Министерство образования и науки Российской Федерации

Государственное образовательное учреждение
высшего профессионального образования
«Московский физико-технический институт (государственный университет)»
Факультет инноваций и высоких технологий
Кафедра анализа данных

Магистерская диссертация

Тема: **Название моей работы (TODO)**

Направление: 010400
Прикладные математика и информатика

Выполнил:
студент 093 группы ____________ Попов М.В.

Научный руководитель:
д.физ.-мат.н., проф.(todo) ____________ Ромащенко А.Е.

г. Москва 2016

Содержание

Введение

(ToDo) Актуальность, новизна, краткая выжимка.

Коммуникационная сложность

1.1 Постановка задачи

Мы будем рассматривать задачи следующего вида: пусть имеется два человека, которые хотят совместно вычислить значение некоторой функции от двух переменных $f(x, y)$. По традиции мы будем называть первого участника игры Алисой, а второго Бобом. Сложность у этой задачи в том, что Алиса знает только значение аргумента x, а Боб значение аргумента y. Алиса и Боб могут обмениваться сообщениями по каналу связи. Требуется вычислить значение $f(x, y)$, переслав по каналу связи минимальное количество информации.

Мы предполагаем, что Алиса и Боб заранее (до того, как им станут известны значения x и y) договариваются о коммуникационном протоколе — о наборе соглашений, какие именно данные и в каком порядке они будут пересылать друг другу при тех или иных значениях x и y.

Опишем теперь всю задачу более формально. Пусть имеются конечные множества X, Y, Z и задана некоторая функция $f : X \times Y \to Z$.

Определение. *Коммуникационным протоколом для вычисления некоторой функции $f : X \times Y \to Z$ называется ориентированное двоичное дерево со следующей разметкой на вершинах и ребрах:*

- *каждая нелистовая вершина помечена буквой A или B;*
  - *у вершин с пометкой A определена функция $g_i : X \to \{0, 1\}$;*
  - *у вершин с пометкой B определена функция $f_j : Y \to \{0, 1\}$;*
- *каждой листовой вершине сопоставлен элемент множеста Z;*
- *каждое ребро помечено* 0 *или* 1.

Пусть Алиса и Боб договорились, что будут действовать по некоторому протоколу $\mathcal{P}$. Затем Алиса получила $x \in X$, а Боб получил $y \in Y$.

Поместим фишку в корневую вершину нашего протокола $\mathcal{P}$ и будем перемещать ее вниз по дереву, последовательно удаляясь от корня, пока она не попадём в один из листьев. Перемещение фишки выполняется следующим образом. Если текущая вершина помечена буквой A это значит, что сейчас очередь Алисы. Она применяет функцию g_i текущей вершины к своему значению x. Алиса отправляет по каналу связи бит равный $g_i(x)$ и перемещает фишку по ребру, помеченному как $g_i(x)$. Боб получает отправленный бит и понимает куда была сдвинута фишка. Для вершин помеченных буквой B поступают аналогично. Когда фишка попадает в лист дерева, записанное там значение $z \in Z$ объявляется результатом выполнения протокола.

Мы говорим, что протокол $\mathcal{P}$ вычисляет функцию $f : X \times Y \to Z$, если для любого $x \in X$ и любого $y \in Y$ при движении из корня по пути, соответствующему заданным x и y, мы попадаем в лист, помеченный $z = f(x, y)$.

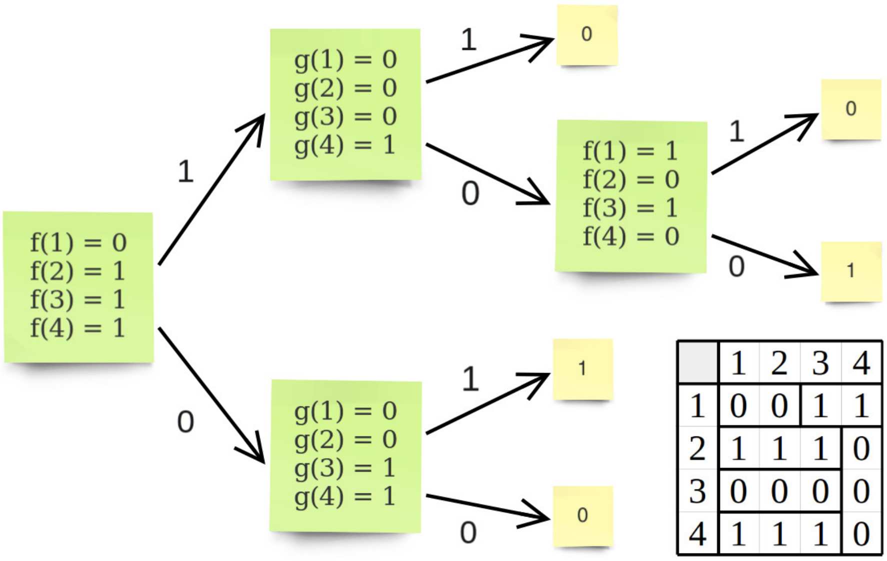
Определение. *Сложностью коммуникационного протокола называется его глубина. Коммуникационной сложностью функции f называется минимальная сложность протокола, вычисляющего f. Мы будем обозначать её $CC(f)$.*

1.2 Одноцветные комбинаторные прямоугольники

Определение. *Множество $S \subset X \times Y$ называется комбинаторным прямоугольником (или просто прямоугольным множеством), если существуют такие $A \subset X$ и $B \subset Y$, что $S = A \times B$.*

Пусть $\mathcal{P}$ некоторый коммуникационный протокол для вычисления функции $f : X \times Y \to Z$ и l один из листьев протокола. Определим S_l, как множество пар $(x, y) \in X \times Y$ таких, что на входе (x, y) Алиса и Боб, следуя протоколу $\mathcal{P}$, приходят в лист l.

Утверждение. *Для всякого коммуникационного протокола $\mathcal{P}$ и для всякого листа l множество S_l является комбинаторным прямоугольником.*



	1	2	3	4
1	0	0	1	1
2	1	1	1	0
3	0	0	0	0
4	1	1	1	0

Рис. 1: Пример протокола и разбиения таблицы значений.

Доказательсво этого утверждения можно прочитать например в [1]. В итоге мы получаем, что коммуникационный протокол для вычисления функции f задаёт разбиение $X \times Y$ - таблицы значений f на прямоугольные множества, соответствующие листьям. Поскольку каждому листу протокола приписано одно значение функции f, эти прямоугольные множества являются одноцветными, то есть во всех точках такого прямоугольного множества функция f принимает одно и то же значение. Например, для $X = Y = \{1, 2, 3, 4\}$, $Z = \{0, 1\}$ и протокола $\mathcal{P}$ (рис 1.) получаем разбиение на 5 одноцветных прямоугольных множества.

Подведем промежуточные итоги: всякий протокол с l листьями (вычисляющий функцию f) задаёт разбиение таблицы значений f на l одноцветных прямоугольных множества. Значит, чтобы доказать, что коммуникационная сложность $CC(f)$ не меньше n, достаточно показать, что таблицу значений невозможно разбить на менее, чем 2^n одноцветных прямоугольных множества.

1.3 Графовая интерпретация

Давайте теперь посмотрим на другое представление множества значений функции f. Рассмотрим полный двудольный граф $G = (X, Y, E)$, ребра которого раскрашены в $|Z|$ цветов. Вершины левой доли соот-

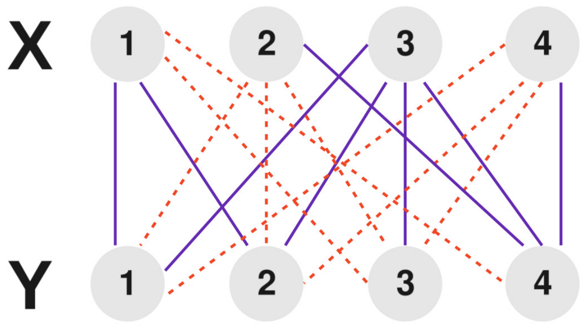


Рис. 2: Графовая интерпретация: синие – 0, красные – 1.

ветствуют элементам множества X, вершины правой доли - элементам множества Y. Ребро $(x, y) \in X \times Y$ имеет цвет $z \in Z$, если $f(x, y) = z$.

Из определения комбинаторного прямоугольника видно, что в графовой интерпертации он является ничем иным, как полным двудольным подграфом. А разбиение таблицы значений f на одноцветные прямоугольные множества это разбиение нашего полного двудольного графа G на одноцветные непересекающиеся биклики (полные двудольные подграфы). Для нашего примера графовую интерпретацию можно посмотреть на рис.2.

Определение. *Бикликовым числом $bcc(G)$ двудольного графа G будем называть наименьшее число биклик, которыми можно покрыть все ребра графа G (Биклики могут пересекаться).*

Для каждого $z \in Z$ определим двудольный граф $G_z = (X, Y, E_z)$, как граф, получающийся из G выкидыванием всех ребер цвета отличного от z, иначе говоря $E_z = \{(x, y) \in X \times Y \mid f(x, y) = z\}$.

Величины $bbc(G_z)$ дают некоторую нижнюю оценку на величину минимального покрытия непересекающимися бикликами, поэтому:

$$2^{CC(f)} \geq \sum_{z \in Z} bcc(G_z)$$

Замечание. *На самом деле величины $bbc(G_z)$ тесно связаны с так называемой недетерминированной коммуникационной сложностью $NCC(f)$.*

Для произовльного множества Z верно:

$$NCC(f) \leq \lceil log_2(\sum_{z \in Z} bcc(G_z)) \rceil + 1$$

а для $Z = \{0, 1\}$:

$$NCC(f) = \lceil log_2(bcc(G_1)) \rceil$$

Подробнее про это можно прочитать, например в [2].

В итоге мы получили мощный иструмент для доказательства нижних оценок коммуникационной сложности. К сожалению задача нахождения величины $bcc(G)$ является NP-полной [3] и точное значение известно только для очень скудного класса графов (например для "Crown graphs"), поэтому напрямую мы не можем использовать эту оценку. В следующей главе я рассмотрю несколько методов, позволяющих для произвольного двудольного графа оценивать снизу величину $bcc(G)$.

Коммуникационная сложность2

2.1 Постановка задачи

деле велич

Список литературы

[1] Kushilevitz Eyal, Nisan Noam. Communication Complexity. Cambridge University press, 2006.

[2] Razborov Alexander. Communication Complexity. In: An Invitation to Mathematics: from Competitions to Research. Springer, 2011.

[3] Garey M.R., Jhonson D.S. Computers and Intractability. W.H. Freeman, San Francisco, 2002.