# Коммуникационная сложность. Лекции в Компьютерном клубе ПОМИ. Часть 2.

1–2 апреля 2017

# Недетерминированная сложность

## Определение

Недетерминированный протокол $\pi$ вычисляет предикат $f : X \times Y \to \{0, 1\}$, если $f(x, y) = 1 \Leftrightarrow \mathrm{P}[\pi(x, y) = 1] > 0$.

## Определение

$N^1(f), \quad N^0(f) = N^1(\neg f)$

# Покрытия прямоугольниками

## Определение

$C^1(f), \quad C^0(f)$

## Теорема

$\log C^1(f) \le N^1(f) \le \log C^1(f) + 2,$
$\log C^0(f) \le N^0(f) \le \log C^0(f) + 2.$

## Теорема

$D(f) \le C^1(f) + 1,\ D(f) \le C^0(f) + 1.$

# Примеры

$C^1(EQ) \geq 2^n$ (методом трудных множеств)
$C^0(EQ) \leq 2n$
$C^1(GT) \geq 2^n - 1$ (методом трудных множеств)
$C^0(GT) \geq 2^n$ (методом трудных множеств)
$C^1(DISJ) \geq 2^n$ (методом трудных множеств)
$C^0(DISJ) \leq n$
$C^1(IP) \geq (2^{2n-1} - 2^{n-1})/2^n \approx 2^{n-1}$ (методом размера прямоугольников)
$C^0(IP) \geq (2^{2n-1} + 2^{n-1})/2^n \approx 2^{n-1}$ (методом размера прямоугольников)

# Теоремы Aho-Ullman-Yannakakis, Halstenberg-Reischuk и Разборова

**Теорема (Aho-Ullman-Yannakakis, Halstenberg-Reischuk)**

$D(f) \leq (\log C^1(f) + 1)(\log C^0(f) + 2)$.

**Определение**

$\mathrm{DISJ}_{nk}$ обозначает сужение предиката DISJ на $k$-элементные подмножества $n$-элементного множества.

**Теорема (Разборов)**

$D(DISJ_{nk}) \geq \log \binom{n}{k} \approx k \log n$,
$N^0(DISJ_{nk}) \leq \log n$,
$N^1(DISJ_{nk}) \leq O(k + \log \log n)$.

# Безошибочные вероятностные протоколы и протоколы с односторонней ошибкой

$$N^1(f) \leq R_\varepsilon^1(f) \leq \frac{R_0(f)}{\varepsilon}$$

$$N^0(f) \leq R_\varepsilon^0(f) \leq \frac{R_0(f)}{\varepsilon}$$

$$R_0(f) \leq \frac{R_\varepsilon^1(f) + R_\varepsilon^0(f)}{1-\varepsilon}$$

Теорема (Хостад-Вигдерсон)

$$R_{1/2}^{1,\text{pub}}(DISJ_{nk}) = O(k), \qquad R_{3/4}^{0,\text{pub}}(DISJ_{nk}) = O(k + \log n)$$

# Статистическое расстояние, дивергенция

$P, Q$ — распределения вероятности на $S$.

$$d(P,Q) = \max\{|P(A) - Q(A)| \mid A \subset S\}$$
$$= \sum_{x \in S} |P(x) - Q(x)|/2.$$

$$P\|Q = \sum_{x \in S} P(x) \log(P(x)/Q(x))$$

## Теорема

$P\|Q \geq 0$
$P\|Q = 0 \Leftrightarrow P = Q$

## Теорема (неравенство Пинскера)

$d(P,Q) \leq \sqrt{\frac{\ln 2 (P\|Q)}{2}}$

# Свойства статистического расстояния и дивергенции

$d(A, C) \leq d(A, B) + d(B, C)$ (неравенство треугольника)

$d(A \times C, B \times C) = d(A, B)$ (аналогично для дивергенции)

$d(A \times B, A' \times B') \leq d(A, A') + d(B, B')$ (для дивергенции неравенство становится равенством)

$A\,|_{B=b}$ — случайная величина, распределённая как $A$ при условии $B = b$.

$d(AC, BC) = \mathrm{E}_{c \leftarrow C}\, d(A\,|_{C=c}, B\,|_{C=c})$ (аналогично для дивергенции)

$d(AB, A \times B) = \mathrm{E}_{b \leftarrow B}\, d(A\,|_{B=b}, A)$ (аналогично для дивергенции)

# Энтропия Шеннона

$H(P) = -\sum_{x \in S} P(x) \log P(x)$

## Теорема (Шеннон)

Пусть дано префиксное кодирование $x \mapsto C(x)$. Тогда $\sum_{x \in S} P(x)(\text{длина } C(x)) \geq H(P)$.

- $H(P) \geq \min_{x \in S} \log(1/P(x))$
- $H(P) \leq \log |S|$
- В обоих неравенствах равенство достигается только для равномерного распределения

# Пример простого применения энтропии Шеннона

Для предикатов EQ, GT, DISJ существует детерминированный протокол, который в среднем передаёт константу бит.

## Теорема

Однако, любой протокол вычисления IP передает в среднем не менее $n$ бит.

## Теорема

Хотя, для предикатов EQ, GT, DISJ также существуют распределения вероятностей на входах, для которых любой протокол в среднем передает не менее $n$ бит.

# Количество информации

$A, B$ — случайные величины с совместным распределением

## Определение (количество информации)

$I(A : B) = H(A) + H(B) - H(A, B) = (A, B) \parallel (A \times B)$,
где $A \times B$ — независимые копии $A, B$.

## Теорема

$I(A : B) \geq 0$, то есть $H(A, B) \leq H(A) + H(B)$.
$I(A : B) = 0 \Leftrightarrow A, B$ независимы.

## Теорема

$d((A, B), (A \times B)) \leq \sqrt{\frac{I(A:B) \ln 2}{2}}$.
(Неравенство Пинскера.)

# Условная энтропия Шеннона

## Определение

$H(A \mid B)$ — среднее значение $H(A \mid_{B=y})$:

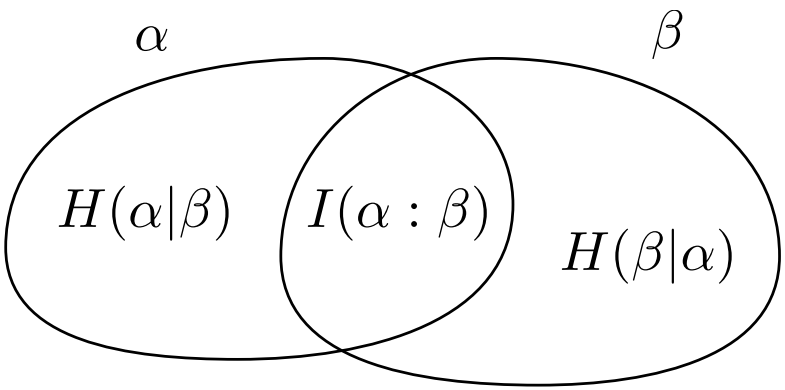
$$\begin{aligned} H(A \mid B) &= \mathrm{E}_{b \leftarrow B}\, H(A \mid_{B=b}) \\ &= H(A, B) - H(B). \end{aligned}$$

Выражение $I(B : A)$ через условную энтропию:

$$I(B : A) = H(A) - H(A \mid B).$$

Поскольку $H(A \mid B) \geq 0$, из этого следует $I(B : A) \leq H(A)$.
А поскольку $I(B : A) \geq 0$ из этого следует $H(A \mid B) \leq H(A)$.

# Энтропии двух случайных величин



# Релятивизация

Частичное усреднение:

$H(A \mid BC) = E_{c \leftarrow C} H(A \mid B, C = c)$.

Релятивизация неравенства

$H(A \mid B) \leq H(A)$

даёт неравенство

$$H(A \mid BC) \leq H(A \mid C)$$

Релятивизация неравенства

$H(AB) = H(A) + H(B \mid A)$

даёт неравенство

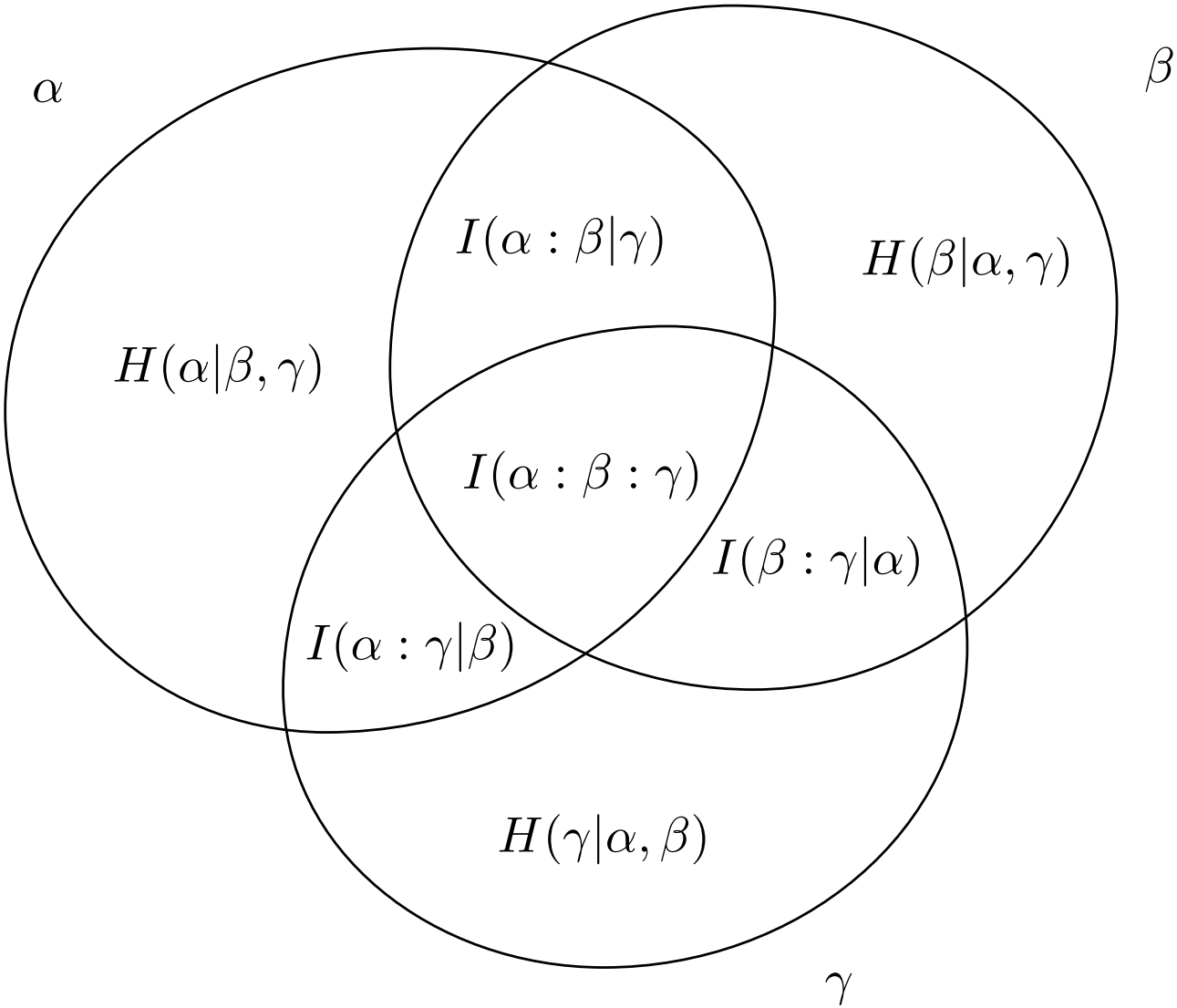
$$H(AB \mid C) = H(A \mid C) + H(B \mid AC)$$

Определение

$$\begin{aligned} I(B : A \mid C) &= H(A \mid C) + H(B \mid C) - H(AB \mid C) \\ &= H(A \mid C) - H(A \mid BC). \end{aligned}$$

$I(A : B \mid CD) = E_{d \leftarrow D} H(A : B \mid C, D = d)$.

# Энтропии трёх случайных величин



# Цепное правило

$$I(A : BC \mid D) = I(A : B \mid D) + I(A : C \mid BD).$$

Получается вычитанием из равенства

$$H(BC \mid D) = H(B \mid D) - H(C \mid BD)$$

его релятивизации

$$H(BC \mid AD) = H(B \mid AD) - H(C \mid ABD)$$

Обобщение цепного правила:

$$I(A : B_1 \dots B_n \mid D) = \sum_{i=1}^{n} I(A : B_i \mid DB_1 \dots B_{i-1}).$$

# Изменение количества общей информации при добавлении условия

$I(A : B)$ vs. $I(A : B \mid C)$

Может быть по-разному:

Пример 1. $A, B$ — случайные независимые биты, $C = A \oplus B$.
$I(A : B) < I(A : B \mid C)$

Пример 2. $A = B = C$ — случайный бит.
$I(A : B) > I(A : B \mid C)$.

Гарантией неравенства $I(A : B) \leq I(A : B \mid C)$ является независимость $A$ или $B$ от $C$:

$$I(A : B) \leq I(A : CB) = I(A : C) + I(A : B \mid C) = I(A : B \mid C).$$

# Информационная сложность коммуникационных протоколов

Неформально, это — сколько бит информации узнают участники о входах друг друга в ходе исполнения протокола.

$$\begin{aligned} IC_\mu(\pi) &= I(X : MR \,|\, Y) + I(Y : MR \,|\, X) \\ &= I(X : M \,|\, YR) + I(Y : M \,|\, XR), \end{aligned}$$

где
$(X, Y)$ — входная пара (с распределением $\mu$),
$R$ — общие случайные биты,
$M$ — коммуникация (последовательность переданных бит).

# Информационная сложность и средняя длина коммуникации

**Теорема**

$$IC_\mu(\pi) \le \text{средняя длина коммуникации } \pi \text{ на входе } x, y \text{ относительно } \mu$$

**Доказательство.**

$$\begin{aligned}
&I(X : M \mid YR) + I(Y : M \mid XR) \\
&= \sum_i \big(I(X : M_i \mid YRM_{<i}) + I(Y : M_i \mid XRM_{<i})\big) \\
&= \sum_i E_u\big(I(X : M_i \mid YRM_{<i} = u) + I(Y : M_i \mid XRM_{<i} = u)\big) \\
&\le \sum_i E_u H(M_i \mid RM_{<i} = u) = \sum_i H(M_i \mid RM_{<i}) \\
&= H(M \mid R) \le \text{средняя длина } M.
\end{aligned}$$

□

# Direct sum problem

$f^n(x_1 \dots x_n, y_1 \dots y_n) = f(x_1, y_1) \dots f(x_n, y_n)$

$\pi$, вычисляющий $f$ $\quad \mapsto \quad$ $\pi'$, вычисляющий $f^n$

- $CC(\pi') = CC(\pi)n$,
- $\mathrm{P}_{(x,y)\leftarrow\mu}[\pi(x,y) \neq f(x,y)] \leq \varepsilon \Longrightarrow$
  $\mathrm{P}_{(x,y)\leftarrow\mu^n}[\pi'(x_1 \dots x_n, y_1 \dots y_n)_i \neq f(x_i, y_i)] \leq \varepsilon.$

Можно ли переделать любой протокол $\pi$, вычисляющий $f^n$ в любой координате на $1-\varepsilon$ доле входных пар относительно $\mu^n$, в протокол $\pi'$ высоты $CC(\pi)/n$, вычисляющий $f$ на $1-\varepsilon$ доле входных пар (относительно $\mu$)?

# Direct sum problem для информационной сложности

## Теорема

$\pi$, вычисляющий $f^n$ $\quad\mapsto\quad$ $\pi'$, вычисляющий $f$:

- $IC_\mu(\pi') = IC_{\mu^n}(\pi)/n$,
- $\mathrm{P}_{(x,y)\leftarrow\mu^n}[\pi(x_1 \dots x_n, y_1 \dots y_n)_i \neq f(x_i, y_i)] \leq \varepsilon \Longrightarrow$ $\mathrm{P}_{(x,y)\leftarrow\mu}[\pi(x, y) \neq f(x, y)] \leq \varepsilon$.

Протокол $\pi'(x, y)$:

1. Выбираем случайное $k \in \{1, \dots, n\}$ (пользуясь общими случайными битами)
2. Выбираем слова $u, v$ длин $k-1$ и $n-k$, соответственно (пользуясь общими случайными битами).
3. Алиса приватно выбирает случайное слово $v'$, а Боб — случайное слово $u'$
4. Запускаем протокол $\pi$ на входах $uxv'$ и $u'yv$.

# Доказательство

Ключевой факт: $X_k$ и $Y_{<k}$ независимы при известном $X_{<k}Y_{\geq k}$
(и аналогично для $Y_k$ и $X_{>k}$)

$$IC_\mu(\pi') = I(X_K : M \mid Y_K X_{<K} Y_{>K} K) + I(Y_K : M \mid X_K X_{<K} Y_{>K} K)$$

$$= (1/n) \sum_{k=1}^{n} \Big( I(X_k : M \mid Y_k X_{<k} Y_{>k}, K = k) + I(Y_k : M \mid X_k X_{<k} Y_{>k}, K =$$

$$= (1/n) \sum_{k=1}^{n} \Big( I(X_k : M \mid X_{<k} Y_{\geq k}) + I(Y_k : M \mid X_{\leq k} Y_{>k}) \Big)$$

$$\leq (1/n) \sum_{k=1}^{n} \Big( I(X_k : M \mid X_{<k} Y) + I(Y_k : M \mid X Y_{>k}) \Big)$$

$$= (1/n) \big( I(X : M \mid Y) + I(Y : M \mid X) \big)$$

# Доказательство

Ключевой факт: $X_k$ и $Y_{<k}$ независимы при известном $X_{<k}Y_{\geq k}$ (и аналогично для $Y_k$ и $X_{>k}$)

$$
\begin{aligned}
IC_\mu(\pi') &= I(X_K : M \mid Y_K X_{<K} Y_{>K} K) + I(Y_K : M \mid X_K X_{<K} Y_{>K} K) \\
&= (1/n) \sum_{k=1}^{n} \Big( I(X_k : M \mid Y_k X_{<k} Y_{>k}) + I(Y_k : M \mid X_k X_{<k} Y_{>k}) \Big), \\
&= (1/n) \sum_{k=1}^{n} \Big( I(X_k : M \mid X_{<k} Y_{\geq k}) + I(Y_k : M \mid X_{\leq k} Y_{>k}) \Big) \\
&\leq (1/n) \sum_{k=1}^{n} \Big( I(X_k : M \mid X_{<k} Y) + I(Y_k : M \mid X Y_{>k}) \Big) \\
&= (1/n) \big( I(X : M \mid Y) + I(Y : M \mid X) \big)
\end{aligned}
$$

# Нижняя оценка вероятностной сложности DISJ

## Теорема

$R^{\text{pub}}_{1/4-\delta}(DISJ) = \Omega(\delta^2 n)$.

Случайные величины $(A, B)$ и $(\tilde{A}, \tilde{B})$:

| $A \backslash B$ | 0 | 1 |
|---|---|---|
| 0 | 1/4 | 1/4 |
| 1 | 1/4 | 1/4 |

| $\tilde{A} \backslash \tilde{B}$ | 0 | 1 |
|---|---|---|
| 0 | 1/3 | 1/3 |
| 1 | 1/3 | 0 |

Трудное распределение на $\{0,1\}^n \times \{0,1\}^n$ для DISJ:

$X = \underbrace{\tilde{A}\tilde{A}\tilde{A}\tilde{A}\tilde{A}A\tilde{A}\tilde{A}\tilde{A}\tilde{A}}_{n}$, $Y = \underbrace{\tilde{B}\tilde{B}\tilde{B}\tilde{B}\tilde{B}B\tilde{B}\tilde{B}\tilde{B}\tilde{B}}_{n}$ (пара $A, B$ стоит на случайно выбранном месте $K$).

$\tilde{X} = \underbrace{\tilde{A}\tilde{A}\tilde{A}\tilde{A}\tilde{A}\tilde{A}\tilde{A}\tilde{A}\tilde{A}\tilde{A}}_{n}$, $\tilde{Y} = \underbrace{\tilde{B}\tilde{B}\tilde{B}\tilde{B}\tilde{B}\tilde{B}\tilde{B}\tilde{B}\tilde{B}\tilde{B}}_{n}$.

# Доказательство

Пусть $\pi$ вычисляет $DISJ(X, Y) = A \wedge B$ с ошибкой $1/4 - \delta$.

$\pi \mapsto \pi'$ (как в Direct sum problem):

- $\pi'$ вычисляет $A \wedge B$ с ошибкой $1/4 - \delta$
- $IC_{\tilde{A},\tilde{B}}(\pi') \leq IC_{\tilde{X},\tilde{Y}}(\pi)/n \leq CC(\pi)/n$

# Доказательство

## Лемма

$\mathrm{P}[\pi'(A,B) = A \wedge B] \le 3/4 + O\left(\sqrt{IC_{\tilde{A},\tilde{B}}(\pi')}\right)$

*Доказательство леммы для протоколов без общих случайных битов:*

## Определение

$M$ — коммуникация $\pi'$ на входе $A, B$.

$$\begin{aligned}\mathrm{P}[\pi'(A,B) = A \wedge B] &= \mathrm{E}_{l \leftarrow M}\, \mathrm{P}[(\text{пометка } l) \;= A \wedge B \mid M = l] \\ &\le \mathrm{E}_{l \leftarrow M}\left(3/4 + d(AB, (AB \mid_{M=l}))\right).\end{aligned}$$

Ключевой момент: $A$ и $B$ независимы и также независимы при известном $M$ (свойство прямоугольников) $\Rightarrow$

$$\begin{aligned}d(AB, (AB \mid_{M=l})) &= d(A \times B, (A \mid_{M=l}) \times (B \mid_{M=l})) \\ &\le d(A, (A \mid_{M=l})) + d(B, (B \mid_{M=l})).\end{aligned}$$

# Доказательство

Усредним оба слагаемых отдельно:

$$\mathrm{E}_{I \leftarrow M} d(A, (A \mid_{M=I})) = d(A \times M, AM) \leq \sqrt{I(A : M)},$$
$$\mathrm{E}_{I \leftarrow M} d(B, (B \mid_{M=I})) = d(B \times M, BM) \leq \sqrt{I(B : M)}.$$

Итак,

$$\mathrm{P}[\pi'(A, B) = A \wedge B] \leq 3/4 + \sqrt{I(A : M)} + \sqrt{I(B : M)}.$$

Теперь надо как-то связать $I(A : M)$ с $I(\tilde{A} : \tilde{M}|\tilde{B})$, а $I(B : M)$ с $I(\tilde{B} : \tilde{M}|\tilde{A})$.

# Доказательство

$$
\begin{aligned}
I(\tilde{A} : \tilde{M} \,|\, \tilde{B}) &= I(\tilde{A} : \tilde{M} \,|\, \tilde{B} = 0)\mathrm{P}[\tilde{B} = 0] + I(\tilde{A} : \tilde{M} \,|\, \tilde{B} = 1)\mathrm{P}[\tilde{B} = 1] \\
&= I(\tilde{A} : \tilde{M} \,|\, \tilde{B} = 0)(2/3) \\
&= I(A : M \,|\, B = 0)(2/3)
\end{aligned}
$$

$$
\begin{aligned}
I(A : M \,|\, B = 0) &= H(A \,|\, B = 0) - H(A \,|\, M, B = 0) \\
&= H(A) - H(A \,|\, M) = I(A : M).
\end{aligned}
$$

# Доказательство леммы для протоколов с общими случайными битами:

При фиксации случайных битов получается детерминированный протокол, к которому можно применить доказанное неравенство, затем усреднить и воспользоваться выпуклостью корня:

$$\begin{aligned}\mathrm{P}[\pi'(A,B) = A \wedge B] &= \mathrm{E}_r\, \mathrm{P}[\pi'_r(A,B) = A \wedge B] \\ &\leq 3/4 + \mathrm{E}_r \sqrt{IC_{\tilde{A},\tilde{B}}(\pi'_r)} \\ &\leq 3/4 + \sqrt{\mathrm{E}_r\, IC_{\tilde{A},\tilde{B}}(\pi'_r)} = 3/4 + \sqrt{IC_{\tilde{A},\tilde{B}}(\pi')}.\end{aligned}$$

# Улучшенная нижняя оценка вероятностной сложности DISJ

## Теорема

$R^{\mathsf{pub}}_{1/2-\delta}(DISJ) \geq c\delta^2 n$,

где $c = (4+\sqrt{2})/(7\ln 2) \approx 1.116$.

| $A \backslash B$ | 0 | 1 |
|---|---|---|
| 0 | $(1-\sqrt{1/2})^2$ | $\sqrt{1/2}(1-\sqrt{1/2})$ |
| 1 | $\sqrt{1/2}(1-\sqrt{1/2})$ | $1/2$ |

| $\tilde{A} \backslash \tilde{B}$ | 0 | 1 |
|---|---|---|
| 0 | $d(1-\sqrt{1/2})$ | $d\sqrt{1/2}$ |
| 1 | $d\sqrt{1/2}$ | 0 |

где $d = 2-\sqrt{2}$.

# Сжатие коммуникационных протоколов

Протокол с малой информационной сложностью $\mapsto$ протокол с небольшой средней коммуникацией (для вычисления той же функции $f$)

## Теорема

$\mu$ — распределение на её входах, $\pi$ — вероятностный протокол высоты $h$, $\varepsilon > 0$

$\mapsto$

вероятностный протокол $\pi'$ такой что

- средняя длина коммуникации $= O\left((\sqrt{IC_\mu(\pi)h} + 2)\log(h/\varepsilon)\right)$
- на любой входной паре вероятность правильного ответа не меньше $(1-\varepsilon)$(вероятность правильного ответа $\pi$)

# Построение протокола $\pi'$

Алиса и Боб мысленно двигают фишку, используя в вершине $v$ на высоте $i$ вместо правильного распределения $M_i \mid_{M_{<i}=v, X=x, Y=y}$ распределение $M_i \mid_{M_{<i}=v, X=x}$ (Алиса) и $M_i \mid_{M_{<i}=v, X=x}$ (Боб).

Перед тем, как двигать фишку в самом деле, они запускают $LCP(\varepsilon/h^2)$ для поиска первого расхождения их мысленных путей. Затем они двигают фишку согласованно до первого расхождения плюс один шаг.

# Анализ ошибки протокола $\pi'$

На любой входной паре:
Вероятность правильной работы $\pi$
= (вероятность того, что ни один из $h$ вызовов LCP не выдал ошибки)(вероятность правильной работы $\pi$)
$\geq (1 - h(\varepsilon/h^2))$ (вероятность правильной работы $\pi$)
$\geq (1 - \varepsilon)$ (вероятность правильной работы $\pi$)

# Анализ средней коммуникации протокола $\pi'$

Допустим, что протокол LCP никогда не дает ошибки.

Среднее количество вызовов LCP

$$= 1 + \sum_{i=1}^{h} \mathrm{E}_{m \leftarrow M} \mathrm{E}_{x,y}\, d(M_i \mid_{M_{<i}=m_{<i}, X=x}, M_i \mid_{M_{<i}=m_{<i}, Y=y}),$$

где

$$d(M_i \mid_{M_{<i}=m_{<i}, X=x}, M_i \mid_{M_{<i}=m_{<i}, Y=y})$$

$$= \begin{cases} d(M_i \mid_{M_{<i}=m_{<i}, X=x, Y=y}, M_i \mid_{M_{<i}=m_{<i}, Y=y}) & \text{если ход Алисы} \\ d(M_i \mid_{M_{<i}=m_{<i}, X=x}, M_i \mid_{M_{<i}=m_{<i}, X=x, Y=y}) & \text{если ход Боба} \end{cases}$$

Если ход Алисы, то усредняя по $x$, получим:

$$\leq \sqrt{I(M_i : X \mid M_{<i} = m_{<i}, Y = y)}.$$

Усредняя затем по $y$, получим

$$\leq \mathrm{E}_y \sqrt{I(M_i : X \mid M_{<i} = m_{<i}, Y = y)}$$

$$\leq \sqrt{\mathrm{E}_y\, I(M_i : X \mid M_{<i} = m_{<i}, Y = y)}$$

# Итак, среднее количество вызовов LCP

$$\leq 1 + \sum_{i=1}^{h} \mathrm{E}_{m \leftarrow M} \Big( \sqrt{I(M_i : X \mid M_{<i} = m_{<i}, Y)} + \sqrt{I(M_i : Y \mid M_{<i} = m_{<}$$

$$\leq 1 + \sum_{i=1}^{h} \mathrm{E}_m \sqrt{I(M_i : X \mid M_{<i} = m_{<i}, Y) + I(M_i : Y \mid M_{<i} = m_{<i}, X)}$$

$$\leq 1 + \sum_{i=1}^{h} \sqrt{\mathrm{E}_m \left( I(M_i : X \mid M_{<i} = m_{<i}, Y) + I(M_i : Y \mid M_{<i} = m_{<i}, X) \right)}$$

$$\leq 1 + \sum_{i=1}^{h} \sqrt{I(M_i : X \mid M_{<i}, Y) + I(M_i : Y \mid M_{<i}, X)}$$

$$\leq 1 + \sqrt{h \sum_{i=1}^{h} (I(M_i : X \mid M_{<i}, Y) + I(M_i : Y \mid M_{<i}, X))}$$

$$= 1 + \sqrt{h(I(M : X \mid Y) + I(M : Y \mid X))}.$$