Упрощённое доказательство теоремы Тарского о разрешимости элементарной теории поля действительных чисел*

Алексей Семенов† Андрей Мучник‡

Вопрос о существовании алгоритма, позволяющего среди всех предложений какого-нибудь формального языка выделять истинные, является одним из центральных вопросов математической логики с самого её возникновения в начале XX века. Уже в 10-е и 20-е годы были построены разрешающие алгоритмы для важных логических теорий. В дальнейшем наиболее значительными достижениями в этой области явились теорема Тарского о разрешимости элементарной теории поля действительных чисел и теорема Рабина о разрешимости монадической теории бинарного дерева.

Естествен практический интерес к построению разрешающих алгоритмов. Были предприняты попытки запрограммировать эти алгоритмы на ЭВМ. Это закономерно, так как логические языки являются гибким и естественным средством формализации самых различных задач, не сводящихся к прямым числовым вычислениям. Однако указанные попытки привели к программам, работающим слишком долго даже на очень коротких примерах. Это явление оказалось не случайным, в 70-е годы было обнаружено, что для задач разрешения обычно рассматриваемых логических теорий в принципе не может существовать алгоритмов, недолго работающих на всех исходных данных.

*Институт новых технологий, 109004, Москва, ул. Нижняя Радищевская, 10. Работа выполнена при частичной финансовой поддержке Совета поддержки научных школ при президенте РФ и гранта N 01-01-00505 Российского Фонда Фундаментальных Исследований.

†E-mail: alsemenov@mtu-net.ru, fax: (095)915-69-63.

‡E-mail: muchnik@lpcs.math.msu.ru, fax: (095)915-69-63.

Тем не менее нам кажется, что в этой области имеются некоторые основания для оптимистической точки зрения. Дело в том, что при практическом использовании алгоритма его и не нужно будет применять ко *всем* исходным данным. Исходные данные всегда будут браться из некоторого „осмысленного“ класса, на котором алгоритм может работать приемлемое время.

В связи с этим нам представляется важной задача отыскания новых и доработки старых алгоритмов разрешения логических теорий с тем, чтобы они приобретали наиболее ясный вид, даже если асимптотическая оценка сложности их работы хуже, чем у алгоритмов со сложным описанием. На основе анализа этих алгоритмов и отдельных входящих в их состав процедур, возможно, удастся создать и практически работающие программы.

Дадим необходимые определения.

Формулы элементарного языка строятся индуктивно:

1. Атомная формула — формула. (Для поля действительных чисел атомные формулы имеют вид $\tau < \rho$, $\tau > \rho$, $\tau = \rho$, где τ, ρ — термы. Термы — это многочлены от нескольких переменных с рациональными коэффициентами.)

2. Если A, B — формулы, то $(A \vee B)$, $(A \wedge B)$, $(A \to B)$, $\neg A$, $\exists x A$, $\forall x A$ — формулы.

Элементарной теорией некоторой структуры называется множество всех формул элементарного языка, истинных в данной структуре.

Наиболее широко применимым видом разрешающих алгоритмов являются алгоритмы элиминации кванторов. Поясним неформально идею элиминации. Естественная идея, возникающая при построении алгоритма разрешения логической теории, состоит в том, что работа такого алгоритма должна определяться индукцией по построению формулы. Попытаемся реализовать эту идею. Пусть, например, A и B — предложения, тогда истинность предложения $(A \vee B)$ тривиально определяется истинностью A, B. Таким образом, если алгоритм уже нашёл истинностные значения A и B, то не составляет труда определить истинностное значение $(A \vee B)$. Трудность возникает при рассмотрении формулы $\exists x A$. Чтобы проверить её истинность нужно рассмотреть, вообще говоря, бесконечную дизъюнкцию подстановок в A всевозможных конкретных значений x. Формула A может содержать свободную переменную x.

При дальнейшем разложении формулы на подформулы могут появиться другие свободные переменные. Иногда формуле $\exists x A$, где A может содержать помимо x и другие свободные переменные, удаётся поставить в соответствие эквивалентную ей бескванторную формулу. Эта операция и называется *элиминацией квантора*. Конечно, её можно осуществить не всегда. Но когда она возможна для всех формул и осуществляется алгоритмически, а сама структура конструктивна, то мы получаем алгоритм разрешения элементарной теории. Если формулам A и B уже сопоставлены эквивалентные им бескванторные формулы, то формуле $(A \vee B)$ сопоставлятся дизъюнкция этих бескванторных формул. Аналогично поступаем с остальными пропозициональными связками. Заметим, что специальной конструкции для элиминации квантора $\forall$ не нужно, так как $\forall x A$ эквивалентно $\neg\exists x \neg A$.

Результат об элиминации кванторов в элементарной теории действительных чисел со сложением, умножением и порядком принадлежит А. Тарскому ([1]). При этом время работы алгоритма, разрешающего указанную теорию, было неэлементарным. Таким же было время работы алгоритма, применяющего методы алгебраической геометрии, который был предложен в [2]. В дальнейшем был найден более сложный (в смысле описания) алгоритм, работающий на экспоненциальной зоне ([5]). То, что всякий алгоритм разрешения рассматриваемой теории на бесконечно многих входах работает экспоненциальное время, было доказано в [4]. Целью нашей статьи является максимальное упрощение конструкции алгоритма и обоснования её корректности. С точки же зрения сложности вычисления предлагаемая разрешающая процедура не экономнее, чем была у Тарского. Сообщение о нашей работе появилось в [6]. Простейшим из ранее известных было доказательство П. Коэна из [3].

Прежде чем описывать процедуру элиминации, введём некоторые дополнительные определения.

Фиксируем переменную x и определим четыре операции над термами нашей теории, рассматриваемыми как многочлены от x, коэффициенты которых — многочлены от остальных переменных. Эти операции будут применяться только к многочленам ненулевой по x степени.

1. *Взятие производной*:
   $a_n x^n + a_{n-1} x^{n-1} + \cdots + a_0 \mapsto n a_n x^{n-1} + (n-1) a_{n-1} x^{n-2} + \ldots + a_1$.

2. *Модифицированный остаток*: применяется к двум многочленам

$p(x) = a_n x^n + a_{n-1}x^{n-1} + \ldots + a_0$ и $q(x) = b_m x^m + b_{m-1}x^{m-1} + \ldots + b_0$ таким, что $m \leqslant n$ и $b_m \neq 0$, и даёт такой многочлен $r(x)$, что

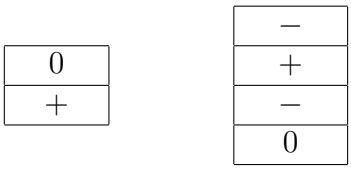
$$b_m^{n-m+1} p(x) = q(x)h(x) + r(x)\,,$$

где $\deg r(x) < \deg q(x)$ и $h(x)$ — некоторый многочлен.

3. *Взятие старшего коэффициента*:
   $a_n x^n + a_{n-1}x^{n-1} + \cdots + a_0 \mapsto a_n$.

4. *Отбрасывание старшего члена*:
   $a_n x^n + a_{n-1}x^{n-1} + \cdots + a_0 \mapsto a_{n-1}x^{n-1} + \cdots + a_0$.

При всяком натуральном $k \geqslant 1$ будем называть *столбцами (высоты k)* элементы множества $\underbrace{\{0,+,-\} \times \ldots \times \{0,+,-\}}_{k}$. Столбцы будем записывать по вертикали, например:



Всякое слово в алфавите столбцов высоты k будем называть *диаграммой высоты k*. В диаграммах естественно выделяются *строки* — слова в алфавите $\{0,+,-\}$.

Пусть задан список из k многочленов от одной переменной с действительными коэффициентами. Сейчас мы сопоставим с ним диаграмму высоты k. Для этого сначала отметим на числовой прямой корни всех ненулевых многочленов списка. Пусть это $\xi_1, \ldots, \xi_n$. Назовём *сегментом* каждое из $2n+1$ множеств $(-\infty, \xi_0)$, $\{\xi_0\}$, (ξ_0, ξ_1), $\ldots$, $\{\xi_n\}$, $(\xi_n, +\infty)$. Возьмём i-ый многочлен списка и выпишем последовательно его знаки на каждом из сегментов (ясно, что внутри сегмента знак многочлена из списка не меняется). Полученное слово в алфавите $\{0,+,-\}$ и будет i-ой строкой строящейся диаграммы.

Для произвольного списка L многочленов от нескольких переменных (с выделенной переменной x) определено понятие *замыкания* $\mathbf{C}L$ этого списка относительно указанных выше четырёх операций. Так как степень по x результата применения каждой операции строго меньше степеней по x аргументов этой операции, то список, являющийся замыканием

конечного списка — конечен. Назовём *основой* $\mathbf{BC}L$ замкнутого списка многочленов подсписок этого списка, образованный многочленами нулевой степени по x. Ясно, что по списку L можно эффективно построить $\mathbf{C}L$ и $\mathbf{BC}L$. Будем обозначать длину списка L через $\delta(L)$.

Для списка многочленов от переменных $x, y_1, \ldots, y_l$ очевидно определяется понятие диаграммы в точке $c_1, \ldots, c_l$: полагаем $y_1 = c_1, \ldots, y_l = c_l$ и рассматриваем диаграмму получающегося семейства многочленов от x с действительными коэффициентами.

Если список состоит из многочленов нулевой по x степени, то его диаграмма в любой точке имеет длину 1.

Основная лемма. *Существует алгоритм, который по произвольному списку L многочленов от x, $y_1, \ldots, y_l$ строит такое отображение, перерабатывающее некоторый диаграммы высоты $\delta(\mathbf{BC}L)$ и длины 1 в диаграммы высоты $\delta(L)$, что диаграмма L в любой точке $\mathbf{c} = (c_1, \ldots, c_l)$ есть образ при построенном отображении диаграммы $\mathbf{BC}L$ в точке $\mathbf{c}$.*

Доказательство. Прежде всего по списку многочленов L построим его замыкание $\mathbf{C}L$ и выделим в нём $\mathbf{BC}L$. Будем теперь по диаграмме для $\mathbf{BC}L$ (имея в виду, что это диаграмма в некоторой неизвестной нам точке $\mathbf{c}$) строить диаграмму для $\mathbf{C}L$; диаграмма для L получится тривиально.

Построение будем вести постепенно, в последовательности, обратной к построению замыкания, то есть добавляя к списку многочлен после того, как все многочлены из замыкания, имеющие меньшую степень, уже добавлены.

Итак, пусть диаграмма для списка U уже построена; как её расширить до диаграммы списка pU? Диаграмма списка pU будет иметь дополнительную строку внизу. Кроме того, она будет содержать дополнительные столбцы, по два на каждый корень многочлена p, не являющийся корнем никаких многочленов из U (которые в $\mathbf{c}$ не равны нулю тождественно по x).

Мы опишем лишь построение диаграммы. Доказательство правильности предлагаемого алгоритма элементарно. Заметим, что в списке U содержится старший коэффициент многочлена p. Пусть отвечающая ему строка диаграммы нулевая. Тогда диаграмма для pU получается из диаграммы для U приписыванием снизу строки, отвечающей многочлену, получаемому из p отбрасыванием старшего члена (такая строка уже была где-то в диаграмме).

Далее будем предполагать, что старший коэффициент многочлена p — ненулевой (и нам известен его знак). Начнём с того, что добавим в диаграмму снизу одну строку и заполним в ней клетки, отвечающие одноточечным сегментам, то есть найдём знаки p в корнях многочленов из U. Это, конечно, сделать нетрудно, используя операцию модифицированного остатка, а именно, деля многочлен p на многочлен, имеющий в данной точке корень.

Ключевыми соображениями теперь будут:

1. строго между двумя корнями многочлена имеется корень его производной;

2. корень многочлена p, в котором p не меняет знак, является корнем его производной;

3. между точками, в которых многочлен p имеет знаки $+$ и $-$, есть корень p.

Исходя из этих соображений, легко построить нужную процедуру, но мы всё же выпишем её полностью. Если в последней строке таблицы оказалась пустая клетка, с которой соседствует с одной стороны $+$, а с другой стороны $+$ или 0, то ставим в эту клетку $+$. Аналогично для $-$. Зная знак старшего коэффициента p и чётность степени p, находим знаки p в $-\infty$ и $+\infty$. Если крайний справа из поставленных знаков в нижней строке есть 0 или совпадает со знаком p в $+\infty$, то заполняем последнюю клетку нижней строки знаком p в $+\infty$. В противном случае заполняем эту клетку тем же знаком, что и предыдущую, и добавляем к диаграмме справа два столбца. Эти столбцы, если отбросить в них самый нижний элемент, совпадают с последним стобцом в диаграмме для U. В нижнюю строку этих столбцов заносим слева 0 и справа знак p в $+\infty$. Совершенно аналогично поступаем с левым краем диаграммы (который соответствует $-\infty$).

Почти так же добавляем столбцы внутрь диаграммы. Именно, если в каком-то столбце γ внизу стоит пустая клетка, а по соседству с ней стоят с одной стороны $+$, а с другой стороны $-$, то этот столбец заменяется на три. Все элементы, кроме нижнего, в этих столбцах те же, что у γ. В нижней строке этих столбцов среднюю клетку заполняем нулём, а каждую из крайних — тем знаком, с которым она соседствует. Лемма доказана. $\square$

Объясним теперь, как проводить элиминацию кванторов в рассматриваемой теории. Пусть дана формула $\exists x \Phi$ (где Φ — бескванторная). Будем считать, что в Φ все атомные формулы приведены к одному из видов $p > 0$, $p = 0$ или $p < 0$. Образуем список L всех многочленов, входящих в Φ. Заметим, что, зная диаграмму списка L в некоторой точке $\mathbf{c}$ и ничего больше не зная об этой точке, мы можем сказать, истинна ли формула $\exists x \Phi$ в этой точке.

Образуем замыкание $\mathbf{C}L$, найдём основу $\mathbf{BC}L$. Согласно лемме достаточно знать диаграмму $\mathbf{BC}L$ в некоторой точке, чтобы найти диаграмму L в этой точке (и, следовательно, определить, истинна ли формула в ней). Таким образом, истинность формулы $\exists x \Phi$ в точке $\mathbf{c}$ определяется знаками некоторых многочленов, не содержащих x, а именно, многочленов из $\mathbf{BC}L$. Это позволяет заменить $\exists x \Phi$ эквивалентной бескванторной формулой. Элиминируемость кванторов доказана.

Список литературы

[1] A. Tarsky. A decision method for elementary algebra and geometry. Univ. of Calif. Press, Berkeley, 1948.

[2] A. A. Seidenberg. A new decision method for elementary algebra. *Annals of Mathematics*, v. 60, 1954, pp. 365–374.

[3] P. J. Cohen. Decision procedures for real and p-adic fields. *Communications on Pure and Applied Mathematics*, v. 22, N 2, 1969, pp. 131–151.

[4] M. J. Fischer, M. O. Rabin. Super-exponential complexity of Presburger arithmetic. *Proc. AMS Symp. on Complexity of Real Computational Processes*, v. 7, 1974.

[5] M. Ben-Or, D. Kozen, J. Reif. The complexity of elementary algebra and geometry. *Journal of Computer and System Sciences*, v. 32, 1986, pp. 251–265.

[6] А. Л. Семенов. Разрешающие алгоритмы для логических теорий. В сборнике: *Кибернетика и вычислительная техника*, М., 1986, вып. 2, с. 134–146.