

## ДОСТАТОЧНОЕ УСЛОВИЕ ПОЛИНОМИАЛЬНОЙ РАЗРЕШИМОСТИ СЛУЧАЙНЫХ 3-КНФ ФОРМУЛ

© 2024 г. С. И. Уваров<sup>1</sup>, \*

Представлено академиком РАН С.Н. Васильевым

Поступило 22.04.2024 г.

После доработки 11.06.2024 г.

Принято к публикации 16.07.2024 г.

Настоящая работа посвящена локализации случайных 3-КНФ формул, полиномиально разрешимых резолюционным алгоритмом. Показано, что случайные формулы с числом дизъюнктов, пропорциональным квадрату числа переменных, с вероятностью близкой к единице полиномиально разрешимы при коэффициенте пропорциональности, превышающем найденный порог.

**Ключевые слова:** 3-КНФ, дизъюнкт, резолюция, алгоритмическая сложность, задача выполнимости

**DOI:** 10.31857/S2686954324040067, **EDN:** YZHCIZ

### 1. ВВЕДЕНИЕ И ФОРМУЛИРОВКА РЕЗУЛЬТАТА

Важность известной проблемы алгоритмической сложности задач, принадлежащих классу NP, ярко представлена в публикации С. Кука [1]. Задача выполнимости формул, заданных в 3-КНФ, — одна из базовых задач в проблематике исследований алгоритмической сложности NP задач (обзоры по исследованию 3-КНФ формул представлены в [2, 3, 4, 5]).

3-КНФ формула — это конъюнкция различных 3-дизъюнктов. Каждый 3-дизъюнкт — это дизъюнкция  $\ell_a \vee \ell_b \vee \ell_c$  трех литералов  $\ell_a, \ell_b, \ell_c$  различных логических переменных  $a, b, c$ , соответственно. Литерал — это сама переменная  $v$  или ее отрицание  $\neg v$ , а общее количество переменных в формуле обозначается через  $N$ . Далее дизъюнкт обозначается как множество литералов, заключенных в фигурные скобки. С учетом того, что каждая переменная может быть представлена двумя литералами, количество  $\mathfrak{M}(N)$  различных 3-дизъюнктов от  $N$  переменных равно  $8 \cdot \binom{N}{3}$ .

В литературе при исследовании различных аспектов задачи выполнимости широко используются случайные 3-КНФ формулы. Случайная

3-КНФ формула содержит  $M_0$  3-дизъюнктов, выбранных равномерно случайным образом из множества, содержащего  $\mathfrak{M}(N)$  дизъюнктов. Отношение числа  $M_0$  дизъюнктов в формуле к числу  $N$  входящих в них разных переменных обозначается через  $R$ .

К настоящему времени (с использованием параметра  $R$ ) для случайных 3-КНФ формул выделено несколько областей в координатах  $N, M_0$ . Если  $R > 4,51$ , то с ростом  $N$  доля выполнимых случайных 3-КНФ формул становится пренебрежимо малой [6, 7]. При  $R < 3,52$  пренебрежимо малой становится доля невыполнимых формул [8]. При “фазовом переходе”  $R \approx 4,3$  доли выполнимых и невыполнимых формул примерно равны [2, 7, 8].

Известна область локализации полиномиально разрешимых выполнимых случайных 3-КНФ формул: при  $R < 1,63$  выполнимость таких формул с вероятностью  $1 - o(1)$  может быть установлена алгоритмом полиномиально ограниченной сложности [2, 9, 10].

В настоящей работе аналитически выделена область локализации полиномиально разрешимых с вероятностью близкой к единице невыполнимых случайных 3-КНФ формул. Это формулы, в которых

$$M_0 \geq 0,061 \cdot N^2. \quad (1)$$

С ростом  $N$  числа переменных вероятность доказательства их невыполнимости алгоритмом полиномиальной сложности увеличивается.

<sup>1</sup> Институт проблем управления Российской академии наук, Москва, Россия

\*E-mail: suvarov@ipu.ru

## 2. ВЫВОД УСЛОВИЯ ПОЛИНОМИАЛЬНОЙ РАЗРЕШИМОСТИ

Известным свойством КНФ формул является возможность резолюционного вывода нового дизъюнкта (резольвента)  $\{b, c, \neg d, e\}$  из пары резольвентов  $\{a, b, c\}$  и  $\{\neg a, c, \neg d, e\}$ , что далее обозначается:  $\{a, b, c\}, \{\neg a, c, \neg d, e\} \vdash \{b, c, \neg d, e\}$ . Резольвента содержит (в единственном числе) все литералы обоих резольвентов за исключением контрарных литералов  $a$  и  $\neg a$ .

Обозначим через  $\mathfrak{R}(S, T, U)$  резолюцию, первый и второй резольвенты которой содержат соответственно  $S$  и  $T$  литералов, а резольвента содержит  $U$  литералов. Если резольвенты содержат  $Q$  общих литералов, а некоторая переменная представлена в резольвентах контрарными литералами, то  $U = S + T - Q - 2$ . При условии, что  $S = T = Q + 1$ , получаем  $U = Q = S - 1 = T - 1$ . В остальных случаях  $U \geq S$  и  $U \geq T$ . Полученный  $U$ -дизъюнкт добавляется в исходную КНФ формулу без изменения ее свойства выполнимости/ невыполнимости. В случае вывода пустого дизъюнкта  $\{\}$  невыполнимость формулы оказывается доказанной. Полнота алгоритма означает, что если не доказана невыполнимость формулы, то она выполнима.

Алгоритмическая сложность резолюционного доказательства невыполнимости 3-КНФ формул имеет экспоненциальную по числу переменных оценку [11, 12]. Такая оценка связана с необходимостью вывода резольвент, число литералов в которых соизмеримо с числом переменных.

Если принудительно ограничить константой максимально допустимое значение  $U$ , алгоритм доказательства невыполнимости КНФ формулы перестает быть полным, а оценка его алгоритмической сложности становится полиномиальной по  $N$ . Неполнота алгоритма состоит в том, что если не доказана невыполнимость формулы, то это еще не означает, что формула обязательно выполнима.

Ограничение  $U \leq 3$  позволяет выполнять только резолюции следующих типов:

- $\mathfrak{R}(3, 3, 3) : \{a, b, c\}, \{\neg a, b, d\} \vdash \{b, c, d\};$
- $\mathfrak{R}(3, 3, 2) : \{a, b, c\}, \{\neg a, b, c\} \vdash \{b, c\};$
- $\mathfrak{R}(3, 2, 3) : \{a, b, c\}, \{\neg a, d\} \vdash \{b, c, d\};$
- $\mathfrak{R}(3, 2, 2) : \{a, b, c\}, \{\neg a, b\} \vdash \{b, c\};$
- $\mathfrak{R}(2, 2, 2) : \{a, b\}, \{\neg a, c\} \vdash \{b, c\};$
- $\mathfrak{R}(2, 2, 1) : \{a, b\}, \{\neg a, b\} \vdash \{b\};$
- $\mathfrak{R}(2, 1, 1) : \{a, b\}, \{\neg a\} \vdash \{b\};$
- $\mathfrak{R}(1, 1, 0) : \{a\}, \{\neg a\} \vdash \{\}.$

Аналитический вывод достаточного условия (1) полиномиальной разрешимости случайных 3-КНФ формул удобно осуществлять, оперируя резольвентами единой длины. С этой целью введена модифицированная резолюция  $\mathfrak{R}^*(3, 3, 2)$  вида  $\{a, b, c\}, \{\neg a, b, c\} \vdash \mathcal{E}(b, c)$ , которая порождает множество  $\mathcal{E}(b, c)$ , состоящее из 3-дизъюнктов  $\{b, c, x\}$ . Здесь  $x$  представляет все литералы логических переменных за исключением литералов переменных  $a, b, c$ . Множество  $\mathcal{E}(b, c)$  содержит  $2(N - 3)$  дизъюнктов.

Рассмотрим возможность доказательства невыполнимости 3-КНФ формул с использованием только двух резолюций:  $\mathfrak{R}(3, 3, 3)$  и  $\mathfrak{R}^*(3, 3, 2)$ . Такое ограничение позволяет считать, что преобразуемая КНФ формула всегда состоит только из 3-дизъюнктов.

Если КНФ формула содержит более, чем  $\mathfrak{M}(N) \cdot 7 / 8$  различных 3-дизъюнктов, то она невыполнима. Действительно, при этом найдутся 3 переменные, из литералов которых построены 8 дизъюнктов, что означает противоречивость формулы.

Рассмотрим алгоритм 1 доказательства невыполнимости 3-КНФ формул, использующий только два упомянутых типа резолюций  $\mathfrak{R}(3, 3, 3)$  и  $\mathfrak{R}^*(3, 3, 2)$ .

### Алгоритм 1

Дано: 3-КНФ формула, содержащая  $M_0$  дизъюнктов. Множество этих дизъюнктов обозначается  $M_0$ .

Найти: Является ли формула невыполнимой (или ее свойство неопределено)?

1. Из резольвентов множества  $M_0$ , резолюциями  $\mathfrak{R}(3, 3, 3)$  и  $\mathfrak{R}^*(3, 3, 2)$ , выводятся новые дизъюнкты. Из полученных новых дизъюнктов формируется множество  $M_1$ .

Полагаем  $k = 1$ .

2. Новые дизъюнкты множества  $M_{k+1}$  выводятся резолюциями  $\mathfrak{R}(3, 3, 3)$  и  $\mathfrak{R}^*(3, 3, 2)$ , причем первыми резольвентами являются дизъюнкты из множества  $M_k$ , а вторыми резольвентами — дизъюнкты из множества  $\bigcup_{j=0}^k M_j$ .

(а) Множество  $M_{k+1}$  пусто. Алгоритм не смог доказать невыполнимость формулы, изначально содержащей  $M_0$  дизъюнктов. **ОСТАНОВ.**

(б) Множество  $M_{k+1}$  не пусто:

i. Число дизъюнктов в множестве  $\bigcup_{j=0}^{k+1} M_j$  превышает число  $\mathfrak{M}(N) \cdot 7 / 8$  — невыполнимость формулы доказана. **ОСТАНОВ.**

ii. Число дизъюнктов в множестве  $\bigcup_{j=0}^{k+1} \mathcal{M}_j$  не превышает число  $\mathfrak{M}(N) \cdot 7/8$ ; полагаем  $k = k + 1$  и повторяем шаг 2.

Оценим ожидаемое число дополнительных 3-дизъюнктов, которые могут быть получены резолюционным правилом из множества  $\mathcal{M}_0$  исходных 3-дизъюнктов. Число дизъюнктов в множестве  $\mathcal{M}_k$  обозначается через  $M_k$ .

Дизъюнкт  $\{a, b, c\}$  резолюциями  $\mathfrak{R}(3, 3, 3)$  в паре с любым из шести дизъюнктов  $\{\neg a, b, x\}$ ,  $\{\neg a, c, x\}$ ,  $\{a, \neg b, x\}$ ,  $\{\neg b, c, x\}$ ,  $\{a, \neg c, x\}$ ,  $\{b, \neg c, x\}$  порождает новый 3-дизъюнкт (резольвенту). В этом случае  $x$  является одним из  $2(N - 3)$  литералов от  $N$  переменных.

Ожидаемое число 3-дизъюнктов в множестве  $\mathcal{M}_1$ , которые порождаются резолюциями  $\mathfrak{R}(3, 3, 3)$  над множеством  $\mathcal{M}_0$  исходных дизъюнктов, оценивается как:

$$M'_1 = 6 \cdot 2(N - 3) \frac{M_0 \cdot M_0}{2 \cdot \mathfrak{M}(N)}.$$

Необходимо учесть возможность применения на множестве  $\mathcal{M}_0$  дизъюнктов резолюции  $\mathfrak{R}^*(3, 3, 2)$ . Резолюция осуществима, если первым резолютом является дизъюнкт  $\{a, b, c\}$ , а вторым — любой из трех дизъюнктов  $\{\neg a, b, c\}$ ,  $\{a, \neg b, c\}$ ,  $\{a, b, \neg c\}$ . Резольвентой, соответственно, будут 2-дизъюнкты из  $\{b, c\}$ ,  $\{a, c\}$ ,  $\{a, b\}$ . С учетом того, что вместо одного 2-дизъюнкта добавляются  $2(N - 3)$  3-дизъюнктов, ожидаемое число добавленных резолюциями  $\mathfrak{R}^*(3, 3, 2)$  3-дизъюнктов оценивается как

$$M''_1 = 3 \cdot 2(N - 3) \frac{M_0 \cdot M_0}{2 \cdot \mathfrak{M}(N)}.$$

Суммарно по резолюциям  $\mathfrak{R}(3, 3, 3)$  и  $\mathfrak{R}^*(3, 3, 2)$  получается

$$\widehat{M}_1 = M'_1 + M''_1 = 9(N - 3) \frac{M_0 \cdot M_0}{\mathfrak{M}(N)}. \quad (2)$$

Учитывая, что порождаемые дизъюнкты могут совпадать с уже имеющимися (при условии, что  $\widehat{M}_{k+1} + \sum_{j=0}^k M_j \leq \mathfrak{M}(N)$ ), при  $k \geq 0$  число дизъюнктов в множестве  $\mathcal{M}_{k+1}$  оценивается как

$$M_{k+1} \geq \widehat{M}_{k+1} \cdot \left( 1 - \frac{\widehat{M}_{k+1} + \sum_{j=0}^k M_j}{\mathfrak{M}(N)} \right). \quad (3)$$

На  $k$ -ом шаге ( $k \geq 1$ ) осуществляется попытка построения новых 3-дизъюнктов на основе ре-

золюций дизъюнктов из множества  $\mathcal{M}_k$  с дизъюнктами как из множества  $\bigcup_{j=0}^{k-1} \mathcal{M}_j$ , так и из множества  $\mathcal{M}_k$ . Полученная рекуррентная формула записывается в виде

$$\widehat{M}_{k+1} = \frac{18(N - 3) \cdot M_k}{\mathfrak{M}(N)} \cdot \left( \sum_{j=0}^{k-1} M_j + \frac{M_k}{2} \right). \quad (4)$$

Далее рассматривается упрощенная система рекуррентных соотношений,  $k \geq 1$ :

$$\begin{aligned} M_1 &= \frac{9(N - 3) \cdot M_0}{\mathfrak{M}(N)} \cdot M_0, \\ M_{k+1} &= \frac{18(N - 3) \cdot M_k}{\mathfrak{M}(N)} \cdot \left( \sum_{j=0}^{k-1} M_j + \frac{M_k}{2} \right). \end{aligned} \quad (5)$$

Для соотношений (5) справедливо следующее утверждение.

**Утверждение 1.** Для любого  $k \geq 1$  при  $M_{k+1} \geq M_k$  справедливо, что  $M_{k+2} \geq M_{k+1}$ .

Условие равенства значений  $M_1$  и  $M_2$  является достаточным условием того, что последовательность  $M_1, M_2, \dots, M_j$ ,  $j > 2$ , является неубывающей и, следовательно, в рекуррентном процессе может быть получено необходимое число 3-дизъюнктов, обеспечивающее доказательство невыполнимости формулы.

Если  $M_1 = M_2$ , и  $M_1$  и  $M_2$  заданы формулами (5) то

$$9(N - 3) \frac{M_0 \cdot M_0}{\mathfrak{M}(N)} = \frac{18(N - 3) \cdot M_1}{\mathfrak{M}(N)} \cdot \left( M_0 + \frac{M_1}{2} \right).$$

При упрощении данного выражения получается  $M_0 \cdot M_0 = 2M_0 \cdot M_1 + M_1 \cdot M_1$ . Далее после подстановки выражения (5) для  $M_1$  следует

$$\begin{aligned} M_0 \cdot M_0 &= 2M_0 \cdot \frac{9(N - 3) \cdot M_0}{\mathfrak{M}(N)} \cdot \\ &\cdot M_0 + \left( \frac{9(N - 3) \cdot M_0}{\mathfrak{M}(N)} \cdot M_0 \right)^2. \end{aligned}$$

Дальнейшее упрощение и обозначение через  $A$  величины  $9(N - 3) \cdot M_0 / \mathfrak{M}(N)$  приводит к квадратному уравнению  $1 = 2A + A^2$  с положительным корнем  $A = \sqrt{2} - 1 \approx 0.414$ . При этом

$$M_0 = \frac{(\sqrt{2} - 1) \cdot \mathfrak{M}(N)}{9(N - 3)}.$$

В представлении  $M_0$  в виде  $r \cdot N^2$  получается

$$r = \frac{8(\sqrt{2} - 1)}{27} \cdot \frac{N(N - 1)(N - 2)}{N^2(N - 3)}, \quad \lim_{N \rightarrow \infty} r \approx 0,061.$$

### 3. ДОКАЗАТЕЛЬСТВО УТВЕРЖДЕНИЯ И ОЦЕНКА ВЕРОЯТНОСТИ

*Доказательство утверждения 1.* Рассматривается разность

$$M_{k+2} - M_{k+1} = \frac{18(N-3)M_{k+1}}{\mathfrak{M}(N)} \left( \sum_{j=0}^k M_j + \frac{M_{k+1}}{2} \right) - \frac{18(N-3)M_k}{\mathfrak{M}(N)} \left( \sum_{j=0}^{k-1} M_j + \frac{M_k}{2} \right).$$

После подстановки в правую часть вместо  $M_{k+1}$  меньшей величины  $M_k$ , получится

$$M_{k+2} - M_{k+1} \geq \frac{18(N-3)M_k}{\mathfrak{M}(N)} \left( \sum_{j=0}^k M_j + \frac{M_k}{2} \right) - \frac{18(N-3)M_k}{\mathfrak{M}(N)} \left( \sum_{j=0}^{k-1} M_j + \frac{M_k}{2} \right).$$

Выполнение упрощающих преобразований приводит к  $M_{k+2} - M_{k+1} \geq 18(N-3) \cdot M_k^2 / \mathfrak{M}(N) \geq 0$ .

Для оценки вероятности того, что условие 1 обеспечивает доказательство (полиномиальной сложности) невыполнимости случайной 3-КНФ формулы, рассмотрим модифицированную систему рекуррентных формул (2) – (4). В этих модифицированных формулах через  $\alpha(M_0)$  обозначено  $M_0$ , а для  $k \geq 1$  через  $\alpha(M_k)$  и  $\gamma(M_k)$  обозначено  $\lfloor \beta \cdot M_k \rfloor$ ,  $1 \geq \beta > 0$  и  $\lfloor 2 \cdot M_k - 1 \rfloor$ , соответственно. Формула (2) для вычисления  $M_1$  остается неизменной, а формула (4) приобретает вид

$$M_{k+1} \geq \widehat{M}_{k+1} \cdot \left( 1 - \frac{\widehat{M}_{k+1} + \sum_{j=0}^k \alpha(M_j)}{\mathfrak{M}(N)} \right). \quad (6)$$

Формула (3) для вычисления  $\widehat{M}_k$  (при  $k > 1$ ) преобразуется в

$$\widehat{M}_{k+1} = \frac{18(N-3) \cdot \alpha(M_k)}{\mathfrak{M}(N)} \cdot \left( \sum_{j=1}^{k-1} \alpha(M_j) + \frac{\alpha(M_k)}{2} \right). \quad (7)$$

Непосредственное вычисление по формулам (2), (4) и (3) демонстрирует то, что для  $N \geq 64$  необходимая для доказательства невыполнимости формулы сумма дизъюнктов набирается за 7 циклов работы алгоритма 1. Вычисление по формулам (2), (6) и (7) при  $\beta = 0,7$  дает 17 циклов.

$$\sum_{k=0}^{17} \alpha(M_k) > \mathfrak{M}(N) \cdot 7 / 8.$$

В свою очередь, это означает, что если в каждом цикле  $k \geq 1$  работы алгоритма будет получено число  $\alpha(M_k)$  новых дизъюнктов, превышающее  $0,7 \cdot M_k$ , невыполнимость формулы будет доказана.

Вероятность  $p_k$  того, что при ожидаемом числе дизъюнктов  $M_k$  и биномиальном распределении  $\text{Bin}(\gamma(M_k), 0,5)$  число новых дизъюнктов будет меньше либо равно  $\alpha(M_k)$ , составляет

$$p_k = 2^{-\gamma(M_k)} \sum_{j=1}^{\alpha(M_k)} \binom{\gamma(M_k)}{j}. \quad (8)$$

Вероятность того, что процесс генерации новых дизъюнктов завершится доказательством невыполнимости 3-КНФ формулы оценивается как:

$$P = \prod_{j=1}^{17} (1 - p_j). \quad (9)$$

Вычисление по формулам (2), (6), (7), (8) и (9) дает следующую оценку вероятности того, что для случайной 3-КНФ формулы, содержащей  $0,061 \cdot N^2$  дизъюнктов, алгоритм 1 будет доказана невыполнимость: при  $N = 256, 1024$  и  $4096$  вероятность составляет  $1 - 10^{-32}$ ,  $1 - 10^{-481}$  и  $1 - 10^{-7500}$ , соответственно.

### СПИСОК ЛИТЕРАТУРЫ

1. Cook S. The Importance of the P versus NP Question // JACM. 2003. V. 50. P. 27–29.
2. Biere A., Heule M., Maaren H., Walsh T. Handbook of Satisfiability (Second Edition). IOS Press, 2021. <https://doi.org/10.3233/FAIA336>
3. Gomes C., Rautz H., Sabharwal A., Selman B. Satisfiability Solvers // Handbook of Knowledge Representation. 2008. P. 89–133.
4. Beame P., Kautz H.A., Sabharwal A. Towards Understanding and Harnessing the Potential of Clause Learning // Journal of Artificial Intelligence Research. 2004. V. 22. P. 319–351.
5. Ganesh V., Vardi M. On the Unreasonable Effectiveness of SAT Solvers // Beyond the Worst-Case Analysis of Algorithms. 2020. P. 547–566.
6. Kirkpatrick S., Selman B. Critical Behavior in the Satisfiability of Random Boolean Expressions // Science. 1994. V. 264. № 5163. P. 1297–1301.
7. Merzard M., Parisi G., Zecchina R. Analytic and Algorithmic Solution of Random Satisfiability Problems // Science. 2002. V. 297. P. 812–815.
8. Kaporis A.C., Kirousis L.M., Laias E.G. The Probabilistic Analysis of a Greedy Satisfiability Algorithm // Random Structures and Algorithms. 2006. V. 28. № 4. P. 444–480.

9. *Broder A.Z., Frieze A.M., Upfal E.* On the Satisfiability and Maximum Satisfiability of Random 3-CNF Formulas // In 4th ACM-SIAM Symposium on Discrete Algorithms. SIAM. 1993. P. 322–330.
10. *Mitzenmacher M.* Tight Thresholds for the Pure Literal Rule // DEC/SRC Technical Note 1997–011. 1997.
11. *Haken A.* The Intractability of Resolution // Theoretical Computer Science. 1985. V. 39. P. 297–308.
12. *Алехнович М.В., Разборов А.А.* Нижние оценки для полиномиального исчисления в случае идеалов, отличных от биномиальных // Труды математического института им. В.А. Стеклова. 2003. Т. 242. С. 23–43.

## SUFFICIENT CONDITION FOR POLYNOMIAL SOLVABILITY OF RANDOM 3-CNF FORMULAS

S. I. Uvarov<sup>a</sup>

Presented by Academician of the RAS S. N. Vassilyev

<sup>a</sup>*V.A. Trapeznikov Institute of Control Sciences of Russian Academy of Sciences, Moscow, Russia*

This paper is devoted to the localisation of random 3-CNF formulas that are polynomially solvable by the resolution algorithm. It is shown that random formulas with the number of clauses proportional to the square of the number of variables, are polynomially solvable with probability close to unity when the proportionality coefficient exceeds the found threshold.

*Keywords:* 3-CNF, clause, resolution, complexity, satisfiability problem