Проблемы логики и методологии науки

НЕЗАВЕРШЕННОСТЬ МАТЕМАТИКИ И АБСОЛЮТНО НЕРАЗРЕШИМЫЕ ПРОБЛЕМЫ*

В.В. Целищев

Статья посвящена одному из аспектов дилеммы Геделя о соотношении человека и компьютера и существовании абсолютно неразрешимых утверждений математики. Рассматривается основная посылка дилеммы, а именно, концепция незавершаемости математики как следствие теорем о неполноте. Делается заключение о незавершенности аргумента Геделя о превосходстве человека над конечной машиной в вопросе о разрешимости математических утверждений.

Ключевые слова: Гедель, математика, компьютер, человек, дилемма

В 1951 г. К. Гедель прочел лекцию в Университете Брауна – одну из серии лекций в честь математика Дж.У. Гиббса. Эта лекция не была опубликована при жизни Геделя, хотя он и намеревался это сделать. Впоследствии она вошла в третий том собрания работ Геделя, извлеченных из его записных книжек (Nachlass) [1]. Публикация этой лекции, которая называется ради краткости Гиббсовской лекцией Геделя, стала важным событием в философии математики. В ней Гедель предложил интригующую дилемму, которая, с его точки зрения, следует из его же второй теоремы о неполноте: «Либо математика незавершаема в этом смысле, а ее очевидные аксиомы никогда не могут быть проявлением (comprised) конечного правила, т.е. человеческий ум (даже в пределах чистой математики) бесконечно превосходит возможности (powers) любой конечной машины, или же существуют абсолютно неразрешимые диофантовые утверждения отмеченного типа» [2].

Корни этой поразительной дилеммы лежат в ряде предположений Геделя о природе математического знания, да и не только математическо-

* Исследования, нашедшие отражение в данной статье, поддержаны грантом Российского гуманитарного научного фонда № 13-03-00073.

го, поскольку речь идет о природе человеческого знания вообще. В частности, Гедель в связи с дилеммой различает объективную и субъективную математику; последнюю он называет человеческой (human) математикой. В значительной степени сама по себе дилемма связана с вопросом о соотношении человеческого разума и машины и может быть интерпретирована как поддержка позиции ментализма против механицизма [3].

Разделение математики на объективную и субъективную поднимает ряд серьезных проблем относительно пределов человеческого знания. Психологически аргументация Геделя важна, потому что такой расплывчатый эпистемологический вопрос, как вопрос о границах знания, Геделем доказывается с опорой на один из самых фундаментальных результатов математической логики. Поэтому для понимания важности этих проблем требуется уточнение терминов и определений. Нужно прежде всего разобраться, какое место в дилемме Геделя занимают понятия эффективной процедуры, а также неразрешимые диофантовы проблемы. Далее, что означает, с его точки зрения, незавершенность математики, каким образом он приходит к различению объективной и субъективной математики и каким же образом дилемма связана с позицией самого Геделя в отношении полемики менталистов и механицистов о природе человеческого разума. Менталисты, в отличие от механицистов, предполагают превосходство человеческого разума над машиной, и само содержание дилеммы Геделя говорит о том, что такой подход к природе человеческого разума стоит у него «на повестке дня», тем более что после опубликования его вышеупомянутой лекции большинство исследователей пришли к мнению, что Гедель является менталистом.

Дизъюнкция Геделя воспринимается им самим как математически установленный факт, который имеет важные философские следствия. Больше того, сам Гедель считает ее переформулировкой его теорем о неполноте. Действительно, можно рассуждать так. Представим дизъюнкцию в другой форме: аксиомы математики схвачены конечным правилом, и не существует абсолютно разрешимых диофантовых проблем. Иными словами, пусть человеческий разум представляет собой конечную машину и не существует абсолютно неразрешимых предложений. Как известно, вместо конечной машины мы можем говорить о машине Тьюринга. Тогда предположение принимает следующий вид: человеческий разум есть машина Тьюринга, и не существует абсолютно неразрешимых диофантовых проблем. Но это утверждение ложно, потому что согласно теоремам о неполноте, для каждой машины Тьюринга есть абсолютно неразрешимое предложение. Признание ложности предыдущего утверждения и представляет собой дизъюнкцию Геделя: человеческий разум превосходит машину. Характер дизъюнкции как установленного математического факта состоит в том, что нельзя отрицать оба члена дизъюнкции.

Но здесь есть неявное предположение, что это исключающая дизъюнкция, т.е. истинным может быть только один ее член. Сам Гедель, судя по всему, придерживался именно такой точки зрения и явно отдавал предпочтение менталистской позиции, состоящей в том, что разум превосходит конечную машину. Однако возможна и такая интерпретация «установленного математического факта», когда оба члена дизъюнкции истинны. Правда, Гедель в примечании не исключил эту возможность, причем заметил, что теоретически следует говорить о трех возможностях. Каковы они?

Во-первых, человеческий разум есть конечная машина, и для него существуют абсолютно неразрешимые диофантовы проблемы.

Во-вторых, человеческий разум бесконечно превосходит конечную машину, и существуют абсолютно неразрешимые диофантовы проблемы.

В-третьих, человеческий разум бесконечно превосходит конечную машину, и не существует абсолютно неразрешимых диофантовых проблем.

Из ранее сказанного следует, что Гедель как менталист предпочитал последний тезис, что подтверждается другими его заметками.

Сама по себе формулировка дилеммы обязана в первую очередь понятию незавершаемости математики. Существование неразрешимого утверждения для любой формальной системы позволяет построить новую формальную систему путем добавления к прежней истинного неразрешимого утверждения в качестве аксиомы. Действительно, теорема Геделя о неполноте (первая) является конструктивной по своему характеру. Если формальная система Φ, содержащая арифметику, ω-непротиворечива, тогда можно эффективно найти предложение G, такое что ни G, ни $\neg G$ не являются теоремами Φ. Далее, если каждая арифметическая теорема Φ истинна, тогда истинно и G. Но в этом случае можно добавить G в качестве новой аксиомы к аксиомам Φ, получая при этом новую систему Φ_1. Для этой новой системы опять-таки эффективно можно получить предложение G_1. И так далее. Таким образом, мы получаем расширение арифметики, и, как мы увидим позднее, такое расширение оказывается бесконечным. В этой связи Гедель говорит о «незавершаемости или неисчерпаемости математики». Вместо ограниченной арифметики мы получаем незавершаемую арифметику.

Принципиальным в философском отношении является вопрос о том, может ли этот процесс получения незавершаемой математики осуществляться конечной машиной или же только человеком. Если это доступно лишь человеку, тогда он действительно превосходит по своим возможностям конечную машину.

Как видно из приведенного выше перечня возникающих в связи с дилеммой вопросов, мы имеем дело с очень глубокой постановкой проблем, имеющих как сугубо философский, так и технический характер. В данной статье мы рассмотрим проблемы, связанные с понятием незавершаемой математики. Но перед обсуждением этих проблем следует уточнить, что понимается под эффективной процедурой и, кроме того, какова связь между понятиями, относящимися к диофантовой проблеме, и непротиворечивостью формальной системы.

Прежде всего, необходимы понятия эффективной процедуры и, соответственно, эффективно заданной формальной системы [4]. Стандартными примерами таких систем являются арифметика Пеано (AP) и теория множеств Цермело – Френкеля (ZF). Как известно, в доказательстве теорем Геделя о неполноте используется арифметизация синтаксиса: в формальной системе S отдельным символам, формулам и последовательностям формул присваиваются согласно определенной процедуре числа. В формальной системе доказательство представляет собой последовательность формул. Для данного числа можно эффективно определить, является ли оно геделевым числом формулы формальной системы. То же относится и к доказательству, поскольку оно есть последовательность формул и его последней формулой является доказуемое утверждение, так что можно эффективно найти, является ли число геделевым числом этого утверждения, или теоремы. Перебором всех целых чисел можно удостовериться, являются ли некоторые из них доказательством теоремы в S.

В эффективно представленной формальной системе S можно сконструировать эффективно вычислимое отношение $\mathrm{Proof}_S(x, y)$, где x есть геделево число доказательства в S формулы, геделево число которой есть y. Непротиворечивая система не содержит ложных утверждений, например отрицания того, что $\neg(0 = 0)$. Утверждение о непротиворечивости системы Con_S формулируется в терминах отношения $\mathrm{Proof}_S(x, y)$, а именно, что для любых доказательств в S не может быть того, чтобы последней формулой было ложное утверждение. В формальном виде Con_S есть $\forall x \neg \mathrm{Proof}_S(x, \neg(0 = 0))$. Теперь рассмотрим пошагово, как получается вторая теорема о неполноте, которая и является основой философских заключений о природе математических утверждений.

1. Гедель показал, что если отношение Proof $_S$ примитивно рекурсивно, оно определяется в языке арифметики Пеано.

2. С.К. Клини показал, что утверждения формы $\forall x\ Px$, если P вычислимо по Тьюрингу, может быть эффективно выражено в форме $\forall x\ Rx$, где R примитивно рекурсивно.

3. Отсюда следует, что $\forall x\ Px$ может быть определимо в языке арифметики.

4. Отсюда следует, что Con_S может быть сконструировано как утверждение языка арифметики.

В этой терминологии и определениях вторая теорема Геделя о неполноте формулируется так:

Если S есть эффективно заданная формальная система, которая содержит РА, и S непротиворечива, тогда Con_S недоказуема в S.

Проблема состоит в том, что вопрос о непротиворечивости формальной системы является вполне законным вопросом, который должен иметь в качестве ответа определенный вариант: система либо непротиворечива, либо противоречива. Поэтому требуется понимание того, как вторая теорема о неполноте соотносится с объективностью утверждения о непротиворечивости. Сам Гедель соотнес этот вопрос с проблемой диофантовых уравнений, т.е. целочисленных решений полиномиальных уравнений с целочисленными коэффициентами. Такой ход мысли Геделя был связан с доказательством им того, что каждое утверждение формы $\forall xRx$, где R примитивно рекурсивно, эквивалентно

$$\forall x_1 \dots x_n\, \exists y_1 \dots y_m [p(x_1 \dots x_n, y_1 \dots y_m) = 0].$$

Другими словами, Con_S может быть приведено к форме диофантова утверждения.

Известно, что любое вычисление может быть закодировано как полиномиальное. То есть для каждой машины Тьюринга существует эквивалентное диофантово уравнение, и свойства решения этого уравнения в точности отражают вычислительные возможности соответствующей машины Тьюринга. Таким образом, формальные системы и диофантовы уравнения на абстрактном уровне идентичны.

Именно это объясняет, почему Гедель говорит об абсолютно неразрешимых диофантовых уравнениях. В дальнейшем мы можем просто говорить о неразрешимых проблемах [5].

Так как же интерпретировать возможность непрерывного расширения системы за счет добавления неразрешимого предложения к исходной системе? Этот вопрос частично фигурирует в Возражении 19 у Р. Пенроуза. Он полагает, что при осмыслении этого результата приходится прибегать к пониманию, которое является прерогативой человека и недоступно машине. «…Процесс продолжается до еще больших ординалов, пока мы все еще способны на каждом последующем этапе понять, как систематизировать все множество геделизаций, которые получили на данный момент. В этом-то и заключается основная проблема: для упомянутых нами “усилий, трудов и напряжений” требуется соответствующее понимание того, как надо систематизировать предыдущие геделизации. Эта систематизация выполнима при условии, что достигаемый к каждому последующему моменту этап будет помечаться так называемым рекурсивным ординалом, что, в сущности, означает, что должен существовать определенный алгоритм, способный такую процедуру генерировать. Однако алгоритмической процедуры, которую можно было бы заложить заранее и которая позволила бы выполнить описанную систематизацию для всех рекурсивных ординалов раз и навсегда, просто-напросто не существует. Нам снова придется прибегать к пониманию [6].

Впервые именно А. Тьюринг показал, что любое Π_1-истинное высказывание можно в некотором смысле доказать с помощью многократной геделизации. Но воспользоваться этим для получения механической процедуры установления Π_1-высказываний нам не удастся по той причине, что механизировать геделизацию нельзя. Но почему нельзя механизировать процесс получения все новых геделевых предложений?

Дж. Лукас полагает, что ситуация с трансфинитным расширением системы говорит в пользу менталистов [7]. Действительно, для механициста вполне естественно включить геделево предложение в машину с другим геделевым предложением, но, конечно, это делает машину другой машиной со своим геделевым предложением, которое машина не может произвести как истинное, а разум может. Поэтому механицист пытается добавлять геделев оператор, который дает всю счетную бесконечность геделевых предложений, что может быть побито новой машиной, включающей в себя геделев оператор. В сущности, это движение от ω, бесконечной последовательности геделевых предложений, производимых геделевым оператором, к $\omega+1$, к следующему трансфинитному ординалу. И по ходу того, как этот процесс продолжается, механицист теряет терпение и вводит новые операторы, пока в ходе процесса не получаем новый ординал, и т.д. Но такие ординалы, не имея предшественников, имеют последователей, точно так же как и любой другой ординал, и разум может «отгеделить» их путем получения нового геделева предложения и считать его истинным, чего не может сделать машина.

Однако ситуация не столь однозначна, как это кажется Лукасу. Дело в том, что каждый новый шаг требует нового имени, – другими словами, мы должны снабдить каждый ординал каким-то именем. В нашей иерархии ω, ω^{ω}, $\omega^{\omega\omega}$ и т.д. мы достигаем ординальное число ε_0. При дальнейшем движении требуются новые имена, порождение которых может осуществляться компьютерной программой. Все рассматриваемые ординальные числа являются конструктивными ординальными числами. Тем не менее их порождение не есть алгоритмический процесс. Согласно теореме Клини – Черча, не существует рекурсивного способа поименования конструктивных ординалов [8]. Причину этого можно усмотреть в том, что при переходе к новой серии ординальных чисел совершается скачок, т.е. в действие вступает нерегулярность. Каждый раз совершается новый креативный шаг, когда мы рассматриваем все до сих пор поименованные ординальные числа, и нам надо собрать их в единое множество, которое может быть использовано для определения нового вида ординала, превосходящего их все.

Д. Хофштадтер поясняет возникающую тут ситуацию следующим образом [9]: когда мы переходим к схеме ω, ω^{ω}, $\omega^{\omega\omega}$, нам нужно новое имя для этого нового шага, т.е. тут возникает нерегулярность. Таким образом, нет алгоритма для получения этого нового имени для следующего ординала. Таким именем является ε_0. Теперь можно полагать, что нерегулярности подобного рода в прогрессии от ординала к ординалу могут быть рассчитаны компьютерной программой. То есть должна быть программа воспроизведения новых имен регулярным образом, и при достижении определенного предела эта программа должна обеспечивать новое имя ординала. Но такая программа не работает. Оказывается, что сами нерегулярности случаются нерегулярным образом, и по этой причине требуется программа второго порядка, т.е. программа, которая делает новые программы для получения новых имен ординалов. Но и этого оказывается недостаточно: необходимой становится программа третьего порядка и т.д., и т.д.

Интуитивная идея, выраженная в формулировке теоремы Клини – Черча, состоит в том, что по мере того как ординалы становятся все больше и больше, возникают нерегулярности, нерегулярности в нерегулярностях, нерегулярности в нерегулярностях в нерегулярностях и т.д. Никакая единая схема, как бы сложна она ни была, не может поименовать все ординалы. Из этого следует, что никакие алгоритмические методы не могут указать, как применять метод Геделя ко всем возможным видам формальных систем. И если не говорить о мистических позициях, тогда следует заключить, что любой человек на некотором этапе просто достигает пределы своей способности к геделизации. Начиная с этого момента формальные системы такой сложности, хотя и неполные по геделевым причинам, будут иметь такую же силу, как и человек.

Следовательно, мы не можем программировать машину так, чтобы произвести имена для всех ординальных чисел. Д. Хофштадтер полагает, что разум может исчерпаться и будет не в состоянии придумывать новые ординалы, и по этой причине невозможно установить искомое менталистом различие между разумом и машиной.

Как видно, сопоставление человека и машины в случае итерирования процесса добавления геделевых предложений к формальной системе вплоть до трансфинитных ординальных чисел представляет значительные трудности для философской интерпретации. Дж. МакКарти продемонстрировал диалектический характер этой проблемы в виде диалога менталиста и механициста [10]:

Менталист: Какую бы формальную аксиоматизацию ни использовала машина, теорема Геделя показывает, как сконструировать в рамках этой аксиоматизации предложение, которое является истинным в предположении обоснованности аксиоматизации, но которое не может быть доказано в этой аксиоматизации.

Механицист: Да, но конструирование этого предложения может быть частью программы, которую машина может присоединить к предыдущей программе для получения новой.

Менталист: Этот процесс может быть итерирован вплоть до трансфинитных ординальных чисел, а ординальные числа, которые использует машина, имеют верхнюю границу. Человек может в принципе определить эту границу путем анализа программы машины.

Механицист: Итерирование процесса конструирования геделевых предложений по ординальным числам требует нотации для ординальных чисел. Это нотация для вычислимых предикатов, но необходимо установить, чтобы вычисление действительно производило тотальное вполне-упорядочение. Таким образом, мы должны рассмотреть доказуемо рекурсивные ординалы. Затем мы должны спросить, какая аксиоматическая система должна быть использована для этих доказательств. Больше того, новые аксиоматические системы, получаемые такой итерацией, зависят от нотации, а не просто от ординального числа, которое определяет нотация.

Очевидно, что точку в этом споре поставить трудно. Но остается впечатление, что возможности человеческого мышления вряд ли зависят от такого рода математических идеализаций, как итерирование операций вплоть до ε_0. Вообще, подобные математические идеализации вряд ли подходят для описания того, как может выглядеть алгоритмизация человеческого мышления, даже в случае математического размышления. Исключение может составлять лишь теория рекурсивных функций, которая в этой связи занимает особое место.

Итак, с одной стороны, можно считать, что оператор геделизации будет следовать человеческой способности производить геделевы предложения посредством рекурсивно поименованной бесконечности. С другой стороны, как отмечает Дж. Уэбб, «нет ни малейшей причины полагать… что машина не сможет смоделировать “изобретательность” ума в этом отношении так далеко, насколько это вообще возможно» [11].

Однако следует повторить, что весь этот вопрос представляется весьма спекулятивным, потому что на уровне ординального числа ε_0 оценка человеческих возможностей столь же спекулятивна, как и оценка компьютерного «мышления». Гораздо более интересным является вопрос о саморефлексии как человека, так и машины. Если утверждать, что существует преимущество человека над машиной, тогда надо признать, что разум всегда способен к рефлексии над рефлексивной формальной моделью. Но тогда способность осуществлять саморефлексию важнее, чем то, кто побеждает в «трансфинитной гонке» [12].

По этой причине есть смысл задать более общий вопрос: что, собственно, представляет собой добавление геделева предложения к исходной формальной системе? Геделево предложение недоказуемо, и поэтому его добавление приводит к расширению системы. Это расширение непротиворечиво, поскольку геделево предложение истинно. Таким образом, мы имеем расширение системы, которое в то же время говорит об исходной системе нечто важное: геделево предложение истинно, если формальная система непротиворечива, т.е. если каждое из утверждений формальной системы истинно. На языке метафоры, геделево предложение «одной крови» с предложениями системы. В определенном смысле присоединение геделева предложения к системе есть способ размышления системы о самой себе, или способ саморефлексии.

Такая саморефлексия может получить оформление в специального рода принципе. Этот принцип называется по вполне понятным причинам рефлексивным. Однако следует говорить скорее о рефлексивных принципах, поскольку эти принципы, имея одинаковую форму, обладают различной силой. Надо заранее сказать, что более общая постановка вопроса с принципами рефлексии вместо просто итерации добавления к системе геделевых предложений не слишком проясняет ситуацию в споре менталиста и механициста. Но с другой стороны, это позволяет установить ряд важных моментов в отношении природы истинности геделева предложения и обоснованности формальной системы. Наиболее интересные вещи подобного рода проявляются в связи с трансфинитной итерацией добавления геделева предложения к системе.

Перед тем как выяснять, представляет ли трансфинитная итерация добавления геделева предложения к системе интерес в связи с проблемой алгоритмизации, рассмотрим, чем являются принципы рефлексии.

С. Феферман дает такое определение принципа рефлексии: «Под принципом рефлексии мы понимаем описание процедуры добавления к некоторому множеству аксиом *A* определенных новых аксиом, общезначимость которых следует из общезначимости (valid) аксиом *A* и которые формально выражают в языке *A* очевидные следствия предположения, что все теоремы *A* общезначимы» [13]. Это означает, что принцип рефлексии представляет собой новую аксиому для теории *A*. Такая аксиома должна быть принята, если предполагается истинность каждой теоремы теории *A*. Что мы получаем при таком добавлении? Если мы добавляем тривиальную истину, тогда мы не получим ничего существенного в расширенной системе. Но если добавлять к системе нетривиальные принципы рефлексии, тогда арифметическая истина будет расширяться неопределенно далеко.

Таким нетривиальным расширением арифметики является использование геделева предложения или же предложения о непротиворечивости системы Con_S. Последнее обстоятельство вполне понятно, поскольку в качестве геделева предложения, истинного, но недоказуемого, выступает утверждение о непротиворечивости системы.

Принципы рефлексии выступают в виде схемы аксиом, «которые выражают, насколько это возможно без использования формального понятия истины, что все доказуемое в системе истинно» [14].

В качестве принципа рефлексии естественно принять Con_S. Присоединение этого утверждения в качестве аксиомы к системе дает непротиворечивую систему. Известно, что в ходе этого процесса принципы рефлексии приобретают разную силу в зависимости от формальной системы. Таким образом, в данном случае мы имеем дело с ситуацией, когда некоторая теория S расширяется путем принятия все более сильных принципов рефлексии. Важно понимать, что при этом указанные принципы становятся настолько сильными, что уже не могут рассматриваться как естественные расширения данной теории. Таким образом, мы вновь имеем ситуацию, когда возможности как человека, так и машины, не могут сопоставляться уже потому, что расширение является неестественным. Общая схема такого расширения:

$$X_S \to X_{S+Y_S},$$

где X – сильный, а Y – более слабый принцип.

Неочевидность принципов на определенном этапе расширения теории обусловлена тем, что мы оправдываем более слабый принцип более сильным. Можно ли считать знание, полученное с помощью более сильных принципов рефлексии, основанием для знания, базирующегося на более слабых принципах? С точки зрения строгой концепции «оснований» это невозможно. Но сильные принципы могут быть использованы в эвристических целях, когда дальнейшее знание нельзя вывести с помощью более слабых принципов. Более сильный принцип рефлексии позволяет нам итерировать расширения через добавления утверждений о непротиворечивости. Мораль всего этого состоит в том, что «если мы хотим быть более осторожными, мы не должны принимать априори принцип рефлексии, но мы можем позволить некоторые его следствия, а именно, итерированные непротиворечивости. Ощущение, что мы всегда можем прогрессировать и делать все более сильные предположения, не отражает нашей особой способности, а просто объясняется постепенным уменьшением нашей веры в их истинность. Таким образом, расплывчатость присутствует в нашем математическом мышлении, но не в процессе дедукции, а при решении того, какие аксиомы мы должны принимать [15].

Возвращаясь к вопросу о незавершаемости математики, естественно предположить, что она является результатом трансфинитного расширения системы путем «механического» добавления к ней соответствующего геделева предложения. Но это означает опять-таки применение соответствующего принципа рефлексии. Каким образом возникает идея алгоритмизации этого процесса?

Далее к исходной теории присоединяется ее геделево предложение. Полученная теория должна быть также непротиворечива, поскольку все ее утверждения, включая новое – геделево предложение, истинны. Для этой теории можно также сконструировать соответствующее геделево предложение и путем его добавления к системе получить новую теорию. Это уже знакомый, описанный ранее процесс, существенной чертой которого является обоснованность каждой получаемой теории. На этом этапе возникает еще одна попытка механизации процесса – теперь уже процесса порождения новых теорий. Например, можно построить объединение всех такого рода теорий, получаемых конечным числом шагов. Ясно, что это будет ω-последовательность теорий.

Дальнейшая механизация требует, как мы уже говорили, скачка в последовательности трансфинитных чисел, и с помощью этих скачков строятся все более мощные теории, соответствующие все большим ординальным числам. Такие теории являются в высшей степени сложными, и надо понять, в какой степени утверждения этих теорий истинны. Мы исходим из того, что в подобном процессе получения новых теорий есть два типа «неопровержимо» истинных утверждений – теоремы арифметики Пеано и утверждения типа $G(S)$, т.е. геделевы предложения. Поскольку в процессе порождения новых теорий исходят из указанных двух типов утверждений, существует искушение полагать все получаемые на этом пути теории непротиворечивыми.

Д. МакКаллох считает, что именно в этом моменте мы наталкиваемся в такой аргументации на подводный камень: «Проблема состоит в том, что для использования теоремы Геделя в получении все более мощных теорий наш математик нуждается во все большей формализации своего собственного мышления и затем совершает скачок к заключению, что сама эта формализация непротиворечива (и, следовательно, соответствующее утверждение Геделя непротиворечиво). Однако если математик формализует слишком многое из своего мышления, включая скачки, тогда результирующая теория будет способна формализовать себя и сделает прыжок к заключению, что ее собственное геделево утверждение истинно. Но это заключение немедленно ведет к противоречию. Поэтому либо (1) математик на некотором этапе прекращает формализацию своего мышления (в этом случае совокупность всех фактов, которые он может доказать, будет в аксиоматической теории), либо (2) он формализует все свое мышление и результирующая теория будет противоречивой (она будет способна доказать свою собственную непротиворечивость)» [16].

Представляется, что это очень важное замечание, потому что при трансфинитном счете совершаются скачки при переходе к новым ординальным числам. До сих пор говорилось о проблеме нахождения имен для таких чисел, но сама природа скачков не принималась во внимание. Между тем вряд ли можно считать формализацию знания непроблематичной в условиях трансфинитного счета. Такого рода ограничение надо понимать не как свидетельство превосходства человека над машиной, а как внутреннее ограничение, присущее мышлению как таковому. Самым характерным свойством мышления является его рефлексивный характер, и попытка получить гарантии непротиворечивости и обоснованности путем формализации мышления обречена на неудачу.

Как уже было указано, сам по себе факт добавления геделева предложения для получения непротиворечивой или обоснованной системы представляет собой важное свойство формальной системы, а именно, принцип рефлексии. Требуемый для трансфинитной последовательности теорий принцип рефлексии рассматривает С. Феферман. Это так называемый принцип рефлексии для ω-правила [17]. Как и ранее, пусть Pr есть предикат доказуемости. В квадратных скобках указывается геделево число формулы. Pr $(\ulcorner \Phi x \urcorner)$ есть формула, утверждающая, что результат подстановки подходящего числа для x в Φx доказуем в теории. Упомянутый принцип рефлексии в этих обозначениях будет иметь следующий вид:

$$\forall x \, \mathrm{Pr} \, (\ulcorner \Phi x \urcorner) \rightarrow \forall x \, \Phi x.$$

Коль скоро принцип рефлексии позволяет присоединять к теории истинное, но не доказуемое в ней предложение, возникает вопрос о формальной возможности постоянной итерации принципа рефлексии через рекурсивные ординалы. Далее мы следуем изложению результатов С. Фефермана в обзоре С. Шапиро [18].

Для только что упомянутых принципов рефлексии если $A(\alpha)$ рекурсивно перечислимо, тогда рекурсивно перечислимо и $A(\alpha+1)$ (где α есть некоторое ординальное число). Пусть ординальные числа обозначены натуральными числами. В частности, число 1 обозначает ординальное число 0, и если n обозначает ординальное число α, тогда 2^n обозначает последующее для α ординальное число $\alpha+1$. Пусть e есть геделево число машины Тьюринга, которая перечисляет числа, обозначающие все более увеличивающуюся последовательность ординальных чисел. Пределом этой последовательности будет $3 \cdot 5^e$. Пусть O будет множеством натуральных чисел, которые обозначают ординальные числа, и если $m \in O$, тогда пусть $|m|$ будет ординальным числом, которое обозначается натуральным числом m. Множество O не является рекурсивно перечислимым.

Пусть P будет принципом рефлексии. Пусть $R(1)$ есть стандартное перечисление теорем теории A. Если $n \in O$, тогда пусть $R(2^n)$ будет перечисление результатов применения принципа рефлексии к $R(n)$. Идея, стоящая за всеми этими обозначениями, такова: для каждого $n \in O$, $R(n)$ есть теория $A|n|$. В данном случае R будет тотальной рекурсивной функцией от натуральных чисел. Если $n \in O$ и базис теории А состоит из истин, тогда $R(n)$ содержит только истины.

А. Тьюринг использовал теории типа $R(n)$ в попытке преодолеть неполноту. В этом случае n пробегает над O. Он показал, что для простого принципа рефлексии Con, если Φ есть истинное Π_1-предложение, тогда имеется $n \in O$ (которое может быть эффективно найдено из Φ), такое что $|n| = \omega + 1$ и Φ будет находится среди теорем $R(n)$. Это означает, что есть способ итерировать геделевскую конструкцию над теориями начиная с A, так что когда мы соберем вместе конечные итерации и построим еще одно геделево предложение, Φ будет разрешимо. Таким образом, для Π_1-предложений имеется определенный вид полноты.

С. Феферман расширил этот результат. С более сильным принципом рефлексии A^*, он показал, что для любого истинного предложения Φ в языке арифметики имеется число $n \in O$, такое что Φ находится среди теорем $R(n)$. То есть для любого предложения Φ имеется способ итерации принципа рефлексии (вплоть до малого трансфинитного уровня) и имеется при этом разрешающая процедура для Φ.

Рассмотрим, как это действует «на практике». Пусть имеется некоторое математическое предложение, в истинностном значении которого мы не уверены; скажем, это может быть проблема Гольдбаха. Математик вычисляет число n и начинает перечислять все теоремы $R(n)$, среди которых может быть и теорема Гольдбаха для некоторого n. Это вполне эффективный процесс. Математик знает, что если догадка Гольдбаха истинна, тогда $n \in$ O и поэтому каждое предложение в $R(n)$ истинно. Однако доказательство Тьюринга показывает, что если догадка Гольдбаха ложна, тогда n не есть в O и, больше того, $R(n)$ противоречиво. То есть результатам перечисления можно верить только в том случае, если догадка Гольдбаха истинна. Это никак не помогает определить истинностное значение догадки Гольдбаха.

То же справедливо для результата Фефермана. Для каждого предложения Φ мы получаем формальную систему F, такую что F обоснована, если Φ истинно, и если Φ истинно, в системе F оно доказуемо. Но этот результат можно было получить простым добавлением Φ в качестве новой аксиомы к теории A. В чем тут дело? Феферман так представляет эту ситуацию: «Всякий раз, когда математик имеет информацию, что $d \in O$, он способен вычислить $R(d)$ и доказать теоремы из $R(d)$. Больше того, если он получает информацию, что $d \in O$, он должен посчитать эти теоремы приемлемыми. К несчастью, по мере продвижения все дальше и дальше в совокупность систем $R(d)$ он больше не может получить знание, которое ему необходимо для решения вопроса о том, является ли выражение $d_0 \in O$ справедливым для любого d_0. Другими словами, для решения этого вопроса он может апеллировать к оракулу [19].

Ключевым в данном вопросе является решение о членстве в O, или же нахождение эффективной нотации для рекурсивных ординалов. В этих терминах проблема упирается в перечисление рекурсивных ординалов. С точки зрения менталистов, процесс добавления геделевых предложений (или же принципов рефлексии Фефермана) можно было бы продолжать «достаточно долго». Но при этом процедура нахождения геделевых предложений становится все более сложной, и тогда менталист полагает, что в такой «сложности» человек ориентируется гораздо лучше машины. Опять-таки, предлагаемый менталистами спор устраивается ими «диалектически», в очной дискуссии. «Снова и снова механицист теряет терпение и встраивает в свою машину оператор, предназначенный для того, чтобы единым махом производить все геделевы предложения, которые являются козырем менталиста. Это усилие состоит в произведении нового предельного ординала. Но такие ординалы, хотя они и не имеют предшествующих, имеют последующие, точно так же как и любые другие ординалы, и разум может перегеделить их произведением геделева предложения для новой версии машины и видеть их истинность, чего не может видеть машина [20].

Согласно теореме Клини – Черча, не существует рекурсивного перечисления всех рекурсивных ординальных чисел. Следствием этого является нерекурсивность O. На основании данных результатов утверждаются различные менталистские позиции. Сильный тезис состоит в том, что человек может перечислить члены O. Более слабый тезис состоит в том, что если идеализированный человек есть данная машина M, которая производит имена для рекурсивных ординальных чисел (например, членов O), эта машина может произвести имя ординального числа, не производимого машиной M. В таком случае человек может выполнять принцип рефлексии дальше, чем M, и поэтому человек не может быть машиной. Но этот тезис весьма расплывчат. Отсюда вытекает, что вряд ли человек может перенумеровать все члены O.

Больше того, если человек может это сделать, тогда он становится всеведущим созданием. Действительно, согласно результату Фефермана, если человек может перечислить все члены O, тогда он может определить истинностное значение любого арифметического предложения. Для этого ему надо систематически порождать выходы $R(n)$ для каждого $n \in O$, используя принцип рефлексии Фефермана. Для каждого арифметического предложения Φ будет либо Φ, либо $\neg\Phi$. Таким образом, если человек может перенумеровать O, он может непогрешимо определять истинностное значение любого арифметического предложения.

Сам Гедель склонялся к такому арифметическому всеведению, отрицая второй член своей знаменитой дизъюнкции, а именно, что существуют абсолютно неразрешимые арифметические предложения. Таким образом, он полагал, что человек не есть машина. Такое всеведение Гедель назвал «рациональным оптимизмом».

Суть спора между менталистами и механицистами состоит, с точки зрения Дж. Лукаса, в соревновании машины и человека: «Все трудности ложатся бременем на механициста, старающегося изобрести машину, которую невозможно перегеделить. Именно механицист прибегает к уловке с предельными ординальными числами и сталкивается с проблемами изобретения для них новой нотации. Человеческому разуму требуется лишь сделать следующий, всегда легкий и непроблематичный шаг и перегеделить все, что предлагает механицист» [21].

Предположим, что механицист производит число e, которое есть код машины Тьюринга, перечисляющей члены O (с обозначенными ординалами в порядке возрастания). Затем вычисляется $3 \cdot 5^e$, которое обозначает следующий ординал, больший чем любой перечисленный машиной M. Но в этом вычислении нет ничего немеханического, точно так же как нет ничего немеханического в выписывании геделева предложения. Лукас также прав в том, что «следующий» ординал – всегда легкий, непроблематичный шаг. Перейти от $| n |$ к $| n+1 |$ – это просто вычислить 2^n.

Лукас полагает, что человек имеет способность к порождению ординальных чисел, которой нет у машины. Так в чем состоит эта способность? Для данной машины *M* Лукас не может сделать «легкого, непроблематичного шага» (достижения знания нового члена *O*) до тех пор, пока он не узнает, что *M* производит только члены *O*. А откуда у Лукаса такое априорное знание? Как всегда, он перекладывает бремя на механициста. Если механицист докажет, что *M* производит только члены *O*, тогда механицист выйдет за пределы *M*.

Но как Лукас переходит от условных предложений к заключениям? Он провозглашает существование некоторого рода озарения, недоступного для машины Тьюринга. Он говорит, что в дуэли с механицистом для рассмотрения всех до сих пор поименованных ординальных чисел требуется креативный шаг. Но что это за креативный шаг? Ясно, что он должен заключаться не в применении алгоритма, т.е. не в конструировании геделева предложения или вычислении нового члена *O*. И как тогда могут помочь в этом теоремы о неполноте?

Лукас говорит, что способность человека распознавать новые ординалы превосходит способность любого формального алгоритма делать это. Ключевое слово здесь – «распознавать». Речь идет об эпистемической способности побеждать механизм, а это нечто большее, чем просто выписать геделево предложение. «Креативный шаг» – это способность видеть, что каждое число в данной последовательности чисел есть член *O* (или обозначает рекурсивный ординал), или способность видеть, что каждая теорема данной формальной теории истинна. Пенроуз в этой связи говорит о человеческой способности к пониманию, которой лишено всякое алгоритмическое устройство.

Человек в отличие от машины способен «передвигаться» в абстрактных понятиях. Хотя Крайзель и Гедель были согласны с тем, что человек – это не машина, они все-таки полагали, что бремя лежит на человеке, который еще должен доказать, что превосходит машину. Но для того чтобы говорить о применении теорем о неполноте, нам потребуется разговор об алгоритмах, а не о креативных шагах.

Интересно сравнить рациональный оптимизм Геделя как техническую концепцию с его общими представлениями о роли рационализма в понимании мира. Хао Ван, имевший тесные отношения с Геделем, записал многие его мысли. Так, в книге «Логическое путешествие: от Геделя к философии» он приводит 14 пунктов, касающихся философских убеждений Геделя [22].

1. Мир рационален.
2. Человеческий разум может в принципе быть развит до более высокого состояния (посредством определенной техники).
3. Есть систематические методы решения всех проблем (включая искусство и др.).
4. Есть другие миры и рациональные существа различных и более высоких видов.
5. Мир, в котором мы живем, является не единственным, в котором мы будем жить или уже жили.
6. Имеется гораздо больше априорного знания, чем известно до сих пор.
7. Развитие человеческой мысли со времени Возрождения является полностью постижимым (durchaus einsichtige).
8. Человеческий разум будет развиваться в различных направлениях.
9. Формальные методы охватывают реальную науку.
10. Материализм ложен.
11. Более высокие существа связаны с другими по аналогии, а не строением.
12. Концепции имеют объективное существование.
13. Существуют научная (точная) философия и теология, которые имеют дело с концепциями высшей абстрактности, и они также очень полезны для науки.
14. Религии, по большей части, плохи, но религия – нет.

Тут же Хао Ван добавляет: «Перечень из 14 пунктов претендует на обрисовку фундаментальных философских представлений Геделя. В нем есть оптимистические представления и догадки. Они далеко выходят за пределы того, “что возможно до того, как появятся новые открытия и изобретения”, как того требовал от философии Витгенштейн. К несчастью, мы очень мало знаем о том, почему Гедель придерживался этих взглядов. Безусловно, центральным является его убеждение в рациональности мира. Это ключевое убеждение представляет собой эмпирическое обобщение его интерпретации человеческого опыта, но то, что мы знаем о его аргументации в этом отношении, вряд ли убедительно» [23].

Как видно, утверждение дилеммы Геделя, которое он считал математически определенным фактом, включает в себя множество предпосылок, носящих откровенно философский характер. Но сам Гедель не считал это пороком или каким-то странным обстоятельством. В этом отношении очень интересным является свидетельство Г. Крайзеля: «…Гедель подчеркивал… как мало ему потребовалось новых математических построений. Оказалось, что надо было только обратить внимание на некоторые достаточно широко известные (философские) различия… Совершенно не испытывая угрызений совести по поводу того, что он, так сказать, “поживился на дармовщинку”, Гедель видел в своих первых успехах реализацию следующей, часто забываемой, но плодотворной общей схемы. Внимательно рассматривая подходящие традиционные философские концепции и вопросы, анализируя их, и, возможно, добавляя чуть-чуть точности, мы безболезненно приходим к нужным понятиям, правильным гипотезам и достаточно простым доказательствам» [24].

С конца 1940-х годов Гедель стал обращаться к философским концепциям особенно часто. Дилемма Геделя, описанная выше, подняла значительное число философских проблем, среди которых различение им субъективной и объективной математики является чрезвычайно философски интересным и важным следствием этой дилеммы.

Примечания

1. *Godel K.* Some basic theorems on the foundations of mathematics and their implications // Godel K. Collected Works. V. III: Unpublished Essays and Lectures / Ed. by S. Feferman et al. – N.Y.: Oxford Univ. Press, 1995. – P. 304–323.

2. Ibil. – P. 310.

3. См.: *Целищев В.В.* Алгоритмизация мышления: Геделевский аргумент. – Новосибирск: Параллель, 2005.

4. См.: *Целищев В.В*. Тезис Черча. – Новосибирск: Параллель, 2008.

5. См.: *Feferman S*. Are there absolutely unsolvable problems? Godel’s dichotomy // Philosophia Mathematica. – 2006. – V. 14, No. 2. – P. 134–152.

6. *Пенроуз Р.* Тени разума. – М., 2003. – С. 187.

7. См.: *Lukas J.* Lukas against mechanism: a rejoinder // Etica & Politica. – 2003. – No. 1.

8. См.: *Kleene S.C.* On notation for ordinal numbers // Journal of Symbolic Logic. – 1938. – V. 3. – P. 150–155.

9. См.: *Hofstadter D*. Godel, Escher, Bach. – Harvester Press, 1979 – P. 475–476.

10. См.: *McCarty J.* Awarness and understanding in computer programms // Psyche. – 1995. – V. 2, No. 11.

11. *Webb J.* Mechanism, Mentalism and Metamathematics. –Dordrecht, Holland: D. Reidel, 1980. – P. 81.

12. См.: *Boajiev D.* Mind versus Godel // Mind versus Computer / Ed. by Gams et al. – IOS Press, 1997. – P. 202–210.

13. *Feferman S.* Transfinite recursive progressions of axiomatic theories // Journal of Symbolic Logic. – 1962. – V. 27. – P. 274.

14. *Feferman S.* Reflecting on incompleteness // Journal of Symbolic Logic. – 1991. – V. 56. – P. 12–13.

15. *Pudlak P.* A note on applicability of incompleteness theorem to human mind // Annals of Pure and Applied Logic. – 1999. – V. 99. – P. 9.

16. *McCullough D.* Can humans escape Godel? // Psyche. – 1995. – V. 2, No. 4.

17. См.: *Feferman S.* Transfinite recursive progressions of axiomatic theories. – P. 259–316.

18. См.: *Shapiro S.* Incompleteness, Mechanism, and Optimism // Bulletin of Symbolic Logic. – 1998. – V. 4, No. 3. – P. 273–302.

19. *Feferman S.* Transfinite recursive progressions of axiomatic theories. – C. 262.

20. *Lukas J.* Minds, Machines, and Godel: A retrospect // Machines and Thought: The Legacy of Alan Turing / Ed. by J. Milican, A. Clark. – Oxford Univ. Press, 1996. – V. 1. – P. 110.

21. Ibid. – P. 113.

22. См.: *Hao Wang*. A Logical Journey: from Godel to Philosophy. –Cambridge: MIT Press, 1996. – P. 316.

23. Ibid.

24. *Крайзель Г.* Биография Курта Геделя. – М., 2003. – С. 7.

Дата поступления 11 февраля 2013 г.

Институт философии и права
СО РАН, г. Новосибирск

director@philosophy.nsc.ru

Tselishchev, V.V. **Incompleteness of mathematics and absolutely unsolvable problems**

The paper deals with one of the aspects of Godel's dilemma of the man-computer relation and the existence of absolutely unsolvable predicates in mathematics. The author discusses the basic premise of the dilemma, viz. the concept of incompleteness of mathematics as a corollary of incompleteness theorems. He concludes that Godel's argument of man's superiority over a finite machine is incomplete as far as solvability of mathematical predicates is concerned.

Keywords: Godel, mathematics, computer, man, dilemma