МАТЕМАТИЧЕСКИЕ ОСНОВЫ ИНФОРМАТИКИ И ПРОГРАММИРОВАНИЯ

УДК 510.52 DOI 10.17223/20710410/64/6

О ГЕНЕРИЧЕСКОЙ СЛОЖНОСТИ РЕШЕНИЯ УРАВНЕНИЙ НАД НАТУРАЛЬНЫМИ ЧИСЛАМИ СО СЛОЖЕНИЕМ[1]

А. Н. Рыбалов

Институт математики им. С. Л. Соболева СО РАН, г. Омск, Россия

E-mail: alexander.rybalov@gmail.com

Изучается генерическая сложность проблемы проверки совместности систем уравнений над натуральными числами со сложением. Доказывается NP-полнота этой проблемы, предлагается полиномиальный генерический алгоритм её решения. Доказывается, что при P $\neq$ NP и P = BPP для проблемы проверки совместности систем уравнений над натуральными числами с нулём не существует сильно генерического полиномиального алгоритма. Для сильно генерического полиномиального алгоритма нет эффективного метода случайной генерации входов, на которых этот алгоритм не может решить проблему.

Ключевые слова: *генерическая сложность, линейные уравнения, натуральные числа.*

ON THE GENERIC COMPLEXITY OF SOLVING EQUATIONS OVER NATURAL NUMBERS WITH ADDITION

A. N. Rybalov

Sobolev Institute of Mathematics, Omsk, Russia

We study the general complexity of the problem of determining the solvability of equations systems over natural numbers with the addition. The NP-completeness of this problem is proved. A polynomial generic algorithm for solving this problem is proposed. It is proved that if P $\neq$ NP and P = BPP, then for the problem of checking the solvability of systems of equations over natural numbers with zero there is no strongly generic polynomial algorithm. For a strongly generic polynomial algorithm, there is no efficient method for random generation of inputs on which the algorithm cannot solve the problem. To prove this theorem, we use the method of generic amplification, which allows us to construct generically hard problems from problems that are hard in the classical sense. The main feature of this method is the cloning technique, which combines the input data of a problem into sufficiently large sets of equivalent input data. Equivalence is understood in the sense that the problem is solved similarly for them.

Keywords: *generic complexity, linear equations, natural numbers.*

[1]Работа выполнена в рамках госзадания ИМ СО РАН, проект FWNF-2022-0003.

Введение

Решение уравнений и систем уравнений над вещественными, комплексными, рациональными, целыми числами является классической темой исследований в различных областях математики в течение уже нескольких тысяч лет. Классическая алгебраическая геометрия изучает множества решений алгебраических уравнений над полями вещественных и комплексных чисел. В рамках диофантовой геометрии и диофантова анализа изучаются решения алгебраических уравнений над целыми и рациональными числами. В XX веке большую роль начали играть вычислительные аспекты этих теорий. Изучение алгоритмических проблем, связанных с определением наличия решения у систем уравнений, а также с нахождением и описанием множества решений, является темой многочисленных теоретических и практических исследований.

Как правило, проблема определения совместности систем уравнений над различными алгебраическими системами является либо неразрешимой, либо имеет большую вычислительную сложность. Даже над нетривиальными конечными алгебраическими системами, например над конечными полями, эта проблема оказывается NP-полной. Это означает, что при условии $\mathrm{P} \neq \mathrm{NP}$ для неё не существует полиномиальных алгоритмов. Поэтому актуальным является изучение генерической сложности [1] данных проблем. В рамках генерического подхода алгоритмическая проблема рассматривается не на всём множестве входов, а на некотором подмножестве «почти всех» входов. Такие входы образуют так называемое генерическое множество. Понятие «почти все» формализуется введением асимптотической плотности на множестве входных данных. Результаты об асимптотической плотности совместных уравнений в свободных группах [2] и в нильпотентных группах [3] получены Р. Гилманом, А. Г. Мясниковым и В. А. Романьковым. Генерическая сложность проблемы совместности уравнений над кольцом целых чисел изучалась А. Н. Рыбаловым [4], а над конечными группами, полями и полугруппами — А. Н. Рыбаловым и А. Н. Шевляковым [5]. Уравнения над моноидом натуральных чисел по сложению в рамках универсальной алгебраической геометрии исследовались в работах А. Н. Шевлякова [6, 7]. С. Л. Кривым [8] рассмотрены критерии совместности системы линейных диофантовых уравнений над множеством натуральных чисел, а также даны верхние оценки для компонент минимального множества решений и алгоритмы построения минимальных порождающих наборов решений для всех таких систем.

В данной работе изучается генерическая сложность проблемы проверки совместности систем уравнений над натуральными числами со сложением.

1. Предварительные сведения

Пусть $\mathbb{N} = \{1, 2, 3, \ldots\}$ — множество натуральных чисел, начинающееся с 1, а $\omega = \{0, 1, 2, \ldots\}$ — множество натуральных чисел с нулём. Будем рассматривать две алгебраические системы $\mathfrak{N}_0 = \langle \omega, + \rangle$ и $\mathfrak{N}_1 = \langle \mathbb{N}, + \rangle$. Нетрудно показать, что любую систему уравнений над натуральными числами со сложением можно привести к эквивалентной системе, в которой каждое уравнение является равенством одного из следующих типов:

1) $x_i = x_j + x_k$;
2) $x_i = 1$.

При этом в k-м уравнении системы могут встречаться только переменные x_i, где $i \leqslant 3k$. Под *системой уравнений* будем понимать систему описанного вида. Размер

такой системы — это число уравнений в ней. Обозначим через $\mathcal{S}$ множество всех систем уравнений.

Лемма 1. Число систем уравнений размера n равно

$$|\mathcal{S}_n| = \prod_{k=1}^{n} (27k^3 + 3k).$$

Доказательство. Для t-го уравнения в системе $S \in \mathcal{S}_n$ существует $(3t)^3$ вариантов выбрать уравнение вида $x_i = x_j + x_k$ и $3t$ вариантов выбрать уравнение вида $x_i = 1$. Итого для t-го уравнения есть $27t^3 + 3t$ вариантов; для всей системы из n уравнений имеем $|\mathcal{S}_n| = \prod_{k=1}^{n} (27k^3 + 3k)$ вариантов. ■

Основные определения генерического подхода можно найти в [1] или в [4]. Определения вычислительных классов P, NP и BPP содержатся в [9].

2. NP-полнота

Теорема 1. Проблемы проверки совместности систем уравнений над $\mathfrak{N}_1$ и $\mathfrak{N}_0$ являются NP-полными.

Доказательство. Принадлежность этих проблем к классу NP следует из того, что в качестве подсказки (решения) можно взять значения переменных, которые делают все уравнения системы истинными. То, что размер минимального решения в двоичной записи ограничен полиномиально от размера системы, доказывается, например, в [10].

Докажем, что к проблеме проверки совместности систем уравнений над $\mathfrak{N}_1$ полиномиально сводится известная NP-полная проблема о выполнимости 3-КНФ, которая заключается в следующем.

3-КНФ $\Phi(x_1, \ldots, x_n)$ — это конъюнкция дизъюнкций вида $(\alpha_1 \vee \alpha_2 \vee \alpha_3)$, где α_i, $i = 1, 2, 3$, есть либо булева переменная x_k, либо её отрицание $\overline{x_k}$. Нужно определить, является ли заданная 3-КНФ $\Phi(x_1, \ldots, x_n)$ выполнимой, то есть существуют ли значения $a_1, \ldots, a_n \in \{0, 1\}$, такие, что $\Phi(a_1, \ldots, a_n) = 1$.

Построим по 3-КНФ $\Phi(x_1, \ldots, x_n)$ систему уравнений над $\mathfrak{N}_1$, которая имеет решение тогда и только тогда, когда $\Phi(x_1, \ldots, x_n)$ выполнима. Для этого каждой булевой переменной из Φ сопоставим две натуральнозначные переменные x и y, а также запишем систему уравнений с новыми вспомогательными переменными t_1, t_2, t_3:

$$\begin{cases} t_1 = 1, \\ t_2 = t_1 + t_1, \\ t_3 = t_1 + t_2, \\ t_3 = x + y. \end{cases}$$

Эта система гарантирует, что x и y могут принимать значения только 1 или 2. Число 2 выполняет роль логической единицы (истина), а 1 — логического нуля (ложь). У переменных роли таковы: x моделирует соответствующую логическую переменную, а y — её отрицание.

Далее смоделируем дизъюнкцию $(\alpha_1 \vee \alpha_2 \vee \alpha_3)$. Для этого возьмём переменные z_1, z_2, z_3 над $\mathbb{N}$, соответствующие логическим переменным (или их отрицаниям) из дизъюнкции, и запишем систему

$$\begin{cases} u_1 = 1, \\ u_2 = u_1 + u_1, \\ u_3 = u_1 + u_2, \\ s = u + u_3, \\ r = z_1 + z_2, \\ s = r + z_3, \end{cases}$$

где u, u_1, u_2, u_3, s, r — новые вспомогательные переменные. Заметим, что эта система совместна над $\mathbb{N}$ тогда и только тогда, когда хотя бы одна из переменных z_1, z_2, z_3 равна 2, что соответствует истинности соответствующей дизъюнкции.

Теперь по каждой дизъюнкции 3-КНФ $\Phi(x_1, \ldots, x_n)$ строим подобные системы уравнений. Затем нумеруем все переменные так, чтобы в k-м уравнении системы встречались только переменные x_i, где $i \leqslant 3k$. В итоге получаем систему S_Φ, которая совместна над $\mathbb{N}$ тогда и только тогда, когда 3-КНФ $\Phi(x_1, \ldots, x_n)$ выполнима. Полиномиальность данной сводимости следует из процесса построения.

Аналогично доказывается NP-полнота для $\mathfrak{N}_0$. ■

3. Решение уравнений над $\mathfrak{N}_1$

Теорема 2. Проблема проверки совместности систем уравнений над $\mathfrak{N}_1$ генерически разрешима за полиномиальное время.

Доказательство. Соответствующий генерический полиномиальный алгоритм работает на системе S следующим образом:

1) ищет в системе S уравнение вида $x_i = x_i + x_j$;
2) если такое уравнение найдётся, то система несовместна над $\mathbb{N}$ и алгоритм выдаёт ответ 0;
3) иначе выдаёт ответ «?».

Покажем, что этот алгоритм даёт неопределённый ответ на пренебрежимом множестве систем уравнений. Обозначим через $\mathcal{L}$ семейство систем, в которых отсутствуют уравнения вида $x_i = x_i + x_j$. Число вариантов выбрать k-е уравнение в системе из множества $\mathcal{L}$ равно $27k^3 + 3k - 9k^2$. Поэтому

$$|\mathcal{L}_n| = \prod_{k=1}^{n} (27k^3 + 3k - 9k^2).$$

По лемме 1

$$|\mathcal{S}_n| = \prod_{k=1}^{n} (27k^3 + 3k).$$

Определение асимптотической плотности ρ и ρ_n можно найти в [1, 4]. Получаем

$$\rho_n(\mathcal{L}) = \frac{|\mathcal{L}_n|}{|\mathcal{S}_n|} = \frac{\prod\limits_{k=1}^{n} (27k^3 + 3k - 9k^2)}{\prod\limits_{k=1}^{n} (27k^3 + 3k)} = \prod_{k=1}^{n} \frac{9k^3 + k - 3k^2}{9k^3 + k} = \prod_{k=1}^{n} \Big(1 - \frac{3k^2}{9k^3 + k}\Big) <$$

$$< \prod_{k=1}^{n} \Big(1 - \frac{3k^2}{9k^3 + k^3}\Big) = \prod_{k=1}^{n} \Big(1 - \frac{3}{10k}\Big) < \prod_{k=1}^{n} \Big(1 - \frac{1}{4k}\Big).$$

Чтобы оценить сверху последнее произведение, возведём его в четвёртую степень:

$$\prod_{k=1}^{n}\Big(1-\frac{1}{4k}\Big)^4 < \prod_{k=1}^{n}\bigg(\Big(1-\frac{1}{4k}\Big)\Big(1-\frac{1}{4k+1}\Big)\Big(1-\frac{1}{4k+2}\Big)\Big(1-\frac{1}{4k+3}\Big)\bigg) = \prod_{k=4}^{4n}\Big(1-\frac{1}{k}\Big) =$$
$$= \prod_{k=4}^{4n}\Big(\frac{k-1}{k}\Big) = \frac{3}{4}\cdot\frac{4}{5}\cdot\ldots\cdot\frac{4n-2}{4n-1}\cdot\frac{4n-1}{4n} = \frac{3}{4n}.$$

Итого получаем

$$\rho(\mathcal{L}) = \lim_{n\to\infty}\frac{|\mathcal{L}_n|}{|\mathcal{S}_n|} \leqslant \lim_{n\to\infty}\sqrt[4]{\frac{3}{4n}} = 0.$$

Теорема 2 доказана. ■

4. Решение уравнений над $\mathfrak{N}_0$

Для произвольной системы уравнений $S = \{S_1, \ldots, S_m\}$ рассмотрим множество систем $eq(S)$, которые получаются добавлением к системе S произвольных уравнений $S_{m+1}, \ldots, S_n$, где l-е уравнение имеет вид $x_i = x_j + x_k$, причём $3m < i, j, k < 3(l+m)$. Очевидно, что любая система из $eq(S)$ совместна над $\mathfrak{N}_0$ тогда и только тогда, когда совместна над $\mathfrak{N}_0$ система S.

Лемма 2. Для любой системы S размера m и любого $n > m$ имеет место оценка

$$\frac{|eq(S)_n|}{|\mathcal{S}_n|} > \frac{1}{2(30n^3)^m}.$$

Доказательство. Пусть $n > m$. Для t-го добавленного к S уравнения вида $x_i = x_j + x_k$, где $3m < i, j, k < 3(t+m)$, имеется $(3t)^3 = 27t^3$ вариантов. Поэтому

$$|eq(S)_n| = \prod_{t=1}^{n-m}(27t^3).$$

По лемме 1

$$\rho_n(eq(S)) = \frac{|eq(S)_n|}{|\mathcal{S}_n|} = \frac{\prod\limits_{k=1}^{n-m}(27k^3)}{\prod\limits_{k=1}^{n}(27k^3+3k)} = \prod_{k=1}^{n-m}\Big(\frac{27k^2}{27k^2+3}\Big)\prod_{k=n-m+1}^{n}\frac{1}{27k^3+3k}.$$

Оценим снизу сначала первое произведение:

$$\prod_{k=1}^{n-m}\Big(\frac{27k^2}{27k^2+3}\Big) = \prod_{k=1}^{n-m}\Big(1-\frac{1}{9k^2+1}\Big) > \prod_{k=2}^{n-m+1}\Big(1-\frac{1}{k^2}\Big) = \prod_{k=2}^{n-m+1}\frac{(k-1)(k+1)}{k^2} =$$
$$= \frac{1\cdot 3}{2^2}\cdot\frac{2\cdot 4}{3^2}\cdot\frac{(n-m-1)(n-m+1)}{(n-m)^2}\cdot\frac{(n-m)(n-m+2)}{(n-m+1)^2} = \frac{n-m+2}{2(n-m+1)} > \frac{1}{2}.$$

Теперь оценим второе произведение:

$$\prod_{k=n-m+1}^{n}\frac{1}{27k^3+3k} > \frac{1}{(27n^3+3n)^m}.$$

Итого получаем

$$\rho_n(eq(S)) = \frac{|eq(S)_n|}{|\mathcal{S}_n|} > \frac{1}{2(27n^3+3n)^m} > \frac{1}{2(30n^3)^m}.$$

Лемма 2 доказана. ■

Теорема 3. Пусть P $\neq$ NP и P = BPP. Тогда для проблемы проверки совместности систем уравнений над $\mathfrak{N}_0$ не существует сильно генерического полиномиального алгоритма.

Доказательство. Допустим, что существует сильно генерический полиномиальный алгоритм $\mathcal{A}$, определяющий совместность систем уравнений над $\mathfrak{N}_0$. Построим вероятностный полиномиальный алгоритм $\mathcal{B}$, решающий эту проблему на всём множестве входов. На системе S размера n алгоритм $\mathcal{B}$ работает следующим образом:

1) генерирует случайно равновероятно систему S' из множества $eq(S)$ размера n^2;
2) запускает алгоритм $\mathcal{A}$ на системе S';
3) если $\mathcal{A}(S') \neq ?$, то алгоритм правильно определяет, совместна ли система S', а вместе с ней и система S;
4) если $\mathcal{A}(S') = ?$, то выдаёт ответ «НЕТ».

Заметим, что полиномиальный вероятностный алгоритм $\mathcal{B}$ выдаёт правильный ответ на шаге 3, а на шаге 4 может выдать неправильный ответ. Надо доказать, что вероятность того, что ответ выдаётся на шаге 4, меньше 1/3.

Оценим вероятность выдачи ответа на шаге 4. Вероятность того, что для S' имеет место $\mathcal{A}(S') = ?$, не больше

$$\frac{|\{S' \in \mathcal{S} : \mathcal{A}(S') = ?\}_{n^2}|}{|eq(S)_{n^2}|} = \frac{|\{S' \in \mathcal{S} : \mathcal{A}(S') = ?\}_{n^2}|}{|\mathcal{S}_{n^2}|} \cdot \frac{|\mathcal{S}_{n^2}|}{|eq(S)_{n^2}|}.$$

Так как множество $\{S' \in \mathcal{S} : \mathcal{A}(S') = ?\}$ сильно пренебрежимое, существует константа $\alpha > 0$, такая, что

$$\frac{|\{S' \in \mathcal{S} : \mathcal{A}(S') = ?\}_{n^2}|}{|\mathcal{S}_{n^2}|} < \frac{1}{2^{\alpha n^2}}$$

для любого n. По лемме 2

$$\frac{|\mathcal{S}_{n^2}|}{|eq(S)_{n^2}|} < 2(30n^6)^n.$$

Поэтому искомая вероятность ответа на шаге 4 не больше

$$\frac{2(30n^6)^n}{2^{\alpha n^2}} \cdot \frac{2^{1+\log 30n+6n\log n}}{2^{\alpha n^2}} = \frac{1}{2^{\alpha n^2 - 6n\log n - \log 30n - 1}} < \frac{1}{3}$$

при больших n.

Таким образом, проблема проверки совместности систем уравнений над $\mathfrak{N}_0$ принадлежит классу BPP. А так как BPP = P, то она принадлежит классу P. А это, по теореме 1, противоречит условию P $\neq$ NP. ■

Автор выражает искреннюю благодарность рецензенту за полезные замечания и предложения по улучшению текста статьи.

ЛИТЕРАТУРА

1. *Kapovich I., Miasnikov A., Schupp P., and Shpilrain V.* Generic-case complexity, decision problems in group theory and random walks // J. Algebra. 2003. V. 264. No. 2. P. 665–694.
2. *Gilman R. H., Myasnikov A., and Roman'kov V.* Random equations in free groups // Groups Complexity Cryptology. 2011. V. 3. No. 2. P. 257–284.
3. *Gilman R. H., Myasnikov A., and Roman'kov V.* Random equations in nilpotent groups // J. Algebra. 2012. V. 352. No. 1. P. 192–214.
4. *Rybalov A.* Generic complexity of the Diophantine problem // Groups Complexity Cryptology. 2013. V. 5. No. 1. P. 25–30.

5. *Rybalov A. and Shevlyakov A.* Generic complexity of solving of equations in finite groups, semigroups and fields // J. Physics: Conf. Ser. 2021. V. 1901. Article 012047. 8 p.
6. *Shevlyakov A.* Algebraic geometry over the additive monoid of natural numbers: The classifcation of coordinate monoids // Groups Complexity Cryptology. 2010. V. 2. No. 1. P. 91–111.
7. *Шевляков А. Н.* Алгебраическая геометрия над моноидом натуральных чисел. Неприводимые алгебраические множества // Труды Института математики и механики УрО РАН. 2010. Т. 16. № 2. С. 258–269.
8. *Kryvyi S. L.* Compatibility of systems of linear constraints over the set of natural numbers // Cybernetics Systems Analysis. 2002. V. 38. No. 1. P. 17–29.
9. *Китаев А., Шень А., Вялый М.* Классические и квантовые вычисления. М.: МЦНМО, ЧеРо, 1999. 192 с.
10. *Схрейвер А.* Теория линейного и целочисленного программирования. М.: Мир, 1991. 360 с.

REFERENCES

1. *Kapovich I., Miasnikov A., Schupp P., and Shpilrain V.* Generic-case complexity, decision problems in group theory and random walks. J. Algebra, 2003, vol. 264, no. 2, pp. 665–694.
2. *Gilman R. H., Myasnikov A., and Roman'kov V.* Random equations in free groups. Groups Complexity Cryptology, 2011, vol. 3, no. 2, pp. 257–284.
3. *Gilman R. H., Myasnikov A., and Roman'kov V.* Random equations in nilpotent groups. J. Algebra, 2012, vol. 352, no. 1, pp. 192–214.
4. *Rybalov A.* Generic complexity of the Diophantine problem. Groups Complexity Cryptology, 2013, vol. 5, no. 1, pp. 25–30.
5. *Rybalov A. and Shevlyakov A.* Generic complexity of solving of equations in finite groups, semigroups and fields. J. Physics: Conf. Ser., 2021, vol. 1901, Article 012047, 8 p.
6. *Shevlyakov A.* Algebraic geometry over the additive monoid of natural numbers: The classifcation of coordinate monoids. Groups Complexity Cryptology, 2010, vol. 2, no. 1, pp. 91–111.
7. *Shevlyakov A. N.* Algebraicheskaya geometriya nad monoidom natural'nykh chisel. Neprivodimye algebraicheskie mnozhestva [Algebraic geometry over the monoid of natural numbers. Irreducible algebraic sets]. Trudy Instituta Matematiki i Mekhaniki UrO RAN, 2010, vol. 16, no. 2, pp. 258–269. (in Russian)
8. *Kryvyi S. L.* Compatibility of systems of linear constraints over the set of natural numbers. Cybernetics Systems Analysis, 2002, vol. 38, no. 1, pp. 17–29.
9. *Kitaev A., Shen' A., and Vyalyy M.* Klassicheskie i kvantovye vychisleniya [Classical and Quantum Computations]. Moscow, MCCME Publ., 1999. 192 p. (in Russian)
10. *Schrijver A.* Theory of Linear and Integer Programming. Wiley, 1998. 484 p.