2 марта 2012 года

Коммуникационная сложность

Рассмотрим следующую модель вычислений: Алиса и Боб хотят совместно найти значение функции $f\colon X \times Y \to Z$, где $X,\ Y,\ Z$ – конечные множества. При этом Алисе известно только значение первого аргумента $x \in X$, а Бобу только значение второго аргумента $y \in Y$. Поскольку и Алиса, и Боб знают значение только одного из двух аргументов функции, для получения значение $f(x,y)$ им приходится обмениваться данными. Предполагается, что имеется канал связи, по которому они могут посылать друг другу информацию (последовательности битов).

Обмен информацией между Алисой и Бобом должен происходить по заранее фиксированному набору правил, называемому *протоколом.* Будем считать, что в ходе выполнения протокола Алиса и Боб по очереди посылают друг другу однобитовые сообщения — каждое сообщение это 0 или 1. В протоколе указывается, кто посылает первое сообщение, кто посылает второе сообщение (в зависимости от значения бита, переданного на первом шаге), и т.д. Кроме того, протокол определяет, какое именно сообщение посылает Алиса или Боб на каждом шаге (в зависимости от их входных данных Алисы или Боба соответственно, а также от сообщений, уже переданных к этому моменту). Наконец, в конце обмена сообщениями протокол должен определять, чему равно значение функции $f(x,y)$. Мы будем требовать, чтобы значение функции однозначно определялось по последовательности битов, переданных Алисой и Бобом друг другом. (Это означает, что сторонний наблюдатель, знающий протокол, но не знающий ни x, ни y, может узнать $f(x,y)$, наблюдая за обменом сообщениями между Алисой и Бобом).

Общее количество битов, которыми обмениваются Алиса и Боб при выполнении протокола, может зависеть от x и y. Мы называем *сложностью* протокола максимальное число битов, которые пересылают друг другу Алиса и Боб (максимум среди всех пар входов $(x,y) \in X \times Y$). При этом мы никак не учитываем сложность вычислений, которые должны проделать Алиса и Боб, следуя протоколу — подсчитывается только число битов, переданных по каналу связи. *Коммуникационной сложностью* функции f называется минимальная сложность протокола, вычисляющего f. Коммуникационная сложность функции f обозначается $\mathrm{cc}(f)$.

Данное выше определение имеет наглядную геометрическую интерпретацию. Коммуникационный протокол удобно представлять себе в виду двоичного дерева с корнем. Каждая внутренняя вершина дерева помечена буквой A или B (ход Алисы или ход Боба). Из каждой внутренней вершины дерева выходят по два ребра (по направлению от корня к листьям); одно из этих рёбер помечается нулём, а другое единицей. Далее, каждой вершине v_i с пометкой A приписывается функция $\phi_i : A \to \{0,1\}$, а каждой вершине v_j с пометкой B приписывается функция $\psi_j : B \to \{0,1\}$. Наконец, каждому листу дерева приписывается некоторый элемент из Z. Выполнение протокола соответствует движению по дереву от корня к листу. На очередном шаге (в вершине v_s) вычисляется значение соответствующей функции: если вершина помечена буквой A, то Алиса вычисляет $\phi_s(x)$; если же вершина помечена буквой B, то Боб вычисляет $\psi_s(y)$. Вычислив очередное значение $\phi_s(x)$ (или $\psi_s(y)$), Алиса (соответственно, Боб) пересылает его своему партнёру. При этом мы перемещаемся в следующую вершину ниже по дереву (в левого или правого потомка текущей вершины, в зависимости от вычисленного значения ϕ_i или ψ_j). Когда мы попадаем в лист дерева, мы объявляем приписанное данному листу $z \in Z$ значением функции, и протокол завершается.

Сложностью протокола мы назвали максимальное количество передаваемых битов. В описанной геометрической интерпретации сложность протокола равна глубине дерева — максимальному расстоянию от корня до листа.

1. Для всякой функции $f\colon X \times Y \to Z$ докажите, что $\mathrm{cc}(f) \leqslant \lceil \log |X| \rceil + \lceil \log |Z| \rceil$.

2. Пусть Алиса и Боб получают подмножества $x, y \subseteq \{1, \ldots, n\}$ и хотят вычислить функцию $\mathrm{MAX}(x,y) = \max(x \cup y)$. Доказать, что **а)** $\mathrm{cc}(MAX) \leqslant 2\lceil \log n \rceil$; **б)** $\mathrm{cc}(MAX) \leqslant \frac{3}{2}\lceil \log n \rceil$; **в)** $\mathrm{cc}(MAX) \leqslant \lceil \log n \rceil + O(\sqrt{\log n})$.

3. Пусть Алиса и Боб получают подмножества $x, y \subseteq \{1, \dots, n\}$. Функцию $\mathrm{AVG}(x, y)$ равна среднему арифметическому чисел в мультимножестве (то есть в множестве, в которое каждый элемент может входить несколько раз) $x \cup y$. Доказать, что $\mathrm{cc}(\mathrm{AVG}) = O(\log n)$.

4. Пусть Алиса и Боб получают подмножества $x, y \subseteq \{1, \dots, n\}$. Функция $\mathrm{MED}(x, y)$ определяется как медиана (средний элемент) в мультимножестве $x \cup y$. Доказать, что **а)** $\mathrm{cc}(\mathrm{MED}) = O(\log^2 n)$; **б)** $\mathrm{cc}(\mathrm{MED}) = O(\log n)$.

5. Пусть $x, y \in \{0, 1\}^n$. Функция $\mathrm{EQ}(x, y)$ равна 1 тогда и только тогда, когда $x = y$. Докажите, что $\mathrm{cc}(EQ) = n + 1$.

6. Пусть $x, y \in \{0, 1\}^n$. Функция $\mathrm{GT}(x, y)$ равна 1 тогда и только тогда, когда $x \geqslant y$, как числа в двоичной записи. Найдите, чему равна $\mathrm{cc}(GT)$.

7. Пусть $x, y \subseteq \{1, \dots, n\}$. Функция $\mathrm{DISJ}(x, y)$ равна 1 тогда и только тогда, когда $x \cap y = \emptyset$. Найдите, чему равна $\mathrm{cc}(DISJ)$.

8. Пусть функция $f \colon X \times Y \to Z$ сюръективна. Докажите, что $cc(f) \geqslant \log |Z|$.

Пусть $f \colon X \times Y \to \{0, 1\}$. Множество $S \subseteq X \times Y$ называется *трудным множеством*, если существует $z \in \{0, 1\}$, такое что

- Для всех $(x, y) \in S$ верно $f(x, y) = z$.
- Для всяких различных $(x_1, y_1) \in S$ и $(x_2, y_2) \in S$ верно $f(x_1, y_2) \neq z$ или $f(x_2, y_1) \neq z$.

9. а) Докажите, что если у f есть трудное множество размера t, то $\mathrm{cc}(f) \geqslant \log_2 t$. **б)** Решите три предыдущие задачи с помощью трудных множеств.

Пороговый элемент задается набором целых чисел $w_1, \dots, w_n, \theta$. На входных булевых переменных $x_1, \dots, x_n$ он вычисляет результат сравнения $\sum_{i=1}^n w_i x_i > \theta$, то есть выдает 1, если неравенство верно, и 0 иначе. *Весом* порогового элемента называется $\sum |w_i|$.

10. Докажите, что функцию GT можно вычислить пороговым элементом.

11. Пусть функция GT вычисляется булевой схемой размера S из пороговых элементов. Пусть W максимальный вес порогового элемента в этой схеме. Докажите, что $S \cdot \log W = \Omega(n)$.

Пусть протокол разрешает каждому из Алисы и Боба в процессе обмена сообщениями подбрасывать монету и действовать по-разному при разных исходах бросания. Назовем такой протокол *вероятностным.* Будем говорить, что такой коммуникационный протокол вычисляет функцию $f : X \times Y \to Z$ с вероятностью ошибки менее $1/3$, если для любых $x \in X$ и $y \in Y$ в результате выполнения протокола Алиса и Боб узнают значение $f(x, y)$ с вероятностью не менее $2/3$ (вероятность берётся по случайным бросаниям монет Алисы и Боба).

12. Докажите, что функция $\mathrm{EQ}(x, y)$ имеет вероятностную коммуникационную сложность $O(\log n)$.

13. Пусть $x, y \in \{0, 1\}^n$ и пусть заранее известно, что хэмминговское расстояние между x и y (число позиций, в которых x и y различаются) не превосходит $0.1n$. Алиса хочет сообщить своё слово Бобу (иначе говоря, Бобу нужно вычислить значение функции $f(x, y) = x$). Докажите, что вероятностная коммуникационная сложность этой задачи не меньше $0.01n$. *[Замечание: В этой задаче мы поменяли "правила игры": здесь не предполагается, что посторонний наблюдатель может определить значение x по одним только передаваемым Алисой и Бобом сообщениям, не зная заранее y.]*

14. Пусть коммуникационная сложность вычисления некоторой функции $f : X \times Y \to Z$ равна k. Докажите, что существует однораундовый коммуникационный протокол сложности не более 2^k, при выполнении которого Боб узнает значение $f(x, y)$. (В однораундовом протоколе вся информация передаётся только в одном направлении – от Алиса к Бобу; затем Боб вычисляет ответ.)

15. Коммуникационная задача NBA имеет следующее неформальное описание. Алиса помнит название команды, которая выиграла в баскетбольном матче (но не помнит название проигравшей команды). Боб знает название обеих команд, которые участвовали в игре, но не знает, какая из них победила. Требуется организовать протокол, по результатам которого Боб узнает команду-победительницу.

Опишем задачу формально. Алисе дано двоичное слово $x \in \{0, 1\}^n$, а Бобу пара слов $y = [u, v] \in \{0, 1\}^n \times \{0, 1\}^n$; при этом известно, что $x = u$ или $x = v$. Требуется организовать коммуникационный протокол, при выполнении которого Боб узнает x.

(а) Какова минимальная сложность 1-раундового протокола для этой задачи? (Алиса посылает Бобу сообщения, по которым Боб определяет значение x.)

(б) Какова минимальная сложность 2-раундового протокола для этой задачи? (Сначала Боб посылает Алисе некоторое сообщение, затем Алиса посылает ответ Бобу. В результате Боб определяет значение x.)

16. Пусть $x, y \in \{0, 1\}^n$ и известно, что хэмминговское расстояние между x и y (число позиций, в которых слова x и y различаются) не превосходит некоторой константы d. Алиса хочет узнать слово Боба y, а Боб хочет узнать слово Алисы x (формально это значит, что Алиса и Боб хотят вычислить значение функции $f(x, y) = [x, y]$). Докажите, что детерминированная коммуникационная сложность этой задачи равна $O(\log n)$. Дополнительный вопрос: как мультипликативная константа в ответе $O(\log n)$ зависит от d ?