Нижняя оценка вероятностной сложности предиката DISJ

3 апреля 2017 г.

Предикат DISJ определен на парах n-битовых строк x, y формулой $\bigvee_{i=1}^{n}(x_i \wedge y_i)$.

Theorem 1. *Любой вероятностный протокол с общими случайными битами, вычисляющий предикат DISJ с вероятностью ошибки не более $1/2-\delta \leqslant 1/2$ (на любом входе), передает в худшем случае не менее $c\delta^2 n$ бит, где $c = (2-\sqrt{2})/\ln 2 \approx 0.845$.*

Доказательство. Пусть дан вероятностный протокол, вычисляющий DISJ с ошибкой не более $1/2-\delta$, которые передаёт в худшем случае не более l бит. Нам нужно показать, что $l/n \geqslant c\delta^2$. Доказательство проходит в три этапа.

На первом этапе мы пользуемся трюком Яо и определяем «трудное» распределение на парах двоичных слов длины n. По теореме Яо существует детерминированный протокол высоты l, вычисляющий предикат DISJ на не менее чем $1/2+\delta$ всех входов относительно этого распределения. На втором этапе мы строим вероятностный протокол с общими и частными случайными битами, который вычисляет конъюнкция двух битов a, b с вероятностью ошибки не более $1/2-\delta$ и информационной сложностью не более l/n. При этом при подсчете вероятности ошибки и информационной сложности мы будем использовать разные, но похожие, распределения на входах. На третьем этапе мы доказываем, что информационная сложность такого протокола не может быть меньше $c\delta^2$.

Сначала объясним, какие распределения на парах битов мы используем. Случайную величину с первым распределением будем называть

(A, B), а случайную величину со вторым распределением — $(\tilde{A}, \tilde{B})$. Нам нужно, чтобы эти случайные величины имели следующие свойства:

- A и B независимы;
- $P[A \wedge B = 1] = 1/2$;
- $P[\tilde{A} \wedge \tilde{B} = 1] = 0$;
- $P[\tilde{A} = 0] = P[\tilde{B} = 0] > 0$;
- распределения случайных величин $A \,|_{B=0}$ и $\tilde{A} \,|_{\tilde{B}=0}$ совпадают;
- распределения случайных величин $B \,|_{A=0}$ и $\tilde{B} \,|_{\tilde{A}=0}$ совпадают.

Для получения какой-то линейной нижней оценки необязательно, чтобы в четвертом пункте вероятности были равны — достаточно, чтобы они обе были положительными (иначе бессмысленно говорить об условных распределениях). Также, необязательно, чтобы во втором пункте была $1/2$; если во втором пункте вместо $1/2$ стоит некоторое положительное число p, то этим методом мы сможем докажем линейную нижнюю оценку для протоколов с вероятностью ошибки не более $p - \delta$ для любого δ, например, для $\delta = p/2$. Из этого с помощью amplification нетрудно вывести некоторую линейную оценку для протоколов с вероятностью ошибки сколь угодно близкой к $1/2$. Например, можно положить распределение пары (A, B) равномерным, а распределение пары $(\tilde{A}, \tilde{B})$ — равномерным на множестве из трех пар $(0, 0)$, $(0, 1)$ и $(1, 0)$. Тогда второй пункт выполнен для $p = 1/4$ и мы докажем теорему для $1/4$ вместо $1/2$. А отсюда выведем теорему для $1/2 - \delta$, но с значительно меньшей константой c.

Но нетрудно построить и распределения (A, B), $(\tilde{A}, \tilde{B})$ с указанными свойствами (на самом деле эти свойства однозначно определяют распределения). В самом деле, возьмём независимые A, B с

$$P[A = 1] = P[B = 1] = \sqrt{1/2}.$$

Это гарантирует первые два требования. Затем определим распределение $(\tilde{A}, \tilde{B})$, так чтобы вероятности пар $(0, 0)$, $(0, 1)$ и $(1, 0)$ были пропорциональны $\sqrt{1/2}$, $1 - \sqrt{1/2}$ и $1 - \sqrt{1/2}$, а коэффициент пропорциональности выбран так, чтобы сумма вероятностей этих трех пар была равна

1. Это гарантирует выполнение остальных четырех требований. Можно подсчитать, что коэффициент пропорциональности равен $2-\sqrt{2}$, а значит $P[\tilde{A}=0]=P[\tilde{B}=0]=2-\sqrt{2}$.

Первый этап: построение трудного распределения вероятностей на парах двоичных слов длины n. Выбираем случайное $i \in \{1,\dots,n\}$ и выбираем пару битов (x_i, y_i) с распределением, как у пары (A,B), а остальные пары битов (x_j, y_j) — с распределением, как у пары $(\tilde{A},\tilde{B})$, независимо от пары (x_i, y_i) и друг от друга. Получившуюся пару (зависимых) случайных величин $(x_1 \dots x_n, y_1 \dots y_n)$ будем обозначать через (X,Y). А случайную выбранное i будем обозначать через I. Мы определили совместно распределенные случайные величины X, Y, I.

По лемме Яо существует детерминированный протокол π высоты l, вычисляющий DISJ с вероятностью ошибки не более $1/2-\delta$ на случайной паре слов (X,Y).

Второй этап: построение вероятностного протокола вычисления конъюнкции двух битов. Этот протокол будет получен из протокола π, построенного на первом этапе. Пользуясь общим источником случайности, Алиса и Боб выбирают случайное $i \in \{1,\dots,n\}$ (с равномерным распределением) и случайные двоичные слова u, v длин $i-1$ и $n-i$, соответственно. Затем Алиса приватно выбирает случайное слово v', а Боб — случайное слово u', длин $n-i$ и $i-1$, соответственно, и они запускают протокол π на словах, $x = uav'$ и $y = u'bv$, где a, b — данные им входные биты.

Пусть входные биты a, b распределены как A, B. Слова u, v, u', v' надо выбирать так, чтобы полученное распределение на парах x, y было, как у пары X, Y. Для этого u, v выбираются по следующему распределению: каждый бит слова u слов имеет то же, распределение, что $\tilde{A}$, и разные биты независимы, каждый бит слова v слов имеет то же, распределение, что $\tilde{B}$, и разные биты независимы. Бит с номером j слова v' выбирается независимо от других битов по распределению $\tilde{A} \,|\, _{\tilde{B}=u_j}$, а бит с номером j слова u' — независимо от других битов по распределению $\tilde{B} \,|\, _{\tilde{A}=v_j}$.

Обозначим построенный протокол через π'. По построению распределения (X,Y,I) значение предиката DISJ на паре X, Y всегда совпадает с $X_I \wedge Y_I$. Поэтому вероятность ошибки протокола π' на случайной паре A, B совпадает с вероятностью ошибки π на случайной паре X, Y, и значит не превосходит $1/2-\delta$. Докажем, что информационная сложность π' относительно распределения $(\tilde{A},\tilde{B})$ не больше l/n.

Обозначим через $\tilde{X}, \tilde{Y}$ пару случайных величин со следующим распределением. Случайная величина $(\tilde{X}, \tilde{Y})$ определяется так же, как (X, Y), с тем лишь различием, что *все пары* (x_j, y_j) выбираются независимо друг от друга с распределением $(\tilde{A}, \tilde{B})$. Предикат DISJ на паре $(\tilde{X}, \tilde{Y})$ всегда равен 0. Будем обозначать через $\tilde{M}$ случайную величину, равную сообщению в протоколе π на случайном входе $\tilde{X}, \tilde{Y}$.

По построению информационная сложность π' равна среднему по i значению случайной величины

$$I_{\tilde{A},\tilde{B}}(\pi') = I(\tilde{X} : \tilde{M} \,|\, \tilde{Y}_i, \tilde{X}_{<i}, \tilde{Y}_{>i}) + I(\tilde{Y} : \tilde{M} \,|\, \tilde{X}_i, \tilde{X}_{<i}, \tilde{Y}_{>i}).$$

Здесь через $x_{<i}$ обозначено начало x длины $i-1$ и аналогично понимается $y_{>i}$.

Рассмотрим первое слагаемое этой суммы. Его среднее значение равно

$$(1/n) \sum_{i=1}^{n} I(\tilde{X}_i : \tilde{M} \,|\, \tilde{X}_{<i} \tilde{Y}_{\geqslant i}).$$

Мы утверждаем, что сумма в этой формуле не превосходит $I(\tilde{X} : \tilde{M} \,|\, \tilde{Y})$. Чтобы это увидеть, применим к последней величине цепное правило:

$$I(\tilde{X} : \tilde{M} \,|\, \tilde{Y}) = \sum_i I(\tilde{X}_i : \tilde{M} \,|\, \tilde{X}_{<i} \tilde{Y}).$$

В этой сумме i-ое слагаемое можно переписать, как

$$I(\tilde{X}_i : \tilde{M} \tilde{Y}_{<i} \,|\, \tilde{X}_{<i} \tilde{Y}_{\geqslant i}),$$

поскольку $\tilde{X}_i$ и $\tilde{Y}_{<i}$ независимы при известном $\tilde{X}_{<i} \tilde{Y}_{\geqslant i}$. (Мы используем общее цепное правило $I(P : QR \,|\, S) = I(P : R \,|\, S) + I(P : Q \,|\, SR)$. В нашем случае $I(P : R \,|\, S) = I(\tilde{X}_i : \tilde{Y}_{<i} \,|\, \tilde{X}_{<i} \tilde{Y}_{\geqslant i}) = 0$, а $Q = \tilde{M}$.) Наконец, последнее не меньше, чем

$$I(\tilde{X}_i : \tilde{M} \,|\, \tilde{X}_{<i} \tilde{Y}_{\geqslant i}).$$

Таким образом

$$\sum_{i=1}^{n} I(\tilde{X}_i : \tilde{M} \,|\, \tilde{X}_{<i} \tilde{Y}_{\geqslant i}) \leqslant I(\tilde{X} : \tilde{M} \,|\, \tilde{Y}).$$

Аналогичным образом можно доказать и неравенство

$$\sum_{i=1}^{n} I(\tilde{Y}_i : \tilde{M} \mid \tilde{X}_{\leqslant i}\tilde{Y}_{>i}) \leqslant I(\tilde{Y} : \tilde{M} \mid \tilde{X}).$$

Для этого нужно применять цепное правило в другом порядке. Поэтому $I_{\tilde{A},\tilde{B}}(\pi')$ не больше $I_{\tilde{X},\tilde{Y}}(\pi)/n \leqslant l/n$.

Третий этап. Итак, мы построили вероятностный протокол с общими и частными случайными битами. Будем обозначать его той же буквой π, что и исходный протокол. Это не вызовет путаницы, поскольку про исходный протокол можно забыть. В протоколе π Алиса и Боб получают по биту, причем протокол выдает правильный ответ на случайных битах A, B с вероятностью не менее $1/2 - \delta$:

$$P[\pi(A, B) = A \wedge B] \geqslant 1/2 + \delta,$$

а его разглашение на случайных битах $\tilde{A}, \tilde{B}$ не больше l/n:

$$I_{\tilde{A}\tilde{B}}(\pi) = I(\tilde{A} : \tilde{M} \mid \tilde{B}, R) + I(\tilde{B} : \tilde{M} \mid \tilde{A}, R) \leqslant l/n.$$

Здесь R обозначает общие случайные биты, а $\tilde{M}$ — случайную величину равную сообщениям протокола на входе $\tilde{A}, \tilde{B}$ с общими случайными битами R (поскольку еще имеются частные случайные биты, $\tilde{M}$ не является функцией от $\tilde{A}, \tilde{B}, R$).

Чтобы завершить доказательство, нам надо связать вероятность ошибки π на входе (A, B) с информационной сложностью π относительно $\tilde{A}, \tilde{B}$, доказав неравенство

$$P[\pi(A, B) = A \wedge B] \leqslant 1/2 + \sqrt{I_{\tilde{A}\tilde{B}}(\pi)/d}.$$

где $d = (2 - \sqrt{2})/\ln 2$. Это неравенство верно для любых протоколов с общими и частными случайными битами. Достаточно доказать его только для протоколов без общих случайных битов и затем воспользоваться вогнутостью корня. В самом деле, при фиксации $R = r$ протокол π превращается в протокол π_r без общих случайных битов. Вероятность ошибки π есть средняя по r вероятность ошибки протокола π_r. Аналогично, информационная сложность π есть средняя информационная сложность

π_r. Поэтому, если неравенство верно для любого r, то

$$\begin{aligned} P[\pi(A,B) = A \wedge B] \leqslant & 1/2 + \mathop{E}_r \sqrt{I_{\tilde{A}\tilde{B}}(\pi_r)/d} \\ \leqslant & 1/2 + \sqrt{\mathop{E}_r I_{\tilde{A}\tilde{B}}(\pi_r)/d} \\ = & 1/2 + \sqrt{I_{\tilde{A}\tilde{B}}(\pi)/d} \end{aligned}$$

(второе неравенство выполнено из-за вогнутости квадратного корня).

Итак, будем считать, что общих случайных бит в протоколе π нет. Поскольку при условии $\tilde{B} = 1$ случайная величина $\tilde{A}$ становится детерминированной (и наоборот),

$$I_{\tilde{A}\tilde{B}}(\pi) = I(\tilde{A} : \tilde{M} \mid \tilde{B} = 0) \cdot P[\tilde{A} = 0] + I(\tilde{B} : \tilde{M} \mid \tilde{A} = 0) \cdot P[\tilde{B} = 0]$$

Вспомним, что $P[\tilde{A} = 0] = P[\tilde{B} = 0] = (2 - \sqrt{2})$, поэтому

$$I(\tilde{A} : \tilde{M} \mid \tilde{B} = 0) + I(\tilde{B} : \tilde{M} \mid \tilde{A} = 0) = I_{\tilde{A}\tilde{B}}(\pi)/(2 - \sqrt{2}).$$

Теперь воспользуемся тем, что распределение $A \mid_{B=0}$ совпадает с распределением $\tilde{A} \mid_{\tilde{B}=0}$. Поскольку распределение на случайных битах в протоколе вообще не зависит от входов Алисы и Боба, отсюда следует, что и распределение $(A, M) \mid_{B=0}$ совпадает с распределением $(\tilde{A}, \tilde{M}) \mid_{\tilde{B}=0}$, где M обозначает сообщения посылаемые протоколом на входе (A, B). Следовательно, мы можем везде выбросить тильду и заключить, что

$$I(A : M \mid B = 0) + I(B : M \mid A = 0) = I_{\tilde{A}\tilde{B}}(\pi)/(2 - \sqrt{2}).$$

Теперь воспользуемся тем, что случайные величины A и B независимы. Более того, они независимы при известном M. В самом деле, зафиксируем некоторое значение m) случайной величины M. Этому значению соответствует некоторый лист протокола π, а ему — некоторый прямоугольник $U \times V$, состоящий из пар $\langle (a, r_{\text{Alice}}), (b, r_{\text{Bob}}) \rangle$. Условие $M = m$ означает, что $\langle (A, R_{\text{Alice}}), (B, R_{\text{Bob}}) \rangle \in U \times V$, то есть, $(A, R_{\text{Alice}}) \in U$ и $(B, R_{\text{Bob}}) \in V$. Поэтому A, B остаются независимыми при условии $M = m$. В самом деле, для любых исходов a, b случайных величин A, B выполнено

$$\begin{aligned} P[A = a \mid B = b, M = m] &= P[A = a \mid B = b, (A, R_{\text{Alice}}) \in U), (B, R_{\text{Bob}}) \in V] \\ &= P[A = a \mid (A, R_{\text{Alice}}) \in U] \\ &= P[A = a \mid (A, R_{\text{Alice}}) \in U, (B, R_{\text{Bob}}) \in V] \\ &= P[A = a \mid M = m]. \end{aligned}$$

Второе и третье равенства выполнены по одной и той же причине: пары (A, R_{Alice}) и (B, R_{Bob}) независимы.

Теперь объясним простую вещь: почему информационное разглашение π не может быть нулевым. Допустим, что

$$I(A : M \mid B = 0) = I(B : M \mid A = 0) = 0,$$

то есть, A и M независимы при условии $B = 0$, а B и M независимы при условии $A = 0$. Докажем, что отсюда следует, что (A, B) и M независимы.

В силу первого равенства для любых m распределение $A \mid_{M=m,B=0}$ такое же, как распределение $A \mid_{B=0}$, то есть, такое же, как распределение A (по построению A и B независимы). В силу второго равенства распределение $B \mid_{M=m,A=0}$ совпадает с распределением B. В силу независимости A и B при известном M, мы получаем, что распределение пары $(A, B) \mid_{M=m}$ то же, что и у пары (A, B).

Вероятность ошибки протокола есть среднее значение по m вклада каждого листа m, равного

$$P[A \wedge B \neq \text{ пометка листа } m \mid M = m].$$

При этом вклад всех пар одинаков и равен $1/2$, поскольку по построению $P[A \wedge B = 1] = 1/2$ и (A, B) независимо с M. Таким образом, вероятность ошибки равна $1/2$, а значит $\delta = 0$ (и доказываемое неравенство тривиально выполнено).

Теперь нам надо «по непрерывности» продолжить это рассуждение на случай произвольного маленького разглашения. Вместо независимости A и M при известном $B = 0$ и независимости B и M при известном $A = 0$ нам теперь известно лишь, что

$$I(A : M \mid B = 0) + I(B : M \mid A = 0)$$

мало.

Фиксируем произвольное m и рассмотрим статистическое расстояние между распределениями $A \mid_{M=m}$ и A. Обозначим его через d_m. Мы утверждаем, что среднее значение d_m не превосходит

$$\sqrt{I(A : M \mid B = 0) \ln 2/2}$$

В самом деле, поскольку A и B независимы при известном M, распределение $A \mid_{M=m}$ совпадает с распределением $A \mid_{B=0,M=m}$. По неравенству

Пинскера в среднем статистическое расстояние от последнего распределения до распределения $A \mid_{B=0}$ не превосходит этого корня, а последнее распределение совпадает с распределением A. Аналогичным образом доказывается, что в среднем статистическое расстояние между распределениями $B \mid_{M=m}$ и B (обозначим его через e_m) не больше

$$\sqrt{I(B : M \mid A = 0) \ln 2/2}$$

Из вогнутости квадратного корня следует, что сумма этих двух корней не больше чем

$$\sqrt{\ln 2(I(A : M \mid B = 0) + I(B : M \mid A = 0))}.$$

Фиксируем опять любое m. Мы уже видели, что если $d_m + e_m = 0$, то распределения $(A, B) \mid_{M=m}$ и (A, B) совпадают, а поэтому вероятность ошибки π при условии $M = m$ равна $1/2$. Теперь мы знаем лишь, что статистические расстояния между распределениями $A \mid_{M=m}$ и A и между $B \mid_{M=m}$ и B не превосходят d_m, e_m, соответственно. Кроме того, A, B независимы при условии $M = m$. Отсюда следует, что статистическое расстояние между распределениями $(A, B) \mid_{M=m}$ и (A, B) не больше $d_m + e_m$, а значит вероятность правильного ответа π при условии $M = m$ не больше $1/2 + d_m + e_m$. Усредняя эту верхнюю оценку по m мы получим, что вероятность правильного ответа π не больше $1/2$ плюс среднее значение суммы $d_m + e_m$. Про это среднее мы уже установили, что оно не больше

$$\sqrt{\ln 2(I(A : M \mid B = 0) + I(B : M \mid A = 0))}.$$

Таким образом,

$$P[\pi(A, B) = A \wedge B] \leqslant \sqrt{\ln 2(I(A : M \mid B = 0) + I(B : M \mid A = 0))},$$

что и требовалось доказать. □