АНАЛИТИЧЕСКАЯ ФИЛОСОФИЯ, ЭПИСТЕМОЛОГИЯ И ФИЛОСОФИЯ НАУКИ

УДК 165
DOI 10.25205/2541-7517-2019-17-3-5-16

Интенсиональность математического дискурса и теорема Лёба

В. В. Целищев

Институт философии и права СО РАН
Новосибирск, Россия

Аннотация

Рассмотрена проблема интенсиональности математического дискурса в свете теоремы Лёба. Исходя из факта эквивалентности теоремы Лёба и Второй теоремы о неполноте Гёделя, а также факта интенсиональности последней, формулируется проблема демонстрации интенсиональности теоремы Лёба. Показано, что эта интенсиональность имеет неявный характер, объясняемый «странностями» (по выражению Булоса) этой теоремы.

Ключевые слова

интенсиональность, теорема Лёба, Вторая теорема Гёделя, доказуемость

Благодарности

Работа выполнена при финансовой поддержке РФФИ (проект № 19-011-00518)

Для цитирования

Целищев В. В. Интенсиональность математического дискурса и теорема Лёба // Сибирский философский журнал. 2019. Т. 17, № 3. С. 5–16. DOI 10.25205/2541-7517-2019-17-3-5-16

Intentionality of Mathematical Discourse and Löb's Theorem

V. V. Tselishchev

Institute of Philosophy and Law SB RAS
Novosibirsk, Russian Federation

Abstract

The article deals with the problem of the intensionality of mathematical discourse in the light of the theorem of Löb. Proceeding from the fact of equivalence of the Löb theorem and the Second Gödel incompleteness theorem, as well as the fact of the intensionality of the latter, the problem of demonstrating the intensionality of Löb's theorem is formulated. It is shown that this intentionality has an implicit character, explained by the "weirdness" (as expressed by G. Boolos) of this theorem.

Keywords

intensionality, Löb theorem, second Gödel theorem, demonstrability

Acknowledgements

The Investigations are supported by Russian Fondation for Fundamental Studies (project no. 19-011-00518)

For citation

Tselishchev V. V. Intentionality of Mathematical Discourse and Löb's Theorem. *Siberian Journal of Philosophy*, 2019, vol. 17, no. 3, p. 5–16. (in Russ.) DOI 10.25205/2541-7517-2019-17-3-5-16

Наиболее известной особенностью гёделевых теорем о неполноте является существование неразрешимого, так называемого «гёделева предложения», которое говорит о том, что оно не доказуемо. Построение Гёделем такого самореферентного предложения осуществляется арифметизацией синтаксиса. Самореферентность напрямую связана с парадоксами, и конструирование гёделева предложения идет бок о бок с формулировкой Парадокса Лжеца. Наличие значительного числа других парадоксов подразумевает интересные метаматематические результаты, самым известным из которых в последние десятилетия является теорема Лёба [Löb, 1955].

Ее появление связано с поиском некоторого рода симметрии по отношению к гёделевскому результату. Пусть имеется формальная система Σ, в которой рассматривается утверждение φ. По Гёделю,

$$\Sigma \vdash \varphi \longleftrightarrow \neg \mathrm{Pr}\,(\ulcorner \varphi \urcorner),$$

где угловые скобки обозначают гёделев номер утверждения φ, а Pr – предикат доказуемости в формальной системе. Эта формула может быть интерпретирована как утверждение φ о своей недоказуемости. Исходя из расплывчатых соображений о симметрии, рассмотрим очень похожую формулу

$$\Sigma \vdash \varphi \longleftrightarrow \mathrm{Pr}\,(\ulcorner \varphi \urcorner),$$

где утверждение φ говорит о своей доказуемости. Л. Генкин поставил вполне допустимый вопрос: если φ у Гёделя говорит о своей недоказуемости и приведенные формулы похожи, является ли φ доказуемой [Henkin, 1952]. Такая (можно сказать, невинная) формулировка вопроса напрямую привела к постановке проблемы интенсиональности математического дискурса. Проблема оказалась трудной, и ее решение связано с рядом догадок и результатов о природе гёделевского доказательства. Гёдель настолько удачно и изобретательно сконструировал свое уникальное неразрешимое предложение (далее – G), что поначалу трудно было понять общность такого рода конструкций. Понадобилось острое внимание философа Рудольфа Карнапа, который заметил, что в основании гёделевских теорем лежит общая формула, ныне называемая диагональной леммой или теоремой о неподвижных точках. Эта формула имеет следующий вид:

$$\Sigma \vdash \varphi \longleftrightarrow \psi \ulcorner \varphi \urcorner.$$

Как видно, отрицание гёделевского предиката доказуемости – частный случай некоторой формулы ψ, которая является неподвижной точкой для φ. Общность определенного рода формулы ψ предполагает некоторую свободу в конструировании аналога гёделевского результата. Другими словами, вместо предиката доказуемости можно использовать другие выражения, смысл которых заключается в схватывании концепции доказуемости другими концептуальными средствами. Тогда ответ на вопрос, доказуема или нет формула φ, упирается в то, какие средства должны быть выбраны для этого «схватывания». Первый важный шаг в этом направлении сделал Г. Крайзель, который сконструировал два предиката, выражающих доказуемость, а также предложение, которое выражает свою собственную доказуемость, способом, отличным о гёделевского [Kreisel, 1953]. Сама по себе зависимость от способа выражения смысла формального предложения говорит об интенсиональном характере контекста.

С формальной точки зрения это говорит о том, что неподвижные точки могут вести себя по-разному. Это является крайне нетипичным обстоятельством в математическом дискурсе. Действительно, разговор об интенсиональности последнее время в существенной степени связан с интенсиональным характером доказательства Второй теоремы Гёделя о неполноте. Можно сказать, что интенсиональность подобного рода не замечалась, поскольку диагональная лемма является формулой, подчиненной экстенсиональной логике, и фактически то, что предложил Крайзель, состоит в указании на то, что можно «играть» с этой формулой, выбирая те или иные интересные неподвижные точки. Кстати, в математической практике так и делается в значительном числе случаев исследования метаматематических результатов. Но это не общий результат, а своего рода промежуточное прозрение. Его важность состоит в том, что он пролил свет на статус самореферентных утверждений в связи с их интенсиональностью.

Холбах и Виссер указали, что самореферентные утверждения напрямую связаны с интенсиональностью, и более того, самореферентность фигурирует в одном из трех уровней интенсиональности. Они предполагают, что конструирование предложений с метатеоретическим содержанием предполагает некоторые базисные предположения: во-первых, кодирование выражений языка натуральными числами должно быть фиксировано; во-вторых, доказуемость и другие предикаты должны быть выражены определенными формулами объект-языка. Наконец, самореференция получается в формальной системе посредством некоторых диагональных выражений [Halbach, Visser, 2014].

Эти требования никоим образом не являются чем-то экстраординарным, но из них пока не видно, на каком этапе появляется интенсиональность. Первый шаг к решению этого вопроса, как уже указывалось выше, сделал Г. Крайзель. Второй шаг связан со Второй теоремой Гёделя о неполноте. Как известно, сам Гёдель не привел ее доказательства, заметив просто, что его доказательство Первой теоремы о неполноте показывает $\Sigma \vdash \mathrm{Con}\,(\Sigma) \rightarrow \neg\, \mathrm{Pr}\,(\ulcorner \varphi \urcorner)$. Содержание этой формулы передается известной формулировкой о том, что если формальная система Σ непротиворечива (Con), тогда есть недоказуемая формула φ. Отсюда следует $\Sigma \vdash \mathrm{Con}\,(\Sigma) \rightarrow \varphi$. Поскольку φ недоказуема, таковой является и $\mathrm{Con}(\Sigma)$. Формализация этого заключения, осуществленная П. Бернайсом, потребовала формулировки так называемых условий выводимости Гильберта – Бернайса, которые включают положения, имеющие эпистемологический интерес. В частности, третье условие в этом списке имеет вид $\Sigma \vdash \mathrm{Pr}\ulcorner \varphi \urcorner \rightarrow \mathrm{Pr}\,\ulcorner (\mathrm{Pr}(\ulcorner \varphi \urcorner)\urcorner)$, т. е. если формула доказуема, то доказуемо, что она доказуема. Такая итерация может считаться как интуитивной, так и несколько искусственной. Несмотря на возможность такой искусственности, и опять-таки искусственности конструкции гёделева предложения, все-таки Вторая теорема может относиться к математическим поискам неразрешимых предложений, и уже есть несколько кандидатов на эту роль, в частности, теорема Гудстейна. И здесь мы имеем глубокую дилемму, связанную с соотношением реальной математики и метаматематики. М. Лёб нашел решение проблемы Генкина, и это решение является чисто метаматематическим. В определенном отношении эта тенденция все большего разрыва реальной математики и метаматематики является крайне интересной. Дело в том, что парадоксы, имеющие прежде всего философский интерес, стали предметом чисто математических исследований, скажем, в аксиоматической теории множеств, с соответствующим угасанием собственно философского интереса к ним. Теорема Лёба возвращает интерес к парадоксам двояким образом: с одной стороны, она предлагает новый парадокс, напоминая об исходном интересе к ним, а с другой – усиливает раскол между реальной математикой и метаматематикой.

Сама теорема формулируется так:

$$\text{если } \Sigma \vdash \mathrm{Pr}\ulcorner\psi\urcorner \to \psi, \text{ тогда } \Sigma \vdash \psi.$$

Своей теоремой Лёб решил вопрос, поставленный Генкиным: если формальная система Σ доказывает «если Σ доказывает ψ, тогда ψ», тогда Σ доказывает ψ (так что предложение Генкина доказуемо в Σ).

Чисто метаматематический характер теоремы Лёба, в отличие от математического, виден в том, что эта теорема – довольно странный принцип для доказательства теорем о натуральных числах. Для того чтобы доказать утверждение А, допустимо предположить в качестве посылки, что φ доказуемо в PA (Арифметике Пеано). Потому что если есть доказательство в PA «если есть доказательство в PA утверждения φ, тогда φ», тогда есть доказательство φ в PA. Этот принцип не имеет известных приложений при доказательстве математических теорем о простых числах или других традиционных математических темах, но используется в метаматематике.

Этот принцип кажется больше помехой, чем просто странным. Как можно допускать в доказательстве φ, что φ доказуемо в PA? В конце концов, что доказуемо в PA, истинно, поэтому разве мы не можем заключить без дальнейшего добавления, что φ истинно, и, таким образом, доказать φ без размышлений вообще? Существенно здесь то, что допустимо предполагать, что φ доказуемо в PA при доказательстве φ, только если размышление, ведущее от предположения, что φ доказуемо в PA, к заключению φ, может быть выполнено на самом деле в рамках PA.

Эта странность подчеркивается К. Сморински в сопоставлении со Второй теоремой Гёделя:

> Там, где гёделевская Вторая теорема просто утверждает недоказуемость непротиворечивости теории в самой теории, теорема Лёба превосходит это, и характеризует эти примеры обоснованности, доказуемые в теории как тривиально доказуемые [Smorynski, 1991. P. 119].

Этот результат, по мнению Сморински, является в высочайшей степени именно метаматематическим. Его философская важность состоит, среди прочего, в том, что он возвращает нас к парадоксам, т. е. к философски мотивированному поиску интенсиональных основ математического мышления.

Исключительно метаматематический характер теоремы Лёба подчеркивается тем, что некоторые ее следствия получаются «игрой» с неподвижными точками, точнее, с их варьированием. Один из вариантов такого варьирования приводит к парадоксу. Пусть мы имеем предложение:

Если я говорю истину, тогда истинно и А.

Если это предложение обозначено как В, тогда мы имеем $B = B \rightarrow A$: В может быть ложным только в случае В истинно, а А ложно. Это означает, что В истинно. В этом случае, $B \rightarrow A$ истинно, и отсюда истинно и А, независимо от того, каково это утверждение на самом деле. В работе Р. Смаллиана этот парадокс обыгрывается как проблема самовыполнимых вер [2013]. В формальном отношении парадокс есть следствие принятия неподвижной точки $\varphi \longleftrightarrow Pr(\ulcorner\varphi\urcorner) \rightarrow \psi$ [Kreisel, 1953]. Возможны и другие неподвижные точки, выбор которых зависит от интересов исследователя. Различия в выборе результируется в различной эффективности доказательств определенных метаматематических результатов. Именно эта очередная вариативность может являться демонстрацией интенсиональности метаматематики. В данном случае, эта интенсиональность проявляется в том, что есть предложения, выражающие определенные стороны самих себя, и просто неподвижные точки. Другими словами, для того, чтобы неподвижные точки давали интересные, хорошо интерпретированные философски, результаты, требуется особая мотивация в выборе средств формализации. Так, Крайзель использует следующую неподвижную точку [Kreisel, 1953]:

$$\Sigma \vdash \varphi \longleftrightarrow Pr(\ulcorner\varphi \rightarrow \psi\urcorner).$$

Сопоставление неподвижных точек Крайзеля и Лёба показывает, что каждая из них имеет свои преимущества в эффективности доказательств разных интересных результатов.

Что демонстрирует парадокс, связанный с теоремой Лёба? В настоящее время в метаматематике широко используются модальные понятия, которые в рамках некоторых систем модальной логики делают доказательства более прозрачными. Вместо предиката доказательства Pr используется модальный оператор $\Box$, который можно интерпретировать по-разному, в зависимости от целей. В частности, это может быть предикатом истины или предикатом доказуемости. В этой нотации теорема Лёба говорит, что если система Σ выполняет условия выводимости Гильберта – Бернайса (необходимые для доказательства Второй теоремы Гёделя о неполноте), тогда если $\Sigma \vdash \Box\varphi \to \varphi$, то $\Sigma \vdash \varphi$.

В модальной нотации условия выводимости Гильберта – Бернайса выглядят так (на самом деле этот вид придан им самим Лёбом, и поэтому иногда они называются условиями выводимости Лёба):

L1. Если $\Sigma \vdash \varphi$, тогда$\Sigma \vdash \Box\varphi$;

L2. $\Sigma \vdash \Box(\varphi \to \psi) \to (\Box\varphi \to \Box\psi)$;

L3. $\Sigma \vdash \Box\varphi \to \Box\Box\varphi$.

Для понимания природы парадокса, ассоциируемого с теоремой Лёба, необходимо проследить «механику» доказательства этой теоремы в связи с условиями выводимости. Ниже представлено формальное доказательство: для простоты опустим непременное упоминание выводимости формулы в системе Σ [Smith, 2013. P. 256].

1.	$\Box\varphi \to \varphi$	Посылка
2.	$\gamma \longleftrightarrow (\Box\gamma \to \varphi)$	Применение к (1) диагональной леммы
3.	$\gamma \to (\Box\gamma \to \varphi)$	(2)
4.	$\Box (\gamma \to (\Box\gamma \to \varphi))$	(3), по L1
5.	$\Box\gamma \to \Box(\Box\gamma \to \varphi)$	(4), по L2
6.	$\Box\gamma \to (\Box\Box\gamma \to \Box\varphi)$	(5), по L2
7.	$\Box\gamma \to \Box\Box\gamma$	по L3
8.	$\Box\gamma \to \Box\varphi$	(6), (7)
9.	$\Box\gamma \to \varphi$	(1), (8)
10.	γ	(2), (9)
11.	$\Box\gamma$	(10), по L1
12.	φ	(9), (11)

Парадоксальность теоремы Лёба лучше всего видна при интерпретации $\Box$ как предиката истины. В этом случае оператор $\Box$ обозначает «истинно», и $\Box\varphi$ – утверждение φ истинно. Здесь парадоксальный момент состоит в том, что φ может означать любое утверждение, даже заведомо ложное. Если мы повторим выше приведенный вывод, только уже с новой интерпретацией оператора $\Box$, то самой подозрительной в отношении возникновения парадокса является как раз строка (2), а именно, $\gamma \longleftrightarrow (\Box\gamma \rightarrow\varphi)$. Если γ – истинное утверждение, тогда в результате всего вывода мы получаем истинность любого утверждения φ. В свою очередь, подозрительность относительно (2) распространятся на утверждение γ, которое мы вводим постулированием. Этот парадокс достаточно серьезен. П. Смит утверждает:

> Не ясно, каков самый лучший способ блокирования этого парадокса, в любом случае, не более ясно, чем в случае Парадокса Лжеца. Без сомнения, есть что-то сомнительное о постулировании утверждения γ, такого, что справедливо $\gamma \longleftrightarrow (\Box\gamma \rightarrow\varphi)$, но *что* в точности? [Smith, 2013. P. 257].

Другими «подозреваемыми» оказываются условия выводимости, в частности, интуитивная посылка, согласно которой если есть доказательство некоторого утверждения, то есть доказательство того, что оно доказано. Эпистемический аспект здесь не является техническим, а на самом деле это инкарнация весьма старых споров. Но как бы то ни было, размышления о допустимости постулирования γ наверняка должны быть отнесены к интенсиональности, если и даже неявной, метаматематических структур. Тот же Смит делает очень важное в этом отношении замечание:

> Теорема Лёба, подобно гёделевской теореме, является не семантическим парадоксом, а ограничительным результатом о неспособности доказывать определенные утверждения о своих собственных доказательных свойствах [Ibid.].

Эта параллель существенна в том отношении, что интенсиональность Второй теоремы Гёделя – общепризнанный факт, в то время как в отношении теоремы Лёба это не совсем очевидно. Здесь мы встречаемся с парадоксальностью второго порядка, потому что из теоремы Лёба можно получить Вторую теорему Гёделя о неполноте, а из теоремы Гёделя следует теорема Лёба. Даже если и есть определенные технические оговорки об установлении такой эквивалентности, в любом случае мы сталкиваемся с затруднительной ситуацией, разрешение которой лежит в области так называемой логики доказательств, например, системе GL.

Однако «странность» теоремы Лёба не ограничивается ее соотношением со Второй теоремой Гёделя о неполноте. Сугубо метаматематический характер теоремы Лёба ставит несколько вопросов, которые требуют разрешения при объяснении предполагаемого интенсионального характера этой теоремы. Г. Булос набросал программу такого рода исследований, из которой можно видеть, насколько велика ее значимость [Boolos, 1996. P. 54–55].

1. Часто трудно понять, как широка математическая пропасть между истиной и доказуемостью. И для тех, кто не понимает этого и не различает истину и доказуемость, выражение $\mathrm{Bew}(\ulcorner S \urcorner) \to S$, доказуемое утверждение по гипотезе теоремы Лёба, может показаться тривиально истинным во *всех* случаях, независимо от того, является ли S истинным или ложным, доказуемым или недоказуемым. Но если S ложно, S лучше не быть доказуемым. Таким образом, S не следует быть всегда доказуемым, при просто условии, что (кажется тривиальным) $\mathrm{Bew}(\ulcorner S \urcorner) \to S$ доказуемо.

2. Bew ведет себя как отрицание. В конце концов, если $\neg S \to S$ доказуемо, тогда доказуемо и S; доказательство S через доказательство $\neg S \to S$ называется доказательством от противного. Больше того, вывод S на основании лишь того, что $(S \to S)$ демонстрируемо, спорен, или же включает круг. Тем, кто сливает истину и доказуемость, может показаться тогда, что теорема Лёба утверждает, что спорные вещи являются допустимой формой размышления в PA.

3. Можно подумать, по крайней мере в одном случае, что PA претендует на обоснованность в отношении недоказуемого предложения S, т. е. что если PA доказывает S, тогда имеет место S. Но теорема Лёба говорит нам, что такого не бывает: PA делает утверждение $\mathrm{Bew}(\ulcorner S \urcorner) \to S$, что оно обосновано в отношении S, только когда консеквент S действительно доказуем. Как выразился Рохит Парих, «PA не могло бы быть более скромным в отношении своей собственной истинности».

4. Было бы естественно предположить, что доказуемость есть вид необходимости, и следовательно, точно так же, как $\Box(\Box p \to p)$ всегда выражает истину, если $\Box$ интерпретируется как «необходимо, что» – потому что тогда $\Box(\Box p \to p)$ говорит, что необходимо истинно, что если утверждение необходимо истинно, оно истинно.

$\mathrm{Bew}(\ulcorner \mathrm{Bew}(\ulcorner S \urcorner \to S) \urcorner$ будет также всегда истинным или, по крайней мере, истинным в некоторых случаях, в которых S ложно, и не истинно только в некоторых исключительных случаях, в которых S действительно доказуемо.

5. Наконец, кажется совсем странным, что утверждение, что если S доказуемо, тогда S истинно, само не доказуемо, в общем. Разве не ясно, для любого S, что S истинно, если доказуемо? Зачем возиться с PA, если ее теоремы ложны? И как такая (кажущаяся ясной) истина не доказуема?

В контексте данной статьи особый интерес представляют п. 3 и 5. Вместе с тем можно предположить, что разрешение остальных вопросов будет напрямую связано интенсиональностью теоремы Лёба.

Список литературы / References

Смаллиан Р. Вовеки неразрешимое. М.: Канон+, 2013.

Smullyan R. Voveki nerazreshimoe [Forever Undecided]. Moscow, Kanon+, 2013. (in Russ.)

Boolos G. The Logic of Provability. Cambridge, Cambridge Uni. Press, 1996, p. 54–55.

Halbach V., Visser A. Self-Reference in Arithmetic I. *Review of Symbolic Logic*, 2014, vol. 7, p. 671–691.

Henkin L. Problem. *Journal of Symbolic Logic*, 1952, vol. 17, p. 160.

Kreisel G. On a Problem of Henkin. In: Proc. Netherlands Acad. Sci., 1953, vol. 56, p. 405–406.

Löb M. H. Solution of a Problem by Leon Henkin. *Journal of Symbolic Logic*, 1955, vol. 20, p. 115–118.

Smith P. An Introduction to Gödel's Theorems. Cambridge, Cambridge Uni. Press, 2013.

Smorynski C. The Development of Self-Reference: Lob ‘s Theorem. In: Perspectives on the History of Mathematical Logic. Ed. by T. Drucker. Birkhäuser, 1991, p. 110–133.

Материал поступил в редколлегию
Received
01.07.2019

Сведения об авторе / Information about the Author

Целищев В. В., Институт философии и права СО РАН (ул. Николаева, 8, Новосибирск, 630090, Россия)

Vitaly V. Tselishchev, Institute of Philosophy and Law SB RAS (8 Nikolaev Str., Novosibirsk, 630090, Russian Federation)

leitval@gmail.com