# Вычислительно трудные задачи и дерандомизация
# Лекция 1: Теорема Разборова

Дмитрий Ицыксон

ПОМИ РАН

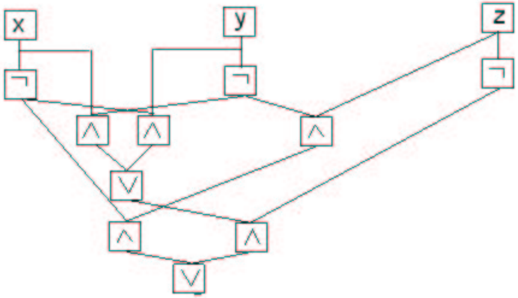
15 февраля 2009

## Содержание курса

- Нижние оценки
  - Теорема Разборова: для монотонных схем
  - Для схем ограниченной глубины
  - Для схем из $Mod_p$ элементов
  - Natural proofs
- Дерандомизация
  - Экспандеры, дисперсеры, экстракторы
  - Уменьшение вероятности ошибки
  - Алгоритм Рейнгольда
- Связь нижних оценок и дерандомизации

Булева схема - это

- Ориентированный граф без циклов
- Ровно одна вершина, из которой не выходит ребер (выход)
- $n$ вершин, в которые не входят ребра
- Все остальные вершины помечены логическими связками $\vee, \wedge, \neg$. (арность связки должна равняться числу входящих ребер)



# Схемная сложность булевых функций

- Размер схемы — это количество вершин в графе, задающем схему.
- Схемная сложность функции — это минимальный размер схемы, вычисляющей эту функцию.
- Схему размера $k$ можно записать с помощью $O(k \log k)$ битов.
- Количество схем размера $2^{\frac{n}{2}}$ не превосходит $2^{O(\frac{n}{2} 2^{n/2})}$.
- Количество булевых функций от $n$ переменных равняется $2^{2^n}$.
- Значит, существует булева функция, сложность которой не менее $2^{\frac{n}{2}}$.
- Открытый вопрос: построить явную функцию большой схемной сложности.

# Монотонные булевы функции и схемы

- $x, y \in \{0,1\}^n$, $x \leq y \iff \forall i, x_i \leq y_i$
- $f : \{0,1\}^n \to \{0,1\}$ монотонная, если $\forall x \leq y, f(x) \leq f(y)$.
- $f$ — монотонная, если при замене 0 на 1 значение $f$ не уменьшается.
- Монотонная схема: все связки $\wedge$ и $\vee$.
- Монотонная схема вычисляет монотонную функцию.
- Любую монотонную функцию (отличную от константы) можно вычислить с помощью монотонных схем.

# Задача о клике

- Граф $G(V, E)$ задан матрицей смежности, $|V| = n$.
- Есть ли в этом графе клика (полный подграф) размера $k$?
- Это **NP**-полная задача
- Монотонная функция
- Индикатор клики: $K \subseteq V$, $\mathcal{C}_K = \bigwedge_{i,j \in K} x_{ij}$
- Монотонная схема для задачи о клике:

$$\bigvee_{K \subset V, |K| = k} \mathcal{C}_K$$

- Размер схемы $C_n^k$.
- Теорема. (Разборов) Монотонная сложность $CLIQUE_{n,k}$ есть $\Omega(2^{\epsilon\sqrt{k}})$ для всех $k \leq n^{\frac{1}{4}}$.

# Схемы из индикаторов

- $S_1, \dots, S_m \subset V$,
$$\bigvee_{i=1}^{m} \mathcal{C}_{S_i},$$
где $\mathcal{C}_K = \bigwedge_{i,j \in K} x_{ij}$
- Ближайшая цель: доказать, что в любой индикаторной схеме обязательно много индикаторов.

# Основные примеры графов

$\mathcal{Y}$-распределение
Выбрать случайное $K \subseteq V$, $|K| = k$, выдать граф, в котором $K$ — клика, а больше ребер нет

$CLIQUE_{n,k}(\mathcal{Y}) = 1$

$\mathcal{N}$-распределение
Раскрасить вершины в $k-1$ цвет случайным образом, ребром соединить вершины разных цветов

$CLIQUE_{n,k}(\mathcal{N}) = 0$

# Поведение индикаторов на основных примерах

Лемма. Для достаточно больших $n$, $k \le n^{1/4}$, $|S| \subseteq V$ либо $\Pr\limits_{G \leftarrow \mathcal{N}}[\mathcal{C}_S(G) = 1] \ge 0.99$, либо $\Pr\limits_{G \leftarrow \mathcal{Y}}[\mathcal{C}_S(G) = 1] \le n^{-\sqrt{k}/20}$.

Доказательство. Пусть $l = \sqrt{k-1}/10$.

- Пусть $|S| \le l$, то $\Pr\limits_{G \leftarrow \mathcal{N}}[\mathcal{C}_S(G) = 1] \ge \frac{(k-1)(k-2)\dots(k-l)}{(k-1)^l} \ge (\frac{k-l}{k-1})^l = (1 - \frac{l-1}{k-1})^l \ge (1 - \frac{1}{10\sqrt{k-1}})^{\sqrt{k-1}/10} \ge 1 - \frac{1}{100}$.
- Пусть $|S| > l$, тогда $\Pr\limits_{G \leftarrow \mathcal{Y}}[\mathcal{C}_S(G) = 1] = \Pr\limits_{K \subseteq V, |K|=k}[S \subseteq K] = \frac{C_{n-|S|}^{k-|S|}}{C_n^k} \le \frac{C_n^{k-|S|}}{C_n^k} \le \frac{C_n^{k-l}}{C_n^k} = \frac{k!(n-k)!}{(k-l)!(n-k+l)!} = \frac{k(k-1)\dots(k-l+1)}{(n-k+1)\dots(n-k+l)} \le (\frac{2k}{n})^l \le \frac{2^l}{n^{0.75l}} \le \frac{1}{n^{0.7l}} < \frac{1}{n^{\sqrt{k}/20}}$.

Следствие. Размер индикаторной схемы для $CLIQUE_{n,k}$ не меньше $n^{\sqrt{k}/20}$.

# Приближение индикаторными схемами

- Пусть схема $C$ размера $s < 2^{\sqrt{k}/100}$ решает $CLIQUE_{n,k}$.
- Будем приближать $C$ идикаторными схемами.
- Пусть $C = f_1, f_2, \dots, f_s$, где $f_k = \left[\begin{array}{l} \text{вход} \\ f_{k'} \vee f_{k''}, \text{ где } k', k'' < k \\ f_{k'} \wedge f_{k''} \end{array}\right.$
- Будем приближать $\tilde{f}_1, \tilde{f}_2, \dots, \tilde{f}_s$, где $\tilde{f}_j$ имеет вид $\bigvee_{i=1}^{m} \mathcal{C}_{S_i}, |S_i| \leq l$.
- $l = \sqrt{k-1}/10, p = 10\sqrt{k}\log n, m = (p-1)^l l!$,
  $m << n^{\sqrt{k}/20}$.
- Функции, которые имеют вид $\bigvee_{i=1}^{m} \mathcal{C}_{S_i}, |S_i| \leq l$, будем называть $(m, l)$-правильными.

# Приближение индикаторными схемами

- Строим $\tilde{f}_k$ индуктивно.
- Если $f_k$ — это вход, то $\tilde{f}_k = f_k$
- Если $f_k = f_{k'} \vee f_{k''}$, то $\tilde{f}_k = \tilde{f}_{k'} \sqcup \tilde{f}_{k''}$.
- Если $f_k = f_{k'} \wedge f_{k''}$, то $\tilde{f}_k = \tilde{f}_{k'} \sqcap \tilde{f}_{k''}$.
- Основные требования на $\sqcup, \sqcap$ для $(m, l)$-правильных $f$ и $g$:

$$\Pr_{G \leftarrow \mathcal{Y}}[f \sqcup g(G) < f \vee g(G)] < \frac{1}{10s} \qquad \Pr_{G \leftarrow \mathcal{N}}[f \sqcup g(G) > f \vee g(G)] < \frac{1}{10s}$$
$$\Pr_{G \leftarrow \mathcal{Y}}[f \sqcap g(G) < f \wedge g(G)] < \frac{1}{10s} \qquad \Pr_{G \leftarrow \mathcal{N}}[f \sqcap g(G) > f \wedge g(G)] < \frac{1}{10s}$$

Итого $\tilde{f}_s$ имеют вид $\bigvee_{i=1}^{m} \mathcal{C}_{S_i}$, где $m << n^{\sqrt{k}/20}$ и $\Pr_{G \leftarrow \mathcal{Y}}[\tilde{f}_s(G) \geq f_s(G)] > 0.9$, $\Pr_{G \leftarrow \mathcal{N}}[\tilde{f}_s(G) \leq f_s(G)] > 0.9$, чего не может быть для индикаторных схем.

## Операция $f \sqcup g$

- $f = \bigvee_{i=1}^{m} C_{S_i}, g = \bigvee_{j=m+1}^{2m} C_{S_j}$
- $h = \bigvee_{i=1}^{2m} C_{S_i}$ — слишком много индикаторов.
- Лемма. (о подсолнухе) Пусть $\mathcal{Z}$ — набор множеств, мощности не более $l$. Если $|\mathcal{Z}| > (p-1)^l l!$, то $\exists Z_1, Z_2, \ldots, Z_p \in \mathcal{Z}$, что $Z_i \cap Z_j = Z$ для всех $1 \leq i, j \leq p$.
- Пока в $h$ есть подсолнух $Z_1, Z_2, \ldots, Z_p$, заменить его на сердцевину $Z$.

# Корректность $f \sqcup g$

1. $\Pr_{G \leftarrow \mathcal{Y}}[f \sqcup g(G) < f \vee g(G)] < \frac{1}{10s}$ — очевидно
2. $\Pr_{G \leftarrow \mathcal{N}}[f \sqcup g(G) > f \vee g(G)] < \frac{1}{10s}$
   - Подсолнух $Z_1, Z_2, \ldots, Z_p$ с сердцевиной $Z$.
   - Рассмотрим распределение $\mathcal{N}$.
   - Событие $B$: все вершины $Z$ покрашены в разный цвет.
   - Событие $A_i$: не все вершины $Z_i$ покрашены в разный цвет.
   - Поскольку $|Z_i| \leq l$, то $\Pr[A_i|B] < \frac{1}{2}$
   - $\Pr[A_1 \wedge A_2 \cdots \wedge A_p \wedge B] \leq \Pr[A_1 \wedge A_2 \cdots \wedge A_p|B] = \prod_{i=1}^{p} \Pr[A_i|B] \leq \frac{1}{2^p} = \frac{1}{n^{10\sqrt{k}}} < \frac{1}{10m^2 s}$ $(s < 2^{\sqrt{k}})$.
   - Было не более $m$ ощипываний подсолнуха.

# Операция $f \sqcap g$

- $f = \bigvee_{i=1}^{m} \mathcal{C}_{S_i}, g = \bigvee_{j=1}^{m} \mathcal{C}_{T_j}$
- $h = \bigvee_{i=1}^{m} \bigvee_{j=1}^{m} \mathcal{C}_{S_i \cup T_j}$ — слишком много индикаторов.
- Выкинем все $\mathcal{C}_S$, если $|S| > l$.
- Пока в $h$ есть подсолнух $Z_1, Z_2, \ldots, Z_p$, заменить его на сердцевину.

# Корректность $f \sqcap g$

1. $\Pr_{G \leftarrow \mathcal{Y}}[f \sqcap g(G) < f \wedge g(G)] < \frac{1}{10s}$ —
   - При ощипывании подсолнуха значение не могло уменьшиться
   - Значит, уменьшилось при выкидывании $\mathcal{C}_S$ при $|S| > l$, но тогда $\Pr_{G \leftarrow \mathcal{Y}}[\mathcal{C}_Z(G) = 1] < \frac{1}{n^{\sqrt{k}/20}} < \frac{1}{10sm^2}$
   - Всего число выкидываний не более $m^2$.
2. $\Pr_{G \leftarrow \mathcal{N}}[f \sqcap g(G) > f \vee g(G)] < \frac{1}{10s}$
   - При удалении множества значение увеличится не могло.
   - Ощипывание подсолнуха оценивали для $f \sqcup g$

# Лемма о подсолнухе

Лемма. (о подсолнухе) Пусть $\mathcal{Z}$ — набор множеств, мощности не более $l$. Если $|\mathcal{Z}| > (p-1)^l l!$, то $\exists Z_1, Z_2, \ldots, Z_p \in \mathcal{Z}$, что $Z_i \cap Z_j = Z$ для всех $1 \le i, j \le p$.

Доказательство.

- Индукция по $l$. База $l = 1$. Все множества различны, значит не пересекаются, можно взять $Z = \emptyset$, $|\mathcal{Z}| \ge p$.
- Переход. Пусть $\mathcal{M}$ — максимальный по включению набор непересекающихся множеств. Для каждого $S \in \mathcal{Z}$ существует $x \in S$, что $x \in \cup\mathcal{M}$. Если $|\mathcal{M}| \ge p$, то $\mathcal{M}$ — подсолнух.
- Значит $|\cup\mathcal{M}| \le (p-1)l$, тогда существует $x \in \cup\mathcal{M}$, который встречается как минимум в $\frac{|\mathcal{Z}|}{l(p-1)}$ множествах $\mathcal{Z}$.
- $S_1, S_2, \ldots, S_t$ содержат этот $x$, $t \ge (p-1)^{l-1}(l-1)!$. Осталось воспользоваться индукционным предположением.