

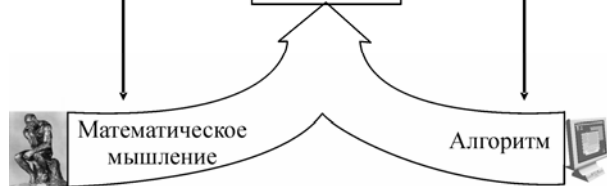
**АЛГОРИТМИЗАЦИЯ ЗНАНИЯ И ЕЕ ПРЕДЕЛЫ (ИССЛЕДОВАНИЕ ПРЕДСТАВЛЕНИЯ  
МЫШЛЕНИЯ В ЛОГИКО-МАТЕМАТИЧЕСКИХ ЯЗЫКАХ ПЕРВОГО  
И ВТОРОГО ПОРЯДКОВ, ВЫЧИСЛИМОСТИ В СВЕТЕ ТЕЗИСА ЧЕРЧА  
И СФЕРЫ ДЕЙСТВИЯ ОГРАНИЧИТЕЛЬНЫХ ТЕОРЕМ ГЕДЕЛЯ И ТАРСКОГО).  
ПРОЕКТ № 125**

**Координаторы:** акад. Ершов Ю. Л., д-р филос. наук Целищев В. В.  
**Исполнители:** ИМ, ИФПР ОИИФФ СО РАН

Конструирование систем искусственного интеллекта определяется в значительной степени, помимо развития технологических информационных технологий, разработкой математических методов построения алгоритмов. Предполагаемое превосходство человеческого мышления над искусственным интеллектом заключается в интуитивной способности человека к «постижению» истин; компьютеру в такой способности отказывается по самым разнообразным причинам, начиная с рудиментарной на сегодняшний день сложности компьютера по сравнению с человеческим мозгом и кончая глубинными философскими препятствиями, обоснованными феноменологией и герменевтикой, в частности, вопросом о том, существует ли разум отдельно от материи и функционирует ли мозг как машина. Поставленная в проекте проблема состоит в том,

можно ли указанное превосходство человеческого мышления доказать строго, опираясь на фундаментальные логико-математические результаты в области моделирования когнитивных процессов. В качестве такого результата рассмотрены теоремы Геделя о неполноте арифметики и эквивалентные им результаты А. Тьюринга. Понятие машины Тьюринга и ассоциированные с ним концепции стали основой не всегда четкой философской дискуссии о соотношении человеческого ума и машинного «интеллекта».

Рассмотрена такая интерпретация упомянутых математических результатов, которая имеет прямое отношение к вопросу о том, можно ли сконструировать вычислительную машину, которая была бы равна по своим возможностям человеческому разуму. Различные ответы на вопрос о возможности алгоритмизации человеческого мышления проанализированы с точки зрения двух антагонистических философских направлений — «ментализма», согласно которому такая алгоритмизация в принципе невозможна, и «механицизма», который полагает деятельность мозга алгоритмическим процессом. Драматическая коллизия двух направлений представлена дилеммой Геделя, напрямую затрагивающей статус математического знания:



Алгоритмизация мышления. Геделевский аргумент.

Algorithmization of reasoning. Godelian argument.

1) человеческий ум не способен к формулировке (или механизации) всех математических интуиций; если они сформулированы, появляются новые, например, о непротиворечивости;

2) либо человеческий ум превосходит все машины (более точно, может решить больше

теоретико-числовых проблем), либо существуют такие теоретико-числовые проблемы, которые неразрешимы для человеческого ума.

Показано, что разрешение дилеммы ведет к различению двух видов математических утверждений — объективных и субъективных. Объективные математические утверждения следует понимать как описание фактов платонистской реальности, а субъективные — доказуемые математические утверждения. Наличие неразрешимых утверждений, если их понимать в абсолютном смысле, и есть свидетельство в пользу такого разделения математических утверждений. Класс субъективных утверждений есть совокупность всех математических теорем, которые могут быть известны с неопровержимой математической определенностью, или же множество познаваемых либо доказуемых арифметических утверждений. Это множество имеет определенные арифметические и вычислительные свойства. Согласно тезису Черча, все арифметические познавательные процедуры или же процедуры доказательства являются эффективными алгоритмами. Механицизм предполагает, что имеется машина Тьюринга, которая перенумеровывает члены этого класса, что означает его рекурсивную перечислимость. Ментализм предполагает, что человеческое мышление не может быть симулировано машиной Тьюринга.

Предположение о том, что процесс приобретения математического знания происходит по некоторому эффективному правилу, т. е. что человеческая способность к доказательству математических утверждений с математической определенностью является механическим процессом, не дает никакой гарантии того, что получаемые при этом результаты будут непротиворечивыми. Дело в том, что если ум уподобляется машине, т. е. представляет собой некоторую формальную систему с эффективными правилами вывода, тогда он не может доказать свою собственную непротиворечивость. Между тем, предложение о его собственной непротиворечивости является чисто математическим предложением, и, согласно второй теореме Геделя о неполноте, доказательство внутри системы ее собственной непротиворечивости невозможно. Таким образом, ум как машина не может полностью

понять своего собственного функционирования.

Менталистское различие между человеческим умом и человеческим мозгом встречается с фундаментальным затруднением, связанным с вынужденным различием ума и мозга. Когда говорится, что человек доказывает теоремы с математической определенностью, что имеется в виду — ум или мозг? Вопрос не только терминологический, поскольку, если эта способность есть результат работы некоторого устройства в мозге, некоторого аналога компьютера как физической машины, тогда первая часть геделевской дилеммы утверждает, что ум превосходит мозг. Это весьма странная точка зрения — человеческий ум превосходит тот самый орган, который является местом его «нахождения». Вторая часть дилеммы утверждает, что человеческий ум, будучи приравнен к машине, не способен постичь некоторые истины математики, которые, таким образом, объявляются абсолютно неразрешимыми предложениями. Другими словами, чистая математика не является полностью продуктом человеческого ума.

Вопрос о принципиальной возможности алгоритмизации математического мышления анализировался через противопоставление семантических и синтаксических характеристик процесса математического мышления. В качестве «решающего эксперимента» была принята проблема понимания человеком и «понимания» компьютером истинности геделевского предложения, не доказуемого в формальной системе (синтаксический аспект), но истинного при стандартной интерпретации формализма (семантический аспект). Исходная проблемная ситуация состояла в предположении, что человек может постичь истинность геделева предложения, в то время как компьютер, ограниченный синтаксическими или алгоритмическими процедурами, не может сделать этого. Анализ проблемы включал следующие исследования по соотношению синтаксиса и семантики формальных систем математики, включая формализованную теорию истины: трансфинитное расширение формальной системы  $F$  путем добавления геделева предложения  $G$  к исходной системе в качестве аксиомы; сопоставление методов получения истинности

и доказуемости в алгоритмической системе; проверка гипотез о наличии неосознанных алгоритмов в человеческом сознании в области математического мышления; анализ концепции истинности геделева предложения в контексте дефляционной теории истины. Выработка критериев «человеческой математики» в сопоставлении с «машинной математикой».

Результаты исследования показали, что теоремы Геделя о неполноте как наиболее доказательный и фундаментальный результат в области математической логики и моделирования мышления не дают оснований для вывода о принципиальных ограничениях в отношении систем искусственного интеллекта по сравнению с человеческим сознанием.

### Основные публикации

1. *Ершов Ю. Л.* Алгебра и логика: старые и новые связи// *Философия науки*. 2004. Т. 4, № 23. С. 132—142.
2. *Целищев В. В.* Нормативность дедуктивного дискурса. Новосибирск: Нонпарель, 2004. 340 с.
3. *Целищев В. В.* Алгоритмизация мышления: геделевский аргумент. Новосибирск: Параллель, 2005. 304 с.
4. *Самохвалов К. Ф.* В защиту предиката существования// *Философия науки*. 2005. Т. 3, № 26. С. 141—146.