Проблемы логики и методологии науки

УДК 164.07

О ДВУХ НЕВЕРНЫХ ДОГМАХ, СВЯЗАННЫХ СО ВТОРОЙ ТЕОРЕМОЙ ГЁДЕЛЯ О НЕПОЛНОТЕ АРИФМЕТИКИ. I*

А.В. Бессонов

Показано, что гёделево доказательство второй теоремы о неполноте формальной арифметики зависит от избранного им предиката доказуемости. С использованием предиката недоказуемости строятся контрпримеры ко второй теореме, из чего следует, что в общем случае вывод второй теоремы не является верным.

Ключевые слова: теоремы Гёделя о неполноте, неадекватность предиката доказуемости, предикат недоказуемости.

Введение

Завершая свою знаменитую статью 1931 г. [1], в которой доказываются ставшими классическими теоремы о неполноте формальной арифметики, К. Гёдель пишет: «Доказательство теоремы XI (вторая теорема о неполноте – *А.Б.*) слово в слово переносится на систему аксиом теории множеств, *M*, и систему аксиом классической математики, *A*, из чего следует: не существует доказательства непротиворечивости для *M* или *A*, которое может быть формализовано соответственно в *M* или в *A*, в предположении, что *M* или *A* непротиворечивы. Я хочу специально отметить, что Теорема XI (и соответствующие результаты для *M* и *A*) не противоречат гильбертовской формалистской точке зрения. Эта точка зрения

---

* Статья публикуется в авторской редакции.

предполагает лишь существование доказательства непротиворечивости, использующего исключительно финитные средства, и возможно, что найдутся финитные доказательства, которые не могут быть выражены в формализме *P* (или *M*, или *A*)» [2].

Следует сказать, что в [1] приведен лишь набросок доказательства второй теоремы о неполноте [3]. Полное доказательство Гёдель планировал опубликовать в следующем номере журнала *Monatshefte für Mathematik und Physik* (что и объясняет использование индекса «I» в названии его работы), однако в дальнейшем он отказался от этих планов [4] и нигде такое доказательство не публиковал. Как свидетельствует Г. Крайзель, Гёделю в 1930–1931 гг. казалось, что в то время не прошло бы и нескольких месяцев, как кто-то другой наткнулся бы на теоремы о полноте и неполноте – его самые знаменитые результаты [5]. Таким образом, сделанные Гёделем выводы из доказательства второй теоремы формулировались им в обстановке острой конкурентной спешки, что вряд ли может свидетельствовать об их глубокой продуманности и обоснованности. Тем не менее, именно эти выводы Гёделя послужили основанием для внедрения в общественное сознание логиков и математиков двух тезисов, которые воспринимаются в самых широких научных кругах как безусловные и бесспорные, иными словами, как догмы. Эти догмы таковы:

I. Арифметика, если она непротиворечива, не может доказать свою непротиворечивость.

II. Вторая теорема Гёделя служит решающим аргументом в доказательстве несостоятельности гильбертовской программы финитного обоснования математики.

Второй тезис получил догматический статус даже несмотря на специальную оговорку Гёделя, поскольку он был воспринят научным сообществом как прямое следствие первой догмы.

В настоящей работе мы покажем, что ни первая, ни вторая догма не являются верными. Работа состоит из двух частей, и в этой части мы сосредоточим свое внимание на первой догме.

1. Первая догма: краткая мифология

Что же представляет собой вторая теорема? Начнем с первоисточника: «Теорема XI. Пусть κ – рекурсивный непротиворечивый класс ФОРМУЛ; тогда ЗАМКНУТАЯ ФОРМУЛА, устанавливающая, что κ

непротиворечив, не является *к*-ДОКАЗУЕМОЙ; в частности, непротиворечивость *P* не доказуема в *P*, в предположении, что *P* непротиворечива (в противном случае, конечно, всякое высказывание доказуемо [в *P*])» [6]. Согласно авторской переформулировке теорем о неполноте от 1963 г. «…во всякой непротиворечивой формальной системе, содержащей определенную порцию финитной теории чисел, имеются неразрешимые арифметические высказывания, к тому же непротиворечивость любой такой системы не может быть доказана в самой системе» [7].

А вот как эта догма подается в культовых книгах по логике. Д. Гильберт и П. Бернайс, первыми давшие полное доказательство второй теоремы [8], проповедуют: «…в случае непротиворечивости формализма *F* не может существовать формализованного в *F* доказательства этой непротиворечивости, т.е. вывода в *F* уже упоминавшейся нами формулы…» [9]. А.А. Френкель и И. Бар-Хилел заклинают: «Никакое предложение, которое можно точным образом интерпретировать как выражающее непротиворечивость какой-либо логистической системы, содержащей арифметику, не может быть доказано в этой системе» [10]. С.К. Клини учит: «Если арифметическая формальная система гл. IV (просто) непротиворечива, то не $\vdash$ Consis; иначе говоря, если указанная система непротиворечива, то не существует доказательства её непротиворечивости, проведенного средствами, формализуемыми в этой системе» [11].

Не отстают и современные адепты первой догмы: «Согласно этой теореме, если формальная система, содержащая арифметику, непротиворечива, то утверждение о её непротиворечивости выразимо в этой системе, но не может быть доказано средствами, формализуемыми в ней» [12]. «Вторая теорема о неполноте устанавливает недоказуемость в теории чисел непротиворечивости теории чисел» [13]. «Вторая теорема Гёделя о неполноте показывает, что никакая непротиворечивая формальная система не может доказать свою собственную непротиворечивость» [14]. «В соответствии со второй теоремой о неполноте, такая формальная система не может доказать, что она сама непротиворечива (в предположении, что она действительно непротиворечива) [15]. «Вторая теорема Гёделя о неполноте говорит приблизительно то, что никакая достаточно богатая рекурсивно аксиоматизируемая математическая теория в классической логике не может доказать свою собственную непротиворечивость» [16].

И подобные заклинания можно множить и множить. Возможно ли остановить это слитное многоголосое пение? Попытка дезавуирования любой догмы, в том числе и научной, представляет собой неподъёмную задачу. Вполне осознавая это, мы все-таки попробуем доказать, что арифметика может доказать свою непротиворечивость. Сразу скажем, что мы не намерены искать ошибки в имеющихся доказательствах теорем о неполноте. Мы пойдем другим путем. Путем, параллельным гёделевому, но который приведёт нас к прямо противоположным выводам. Для этого придётся углубиться в некоторые технические детали гёделевой аргументации.

2. Первая догма, «выразимость» и предикаты доказуемости

В [1] Гёдель рассматривает формальную систему P, аналогичную самой популярной в то время системе аксиом для формализации математики – *Principia Mathematica* Рассела–Уайтхеда. Он указывает, что его результаты относятся и к другим аксиоматическим системам, наиболее интересный пример которых представляет аксиоматика Дедекинда–Пеано, формализованная в языке логики первого порядка. Обозначим эту систему через PA и далее сосредоточим свое внимание именно на ней. Этот выбор соответствует стандартной точке зрения на теоремы Гёделя и интересен для нас тем, что он тесно связан со второй догмой.

Язык PA надстраивается над языком исчисления предикатов первого порядка и содержит дополнительно один константный символ $\mathbf{0}$, один одноместный функциональный символ S, два двухместных функциональных символа $+$ и $\cdot$, а также единственный двухместный предикатный символ $=$. В стандартной интерпретации переменные принимают значения в множестве натуральных чисел, символ $\mathbf{0}$ интерпретируется как ноль, S как операция прибавления единицы («следующий за»), а $+$ и $\cdot$ как сложение и умножение соответственно.

Натуральным числам соответствуют нумералы, построенные путем приписывания к $\mathbf{0}$ слева символа S необходимое число раз, например, числу 3 соответствует нумерал $SSS0$. Нумерал, соответствующий числу n, принято обозначать через $\boldsymbol{n}$.

Аксиоматика состоит из 7 аксиом

$$\forall x(\neg(\mathbf{0} = Sx)),$$

$$\forall x \forall y(Sx = Sy \to x = y),$$

$$\forall x(\neg(x = \mathbf{0}) \to y(x = Sy)),$$

$$\forall x(x + \mathbf{0} = x),$$
$$\forall x \forall y(x + Sy = S(x + y)),$$
$$\forall x(x \cdot \mathbf{0} = \mathbf{0}),$$
$$\forall x \forall y(x \cdot (Sy) = (x \cdot y) + x)$$

и схемы полной индукции

$$(\varphi(\mathbf{0}) \,\&\, \forall x(\varphi(x) \to \varphi(Sx))) \to \forall x \varphi(x),$$

где $\varphi(x)$ – произвольная формула со свободной переменной x.

Как легко убедиться, в PA нет никаких выразительных средств, с помощью которых можно было бы непосредственно формализовать метаязыковые рассуждения о доказуемости, недоказуемости и неразрешимости. Формулировки же теорем о неполноте используют именно эти понятия. Иначе говоря, в этих теоремах речь идет не о свойствах чисел, а о свойствах теории чисел, PA. Как же можно заставить формулы языка арифметики «говорить» о самих себе, о своей (не)доказуемости?

Гёдель использует две основных идеи:

(1) Арифметизация синтаксиса, т.е. кодирование языка формальной арифметики и её логики (гёделева нумерация). При кодировании языковым выражениям PA ставятся в соответствие числа (гёделевы номера) так, чтобы разным выражениям сопоставлялись различные числа, и по выбранному числу можно было бы эффективно восстановить выражение, номером которого данное число является.

(2) Введение понятия «выразимости» («определимости»).

Определение. Предикат $F(x_1, \ldots, x_k)$, заданный на множестве натуральных чисел, называется «выразимым» в PA, если в PA найдется формула $\Phi(x_1, \ldots, x_k)$, такая что для любого набора натуральных чисел $(n_1, \ldots, n_k)$ справедливы условия

(1) если $F(n_1, \ldots, n_k)$ выполняется, то $\vdash \Phi(\boldsymbol{n}_1, \ldots, \boldsymbol{n}_k)$;

(2) если $F(n_1, \ldots, n_k)$ не выполняется, то $\vdash \neg\Phi(\boldsymbol{n}_1, \ldots, \boldsymbol{n}_k)$.

Здесь, как обычно, $\vdash$ означает доказуемость (в PA), а выражение $\ulcorner A \urcorner$ будет обозначать нумерал, соответствующий гёделеву номеру $\lceil A \rceil$ формулы A.

Напомним, что теория называется непротиворечивой, если в ней имеется хотя бы одна недоказуемая формула, или, что эквивалентно, в ней ни для какой формулы A не могут одновременно быть доказаны формулы A и $\neg A$. Теория называется ω-непротиворечивой, если в ней ни для какой формулы $\Phi(x)$ с одной свободной переменной не могут одновременно быть доказуемыми формулы

$$\Phi(1), \Phi(2), \ldots, \neg\forall x\Phi(x).$$

В первой теореме Гёделя о неполноте арифметики утверждается, что если PA ω-непротиворечива, то она неполна. Более точно, в ней доказывается существование некоторой замкнутой формулы G («говорящей» о своей собственной недоказуемости) такой, что ни она, ни её отрицание не доказуемы в PA (такие формулы называются *неразрешимыми*). В соответствии со второй теоремой, если PA непротиворечива, то в ней не доказуема формула, «выражающая» непротиворечивость PA [17]. Как же строится эта формула?

Гёдель вводит предикат доказуемости $\mathrm{Pr}(x, y)$, который выполняется тогда и только тогда, когда x является гёделевым номером некоторой формулы, а y – гёделевым номером её доказательства. Известно, что этот предикат эффективно разрешим. Гёдель установил, что всякий заданный на натуральных числах предикат «выразим» в PA тогда и только тогда, когда он разрешим. Значит, $\mathrm{Pr}(x, y)$ «выразим» в PA с помощью некоторой арифметической формулы $\mathrm{Prov}(x, y)$, т.е. для любых натуральных чисел n, k верны условия:

(1) если $\mathrm{Pr}(n, k)$ выполняется, то $\vdash \mathrm{Prov}(\boldsymbol{n}, \boldsymbol{k})$;
(2) если $\mathrm{Pr}(n, k)$ не выполняется, то $\vdash \neg\mathrm{Prov}(\boldsymbol{n}, \boldsymbol{k})$.

В этих обозначениях определение непротиворечивости выражается арифметической формулой

$$\exists x \forall y \neg \mathrm{Prov}(x, y)\,, \tag{*}$$

которую С. Клини назвал «Consis». (Consis «выражает» факт существования в PA недоказуемой формулы.) Во второй теореме доказывается, ровно следующее: если PA непротиворечива, то формула (*) не может быть доказана в PA. Отсюда Гёдель немедленно делает вывод, что «непротиворечивость P (PA в нашем случае. – *А.Б.*) недоказуема в P в предположении, что P непротиворечива» [18].

Следует указать на то, что широко распространённое буквальное прочтение этих слов Гёделя является ошибочным. Недопустимо смешивать формулу, по-гёделевски «выражающую» непротиворечивость PA, с «выражаемым» ею фактом, а доказательство этой формулы в PA приравнивать к доказательству непротиворечивости самой PA. Гёделю и не нужно доказывать непротиворечивость PA, поскольку он её предполагает! Следует ясно понимать, что во второй теореме речь идет не более чем о недоказуемости некоторой формулы, определенным способом «выражающей» непротиворечивость PA. И даже если бы подобная формула была доказуемой, это никак не гарантировало бы непротиворечивость PA, поскольку в противоречивой теории доказуемы все формулы, в том числе и формула, «выражающая» непротиворечивость теории.

Таким образом, корректное прочтение гёделевской формулировки второй теоремы должно быть таким: «формула “выражающая” непротиворечивость *P*, недоказуема в *P* в предположении, что *P* непротиворечива». Но и уточненная таким образом формулировка может вызывать разночтения: имеется ли в виду любая формула, «выражающая» непротиворечивость или речь идёт о какой-то конкретной формуле. Во второй теореме вывод о недоказуемости формулы, «выражающей» непротиворечивость PA, доказывается исключительно по отношению к гёделевой формуле Consis. Но Гёдель сам совершает ошибку и необоснованно распространяет вывод на любые формулы, «выражающие» непротиворечивость PA. Приведем выдержку из его формулировки так, как она переведена на английский язык: «…**the** SENTENTIAL FORMULA stating that κ is consistent is not κ-PROVABLE; in particular, the **consistency of** *P* **is not provable in** *P*, provided *P* is consistent» (полужирный мой. – *А.Б.*) [19]. Как видно, в первой части предложения речь идет о некой определённой формуле, а заключительная фраза относится уже к любой подобной формуле, поскольку утверждение о недоказуемости непротиворечивости *P* может означать только одно: **любая** формула, «выражающая» непротиворечивость *P*, недоказуема в *P*. Но вполне резонно предположить, что могут существовать отличные от Consis формулы, «выражающие» непротиворечивость PA. Из доказательства же второй теоремы не следует, что в PA не может быть доказана никакая другая (не эквивалентная гёделевой Consis) формула, «выражающая» непротиворечивость PA. Из второй теоремы никак не следует несуществование «*других, чем в теореме Гёделя*, быть может не столь естественных (или естественных как-то по-другому), но все же разрешимых в *P* арифметических выражений непротиворечивости... Странным образом последнее обстоятельство не было замечено в 1931 году» [20].

Эта ошибка Гёделя плавно перекочевала во многие культовые тексты. Приведем пример из классического учебника Э. Мендельсона:

«Если теория *S* непротиворечива, то в ней невыводима и формула Con_S; иными словами, если теория непротиворечива, то в ней невыводима некоторая(! – *А.Б.*) формула, содержательно выражающая непротиворечивость теории *S*. Этот результат носит название второй теоремы Гёделя. Грубо говоря, эта теорема утверждает, что если теория *S* непротиворечива, то доказательство непротиворечивости теории не может быть проведено средствами самой теории *S*, т.е. всякое(!? – *А.Б.*) такое доказательство обязательно должно использовать невыразимые в теории *S* идеи или методы» [21].

От внимания авторов исследований, посвящённых теоремам Гёделя о неполноте, зачастую ускользает различие в природе первой и второй теорем. На самом деле эти две теоремы относятся к разным задачам. Первая решает задачу построения неразрешимой в PA формулы-монстра, и тут чем хуже используемые «выразительные» возможности, тем лучше. Вторая же сформулирована в рамках задачи построения формулы, «выражающей» непротиворечивость PA, и доказательстве недоказуемости в PA любой такой формулы. Гёдель строит Consis исходя из своей неразрешимой формулы G. Вторая теорема по сути состоит в доказательстве того, что в PA доказуема импликация

$$\mathrm{Consis} \to \mathrm{G}.$$

Из этого следует, что если бы формула Consis была доказуемой, то доказуемой была бы и формула G, что по первой теореме означало бы противоречивость PA. Однако, вообще говоря, формула, «выражающая» непротиворечивость PA, не обязательно должна быть эквивалентна гёделевой формуле Consis. Поэтому при оценке степени универсальности второй теоремы следует более внимательно отнестись к выразительным возможностям различных формул.

Но имеются ли вообще в PA формулы, которые можно рассматривать в качестве контрпримера к универсальному истолкованию второй теоремы? Существуют ли выражающие непротиворечивость PA формулы, «быть может не столь естественные (или естественные как-то по-другому)», и все-таки доказуемые в PA? Ирония заключается в том, что в PA тривиально доказуемы формулы, вполне естественно выражающие (без кавычек!) непротиворечивость PA, построенные без арифметизации синтаксиса и без введения понятия «выразимости» и предиката доказуемости. Это, например, все формулы, полученные подстановкой арифметических формул в доказуемую в классическом пропозициональном исчислении (КИВ) схему

$$\neg(A\&\neg A).$$

Поскольку PA содержит КИВ, все эти подстановочные примеры элементарно доказуемы в PA. А так как КИВ полно, т.е. в нем истинность и доказуемость совпадают, то вполне естественно истолковать доказуемую арифметическую формулу вида $\neg(A\&\neg A)$ как выражающую недоказуемость формулы $A\&\neg A$, т.е. выражающую непротиворечивость PA! Выражающую (без кавычек!) настолько же естественно, насколько & выражает конъюнкцию, а $\to$ импликацию.

Для тех, кто не пожелает вырваться за рамки гёделева аппарата, кто понимает вторую теорему исключительно в терминах «выразимости» (в каковых она и доказана Гёделем), рассмотрим свойства этого понятия более подробно. Напомним определение. Предикат $F(x_1, \ldots, x_k)$, заданный на множестве натуральных чисел, «выразим» в PA, если в PA найдется формула $\Phi(x_1, \ldots, x_k)$, такая что для любого набора натуральных чисел $(n_1, \ldots, n_k)$ справедливы условия

(1) если $F(n_1, \ldots, n_k)$ выполняется, то $\vdash \Phi(\boldsymbol{n}_1, \ldots, \boldsymbol{n}_k)$;
(2) если $F(n_1, \ldots, n_k)$ не выполняется, то $\vdash\neg\Phi(\boldsymbol{n}_1, \ldots, \boldsymbol{n}_k)$.

Отметим, что условия (1), (2) дают исчерпывающий список того, что Гёдель требует от «выразимости».

Рассмотрим, какие предикаты могут стоять слева в этом определении, и что могут «выражать» стоящие справа обычные формулы арифметики. Для предикатов достаточно быть определёнными на натуральных числах. При этом, что чрезвычайно важно при арифметизации синтаксиса, числам не запрещается иметь дополнительное, неарифметическое значение. И никаких пределов этой дополнительности не установлено. Так, при определении предиката доказуемости используются фразы «гёделев номер предложения», «гёделев номер вывода» и т.п. Но нумерирование чего-либо не является *свойством* чисел, оно относится к *свойствам использования* чисел. Тот факт, что Гёдель присвоил левой скобке номер 11, не относится к свойствам числа 11, а относится к тому, как это число Гёдель использовал при арифметизации синтаксиса. Использовать же числа можно самыми различными способами, например, для пересчёта яблок, лежащих на столе в данный момент. И предикат «число яблок, лежащих на столе» (обозначим его Ябл(x)) «выразим» в смысле Гёделя арифметической формулой. Действительно, пусть на столе лежат три яблока. Тогда Ябл(x) выполняется при x =

3 и не выполняется при всех других значениях x. Легко понять, что формула $x = \mathbf{3}$ «выражает» предикат Ябл(x) в PA: в PA доказуемы формулы (1) $\mathbf{3 = 3}$ и (2) $\neg(k = \mathbf{3})$ при любом $k \neq \mathbf{3}$ [22], т.е. выполняются все условия «выразимости».

Да, в PA нет никаких яблок, но в этой системе нет также ни гёделевой, ни какой-либо иной нумерации: из каких аксиом PA с логической необходимостью следует существование нумераций? Перефразируя известное выражение У. Куайна, можно сказать, что формулы PA не несут на своих лбах собственные гёделевы номера. Нумерирование (арифметизация) синтаксиса и логики PA является таким же фактическим (произвольным, случайным, внешним, посторонним и т. п.) обстоятельством для PA как формальной системы, что и факт нахождения в данный момент трёх яблок на столе. Последний факт связан с аксиомами PA в столь же малой степени, что и, например, факт использования Гёделем для нумерации левой скобки – числа 11, правой – числа 13, а не наоборот. При этом арифметизация синтаксиса приводит к смешению упоминания и употребления выражений PA, когда выбрав нумерал, являющийся гёделевым номером какой-либо формулы, мы не знаем, что он обозначает: число или формулу. Подобное смешение, на недопустимость которого указывал У. Куайн [23], само по себе является потенциальным источником парадоксов. И те немногие, кто действительно сомневаются в непротиворечивости PA, должны с ещё большей озабоченностью подойти к системе PA + нумерация, т.к. непротиворечивость последней прямо не следует из предположения о непротиворечивости PA. В системе PA + нумерация предложения могут «говорить» сами о себе, т.е. такая система семантически замкнута что, в соответствие с А. Тарским [24], также чревато парадоксами. Всё это заставляет относиться к результатам, полученным в системе PA + нумерация (включая первую теорему о неполноте), с некоторой долей сомнения.

Одна и та же формула арифметики может ничего не «выражать», «выражать» какой-либо предикат в одной нумерации и предикат, отличный от него – в другой. Так, формула арифметики $x = 11$ ничего не «выражает» в случае, когда 11 вообще не является гёделевым номером, «выражает» предикат «быть номером левой скобки» в нумерации самого Гёделя и предикат «быть номером правой скобки» в нумерации, полученной из нумерации Гёделя путем перестановки номеров левой и правой скобок. Из этого примера понятно, что чисто арифметический смысл формулы, «выражающей» какой-либо предикат, может не иметь ничего общего со смыслом самого предиката. Но для нас наибольший интерес представляет то, что непротиворечивость PA могут «выражать» не эквивалентные формулы. Так, Клини приводит следующий пример (с точностью до обозначений). Возьмем определённую выше формулу Prov(x, y) «выражающую» то, что x является геделевым номером формулы, а y – номером её доказательства, и по ней построим формулу Prov$'$(x, y):

$$\mathrm{Prov}(x, y) \,\&\neg\, \mathrm{Prov}(\ulcorner \mathbf{0} = \mathbf{1} \urcorner, y)\,.$$

Поскольку непротиворечивость PA предполагается, второй член конъюнкции становится доказуемой формулой при подстановке любого нумерала вместо y. Если теперь Prov(x, y) выполняется, то $\vdash$ Prov(x, y) по условиям «выразимости», и следовательно $\vdash$ Prov$'$(x, y). Если же Pr(x, y) не выполняется, то $\vdash\neg$Prov(x, y) т.е. доказуем один из дизъюнктов $\neg$ Prov$'$(x, y) значит, и сама эта формула доказуема. Таким образом, Prov$'$(x, y) также «выражает» предикат Pr(x, y). Но в PA доказуемо

$$\forall y \neg((\mathrm{Prov}(\ulcorner \mathbf{0} = \mathbf{1} \urcorner, y) \,\&\neg\, \mathrm{Prov}(\ulcorner \mathbf{0} = \mathbf{1} \urcorner, y))\,,$$

что эквивалентно

$$\exists y \neg((\mathrm{Prov}(\ulcorner \mathbf{0} = \mathbf{1} \urcorner, y) \,\&\neg\, \mathrm{Prov}(\ulcorner \mathbf{0} = \mathbf{1} \urcorner, y))\,,$$

последняя же формула «выражает» недоказуемость в PA формулы $\mathbf{0} = \mathbf{1}$, что и означает непротиворечивость PA [25].

Приведённый пример показывает, что в оригинальной формулировке самого Гёделя вторая теорема о неполноте не верна: арифметика может «доказать свою непротиворечивость» хотя бы посредством формулы Prov$'$(x, y). Поэтому Гильберт и Бернайс [26] доказывают вторую теорему уже налагая на предикат доказуемости дополнительные ограничения. Пусть Proof(x) обозначает формулу ∃yProv(x, y), которая «выражает» факт доказуемости в PA формулы с номером x. Условия на доказуемость, известные под названием «Условия Гильберта–Бернайса», в стандартной современной формулировке (предложенной Лёбом в [27]), выглядят так:

(D1) если $\vdash A$, то $\vdash \mathrm{Proof}(\ulcorner A \urcorner)$.

(D2) $\vdash \mathrm{Proof}(\ulcorner A \urcorner) \to \mathrm{Proof}(\ulcorner \mathrm{Proof}(\ulcorner A \urcorner) \urcorner)$.

(D3) $\vdash \mathrm{Proof}(\ulcorner A \urcorner) \;\&\; \mathrm{Proof}(\ulcorner A \to B \urcorner) \to \mathrm{Proof}(\ulcorner B \urcorner)$.

(D1) является переформулировкой условий «выразимости» предиката доказуемости;

(D2) выражает требование, чтобы (D1) было доказуемо в PA;

(D3) выражает требование замкнутости предиката доказуемости относительно правила вывода Modus Ponens.

Гильберт и Бернайс доказали, что формула, «выражающая» непротиворечивость PA посредством предиката доказуемости, удовлетворяющего условиям (D1)–(D3), действительно недоказуема в PA, если PA непротиворечива (в [28] показано, что вторую теорему можно доказать даже без требования (D3)). Условия (D1)–(D3) считаются «естественными» условиями для предиката доказуемости, поскольку они являются простым переписыванием на языке PA неформального определения доказуемости в PA. Дж. Булос и Р. Джеффри [29] прямо определяют предикат доказуемости как формулу $B(x)$, удовлетворяющую условиям (D1)–(D3). Введённая выше формула Proof (x) удовлетворяет этим условиям, однако, как указывают авторы, «... вообще говоря, предикаты доказуемости могут не иметь прямого отношения к доказуемости...» [30]. И даже относительно такого, более общего по сравнению с гёделевым определения предиката доказуемости, вторая теорема оказывается верной в следующей формулировке. Если $B(x)$ – предикат доказуемости для PA (т.е. $B(x)$ удовлетворяет условиям (D1)–(D3) – *А.Б.*), а PA непротиворечива, то в PA не доказуемо $\neg B(\ulcorner \mathbf{0} = \mathbf{1} \urcorner)$ [31].

Так что же, адепты первой догмы могут торжествовать? Ведь арифметика, если она непротиворечива, в самом деле не может доказать свою непротиворечивость, если последняя «выражается» посредством «естественного» предиката доказуемости! Далее мы покажем, что из этих на первый взгляд естественных требований вытекают совершенно неестественные следствия.

3. Первая догма и предикаты недоказуемости

Адепты первой догмы, как и сам Гёдель, видят «выражение» непротиворечивости PA исключительно посредством предиката доказуемости. Спор между ними может идти лишь относительно «естественности» свойств данного предиката. Но так ли хорош этот способ «выражения» и нельзя ли построить формулы, «выражающие» непротиворечивость PA, используя принципиально иные выразительные средства? Для ответа на поставленные вопросы рассмотрим свойства гёделевой формулы Prov(x, y) более внимательно.

Очевидным следствием второй теоремы, является то, что для любой формулы A в PA не может быть доказано

$$\forall y \neg \mathrm{Prov}(\ulcorner A \urcorner, y). \qquad (*_A)$$

Действительно, из $\vdash \forall y \neg \mathrm{Prov}(\ulcorner A \urcorner, y)$ по закону экзистенциального обобщения

$$F(t) \to \exists x F(x)$$

немедленно следовало бы $\vdash \exists x \forall y \neg \mathrm{Prov}(x, y)$, т.е. $\vdash$ Consis. Но это означает, что в PA невозможно доказать недоказуемость никакой – ни доказуемой, ни недоказуемой формулы. Мы можем усилить этот обескураживающий вывод следующим также почти очевидным утверждением.

ТЕОРЕМА 2+[1] (1) *Если* PA *непротиворечива, то для любой формулы A формула, «выражающая» недоказуемость A, недоказуема в* PA.

(2) *Если* PA ω*-непротиворечива, то для любой недоказуемой в* PA *формулы A формула, «выражающая» недоказуемость A, неразрешима в* PA.

ДОКАЗАТЕЛЬСТВО. (1) Доказано выше.

(2) Пусть A – некоторая недоказуемая в PA формула с гёделевым номером $\ulcorner A \urcorner$. Её недоказуемость «выражается» формулой ($*_A$), которая, как следует из второй теоремы, недоказуема в PA. Рассмотрим отрицание формулы ($*_A$), т. е. формулу

$$\neg \forall y \neg \mathrm{Prov}(\ulcorner A \urcorner, y).$$

Предположим, что эта формула доказуема. Также предположим (как это и делается Гёделем при доказательстве первой теоремы), что PA ω-непротиворечива. Тогда хотя бы одна из формул в перечне

$$\neg \mathrm{Prov}(\ulcorner A \urcorner, \mathbf{1}), \neg \mathrm{Prov}(\ulcorner A \urcorner, \mathbf{2}), \ldots, \neg \forall y \neg \mathrm{Prov}(\ulcorner A \urcorner, y)$$

должна быть недоказуемой. По гипотезе последняя формула в перечне доказуема. Значит, найдётся номер k такой, что формула $\neg$Prov($\ulcorner A \urcorner$, $\boldsymbol{k}$) будет недоказуемой. В таком случае по условиям «выразимости»

[1] В [32] мы назвали это утверждение «Третьей теоремой о неполноте». Название изменено, т.к. термин «третья теорема», хотя и крайне редко, уже использовался другими авторами.

$\neg\mathrm{Pr}(\ulcorner A \urcorner, k)$ не может быть истинным, следовательно, это ложно, откуда вытекает, что $\mathrm{Pr}(\ulcorner A \urcorner, k)$ истинно. По определению предиката $\mathrm{Pr}(x, y)$ последнее означает, что формула A доказуема. Приходим к противоречию, поскольку мы предположили, что A – недоказуемая формула. Доказательство закончено.

Это просто доказываемое утверждение безусловно заслуживает называться теоремой, т.к. по степени общности оно очевидно превосходит первую теорему о неполноте. Однако за крайне редким исключением в публикациях по теоремам Гёделя вы не найдете формулировки п. (1) теоремы 2+. Утверждение же п. (2) вообще не содержится ни в одной из опубликованных работ! И это при том, что в многочисленных публикациях, начиная со знаменитой статьи самого Гёделя, находится место для доказательства предельно тривиального утверждения, согласно которому из ω-непротиворечивости следует обычная непротиворечивость.

Чем же обусловлен этот обет молчания? Не опасением ли бросить тень на первую догму? Не желанием ли несмотря ни на что сохранить священную веру в неё? Ведь по теореме 2+ в PA имеется бесконечное число неразрешимых замкнутых формул, причем неразрешимыми оказываются формулы, «выражающие» самые банальные мета-арифметические утверждения. Например, неразрешима формула

$$\forall y \neg \mathrm{Prov}(\ulcorner \neg(\mathbf{0} = \mathbf{0}) \urcorner, y),$$

«выражающая» недоказуемость формулы $\neg(\mathbf{0} = \mathbf{0})$, и это при (декларируемом Гёделем и его последователями) предположении о непротиворечивости PA! Таким образом, использование «естественного» предиката доказуемости, при котором только и верна вторая теорема о неполноте, неминуемо приводит к п. (2) теоремы 2+, т.е. к выводу о неразрешимости любой формулы, «выражающей» недоказуемость какой-либо недоказуемой формулы PA. Проще говоря, о (не)доказуемости недоказуемых в PA формул вообще ничего вразумительного доказать нельзя, если для «выражения» доказуемости используется исключительно «естественный» предикат доказуемости. В этом-то и состоит великая тайна теорем о неполноте.

Ситуация напоминает известный со времен Парменида парадокс сингулярного существования, когда о несуществующем без противоречия нельзя высказать никакое истинное утверждение, в том числе утверждение о его несуществовании.

Утверждение п. (2) резко контрастирует с тем, что на уровне содержательных, «самых обычных в математике» рассуждений недоказуемость $\neg(\mathbf{0} = \mathbf{0})$ в (принимаемом Гёделем) предположении о непротиворечивости PA доказывается совершенно элементарно методом от противного. Из предположения о доказуемости $\neg(\mathbf{0} = \mathbf{0})$, учитывая, что в PA $\vdash (\mathbf{0} = \mathbf{0})$, следовало бы $\vdash (\mathbf{0} = \mathbf{0})\ \&\ \neg(\mathbf{0} = \mathbf{0})$, т.е. PA была бы противоречивой. Мы можем сделать вывод, что использование «естественного» предиката доказуемости, более или менее пригодного для «выражения» обычных рассуждений о доказуемости, совершенно не подходит для «выражения» рассуждений о недоказуемости. А ведь при доказательствах непротиворечивости или неразрешимости необходима адекватность представления рассуждений именно о недоказуемости: требуется показать недоказуемость одновременно формулы и её отрицания или недоказуемость как формулы, так и её отрицания.

Чрезвычайно важно понимать, что выбор предиката именно доказуемости и соответствующей ему формулы в качестве исключительных средств для представления рассуждений о (не)доказуемости в PA, несмотря на кажущуюся его естественность, Гёделем и его последователями *никак не обосновывается*. Эти средства выбраны абсолютно произвольно. Вот красноречивое свидетельство: «Мы можем и будем ставить эти и другие вопросы, формулировка которых использует понятия, подобные \`утверждающее свою собственную доказуемость’ или \`выражающее недоказуемость $\mathbf{0} = \mathbf{1}$’, как вопросы о различных предложениях, конструируемых некоторым специальным образом именно из $\mathrm{Prov}(y)$, а не из какой-либо другой формулы ...» [33]. «Можем и будем» – вот и все обоснование. При этом вопрос о «выражении» рассуждений о недоказуемости посредством иных, не основанных на предикате доказуемости выразительных средств, никем даже не поднимается. Мы же далее рассмотрим именно этот вопрос.

Для построения формулы, «выражающей» факт непротиворечивости PA, достаточно указать всего на одну недоказуемую в PA формулу, каковая конечно же имеется в предположении, что PA непротиворечива. Естественно попытаться «выразить» этот факт путем введения различных предикатов *недоказуемости.* Рассмотрим предикат $\mathrm{BessNPr}(x)$, который выполняется тогда и только тогда, когда x является гёделевым номером первой пришедшей в голову автору этой статьи в момент написания этой фразы недоказуемой в PA формулы. Заверяю, что таковой оказалась формула $\neg(\mathbf{0} = \mathbf{0})$, очевидно недоказуемая в PA, если последняя непротиворечива. Пусть гёделев номер формулы $\neg(0 = 0)$ равен k и $\boldsymbol{k}$ – соответствующий k нумерал. Покажем, что предикат $\mathrm{BessNPr}(x)$

«выразим» в PA определённой формулой BessNProv(x). В качестве такой можно взять формулу $x = \boldsymbol{k}$. В самом деле, предикат BessNPr(x) выполняется только для значения x, совпадающего с k, и в этом случае в PA $\vdash \boldsymbol{k} = \boldsymbol{k}$, т.е. $\vdash$ BessNProv($\boldsymbol{k}$). При любом другом значении x, скажем, l, не равном k, BessNPr(l) не выполняется, и в этом случае, очевидно, $\vdash\neg$ BessNPr($\boldsymbol{l}$), так как для любого $l \neq k$ в PA доказуемо $\neg(\boldsymbol{l} = \boldsymbol{k})$ (см. §2). Поскольку в PA $\vdash$ BessNProv($\boldsymbol{k}$), по закону экзистенциального обобщения получаем

$$\vdash \exists x \mathrm{BessNProv}(x) \cdot$$

Иными словами, в PA доказуема формула «выражающая» существование в PA первой пришедшей мне в голову недоказуемой в PA формулы. Но из этого следует существование в PA недоказуемой формулы, т.е. следует непротиворечивость PA! Таким образом, PA вполне может «доказать свою непротиворечивость» ровно в том смысле, в каком, как полагается, она не может этого сделать, если вторая теорема о неполноте верна. Можно привести пример и более общего предиката недоказуемости.

Если PA непротиворечива и в ней доказуема формула A, то формула $\neg A$ очевидно недоказуема по определению непротиворечивости. Рассмотрим предикат NPr(x, y), который выполняется тогда и только тогда, когда x является гёделевым номером некоторой формулы, а y – гёделевым номером доказательства отрицания этой формулы. Очевидно, что этот предикат разрешим. Подобную конструкцию во вспомогательных целях использовали, например, Клини [34]) и Б. Россер [35]. Следовательно, предикат NPr(x, y) «выразим» в PA некоторой формулой NProv(x, y), а факт существования в PA недоказуемой формулы может быть «выражен» формулой

$$\exists x \exists y \mathrm{NProv}(x, y). \tag{**}$$

В PA выводима формула ($\mathbf{0} = \mathbf{0}$), значит, выводима и формула $\neg\neg(\mathbf{0} = \mathbf{0})$, поскольку в PA $\vdash(\mathbf{0} = \mathbf{0}) \sim \neg\neg(\mathbf{0} = \mathbf{0})$. Пусть $\boldsymbol{k}$ – нумерал, соответствующий гёделеву номеру формулы $\neg(\mathbf{0} = \mathbf{0})$, а $\boldsymbol{n}$ – нумерал, соответствующий гёделеву номеру вывода формулы $\neg\neg(\mathbf{0} = \mathbf{0})$. Из определения предиката NPr(x, y) следует, что NPr(k, n) истинно. Тогда по условиям «выразимости» в PA доказуема формула NProv($\boldsymbol{k}, \boldsymbol{n}$). Дважды применяя к последней формуле закон экзистенциального обобщения, получаем вывод формулы (**). Таким образом, в PA доказуема формула, «выражающая» существование недоказуемой формулы. Следовательно, арифметика вполне может «доказать свою собственную непротиворечивость» абсолютно в гёделевом смысле этой формулировки.

Чем формула (**) хуже гёделевой формулы Consis? Её построение пошагово совпадает с построением гёделевой формулы, за исключением того, что в ней предикат доказуемости заменён на предикат недоказуемости, «выразимый» так же хорошо, как и предикат Гёделя. Предикат NPr(x, y) определён столь же корректно, что и гёделев предикат NPr(x, y): последний перечисляет доказуемые в PA формулы основываясь на определении правильного вывода, а NPr(x, y) перечисляет недоказуемые формулы основываясь также на определении непротиворечивости. Поэтому предложение (**) вполне можно «точным образом интерпретировать как выражающее непротиворечивость какой-либо логистической системы, содержащей арифметику» [36]. И оно доказуемо в PA.

Конечно, предикат NPr(x, y) перечисляет не все недоказуемые формулы, так как неразрешимые (A недоказуема, но $\neg A$ также не доказуема) формулы как недоказуемые он не опознает. Но для «выражения» непротиворечивости PA этого и не требуется: достаточно указать хотя бы на одну недоказуемую формулу, с чем NPr(x, y) прекрасно справляется.

Зачастую вторая теорема о неполноте формулируется так:

PA непротиворечива в том и только том случае, если в PA не может быть доказана формула, «выражающая» непротиворечивость PA.

Теперь понятно, что такая формулировка неверна. Формулы BessNProv(x) и NProv(x, y) доказуемы в PA и прекрасно «выражают» непротиворечивость PA. Однако из этого противоречивость PA никак не следует. Противоречие согласно гёделевскому доказательству возникло бы только тогда, когда наши формулы, «выражающие» непротиворечивость PA, были бы эквивалентны гёделевой формуле Consis. В такой ситуации было бы, например, доказуемо

$$(**) \rightarrow \mathrm{G}$$

(G – гёделева неразрешимая формула), т.к. по второй теореме $\vdash$Consis $\rightarrow$ G. Из этого, поскольку (**) доказуемо в PA, следовало бы $\vdash$G, что по первой теореме означало бы противоречивость PA. Однако наши формулы, «выражающие» недоказуемость, очевидно не эквивалентны гёделевой формуле NProv(x, y). В самом деле, как мы видели, $\vdash$ NProv($\boldsymbol{k}, \boldsymbol{n}$), где $\boldsymbol{k}$ – нумерал, соответствующий гёделеву номеру формулы $\neg(\mathbf{0} = \mathbf{0})$, а $\boldsymbol{n}$ – нумерал, соответствующий гёделеву номеру вывода формулы $\neg\neg(\mathbf{0} = \mathbf{0})$. Но $\vdash$ Prov($\boldsymbol{k}, \boldsymbol{n}$) означало бы $\vdash\neg(\mathbf{0} = \mathbf{0})$, т.е. противоречивость PA. Поэтому формула (**) не эквивалентна Consis, если PA непротиворечива.

Отметим, что мы доказали формулу, «выражающую» непротиворечивость PA, вообще не используя какой-либо предикат доказуемости, «естественный» или нет. Однако предикаты доказуемости конечно же могут быть определены и в нашей системе. При этом соответствующие формулы могут быть связаны друг с другом. Например, NProv(x, y) – это ни что иное, как Prov($\sigma(x), y$), где Prov(x, y) – гёделева формула доказуемости, а $\sigma(x)$ по гёделевому номеру формулы x определяет гёделев номер отрицания этой формулы. Очевидно, что наш результат останется в силе независимо от того, удовлетворяет или нет «порождающий» предикат доказуемости условиям Гильберта – Бернайса.

Л.Д. Беклемишев в своей обзорной статье, посвящённой теоремам Гёделя о неполноте, пишет: «С нашей точки зрения, еще рано выносить окончательное суждение о том, что представляет собой `правильный’ контекст для второй теоремы Гёделя о неполноте» [37]. С нашей же точки зрения, эта пора уже настала. Пора сделать вывод, что вторая теорема Гёделя в её универсальном прочтении попросту не верна. От этой теоремы остается только нечто вроде «если PA непротиворечива, то в ней имеется некоторая формула, одним из возможных способов «выражающая» непротиворечивость PA и недоказуемая в PA. При этом существуют другие, ничуть не хуже «выражающие» непротиворечивость PA и доказуемые в PA формулы». Поэтому арифметика не может доказать свою непротиворечивость, только если она (или кто-то из её пользователей) очень этого не хочет.

Думается, что вряд ли кто из адептов первой догмы, тех, кто свято верит в то, что в любой содержащей арифметику теории недоказуема любая формула, «выражающая» её непротиворечивость, если эта теория непротиворечива, обрадуется низведению статуса второй теоремы до рядового следствия первой теоремы о неполноте. Ведь значимость второй теоремы, первоначально действительно воспринятой как «довесок» к первой, обусловилась именно её универсальным прочтением. Только при таком прочтении она могла рассматриваться в качестве решающего аргумента неосуществимости гильбертовской программы финитного обоснования математики, т.е. служить основанием для второй догмы. Но у нас не осталось никаких оснований повторять как мантру «арифметика, если она непротиворечива, не может доказать свою собственную непротиворечивость». И это подрывает святую веру во вторую догму. Подробнее об этом – в следующей части работы.

Примечания

1. См.: *Gödel K.* Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme I // Monatshefte für Mathematik und Physik. – 1931. – Bd. 38. – S. 173–198.

2. См.: *Gödel K.* On formally undecidable propositions of Principia mathematica and related systems I. – S. Feferman, J.R. Dawson, S.C. Kleene, G.H. Moore, R.M. Solovay, and J. van Heijenoort (eds.). Kurt Gödel. Collected Works, Vol. 1. – New York: 1986. – P. 195.

3. Ibid.

4. Ibid.

5. См.: *Крайзель Г.* Биография Курта Геделя // Успехи математических наук. – 1988. – Т. 43, Вып. 2(260). – С. 184.

6. См.: *Gödel K.* On formally undecidable propositions of Principia mathematica and related systems I. – P. 193.

7. См.: *Gödel K.* On formally undecidable propositions of Principia mathematica and related systems I. – P. 195.

8. См.: *Hilbert D., Bernays P.* Grundlagen der Mathematik., V. II. – Berlin: 1939. – P. 285–328.

9. См.: *Гильберт Д., Бернайс П.* Основания математики. Теория доказательств. – М.: Наука, 1982. – С. 355.

10. См.: *Френкель А.А., Бар-Хилел И.* Основания теории множеств. – М., 1966. – С. 370.

11. См.: *Клини С.К.* Введение в метаматематику. – М.: Изд-во иностранной литературы, 1957. – С. 190.

12. См.: *Адян С.И.* Предисловие редактора русского перевода. – В кн.: Гильберт Д., Бернайс П. Основания математики. Логические исчисления и формализация арифметики. – М.: Наука, 1979. – С. 14.

13. См.: *Kennedy J.* The Stanford Encyclopedia of Philosophy, http://plato.stanford.edu/ entries/goedel/#SecIncThe, 2011.

14. См.: *Raatikainen P.* On the philosophical relevance of Gödel's incompleteness theorems // Revue Internationale de Philosophie. – 2005. – V. 59. – No. 4. – C. 513.

15. См.: *Raatikainen P.* The Stanford Encyclopedia of Philosophy http://plato.stanford.edu/entries/goedel-incompleteness/, 2013.

16. См.: *Field H.* Truth and the unprovability of consistency // Mind. – 2006. – No.115 (459). – P. 567.

17. См.: *Gödel K.* On formally undecidable propositions of Principia mathematica and related systems I. – Theorems VI, XI.

18. Ibid.

19. Ibid.

20. См.: *Ершов Ю.Л., Самохвалов К.Ф.* Современная философия математики: недомогания и лечение. – Новосибирск, 2007. – С. 112.

21. См.: *Мендельсон Э.* Введение в математическую логику. – М., 1976. – С. 165.

22. Ibid. – С. 132.

23. См.: *Куайн У.* Слово и объект. – М.: Логос, 2000.

24. См.: *Тарский А.* Истина и доказательство // Вопросы философии. – 1972. – № 8. – С. 136–145.

25. См.: *Клини С.К.* Введение в метаматематику. – С. 473.

26. См.: *Hilbert D., Bernays P.* Grundlagen der Mathematik., V. II.

27. См.: *Löb M.H.* Solution of a problem of Leon Henkin // Journal of Symbolic Logic. – 1955. – V. 20. – P. 115–118.

28. См.: *Jeroslow R*. Redundancies in the Hilbert – Bernays derivability conditions for Gödel's second incompleteness theorem // Journal of Symbolic Logic. – 1973 – V. 38. – P. 359–367.

29. См.: *Булос Дж., Джеффри Р*. Вычислимость и логика. – М.: Мир, 1994. – С. 246.

30. Ibid.

31. Ibid. – С. 249.

32. См.: *Бессонов А.В*. К интерпретации теорем Гёделя о неполноте арифметики // Вестник ТГУ, Сер. философия, социология, политология. – 2011. – № 4. – С. 177–189.

33. См.: *Булос Дж., Джеффри Р*. Вычислимость и логика. – С. 245–246.

34. См.: *Клини С.К*. Введение в метаматематику. – С. 186.

35. См.: *Rosser B*. Extensions of some theorems of Gödel and Church // Journal of Symbolic Logic. 1936. – V. 1. – No. 3. – P. 87–91.

36. См.: *Френкель А.А., Бар-Хилел И.* Основания теории множеств. С. 370

37. См.: *Беклемишев Л.Д*. Теоремы Гёделя о неполноте и границы их применимости. I // Успехи математических наук. – 2010. – Т. 65. – Вып. 5(395). – С. 63.

Дата поступления 27.11.2014

Институт философии и права СО РАН,
Новосибирский государственный
университет, г. Новосибирск

trt@academ.org

Bessonov A.V. **Two false dogmas related with Gödel's second incompleteness theorem. I**

It is shown that Gödel's proof of the second incompleteness theorem for formal arithmetic depends on a chosen provability predicate. We use an unprovability predicate to construct counterexamples to the second theorem and prove that in the generall case the conclusion of the second theorem is not true.

Keywords: Gödel's incompleteness theorems, inadequacy of a provability predicate, unprovability predicate.