible events and objects. There is a need to apply the theory of worlds to genuine, concrete, but impossible worlds. One way to do this is to reject Lewis’s classical negation on the grounds that it leads to problems of completeness and inconsistency inside the worlds. The proposed extension for impossibility is compatible with Lewis’s extensional metaphysics, although it leads to some loss for description completeness in semantics.

Keywords

realism, dialetism, contradiction, possible world, extensional semantics, counterfactual statements, classical logic, philosophy, metaphysics

For citation

Karpovich V. N. Rational Reasons, Counterfactual Statements and “Impossible Worlds” in the Philosophical Justifications of Thought Experiments. *Siberian Journal of Philosophy*, 2021, vol. 19, no. 3, pp. 33–43. (in Russ.) DOI 10.25205/2541-7517-2021-19-3-33-43

Философский реализм в отношении возможных миров философски и методологически значим и применяется в философских построениях, а допущение невозможных миров при этом обычно выполняет инструментальную, техническую цель. В связи с этим важно понятие «диалетейя», обозначающее утверждение A такое, что верны и оно само, и его отрицание, $\neg A$. Если возможные миры допускают индивидуацию суждений и предикатов для наборов миров и их составляющих, невозможные миры расширяют этот прием до необходимых и невозможных, в частности, содержащих диалетейи. Возможные миры допускают оценку контрфактических суждений на истинность, а невозможные миры распространяют этот прием на сослагательные наклонения с ложным и даже невозможным антецедентом, т. е. допускают рассмотрение невозможного.

В литературе есть предложения допустить невозможные миры как прием мысленного эксперимента, но тогда реализм Дэвида Льюиса по поводу возможных миров, приспособленный для анализа законов природы, не укладывается в возникающую философскую концепцию. Действительно, Льюис (1986) сам отверг такое расширение своей теории. Вместе с тем, идея невозможных миров становится все более популярной в философских теориях, и это можно считать затруднением для радикального реализма. Как сказано у Сталнакера, «...если ...модальный реализм не может найти место для невозможных миров, тем хуже для ...модального реализма» [Stalnaker, 1996, p. 199]. В связи с этим необходимо очертить концептуальное пространство для подлинно невозможных миров с помощью аргументов самого Льюиса, которые он предлагал против допущения невозможных миров.

Основная идея в пользу допущения невозможных миров проистекает из так называемой «зернистости» (*granularity*). Возможные миры помогают индивидуализировать отдельные суждения в границах возможных миров, но при этом они предлагают слишком грубое (дискретное) описание, что не позволяет адекватно различать все интуитивно различные мыслительные конструкции.

Приведем примеры. В математике пытались определить «квадратуру круга», вписать круг в квадрат, тем самым клали в основу поиска решения противоречие, а значит, связывали несвязуемое. $1 + 1 = 5$ так же нелепо в возможном мире, как и утверждение, что человек там может быть одновременно мертв и не мертв. При этом обозначенные ситуации в принципе содержательно различны, и можно обдумывать каждую из них без связи с другой. Ситуация аналогична, когда рассматривают аналитически истинные, логически необходимые утверждения, которые обычно считают истинными во всех мирах. И в них можно считать, что $5 + 7 = 12$, но при этом не знать, что утренняя звезда – она же вечерняя, и можно не осознавать всех следствий из принятых утверждений, т. е. теорию в ее полноте, поскольку минимальное требование к теориям – это дедуктивное замыкание.

Таким образом, использование невозможных миров на самом деле позволяет различать интуитивно различные суждения за пределами действительного мира (см: [Barwise, 1997; Berto, 2010; Jago, 2014; Priest, 2016]).

Еще одна ключевая мотивация для рассмотрения невозможных миров – это желание построить семантику для невозможных условных выражений, которая не делает их все истинными из-за противоречивости оснований. Согласно стандартной семантике Льюиса – Сталнакера, контрфактические высказывания (а таковы законы природы!) истинны только тогда, когда все миры, наиболее близкие к нашему (или, что более важно, похожие на наш), в которых истинен

антецедент (основание) импликации, обеспечивают и консеквент (следствие). Это могут быть даже просто миры, в которых основание просто ложно, но не по предположению, а по факту (см.: [Lewis, 1973; Stalnaker, 1968]). При этом без невозможных миров такая же трактовка не распространяется на них, поскольку основания таких импликаций ложны во всех мирах, и следовательно, они тривиально удовлетворяют условиям истинности таких условных суждений, в том числе законов природы. Понятно, что то же самое верно применительно в случае истинности консеквента (следствия), в силу особенностей так называемой материальной импликации, когда действуют принципы «из лжи все что угодно» и «истина из чего угодно».

Философы (и математики) довольно часто допускают невозможное, как указывает Дэниел Нолан [Nolan, 1997, pp. 544–546], и часто используют такие предположения в логике и метафизике. К такому приему прибегают для оценки взаимоисключающих метафизических теорий или дедуктивных систем, например, чтобы оценить их следствия. И хотя с эпистемологической точки зрения это возможно и оправдано, если одна из таких теорий верна, то, по-видимому, она верна с необходимостью, и, следовательно, оппоненты обязательно неправы, то есть их предположения описывают невозможный мир. При всем этом само содержание этих теорий нетривиально, а в самой теории все следствия из предположений не являются странными, поскольку обоснованы этими гипотезами.

Кажется маловероятным, что два философа сильно расходятся во мнениях, скажем, о поведении отрицания или о природе свойств, но все же, по крайней мере, один из них произносит очевидную истину. Более разумной интерпретацией было бы сказать, что философы, математики, логики и т. п. способны вести значимые споры по таким вопросам и выдвигать иногда истинные, а иногда ложные утверждения. Берит Бругард и Джо Салерно выразили это так: «Возможно, большая часть философии бессодержательна, неинформативна и ошибочна. Но если это так, то это не из-за систематического неправильного использования контрфактических суждений» [Brogaard, Salerno, 2013, p. 644]. Наоборот, невозможные миры допускают содержательные философские обсуждения и расширяют стандартную трактовку контрфактических предположений на солагательные наклонения всех видов.

Несмотря на это привлекательное использование невозможных миров, Дэвид Льюис в известной сноске своей работы «О множественности миров» (1986) утверждает, что такие миры исключаются его «подлинно реалистической» метафизикой миров «подлинным реализмом». Для этого подлинного реализма «каким бы мир ни был, это все-таки мир» [Lewis, 1986, p. 75], а миры являются подлинными (или конкретными) целостностями, которые не отличаются от нашего

мира по своей сути. Результирующая структура включает в себя базовую онтологию индивидов, множеств и их комбинаций, оформленных с помощью классической логики. Аргумент против невозможных миров обращается к алетической природе репрезентации Льюиса, а именно к тому, что должно быть истинным в каком-то из допустимых миров. Для того чтобы что-то было истинным в таком мире или по его законам, оно истинно в буквальном смысле слова при квантификации по всем объектам этого мира. Вот как это обосновывает сам Льюис:

«Обсуждение ограничивающих модификаторов (кванторов или функторов с заданной областью действия. – В. К) позволяет мне сказать, почему мне не нужны невозможные миры в дополнение к возможным. Для разъяснения ситуации предположим, что путешественники рассказывают о каком-то месте в этом мире, где истинны противоречащие утверждения – волшебной горе, далеко в джунглях. Якобы у нас есть истины вида «На горе верны и Р, и не Р». Но если «на горе» – это ограничивающий модификатор, который задает ограничение областей неявной и явной квантификации определенной части всего, что есть в области, то он не влияет на функциональные связи истинности. Тогда порядок модификатора и связок не имеет значения. «На горе и Р, и Q» эквивалентно конъюнкции «на горе Р и на горе Q»; точно так же «на горе не Р» эквивалентно «на горе нет Р»; складываем полученное вместе и получаем, что предполагаемая истина «на горе Р и не Р» эквивалентно явному противоречию «на горе Р» и «на горе нет Р». Получается, что нет разницы между противоречием внутри области действия модификатора и явным противоречием. Это значит, что рассказ о чудесах с противоречиями, которые происходят на горе, ничем не отличается от противоречия как такового. Но нет предмета, каким бы волшебным он ни был, о котором вы можете высказать истину, противореча при этом самому себе. Поэтому нет горы, где противоречия истинны» [Lewis, 1986, p. 7].

Учитывая то, как устроены подлинные миры, модификатор «в w» (здесь w – один из миров) логически прозрачен. Значит, следуя Льюису (1968), где схема перевода для функциональных связей истинности изложена полностью, условия истинно-в-w для конъюнкции и отрицания заключаются в следующем:

- [1] в w $(A \wedge B)$ если и только если $[(\text{в w } A) \wedge (\text{в w } B)]$
 (буквы А и В будут обозначать суждения, если не сказано другое);
 [2] в w $(\neg A)$ если и только если $\neg(\text{в w } A)$.

Это дает нам посылку (i):

- (i) $[\text{at w } (A \wedge \neg A)]$ если и только если $[(\text{в w } A) \wedge \neg(\text{в w } A)]$.

Подставляем предположение (ii) относительно противоречивости мира:

- (ii) в w $(A \wedge \neg A)$,

получим в качестве заключения прямое противоречие:

(iii) $(\forall w A) \wedge \neg(\forall w A)$.

Теперь, используя правило «от противного», Льюис отвергает предположение (ii) и тем самым существование невозможных миров.

Хотя правильность предъявленных выводов очевидна с классической точки зрения, возникает вопрос: а все ли правильно, когда такие рассуждения применяются к невозможным мирам? Возможность легко оспорить аргумент Льюиса дает посылка (i).

Этот способ основан на возражениях к определению истины для логических связей у Льюиса, в частности конъюнкции и отрицания, поскольку именно с помощью такого определения истины Льюис объясняет, что в его теории нет места для невозможных миров. При ближайшем рассмотрении неясно, можно ли придерживаться [2], рассматривая гипотезу невозможных миров, потому что это ставит под вопрос саму гипотезу несовместимых и неполных миров (см.: [Lucan, 1994, p. 40; Routley, 1980, §1.10; Stalnaker, 1996; Priest, 1997, p. 485]).

[2] – это всего лишь классические условия истинности для отрицания, применяемых к понятию «истина в w». Льюис утверждает, что миры непротиворечивы и полны. Однако в невозможных мирах это не получится, они по условию противоречивы.

В этом духе Льюис указывает, что [2] не выполняется, если миры больше похожи на выдуманные истории:

«Иван говорит, что не Р» и «неверно, что Иван говорит Р» независимы: либо оба истинны, либо только одно истинно, либо ни одно из них не истинно. Если бы миры были похожи на сказки... действительно было бы место для миров, в которых противоречия истинны. Печальная правда о многозначности и непоследовательности этих миров сама по себе ничему бы не противоречила бы» [Lewis, 1986, p. 7].

Это верно. Но верно и то, что миры Льюиса содержат конкретные, подлинные фрагменты реальности, и по ним можно квантифицировать, а значит, не очевидно, почему это делает возможным шаг [2]. Рассмотрим [5]:

[5] Иван говорит $(\neg A)$ тогда и только тогда, когда $\neg(\text{Иван говорит } A)$.

Почему [5] ложно? Здесь условия истинности простые: нечто ложно, если значения истинности левой и правой сторон не совпадают, то есть, если оба, одно или ни одно не могут быть истинными. Такое может быть? «Иван говорит $(\neg A)$ » может быть правдой, а « $\neg(\text{Иван говорит } A)$ » ложью, если Иван противоречит сам себе. Точно так же « $\neg(\text{Иван говорит } A)$ » может быть истинным, а «Иван

говорит $(\neg A)$ » ложно, когда Иван не может произнести или опровергнуть A . Получается, что [5] ложно, поскольку утверждение Ивана может быть непоследовательным или неполным. Но ведь как раз это и подразумевается для невозможных миров, реально существующих или мыслимых. В таких мирах [2] может не быть истинным по тем же причинам, что и для высказывания [5]. Сохранение [2] при выводе посылки (i) сразу предопределяет отказ от гипотезы невозможных миров, существующих или лишь предполагаемых.

Здесь можно возразить, что [2] выражает лишь собственное определение Льюиса того, как себя ведет отрицание-в- w . Хотя и важно сохранить как можно больше исходных принципов реализма, но естественно допустить, что расширение такой классической теории до логически невозможного потребует некоторого пересмотра логики. При этом [2] главный кандидат на исключение, поскольку именно в нем утверждается, что все миры ведут себя классически.

Здесь возникает другое возражение: отрицание [2] означает отказ от классической логики, а значит, существенный отказ от основных логических и метафизических положений исходной теории самого Льюиса. Однако отказаться нужно лишь от классической схемы перевода для отрицания в мире, а не отрицания как такового. Нужно лишь сохранить классическое «прямое» отрицание в рассуждениях о мирах (не в мирах), чтобы там оно было классическим, и выглядело обычно:

[6] $\neg B$ is истинно, если и только если B неверно [Priest, 1997].

Понятно также, что от [2] нужно отказаться, чтобы сохранить [6]. [2] может быть прочитано как эквивалентность лжи-в- w для предложения A и ложность другого предложения B применительно к w , где $[B] = [w \ A]$. Тогда, если $(w \ A)$ и $(w \ \neg A)$, для некоторого w , [2] дает нам $(w \ A)$ и $\neg(w \ A)$, т.е. простое $B \wedge \neg B$, что несовместимо с [6]. Отказ от [2] – именно то, что предупреждает этот шаг, делает его невозможным, и позволяет языку теории и, следовательно, трактовке истины, оставаться классической.

У Льюиса есть и независимые веские причины придерживаться [2], то есть классических условий истинности отрицания в- w .

Самый сильный аргумент в пользу сохранения [2] основан на экстенциональной метафизике подлинного реализма, в частности, на его алетической природе репрезентации. Действительный мир w представляет пропозицию как «истинную в w », поскольку это высказывание является истиной о w , то есть буквально описывает w . В этом смысле предложения являются для Льюиса своего рода свойством, предикатом для миров. Он так и говорит:

«Я отождествляю суждения с определенными свойствами, а именно с теми, которые реализуются (представлены) только в возможных мирах» [Lewis, 1986, p. 53].

Таким образом, можно считать, что репрезентация представляет собой своего рода предикацию для Льюиса, простую предикацию определенного предложения А миру w согласно [7]:

[7] в w А тогда и только тогда, когда Aw .

[7] говорит, что А истинно в w только тогда, когда А действительно описывает w . Точно так же "А ложно в w " выглядит теперь так:

[8] при w $\neg A$ тогда и только тогда, когда $\neg Aw$;

[8] просто говорит, что $\neg A$ истинно в мире w , когда неверно, что А там истинно. Более точно:

[9] в w ($\neg A$) тогда и только тогда, когда $\neg(Aw)$.

[9] теперь просто повторение [2]:

[2] в w ($\neg A$) тогда и только тогда, когда $\neg(w \text{ в } A)$ Но если [9] является условием истины для отрицания, которое вытекает непосредственно из [7], а именно из метафизической трактовки репрезентации у Льюиса, то [2] является неотъемлемой частью подлинного реализма, не оставляя места для подлинно невозможных миров.

Однако, строго говоря, базовое понятие алетической репрезентации, зафиксированное в [7] (как буквальное присвоение свойства миру) не дает нам [9]. [7] само по себе ничего не говорит о лжи и остальных истинно-функциональных связках. Возможно, условие истинно-в- w для отрицания [8], которое строго вытекает из [7], не позволяет различать [9] и более слабый принцип, а именно [10]:

[10] в w ($\neg A$) тогда и только тогда, когда $(\neg A)w$.

Это непрозрачное, узкое толкование отрицания как модификатора свойства, зафиксированного в [10], по-прежнему совместимо с основной экстенциональной метафизикой истины Льюиса, представленной в [7]. [10] просто утверждает, что предикат ($\neg A$) действительно может быть приписан миру w . Это оставляет место для настоящих невозможных миров, в которых противоречия не перетекают на внешний язык описания миров. Таким образом блокируются те шаги, которые реализуют посылку (i) аргумента Льюиса.

Понятие истины в мире только как буквальной предикации, как это представлено в [7], не дает нам полной схемы перевода Льюиса, и это легко увидеть, обратившись к другим логическим связкам. Возьмем, к примеру, «и», конъюнк-

цию. Более слабая, логически непрозрачная схема для «и», вытекающая из [6], дает нам не [1] выше, а [11]:

[11] в $w (A \wedge B)$ тогда и только тогда, когда $(A \wedge B)w$.

Чтобы получить полную схему перевода для конъюнкции, то есть [1]:

[1] в $w (A \wedge B)$ тогда и только тогда, когда $[(w A) \wedge (w B)]$,

нам нужен еще один шаг, а именно экспортировать «и» и разбить суждение на части. Парадоксально, но основная метафизика репрезентации у Льюиса не требует этого шага, поскольку трактует истину как оценку конкретной области и буквально приписывает свойство мирам. Это всего лишь прием в его экстенциональной классической семантике. Метафизика алетической репрезентации Льюиса дает нам в лучшем случае такое соотношение для любой формулы φ :

[12] в $w \varphi$ тогда и только тогда, когда φw .

Анализ формулы не описан и невозможен. Экстенциональный характер логических связей простирается только до этих границ, т. е. до объявления, что связки понимаются и действуют классически во всех мирах. Но ведь по предположению невозможности мира это не обязательно так, а значит, нет и обязательства поступать именно таким образом.

В результате, приняв эту непрозрачную схему перевода для логических связей на базовом уровне, мы можем рассматривать все виды действительно реальных невозможных миров. У нас могут быть миры, например, в которых не работают даже простейшие логические истины, например, мир, в котором $A \wedge B$ истинно, но при этом A не является истинным (хотя не обязательно ложно), и другие подобные логические особенности, пока не задана логическая структура миров на базовом уровне. Там, где мы имеем дело только с мирами, которые непротиворечивы и полны, мы можем просто усилить теорию и вернуться к первоначальному классическому определению истины Льюиса.

Общий вывод такой: есть определенные преимущества для рассмотренного расширения теории возможных миров за счет включения невозможных ситуаций в метафизическую картину Льюиса, поскольку она выстроена для объяснения законов как утверждений в сослагательном наклонении: если бы вещество имело химическую формулу H_2O , то оно обладало бы определенными свойствами. Однако не хватает возможности более точного процесса описания структуры мира, как это сделано для действительного мира, например, в определении истины Тарского. Можно ли предложить механизм для преодоления «зернистости»

возможного мира, не очень ясно. Возможно, для этого потребуются какие-то дополнительные допущения, и, возможно, не очень фантастические.

Список литературы / References

- Barwise J.** Information and impossibilities. *Notre Dame Journal of Formal Logic*, 1997, vol. 38(4), pp. 488–515.
- Brogaard B., Salerno J.** Remarks on counterpossibles. *Synthese*, 2013, vol. 190 (4), pp. 639–660.
- Jago M.** Constructing worlds. *Synthese*, 2012, vol. 189(1), pp. 59–74.
- Jago M.** *The Impossible: An Essay on Hyperintensionality*. Oxford, 2014.
- Jago M.** Impossible worlds. *Noûs*, 2015, vol. 49 (4), pp. 713–728.
- Lewis D.** Counterpart theory and quantified modal logic. *Journal of Philosophy*, 1968, vol. 66, pp. 113–126.
- Lewis D.** *Counterfactuals*. Oxford, Blackwell, 1973.
- Lewis D.** *On the Plurality of Worlds*. Oxford, Blackwell, 1986.
- Lycan W.** *Modality and Meaning*. Dordrecht, Kluwer, 1994
- Mares E.** Who's afraid of impossible worlds? *Notre Dame Journal of Formal Logic*, 1997, vol. 38 (4), pp. 516–526.
- Nolan D.** Impossible worlds: a modest approach. *Notre Dame Journal of Formal Logic*, 1997, vol. 38 (4), pp. 535–572.
- Priest G.** Editor's introduction. *Notre Dame Journal of Formal Logic*, 1997, vol. 38 (4), pp. 481–487.
- Priest G.** *Towards Non-Being: The Logic And Metaphysics of Intentionality*. Oxford Uni. Press, 2016.
- Restall G.** Ways things can't be. *Notre Dame Journal of Formal Logic*, 1997, vol. 38 (4), pp. 583–596.
- Routley R.** *Exploring Meinong's Jungle And Beyond*. Canberra, Australian National Uni. Press, 1980.
- Stalnaker R.** Impossibilities. *Philosophical Topics*, 1996, vol. 24, pp. 193–204.
- Stalnaker R.** A Theory of Conditionals. In: N. Rescher (ed.). *Studies in Logical Theory*. Oxford, Blackwell, 1968, pp. 98–112.
- Vander Laan D.** The ontology of impossible worlds. *Notre Dame Journal of Formal Logic*, 1997, vol. 38 (4), pp. 597–620.
- Varzi A.** Inconsistency without contradiction. *Notre Dame Journal of Formal Logic*, 1997, vol. 38 (4), pp. 621–639.
- Zalta E.** A classically-based theory of impossible worlds. *Notre Dame Journal of Formal Logic*, 1997, vol. 38 (4), pp. 640–660.

Информация об авторе

Валентин Никонович Карпович

доктор философских наук

ведущий научный сотрудник Института философии и права СО РАН

Information about the Author

Valentin N. Karpovich

Doctor of Sciences (Philosophy)

Leading Research Assistant, Institute of Philosophy and Law SB RAS

Статья поступила в редакцию 11.09.2021;

одобрена после рецензирования 28.09.2021; принята к публикации 28.09.2021

The article was submitted 11.09.2021;

approved after reviewing 28.09.2021; accepted for publication 28.09.2021