ВЕСТНИК ТОМСКОГО ГОСУДАРСТВЕННОГО УНИВЕРСИТЕТА

2011 Философия. Социология. Политология №4(16)

МОНОЛОГИ, ДИАЛОГИ, ДИСКУССИИ

ОТ РЕДКОЛЕГИИ: Материалы, публикуемые в рубрике «МОНОЛОГИ, ДИАЛОГИ, ДИСКУССИИ», являются *ДИСКУССИОННЫМИ*, т.е. мнение редакционной коллегии журнала «ВЕСТНИК ТОМСКОГО ГОСУДАРСТВЕННОГО УНИВЕРСИТЕТА. ФИЛОСОФИЯ. СОЦИОЛОГИЯ. ПОЛИТОЛОГИЯ» может не совпадать с мнением автора идущей ниже статьи. Рубрика предполагает возможность широкого обсуждения с привлечением мнений пропонентов и оппонентов. Все поступившие в редколлегию материалы относительно идущей ниже статьи будут рассмотрены и, по возможности, опубликованы.

УДК 164.07

А.В. Бессонов

К ИНТЕРПРЕТАЦИИ ТЕОРЕМ ГЁДЕЛЯ О НЕПОЛНОТЕ АРИФМЕТИКИ[1]

Опровергается общепринятое универсальное ограничительное истолкование знаменитых теорем К. Гёделя о неполноте арифметики. Приводятся контрпримеры ко второй теореме, показывается ограниченность используемых Гёделем выразительных средств. В рамках гёделева подхода доказывается третья теорема о неполноте, по которой неразрешимыми оказываются самые обычные в (мета)арифметике суждения, причём таких суждений бесконечно много. Тем самым обосновывается вывод о принципиальной неадекватности гёделева представления знания, из чего следует неправомерность переноса полученных в таком представлении выводов на содержательное знание.

Ключевые слова: *теоремы Гёделя о неполноте, неразрешимость, предикат недоказуемости, третья теорема о неполноте, представление знания.*

Почти общепризнано, что знаменитые теоремы К. Гёделя о неполноте арифметики [1, 2] свидетельствуют о принципиальной ограниченности метода формализации даже в области математического знания. Они рассматриваются как решающий аргумент о несостоятельности гильбертовской финитистской программы [3] обоснования математики (с такой оценкой в свое время был согласен и сам Д. Гильберт). Более того, на них основывается вывод о принципиальной неполноте любых более богатых по сравнению с арифметикой теорий и в других отраслях знания.

Убеждение в сверхзначимости теорем Гёделя о неполноте формальной арифметики давно вышло за пределы математики и ныне тотально распространено на философию, гносеологию [4], социологию [5], юриспуденцию и т.д. и является «почти неисчерпаемым источником интеллектуальных злоупотреблений» [6]. Между тем имеются свидетельства того, что А.Н. Колмогоров не верил в распространение этих теорем на все известные теории при любых их построениях [7]. Я. Хинтикка предложил «независимо-

[1] Исследования, результаты которых отражены в данной статье, поддержаны Междисциплинарным интеграционным проектом Сибирского отделения РАН № 47.

дружественную кванторную логику», в которой теоремы Гёделя не выполняются [8]. Да и сам Гёдель вовсе не был убеждён в величии и универсальности своих результатов и следующих из них выводов [9]. Не покушаясь на собственно математическую сторону доказательств, а сосредоточившись на критике общепринятого истолкования этих теорем, мы покажем, что их значимость чрезвычайно преувеличена даже в родной для них области логики и математики.

§1. Теоремы Гёделя о неполноте: комментарии

Приведём краткое изложение гёделевых доказательств теорем о неполноте формальной арифметики[2].

Напомним, что 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 формальная арифметика Пеано (PA) ω-непротиворечива, то она неполна[2]. Более точно, в ней доказывается существование некоторой замкнутой формулы («говорящей» о своей собственной недоказуемости) такой, что ни она, ни ее отрицание не доказуемы в PA (такие формулы называются неразрешимыми). В соответствии со второй теоремой Гёделя, если PA непротиворечива, то в ней не доказуема формула, выражающая непротиворечивость PA [2. Theorems VI, XI].

Вторая теорема обычно доказывается как следствие первой, но при этом ей придаётся особый по сравнению с первой статус. Считается, что вторая теорема даёт конкретный пример неразрешимой формулы, и самое главное, что в ней устанавливается фундаментальный факт невозможности доказательства непротиворечивости арифметики средствами самой арифметики: «арифметика не может доказать свою собственную непротиворечивость».

Эта широко используемая метафорическая формулировка нуждается в пояснении. То, что никакая теория не в состоянии обосновать саму себя, известно с античных времен независимо от теорем Гёделя. Следует понимать, что во второй теореме речь идет лишь о недоказуемости некоторой формулы (тем или иным 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 (имеющаяся в PA) формула, «выражающая» формулировку второй теоремы: «Если PA непротиворечива, то формула, выражаю-

[1] Подчеркнём, что в данной работе мы рассматриваем именно гёделевы доказательства в том виде, в котором они изложены в [2], только в более современной и удобной нотации (предложенной Р. Карнапом).

[2] Д.Б. Россер показал, что для доказательства первой теоремы достаточно предположить непротиворечивость PA [10], но для нашего изложения это несущественно.

щая непротиворечивость PA, в ней не доказуема». Рассмотрим вторую теорему более внимательно.

Доказательства Гёделя основаны на кодировании языка формальной арифметики и её логики (гёделева нумерация), а также на введённом им понятии «определимости» («выразимости»).

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

если $F(n_1,\dots,n_k)$ выполняется, то $\vdash \Phi(n_1,\dots,n_k)$;

если $F(n_1,\dots,n_k)$ не выполняется, то $\vdash \neg\Phi(n_1,\dots,n_k)$.

Здесь, как обычно, $\vdash$ означает доказуемость (в PA). Для сокращения записи условимся вместо нумералов в PA использовать обычные цифры, т.е., например, вместо $0'''$ писать 3, $\ulcorner A \urcorner$ будет обозначать гёделев номер формулы A.

Содержательно определение непротиворечивости теории может быть дано самыми разными способами. Обычно используется следующее определение:

(1) теория T непротиворечива, если в ней ни для какой формулы A не могут одновременно быть доказаны формулы A и $\neg A$.

Мы, как и сам Гёдель [2. Theorem XI], будем использовать другое содержательное определение непротиворечивости:

(2) теория T непротиворечива, если в ней имеется хотя бы одна недоказуемая формула.

Как известно, эти два определения эквивалентны: если теория T непротиворечива в соответствии с первым определением, то она непротиворечива в соответствии со вторым, и наоборот. В самом деле, из (1) следует, что для любой формулы A либо она, либо её отрицание не доказуемо, т.е. существует хотя бы одна недоказуемая формула, т.е. следует (2). Пусть верно (2). Из этого следует (1), так как в противном случае в T была бы доказуемой какая-то формула вместе с её отрицанием, и по правилу $A \& \neg A \vdash B$ доказуемой была бы любая формула.

При доказательстве теорем о неполноте синтаксис и логика PA нумеруются посредством гёделевой нумерации, после чего вводится предикат доказуемости $\Pr(x,y)$, 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 «определим» в PA тогда и только тогда, когда он разрешим. Значит, $\Pr(x,y)$ «определим» в PA с помощью некоторой арифметической формулы $\mathrm{Prov}(x,y)$, т.е. для любых натуральных чисел n, k верны условия:

если $\Pr(n,k)$ выполняется, то $\vdash \mathrm{Prov}(n,k)$;

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

В этих обозначениях определение непротиворечивости (2) выражается арифметической формулой

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

а в силу эквивалентности определений непротиворечивости (1) и (2), формула $(*)$ (которую С. Клини назвал «Consis») «выражает» непротиворечивость PA независимо от того, какое именно содержательное определение непротиворечивости имеется в виду.

Во второй теореме Гёделя доказывается, что если PA непротиворечива, то формула $(*)$ не может быть доказана в PA. Отсюда Гёдель немедленно делает вывод, что «непротиворечивость P (PA в наших обозначениях. – *А.Б.*) недоказуема в P» [2. P. 193]. Отметим, что во второй теореме этот вывод доказывается исключительно по отношению к гёделевой формуле Consis. Однако Гёдель, видимо, предполагая, что других формул, «выражающих» непротиворечивость PA, нет, тут же максимально обобщает вывод второй теоремы: «Доказательство теоремы XI 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» [2. P. 195].

Очевидно, что подобное обобщение второй теоремой никак не обосновано. Доказательства второй теоремы явно недостаточно для вывода, что в PA не может быть доказана никакая другая (не эквивалентная гёделевой Consis) формула, «выражающая» непротиворечивость PA. Ведь даже студенту первого курса мехмата известно, что из существования формулы, удовлетворяющей некоторому свойству («выражать» непротиворечивость PA и быть недоказуемой), логически не следует, что все формулы должны удовлетворять этому свойству. Из второй теоремы никак не следует несуществование «*других, чем в теореме Гёделя*, быть может не столь естественных (или естественных как-то по-другому), но всё же разрешимых в P арифметических выражений непротиворечивости... Странным образом последнее обстоятельство не было замечено в 1931 году» [11. С. 112].

Вопрос о ином способе построения Consis вообще не поднимается ни самим Гёделем, ни многими его последователями. Приводим цитаты без купюр:

«Если арифметическая формальная система гл. IV (просто) непротиворечива, то не $\vdash$ Consis; иначе говоря, если указанная система непротиворечива, то не существует (!? – *А.Б.*) доказательства её непротиворечивости, проведённого средствами, формализуемыми в этой системе (Вторая теорема Гёделя)» [12. С. 190].

«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» [13. С. 165].

Следует сказать, что Мендельсон все-таки оговаривается, указывая на результат С. Фефермана [14. P. 35–92]: «... как показал Феферман [1960], существует некоторый приемлемый способ построения Con_S, при котором $\vdash_S Con_S$. Итак, следует уточнить формулировку теоремы». Уточнение сводится к тому, что в формулировке второй теоремы фраза «некоторая формула, выражающая...» заменяется на «некоторая RE-формула, выражающая...» (RE-формула – это формула вида $\exists y_1 \ldots \exists y_k A$, где $k \geqslant 0$ и A – формула «выражающая» какую-либо примитивно рекурсивную функцию) [13. С. 166–167]. Значит, в PA все-таки существует «выражающая» непротиворечивость PA формула (пусть и не RE-формула), и она доказуема в PA, т.е. PA все-таки может «доказать свою непротиворечивость»!

Еще один контрпример ко второй теореме Гёделя построил К.Ф. Самохвалов [15]. Однако и он, и Феферман строят свои формулы, «выражающие» непротиворечивость PA, как производные от гёделевых предиката доказуемости и формулы Consis. А нельзя ли построить формулы, «выражающие» непротиворечивость PA и доказуемые в ней, используя принципиально иные выразительные средства?

§2. Вторая теорема Гёделя: контрпримеры

Для ответа на поставленный вопрос рассмотрим свойства гёделевых формулы $\mathrm{Prov}(x, y)$ и предиката $\mathrm{Pr}(x, y)$.

Из второй теоремы, как известно, следует, 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. $\vdash (*)$. Но это означает, что в PA невозможно доказать недоказуемость никакой – ни доказуемой, ни недоказуемой, ни неразрешимой (если таковые есть) формулы. Этот обескураживающий факт весьма просто объясняется выбором базового в доказательствах теорем о неполноте предиката $\mathrm{Pr}(x, y)$. Его разрешающая процедура такова:

Выберем пару натуральных чисел (x, y) и проверим, является ли x гёделевым номером формулы, а y – гёделевым номером доказательства. По гёделевым номерам восстановим формулу с номером x и доказательство с номером y. Доказательство представляет собой последовательность формул. Рассмотрим заключительную формулу этой последовательности и сравним её с формулой, имеющей номер x. Если эти формулы совпадут, то $\mathrm{Pr}(x, y)$ истинно.

Как по этому алгоритму установить доказуемость формулы – понятно. А как установить недоказуемость? Восстановив по номеру x формулу A, мы должны последовательно перебрать номера всех доказательств, восстановить их и убедиться, что во всех случаях заключительная формула доказательства отличается от A. Ясно, что если РА непротиворечива и A доказуема, то установить недоказуемость A нельзя. Если же A недоказуема, то этот процесс никогда не остановится, даже если заключительной формулой какого-то вывода окажется формула $\neg A$: она также отличается от A. Требование ω-непротиворечивости в доказательстве первой теоремы даёт дополнительную (как следует из результата Б. Россера, даже избыточную) гарантию того, что недоказуемость A не может быть установлена каким-либо иным образом.

Отсюда следует, что подлинной причиной недоказуемости неразрешимой гёделевой формулы является не какое-то врождённое свойство РА, а то, что базовый предикат $\mathrm{Pr}(x, y)$, более или менее пригодный для представления рассуждений о доказуемости, совершенно не подходит для представления рассуждений о недоказуемости. Этот предикат мажет одной краской все формулы, «выражающие» недоказуемость какой бы то ни было формулы, а не только те, которые «говорят» о своей собственной недоказуемости или о непротиворечивости РА в целом. А ведь при доказательствах непротиворечивости или неразрешимости требуется адекватность представления рассуждений именно о недоказуемости! Таким образом, попытки установить недоказуемость какой-либо формулы РА, используя исключительно гёделевы выразительные средства, подобны попыткам выкопать яму зубочисткой, для этого заведомо не приспособленной. А может, вместо зубочистки прибегнуть к лопате или экскаватору?

Чрезвычайно важно понимать, что выбор предиката $\mathrm{Pr}(x, y)$ и соответствующей ему формулы $\mathrm{Prov}(x, y)$ (или производных от них) в качестве исключительных средств для представления рассуждений о (не)доказуемости в РА, несмотря на кажущуюся его естественность, Гёделем и его последователями *никак* не обосновывается. Эти средства выбраны абсолютно произвольно, при этом вопрос о возможности использования принципиально иных выразительных средств даже не поднимается. Мы же далее рассмотрим именно этот вопрос.

По алгоритму разрешения предиката $\mathrm{Pr}(x, y)$ для установления недоказуемости формулы требуется осуществить бесконечный перебор номеров всех выводов. Но для построения формулы, «выражающей» факт непротиворечивости РА, достаточно указать всего на одну недоказуемую в РА формулу, которая, конечно же, имеется в предположении, что РА непротиворечива. Отсюда просто напрашивается мысль о введении предиката недоказуемости.

Рассмотрим предикат $\mathrm{BessNPr}(x)$, который выполняется тогда и только тогда, когда x является гёделевым номером первой пришедшей в голову автору этой статьи недоказуемой в РА формулы. Заверяю, что это была формула $\neg(0 = 0)$, очевидно недоказуемая в РА, если последняя непротиворечива.

Этот предикат «определим» в PA формулой $\mathrm{BessNProv}(x)$, эквивалентной формуле $x = \ulcorner\neg(0=0)\urcorner$. В самом деле, предикат $\mathrm{BessNPr}(x)$ выполняется только для значения x, совпадающего с $\ulcorner\neg(0=0)\urcorner$, и в этом случае в PA $\vdash \ulcorner\neg(0=0)\urcorner = \ulcorner\neg(0=0)\urcorner$, т.е. $\vdash \mathrm{BessNProv}(\ulcorner\neg(0=0)\urcorner)$. При любом другом значении x, скажем, l, не равном $\ulcorner\neg(0=0)\urcorner$, $\mathrm{BessNPr}(l)$ не выполняется, и в этом случае, очевидно, $\vdash \neg\mathrm{BessNProv}(l)$, так как для любого $l \neq \ulcorner\neg(0=0)\urcorner$ в PA доказуемо $\neg(l = \ulcorner\neg(0=0)\urcorner)$. Поскольку в PA $\vdash \mathrm{BessNProv}(\ulcorner\neg(0=0)\urcorner)$, по закону экзистенциального обобщения получаем

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

Иными словами, в PA доказуема формула, «выражающая» существование в PA первой пришедшей в голову автору 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.

Для тех, кто посчитает приведённый контрпример ко второй теореме незначимым на том основании, что в PA нет никакого «автора», следует сказать, что в PA нет также ни гёделевой, ни какой-либо иной нумерации: из каких аксиом PA следует существование нумераций? 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 пришествия в голову А.В. Бессонова некоей недоказуемой в PA формулы. 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. Можно привести и более математичные примеры.

Рассмотрим предикат $\mathrm{MinNPr}(x)$, который выполняется тогда и только тогда, когда x является наименьшим гёделевым номером недоказуемой в PA формулы. Очевидно, что такой номер n существует, поскольку минимум ищется в конечном множестве $\overline{0, \ulcorner\neg(0=0)\urcorner}$. Проводя те же рассуждения, что и в случае предиката $\mathrm{BessNPr}(x)$, получаем, что предикат $\mathrm{MinNPr}(x)$ «определим» в PA формулой $\mathrm{MinNProv}(x)$, эквивалентной формуле $x = n$, и что в PA доказуема формула

$$\exists x \mathrm{MinNProv}(x),$$

«выражающая» существование в PA недоказуемой формулы с наименьшим гёделевым номером. Но из этого также следует существование в PA недоказуемой формулы, т.е. непротиворечивость PA.

Чем приведённые выше формулы, «выражающие» непротиворечивость PA, хуже гёделевой формулы Consis? А.А. Френкель и И. Бар-Хилел формулируют следствие второй теоремы Гёделя так: «никакое предложение, которое можно *точным образом* (курсив наш. – *А.Б.*) интерпретировать как выражающее непротиворечивость какой-либо логистической системы, содержащей арифметику, не может быть доказано в этой системе» [16]. Очевидно, что наши формулы можно интерпретировать как выражающие непротиворечивость PA ничуть не менее точным образом, чем гёделев Consis. Эти формулы вполне по-гёделевски «выражают» факт существования в PA хотя бы одной недоказуемой формулы, что и означает непротиворечивость PA.

Г. Крайзель уточняет формулировку второй теоремы следующим образом: «Если система S непротиворечива, а относительно формулы **A** в S может быть доказано, что **A** выражает непротиворечивость S, то **A** не может быть доказана в S» [17]. Но как доказать в PA, что сам гёделев Consis выражает непротиворечивость PA? Если Крайзель имеет в виду требование, чтобы любая формула, выражающая непротиворечивость PA, должна быть доказуемо эквивалентной в PA гёделевой формуле Consis или её производным (в духе Фефермана), то это напоминает ситуацию, когда требуется доказать, что лопатой или ковшом экскаватора тоже можно чистить зубы.

Приведённые выше контрпримеры построены с использованием сингулярных, т.е. выполнимых относительно одной формулы предикатов. Можно привести пример и более общего предиката недоказуемости. Назовём формулу A «заведомо недоказуемой», если в PA доказуема формула $\neg A$. Ясно, что из заведомой недоказуемости формулы A следует недоказуемость A, если PA непротиворечива. Рассмотрим предикат $\mathrm{NPr}(x, y)$, 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. Его разрешающая процедура отличается от таковой для гёделева предиката $\mathrm{Pr}(x, y)$ только тем, что заключительная формула вывода с номером y сравнивается не с самой формулой, имеющей номер x, а с её отрицанием. Следовательно, предикат $\mathrm{NPr}(x, y)$ «определим» в PA некоторой формулой $\mathrm{NProv}(x, y)$, а факт существования в PA заведомо недоказуемой формулы «выражается» формулой

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

В PA выводима формула $(0 = 0)$, значит, выводима и формула $\neg\neg(0 = 0)$, поскольку в PA $\vdash ((0 = 0) \sim \neg\neg(0 = 0))$. Пусть n – гёделев номер вывода формулы $\neg\neg(0 = 0)$. Из определения предиката $\mathrm{NPr}(x, y)$ следует, что $\mathrm{NPr}(\ulcorner\neg(0 = 0)\urcorner, n)$ истинно. Тогда по условиям «определимости» в PA доказуема формула $\mathrm{NProv}(\ulcorner\neg(0 = 0)\urcorner, n)$. Дважды применяя к последней формуле закон экзистенциального обобщения, получаем вывод формулы $(**)$. Таким образом, в PA доказуема формула, «выражающая» существование заведомо недоказуемой, а значит, и просто недоказуемой формулы. Поэтому арифметика вполне может «доказать свою собственную непротиворечивость».

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 в приведённых контрпримерах доказуемы в PA и прекрасно выражают непротиворечивость PA. Однако из этого противоречивость PA никак не следует.

Заметим, что в приведённых примерах использованы вполне RE-формулы. Поэтому и в переформулировке Мендельсона–Фефермана вторая теорема о неполноте не имеет универсального значения. От этой теоремы остаётся 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». Вторая теорема, таким образом, лишается ошибочно приписываемого ей пафосного значения и должна рассматриваться просто как одно из рядовых следствий первой теоремы о неполноте. В то же время это следствие, как мы покажем ниже, заставляет пересмотреть общепринятое универсальное истолкование первой теоремы.

§3. Неадекватность гёделева представления

Под представлением знаний обычно понимают задачу структурирования знаний с целью формализации процессов обработки информации в определённой проблемной области. Эта задача актуальна для когнитологии, информатики, а также в исследованиях по искусственному интеллекту. К. Гёдель, сформулировав и развив метод арифметизации синтаксиса и логики научной теории, по существу дал первое систематическое решение задачи представления знания. Заслуга Гёделя в этом, а также значимость аппарата кодирования, созданного им при доказательстве теорем о неполноте, несомненны.

Вместе с тем важнейшим требованием к решению задачи представления знания является требование логической адекватности: представление должно обладать способностью распознавать все существенные отличия, заложенные в представляемом знании. В частности, в представлении не должны устанавливаться результаты, заведомо ложные при обратной перекодировке. В самом деле перенос на представляемое знание выводов, полученных в неадекватном представлении, может приводить к очевидно ложным заключениям. Так, рисуя окружающую действительность исключительно голубой краской, мы получаем более-менее адекватное представление о небе. Но перенос на представляемое знание верного в таком представлении вывода о том, что все мужчины – голубые, вряд ли окажется приемлемым для широкой научной общественности.

Вернёмся к поставленному в начале статьи вопросу о значимости теорем Гёделя. Очевидно, что правомерность переноса на содержательную математику и тем более на иные отрасли знания выводов о принципиальной неполноте теорий основывается на молчаливом предположении об адекватности способа представления знания, используемом при доказательстве теорем Гёделя. Но насколько верно это предположение?

Теоремы Гёделя изначально относятся к PA, формальной арифметике, которую саму можно рассматривать как некоторое представление неформализованного содержательного математического знания. При этом адекватность собственно PA в этом качестве сомнению обычно не подвергается. В PA как таковой нет, да и не должно быть выразительных средств, с помощью которых можно было бы формализовать рассуждения о доказуемости, недоказуемости и неразрешимости. Для представления подобных рассуждений Гёдель использует три основных положения:

I. Нумерирование (арфметизация синтаксиса).

II. Введение определённого понятия «выразимости», с помощью которого содержательные утверждения транслируются в PA.

III. Выбор определённых выразительных средств (базовый предикат $\mathrm{Pr}(x, y)$ и соответствующая ему формула $\mathrm{Prov}(x, y)$).

Понятно, что для обоснования универсального истолкования гёделевых теорем каждый из этих пунктов должен быть положительно оценен в плане его универсальности и адекватности. Оставляя пока в стороне п. I и II, оценим п. III.

Как мы убедились, избранные Гёделем выразительные средства далеко не уникальны: формула $\mathrm{NProv}(x, y)$ ничуть не хуже «выражает» определённые рассуждения о (не)доказуемости, чем гёделева формула $\mathrm{Prov}(x, y)$. При этом утверждение второй теоремы оказывается предикатно зависимым: основанная на $\mathrm{Prov}(x, y)$ формула, «выражающая» непротиворечивость PA, недоказуема, а основанная на $\mathrm{NProv}(x, y)$ – элементарно доказуема. Но это прямо опровергает общепринятую универсальную трактовку второй теоремы о неполноте: «В *любой* теории, содержащей арифметику, недоказуема *любая* формула, выражающая её непротиворечивость, если арифметика непротиворечива».

С адекватностью гёделевого представления, его соответствием содержательному математическому знанию дело обстоит ещё хуже. Как уже было сказано, следствием второй теоремы является то, что в PA ни для какой формулы A не может быть доказана «выражающая» её недоказуемость формула $(*_A)$:

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

Этот факт известен, но, по-видимому, не известно, что следствием его при использовании для «выражения» рассуждений о доказуемости исключительно формулы $\mathrm{Prov}(x, y)$ (т.е. в тех рамках, в которых Гёделем и проводятся доказательства теорем о неполноте PA) является следующая

ТЕОРЕМА (третья теорема о неполноте). *Если* PA *непротиворечива, то для любой формулы* A *формула, «выражающая» недоказуемость* A, *недоказуема в* PA.

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

ДОКАЗАТЕЛЬСТВО. Пусть A – некоторая недоказуемая в PA формула.

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

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

должна быть недоказуемой. По гипотезе последняя формула в перечне доказуема. Значит, найдётся номер k такой, что формула $\neg\mathrm{Prov}(\ulcorner A\urcorner,k)$ будет недоказуемой. В таком случае по условиям определимости формула $\neg\mathrm{Pr}(\ulcorner A\urcorner,k)$ не может быть истинной, следовательно, она ложна, откуда вытекает, что формула $\mathrm{Pr}(\ulcorner A\urcorner,k)$ истинна. По определению предиката Pr последнее означает, что формула *A* доказуема. Приходим к противоречию, поскольку мы предположили, что *A* – недоказуемая формула.

Из этой теоремы следует существование бесконечного числа неразрешимых в PA замкнутых формул, причём вовсе не обязательно производных от гёделевой неразрешимой формулы, либо «выражающих» автоссылки, сходные с «говорит о своей собственной недоказуемости», или суждения о PA в целом, подобные «говорит о непротиворечивости PA». Неразрешимыми оказываются формулы, выражающие самые банальные утверждения. Например, из теоремы следует, что формула

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

«выражающая» недоказуемость формулы $\neg(0=0)$, неразрешима в PA.

Математики могут смириться с тем, что некая загадочная, существующая «в принципе» и интуитивно необозримая формула, «выражающая» свою собственную недоказуемость, неразрешима в PA. Так же, основываясь на смутных воспоминаниях из университетского курса философии, они могут согласиться с тем, что «арифметика не может доказать свою непротиворечивость». Но укажите на нормального математика, который согласится считать неразрешимым суждение о недоказуемости формулы $\neg(0=0)$! Ведь на самом деле недоказуемость этой формулы доказывается совершенно элементарно методом от противного: если бы формула $\neg(0=0)$ была доказуемой, то, учитывая доказуемость в PA формулы $(0=0)$, по правилам пропозициональной логики доказуемой была бы и формула $(0=0)\&\neg(0=0)$, т.е. арифметика была бы противоречивой.

Подчеркнем, что правомерность подобного рассуждения в содержательной математике никем не оспаривается. В нём не используется ни аксиома индукции, ни тем более трансфинитная индукция. С применением закона исключённого третьего в данном рассуждении согласятся даже интуиционисты, поскольку в нём рассматривается лишь конечное множество формул (две формулы) и применяется никем не оспариваемое правило вывода пропозициональной логики $A, B \vdash A\&B$.

Но грош цена представлению содержательного математического знания (а следовательно, и полученным в данном представлении результатам), в котором теоремно устанавливается невозможность доказать формулу, «выражающую» недоказуемость $\neg(0=0)$ в предположении о непротиворечивости арифметики! А во сколько мы оценили бы представление содержательного математического знания посредством некоей «формальной арифметики», в которой было бы доказуемо $2\times 2=5$ или оказывалась бы неразрешимой формула $1+1=2$? Поскольку недоказуемость формулы $\neg(0=0)$ элементарно и бесспорно устанавливается в содержательной арифметике, в гёделевом представлении оказываются доказуемыми *заведомо ложные* утверждения, прямо противоречащие самой обыденной математической практике. Но представление, в котором доказываются ложные в представляемом знании суждения, не отвечает требованию логической адекватности и безусловно должно быть отвергнуто.

Сказанное косвенно опровергает и первую теорему о неполноте. Из первой теоремы следует вторая, из неё – третья, следствия которой (например, неразрешимость формулы, «выражающей» недоказуемость формулы $\neg(0=0)$) совершенно неприемлемы. Но признав гёделево доказательство существования в PA неразрешимой замкнутой формулы, мы обязаны признать и неразрешимость формулы, «выражающей» недоказуемость $\neg(0=0)$! Значит, предположив, что первая теорема верна, мы получим абсурдные по отношению к представляемому знанию выводы. Таким образом, первая теорема опровергается буквальным reductio ad absurdum.

Суммируя два последних абзаца, можно заключить, что утверждения теорем Гёделя о неполноте PA на самом деле относятся не к PA, а к гёделевому представлению рассуждений о доказуемости и недоказуемости в PA, свидетельствуя о его очевидной неадекватности. Более полные формулировки теорем должны выглядеть приблизительно следующим образом.

Исключительно в рамках некоторой фиксированной нумерации, исключительно в рамках введённого Гёделем понятия «выразимости», при использовании в качестве базовых для представления рассуждений о (не)доказуемости исключительно предиката $\mathrm{Pr}(x, y)$ и «определяющей» его формулы $\mathrm{Prov}(x, y)$, можно доказать, что PA содержит неразрешимую формулу и в ней недоказуема формула, «выражающая» непротиворечивость PA.

Универсальная значимость указанных посылок весьма сомнительна. К тому же, как показано выше, эти посылки приводят к абсурдным следствиям.

Как мы уже отмечали, перенесение на представляемое знание результатов, доказанных в заведомо неадекватном представлении, недопустимо. Поэтому из доказательств теорем Гёделя о неполноте не следуют ни неполнота

PA, ни невозможность доказательства в ней формулы, выражающей её собственную непротиворечивость. Полученные в неадекватном гёделевом представлении выводы о принципиальной неполноте тем более не могут быть перенесены ни на содержательное математическое знание, ни на какие-либо более богатые по сравнению с арифметикой теории. Они также не могут истолковываться как сколько-нибудь значимый аргумент о несостоятельности гильбертовской финитистской программы обоснования математики.

Мы не утверждаем, что наши результаты доказывают полноту PA. Но из них следует, что в теоремах Гёделя о неполноте арифметики не доказано обратное.

В заключение автор выражает искреннюю благодарность участникам совместного философско-математического семинара Отдела математической логики Института математики СО РАН и Отдела философии Института философии и права СО РАН за заинтересованное обсуждение работы и полезные замечания. Отдельная благодарность – д.филос.н. К.Ф. Самохвалову, чья критика и стимулирующее влияние во многом способствовали осуществлению данного исследования.

Литература

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

2. *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. 145–195.

3. *Гильберт Д.* Основания геометрии. М., 1948. С. 322–399.

4. *Пенроуз Р.* Тени разума: в поисках науки о сознании. М., 2005.

5. *Debray R.* Critique de la raison politique. Paris, 1981.

6. *Сокал А., Брикмон Ж.* Интеллектуальные уловки. Критика философии постмодерна. М., 2002.

7. *Кузичев А.С.* Новые колмогоровские теоретико-множественные основания современной математики [Электроннный ресурс]. Режим доступа: http://kuzichev.exponenta.ru/

8. *Hintikka Ja.* The principles of mathematics revisited. London, 1996.

9. *Крайзель Г.* Биография Курта Гёделя // Успехи математических наук. 1988. Т. 43, вып. 2(260). С. 175–216.

10. *Rosser B.* Extensions of some theorems of Gödel and Church // Journal of Symbolic Logic. 1936. Vol. 1, № 3. P. 87–91.

11. *Ершов Ю.Л., Самохвалов К.Ф.* Современная философия математики: недомогания и лечение. Новосибирск, 2007.

12. *Клини С.* Введение в метаматематику. 1957.

13. *Мендельсон Э.* Введение в математическую логику. М., 1976.

14. *Feferman S.* Arithmetization of metamathematics in a general setting // Fundam. Math. 1960. Vol. 49. P. 35–92.

15. *Самохвалов К.Ф.* Уточнения обычной интерпретации теорем Геделя о неполноте и понятия рекурсивной перечислимости // Проблемы логики и методологии науки. Новосибирск, 1982. С. 42–57.

16. *Френкель А.А., Бар-Хилел И.* Основания теории множеств. М., 1966.

17. *Kreisel G.* Rewiew of Feferman's «Arithmetization etc» // Mathematical Reviews. 1963. Vol. 25, № 5. P. 938–939.