

Сложность вычислений
Лекция 4: Естественные доказательства
(natural proofs)

Дмитрий Ицксон

ПОМИ РАН

3 марта 2011

Complexity theory's Waterloo

Почему не получается доказать $\mathbf{P} \neq \mathbf{NP}$?

- 1 (Baker, Gill, Solovay 1975) Существуют такие оракулы A, B , что $\mathbf{P}^A = \mathbf{NP}^A$ и $\mathbf{P}^B \neq \mathbf{NP}^B$.
 - Техника доказательства $\mathbf{P} \neq \mathbf{NP}$ должна быть нерелятивизируема.
 - Диагонализации не хватает
- 2 (Razborov, Rudich, 1994) Если существуют **естественное** комбинаторное доказательство нетривиальной нижней оценки на схемную сложность явной функции, то не существует субэкспоненциальных односторонних функций.
- 3 (Aaronson, Wigderson, 2007) Алгебраизации не хватает, чтобы доказать $\mathbf{P} \neq \mathbf{NP}$
 - $\mathbf{IP} = \mathbf{PSPACE}$
 - \mathbf{PCP} -теорема.

Почему схемы?

- Схема — это помеченный ориентированный граф.
- Простой комбинаторный объект.
- Комбинаторные методы позволяют доказывать нижние оценки в ограниченных моделях.
- Из нижних оценок на схемы, следуют нижние оценки для алгоритмов.
- Нижние оценки на схемы — понятный комбинаторный путь доказательства $P \neq NP$.

Естественные доказательства

Мы хотим доказать, что функция $f : \{0, 1\}^n \rightarrow \{0, 1\}$ не имеет схем размера n^c . **Метод доказательства** : построение **естественного предиката** \mathcal{P} , заданного на булевых функциях:

- 1 (n^c -полезность) $\mathcal{P}(f) = 1$, но $\mathcal{P}(g) = 0$ для всех $g \in \mathbf{SIZE}(n^c)$.
- 2 (Конструктивность) Предикат \mathcal{P} вычислим за полиномиальное время по таблице истинности функции $g : \{0, 1\}^n \rightarrow \{0, 1\}$. (Т.е., вычислим за время 2^{dn}).
- 3 (Объемность) Для случайной функции $g : \{0, 1\}^n \rightarrow \{0, 1\}$ с вероятностью $\frac{1}{poly(n)}$ выполняется $\mathcal{P}(g) = 1$.

Теорема. (Natural proofs) Предположим, что существуют субэкспоненциальные односторонние функции. Тогда $\exists c > 0$, что не существует n^c -полезных естественных предикатов \mathcal{P} .

Субэкспоненциальные односторонние функции

Определение. Функция $f : \{0, 1\}^* \rightarrow \{0, 1\}^*$ называется **субэкспоненциально** односторонней, если

- f вычислима за полиномиальное время
- f — честная: $|x| \leq \text{poly}(|f(x)|)$
- $\exists \epsilon > 0$, для всех $c > 0$, для любого вероятностного алгоритма B , работающего **полиномиальное время** **время** 2^{n^ϵ} , для всех достаточно больших n выполняется, что

$$\Pr_{x \leftarrow U(\{0,1\}^n)} [B(f(x)) \in f^{-1}(f(x))] \leq \frac{1}{n^c}$$

Пример естественных доказательств

- Мы доказали, что *Parity* нельзя решить полиномиальными по размеру схемами из \neg, \vee, \wedge константной глубины.
- Доказательство проходит для любой функции, которая не обращается в константу при подстановке $n - n^\epsilon$ входных переменных.
- $\mathcal{P}(f) = 1 \iff f$ не обращается в константу при подстановки $n - n^\epsilon$ входных переменных.
- n^ϵ -полезность уже доказывали.
- Конструктивность: по таблице истинности за полиномиальное время проверить \mathcal{P} легко.
- Объемность...

Объемность

- Оценим количество **неустойчивых** функций, которые обращаются в константу от подстановки $n - n^\epsilon$ переменных.
- Всего булевых функций 2^{2^n} : каждая задается таблицей истинности.
- Неустойчивую функцию можно задать так:
 - 1 таблица истинности, из которой выкинуты 2^{n^ϵ} строчек.
 - 2 Номер сочетания (от 1 до $C_n^{n^\epsilon}$)
 - 3 Значение выкинутых строчек
- Доля неустойчивых функций:

$$\frac{C_n^{n^\epsilon} 2^{2^n - 2^{n^\epsilon}} \text{poly}(n)}{2^{2^n}} = \frac{C_n^{n^\epsilon} \text{poly}(n)}{2^{2^{n^\epsilon}}} < 0.01$$

Почему конструктивность?

- Наивный аргумент: предикат f вычисляет \exists — SAT является конструктивным
- Очень часто комбинаторные аргументы конструктивны
- Вероятностный метод:
 - Многие результаты со временем становятся конструктивными, например LLL
 - Конструктивным должен быть предикат

Почему объемность?

- Пусть $f_0 : \{0, 1\}^n \rightarrow \{0, 1\}$ имеет сложность $\geq S$
- Тогда хотя бы половина функций $\{0, 1\}^n \rightarrow \{0, 1\}$ имеют сложность $\geq \frac{S}{2} - 10$.
- Пусть g — случайная функция. Если сложность g и $g \oplus f$ $< \frac{S}{2} - 10$, то сложность $f = g \oplus (g \oplus f)$ менее S .

Доказательство

Теорема. Если существуют субэкспоненциальные односторонние функции, то для некоторого c не существует n^c -полезных естественных предикатов.

Доказательство.

- Из криптографии: \exists односторонние функции $\implies \exists$ семейство псевдослучайных функций.
- Аналогично: \exists субэксп. односторонние функции $\implies \exists$ семейство субэксп. псевдослучайных функций.
- $\{f_s\}_{s \in \{0,1\}^*}$, если $s \in \{0,1\}^m$, то $f_s : \{0,1\}^m \rightarrow \{0,1\}$
 - 1 $f_s(x)$ вычисляется за полиномиальное время
 - 2 $\forall d$ любой вероятностный алгоритм B , работающий 2^{n^c} шагов $|\Pr_{s \leftarrow \{0,1\}^m}[B^{f_s} = 1] - \Pr_{g \leftarrow \text{rand.}}[B^g = 1]| < \frac{1}{m^d}$

Доказательство(продолжение)

- $n = m^{\epsilon/2}$, $g_s : \{0, 1\}^n \rightarrow \{0, 1\}$.
- $g_s(x) = f_s(x0^{m-n})$
- g_s неотличима от случайной за время $2^{m^\epsilon} = 2^{n^2}$
- g_s вычислима за полиномиальное время \implies вычислима схемой размера n^c .
- n^c -полезный естественный предикат \mathcal{P} принимает 0 на функциях, вычисляемых схемами размера n^c и принимает 1 на $\frac{1}{poly}$ случайных функций.

О теореме Разборова-Смоленского

- Хорошо приблизили схему ограниченной глубины многочленом маленькой степени.
- Показали, что многочлены маленькой степени плохо вычисляют MOD_2 .
- Какой естественный предикат?
- $\mathcal{P}(f)$: f плохо вычисляются многочленами малой степени.
- Объемность есть, но конструктивности нет.
- Вспомним идею доказательства...

Вспоминаем идею доказательства...

- $y_1 y_2 \dots y_n$ на $\{-1, 1\}^n \setminus Z$ представили многочленом p маленькой степени d .
- Любую функцию из $\{-1, 1\}^n \setminus Z \rightarrow \{-1, 0, 1\}$ представили в виде суммы “мономов” $p l_1 + l_2$, где $\deg l_1, \deg l_2 \leq \frac{n}{2}$.
- Отсюда $3^{2^n - |Z|} \leq 3^{C_n^1 + C_n^2 + \dots + C_n^{\frac{n}{2} + d}} \leq 3^{\frac{49}{50} 2^n}$.
- $|Z| \geq \frac{1}{50} 2^n$
- А что будет для произвольной функции f ?
- Пусть $f : \{-1, 1\}^n \rightarrow \{-1, 0, 1\}$ задается полилинейным многочленом p_f .
- Попытка 2: \mathcal{P} : любой многочлен q может быть записан в виде $p_f l_1 + l_2$, где $\deg l_1, \deg l_2 \leq \frac{n}{2}$.

Оестествление

- Попытка 2: $\mathcal{P}(f) = 1 \iff$ любой многочлен q может быть записан в виде $p_f l_1 + l_2$, где $\deg l_1, \deg l_2 \leq \frac{n}{2}$.
- L — линейное пространство многочленов степени $< \frac{n}{2}$,
 T — линейное пространство многочленов, все мономы в которых имеют степень $> \frac{n}{2}$ (пусть n — нечетное).
- $L \oplus T$ — все многочлены.
- $\mathcal{P}(f) = 1 \iff$ отображение $\pi_f : L \rightarrow T$, $\pi_f(l) = (p_f l)|_T$ взаимно-однозначное.
- Достаточно проверить, что матрица, соответствующая p_f невырожденная. Конструктивность есть!
- Проблемы с доказательством объемности...

Оестествление: попытка 3

- $\mathcal{P}(f) = 1 \iff \dim(p_f L + L) \geq 2^n(\frac{1}{2} + \epsilon)$
- Если $\epsilon = \frac{1}{2}$, то все как раньше.
- Если $\epsilon > 0$, то как минимум $3^{2^n(\frac{1}{2} + \epsilon) - |Z|}$ функций могут быть представлены многочленом степени $\frac{n}{2} + d$.
- Пусть $\epsilon = \frac{1}{4}$, проверим объемность!
- Покажем, что для любой f , либо $\mathcal{P}(f) = 1$, либо $\mathcal{P}(x_1 \oplus x_2 \oplus \dots \oplus x_n \oplus f) = 1$
- Пусть $\dim(p_f L + L) < \frac{3}{4}2^n$.

$$\begin{aligned} \dim(x_1 \cdots x_n p_f L + L) - \dim L &= \dim(x_1 \cdots x_n L + p_f L) - \dim(p_f L) \geq \\ &\dim(x_1 \cdots x_n L + p_f L + L) - \dim(p_f L + L) \\ &= \dim(T + L) - \dim(p_f L + L) \geq \frac{1}{4}2^n \end{aligned}$$